Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v6n2p260

Estimating the Five Parameter Lambda Distribution Using Moment Based Methods

By Mahdi,Cadogan

October 14, 2013

Electronic Journal of Applied Statistical Analysis
Vol. 6, Issue 2, 2013, 260-279
DOI: 10.1285/i20705948v6n2p260

Estimating the Five Parameter Lambda Distribution Using Moment Based Methods

Smail Mahdi* and Kimberley Cadogan

Computer Science, Mathematics and Physics Department, University of the West Indies, Barbados

October 14, 2013

With a flexible probability density function (p.d.f) and five parameters at its disposal, the five parameter lambda distribution (FPLD) is suitable for distributional modelling. However, little research has been carried out on this distribution to date. And although the most recent published work focuses on how to apply newly developed estimation techniques, the literature does not address how to accomplish parametric estimation using existing well-established estimation methods. Hence, this research shows how to estimate the FPLD using the methods of moments, probability weighted moments (PWMs) and linear moments (L-moments) with the specific goal of determining whether any one method is superior to the others. To illustrate the proposed methods, the FPLD was fitted to the Standard Normal distribution. The results show that Standard Normal distribution was easily approximated by the FPLD using all three estimation techniques. Overall, the methods of PWMs and L-moments were deemed to be superior to the method of moments despite the fact that neither outperformed the other according to the goodness of fit tests.

keywords: lambda, probability weighted moments, method of moments, linear moments, normal distribution.

1 Introduction

In recent times, the practice of fitting a probability distribution using two-parameter models like the normal distribution has been overused. This development is particularly alarming because it gives the misleading impression that any empirical distribution can be summarised by just two characteristics, however, this is not always the case. For instance, while the centre and variability of a distribution is captured by the two parameters, information on the shape of the distribution

*Corresponding authors: smail.mahdi@cavehill.uwi.edu, kim.cadogan@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

has been lost. Thus, a two parameter model is a grossly inadequate means of encapsulating all the information contained within a distribution. In order to go beyond these characteristics to include say skewness or kurtosis, a new model with at least three parameters readily available is required. As its name suggests, the Five Parameter Lambda Distribution (FPLD) meets this basic requirement with a count of five parameters. This makes it an attractive option for distributional modelling.

Statistical distributions are usually defined by either a probability density function (p.d.f) or cumulative distribution function (c.d.f), however, the FPLD is instead specified by a quantile function

$$Q(p) = \lambda_1 + \lambda_2 p^{\lambda_4} - \lambda_3 (1-p)^{\lambda_5}, \tag{1}$$

where λ_1 is the location parameter, $\lambda_{2,}$ λ_3 are linear scale parameters and the parameters $\lambda_{4,}$ λ_5 determine the shape of the quantile function. In terms of application, Equation (1) is relevant for not only simulations, but also for order statistics, optimal grouping, inequality measures, heavy tail behaviour analysis, loss distributions, osculatory interpolation and Quantile-Quantile plotting (Tarsitano, 2005). Our research will however focus on an alternative form that was originally suggested by Gilchrist (2000):

$$Q(p) = \lambda_1 + \frac{\lambda_2}{2}\left[(1-\lambda_3)\left(\frac{p^{\lambda_4}-1}{\lambda_4}\right) - (1+\lambda_3)\left(\frac{(1-p)^{\lambda_5}-1}{\lambda_5}\right)\right]. \tag{2}$$

In Equation (2), λ_1 and λ_3 and are the respective location and skew parameters; albeit since λ_3 acts a relative weight of the tail, it also influences the distribution's shape which is usually determined independently by λ_4 and λ_5. The parameter λ_2 behaves as a multiplier to the quantile function of the transformed random variable $Z = X - \lambda_1$; λ_2 is therefore the scale parameter. Observe that when $\lambda_2 = 0$, the FPLD degenerates to a one-point distribution $Q(p) = \lambda_1$. Equation (2) is therefore valid only when $\lambda_2 > 0$ and $-1 \leq \lambda_3 \leq 1$ since these conditions ensure the equation is a continuous and monotonically increasing function of p (Tarsitano, 2010).

The fact that $F(x) = p$ and $x = Q(p)$ implies

$$f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} = \frac{\mathrm{d}p}{\mathrm{d}Q(p)} = \left(\frac{\mathrm{d}Q(p)}{\mathrm{d}p}\right)^{-1}. \tag{3}$$

Therefore, the p.d.f of the FPLD is derived from the derivative of Equation (2) to be

$$f(x) \quad = \quad \frac{2}{\lambda_2\left[(1-\lambda_3)\,p^{\lambda_4-1} + (1+\lambda_3)\left((1-p)^{\lambda_5-1}\right)\right]}. \tag{4}$$

According to Tarsitano (2010), the density will be zeromodal if $\{\max(\lambda_{4,}\lambda_5) > 1 \ \wedge\ \min(\lambda_{4,}\lambda_5) < 1\}$, unimodal with continuous tails if $\{\max(\lambda_{4,}\lambda_5) < 1\}$, unimodal with truncated tails if $\{\min(\lambda_{4,}\lambda_5) > 2\}$, U-shaped if $\lambda_4 > 1, \lambda_5 > 2$ and S-shaped if $\{\max(\lambda_{4,}\lambda_5) > 2 \ \wedge\ \min(\lambda_{4,}\lambda_5) > 1\}$. Table 1 gives a synopsis of the behavior of the distribution (Tarsitano, 2010).

To date only two papers, both written by Tarsitano (2005, 2010), were found at the time of writing this literature review. Tarsitano (2005) conducted a comprehensive study of Equation (1)

