REVSTAT – Statistical Journal
Volume 16, Number 4, October 2018, 429–443

PARAMETER ESTIMATION FOR THE LOG-LOGISTIC DISTRIBUTION BASED ON ORDER STATISTICS

Authors: Mohammad Ahsanullah
– Department of Management Sciences, Rider University, New Jersey, USA
ahsan@rider.edu

Ayman Alzaatreh
– Department of Mathematics and Statistics, American University of Sharjah, Sharjah, UAE
aalzaatreh@aus.edu

Received: April 2016 Revised: September 2016 Accepted: October 2016

Abstract:

- In this paper, we discuss the moments and product moments of the order statistics in a sample of size n drawn from the log-logistic distribution. We provide more compact forms for the mean, variance and covariance of order statistics. Parameter estimation for the log-logistic distribution based on order statistics is studied. In particular, best linear unbiased estimators (BLUEs) for the location and scale parameters for the log-logistic distribution with known shape parameter are studied. Hill estimator is proposed for estimating the shape parameter.

Key-Words:

- *log-logistic distribution; moments; order statistics; best linear unbiased estimators; recurrence relations; Hill estimator.*

AMS Subject Classification:

- 62E10, 62F10.

1. INTRODUCTION

The probability density function (pdf) of the log-logistic distribution with unit scale parameter is given by

$$f(x) = \frac{\alpha\, x^{\alpha-1}}{(1+x^\alpha)^2}, \quad x \geq 0, \tag{1.1}$$

where α is a positive real number. A random variable X that follows the density function in (1.1) is denoted as $X \sim$ log-logistic(α). The cumulative distribution (cdf) and quantile functions of the log-logistic distribution, respectively, are

$$F(x) = \frac{x^\alpha}{1+x^\alpha}, \quad x \geq 0. \tag{1.2}$$

and

$$F^{-1}(x) = \left(\frac{x}{1-x}\right)^{1/\alpha}, \quad 0 < x < 1. \tag{1.3}$$

The k-th moments of the log-logistic distribution in (1.1) can be easily computed as

$$\mu'_k = B\left(1-\frac{k}{\alpha}, 1+\frac{k}{\alpha}\right), \tag{1.4}$$

where $B(.,.)$ is the beta function.
Note that the k-th moment exists iff $\alpha > k$. A more compact form of (1.4) can be derived using the fact that $\Gamma(z)\,\Gamma(1-z) = \pi\,\csc(\pi\, z)$ (Abramowitz and Stegun, 1964) as follows

$$\mu'_k = \Gamma(1-k/\alpha)\,\Gamma(1+k/\alpha) = \frac{k\pi}{\alpha}\csc\frac{k\pi}{\alpha}, \quad \alpha > k. \tag{1.5}$$

Therefore,

$$E(X) = (\pi/\alpha)\csc(\pi/\alpha) \text{ and } Var(X) = (\pi/\alpha)\{2\csc(2\pi/\alpha) - (\pi/\alpha)\csc^2(\pi/\alpha)\}.$$

The log-logistic distribution is a well-known distribution and it is used in different fields of study such as survival analysis, hydrology and economy. For some applications of the log-logistic distribution we refer the reader to Shoukri *et al.* [23], Bennett [10], Collett [11] and Ashkar and Mahdi [7]. It is also known that the log-logistic distribution provides good approximation to the normal and the log-normal distributions. The log-logistic distribution has been studied by many researchers such as Shah and Dave [22], Tadikamalla and Johnson [24], Ragab and Green [21], Voorn [25] and Ali and Khan [4]. Ragab and Green [21] studied some properties of the order statistics from the log-logistic distribution. Ali and Khan [4] obtained several recurrence relations for the moments of order statistics. Voorn [25] characterized the log-logistic distribution based on extreme related stability with random sample size. In this paper, we discuss the moments of order statistics for the log-logistic distribution. We review some known results and provide a more compact expression for calculating the covariance between two order statistics. Also, we discuss the parameter estimation of the log-logistic distribution based on order statistics.

2. SOME RESULTS FOR THE MOMENTS OF ORDER STATISTICS

Let $X_1, X_2, ..., X_n$ be n independent copies of a random variable X that follows log-logistic(α). Let $\mathrm{X}_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ be the corresponding order statistics. Then from (1.1) and (1.2), the pdf of the r^{th} order statistics is given by

$$f_{r:n}(x) = C_{r:n} \frac{\alpha\, x^{\alpha r-1}}{(1+x^{\alpha})^{n+1}}, \quad x \geq 0, \tag{2.1}$$

where $C_{r:n} = \frac{n!}{(r-1)!(n-r)!}$.
The k^{th} moments of $X_{r:n}$ can be easily derived from (2.1) as

$$\alpha_{r:n}^{(k)} = C_{r:n}\, B\left(n-r+1-\frac{k}{\alpha}, r+\frac{k}{\alpha}\right), \quad \alpha > k, \tag{2.2}$$

