Moments of $L$-statistics: A Divided Difference Approach

By

Girdhar. G. Agarwal

and

Rashmi Pant
Department of Statistics
Lucknow University
Lucknow
India

Corresponding author:
Rashmi Pant
e-mail: rashmi6380@yahoo.co.uk
Moments of Some $L$-statistics: A Divided Difference Approach

ABSTRACT

Owing to their extensive application in the field of robust estimation, the derivation of the distributions of linear combinations of order statistics or $L$-statistics and the
computation of their moments has been approached in several ways. In this paper we use the properties of divided differences to obtain expressions for moments of some order statistics which arise as special cases of L-statistics. Expectations of some well known L-statistics such as trimmed mean and winsorised mean for the pareto distribution are computed. The study also undertakes the computation of L-moments which are expectations of certain linear combinations of order statistics. The algorithms have been implemented using some well known continuous distributions as examples. The results are compared with known values wherever they are available in literature.

1. INTRODUCTION

Suppose \( X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)} \) are the order statistics from a population with continuous distribution function \( F(x) \) and \( \{a_i\}_{i=1}^n \) is a sequence of real numbers. The statistic

\[
L = \sum_{i=1}^{n} a_i X_{(i)}
\]  

(1.1)

is called a linear combination of order statistics (or L-statistics) from the distribution \( F \).

The choice of the constants \( a_i \)'s in (1.1) determines the properties and functionality of \( L \).

L-statistics have witnessed prolific development in the field of robust estimation (Tukey, 1960), (Crow & Siddiqui, 1967), (Hodges, 1967). L-statistics have been studied by many authors for different sets of \( a_i \)'s and for different random variables (Olds, 1952; Dempster & Kleyle, 1968; Stigler,1974). With the help of a generalized form of the Gauss-Markov theorem, Lloyd (1952) has used L-statistics for the estimation of location and scale parameters. This approach has been found to be immensely useful in the face of
non-normality of the distribution under study. Eisenberger & Posner (1965) have employed $L$-statistics for the compression of data in space telemetry. Welsh and Morrison (1990) have applied $L$-estimation for the study of the evolution of our galaxy using stellar velocity distributions. Sample sizes are small in such studies and there is a need for robust techniques because velocity distributions are not always Gaussian. In censored sampling from the uniform distribution on the interval $(\alpha - \beta, \alpha + \beta)$, the best linear unbiased estimators of $\alpha$ & $\beta$ are linear functions of Uniform variates (Aggarwal and Balakrishnan, 1998). Some recent applications of $L$-statistics include use in stratified multistage sampling (Shao, 1994), study of the visual acuity of patients with certain eye disorders (Viana, 1998), analysis of cross-over designs in clinical trials. (Putt and Chinchilli, 2000) and analysis of control charts (Elamir and Seheult, 2001).

Owing to various computational difficulties, exact expressions for the moments of $L$-statistics are not always available. Several methods have been used to study the moments of $L$-statistics such as bootstrapping (Houston and Ernst, 2000), bounds and approximations (Rychlik, 2004). Arnold and Groeneveld’s (1979) pioneering work on the bounds of expectations of $L$-statistics paved the way for subsequent work (Aven, 1985), (Lefevre, 1986), (Balakrishnan et al, 2003), (Papadatos and Rychlik, 2004).

A very interesting approach to the study of $L$-statistics is the representation of B-splines as the densities of linear combinations of order statistics from the uniform distribution on the interval $(0,1)$ (called LCUOS). This was given independently by Karlin and Micchelli (1986) and Ignatov and Kaishev (1989).

We give a brief outline of their method.

The density $f_L(x)$ of the quantity in (1.1) is given by
\[ M_n(x; x_0, x_1, \ldots, x_n) = n[x_0, x_1, \ldots, x_n](-x)^{n-1} \]  \hspace{1cm} (1.2)

where \( x_0, x_1, \ldots, x_n \) are specified by the relation

\[ x_i = \sum_{j=0}^{i} a_{n+1-j} \quad , i = 0, \ldots, n \]  \hspace{1cm} (1.3)

The function in (1.2) is the B-spline introduced by (Curry and Schoenberg, 1966) and the right hand side of (1.2) is the divided difference of order \( n \) of the truncated power function \( g(y) = n(y-x)^{n-1} \) at the points \( y = x_0, x_1, \ldots, x_n \).

