Sankhyā : The Indian Journal of Statistics
2007, Volume 69, Part 4, pp. 870-884

Linear Estimation of Scale Parameter for Logistic Distribution Based on Consecutive Order Statistics

Patrick G.O. Weke
University of Nairobi, Nairobi, Kenya

Abstract

Linear estimation of the scale parameter of the logistic population based on the sum of consecutive order statistics, when the location parameter is unknown, is discussed. A method based on a pair of single spacing and 'zero-one' weights rather than the optimum weights is presented. Limited simulations indicate small bias and variance of the estimator, and reasonably high relative efficiencies with respect to the Cramer-Rao lower bound and best linear unbiased estimators (BLUE's) for small sample sizes.

AMS (2000) *subject classification.* Primary 62F10.
Keywords and phrases. Order statistics, logistic distribution, linear estimation, relative efficiency.

1 Introduction

The logistic distribution arises frequently in statistical modelling. It has been used in the analysis of survival data, graduation of mortality statistics and is used in some applications as a substitute for the normal distribution (Balakrishnan and Cohen, 1991).

Let $X_{1:n} \leq \cdots \leq X_{n:n}$ denote the order statistics from a random sample of size n from the logistic distribution whose cumulative distribution function is

$$F(x;\mu,\sigma) = \left[1+\exp\left\{-\pi(x-\mu)/\sigma\sqrt{3}\right\}\right]^{-1}; \quad -\infty < x < \infty,$$
$$-\infty < \mu < \infty, \ \sigma > 0.$$

The distribution is absolutely continuous, symmetric about the location parameter μ and has scale parameter σ. There has been much work on the

estimation of the location and scale parameters of a distribution through linear functions of order statistics. Optimal weights for these linear combinations can be calculated in the case of a given distribution (Balakrishnan and Cohen, 1991). Dixon (1957) introduced the notion of using 'zero-one' weights rather than the optimum weights and noted that high efficiencies are achievable. This work has been extended by many authors, notably for the logistic distribution (Raghunandanan and Srinivasan, 1970) and the normal distribution (Wang, 1980). Raghunandanan and Srinivasan (1970) constructed simplified estimators of the location and scale parameters for complete and symmetrically censored samples for sample sizes $4 \leq n \leq 20$.

The principal objective of this paper is to develop, based on the work of Wang (1980), Weke et al. (2001) and Weke (2005), a simplified linear estimator of the scale parameter of the population when its location parameter is unknown. The proposed procedure is applicable to both censored and uncensored samples and for large sample sizes. Bias and variance of this estimator are examined through Monte Carlo simulations.

2 Deriving Expectation Formulae

Let $U_1, U_2, \ldots, U_n$ be a random sample of size n from the uniform $U(0,1)$ distribution with pdf $f(u) = 1, \ 0 \leq u \leq 1$. Then, the density function of the i-th order statistics $U_{i:n}, \ 1 \leq i \leq n$, is given by

$$f_i(u;n) = \frac{\Gamma(n+1)}{\Gamma(i)\Gamma(n-i+1)} u^{i-1}(1-u)^{n-i}, \quad 0 \leq u \leq 1. \tag{2.1}$$

In the sequel, we will work with some adjustments to the values of i and n, and a corresponding adjustment to the beta distribution given in (2.1). If α and β are positive real numbers such that $1+\alpha \leq i \leq n-\beta$, we consider the modified pdf

$$f_{i-\alpha}(u;n-\alpha-\beta) = \frac{\Gamma(n-\alpha-\beta+1)}{\Gamma(i-\alpha)\Gamma(n-i-\beta+1)} u^{i-\alpha-1}(1-u)^{n-i-\beta}, \quad 0 \leq u \leq 1. \tag{2.2}$$

Let $i' = i-\alpha$ and $n' = n-\alpha-\beta$. Let $U_{i':n'}$ denote a random variable having the pdf of (2.2). The mean of this pdf is

$$\pi_{i'} = E(U_{i':n'}) = \frac{i'}{n'+1}, \tag{2.3}$$

and the central moment of order k is given by

$$\mu_k = E[(U_{i':n'} - \pi_{i'})^k] = \frac{(k-1)\pi_{i'}(1-\pi_{i'})}{n'+2} E\left[(U_{i+1:n+2} - \pi_{i'})^{k-2}\right]. \tag{2.4}$$

To see this, note that

$$\begin{aligned}\mu_k &= \int_0^1 (u-\pi_{i'})^{k-1}(u-\pi_{i'})\frac{\Gamma(n'+1)}{\Gamma(i')\Gamma(n'-i'+1)}u^{i'-1}(1-u)^{n'-i'}du \\ &= -\frac{\pi_{i'}(1-\pi_{i'})}{n'+2}\int_0^1 (u-\pi_{i'})^{k-1}\frac{\Gamma(n'+3)}{\Gamma(i'+1)\Gamma(n'-i'+2)} \\ &\qquad \times[i'(1-u)-(n'-i'+1)u]u^{i'-1}(1-u)^{n'-i'}du \\ &\qquad \text{(integrating by parts)} \\ &= \frac{(k-1)\pi_{i'}(1-\pi_{i'})}{n'+2}\int_0^1 (u-\pi_{i'})^{k-2}\frac{\Gamma(n'+3)}{\Gamma(i'+1)\Gamma(n'-i'+2)} \\ &\qquad \times u^{i'}(1-u)^{n'-i'+1}du.\end{aligned}$$

Hence, the following results may easily be obtained.