Table 1: Behaviour and shapes of the FPLD

Cases	Condition	Comment
Special	$\lambda_3 < 0$ and $\lambda_4 = \lambda_5$	Skewed to the left (i.e. negatively skewed)
	$\lambda_3 = 0$ and $\lambda_4 = \lambda_5$	Symmetric about λ_1
	$\lambda_3 > 0$ and $\lambda_4 = \lambda_5$	Skewed to the right (i.e. positively skewed)
	$\lambda_3 = -1$	λ_4 controls the kurtosis of left tail
	$\lambda_3 = 1$	λ_5 controls the kurtosis of right tail
General	$\lambda_3 = -1$ and $\lambda_4 = 1$	Uniform distribution
	$\lambda_3 = 1$ and $\lambda_5 = 1$	Uniform distribution
	$\lambda_4 = 1$ and $\lambda_5 = 1$	Uniform distribution
	$\lambda_3 = 0$ and $\lambda_4, \lambda_5 = 2$	Uniform distribution
	$\lambda_4 \to 0$ and $\lambda_5 \to 0$	Skew logistic distribution
	$\lambda_4 \to \infty$ and $\lambda_5 \to 0$	Exponential distribution
	$\lambda_4 \to 0$ and $\lambda_5 \to \infty$	Reflected exponential distribution
	$\lambda_4 \to \infty$ and $\|\lambda_5\| < \infty$	Generalized Pareto distribution
	$\|\lambda_4\| < \infty$ and $\lambda_5 \to \infty$	Power-function distribution

by estimating it using six methods: percentile, moments, probability weighted moments, minimum Cramér-Von Mises, maximum likelihood and pseudo least squares. The purpose of his investigations was to identify the most appropriate method for estimating the FPLD from data grouped in histograms or frequency tables. This was achieved by using Monte Carlo simulations for nine combinations of sample sizes and a number of classes. The pseudo least squares method was deemed the best suited candidate. Tarsitano (2010) lamented the non-evolution of parametric estimation techniques despite advances in computing and technology. To advance the literature and correct this perceived deficiency, Tarsitano (2010) devised two procedures based on nonlinear least squares and least absolute deviations. In either approach, the linear parameters of Equation (2) were first estimated for fixed values of the nonlinear parameters and then optimized using a controlled random search algorithm that terminated when the prefixed tolerance was exceeded. Like its forerunner ordinary least squares, nonlinear least squares' objective is to minimize the sum of the squared difference between each ordered statistic (or observation) and its mathematical expectation. However, instead of using an approximation of the expected value, the mean is computed from a closed-form expression. In comparison, the method of least absolute deviations adopts the median instead of the mean as the measure of location. Subsequently, this method is less susceptible to outliers and influential data points. Moreover, instead of minimizing the squared difference, the objective becomes to minimize the absolute deviation or error for all the ordered statistics, hence the name. Tarsitano (2010) applied both schemes to simulated and real data sets and found that the nonlinear least squares' estimators outperformed the least absolute deviation's estimators in terms of their bias and mean squared error. Nonetheless, in larger samples, both methods were found to have performed comparably.

1.1 Research Objective and Organization of Paper

The purpose of this work is to demonstrate the estimation of the FPLD using the selected methods of moments, probability weighted moments (PWMs), and linear moments (L-moments) and to determine whether any one method is superior in fitting the model. This paper is organized as follows: Section 2 describes how to apply the methods of moments, PWMs and L-moments to fit the FPLD; this is followed by a demonstration of FPLD modelling using samples of data randomly generated from the Standard Normal distribution in Section 4; after which we offer our conclusions and recommendations.

2 Estimation Methods

The selected estimation methods, namely: the methods of moments, PWMs and L-moments, were originally developed to be used with distributions defined by either a p.d.f or c.d.f. These methods are aptly called "matching methods" because they seek to match specific population properties of the distribution being fitted, in this case the moments, PWMs and L-moments, with their corresponding properties computed from a sample of data. As a result, we describe rather extensively how to adapt these methods to fit the FPLD model which is readily defined by a quantile function. However, due to the fact that the FPLD quantile function is implicitly defined, suitable initial values for each procedure must be found and then improved under an appropriate optimization scheme.

2.1 Method of Moments

In their introductory paper, Ramberg and Schmeiser (1974) described an algorithm that exclusively used moments to estimate the parameters of the GLD by simply equating the first four theoretical moments of the distribution to the sample moments of the empirical data. Therefore the parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 of Equation (2) can be similarly found by equating the mean α_1, variance α_2, skewness α_3, kurtosis α_4 and the fifth standardised moment α_5 of the FPLD to the empirical mean $\hat{\alpha}_1$, variance $\hat{\alpha}_2$, skewness $\hat{\alpha}_3$, kurtosis $\hat{\alpha}_4$ and the fifth standardised moment $\hat{\alpha}_5$ of a sample. That is, these parameters are computed such that:

$$\alpha_1 \equiv E(X) = \hat{\alpha}_1 \tag{5}$$

$$\alpha_2 \equiv E\left((X-\mu)^2\right) = \hat{\alpha}_2 \tag{6}$$

$$\alpha_3 \equiv \frac{E\left((X-\mu)^3\right)}{\sigma^3} = \hat{\alpha}_3 \tag{7}$$

$$\alpha_4 \equiv \frac{E\left((X-\mu)^4\right)}{\sigma^4} = \hat{\alpha}_4 \tag{8}$$

$$\alpha_5 \equiv \frac{E\left((X-\mu)^5\right)}{\sigma^5} = \hat{\alpha}_5 \tag{9}$$

Rewriting the Equation (2) as

$$X = \lambda_1 + \frac{\lambda_2}{2}\left\{k_1 p^{\lambda_4} - k_1 - k_2(1-p)^{\lambda_5} + k_2\right\}$$

where $k_1 = \frac{1-\lambda_3}{\lambda_4}$ and $k_2 = \frac{1+\lambda_3}{\lambda_5}$, observe that a linear transformation $Z = \frac{2(X-\lambda_1)}{\lambda_2} + k_1 - k_2 = k_1 p^{\lambda_4} - k_2(1-p)^{\lambda_5}$ simplifies our calculations as the noncentral moments $E(X^r)$ are now linearly related to $E(Z^r)$ by

$$E(X^r) = E\left(\left[\lambda_1 + \frac{\lambda_2}{2}\{Z - k_1 + k_2\}\right]^r\right). \quad (10)$$

Given

$$\begin{aligned} E(Z^r) &= \int_{-\infty}^{\infty} z^r f(z)\, \mathrm{d}z \\ &= \int_0^1 \left(k_1 p^{\lambda_4} - k_2(1-p)^{\lambda_5}\right)^r \mathrm{d}p, \end{aligned} \quad (11)$$

expanding the integrand on the right using the binomial theorem gives

$$\begin{aligned} E(Z^r) &= \int_0^1 \left\{\sum_{j=0}^{r} \binom{r}{j} k_1^{(r-j)} p^{\lambda_4(r-j)} (-1)^j k_2^j (1-p)^{\lambda_5 j}\right\} \mathrm{d}p \\ &= \sum_{j=0}^{r} \left[(-1)^j \binom{r}{j} k_1^{(r-j)} k_2^j \mathrm{B}(\lambda_4(r-j)+1, \lambda_5 j + 1)\right]. \end{aligned} \quad (12)$$