Subsequently expressions for moments of LCUOS were obtained by Agarwal et al (2002).

In the next section we discuss and use their results to obtain the expressions for moments of some well known \( L \)-statistics from some continuous distributions.

### 2. MOMENTS AND DIVIDED DIFFERENCES

**Theorem 1:** (Agarwal et al, 2002): The \( m \)-th raw moment of the \( r \)-th order statistic \( X_{(r)} \) is given by

\[ \mu_{r,m} = n! [x_0, x_1, \ldots, x_n] h \quad , r=1, \ldots n; \quad m=1,2, \ldots \]

where \( h^{(m)}(x) = [G(x)]^m \) and \( G(x) \) is the inverse cumulative distribution function.

**Proof:** If the random variable \( X \) has the cumulative distribution function \( F(x) \), then it is well known that the probability integral transform \( Y = F(X) \) has the uniform distribution over \((0,1)\). This transform is order preserving so we have for the order statistics \( X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)} \) from an arbitrary distribution with continuous distribution function \( F(x) \) the order statistics \( Y_{(i)} = F(X_{(i)}) \) such that
$Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(n)}$ are order statistics from the uniform $(0,1)$ distribution.

The $m$-th moment of the $r$-th order statistic is defined as

$$E(X_{r,n}^m) = \mu_{r,n}^m = \int_0^1 [F^{-1}(x)]^m f_{Y_{(r)}}(x) \, dx$$

; $r=1,2,\ldots,n$ & $m=1,2,\ldots$

where $f_{Y_{(r)}}(x)$ is the density of $r$-th order statistic from U(0,1).

We know from (1.2) $f_{Y_{(r)}}(x)$ can be represented by the B-spline $M_n(x; x_0,\ldots,x_n)$. Hence

$$\mu_{r,n}^m = E[(X_{(r)})^m] = E[(G(Y_{(r)}))^m] = \int_0^1 [G(x)]^m M_n(x; x_0,\ldots,x_n) \, dx$$

where $G \equiv F^{-1}$.

A simple application of Peano kernel theorem (Davis, 1963, page 69) gives us

$$\mu_{r,n}^m = n! [x_0, x_1, \ldots, x_n] h$$

(2.1)

where $h^{(m)}(x) = [G(x)]^m$.

Although the authors do not state it, we wish to emphasize here that the $x_i$'s in (2.1) will always be a sequence of 0's and 1's depending on the value of $n$ and $r$.

This is easily seen since here

$$L = X_{(i)}$$

and on comparison with (1.1) we see that

$$a_j = \begin{cases} 0 & \text{for } j \neq i \\ 1 & \text{for } j = i \end{cases}$$

Using (1.2), this gives us

$$x_i = \begin{cases} 0 & \text{for } i = 0,\ldots,n-r \\ 1 & \text{for } i = n-r+1,\ldots,n \end{cases}$$
Hence (2.1) can be written as

\[ \mu_{r,n}^{m} = n! [0, \ldots, 0, 1, \ldots, 1]_{h}^{r} \quad , \quad r=1, \ldots, n \quad \& \quad m=1, 2, \ldots \]  

(2.2)

The divided difference on the right side of equation (2.2) can be explicitly evaluated using various representations (Davis, 1963). We use the following recursion relations which express higher order differences in terms of the differences of lower order (de Boor, 1978).

\[
[x_{0}, x_{1}, \ldots, x_{n}]h = \begin{cases} 
\frac{h^{(a)}(x) / n!}{x_{0} - x_{n}} & \text{if } x_{0} = \ldots = x_{n} = x \text{ and } h \in C^{(a)} \\
\frac{[x_{0}, \ldots, x_{r-1}, x_{r+1}, \ldots, x_{n}]h - [x_{0}, \ldots, x_{r-1}, x_{r+1}, \ldots, x_{n}]h_{x_{r}} - x_{r}}{x_{s} - x_{r}} & \text{if } x_{s} \neq x_{r}
\end{cases}
\]

(2.3)