(i) $E(U_{i':n'}) = \pi_{i'}$. (2.5)

(ii) $\mu_2 = \dfrac{1}{n'+2}\pi_{i'}(1-\pi_{i'})$. (2.6)

(iii) $\mu_3 = \dfrac{2}{(n'+2)(n'+3)}\pi_{i'}(1-\pi_{i'})(1-2\pi_{i'})$. (2.7)

(iv) $\mu_4 = \dfrac{3\pi_{i'}(1-\pi_{i'})}{n'+2}\left\{\dfrac{\pi_{i',1}(1-\pi_{i',1})}{n'+4}+\left(\dfrac{1-2\pi_{i'}}{n'+3}\right)^2\right\}$, (2.8)

where $\pi_{i',1} = \dfrac{i'+1}{n+3}$.

(v) For any non-negative integer r, it is possible to find a quantity M, which does not depend on n' and i', such that γ_r, the absolute central moment of order r of $U_{i':n'}$, satisfies

$$\begin{aligned}\gamma_r &< \frac{M}{n^{r/2}}, &&\text{if } r \text{ is even} \\ \text{and} \quad \gamma_r &< \frac{M}{n^{(r+1)/2}}, &&\text{if } r \text{ is odd.}\end{aligned} \tag{2.9}$$

The inequalities in (2.9) can be proved by induction from (2.4).

Let $X_{1:n}, \ldots, X_{n:n}$ be order statistics from a continuous and strictly increasing distribution function F.

DEFINITION 1. For any fixed $c \in [0,1]$, a sequence of order statistics $\{X_{i_n:n}\}$, $n = 1, 2, \ldots$, is called a *c-sequence*, if $i_n/n \to c$ as $n \to \infty$.

DEFINITION 2. For $\varepsilon > 0$, an ε-neighbourhood of the convergent sequence ζ_r', $r = 1, 2, \ldots$, is defined as the set of all points ζ satisfying the inequalities $|\zeta - \zeta_r'| \le \varepsilon$, for all r.

Let G denote the inverse of the distribution function F. Let H denote the function

$$H(u) = u^\alpha (1-u)^\beta, \qquad 0 \le u \le 1, \tag{2.10}$$

where α and β are positive constants. We use the notations

$$L_r(u) = u(1-u)\frac{G^{(r+1)}(u)}{G^{(r)}(u)}, \qquad r = 1,2,3, \tag{2.11}$$

and

$$K_r(u) = u^r(1-u)^r\frac{H^{(r)}(u)}{H(u)}, \qquad r = 1,2,3, \tag{2.12}$$

where $G^{(r)}(u)$ and $H^{(r)}(u)$ denote the r-th order derivatives of $G(u)$ and $H(u)$, respectively. Further, let

$$Q_1(\pi_{i'};\alpha,\beta) = K_1(\pi_{i'}) + \frac{1}{2}L_1(\pi_{i'}) \quad \text{and} \tag{2.13}$$

$$\begin{aligned}Q_2(\pi_{i'};\alpha,\beta) &= -\frac{1}{2}K_2(\pi_{i'})\left(2\frac{K_1(\pi_{i'})}{L_1(\pi_{i'})}+1\right) \\ &\quad +\frac{2}{3}(1-2\pi_{i'})\left[3\frac{K_2(\pi_{i'})}{L_1(\pi_{i'})}+3K_1(\pi_{i'})+L_2(\pi_{i'})\right] \\ &\quad +\frac{1}{4}\left[\frac{\pi_{i,1}(1-\pi_{i,1}')+(n'+3)^{-1}(1-2\pi_{i'})^2}{\pi_{i'}(1-\pi_{i'})}\right] \\ &\quad \times 4\left[\frac{K_3(\pi_{i'})}{L_1(\pi_{i'})}+6K_2(\pi_{i'})+4L_2(\pi_{i'})K_1(\pi_{i'})+L_2(\pi_{i'})L_3(\pi_{i'})\right]\end{aligned} \tag{2.14}$$

THEOREM 2.1. *Let $X_{i:n}, \ldots X_{n:n}$ be order statistics from the distribution F, such that the following conditions hold.*

(a) For some positive real numbers α and β, the inverse G of the distribution function F is such that $G(u)u^\alpha(1-u)^\beta$ is bounded in the closed interval $[0,1]$;

(b) $G(u)$ and its first four derivatives are continuous and bounded in the closed interval $[0,1]$, and $G^{(r)}$ is bounded in the open interval $(0,1)$, possibly with the exception of a finite number of points;

(c) $\{X_{i_n:n}\}$ is a c-sequence, where $c \in (0,1)$; and

(d) $0 \le \alpha < i_n$ and $0 \le \beta < n - i_n + 1$.