Similarly as in (1.5), one can show that

$$\alpha_{r:n}^{(k)} = \frac{(-1)^r \pi \ \csc \frac{k\pi}{\alpha}}{(r-1)!(n-r)!} \prod_{i=1}^{n} \left(i-r-\frac{k}{\alpha}\right), \quad \alpha > k. \tag{2.3}$$

Note that when $r = n = 1$, $\alpha_{1:1}^{(k)} = B\left(1-\frac{k}{\alpha}, 1+\frac{k}{\alpha}\right)$ which agrees with (1.4). From (2.2), the first and second moments of $X_{r:n}$ are, respectively, given by

$$\alpha_{r:n}^{(1)} = \frac{(-1)^r \pi \ \csc \frac{\pi}{\alpha}}{(r-1)!(n-r)!} \prod_{i=1}^{n} \left(i-r-\frac{1}{\alpha}\right), \quad \alpha > 1, \tag{2.4}$$

and

$$\alpha_{r:n}^{(2)} = \frac{(-1)^r \pi \ \csc \frac{2\pi}{\alpha}}{(r-1)!(n-r)!} \prod_{i=1}^{n} \left(i-r-\frac{2}{\alpha}\right), \quad \alpha > 2. \tag{2.5}$$

It is interesting to note that (2.3) can be used easily to derive several recurrence relations for the moments of order statistics. Some of these recurrence relations already exist in the literature. Below, we provide some of these recurrence relations.

I. From (2.3), we can write

$$
\begin{aligned}
\alpha_{r:n}^{(k)} &= \frac{-1}{r-1}\,\frac{(-1)^{r-1}\pi\ \csc\frac{k\pi}{\alpha}}{(r-2)!(n-r)!}\prod_{i=0}^{n-1}\left(i-(r-1)-\frac{k}{\alpha}\right)\\
&= \frac{r-1+k/\alpha}{r-1}\,\frac{(-1)^{r-1}\pi\ \csc\frac{k\pi}{\alpha}}{(r-2)!(n-r)!}\prod_{i=1}^{n-1}\left(i-(r-1)-\frac{k}{\alpha}\right)\\
&= \left[1+\frac{k}{\alpha(r-1)}\right]\alpha_{r-1:n-1}^{(k)}, \quad 2\le r\le n.
\end{aligned}
\tag{2.6}
$$

Note that the recurrence relation in (2.6) was first appeared in Ragab and Green (1984).

II. If $r=1$ in (2.3), then

$$
\begin{aligned}
\alpha_{1:n}^{(k)} &= \frac{-\pi\ \csc\frac{k\pi}{\alpha}}{(n-1)(n-2)!}\left(n-1-\frac{k}{\alpha}\right)\prod_{i=1}^{n-1}\left(i-1-\frac{k}{\alpha}\right)\\
&= \left[1-\frac{k}{\alpha(n-1)}\right]\alpha_{1:n-1}^{(k)}, \quad n\ge 2.
\end{aligned}
\tag{2.7}
$$

The recurrence relation in (2.7) first appeared in Ali and Khan (1987).

III. For $m\in\mathbb{N}$, (2.3) implies

$$
\begin{aligned}
\alpha_{r:n}^{(k-m\alpha)} &= \frac{(-1)^{r}\pi\ \csc\left(\frac{k}{\alpha}-m\right)\pi}{(r-1)!(n-r)!}\prod_{i=1}^{n}\left(i-r-\frac{k}{\alpha}+m\right)\\
&= \frac{(r-m-1)!(n-r+m)!}{(r-1)!(n-r)!}\,\frac{(-1)^{r-m}\pi\ \csc\frac{k\pi}{\alpha}}{(r-m-1)!(n-r+m)!}\\
&\quad\times\prod_{i=1}^{n}\left(i-(r-m)-\frac{k}{\alpha}\right)\\
&= \frac{(r-m-1)!(n-r+m)!}{(r-1)!(n-r)!}\,\alpha_{r-m:n}^{(k)}, \quad m+1\le r\le n.
\end{aligned}
\tag{2.8}
$$

When $m=1$, (2.8) reduces to the recurrence relation given by Ali and Khan (1987) as $\alpha_{r:n}^{(k-\alpha)} = \frac{n-r+1}{r-1}\,\alpha_{r-1:n}^{(k)}, \quad 2\le r\le n.$