3. MOMENTS OF EXTREME ORDER STATISTICS

1. Smallest order statistic

Using the above approach the m-th order moment for the smallest order statistic \( X_{(1)} \) from a continuous distribution \( F(x) \) would be

\[
E(X_{(1)}^{m}) = \mu_{1,n}^{m} = n! [0, 0, \ldots, 1]_{h}
\]

\[
= n! \left[ [0, 0, \ldots, 0, 1]_{n} - [0, 0, \ldots, 0]_{n} \right] \frac{1}{1 - 0}
\]

\[
= n! \left[ [0, 0, \ldots, 0, 1]_{n} - h^{(n-1)}(0) \right] \frac{1}{(n-1)!}
\]

Applying (2.3) recursively we have
\[= n! \left\{ [0,1]h - [0,0]h - [0,0,0]h - \ldots - [0,\ldots,0]h \right\}\]

\[\Rightarrow \mu_{n,n}^m = n! \left\{ h(1) - h(0) - \sum_{i=1}^{n-1} \frac{h^{(i)}(0)}{i!} \right\}\]

(3.1)

2. Largest Order Statistic

The \(m\)-th order moment for \(X_{(n)}\) is

\[E(X_{n,n}^m) = \mu_{n,n}^m = n![0,1,\ldots,1]h\]

Where \(h^{(m)}(x) = [G(x)]^m \quad m=1,2,\ldots\)

Using property (2.3) of divided differences, we have

\[\mu_{n,n}^m = n! \left\{ \frac{[1,\ldots,1]h - [0,1,\ldots,1]h}{1 - 0} \right\}\]

Proceeding recursively we have

\[\mu_{n,n}^m = n! \left\{ \frac{[1,\ldots,1]h - [1,\ldots,1]h + [1,\ldots,1]h - \ldots + (-1)^{n-2}[1,1]h + (-1)^{n-1}[0,1]h}{n-1} \right\}\]

\[= n! \left\{ \frac{h^{(n-1)}(1)}{(n-1)!} - \frac{h^{(n-2)}(1)}{(n-2)!} + \frac{h^{(n-3)}(1)}{(n-3)!} - \ldots + (-1)^{n-2} \frac{h^{(1)}(1)}{1!} + (-1)^{n-1}(h(1) - h(0)) \right\}\]

Hence

\[\mu_{n,n}^m = n! \left\{ \sum_{i=1}^{n} (-1)^{i-1} \frac{h^{(n-i)}(1)}{(n-i)!} + (-1)^n h(0) \right\}\]

(3.2)
In general the $m$-th moment for the $r$-th order statistic can be written as

$$
\mu_{r,m}^m = n! \left\{ \frac{(-1)^r}{(r-1)!} \sum_{k=0}^{n-r} \frac{(n-k)!}{(n-r-k)!k!} h^{(k)}(0) + \sum_{k=0}^{r-1} \frac{(-1)^{r-1-k}}{k!(r-1-k)!} \frac{(n-k+1)!}{h^{(k)}(1)} \right\}
$$

(3.3)

In the following sections we have implemented the algorithms proposed in section 3 using computer programs written in the Matlab programming language. Sample programs are provided in the appendix.

4. EXAMPLE

1. Pareto Distribution

Many socio-economic and other naturally occurring quantities are distributed according to certain statistical distributions with very long tails. Examples of some of these empirical phenomena are distributions of city population sizes, occurrence of natural resources, stock price fluctuations, insurance risk, word frequencies, business mortality, personal income, service times in queuing systems, error clustering in communication circuits, and distribution of oil fields in specific areas by size (Goldberg, 1967) etc. The Pareto distribution has played a key role in the study of such phenomena.

The distribution function studied here is

$$
F_x(x) = 1 - \left( \frac{\gamma}{x} \right)^\alpha, \quad \gamma, \alpha > 0 \text{ & } x \geq \gamma
$$

(4.1)

The inverse distribution function is

$$
G(x) = \left( \frac{\gamma}{1-x} \right)^{1/\alpha}, \quad \gamma, \alpha > 0 \text{ & } 0 < x < 1
$$
Malik (1966) has published tables for the moments of order statistics from the Pareto distribution for specific sample sizes \( n \leq 12 \). We present similar tables (4.1) and (4.2) for the mean and variance of the extreme order statistics, but for different values of the parameters and sample sizes upto 20. The maximum error is of the order of \( 10^{-11} \) for the mean and of \( 10^{-8} \) for the variance.