Then, with the simplified notations: i for the index i_n, i' for $i-\alpha$ and n' for $n-\alpha-\beta$, the following holds.

$$\begin{aligned} E(X_{i:n}) &= G(\pi_{i'}) + \frac{1}{n'+2} G^{(1)}(\pi_{i'}) Q_1(\pi_{i'};\alpha,\beta) \\ &+ \frac{1}{2(n'+1)^2} G^{(1)}(\pi_{i'}) Q_2(\pi_{i'};\alpha,\beta) + O\left[(n'+1)^{-3}\right] \quad (2.15) \end{aligned}$$

for all $1 \le i \le n$.

PROOF.

$$\begin{aligned} E(X_{i:n}) &= \frac{n!}{(i-1)!(n-i)!} \int_{-\infty}^{\infty} x[F(x)]^{i-1}[1-F(x)]^{n-i} f(x)dx \quad (2.16) \\ &= \frac{n!}{(i-1)!(n-i)!} \int_0^1 G(u) u^{i-1}(1-u)^{n-i} du \\ &= \frac{n!}{(i-1)!(n-i)!} \int_0^1 G(u)H(u) u^{i-\alpha-1}(1-u)^{n-i-\beta} du \\ &= \frac{\int_0^1 G(u)H(u) f_{i'}(u;n')du}{\int_0^1 H(u) f_{i'}(u;n')du} = \frac{E(GH)}{E(H)}. \quad (2.17) \end{aligned}$$

The expectations in the last expression in (2.17) are calculated by using the pdf in (2.2).

From condition (a), $G(u)H(u)$ in (2.17) is bounded. From (d), $f_{i'}(u;n')$ may be considered as a pdf, and therefore the integral in the numerator exists and is bounded. For similar reasons, the integral in the denominator of (2.17) exists and bounded, and its value is greater than zero. Hence, the ratio in (2.17) is well defined.

From (c), $\pi_{i'}$ is located in an ε-neighbourhood of c, which ensures that (d) is satisfied for large n. Denote an ε-neighbourhood of $\{\pi_{i'}\}$ by Ω_1 and

the remainder of the interval $[0,1]$ by Ω_2. Separation of the numerator of (2.17) into two parts leads to

$$\begin{aligned} \int_0^1 & G(u)H(u) f_{i'}(u;n')du \\ &= \int_{\Omega_1} G(u)H(u) f_{i'}(u;n')du + \int_{\Omega_2} G(u)H(u) f_{i'}(u;n')du \\ &= J_1 + J_2 \quad \text{(say)}. \end{aligned}$$

By first considering J_2, and using the fact that GH is bounded, we have, by Chebyschev's inequality,

$$|J_2| \le M\varepsilon^{-k}\gamma_k, \quad (2.18)$$

where the quantity M is a positive number and k can be any positive integer. Now, using condition (b), the integral J_1 can be expanded in Taylor series as

$$\begin{aligned} G(u)H(u) &= \sum_{r=0}^{4} \frac{1}{r!} [G(u)H(u)]^{(r)}\Big|_{u=\pi_{i'}} (u-\pi_{i'})^r \\ &+ \frac{1}{5!} [G(u)H(u)]^{(5)}\Big|_{u=u'} (u-\pi_{i'})^5, \quad (2.19) \end{aligned}$$

where u' is some number in Ω_1.

Since

$$\mu_r = \int_{\Omega_1} (u-\pi_{i'})^r f_{i'}(u;n')du + \int_{\Omega_2} (u-\pi_{i'})^r f_{i'}(u;n')du,$$

we have

$$J_1 = \sum_{r=0}^{4} \frac{1}{r!} (GH)^{(r)}\Big|_{u=\pi_{i'}} \cdot \mu_r - H_1 + H_2, \quad (2.20)$$

where

$$\begin{aligned} H_1 &= \sum_{r=0}^{4} \frac{1}{r!} (GH)^{(r)}\Big|_{u=\pi_{i'}} \cdot \int_{\Omega_2} (u-\pi_{i'})^r f_{i'}(u;n')du \quad (2.21) \\ \text{and } H_2 &= \frac{1}{5!} (GH)^5\Big|_{u=u'} \cdot \int_{\Omega_2} (u-\pi_{i'})^5 f_{i'}(u;n')du. \quad (2.22) \end{aligned}$$

For the same reasoning as for J_2 in (2.18), we have

$$|H_1| < M\gamma_5 \quad (2.23)$$

for all positive integer k, where M is a positive number. Since $H(\cdot)$ is positive definite and $(GH)^{(5)}|_{u=u'}$ is bounded due to condition (b), we also have

$$|H_2| < M\gamma_5, \tag{2.24}$$

where M is a positive number.

Let $k = 5$ in (2.18) and (2.23). By considering (2.17), (2.18), (2.20), (2.23), (2.24) and (2.9), we obtain

$$E(GH) = \sum_{r=0}^{4} \frac{1}{r!}(G(u)H(u))^{(r)}\Big|_{u=\pi_{i'}} \cdot \mu_k + O\left[(n'+1)^{-3}\right]. \tag{2.25}$$

Similarly, the denominator in (2.17) simplifies to

$$E(H) = \sum_{r=0}^{4} \frac{1}{r!}(H)^{(r)}\Big|_{u=\pi_{i'}} \cdot \mu_k + O\left[(n'+1)^{-3}\right]. \tag{2.26}$$

