International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 6 (Special Issue), 2007, pp 101-107

On the Positivity of Best Linear Unbiased Estimators of Scale Based on Three Order Statistics

Dale Umbach
Dept. of Mathematical Sciences
Ball State University
Muncie, IN 47306, USA
E-mail:dumbach@bsu.edu

[Received November 4, 2005; Revised June 23, 2007; Accepted July 8, 2007]

Abstract

Suppose that $X_{n_1},\ X_{n_2},\ X_{n_3}$ are three order statistics with $1 \leq n_1 < n_2 < n_3 \leq n$ from a random sample of size n from a continuous location/scale distribution. Let μ_1 and σ_1^2 be the mean and variance of the spacing $X_{n_2} - X_{n_1}$. Let μ_2 and σ_2^2 be the mean and variance of the spacing $X_{n_3} - X_{n_2}$. Let ρ be the correlation between the spacings $X_{n_2} - X_{n_1}$ and $X_{n_3} - X_{n_2}$. It is established that the Best Linear Unbiased Estimator for the scale parameter is positive with probability 1 if, and only if, $\mu_1\sigma_2 \geq \rho\mu_2\sigma_1$ and $\mu_2\sigma_1 \geq \rho\mu_1\sigma_2$. Related results are given for the asymptotic case.

Keywords and Phrases: Scale parameter, Positivity, BLUE, ABLUE, Spacings

AMS Classification: 62F10.

1 Introduction

Let $X_{n_1} < X_{n_2} < X_{n_3}$, for $1 \leq n_1 < n_2 < n_3 \leq n$ represent three order statistics of a random sample of size n from a distribution with location parameter λ and scale parameter δ whose distribution function is given by $F_{\lambda,\delta}(x) = F((x-\lambda)/\delta)$, for some distribution function F, with associated density f. Here we assume $-\infty < \lambda < \infty$ and $0 < \delta < \infty$. We consider linear estimators of δ based on these order statistics, say $\hat{\delta} = \sum_{i=1}^{3} b_i X_{n_i}$. Since $\delta > 0$ in this model, a desirable property of $\hat{\delta}$ would be that $\Pr(\hat{\delta} > 0) = 1$. However, there are reasonable estimators of δ that do not have this property. See Sarkar and Wang (1998) for a good discussion of this.

Bai, Sarkar, and Wang (1997) establish the positivity of the Best Linear Unbiased Estimator (BLUE) of scale for arbitrary sample size for the class of log-convex distributions. The positivity of the BLUE and the Asymptotically Best Linear Unbiased Estimator (ABLUE) for samples of size 2 for arbitrary distributions are easy to establish as is pointed out in their paper, as well. However, while generally accepted as true, neither the positivity of the BLUE nor the ABLUE for scale have been established outside of the log-convex case for sample sizes exceeding 2.

In this paper, we prove the following, which presents necessary and sufficient conditions for $\Pr(\hat{\delta} > 0) = 1$.

Theorem 1: With $X_{n_1}, X_{n_2}, X_{n_3}$ as above, let $\mu_1 = \mathrm{E}(X_{n_2} - X_{n_1})$, $\sigma_1^2 = \mathrm{Var}(X_{n_2} - X_{n_1})$, $\mu_2 = \mathrm{E}(X_{n_3} - X_{n_2})$, $\sigma_2^2 = \mathrm{Var}(X_{n_3} - X_{n_2})$, and $\rho = \mathrm{Cor}((X_{n_2} - X_{n_1}), (X_{n_3} - X_{n_2}))$. Then, the BLUE for δ is positive with probability 1 if, and only if,

$$\frac{\sigma_1\,\mu_2}{\mu_1\,\sigma_2} \geq \rho \qquad \text{and} \qquad \frac{\sigma_2\,\mu_1}{\mu_2\,\sigma_1} \geq \rho. \tag{1}$$

The spacings of many commonly encountered distributions are negatively correlated and the spacings of order statistics from the exponential distribution are uncorrelated. Such distributions satisfy the conditions of the theorem, thereby guaranteeing positivity. However, for some distributions the spacings are positively correlated. The gamma distribution with shape parameter 1/2 provides an example. So, consider