**Table 4.1:** Mean and variance of smallest order statistic from pareto distribution with parameters \( \gamma=2 \) and \( \alpha=2 \)

<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.66666667</td>
<td>0.88888887</td>
</tr>
<tr>
<td>4</td>
<td>2.28571429</td>
<td>0.1084354</td>
</tr>
<tr>
<td>6</td>
<td>2.18181818</td>
<td>0.0396947</td>
</tr>
<tr>
<td>8</td>
<td>2.13333333</td>
<td>0.02031774</td>
</tr>
<tr>
<td>10</td>
<td>2.10526316</td>
<td>0.01231258</td>
</tr>
<tr>
<td>12</td>
<td>2.08695652</td>
<td>0.00825216</td>
</tr>
<tr>
<td>14</td>
<td>2.07407407</td>
<td>0.00591727</td>
</tr>
<tr>
<td>16</td>
<td>2.06451613</td>
<td>0.00445760</td>
</tr>
<tr>
<td>18</td>
<td>2.05714286</td>
<td>0.00349206</td>
</tr>
<tr>
<td>20</td>
<td>2.05128205</td>
<td>0.00283080</td>
</tr>
</tbody>
</table>

**Table 4.2:** Mean and variance of largest order statistic from pareto distribution with parameters \( \gamma=2 \) and \( \alpha=4 \)

<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3.04761888</td>
<td>1.37852584</td>
</tr>
<tr>
<td>4</td>
<td>3.54632001</td>
<td>2.05186582</td>
</tr>
</tbody>
</table>
6  3.89527127  2.55798346
8  4.16984820  2.98033512
10 4.39896061  3.35020096
12 4.59703342  3.68325883
14 4.77239250  3.98867393
16 4.93031658  4.27234821
18 5.07438077  4.53834469
20 5.20712816  4.78960088

5. EXPECTATION OF THE TRIMMED MEAN AND WINSORISED MEAN

Increasing interest in robust estimation of location parameters is spurred by the fact that the sample mean although efficient under normality assumptions, is seriously affected by slight non-normality, being particularly sensitive to outliers (extreme observations). Slight departure from normality towards a heavy-tailed distribution can substantially lower the power of methods for comparing means because heavy-tailed distributions inflate the standard error of the mean. Some important $L$-statistics that are used in robust estimation and study of outliers are the trimmed mean and the winsorised mean. One of the simplest ways to make the arithmetic mean less sensitive to outliers is to delete or ‘trim’ a proportion of ordered observations on one or both sides of the sample.

The $k$-th trimmed mean is defined as

$$L = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} X_{(i)}$$
Table 5.1 below gives the expectations of the trimmed mean for the pareto distribution with parameters $\gamma=2$ and $\alpha=2$. The maximum error is of the order of $10^{-13}$. 
Table 5.1: Expected values of $k$-th trimmed mean for sample size 4(2)20 and parameters $\gamma=2$ and $\alpha=2$.

<table>
<thead>
<tr>
<th></th>
<th>n=4</th>
<th>n=6</th>
<th>n=8</th>
<th>n=10</th>
<th>n=12</th>
<th>n=14</th>
<th>n=16</th>
<th>n=18</th>
<th>n=20</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.9486</td>
<td>2.9613</td>
<td>2.9809</td>
<td>3.0026</td>
<td>3.0244</td>
<td>3.0455</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.9264</td>
<td>2.9360</td>
<td>2.9515</td>
<td>2.9694</td>
<td>2.9879</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.9111</td>
<td>2.9186</td>
<td>2.9312</td>
<td>2.9462</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.9000</td>
<td>2.9059</td>
<td>2.9164</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.8915</td>
<td>2.8964</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.8848</td>
<td></td>
</tr>
</tbody>
</table>
The $k$-th Winsorised mean is

$$L = \frac{(k + 1)(X_{(k+1)} + X_{(n-k)}) + \sum_{i=k+2}^{n-k-1} X_{(i)}}{n} \quad 0 < k < \frac{n - 1}{2}$$

where $n$ is even.