Since $\mu_1 = 0$, formula (2.17), together with (2.25) and (2.26), yields

$$\begin{aligned}
&E(X_{i:n})\\
&= \frac{GH + \frac{1}{2!}\mu_2(GH)^{(2)} + \frac{1}{3!}\mu_3(GH)^{(3)} + \frac{1}{4!}\mu_4(GH)^{(4)} + O[(n'+1)^{-3}]}{H + \frac{1}{2!}\mu_2 H^{(2)} \frac{1}{3!}\mu_3 H^{(3)} + \frac{1}{4!}\mu_4 H^{(4)} + O[(n'+1)^{-3}]} \qquad (2.27)\\
&= G\left[1 + \frac{1}{2!}\mu_2\frac{(GH)^{(2)}}{GH} + \frac{1}{3!}\mu_3\frac{(GH)^{(3)}}{GH} + \frac{1}{4!}\mu_4\frac{(GH)^{(4)}}{GH}\right]\\
&\quad\times\left[1 - \frac{1}{2!}\mu_2\frac{H^{(2)}}{H} + \left(\frac{1}{2!}\mu_2\frac{H^{(2)}}{H}\right)^2 - \frac{1}{3!}\mu_3\frac{H^{(3)}}{H} - \frac{1}{4!}\mu_4\frac{H^{(4)}}{H}\right] + O[(n'+1)^{-3}]\\
&= G\left[1 + \frac{1}{2!}\mu_2\left(\frac{(GH)^{(2)}}{GH} - \frac{H^{(2)}}{H}\right) - \left(\frac{1}{2}\mu_2\right)^2\left(\frac{(GH)^{(2)}H^{(2)}}{GH.H} - \left(\frac{H^{(2)}}{H}\right)^2\right)\right.\\
&\quad\left. + \frac{1}{3!}\mu_3\left(\frac{(GH)^{(3)}}{GH} - \frac{H^{(3)}}{H}\right) + \frac{1}{4!}\mu_4\left(\frac{(GH)^{(4)}}{GH} - \frac{H^{(4)}}{H}\right)\right] + O[(n'+1)^{-3}]\\
&= G\left[1 + \frac{1}{2!}\mu_2\left(2\frac{G^{(1)}H^{(1)}}{GH} + \frac{G^{(2)}}{G}\right) - \left(\frac{1}{2}\mu_2\right)^2\left(2\frac{G^{(1)}H^{(1)}H^{(2)}}{GH^2} + \frac{G^{(2)}H^{(2)}}{GH}\right)\right.\\
&\quad + \frac{1}{3!}\mu_3\left(3\frac{G^{(1)}H^{(2)}}{GH} + 3\frac{G^{(2)}H^{(1)}}{GH} + \frac{G^{(3)}}{G}\right)\\
&\quad\left. + \frac{1}{4!}\mu_4\left(4\frac{G^{(1)}H^{(3)}}{GH} + 6\frac{G^{(2)}H^{(2)}}{GH} + 4\frac{G^{(3)}H^{(1)}}{GH} + \frac{G^{(4)}}{G}\right)\right] + O[(n'+1)^{-3}], \qquad (2.28)
\end{aligned}$$

where all the functions are evaluated at $\pi_{i'}$. Multiplying the terms within the bracket in the above equation by G, and by taking out $G^{(2)}$ from the third, the fourth and the fifth terms, we obtain

$$\begin{aligned}
&E(X_{i:n})\\
&= G + \mu_2 G^{(1)}\left[\frac{H^{(1)}}{H} + \frac{1}{2}\frac{G^{(2)}}{G^{(1)}}\right] + \frac{1}{2}G^{(2)}\left[-\mu_2^2\frac{H^{(2)}}{H}\left(2\frac{G^{(1)}H^{(1)}}{G^{(2)}H} + 1\right)\right.\\
&\quad + \frac{1}{3}\mu_3\left(3\frac{G^{(1)}H^{(2)}}{G^{(2)}H} + 3\frac{H^{(1)}}{H} + \frac{G^{(3)}}{G^{(2)}}\right)\\
&\quad\left. + \frac{1}{12}\mu_4\left(4\frac{G^{(1)}H^{(3)}}{G^{(2)}H} + 6\frac{H^{(2)}}{H} + 4\frac{G^{(3)}H^{(1)}}{G^{(2)}H} + \frac{G^{(4)}}{G^{(3)}}\cdot\frac{G^{(3)}}{G^{(2)}}\right)\right] + O\left[(n'+1)^{-3}\right].
\end{aligned}$$

Use of Equations (2.6)–(2.8) and (2.11)–(2.14) gives

$$\begin{aligned}
E(X_{i:n}) \;=\;& G(\pi_{i'}) + \frac{1}{n'+2}G^{(1)}(\pi_{i'})Q_1(\pi_{i'};\alpha,\beta)\\
&+ \frac{1}{2(n'+1)^2}G^{(1)}(\pi_{i'})Q_2(\pi_{i'};\alpha,\beta) + O\left[(n'+1)^{-3}\right],
\end{aligned} \tag{2.29}$$

which completes the proof. □

Note that Weke *et al.* (2001) used the above approach but expanded $G(u)H(u)$ only up to the second term.