IV. Another form of (2.8) can be derived as follows

$$
\begin{aligned}
\alpha_{r:n}^{(k-m\alpha)} &= \frac{(-1)^{r+m}\pi\ \csc\frac{k\pi}{\alpha}}{(r-1)!(n-r)!}\prod_{i=1}^{n}\left(i+m-r-\frac{k}{\alpha}\right)\\
&= \frac{(-1)^{r+m}\pi\ \csc\frac{k\pi}{\alpha}}{(r-1)!(n-r)!}\prod_{i=m+1}^{n+m}\left(i-r-\frac{k}{\alpha}\right)\\
&= \frac{(-1)^{m}\prod_{i=n+1}^{n+m}\left(i-r-\frac{k}{\alpha}\right)}{\prod_{i=1}^{m}\left(i-r-\frac{k}{\alpha}\right)}\,\alpha_{r:n}^{(k)}, \quad m+1\le r\le n.
\end{aligned}
\tag{2.9}
$$

V. From (2.8) and (2.9), we get

$$\alpha_{r:n}^{(k)} = \frac{(-1)^m\,(r-m-1)!(n-r+m)!\,\prod_{i=1}^{m}\left(i-r-\frac{k}{\alpha}\right)}{(r-1)!(n-r)!\,\prod_{i=n+1}^{n+m}\left(i-r-\frac{k}{\alpha}\right)}\alpha_{r-m:n}^{(k)}, \qquad +1 \le r \le n. \tag{2.10}$$

3. COVARIANCE BETWEEN ORDER STATISTICS

To calculate the covariance between $X_{r:n}$ and $X_{s:n}$, consider the joint pdf of $X_{r:n}$ and $X_{s:n}$, $1 \le r < s \le n$ as follows

$$f_{r,s:n}(x,y) = \alpha^2\, C_{r,s:n}\,\frac{x^{\alpha r-1}y^{\alpha-1}(y^\alpha - x^\alpha)^{s-r-1}}{(1+x^\alpha)^s(1+y^\alpha)^{n-r+1}}, \quad 0 \le x \le y < \infty, \tag{3.1}$$

where $C_{r,s:n} = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}$.
Therefore the product moments, $\alpha_{r,s:n} = E(X_{r:n}Y_{s:n})$, can be written as

$$\alpha_{r,s:n} = \alpha^2\, C_{r,s:n}\int_0^\infty\int_0^y \frac{x^{\alpha r}y^\alpha(y^\alpha - x^\alpha)^{s-r-1}}{(1+x^\alpha)^s(1+y^\alpha)^{n-r+1}}dxdy. \tag{3.2}$$

On using the substitution $u = x^\alpha$ and $v = y^\alpha$, (3.2) reduces to

$$\alpha_{r,s:n} = C_{r,s:n}\int_0^\infty \frac{v^{\frac{1}{\alpha}}}{(1+v)^{n-r+1}}\underbrace{\left(\int_0^v \frac{u^{r+\frac{1}{\alpha}-1}(v-u)^{s-r-1}}{(1+u)^s}du\right)}_{I}dv. \tag{3.3}$$

By using the substitution $t = \frac{u}{v}$, it is not difficult to show that I can be simplified to

$$I = v^{s+\frac{1}{\alpha}-1}B\left(r+\frac{1}{\alpha}, s-r\right){}_2F_1\left(s, r+\frac{1}{\alpha}, s+\frac{1}{\alpha}; -v\right), \tag{3.4}$$

where ${}_pF_q$ is the generalized hypergeometric function defined as

$${}_pF_q(a_1,...,a_p;b_1,...,b_q;x) = \sum_{k=0}^{\infty}\frac{(a_1)_k\cdots(a_p)_k}{(b_1)_k\cdots(b_q)_k}\,\frac{x^k}{n!}.$$

Using the Pfaff transformation, ${}_2F_1(a,b;c;x) = (1-x)^{-a}\,{}_2F_1(a,c-b;c;\frac{x}{x-1})$, we have

$${}_2F_1\left(s, r+\frac{1}{\alpha}, s+\frac{1}{\alpha}; -v\right) = (1+v)^{-r-\frac{1}{\alpha}}{}_2F_1\left(\frac{1}{\alpha}, r+\frac{1}{\alpha}, s+\frac{1}{\alpha}; \frac{v}{1+v}\right). \tag{3.5}$$

Now, using (3.4), (3.5) and the substitution $w = \frac{v}{1+v}$, (3.3) reduces to

$$\alpha_{r,s:n} = C_{r,s:n}\,B\left(r+\frac{1}{\alpha}, s-r\right)\int_0^1 w^{s+\frac{2}{\alpha}-1}(1-w)^{n-s-\frac{1}{\alpha}}\,{}_2F_1\left(\frac{1}{\alpha}, r+\frac{1}{\alpha}, s+\frac{1}{\alpha}; w\right). \tag{3.6}$$

On using the identity [Gradshteyn and Ryzhik, [14], p. 813]

$$\int_0^1 x^{\rho-1}(1-x)^{\sigma-1}\,{}_2F_1(\alpha,\beta,\gamma;x)dx = B(\rho,\sigma)_3F_2(\alpha,\beta,\rho;\gamma,\rho+\sigma;1),$$

the product moments of the log-logistic distribution can be written as

$$\tag{3.7} \begin{aligned} \alpha_{r,s:n} &= C_{r,s:n}\,B\left(r+\frac{1}{\alpha},s-r\right)\;B\left(s+\frac{2}{\alpha},n-s-\frac{1}{\alpha}+1\right) \\ &\quad\times {}_3F_2\left(\frac{1}{\alpha},r+\frac{1}{\alpha},s+\frac{2}{\alpha};s+\frac{1}{\alpha},n+\frac{1}{\alpha}+1;1\right). \end{aligned}$$

It is clear from (3.7) that $\alpha_{r,s:n}$ exists for all $\alpha > 1$.