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x^{-1/2}e^{-x}/\sqrt{\pi} & \text{for } 0 < x. \end{cases}$$

For a sample of size $n = 3$ from this distribution we have $\mathrm{Cov}((X_{(2)} - X_{(1)}), (X_{(3)} - X_{(2)})) = 0.0201875$. We note, however, that $\mu_1 = 0.255873$, $\sigma_1 = 0.255873$, $\mu_2 = 0.699057$, and $\sigma_2 = 0.799535$. However, both $\frac{\sigma_1\,\mu_2}{\mu_1\,\sigma_2} = 1.12054$ and $\frac{\sigma_2\,\mu_1}{\mu_2\,\sigma_1} = 0.892425$ are greater than $\rho = 0.076996$. Thus, for this distribution the BLUE for scale based on three order statistics will be positive with probability 1, even though the spacings are positively correlated.

The following lemma provides a foundation on which to attack the problem in both the exact and asymptotic cases. It allows one to attack the positivity problem by considering the sign of the partial sums of the coefficients of the estimator of scale. The proof can be found in Umbach (2002).

Lemma: Let $\mathbf{d} = (d_1, d_2, \ldots, d_k)' \in \Re^k$ such that $\sum_{i=1}^{k} d_i = 0$. Then $\mathbf{d} \cdot \mathbf{x} \geq 0$ for all $\mathbf{x} = (x_1, x_2, \ldots, x_k)' \in \Re^k$ with $x_1 \leq x_2 \leq \cdots \leq x_k$ if, and only if, $\sum_{i=1}^{r} d_i \leq 0$ for $r = 1, 2\ \ldots, k-1$.

2 The Exact Case

Let $m_i = \mathrm{E}_{0,1}(X_{n_i})$ and $s_{ij} = \mathrm{Cov}_{0,1}(X_{n_i}, X_{n_j})$ for $i = 1,\ 2,\ 3$ and $j = 1,\ \ldots, i$ where the (0,1) indicates that the moments of these order statistics are calculated using the standardized common distribution $F_{0,1}(x) = F(x)$. For $\sum_{i=1}^{3} b_i X_{n_i}$ to be unbiased for δ we must have

$$\delta = \mathrm{E}(\sum_{i=1}^{3} b_i X_{n_i}) = \sum_{i=1}^{3} b_i(\lambda + \delta m_i) = \lambda \sum_{i=1}^{3} b_i + \delta \sum_{i=1}^{3} b_i m_i$$

for all λ and δ. Thus, we must have $\sum_{i=1}^{3} b_i = 0$ and $\sum_{i=1}^{3} b_i m_i = 1$. Using these two equations, we can solve for b_1 and b_3 in terms of b_2 as

$$b_1 = \frac{-1}{m_3 - m_1} - b_2 \frac{m_3 - m_2}{m_3 - m_1} \tag{2}$$

$$b_3 = \frac{1}{m_3 - m_1} - b_2 \frac{m_2 - m_1}{m_3 - m_1}. \tag{3}$$

Since unbiasedness requires $b_1 + b_2 + b_3 = 0$, the lemma gives $\sum_{i=1}^{3} b_i X_{n_i} \geq 0$ with probability 1 if and only if $b_1 \leq 0$ and $b_1 + b_2 \leq 0$. Writing these inequalities in terms of b_2 yields

$$\frac{-1}{m_3 - m_2} \leq b_2 \leq \frac{1}{m_2 - m_1}. \tag{4}$$

Thus, Theorem 1 is proved by showing that b_2 for the BLUE satisfies this inequality precisely when (1) holds.

Let $g(b_2) = \mathrm{Var}(\sum_{i=1}^{3} b_i X_{n_i})$ with b_1 as in (2) and b_3 as in (3). The value of b_2 for the BLUE of δ, say b_2^*, minimizes the quadratic form $g(b_2)$. It can be shown that $(m_3 - m_1)^2 g(b_2) =$

$$\begin{aligned}
&b_2^2 \left\{ m_3^2(-2s_{12} + s_{22}) + 2m_2m_3(s_{12} + s_{13} - s_{23}) + m_2^2(-2s_{13} + s_{33}) \right.\\
&+m_1^2(s_{22} - 2s_{23} + s_{33}) - 2m_1m_3(-s_{12} + s_{13} + s_{22} - s_{23})\\
&\left. -2m_1m_2(s_{12} - s_{13} - s_{23} + s_{33}) \right\} + (1 + b_2(m_3 - m_2))^2 s_{11}\\
&+2b_2(-m_3(s_{12} + s_{13} - s_{23}) + m_2(2s_{13} - s_{33}) + m_1(s_{12} - s_{13} - s_{23} + s_{33}))\\
&-2s_{13} + s_{33}.
\end{aligned}$$