DEFINITION 3. If, for arbitrarily small $\varepsilon > 0$, the function $G(\cdot)$ satisfies

$$G(u) = -c_0\left(\ln\frac{1}{u}\right)^k\left[1 + O\left\{\left(\ln\frac{1}{u}\right)^{-1+\varepsilon}\right\}\right] \tag{2.30}$$

or

$$G(u) = c_0\left(\ln\frac{1}{1-u}\right)^k\left[1 + O\left\{\left(\ln\frac{1}{1-u}\right)^{-1+\varepsilon}\right\}\right] \tag{2.31}$$

for some positive constants c_0 and k, then $G(\cdot)$ is called *asymptotic logarithm transform* (abbreviated as ALT) at $u = 0$ and $u = 1$, and a random variable having distribution G^{-1} is called an AL-variable.

Furthermore, if the error terms in formulae (2.30) and (2.31) satisfy

$$\left[O\left\{\left(\ln\frac{1}{u}\right)^{-1+\varepsilon}\right\}\right]^{(r)} = \left[\left\{\left(\ln\frac{1}{u}\right)^{-1}\right\}^{(r)}\right] O\left[\left(\ln\frac{1}{u}\right)^{-1+\varepsilon}\right], \quad \text{for } r = 1,2,3,4, \tag{2.32}$$

and

$$\left[O\left\{\left(\ln\frac{1}{1-u}\right)^{-1+\varepsilon}\right\}\right]^{(r)} = \left[\left\{\left(\ln\frac{1}{1-u}\right)^{-1}\right\}^{(r)}\right] O\left[\left(\ln\frac{1}{1-u}\right)^{-1+\varepsilon}\right], \quad \text{for } r = 1,2,3,4, \tag{2.33}$$

then from the relations (2.30) and (2.31), it can be deduced that the r-th derivatives of $G(u)$ satisfy

$$G^{(r)}(u) = -c_0\left[\left(\ln\frac{1}{u}\right)^{k}\right]^{(r)}\left[1+O\left\{\left(\ln\frac{1}{u}\right)^{-1+\varepsilon}\right\}\right], \quad \text{for } r = 1,2,3,4 \tag{2.34}$$

and

$$G^{(r)}(u) = c_0\left[\left(\ln\frac{1}{1-u}\right)^{k}\right]^{(r)}\left[1+O\left\{\left(\ln\frac{1}{1-u}\right)^{-1+\varepsilon}\right\}\right], \quad \text{for } r=1,2,3,4. \tag{2.35}$$

Let $L_r(0) = \lim_{u\to 0+} L_r(u)$, and $L_r(1) = \lim_{u\to 1-} L_r(u)$, $r = 1,2,3$. Then, it can be obtained from the relations (2.34) and (2.35) that

$$L_r(0) = -r, \quad r = 1,2,3 \tag{2.36}$$

and

$$L_r(1) = r, \quad r = 1,2,3. \tag{2.37}$$

Let $K_r(0) = \lim_{u\to 0+} K_r(u)$ and $K_r(1) = \lim_{u\to 1-} K_r(u)$, $r = 1,2,3$. Then, it can be obtained by using the relation (2.12) that

$$K_1(0) = \alpha, \quad K_2(0) = \alpha(\alpha-1) \quad \text{and} \quad K_3(0) = \alpha(\alpha-1)(\alpha-2) \tag{2.38}$$

and

$$K_1(1) = -\beta, \quad K_2(1) = \beta(\beta-1) \quad \text{and} \quad K_3(1) = -\beta(\beta-1)(\beta-2). \tag{2.39}$$

Thus, it can be derived from relations (2.36)–(2.39) and (2.13) that

$$\begin{aligned} Q_1(c;\alpha,\beta) &= \alpha - 1/2 + O(c) &(2.40)\\ &= -\beta - 1/2 + O(1-c). &(2.41)\end{aligned}$$

By choosing $\alpha = \beta = 1/2$, we have

$$Q_1(c;1/2,1/2) = O[c(1-c)]. \tag{2.42}$$

Similarly, by considering equations (2.36)–(2.39) and (2.14), the following equation is obtained

$$\begin{aligned} Q_2(\pi_{i'};1/2,1/2) &= \frac{1}{6}\left(1-\frac{3}{n}\right) - \frac{1}{4}\left[\left(1-\frac{10}{n}\right) + \frac{2}{\Delta_t}\left(1-\frac{6}{n}\right)\right] \\ &\quad + O\left[c(1-c)\right] + O(n^{-1}), &(2.43)\end{aligned}$$

where $\Delta_t = \min[i-1/2, n-i+1/2]$.

Based on the preceding work, the theorem below is stated and proved.

THEOREM 2.2. *Suppose that conditions (a), (b) and (d) in Theorem 2.1 are satisfied, and*

(a) $G(u)$ is ALT at both $u = 0$ and $u = 1$,

(b) equations (2.36) and (2.37) hold, and

(c) i, the integer part of $nc + 0.5$, satisfies $0 < i < n$.