The expectations of $L$ for different sample sizes and all values of $k$ are computed in table 5.2 and 5.3. The exponential distribution with parameter 1 and the pareto distribution considered in (4.1) serve as the examples. The maximum error is of the order of $10^{-11}$ for the pareto distribution and of $10^{-15}$ for the exponential distribution.
Table 5.2: Expectations of the Winsorised mean for the exponential distribution with parameter 1

<table>
<thead>
<tr>
<th>n=8</th>
<th>n=10</th>
<th>n=12</th>
<th>n=14</th>
<th>n=16</th>
<th>n=18</th>
<th>n=20</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.07613</td>
<td>12.26338</td>
<td>18.63402</td>
<td>25.17515</td>
<td>31.86190</td>
<td>38.67107</td>
<td>45.58341</td>
</tr>
</tbody>
</table>

- | 7.45628 | 15.01646 | 22.76491 | 30.70290 | 38.81234 | 47.07224 |
- | - | 8.83845 | 17.77305 | 26.89410 | 36.21443 | 45.72347 |
- | - | - | 10.22178 | 20.53245 | 31.02456 | 41.71966 |
- | - | - | - | 11.60585 | 23.29406 | 35.15774 |
- | - | - | - | - | 12.99039 | 26.05681 |
- | - | - | - | - | - | 14.37570 |
Table 5.3: Expectations of the Winsorised mean for the pareto distribution with parameters $\gamma = 2$ and $\alpha = 2$

<table>
<thead>
<tr>
<th>n=8</th>
<th>n=10</th>
<th>n=12</th>
<th>n=14</th>
<th>n=16</th>
<th>n=18</th>
<th>n=20</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.9680</td>
<td>39.9600</td>
<td>47.9520</td>
<td>55.9440</td>
<td>63.9360</td>
<td>71.9280</td>
<td>79.9200</td>
</tr>
<tr>
<td>39.3648</td>
<td>53.0878</td>
<td>67.0094</td>
<td>81.0866</td>
<td>95.2897</td>
<td>109.5970</td>
<td>123.9926</td>
</tr>
<tr>
<td>36.8784</td>
<td>55.9178</td>
<td>75.3422</td>
<td>95.0844</td>
<td>115.0908</td>
<td>135.3199</td>
<td>155.7392</td>
</tr>
<tr>
<td>23.8695</td>
<td>48.0198</td>
<td>72.6378</td>
<td>97.7017</td>
<td>123.1589</td>
<td>148.9581</td>
<td>175.0554</td>
</tr>
<tr>
<td>-</td>
<td>29.4858</td>
<td>59.2251</td>
<td>89.4265</td>
<td>120.1021</td>
<td>151.2177</td>
<td>182.7302</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>35.1170</td>
<td>70.4630</td>
<td>106.2553</td>
<td>142.5319</td>
<td>179.2762</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>40.7560</td>
<td>81.7203</td>
<td>123.1113</td>
<td>164.9860</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>46.3999</td>
<td>92.9903</td>
<td>139.9869</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>52.0468</td>
<td>104.2689</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>57.6958</td>
</tr>
</tbody>
</table>
6. COMPUTING L-MOMENTS

Conventional moments are plagued by two major drawbacks. First, they do not always impart easily interpreted information about shape of the distribution (especially moments of third and higher order). Secondly estimates of parameters of distributions fitted by moments are often less accurate than those fitted by other methods such as maximum likelihood. Hosking (1990) introduced $L$-moments which are used for the characterization of distributions, the summarization of observed data sets, the fitting of probability distributions to observed samples and the testing of hypotheses for distributional form. A formal definition follows.

**Definition 6.1:** Suppose $X$ is a real-valued random variable with cumulative distribution function $F(x)$ and a finite mean. Let $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ be the order statistics of a random sample of size $n$ drawn from $F$. The $r$-th $L$-moment is defined as

$$
\lambda_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} EX_{(r-k)r} \quad r = 1, 2, \ldots n
$$

(6.1)

Thus the first four moments are:

$$
\lambda_1 = E(X)
$$

$$
\lambda_2 = \frac{1}{2} [E(X_{(2)} - X_{(1)})]
$$

$$
\lambda_3 = \frac{1}{3} [E(X_{(3)} - 2X_{(2)} + X_{(1)})]
$$

(6.2)

$$
\lambda_4 = \frac{1}{4} [E(X_{(4)} - 3X_{(3)} + 3X_{(2)} - X_{(1)})]
$$
Higher order $L$-moments are expressed as $L$-moment ratios to make them independent of the unit of measurement.