It is noteworthy to mention that one can use some existing recurrence relations in the literature to compute $\alpha_{r,s:n}$ in a more efficient way. For example, Joshi and Balakrishnan (1982) show that for any continuous distribution, the following recurrence relation holds

$$\tag{3.8} \begin{aligned} \alpha_{r,n:n} &= \sum_{i=1}^{n-r}(-1)^{n-r-i}\binom{n}{n-i}\binom{n-i-1}{r-1}\alpha_{n-i:n-i}\alpha_{i:i} \\ &\quad-\sum_{\ell=0}^{r-1}(-1)^{n-\ell}\binom{n}{\ell}\alpha_{1,n-r+1:n-\ell}, \quad 1\le r\le n-1. \end{aligned}$$

Also, Ali and Khan (1987) show the following recurrence relation for the log-logistic distribution,

$$\tag{3.9} \begin{aligned} \alpha_{r,s:n} &= \alpha_{r,s-1:n}+\left(\frac{n}{n-s+1}\right)\left(1-\frac{1}{\alpha(n-s)}\right)\alpha_{r,s:n-1} \\ &\quad-\frac{n}{n-s+1}\alpha_{r,s-1:n-1}, \quad 1\le r<s\le n-1. \end{aligned}$$

The covariance $\beta_{r,s:n} = \alpha_{r,s:n} - \alpha_{r:n}\,\alpha_{s:n}$, can be obtained from equations (2.4), (2.5) and (3.7). Note that when $r = s$, the variances $\beta_{r,r:n} = \alpha^2_{r:n} - (\alpha_{r:n})^2$. The recurrence relations in (23) and (24) can be also used in these calculations.

4. PARAMETER ESTIMATION FOR THE LOG-LOGISTIC DISTRIBUTION

In this section, we discuss the parameter estimation for the log-logistic distribution based on order statistics.

4.1. Estimation of location and scale parameters

Let $Y_1, Y_2, ..., Y_n$ be a random sample of size n from the log-logistic$(\alpha, \theta_1, \theta_2)$, where θ_1 is the location parameter and $\theta_2 > 0$ is the scale parameter. I.e. $f(y) = \alpha\,\theta_2^{-1}\left(\frac{y-\theta_1}{\theta_2}\right)^{\alpha-1}\left(1+\left(\frac{y-\theta_1}{\theta_2}\right)^{\alpha}\right)^{-2}$, $y \geq \theta_1$. In this section, we compute the best linear unbiased estimators (BLUEs) for θ_1 and θ_2 when the shape parameter α is known. Let $X = (Y - \theta_1)/\theta_2$. When α is known, the mean, $\alpha_{r:n}$, and the covariance, $\beta_{r,s:n}$, of order statistics are completely known and free of parameters. The estimators for θ_1 and θ_2 are derived based on weighted regression on the quantile-quantile plot of order statistics against their expected value. The weights depend on the covariance matrix of the order statistics. The estimations of location and scale parameters based on order statistics were originally introduced by Lloyd [20]. Several authors including Arnold *et al.* ([6], p. 17) and Ahsanullah *et al.* ([3], p. 154) used Lloyd's method to obtain best linear unbiased estimator (BLUE) of the location and scale parameters for probability distributions. The BLUEs of θ_1 and θ_2 can be computed as follows [see Arnold *et al.* ([6], pp. 171–173) and Ahsanullah *et al.* ([3], p. 154)]

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{A'}\,\boldsymbol{\Sigma^{-1}}\,\boldsymbol{A})^{-1}\,\boldsymbol{A'}\,\boldsymbol{\Sigma^{-1}}\,\boldsymbol{Y},$$

where $\boldsymbol{A'}$ denotes to the transpose of $\boldsymbol{A}$, $\boldsymbol{A} = (\mathbf{1}, \boldsymbol{\mu})$, $\mathbf{1} = (1, 1, ..., 1)'_{1\times n}$, $\boldsymbol{\mu} = (\alpha_{1:n}, \alpha_{2:n}, ..., \alpha_{n:n})'$, $\hat{\boldsymbol{\theta}} = (\hat{\theta_1}, \hat{\theta_2})'$, $\boldsymbol{\Sigma} = ((\beta_{r,s:n}))_{n\times n}$ and $\boldsymbol{Y} = (Y_{1:n}, Y_{2:n}, ..., Y_{n:n})'$. Alternatively,