Then,

$$\begin{aligned} &E(X_{i:n} + X_{i+1:n}) \\ &= 2\left[G(c) + G^{(1)}(c)\left\{\left(\frac{i}{n} - c\right) + \frac{O[c(1-c)]}{n+1}\right\}\right. \\ &\quad \left. + \frac{G^{(2)}(c)}{2n^2}\left\{\frac{1}{6} + (i-nc)^2 - \frac{1}{4}\left(\frac{1}{\Delta_i} + \frac{1}{\Delta_{i+1}}\right) + [1+2(i-nc)]O[c(1-c)]\right\}\right] \\ &\quad + O(n^{-3}). &(2.44)\end{aligned}$$

PROOF. From condition (c), $\{X_{i:n}\}$ and $\{X_{i+1:n}\}$ are c-sequences. Suppose that conditions (a), (b) and (d) in Theorem 2.1 are satisfied. Using

Equation (2.29), the following equation is obtained.

$$\begin{aligned} &E(X_{i:n}+X_{i+1:n}) \\ &= \sum_{k=i'}^{i'+1}\left\{G(\pi_k)+\frac{1}{n'+2}G^{(1)}(\pi_k)Q_1(\pi_k;\alpha,\beta)+\frac{1}{2(n'+1)^2}G^{(2)}(\pi_k)Q_2(\pi_k;\alpha,\beta)\right\} \\ &\quad +O\left[(n'+1)^{-3}\right]. \end{aligned} \tag{2.45}$$

Suppose that $\alpha=\beta=1/2$. By using conditions (a) and (b) of the theorem and substituting equations (2.42) and (2.43) into relation (2.45), we obtain

$$\begin{aligned} E(X_{i:n}+X_{i+1:n}) &= \sum_{k=i'}^{i'+1}\left[G(\pi_k)+\frac{1}{n+1}G^{(1)}(\pi_k)O[c(1-c)]\right. \\ &\quad \left.+\frac{1}{2n^2}G^{(2)}(\pi_k)\left\{-\frac{1}{12}-\frac{1}{2\Delta_k}+O[c(1-c)]\right\}\right] \\ &\quad +O\left(n^{-3}\right). \end{aligned} \tag{2.46}$$

Expanding $G(\pi_k)$, $G^{(1)}(\pi_k)$ and $G^{(2)}(\pi_k)$ at $\pi_k=c$ up to the second, the first and the constant terms, respectively, and putting together terms with higher order in $O(n^{-3})$, we obtain

$$\begin{aligned} &E(X_{i:n}+X_{i+1:n}) \\ &= 2G(c)+G^{(1)}(c)(\pi_{i'}+\pi_{i'+1}-2c)+\frac{1}{2}G^{(2)}(c)\left[(\pi_{i'}-c)^2+(\pi_{i'+1}-c)^2\right] \\ &\quad +2G'(c)\frac{O[c(1-c)]}{n+1}+G^{(2)}(c)(\pi_{i'}+\pi_{i'+1}-2c)\frac{O[c(1-c)]}{n+1} \\ &\quad +2\frac{1}{2n^2}G^{(2)}(c)\left\{-\frac{1}{12}-\frac{1}{4}\left(\frac{1}{\Delta_i}+\frac{1}{\Delta_{i+1}}\right)+O[c(1-c)]\right\}+O(n^{-3}) \\ &= 2\left[G(c)+G^{(1)}(c)\left\{\left(\frac{i}{n}-c\right)+\frac{O[c(1-c)]}{n+1}\right\}\right. \\ &\quad \left.+\frac{1}{2n^2}G^{(2)}(c)\left\{\frac{1}{6}+(i-nc)^2-\frac{1}{4}\left(\frac{1}{\Delta_i}+\frac{1}{\Delta_{i+1}}\right)+[1+2(i-nc)]O[c(1-c)]\right\}\right] \\ &\quad +O(n^{-3}). \end{aligned} \tag{2.47}$$

Hence the proof. □

Finally, it can easily be seen that for the standard logistic distribution,

$$\begin{aligned} G(u) &= \log\left(\frac{u}{1-u}\right) = \log\left(\frac{u}{1-u}\right)-\left(\frac{1}{u}\right) \\ &= c_0\left(\ln\frac{u}{1-u}\right)^k\left[1+O\left\{\ln\left(\frac{1}{1-u}\right)^{-1+\varepsilon}\right\}\right] \\ &\quad -c_0\left(\ln\frac{1}{u}\right)^k\left[1+O\left\{\ln\left(\frac{1}{u}\right)^{-1+\varepsilon}\right\}\right], \end{aligned} \tag{2.48}$$

for $c_0=k=1$. Hence, $G(u)$ is asymptotic logarithm transform at both $u=0$ and $u=1$. Hence, it follows that the function GH is bounded. Consequently, it is concluded that a random variable having the logistic distribution is an AL-variable. The inverse function, $G(u)$, of the logistic distribution also satisfies condition (b) in Theorem 2.2. Thus if i is the integer part of $nc+0.5$, then equation (2.44) can be applied.