$$\tau_r = \frac{\lambda_r}{\lambda_2}, \quad r \geq 3 \quad (3.5.3)$$

$L$-moments are thus linear combinations of order statistics. $\lambda_1, \lambda_2, \tau_3, \tau_4$ are regarded as measures of location, scale, skewness and kurtosis. Furthermore, a distribution having finite mean is uniquely determined by its $L$-moments (Hosking, 2006). The main advantage of $L$-moments over conventional moments is that they suffer less from the effects of sampling variability and they are more robust to outliers in data. Better inferences about the underlying probability distribution can be made even when small samples are available.

We can now use the method of divided differences to compute the $L$-moments for some well known distributions and the results when compared with those obtained by Hosking (1990), are found to be fairly accurate.

Table 6.1: $L$-moments of some continuous distributions computed using the divided difference method.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\tau_3$</th>
<th>$\tau_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>0</td>
<td>0.551281459</td>
<td>0</td>
<td>0.166635306</td>
</tr>
<tr>
<td>Generalised</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extreme-Value</td>
<td>-0.249945375</td>
<td>0.374944375</td>
<td>-0.629571349</td>
<td>0.398052562</td>
</tr>
<tr>
<td>Exponential</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>1.999932738</td>
<td>0.666643404</td>
<td>0.333332913</td>
<td>0.166595305</td>
</tr>
<tr>
<td>0.5</td>
<td>0.999966369</td>
<td>0.333321702</td>
<td>0.333332913</td>
<td>0.166635306</td>
</tr>
<tr>
<td>0.75</td>
<td>0.666644246</td>
<td>0.222214468</td>
<td>0.333332913</td>
<td>0.166635306</td>
</tr>
<tr>
<td>1</td>
<td>0.499983184</td>
<td>0.166660851</td>
<td>0.333332913</td>
<td>0.166635306</td>
</tr>
<tr>
<td>1.25</td>
<td>0.399986548</td>
<td>0.133328681</td>
<td>0.333332913</td>
<td>0.166635306</td>
</tr>
<tr>
<td>1.5</td>
<td>0.333322123</td>
<td>0.111107234</td>
<td>0.333332913</td>
<td>0.166635306</td>
</tr>
<tr>
<td>2</td>
<td>0.249991592</td>
<td>0.083330426</td>
<td>0.333332913</td>
<td>0.166635306</td>
</tr>
<tr>
<td>Pareto</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma=2,a=2$</td>
<td>3.995998</td>
<td>1.329327328</td>
<td>0.332664663</td>
<td>1.329327328</td>
</tr>
<tr>
<td>$\gamma=4,a=2$</td>
<td>7.991996</td>
<td>2.658654656</td>
<td>0.332664663</td>
<td>0.199196792</td>
</tr>
</tbody>
</table>
Theoretically L-moments can characterize a wider range of distributions than conventional moments (Hosking, 2006). Practically they are less subject to bias in estimation and they approximate their asymptotic normal distribution more closely.

REFERENCES


Hosking, J.R.M. (1992), Moments or L-moments? An Example Comparing two measures of Distribution Shape. American Statistician, 46,[3]:186-189


APPENDIX

1. Program to compute the mean and variance of smallest order statistic from the pareto distribution using equation (3.1).

```matlab
clear;
syms x;                  % symbolic variable
k=2;a=2;                 % parameters of distribution
mu1=zeros(1,10);         % first raw moment
mu2=zeros(1,10);         % second raw moment
for m=1:2
    counter=1;
    Gx=(k/(1-x)^(1/a))^m;        % inverse distribution function
    for n=2:2:20
        h1=zeros(1,n+1);         % value of hx at x=1
        h0=zeros(1,n+1);         % value of hx at x=0
        hx(1)=Gx;
        f=inline(char(hx(1)));  
        h1(n+1)=f(0.9999999999);
        h0(n+1)=f(0);
        for i=2:n+1
            hx(i)=int(hx(i-1));
            f=inline(char(hx(i)));  
            h1(n+2-i)=f(0.9999999999);
            h0(n+2-i)=f(0);
        end
    end
    %Values of the moment of smallest order statistic calculated using
    %equation 3.1
    if m==1
        for j=2:n
            mu1(counter)=mu1(counter)+h0(j)/factorial(j-1);
        end
        mu1(counter)=factorial(n)*(h1(1)-h0(1)-mu1(counter));
    else
        for j=2:n
            mu2(counter)=mu2(counter)+h0(j)/factorial(j-1);
        end
        mu2(counter)=factorial(n)*(h1(1)-h0(1)-mu2(counter));
    end
    obtained value of second moment
end
counter=counter+1;
end
```