The derivative is given by $(m_3 - m_1)^2 g'(b_2) =$

$$\begin{aligned}
&2b_2 \left\{ m_3^2(s_{11} - 2s_{12} + s_{22}) + 2m_2m_3(s_{12} + s_{13} - s_{23} - s_{11}) + m_2^2(s_{11} - 2s_{13} + s_{33}) \right.\\
&+m_1^2(s_{22} - 2s_{23} + s_{33}) - 2m_1m_3(-s_{12} + s_{13} + s_{22} - s_{23})\\
&\left. -2m_1m_2(s_{12} - s_{13} - s_{23} + s_{33}) \right\}\\
&-2m_3(s_{12} + s_{13} - s_{23} - s_{11}) + 2m_2(2s_{13} - s_{33} - s_{11}) + 2m_1(s_{12} - s_{13} - s_{23} + s_{33}).
\end{aligned}$$

Thus, (4) will hold for the BLUE if, and only if, $g'(-1/(m_3-m_2)) \leq 0$ and $g'(1/(m_2-m_1)) \geq 0$.

A bit of algebra yields $(m_3-m_1)(m_3-m_2)g'(-1/(m_3-m_2))/2 =$

$$m_3(s_{12}-s_{13}-s_{22}+s_{23})+m_2(-s_{12}+s_{13}+s_{23}-s_{33})+m_1(s_{22}-2s_{23}+s_{33}) \tag{5}$$

and $(m_3-m_1)(m_2-m_1)g'(1/(m_2-m_1))/2 =$

$$m_3(s_{11}-2s_{12}+s_{22})+m_2(-s_{11}+s_{12}+s_{13}-s_{23})+m_1(s_{12}-s_{13}-s_{22}+s_{23}) \tag{6}$$

Since $m_1 < m_2 < m_3$, we see that $g'(-1/(m_3-m_2))$ will be negative precisely when (5) is negative and $g'(1/(m_2-m_1))$ will be positive precisely when (6) is positive.

Now, note that

$$\begin{aligned}
&\mathrm{E}_{0,1}\left[\mu_2(X_{n_2}-X_{n_1})^2-\mu_1(X_{n_2}-X_{n_1})(X_{n_3}-X_{n_2})\right]\\
&=(m_3-m_2)(s_{22}-2s_{12}+s_{11})-(m_2-m_1)(s_{23}-s_{22}-s_{13}+s_{12})\\
&=m_3(s_{22}-2s_{12}+s_{11})+m_2(s_{12}-s_{11}-s_{23}+s_{13})+m_1(s_{23}-s_{22}-s_{13}+s_{12}),
\end{aligned}$$

which is identical to (6). Thus $g'(1/(m_2-m_1))$ will be positive precisely when

$$\mathrm{E}_{0,1}\left[\mu_2(X_{n_2}-X_{n_1})^2-\mu_1(X_{n_2}-X_{n_1})(X_{n_3}-X_{n_2})\right] \;\geq\; 0.$$