However, $E(Z^r)$ exists if and only if the beta function $B(\cdot)$ is defined, that is,

$$\lambda_4(r-j) + 1 > 0 \text{ and } \lambda_5 j + 1 > 0$$

for all $j = 0, 1, 2, \ldots, r$. This condition prevails only when $\lambda_4 > -\frac{1}{r}$ and $\lambda_5 > -\frac{1}{r}$, and since our intention is to calculate the first five moments of the FPLD, this means the shape parameters must be restricted to $\lambda_4 > -\frac{1}{5}$ and $\lambda_5 > -\frac{1}{5}$. Therefore expanding Equation (12) for $r = 1, 2, 3, 4, 5$ yields:

$$
\begin{aligned}
E(Z) &= \mathscr{A} && (13)\\
&= k_1 \mathrm{B}(1+\lambda_4, 1) - k_2 \mathrm{B}(1, 1+\lambda_5);\\
E\left(Z^2\right) &= \mathscr{B} && (14)\\
&= k_1^2 \mathrm{B}(1+2\lambda_4, 1) - 2k_1k_2 \mathrm{B}(1+\lambda_4, 1+\lambda_5) + k_2^2 \mathrm{B}(1, 1+2\lambda_5);\\
E\left(Z^3\right) &= \mathscr{C} && (15)\\
&= k_1^3 \mathrm{B}(1+3\lambda_4, 1) - 3k_1^2k_2 \mathrm{B}(1+2\lambda_4, 1+\lambda_5)\\
&\quad + 3k_1k_2^2 \mathrm{B}(1+\lambda_4, 1+2\lambda_5) - k_2^3 \mathrm{B}(1, 1+3\lambda_5);\\
E\left(Z^4\right) &= \mathscr{D} && (16)\\
&= k_1^4 \mathrm{B}(1+4\lambda_4, 1) - 4k_1^3k_2 \mathrm{B}(1+3\lambda_4, 1+\lambda_5)\\
&\quad + 6k_1^2k_2^2 \mathrm{B}(1+2\lambda_4, 1+2\lambda_5) - 4k_1k_2^3 \mathrm{B}(1+\lambda_4, 1+3\lambda_5)\\
&\quad + k_2^4 \mathrm{B}(1, 1+4\lambda_5);\\
E\left(Z^5\right) &= \mathscr{E} && (17)\\
&= k_1^5 \mathrm{B}(1+5\lambda_4, 1) - 5k_1^4k_2 \mathrm{B}(1+4\lambda_4, 1+\lambda_5)\\
&\quad + 10k_1^3k_2^2 \mathrm{B}(1+3\lambda_4, 1+2\lambda_5) - 10k_1^2k_2^3 \mathrm{B}(1+2\lambda_4, 1+3\lambda_5)\\
&\quad + 5k_1k_2^4 \mathrm{B}(1+\lambda_4, 1+4\lambda_5) - k_2^5 \mathrm{B}(1, 1+5\lambda_5).
\end{aligned}
$$

and subsequently substituting Equation (13) and $r = 1$ into Equation (10) gives

$$\alpha_1 \quad = \quad \lambda_1 + \frac{\lambda_2}{2}\left\{\mathscr{A} - k_1 + k_2\right\}. \tag{18}$$

Since $\mu = \alpha_1$,

$$X - \mu = \frac{\lambda_2}{2}\left\{Z - \mathscr{A}\right\}$$

so that in general

$$E\left([X-\mu]^r\right) = \frac{\lambda_2^r}{2^r} E\left([Z-\mathscr{A}]^r\right). \tag{19}$$

Thus, using the binomial expansions for $r = 1, 2, 3, 4, 5$ and appropriate substitutions of Equation (13) through Equation (17) yields the following expressions for the moments of the FPLD:

$$\alpha_2 = \frac{\lambda_2^2\left(\mathscr{B}-\mathscr{A}^2\right)}{4}, \tag{20}$$

$$\alpha_3 = \frac{\lambda_2^3\left(\mathscr{C}-3\mathscr{A}\mathscr{B}+2\mathscr{A}^3\right)}{8\sigma^3}, \tag{21}$$

$$\alpha_4 = \frac{\lambda_2^4\left(\mathscr{D}-4\mathscr{A}\mathscr{C}+6\mathscr{A}^2\mathscr{B}-3\mathscr{A}^4\right)}{16\sigma^4}, \tag{22}$$

$$\alpha_5 = \frac{\lambda_2^5\left(\mathscr{E}-5\mathscr{A}\mathscr{D}+10\mathscr{A}^2\mathscr{C}-10\mathscr{A}^3\mathscr{B}+4\mathscr{A}^5\right)}{32\sigma^5}. \tag{23}$$

But rearranging Equation (20) and generalising gives $\left({}^{\lambda_2}/_{2\sigma}\right)^i = \left(\mathscr{B}-\mathscr{A}^2\right)^{-i/2}$ for any integer i. Hence, when $i = 3,4$ and 5, Equation (21), Equation (22) and Equation (23) simplify to

$$\alpha_3 = \frac{\mathscr{C}-3\mathscr{A}\mathscr{B}+2\mathscr{A}^3}{\left(\mathscr{B}-\mathscr{A}^2\right)^{3/2}}, \tag{24}$$

$$\alpha_4 = \frac{\mathscr{D}-4\mathscr{A}\mathscr{C}+6\mathscr{A}^2\mathscr{B}-3\mathscr{A}^4}{\left(\mathscr{B}-\mathscr{A}^2\right)^2} \text{ and} \tag{25}$$

$$\alpha_5 = \frac{\mathscr{E}-5\mathscr{A}\mathscr{D}+10\mathscr{A}^2\mathscr{C}-10\mathscr{A}^3\mathscr{B}+4\mathscr{A}^5}{\left(\mathscr{B}-\mathscr{A}^2\right)^{5/2}}. \tag{26}$$