$$\hat{\theta_1} = -\boldsymbol{\mu}'\,\boldsymbol{\Gamma}\,\boldsymbol{Y} \quad\text{and}\quad \hat{\theta_2} = \mathbf{1}'\,\boldsymbol{\Gamma}\,\boldsymbol{Y},$$

where $\boldsymbol{\Gamma = \Sigma^{-1}(1\mu' - \mu\,1')\Sigma^{-1}/\Delta}$ and $\boldsymbol{\Delta = (1'\Sigma^{-1}1)(\mu'\Sigma^{-1}\mu) - (1'\Sigma^{-1}\mu)^2}$. The coefficient matrix $\boldsymbol{C} = (\boldsymbol{A'}\,\boldsymbol{\Sigma^{-1}}\,\boldsymbol{A})^{-1}\,\boldsymbol{A'}\,\boldsymbol{\Sigma^{-1}}$ can be obtained using $\alpha_{r:n}$, $\alpha_{r:n}^{(2)}$, $\alpha_{r,s:n}$ and $\beta_{r,s:n}$ from previous section. The covariance matrix of the estimators can be computed in terms of θ_2 as follows

$$\boldsymbol{Cov(\hat{\theta}) = (A'\,\Sigma^{-1}\,A)}^{-1}\theta_2^2. \tag{4.1}$$

In particular,

$$\begin{aligned}
var(\hat{\theta_1}) &= \theta_2^2\boldsymbol{\mu'\Sigma^{-1}\,\mu/\Delta},\\
var(\hat{\theta_2}) &= \theta_2^2\boldsymbol{1'\Sigma^{-1}\,1/\Delta},\\
Cov(\hat{\theta_1}, \hat{\theta_2}) &= -\theta_2^2\boldsymbol{\mu'\Sigma^{-1}1/\Delta}.
\end{aligned}$$

Equation (4.1) is used to compute the variance and covariance of $\hat{\theta_1}$ and $\hat{\theta_2}$ in terms of θ_2. The coefficients and covariances for computing the BLUE of $\hat{\boldsymbol{\theta}}$ for various values of the shape parameter α and sample sizes up to 10 are available on `https://sites.google.com/site/statisticsmanagementservices/`.

4.2. Estimation of the shape parameter

In real life situations, we encounter unknown value for the shape parameter α. In order to use the The BLUEs for θ_1 and θ_2, we first estimate the shape parameter α.

Lemma 4.1. *The log-logistic distribution is a member of the Pareto-type distributions with tail index α.*

Proof: Note that $1 - F(x) = \frac{1}{1+x^\alpha} = x^{-\alpha}\ell(x)$, where $\ell(x) = 1 - x^{-\alpha} + x^{-2\alpha} + \cdots$ is slowly varying function at infinity. To see this, for any $\lambda > 0$, $\frac{\ell(\lambda x)}{\ell(x)} \longrightarrow 1$ as $x \longrightarrow \infty$. Hence $F(x)$ constitutes a Pareto-type distribution with tail index α. $\square$

Several estimators for the heavy tail index α exist in the literature. For example, a family of kernel estimators for α was proposed by Csorgo, Deheuvels and Mason [12]. Bacro and Brito [8] and De Hann [13] proposed estimators for α which are members of the family of kernel estimators. For more information, we refer the reader to the paper by Beirlant *et al.* [9] and Gomes and Henriques-Rodrigues [17]. The most popular estimator for α is the Hill estimator proposed by Hill [18] as follows:
Let $X_1, X_2, ..., X_n$ be n independent random sample from log-logistic$(\alpha, \theta_1, \theta_2)$. Let $\mathrm{X}_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ be the corresponding order statistics. The Hill estimator for α based on upper k order statistics is given by

$$\hat{\alpha} = \frac{1}{H_{k,n}}, \quad H_{k,n} = \frac{1}{k}\sum_{j=1}^{k} \log \frac{X_{n-j+1,n}}{X_{n-k,n}}. \tag{4.2}$$

Although the Hill estimator is scale invariant, it is not shift invariant. Aban and Meerschaert [1] proposed a modification of Hill estimator in order to make it both shift and scale invariant as follows:

$$\hat{\alpha}^{-1} = \frac{1}{k}\sum_{j=1}^{k} \log \frac{X_{n-j+1,n} - \hat{s}}{X_{n-k,n} - \hat{s}},$$

where the sift $\hat{s}$ satisfies the equation

$$\hat{\alpha}(X_{n-k,n} - \hat{s})^{-1} = \frac{\hat{\alpha}+1}{k}\sum_{j=1}^{k}(X_{n-j+1,n} - \hat{s})^{-1}, \quad \hat{s} < X_{n-k,n}.$$

In general, the modified Hill estimator results in large variation of the sampling distribution in compared with the Hill estimator. In our case, based on various simulated random samples with different sample sizes from $X \sim$ log-logistic$(\alpha, \theta_1, \theta_2)$, the modified Hill estimator produces poor estimate for the parameter α. Therefore, we decided to shift the random sample by the sample minimum, $X_{1,n}$, and then use the Hill estimator to estimate α. This is justified since the lower end of the distribution is finite. For an interesting discussion of this topic see Araújo-Santos *et al.* [5] and Gomes *et al.* [15]. The results showed good estimate to the shape parameter α (see Table 1).