% Theoretical Results and Errors %
for c=1:10
    theory1(c)=((gamma(2*c+1)*gamma(2*c-(1/a)))/(gamma(2*c-(1/a)+1)*gamma(2*c)))*k;  % theoretical value of mean
    error1(c)=abs(theory1(c)-mu1(c));
    pererror1(c)=error1(c)/theory1(c)*100;
    var(c)=mu2(c)-(mu1(c))^2;         % variance
    theory2(c)=((gamma(2*c+1)*gamma(2*c-(2/a)))/(gamma(2*c-(2/a)+1)*gamma(2*c)))*(k^2);  % theoretical value of second
    tvar(c)=theory2(c)-theory1(c)^2;                       % theoretical value of variance
    error2(c)=abs(tvar(c)-var(c));
    pererror2(c)=error2(c)/tvar(c)*100;
end

% Print Results %
for e=1:10
    disp(sprintf('%2d %12.8f %12.8f %15.8f %15.8f',2*e,mu1(e),theory1(e),error1(e),pererror1(e)))
end

2. Program to compute the trimmed mean from the pareto distribution using equation (3.3).

% trimexp.m
 clear;
 sym x;  % symbolic variable
 m=1;k=2;a=2;  % order of the moment
 Gx=(k/(1-x)^(1/a))^m;  % inverse distribution function
 E=zeros(10,9);
 Et=zeros(10,9);
 counter=1;
 for n=4:2:20
    h1=zeros(1,n+1);  % value of hx at x=1

\begin{verbatim}
% value of \( h_x \) at \( x=0 \)
% \( h_x(1) = G_x \);
\texttt{f} = \texttt{inline(char(\texttt{hx}(1)))};
\texttt{h}1(\texttt{n+1}) = \texttt{f}(0.999999);
\texttt{h}0(\texttt{n+1}) = \texttt{f}(0.000001);
\texttt{for} \texttt{i}=\texttt{2} : \texttt{n+1}
    \texttt{hx(i)} = \texttt{int(hx(i-1))};
    \texttt{f} = \texttt{inline(char(hx(i)))};
    \texttt{h}1(n+2-i) = \texttt{f}(0.999999);
    \texttt{h}0(n+2-i) = \texttt{f}(0.000001);
\texttt{end}
\texttt{r}=\texttt{0};
\texttt{while} (\texttt{r} < \texttt{n/2})
    \texttt{for} \texttt{rr}=\texttt{r+1} : \texttt{n-r}
        \texttt{mu}0=\texttt{0}; \texttt{mu}0=\texttt{0};
        \texttt{for} \texttt{k}=\texttt{0} : \texttt{n-rr}
            \texttt{mu}0=\texttt{mu}0+\texttt{h}0(k+1)*\texttt{factorial(n-k-1)}/\texttt{factorial(n-rr-k)}*\texttt{factorial(k)};
        \texttt{end}
        \texttt{for} \texttt{k}=\texttt{0} : \texttt{rr-1}
            \texttt{mu}1=\texttt{mu}1+(-1)^{\texttt{rr-1-k}}*\texttt{factorial(n-k-1)}/\texttt{factorial(k)*factorial(rr-1-k)}*\texttt{h}1(k+1);
        \texttt{end}
        \texttt{mu}(\texttt{rr})=\texttt{factorial(n)*((-1)^rr*mu0/factorial(rr-1))/gamma(n+1)*gamma(n-rr-1-m/a)/gamma(n+rr-m/a)*gamma(n-rr-1)};
        \texttt{E}(r+1,counter)=E(r+1,counter)+\texttt{mu}(\texttt{rr})
        \texttt{Et}(r+1,counter)=Et(r+1,counter)+2*(((gamma(n+1)*gamma(n-rr-1-m/a))/gamma(n+rr-m/a)*gamma(n-rr-1)));
        \texttt{E}(r+1,counter)=1/(n-2*r)*E(r+1,counter);
        \texttt{Et}(r+1,counter)=1/(n-2*r)*Et(r+1,counter);
    \texttt{error}(r+1,counter)=abs(Et(r+1,counter)-E(r+1,counter));
    \texttt{r}=\texttt{r+1};
\texttt{end}
counter=counter+1;
\texttt{end}
\end{verbatim}