Notice that $\mathscr{A}$ (Equation (13)) to $\mathscr{E}$ (Equation (17)) are functions of the parameters λ_3, λ_4 and λ_5 only. This makes finding the estimates somewhat easier since we are essentially attempting to solve the reduced system of nonlinear equations:

$$\begin{aligned} \alpha_3 &= \hat{\alpha}_3 \\ \alpha_4 &= \hat{\alpha}_4 \\ \alpha_5 &= \hat{\alpha}_5 \end{aligned}$$

for λ_3, λ_4 and λ_5 in the subregion delimited by $-1 \leq \lambda_3 \leq 1$ and $-0.20 < \lambda_4, \lambda_5 \leq 3$. Tarsitano (2010) identified the region $(-0.999, 3) \times (-0.999, 3)$ as being promising for solutions however the restrictions were enforced to ensure the moments existed. Finally once the values for λ_3, λ_4 and λ_5 are found, the remaining parameters are then computed from the formulas below:

$$\lambda_2 = \sqrt{\frac{4\hat{\alpha}_2}{\mathscr{B}-\mathscr{A}^2}} \qquad \text{and} \qquad \lambda_1 = \hat{\alpha}_1 - \frac{\lambda_2}{2}\left(A - k_1 + k_2\right).$$

2.2 Probability Weighted Moments

Let X be a real-valued random variable with cumulative distribution function $F(x)$ and quantile function $Q(x)$. Greenwood et al. (1979) defined PWMs of order t,r,s as

$$M_{t,r,s} = \mathrm{E}\left(X^t \{F(X)\}^r \{1-F(X)\}^s\right)$$

where t,r,s are real numbers. However, since $X = Q(P) = F^{-1}(P)$,

$$M_{t,r,s} = \mathrm{E}\left(Q^t(P) \{P\}^r \{1-P\}^s\right)$$

in terms of the quantile function; putting $t=1$ and $s=0$ gives the unique set of beta-PWMs of order r

$$\beta_r = \mathrm{E}\left(Q(P)P^r\right) = \int_0^1 Q(p)\, p^r \,\mathrm{d}p. \tag{27}$$

The method of PWMs equates population and sample PWMs to obtain a number of equations that are solved for the estimates of the parameters of a distribution. Since the number of equations needed depends on the numbers of parameters in the model being fitted, estimation of the FPLD by this method requires our interest in at least the first five PWMs of this distribution. Evaluation of the integral in Equation (27) gives

$$\begin{aligned}\beta_r &= \int_0^1 \left[\lambda_1 + \frac{\lambda_2}{2}\left\{k_1\left(p^{\lambda_4}-1\right) - k_2\left((1-p)^{\lambda_5}-1\right)\right\}\right] p^r \,\mathrm{d}p \\ &= \frac{\lambda_1}{r+1} + \frac{\lambda_2}{2}\left\{k_1\left(\frac{1}{\lambda_4+r+1} - \frac{1}{r+1}\right) - k_2\left(\mathrm{B}(r+1,\lambda_5+1) - \frac{1}{r+1}\right)\right\}.\end{aligned} \tag{28}$$

Thus, substituting $r = 0,1,2,3,4$ into Equation (28) and simplifying yields the following expressions for the FPLD population PWMs:

$$\beta_0 = \lambda_1 + \frac{\lambda_2(1-\lambda_3)}{2\lambda_4}\left(\frac{1}{\lambda_4+1} - 1\right) - \frac{\lambda_2(1+\lambda_3)}{2\lambda_5}\left(\mathrm{B}(1,\lambda_5+1) - 1\right) \tag{29}$$

$$\beta_1 = \frac{\lambda_1}{2} + \frac{\lambda_2(1-\lambda_3)}{2\lambda_4}\left(\frac{1}{\lambda_4+2} - \frac{1}{2}\right) - \frac{\lambda_2(1+\lambda_3)}{2\lambda_5}\left(\mathrm{B}(2,\lambda_5+1) - \frac{1}{2}\right) \tag{30}$$

$$\beta_2 = \frac{\lambda_1}{3} + \frac{\lambda_2(1-\lambda_3)}{2\lambda_4}\left(\frac{1}{\lambda_4+3} - \frac{1}{3}\right) - \frac{\lambda_2(1+\lambda_3)}{2\lambda_5}\left(\mathrm{B}(3,\lambda_5+1) - \frac{1}{3}\right) \tag{31}$$

$$\beta_3 = \frac{\lambda_1}{4} + \frac{\lambda_2(1-\lambda_3)}{2\lambda_4}\left(\frac{1}{\lambda_4+4} - \frac{1}{4}\right) - \frac{\lambda_2(1+\lambda_3)}{2\lambda_5}\left(\mathrm{B}(4,\lambda_5+1) - \frac{1}{4}\right) \tag{32}$$

$$\beta_4 = \frac{\lambda_1}{5} + \frac{\lambda_2(1-\lambda_3)}{2\lambda_4}\left(\frac{1}{\lambda_4+5} - \frac{1}{5}\right) - \frac{\lambda_2(1+\lambda_3)}{2\lambda_5}\left(\mathrm{B}(5,\lambda_5+1) - \frac{1}{5}\right) \tag{33}$$

Now let $x_{i:n}$ represent the value of the i^{th} ordered statistic of a sample of size n. Then the sample PWMs $\hat{\beta}_r$ are given by

$$\hat{\beta}_r = \frac{1}{n}\sum_i (p_i^r)\, x_{i:n},$$

where $p_i = \frac{i}{n}$ is taken to be the probability of the the i^{th} ordered statistic occurring. Thus, equating Equation (29) through Equation (33) with their corresponding $\hat{\beta}_r$-statistics produces a system of five equations

$$\beta_r = \hat{\beta}_r, \quad r = 0, 1, 2, 3, 4, \tag{34}$$

that must be solved simultaneously for estimates of the parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 in order to fit the FPLD to the sample distribution. However, since this nonlinear system is too complex to solve analytically, the estimates found at the end of the method of moments procedure were used to start the numerical algorithm. The final output of the algorithm was then checked to ensure that the obtained PWMs estimates for the FPLD satisfied the constraints $\hat{\lambda}_2 > 0$ and $-1 \le \hat{\lambda}_3 \le 1$.