4.3. Monte Carlo simulation study

In this subsection, we generate different random samples with various sizes, $n = 100$, 500, 1,000 and 10,000. The simulation study is repeated 1,000 times for four groups of parameters:

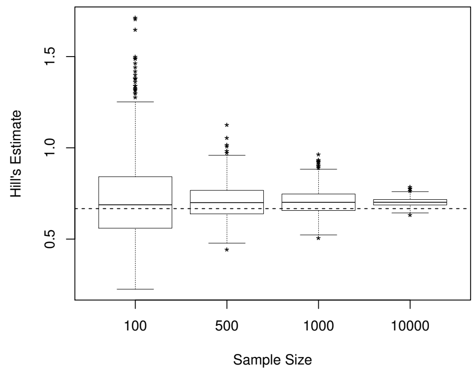
$$\begin{aligned}
\mathrm{I}:&\ \alpha = 0.5,\ \theta_1 = 1,\ \theta_2 = 1,\\
\mathrm{II}:&\ \alpha = 1.5,\ \theta_1 = 0,\ \theta_2 = 1,\\
\mathrm{III}:&\ \alpha = 2.5,\ \theta_1 = 2,\ \theta_2 = 3,\\
\mathrm{IV}:&\ \alpha = 4,\quad \theta_1 = 2,\ \theta_2 = 0.5.
\end{aligned}$$

For each parameter combination, we generate random samples, Y_i, $i = 1, ..., n$ from log-logistic$(\alpha, \theta_1, \theta_2)$. We assume the random sample $X_i = Y_i - Y_{1,n}$ follows log-logistic$(\alpha, 0, \theta_2)$. Then we estimate α using the Hill estimator in equation (4.2). Gomes and Guillou [16] have given an interesting discussion about the choice of k. It is known that the bias of the estimator of the index parameter increases as k increases and the variance of the index estimator increases if k is small. The choice of k is a question between the choice of bias and variance.

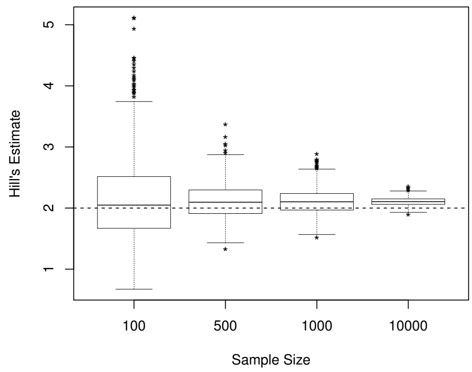
Table 1: Mean, median and standard deviation for $\hat{\alpha}$ using the Hill estimate.

Sample Size	Summary Statistics	Group I	Group II	Group III	Group IV
		$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\alpha}$
100	Median	0.4881	1.4539	2.3275	3.3510
	Mean	0.5158	1.5371	2.4615	3.5460
	Standard Deviation	0.1716	0.5094	0.8067	1.1494
500	Median	0.4771	1.4287	2.3320	3.4810
	Mean	0.4827	1.4450	2.3600	3.5150
	Standard Deviation	0.0690	0.2063	0.3350	0.4995
1000	Median	0.4758	1.4256	2.3390	3.5260
	Mean	0.4786	1.4342	2.3540	3.5460
	Standard Deviation	0.0473	0.1418	0.2330	0.3584
10000	Median	0.4750	1.4246	2.3580	3.6438
	Mean	0.4754	1.4258	2.3616	3.6460
	Standard Deviation	0.0147	0.0441	0.0733	0.1211

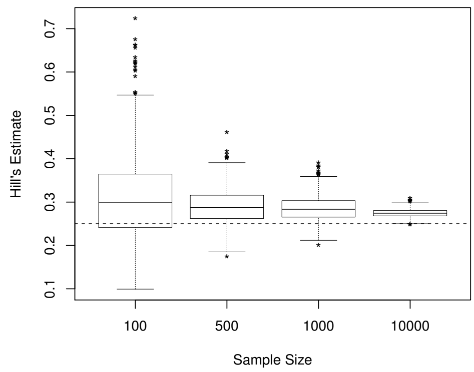
We have taken $k = 10\%$ of the sample size with $n > 100$. The simulation results in Table 1 show that as the parameter α increases, the absolute bias and standard deviation increase. Overall, the Hill estimator performs well in estimating the shape parameter α. Figures 1–4 represent the Boxplots for the observed sampling distributions of the Hill estimate for different sample sizes. These Figures indicate that the observed distributions are approximately normal and centered roughly at α^{-1}.