3 Estimator of Standard Deviation

The method discussed here is based on a single spacing and the expectation of the sum of consecutive order statistics in a sample of size n. Let a spacing c be defined in relation to the rank i of the order statistic $X_{i:n}$ and sample size n, so that i is the integer part of $nc+0.5$. Since $\lim_{n\to\infty} i/n=c$, it can therefore be easily shown that $\hat{\sigma}_1$, the desired estimator of σ, is asymptotically unbiased. Using the fact that the inverse function is

$$G(u)=F^{-1}(u;0,1)=\tau\ln\left(\frac{u}{1-u}\right),\quad \tau=\frac{\sqrt{3}}{\pi},\quad 0<u<1, \tag{3.1}$$

and by noting that $\frac{i-0.5}{n}\le c\le\frac{i+0.5}{n}$, the value $c^*=i/n$ is chosen as the expansion point for the expectation in (2.44). Thus, equation (2.44) simplifies to

$$\begin{aligned} E(X_{i:n}+X_{i+1:n}) &= 2\left[G(c^*)+\frac{G^{(2)}(c^*)}{2n^2}\left\{\frac{1}{6}-\frac{1}{4}\left(\frac{1}{\Delta_i}+\frac{1}{\Delta_{i+1}}\right)\right\}\right] \\ &\quad +O(n^{-3}). \end{aligned} \tag{3.2}$$

An approximate expression for $-E(X_{i:n}+X_{i+1:n})$ is

$$E_i=2\tau\left[\ln\left(\frac{n-i}{i}\right)+\frac{n(n-2i)}{2i^2(n-i)^2}\left\{\frac{1}{6}-\frac{1}{4}\left(\frac{1}{\Delta_i}+\frac{1}{\Delta_{i+1}}\right)\right\}\right]. \tag{3.3}$$

Then, $E_i = -E_{n-i}$ due to symmetry.

Let R_i denote the i-th sample quasi-range:

$$R_i = X_{n-i:n} - X_{i+1:n}, \quad 1 \leq i < \frac{n-1}{2}. \tag{3.4}$$

It follows that if $X_{1:n}, \ldots, X_{n:n}$ are order statistics from $F(x; \mu, \sigma)$, then the expression given in (3.3) is an approximation of $E(R_i + R_{i-1})/(2\sigma)$. Based on the above discussion, the statistic

$$\hat{\sigma}_1 = \frac{R_i + R_{i-1}}{2E_i} \tag{3.5}$$

is proposed as an estimator of the scale parameter σ. The index i depends on the constant c through the relation that i is the integer part of $nc + 0.5$. The spacing value c should be chosen from the interval $(0, 1/2)$ such that the variance of $\hat{\sigma}_1$ in (3.5) is small.

4 Discussion

The estimator $\hat{\sigma}_1$ given in (3.5) is similar to the estimator proposed by Ogawa (1951) in the case of known μ, which is based on $X_{n-i:n} - X_{i:n}$, i.e., difference between two order statistics. Gupta and Gnanadesikan (1966) showed that this estimator has the smallest variance when $\frac{i}{n} = 0.10293$. Monte Carlo simulations indicate that the choice of i as the integer part of $nc + 0.5$ with $c = 0.10293$, generally minimizes the variance of $\hat{\sigma}_1$ given in (3.5). This value of c is used in (3.5) for a comparative study.

The estimator $\hat{\sigma}_1$ is compared with the estimator $\hat{\sigma}$ of Chan et al. (1971), which is also based on four order statistics and, together with a corresponding estimator of μ, has the largest joint efficacy among all linear estimators based on four order statistics. We denote by $V(\hat{\sigma}_1)$ and $V(\hat{\sigma})$ the variances of $\hat{\sigma}_1$ and $\hat{\sigma}$, respectively.

Table 1 gives the bias and the variance of the estimator $\hat{\sigma}_1$, the variance of $\hat{\sigma}$ and the efficiencies relative to Cramer-Rao lower bound, the best linear unbiased estimator (BLUE) and the 'zero-one' linear estimators of the scale parameter for various values of i and n, computed from a simulation study with 200 runs. Note that the Cramer-Rao lower bound on the variance of an unbiased estimator of σ is $9/\{n(3+\pi^2)\}$ (Gupta and Gnanadesikan, 1966), the variance of the BLUE is given in Gupta et al. (1967) and the variance of the 'zero-one' linear estimator is given in Balakrishnan and Cohen (1991, p.255) for $2 \leq n \leq 20$.

TABLE 1. BIAS, VARIANCE AND EFFICIENCIES FOR VARIOUS i AND n

n	i	$Bias^2/\sigma^2$	$V(\hat{\sigma}_1)/\sigma^2$	$V(\hat{\sigma})/\sigma^2$	eff_1	eff_2	$eff_{0\text{-}1}$
5	1	0.0002	0.1720	0.1704	0.813	0.991	1.000
6	1	0.0001	0.1377	0.1429	0.847	0.995	0.999
7	1	0.0000	0.1160	0.1232	0.861	0.987	0.999
8	1	0.0000	0.1011	0.1052	0.865	0.973	0.999
9	1	0.0000	0.0903	0.0918	0.861	0.956	0.978
10	1	0.0000	0.0820	0.0850	0.853	0.937	0.958
11	1	0.0000	0.0754	0.0788	0.843	0.917	0.943
12	1	0.0000	0.0701	0.0711	0.831	0.898	0.920
13	1	0.0000	0.0658	0.0648	0.818	0.878	0.900
14	1	0.0000	0.0621	0.0634	0.805	0.860	0.881
15	2	0.0000	0.0600	0.0584	0.777	0.826	0.862
16	2	0.0000	0.0559	0.0541	0.782	0.828	0.843
17	2	0.0000	0.0525	0.0527	0.784	0.827	0.830
18	2	0.0000	0.0495	0.0491	0.784	0.825	0.827
19	2	0.0000	0.0470	0.0460	0.783	0.822	0.826
20	2	0.0000	0.0448	0.0439	0.781	0.818	0.818
25	3	0.0000	0.0372	0.0352	0.751	0.779	
30	3	0.0000	0.0310		0.752	0.775	
35	4	0.0000	0.0271		0.736	0.755	
40	4	0.0000	0.0238		0.736	0.752	
45	5	0.0000	0.0214		0.726	0.741	
50	5	0.0000	0.0193		0.726	0.737	
55	6	0.0000	0.0176		0.720	0.730	
60	6	0.0000	0.0162		0.720	0.718	
65	7	0.0000	0.0150		0.715	0.714	
∞		0.0000	$1.0227/n$		0.684		