2.3 Linear Moments

Although L-moments are linear functions of the expectations of order statistics, Hosking (1986) showed that they can be alternatively defined via the beta- PWMs by

$$\Lambda_{r+1} = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \beta_k, \quad r = 0, 1, 2, \ldots$$

which is in turn equivalent to

$$\Lambda_1 = \beta_0 \text{ and } \Lambda_r = \sum_{j=0}^{r-1} (p_{r,j})\, \beta_j, \quad r = 2, 3, \ldots \tag{35}$$

where

$$p_{r,j} = (-1)^{r-1-j} \binom{r-1}{j} \binom{r+j-1}{j} = \frac{(-1)^{r-1-j}(r+j-1)!}{(j!)^2 (r-j-1)!}.$$

Like PWMs, L-moments exist once the mean is real and finite. Hence, a distribution may be uniquely specified by its L-moments. Also, as with conventional moments, it is convenient to standardize the higher order L-moments Λ_r, $r \ge 3$ so that they are independent of the units of measurement of X. Accordingly, the L-moment ratios of X (abbreviated L-ratios) are defined to be

$$\tau_r = \frac{\Lambda_r}{\Lambda_2}, \quad r = 3, 4, 5, \ldots \tag{36}$$

so that using Equation (27) and Equation (35) the first five L-moments of the FPLD are:

$$\begin{aligned}\Lambda_1 &= \beta_0 \\ &= \int_0^1 Q(p)\,\mathrm{d}p \\ &= \lambda_1 + \frac{\lambda_2}{2}\left\{\frac{(1+\lambda_3)(\lambda_4+1)-(1-\lambda_3)(\lambda_5+1)}{(\lambda_4+1)(\lambda_5+1)}\right\}\end{aligned} \tag{37}$$

$$\begin{aligned}\Lambda_2 &= 2\beta_1 - \beta_0 \\ &= \int_0^1 Q(p)\cdot(2p-1)\,\mathrm{d}p \\ &= \frac{\lambda_2}{2}\left\{\frac{(1-\lambda_3)(\lambda_5+1)(\lambda_5+2)+(1+\lambda_3)(\lambda_4+1)(\lambda_4+2)}{(\lambda_4+1)(\lambda_4+2)(\lambda_5+1)(\lambda_5+2)}\right\}\end{aligned} \tag{38}$$

$$\begin{aligned}\tau_3 &= \frac{6\beta_2 - 6\beta_1 + \beta_0}{\Lambda_2} \\ &= \frac{\int_0^1 Q(p)\cdot\left(6p^2-6p+1\right)\,\mathrm{d}p}{\Lambda_2} \\ &= \frac{(1-\lambda_3)(\lambda_4-1)(\lambda_5+1)(\lambda_5+2)(\lambda_5+3)-(1+\lambda_3)(\lambda_5-1)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)}{(\lambda_4+3)(\lambda_5+3)\left\{A\right\}}\end{aligned} \tag{39}$$

$$\begin{aligned}\tau_4 &= \frac{20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0}{\Lambda_2} \\ &= \frac{\int_0^1 Q(p)\cdot\left(20p^3-30p^2+12p-1\right)\,\mathrm{d}p}{\Lambda_2} \\ &= \frac{(1-\lambda_3)(\lambda_4-1)(\lambda_4-2)E+(1+\lambda_3)(\lambda_5-1)(\lambda_5-2)D}{(\lambda_4+3)(\lambda_4+4)(\lambda_5+3)(\lambda_5+4)\left\{A\right\}}\end{aligned} \tag{40}$$

$$\begin{aligned}\tau_5 &= \frac{70\beta_4 - 140\beta_3 + 90\beta_2 - 20\beta_1 + \beta_0}{\Lambda_2} \\ &= \frac{\int_0^1 Q(p)\cdot\left(70p^4-140p^3+90p^2-20p+1\right)\,\mathrm{d}p}{\Lambda_2} \\ &= \frac{(1-\lambda_3)(\lambda_4-1)(\lambda_4-2)(\lambda_4-3)S-(1+\lambda_3)(\lambda_5-1)(\lambda_5-2)(\lambda_5-3)R}{(\lambda_4+3)(\lambda_4+4)(\lambda_4+5)(\lambda_5+3)(\lambda_5+4)(\lambda_5+5)\{A\}}\end{aligned} \tag{41}$$

where

$$\begin{aligned}
A &= (1-\lambda_3)(\lambda_5+1)(\lambda_5+2)+(1+\lambda_3)(\lambda_4+1)(\lambda_4+2),\\
D &= (\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4),\\
E &= (\lambda_5+1)(\lambda_5+2)(\lambda_5+3)(\lambda_5+4),\\
R &= D(\lambda_4+5),\\
S &= E(\lambda_5+5).
\end{aligned}$$

In many respects, L-moments are analogous to conventional central moments. For example, the first L-moment Λ_1 is the mean and the second L-moment Λ_2 is a scalar multiple of Gini's mean difference which is often used as a measure of dispersion. In addition, the higher order L-moment ratios τ_r, in particular, τ_3 and τ_4, measure skewness and kurtosis respectively. For a more detailed discussion of the interpretation of $\Lambda_1, \Lambda_2, \Lambda_3$ and Λ_4, see Hosking (1986, 1990).

The sample L-moments are computed from the sample as

$$l_r = \binom{n}{r}^{-1} \sum\sum_{i_1<i_2<\cdots<i_r}\cdots\sum \quad r^{-1}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} x_{i_{r-k}}$$

and the sample L-ratios denoted by $t_r = \frac{l_r}{l_2}$. However, these computations were automated by the function `lmom.ub` from the lmomco package in R.

Finally, equating the first five L-moments of the FPLD to those of the distribution of X produces the system of equations:

$$\begin{aligned}
\Lambda_r &= l_r, \quad r=1,2\\
\tau_r &= t_r, \quad r=3,4,5
\end{aligned}$$

which must be solved simultaneously in order to obtain the set of lambda values that approximate the unknown distribution of the random sample of data. Once again the method of moments' estimates were used as the initial points of the numerical procedure.

3 Assessing the Quality of Fit

While it is natural to speculate whether a particular set of λ_i-estimates provides a good fit to a distribution, proving it does is often difficult because the underlying distribution is not actually known. For this reason graphical tests were opted for as a direct means of determining the goodness of fit because they easily display the agreement between distributional fits. To evaluate the suitability of potential fits, the p.d.f of the FPLD fit is plotted against its theoretical counterpart in order to visualize the proximity or closeness of both models. These overplots of p.d.f.s were supplemented with the more powerful Quantile-Quantile (Q-Q) plots to help overcome any indecisiveness when discriminating between fits. A Q-Q plot is a fit-observation diagram which compares two probability distributions by plotting their quantiles against each other, that is, the observed data is plotted against the values estimated from the fitted distribution.

If the fitted distribution is the exact parent distribution, this relationship should appear as a straight line through the origin with a 45° slope (i.e. the line $y = x$). Therefore, if the two distributions are similar, then points in the Q-Q plot will lie approximately on the line. Hence the suitability of a distribution fit may be judged effectively by the correspondence between quantile estimates and those given by the straight line.