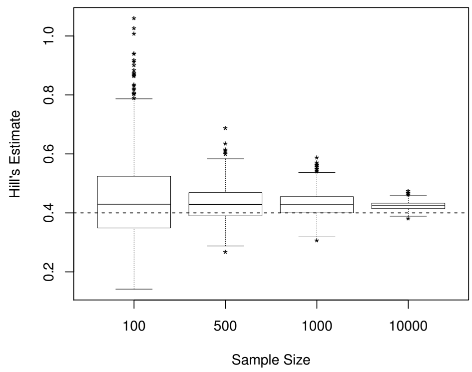
(a) Group I



(b) Group II



(a) Group III



(b) Group IV

Figure 1: Boxplots for the observed sampling distributions of $\hat{\alpha}^{-1}$. Dashed line represents the true parameter α^{-1}.

4.4. Numerical Example

In this subsection, we illustrate the use of Hill estimator and the BLUE's for estimating the three-parameter log-logistic distribution. We simulate a random sample with $n = 30$ observations from log-logistic distribution with parameters $\alpha = 4$, $\theta_1 = 2$ and $\theta_2 = 3$. The simulated data is given below:

5.80310, 6.88820, 6.00730, 7.01140, 4.87250, 4.00560, 4.49970, 5.02880, 5.83690, 11.40110, 3.30511, 3.95312, 5.87513, 2.55114, 4.68615, 4.88916, 4.67717, 4.71818, 4.05190, 8.31920, 4.86421, 4.50422, 8.89623, 5.74124, 5.48125, 4.68226, 5.70127, 5.13528, 4.20729, 4.95430.

Using similar approach as in subsection 4.3, the estimated value of α based on the Hill estimator is $\hat{\alpha} = 3.350$. Based on this value and the sample size of $n = 30$, the coefficient matrix, $\boldsymbol{C} = (\boldsymbol{A'}\,\boldsymbol{\Sigma^{-1}}\,\boldsymbol{A})^{-1}\,\boldsymbol{A'}\,\boldsymbol{\Sigma^{-1}}$, and the covariance, $\boldsymbol{Cov(\hat{\theta}) = (A'\,\Sigma^{-1}\,A)}^{-1}\theta_2^2$, can be calculated using $\alpha_{r:n}$, $\alpha_{r:n}^{(2)}$, $\alpha_{r,s:n}$ in equations (9), (10) and (22) respectively. These coefficients for computing the BLUE's for θ_1 and θ_2 and the covariance matrix are provided below:

$$\boldsymbol{C_{\theta_1}} = \begin{pmatrix} 0.6077 \\ 0.25102 \\ -0.08301 \\ 0.66281 \\ -0.17817 \\ -0.09813 \\ 0.27104 \\ -0.02089 \\ 0.0082 \\ 0.07347 \\ -0.18869 \\ 0.48597 \\ -0.09435 \\ -0.48662 \\ -0.0126 \\ 0.51641 \\ -0.302 \\ 0.43571 \\ -0.66808 \\ 0.0207 \\ 0.11805 \\ -0.16527 \\ 0.00062 \\ -0.01915 \\ -0.07406 \\ 0.00442 \\ -0.00919 \\ -0.04542 \\ -0.00744 \\ -0.00303 \end{pmatrix}, \qquad \boldsymbol{C_{\theta_2}} = \begin{pmatrix} -0.63961 \\ -0.23399 \\ 0.09682 \\ -0.64294 \\ 0.15847 \\ -0.01285 \\ 0.12400 \\ 0.00067 \\ -0.00583 \\ -0.01679 \\ 0.15861 \\ -0.09637 \\ 0.00875 \\ 0.44036 \\ 0.02531 \\ -0.46079 \\ 0.23208 \\ -0.21193 \\ 0.70721 \\ -0.02347 \\ -0.04269 \\ 0.18111 \\ -0.01123 \\ 0.07461 \\ 0.08937 \\ -0.00810 \\ 0.05792 \\ 0.03363 \\ 0.01434 \\ 0.00333 \end{pmatrix},$$

$$\boldsymbol{Cov(\hat{\theta})} = \begin{pmatrix} 0.01636 & -0.01755 \\ -0.01755 & 0.02777 \end{pmatrix}.$$

Therefore, the BLUE's for θ_1 and θ_2 and the estimated covariances are evaluated to be

$$\hat{\theta_1} = 1.98287, \quad \hat{\theta_2} = 3.09528,$$

$$Var(\hat{\theta_1}) = 0.15676, \quad Var(\hat{\theta_2}) = 0.26606$$

and

$$Cov(\hat{\theta_1}, \hat{\theta_2}) = -0.16812.$$

5. CONCLUDING REMARKS

In this paper, the moments and product moments of the order statistics in a sample of size n drawn from the log-logistic distribution are discussed. We also provided in the same section more compact formulas for the means, variances and covariances of order statistics. Best linear unbiased estimators (BLUEs) for the location and scale parameters for the log-logistic distribution with known shape parameter based on order statistics are studied. The Hill estimator is proposed for estimating the shape parameter.