It is seen from Table 1 that for large n, the efficiency of $\hat{\sigma}_1$ with respect to the Cramer Rao lower bound is 0.684. This value coincides with the small sample efficiency of the estimator of Ogawa (1951) with optimal i. Thus, there is a gain for small samples. The variance of $\hat{\sigma}_1$ is comparable to that of $\hat{\sigma}$, while the three efficiencies are not too small. Given that the BLUE, the zero-one estimator, and even the estimator of Chen et al. (1971) are based on complex computations and/or look-up tables for ranks/coefficients, the estimator $\hat{\sigma}_1$ appears to be an attractive alternative.

References

BALAKRISHNAN, N. and COHEN, A.C. (1991). *Order Statistics and Inference: Estimation Methods*, Academic Press, Boston.

CHAN, L.K., CHAN, N.N. and MEAD, E.R. (1971). Best linear unbiased estimates of the parameters of the logistic distribution based on selected order statistics. *J. Amer. Statist. Assoc.*, **66**, 889–892.

DIXON, W.J. (1957). Estimates of the mean and standard deviation of a normal population. *Ann. Math. Statist.*, **28**, 806–809.

GUPTA, S.S. and GNANADESIKAN, M. (1966). Estimation of the parameters of the logistic distribution, *Biometrika*, **53**, 565–570.

GUPTA, S.S., QUREISHI, A.S. and SHAH, B.K. (1967). Best linear unbiased estimators of the parameters of the logistic distribution using order statistics, *Technometrics*, **9**, 43–56.

OGAWA, J. (1951). Contributions to the theory of systematic statistics I, *Osaka Math. J.*, **3**, 175–213.

RAGHUNANDANAN, K. and SRINIVASAN, R. (1970). Simplified estimation of parameters in a logistic distribution, *Biometrika*, **57**, 677–679.

WANG, C.-G. (1980). The expectation of the sum of consecutive AL-variables and the double 'zero-one' linear estimates for the standard deviation of normal population. *J. Harbin Inst. Tech.*, **1**, 1–33 (in Chinese).

WEKE, P.G.O., WANG, C.-G. and WU, C.-X. (2001). Nearly best linear etimates of the logistic parameters based on complete ordered statistics. *J. Harbin Inst. Tech.*, **8**, 178–183.

WEKE, P.G.O. (2005). Linear estimation of standard deviation of logistic distribution: Theory and algorithm. *Far East J. Theor. Statist.* **15**, 199–214.

PATRICK G.O. WEKE
ACTUARIAL SCIENCE AND FINANCIAL MATHEMATICS DIVISION
SCHOOL OF MATHEMATICS
UNIVERSITY OF NAIROBI
P.O. BOX 30197, GPO 00100
NAIROBI, KENYA
E-mail: pweke@uonbi.ac.ke

Paper received January 2005; revised January 2007.

Sankhyā : The Indian Journal of Statistics
2007, Volume 69, Part 4, p. 885-886

Book Reviews

Modeling Financial Time Series with S-PLUS (2nd Edition)
ERIC ZIVOT AND JIAHUI WANG
(2006) Springer, (xxii)+998 pp.
Price € 59.95, $ 69.95, £ 46.00, ISBN 10-387-27965-2

Financial time series is a rapidly growing subject. Zivot and Wang deliver an impressive book covering many relevant topics on theoretical and empirical financial econometrics, statistics and time series. This book is the second edition of the previous version (Zivot and Wang, 2003). A few chapters, viz., chapters 18 through 23 are new and cover nonlinear time series models, copulas, continuous-time models for financial time series, semi-parametric and non-parametric conditional density models, and efficient method of moments.

The target audience comprises practitioners, researchers and students in empirical finance and financial econometrics. The book is of great help to practitioners looking for examples and tools to analyse their own time series. The book covers a vast area of ongoing research in financial econometrics and provides material on classical and modern, univariate and multivariate time series modeling and estimation methods. It covers an extensive and exhaustive area on unit root tests, co-integration tests, ARCH/GARCH (univariate as well as multivariate), extreme value theory, long memory, GMM etc. It provides a comprehensive list of references.

As it is based on S-PLUS, it is necessary to buy S-PLUS license with the *S+FinMetrics* module, which may require a significant investment. Basic familiarity with S-PLUS is also required. The related discussion in the book suffers somewhat from lack of uniformity and integrity as the libraries are taken from various sources.

In summary, the book should be useful for practitioners in financial econometrics, and also very useful as an easy source of references.