In addition to these visual aids, the distance between p.d.f.s,

$$\max_{1 \leq i \leq 500} \left| f(x_i) - \hat{f}(x_i) \right|,$$

was computed to measure the maximum distance between the p.d.f.s of two distributions. In this case, the statistic obtained reflects the worst case scenario of the proximity between the true and fitted distribution's p.d.f.s.

4 Application

Real data can be obtained from any number of statistical distributions, it is therefore crucial that the FPLD provides reasonably good fits to a multitude of distributions. We thus illustrate how to fit to the Standard Normal distribution in order to demonstrate the accuracy and reliability of the FPLD approximation. The normal distribution was selected because of its frequent use in statistics.

The first five analytical moments and L-moments of a normal distribution with mean μ and variance σ^2 $(\sigma > 0)$ are derived to be:

$$\alpha_1 = \mu, \quad \alpha_2 = \sigma^2, \quad \alpha_3 = 0, \quad \alpha_4 = 3, \quad \alpha_5 = 0 \tag{42}$$

and

$$\Lambda_1 = \mu, \quad \Lambda_2 = \pi^{-\frac{1}{2}}\sigma, \quad \tau_3 = 0, \quad \tau_4 = 0.122602, \quad \tau_5 = 0 \tag{43}$$

respectively. Since $\alpha_3, \alpha_4, \alpha_5, \tau_3, \tau_4$ and τ_5 are independent of μ and σ, these quantities will be unaffected by changes in those parameters. Hence, if the Standard Normal distribution is approximated by the FPLD, then the skew and shape parameters λ_3, λ_4 and λ_5 obtained from the fit can be used to model all other normal distributions provided of course the location and scale parameters are adjusted appropriately. On the other hand, note that the first two moments α_1, α_2 (Equation (42)) and L-moments Λ_1, Λ_2 (Equation (43)) are directly proportional to μ and σ. This means that the mean μ and variance σ^2 of any normal distribution can be computed directly from the sample/population moments and L-moments once their numerical values are known. Hosking (1986) indicates there is an explicit linear relationship between the parameters of a normal distribution and its first four PWMs, mainly:

Table 2: Parameters of the selected theoretical distribution

Method	First five parameters of the Standard Normal distribution				
Moments	$\alpha_1 = 0$	$\alpha_2 = 1$	$\alpha_3 = 0$	$\alpha_4 = 3$	$\alpha_5 = 0$
PWMs	$\beta_0 = 0$	$\beta_1 = 0.282095$	$\beta_2 = 0.282095$	$\beta_3 = 0.257344$	$\beta_4 = 0.232593$
L-moments	$\Lambda_1 = 0$	$\Lambda_2 = 1/\sqrt{\pi}$	$\tau_3 = 0$	$\tau_4 = 0.122602$	$\tau_5 = 0$

$$\beta_0 = \mu \tag{44}$$

$$2\beta_1 - \beta_0 = \pi^{-\frac{1}{2}}\sigma \tag{45}$$

$$3\beta_2 - \beta_0 = \frac{3}{2}\pi^{-\frac{1}{2}}\sigma \tag{46}$$

$$4\beta_3 - \beta_0 = 6\pi^{-\frac{3}{2}}\arctan\left(\sqrt{2}\right)\sigma. \tag{47}$$

However, "PWMs of higher order do not in general have analytical expressions, but can be found using tables of expected values of normal order statistics" (Hosking, 1986). Nevertheless, using Equation (44) in conjunction with any one of the other equations that appears in the system above, the parameters of the normal distribution can still be estimated from computed PWMs values. For example, the first five PWMs of a pseudo random sample of size 1000 generated from a normal distribution with mean $\mu = 6$ and standard deviation $\sigma = 0.5$ were calculated to be:

$$\hat{\beta}_0 = 6.018571, \quad \hat{\beta}_1 = 3.149902, \quad \hat{\beta}_2 = 2.146287, \quad \hat{\beta}_3 = 1.632156, \quad \hat{\beta}_4 = 1.318790.$$

Thus, Equation (44) implies

$$\hat{\mu} = \hat{\beta}_0 \approx 6.0 \quad (1 \text{ d.p.}), \tag{48}$$

and rearranging Equation (45) to make σ the subject gives

$$\hat{\sigma} = \frac{2\hat{\beta}_1 - \hat{\beta}_0}{\pi^{-\frac{1}{2}}} \approx 0.5 \quad (1 \text{ d.p.}). \tag{49}$$

Both results suggest the data must follow a distribution with estimates $\hat{\mu} = 6$ and $\hat{\sigma}^2 = 0.5^2 = 0.25$. This agrees with our scenario as the sample was drawn from a $N(6, 0.25)$ population.

Previously, we discussed how to model the moments, PWMs and L-moments of the distribution from a sample of randomly generated data or real data where the theoretical distribution is fully-known. However, these quantities can also be computed analytically and subsequent substitution of $\mu = 0$ and $\sigma = 1$ into Equation (42) and Equation (43) gives the theoretical population moments and L-moment of the Standard Normal distribution. In order to obtain the PWMs of the Standard Normal distribution, the linear relationship between PWMs and L-moments was exploited. That is, using the numerical values of analytical L-moments for $N(0,1)$ presented in Table 2, Equation (35) was expanded up to the first five terms and the generated linear system solved simultaneously for the unknown PWMs.

Finally, the FPLD was fitted to the Standard Normal distribution by simply replacing the sample moments, PWMs and L-moments in the system of equations under each estimation routine with the appropriate population parameters reported in Table 2.

5 The Results

The program `fitStdNormMoM` yielded three distinct fits of the FPLD to the Standard Normal distribution $N(0,1)$:

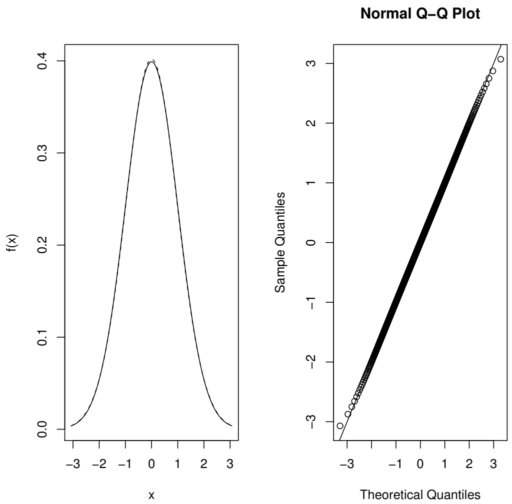
$$FPLD_{MoM1}(0, 1.3665, 0, 0.1349, 0.1349),$$
$$FPLD_{MoM2}(0, 24.8878, 0, 5.2029, 5.2029)$$

and

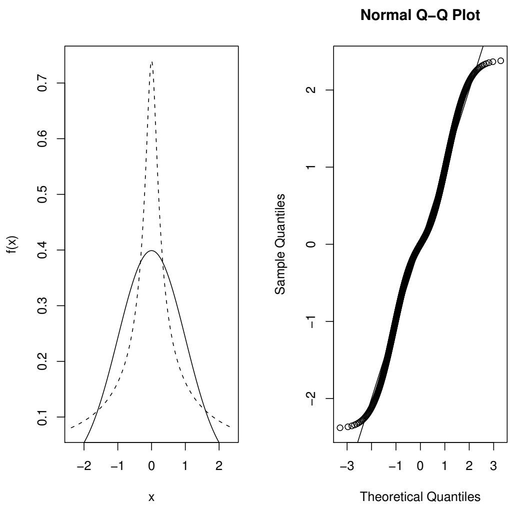
$$FPLD_{MoM3}(0, 24.8879, 0, 5.2029, 5.2029).$$

Note that the above output adheres to the pattern established in Table 1 for symmetrical distributions. That is, $\hat{\lambda}_3 = 0$ and $\hat{\lambda}_4 = \hat{\lambda}_5$. But even though all three fits accurately found the mean ($\mu = \hat{\lambda}_1 = 0$) of the Standard Normal distribution, only the first fit $FPLD_{MoM1}$ gives a reasonable approximation of the variance ($\sigma = \hat{\lambda}_2 \approx 1$). As such, this observation suggests $FPLD_{MoM2}$ and $FPLD_{MoM3}$ are poor models of the Standard Normal, a deduction corroborated by Figure 1(b) and the "Moments" rows in Table 3. Given $FPLD_{MoM2}$ and $FPLD_{MoM3}$ are essentially the same fit, both models will be depicted by the same plot instead of drawing two separate diagnostic figures.

Observe in Figure 1(a) that the p.d.f.s of the $N(0,1)$ and $FPLD_{MoM1}$ are almost identical. Here, the fitted p.d.f peaks slightly above the Standard Normal's curve towards the center of the distribution. In contrast, Figure 1(b) shows the fitted FPLD models are leptokurtic which accounts for the great disparity between the shapes of the $N(0,1)$ p.d.f and its approximations $FPLD_{MoM2}$ and $FPLD_{MoM3}$. Differences between the p.d.f.s and quantiles are furthermore captured by the numerical measures presented in Table 3. As the "Moments" rows in Table 3 suggest, of the three fits obtained from the method of moments estimation, the first fit $FPLD_{MoM1}$ appears to be the best model of the Standard Normal distribution because its goodness of fit metric was the smallest. This deduction is further supported by the respective Q-Q plots. For $FPLD_{MoM1}$ (see Figure 1(a)), the quantiles mostly fall on the straight line $y = x$, whereas for $FPLD_{MoM2}$ and $FPLD_{MoM3}$ (see Figure 1(b)), the plot oscillates or "snakes" (albeit tightly) about the said line. Hence, according to the method of moments, the estimate $\hat{\lambda} = (0, 1.3665, 0, 0.1349, 0.1349)$ renders the best approximation of Standard Normal distribution by FPLD.

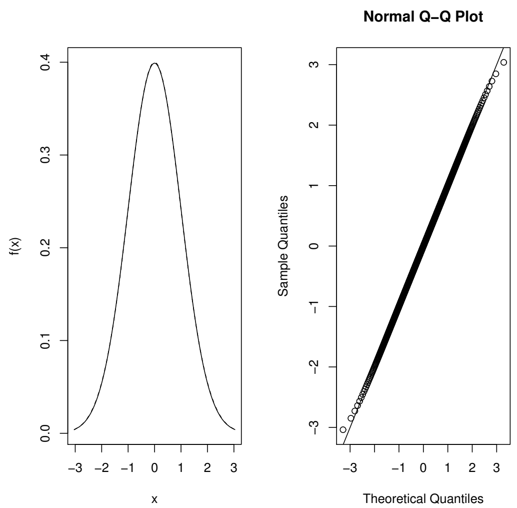


(a) Graph of pd.f.s for the $N(0,1)$ and $FPLD_{MoM1}$ (left); associated Q-Q plot (right)

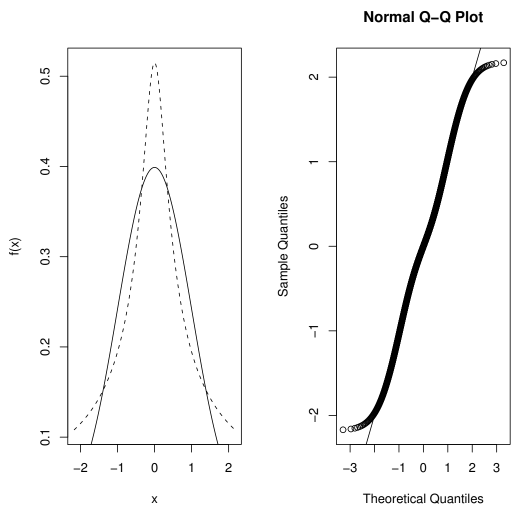


(b) Graph of pd.f.s for the $N(0,1)$, $FPLD_{MoM2}$ and $FPLD_{MoM3}$ (left); associated Q-Q plot (right).

Figure 1: Graphical assessment of the FPLD fit to the Standard Normal via the method of moments.



(a) Graph of pd.f.s for the $N(0,1)$, $FPLD_{PWM1}$ and $FPLD_{LMOM1}$ (left); associated Q-Q plot (right).



(b) Graph of pd.f.s for the $N(0,1)$, $FPLD_{PWM2}$ and $FPLD_{LMOM2}$ (left); associated Q-Q plot (right).

Figure 2: Graphical assessment of the FPLD fit to the Standard Normal via the method of PWMs and L-moments.