ACKNOWLEDGMENTS

The authors are grateful for the comments and suggestions by the referee and the Associate Editor. Their comments and suggestions greatly improved the article.

REFERENCES

[1] ABAN, I.B. and MEERSCHAERT, M.M. (2001). Shifted Hill's estimator for heavy tails, *Communications in Statistics-Simulation and Computation*, **30**, 949–962.

[2] ABRAMOWITZ, M. and STEGUN, I.A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Vol. 55, Dover Publication, Inc., New York.

[3] AHSANULLAH, M.; NEVZOROV, V.B. and SHAKIL, M. (2013). *Order statistics*, Atlantis Press, Paris, France.

[4] ALI, M.M. and KHAN, A.H. (1987). On order statistics from the log-logistic distribution, *Journal of Statistical Planning and Inference*, **17**, 103–108.

[5] Araújo-Santos, P.; Fraga Alves, M.I. and Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and high quantile estimation, *Revstat*, **4**, 227–247.

[6] Arnold, B.C.; Balakrishnan, N. and Nagaraja, H.N. (1992). *A first course in order statistics*, Vol. 54, SIAM.

[7] Ashkar, F. and Mahdi, S. (2006). Fitting the log-logistic distribution by generalized moments, *Journal of Hydrology*, **238**, 694–703.

[8] Bacro, J.N. and Brito, M. (1993). Strong limiting behaviour of a simple tail Pareto-index estimator, *Statistics and Decisions*, **3**, 133–134.

[9] Beirlant, J.; Vynckier, P. and Teugels, J.L. (1996). Tail index estimation, Pareto quantile plots, and regression diagnostics, *Journal of the American statistical Association*, **91**, 1659–1667.

[10] Bennett, S. (1983). Log-Logistic Regression Models for Survival Data, *Journal of the Royal Statistical Society, Series C*, **32**, 165–171.

[11] Collett, D. (2003). *Modelling Survival Data in Medical Research*, 2nd edition, CRC Press.

[12] Csorgo, S.; Deheuvels, P. and Mason, D. (1985). Kernel estimates of the tail index of a distribution, *The Annals of Statistics*, **13**, 1050–1077.

[13] De Haan, L. (1981). Estimation of the minimum of a function using order statistics, *Journal of the American Statistical Association*, **76**, 467–469.

[14] Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of Integrals, Series and Products*, 7th edition, Academic, New York, USA.

[15] Gomes, M.I.; Fraga Alves, M.I. and Araújo-Santos, P. (2008). PORT Hill and moment estimators for heavy-tailed models, *Communications in Statistics-Simulation and Computation*, **37**, 1281–1306.

[16] Gomes, M.I. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: A review, *International Statistical Review*, **83**, 263–292.

[17] Gomes, M.I. and Henriques-Rodrigues, L. (2016). Competitive estimators of extreme value Index, *Statistics and Probability Letters*, **67**, 128–135.

[18] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, **3**, 1163–1174.

[19] Joshi, P.C. and Balakrishnan, N. (1982). Recurrence relations and identities for the product moments of order statistics, *Sankhya: The Indian Journal of Statistics, Series B*, **44**, 39–49.

[20] Lloyd, E.H. (1952). Least squares estimation of location and scale parameters using order statistics, *Biometrika*, **39**, 88–95.

[21] Ragab, A. and Green J. (1984). On order statistics from the log-logistic distribution and their properties, *Communications in Statistics-Theory and Methods*, **13**, 2713–2724.

[22] Shah, B.K. and Dave, P.H. (1963). A note on log-logistic distribution, *Journal of the MS University of Baroda*, **12**, 15–20.

[23] Shoukri, M.M.; Mian, I.U.M. and Tracy, D.S. (1988). Sampling Properties of Estimators of the Log-Logistic Distribution with Application to Canadian Precipitation Data, *The Canadian Journal of Statistics*, **16**, 223–236.

[24] Tadikamalla, P.R. and Johnson, N.L. (1982). Systems of frequency curves generated by transformations of logistic variables, *Biometrika*, **69**, 461–465.

[25] Voorn, W.J. (1987). Characterization of the logistic and log-logistic distributions by extreme value related stability with random sample size, *Journal of Applied Probability*, **24**, 838–851.