3. Program to compute the winsorised mean from the pareto distribution using equation (3.3).

%%%%%%%%%%%%%%%%%%%
\% \texttt{winpareto.m} June 12 2007 \% 
\% Author Rashmi Pant \% 
\% Winsorised mean \% 
\% of the pareto dist. \% 
%%%%%%%%%%%%%%%%%%%
clear;
syms \( x \); \hspace{1cm} \% symbolic variable
\texttt{m}=1;\texttt{k}=2;\texttt{a}=2; \hspace{1cm} \% order of the moment
Gx=(\texttt{k}/(\texttt{1-x})^{(\texttt{1/a})})^{\texttt{m}}; \hspace{1cm} \% value of the function \( G_x \)
E=\texttt{zeros}(10,7);
Et=\texttt{zeros}(10,7);
counter=1;
for n=8:2:20
    h1=zeros(1,n+1);         % value of hx at x=1
    h0=zeros(1,n+1);         % value of hx at x=0
    hx(1)=Gx;
    f=inline(char(hx(1)));
    h1(n+1)=f(0.999999);
    h0(n+1)=f(0.000001);
    for i=2:n+1
        hx(i)=int(hx(i-1));
        f=inline(char(hx(i)));
        h1(n+2-i)=f(0.999999);
        h0(n+2-i)=f(0.000001);
    end
    r=0;
    while (r < (n-1)/2)
        for rr=r+2:n-r-1
            mu=mvalue(rr,n,h1,h0);
            Et(r+1,counter)=Et(r+1,counter)+mu;
            for j=n-rr+1:n
                Et(r+1,counter)=Et(r+1,counter)+1/j;
            end
        end
        mur=mvalue(r+1,n,h1,h0);
        munr=mvalue(n-r,n,h1,h0);
        E(r+1,counter)=(r+1)*(mur+mnr+E(r+1,counter));
        r=r+1;
    end
    counter=counter+1;
end

4. Program to compute the L-moments using the divide difference approach. Sample program for the pareto distribution.

% Program to compute the L-moments using the divide difference approach. Sample program for the pareto distribution.

% lmoments.m
% Compute L-moments
% PARETO

clear;

% symbolic variable
k=4;a=2;
Gx=(k/(1-x))^(1/a));
E=zeros(6,1);
counter=1;
for n=1:6
    h1=zeros(1,n+1);         % value of hx at x=1
    h0=zeros(1,n+1);         % value of hx at x=0
    hx(1)=Gx;
    f=inline(char(hx(1)));
    h1(n+1)=f(0.999999);
    h0(n+1)=f(0.000001);
    for i=2:n+1
        hx(i)=int(hx(i-1));
    end
end
f=inline(char(hx(i)));  
end

h1(n+2-i)=f(0.999999);  
h0(n+2-i)=f(0.000001);

for rr=0:n-1
   mu=muvalue(n-rr,n,h1,h0);
   E(counter,1)=E(counter,1)+(-1)^rr*((factorial(n-1))/(factorial(rr)*factorial(n-1-rr)))*mu;
end

E(counter,1)=1/n*E(counter,1);
counter=counter+1;
end
E

5. Program muvalue.m which computes the m-th moment of r-th order statistic equation (3.3).

function mu = muvalue(n,n,h1,h0);
mu=0; mu1=0; mu0=0;
for k=0:n-rr
   mu0=mu0+h0(k+1)*factorial(n-k-1)/(factorial(n-rr-k)*factorial(k));
end
for k=0:rr-1
   mu1=mu1+(-1)^(rr-1-k)*factorial(n-k-1)/(factorial(k)*factorial(rr-1-k))*h1(k+1);
end
mu=factorial(n)*((-1)^rr*mu0/factorial(rr-1)+mu1/factorial(n-rr));