Table 3: Numerical measures of the goodness of fit of the FPLD to $N(0,1)$

Method	Fit #	Proximity of p.d.f.s
Moments	1	0.0028201
	2	0.34101
	3	0.34100
PWMs	1	0.00097457
	2	0.11596
L-Moments	1	0.00097362
	2	0.11597

Using this estimate the procedures `fitStdNormPWM` and `fitStdNormLMOM` each yielded two viable FPLD models. They are

$$FPLD_{PWM1}(0, 1.3794, 0, 0.1416, 0.1416)$$

and

$$FPLD_{PWM2}(0, 18.5495, 0, 4.2557, 4.2557),$$

for the method of probability weighted moments and,

$$FPLD_{LMOM1}(0, 1.3794, 0, 0.1416, 0.1416)$$

and

$$FPLD_{LMOM2}(0, 18.5494, 0, 4.2557, 4.2557),$$

for the method of linear moments. Since both methods are theoretically linear equivalent, it is no surprise that these estimates are essentially the same. The closeness between the respective p.d.f.s and quantile functions is again quantified in Table 3. As the indices in "PWMs" and "L-moments" rows of Table 3 show, $FPLD_{PWM1}$ and $FPLD_{LMOM1}$ improve on the $FPLD_{MoM1}$ fit to give a better approximation of the Standard Normal distribution. This fitting improvement is exemplified in Figure 2(a) where the fitted p.d.f.s are practically indiscernible from the Standard Normal density curve. On the other hand, the second pair of fits $FPLD_{PWM2}$ and $FPLD_{LMOM2}$ were found to be less desirable for the same reasons given in the case of $FPLD_{MoM2}$ and $FPLD_{MoM3}$. In spite of this, it must be pointed out that both fits still managed to close the gap between the p.d.f.s. Referring to Figure 1(b) and Figure 2(b), observe that the peak of the fitted density plot fell from just above 0.7 to around 0.5 and is noticeably closer to the maximum of $N(0,1)$. Once again, the linearity of the Q-Q plots for $FPLD_{PWM1}$ and $FPLD_{LMOM1}$ (see Figure 2(a)) strongly suggest they are the better estimates of the Standard Normal distribution for the respective estimation techniques. Thereby, according to the methods of PWMs and L-moments, the estimate $\hat{\lambda} = (0, 1.3794, 0, 0.1416, 0.1416)$ yields the better FPLD approximation.

Ultimately, in the case of the Standard Normal distribution, the method of moments was ruled inferior to the alternative methods used to estimate the FPLD distribution. Consequently, $N(0,1) \approx FPLD(0, 1.3794, 0, 0.1416, 0.1416)$. In other words, the FPLD imitates the shape of the Standard Normal distribution whenever its parameters assume the PWMs and L-moments estimate. This conclusion was deduced from, and supported by the selected goodness of fit tests. Table 4 provides further verification of its reliability with a comparison of the cumulative probabilities from the theoretical Standard Normal distribution and its approximating FPLD model at several values of a random variable X.

It remains inconclusive as to whether PWMs or L-moments provided the closest fit to $N(0,1)$ because neither method dominated by having the lowest indices across their respective rows in Table 3. Consequently, since neither of the first fits $FPLD_{PWM1}$ or $FPLD_{LMOM1}$were prominently ahead of the other, either estimation technique can be thought of as the best candidate for the job of modelling the Standard Normal distribution.

Table 4: Comparison of probabilities from $N(0,1)$ and its FPLD approximation

	Cumulative Probability	
X	Exact	Approximated
-2.77	0.0028	0.0026
-0.5	0.3085	0.3082
0	0.5000	0.5000
0.3	0.6179	0.6181
1.44	0.8729	0.8727
3	0.9987	0.9988

6 Conclusion

The FPLD is an extremely useful model to have in a researcher's toolbox when conducting empirical work such as statistical modelling in the area of data analysis because its five parameters makes it fully capable of succinctly expressing the essential features of many statistical distributions. It has also been employed across a diverse number of situations in applications ranging from as far as Monte Carlo simulations to Q-Q plotting due to the flexible nature of both its quantile and probability density functions (Tarsitano, 2010).

In spite of this, the literature does not currently address how to use well-established estimation techniques to estimate the parameters of this distribution, to be more specific the Gilchrist (2000) parametrisation of the FPLD quantile function. Thus the purpose of this work was to demonstrate how to execute a parametric estimation of the FPLD by applying the methods of moments, probability weighted moments and linear moments. Moreover, the aim was to determine whether any particular method was superior to the others. To facilitate this objective, the FPLD was fitted to the Standard Normal distribution whereby the determination of the appropriateness of a fit was ascertained from selected goodness of fit tests, namely Q-Q plots, p.d.f overplots and the distance between superimposed p.d.f.s.

The results indicated that both the methods of PWMs and L-moments were better than the method of moments at fitting the estimated FPLD function to the chosen theoretical distribution. This was evident from the low metrics of the proximity between p.d.f.s in addition to the linearity of the Q-Q plots. However, neither method clearly outperformed the other, and so they were instead deemed to perform equally well in terms of their ability to approximate the Standard Normal distribution. As a result, even though the method of moments was useful for obtaining initial values for the optimisation schemes for PWMs and L-moments, it was overall the worst performer among the selected estimation methods. It is therefore the conclusion of this work that the methods of PWMs and L-moment are both superior to the method of moments for the purpose of estimating the FPLD.

In conclusion, the above findings show that the FPLD is extremely useful for applications in statistical analysis given the flexible nature of its p.d.f. It is therefore recommended that the FPLD be employed in practical experiments to model both theoretical and empirical distributions. In the latter case, particularly if the underlying distribution is unknown, it is highly recommended that the user conduct a preliminary analysis of the data, for example consult a histogram, to identify a possible statistical model before executing the methodology outline in this work. Furthermore, in light of the fact that neither of the PWMs nor L-moments methods seemed to outperform each other, the authors recommend the method of L-moments as the preferred method when fitting the FPLD to any statistical distribution. This is because it was computationally easier to apply and the estimates obtained can be interpreted to give the defining characteristics of the underlying distribution, such as the mean, variance or skewness.

References

Gilchrist, W. G. (2000). *Statistical modelling with quantile functions*. Chapman & Hall/CRC, Boca Raton.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15:1049–1054.

Hosking, J. (1986). *The theory of probability weighted moments*. IBM Research Division, Yorktown Heights.

Hosking, J. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):105–124.

Ramberg, J. S. and Schmeiser, B. W. (1974). An approximate method for generating asymmetric random variables. *Communications of the ACM*, 17(2):78–82.

Tarsitano, A. (2005). Estimation of the generalized lambda distribution parameters for grouped data. *Communications in Statistics: Theory and Methods*, 34(8):1689–1709.

Tarsitano, A. (2010). Comparing estimation methods for the FPLD.