This is precisely when

$$\mu_2(\sigma_1^2+\mu_1^2) \;\geq\; \mu_1\rho\sigma_1\sigma_2+\mu_1^2\mu_2,$$

which is precisely when

$$\mu_2\sigma_1^2 \;\geq\; \mu_1\rho\sigma_1\sigma_2,$$

which is precisely when

$$\frac{\sigma_1\mu_2}{\mu_1\sigma_2} \geq \rho.$$

In a similar fashion, one can show that $g'(-1/(m_3-m_2)) \leq 0$ precisely when

$$\frac{\sigma_2\mu_1}{\mu_2\sigma_1} \geq \rho.$$

3 The Asymptotic Case

The asymptotic case is based a fixed set (p_1, p_2, p_3) for which $0 < p_1 < p_2 < p_3 < 1$. Such a set is called a spacing of the order statistics to be used for estimation. Suppose that $n_i = [n\,p_i] + 1$, $u_i = F^{-1}(p_i)$, and $f_i = f(u_i)$ for $i = 1,\ 2,\ 3$, where $[\cdot]$ is the greatest integer function. If f and f' are continuous in a neighborhood of u_i for $i = 1,\ 2,\ 3$, then Mosteller (1946) gives asymptotic first moments of $X_{n_1},\ X_{n_2},\ X_{n_3}$ as $\lambda + \delta\, u_i$ and asymptotic second moments by

$$\mathrm{Cov}(X_{n_i}, X_{n_j}) = \frac{\delta^2}{n}\frac{p_i(1-p_j)}{f_i\, f_j} \text{ for } i \le j = 1,\ 2,\ 3.$$

Using this result, Theorem 1 can be used with only slight modification. In particular, u_i now plays the role of m_i. The quadratic form g that is to be minimized is now the asymptotic version, but still has the same form. Thus, we may simply replace all of the moments by their asymptotic counterparts. Thus, we see that the ABLUE for scale is positive with probability 1 if, and only if, both $\frac{\mu_1\sigma_2}{\mu_2\sigma_1}$ and $\frac{\mu_2\sigma_1}{\mu_1\sigma_2}$ are greater than or equal to ρ, where

$$\begin{aligned}
\mu_1 &= u_2 - u_1\\
\mu_2 &= u_3 - u_2\\
\sigma_1^2 &= \frac{p_1(1-p_1)}{f_1^2} + \frac{p_2(1-p_2)}{f_2^2} - 2\frac{p_1(1-p_2)}{f_1\, f_2}\\
\sigma_2^2 &= \frac{p_2(1-p_2)}{f_2^2} + \frac{p_3(1-p_3)}{f_3^2} - 2\frac{p_2(1-p_3)}{f_2\, f_3}\\
\rho\sigma_1\sigma_2 &= \frac{p_2(1-p_3)}{f_2\, f_3} - \frac{p_1(1-p_3)}{f_1\, f_3} - \frac{p_2(1-p_2)}{f_2^2} + \frac{p_1(1-p_2)}{f_1\, f_2}.
\end{aligned}$$

Unfortunately, the expressions do not simplify much after these substitutions are made. However, the Lemma can be used in a different way to get the results in Theorem 2. It is based on the fact that we have useful expressions for the coefficients in the asymptotic case. Using Sarhan and Greenberg (1962), let $p_0 = 0$, $p_4 = 1$, $f_0 = f_0\, u_0 = f_4 = f_4\, u_4 = 0$,

$$\begin{aligned}
K_1 &= \sum_{i=1}^{4} \frac{(f_i - f_{i-1})^2}{p_i - p_{i-1}},\\
K_2 &= \sum_{i=1}^{4} \frac{(f_i u_i - f_{i-1} u_{i-1})^2}{p_i - p_{i-1}},\\
K_3 &= \sum_{i=1}^{4} \frac{(f_i u_i - f_{i-1} u_{i-1})(f_i - f_{i-1})}{p_i - p_{i-1}}, \text{ and}\\
\Delta &= K_1\, K_2 - K_3^2.
\end{aligned}$$

They showed that the coefficients in the ABLUE of the scale parameter can be written as

$$\begin{aligned} b_i &= \frac{f_i K_1}{\Delta}\left(\frac{f_i u_i - f_{i-1}u_{i-1}}{p_i - p_{i-1}} - \frac{f_{i+1}u_{i+1} - f_i u_i}{p_{i+1} - p_i}\right) \\ &\quad - \frac{f_i K_3}{\Delta}\left(\frac{f_i - f_{i-1}}{p_i - p_{i-1}} - \frac{f_{i+1} - f_i}{p_{i+1} - p_i}\right) \text{ for } i = 1, 2, 3. \end{aligned} \tag{7}$$

Theorem 2: With X_{n_1}, X_{n_2}, and X_{n_3} as above, the ABLUE for δ is positive with probability 1 if, and only if, both

$$f_1 p_2 (K_1 u_1 - K_3) \leq f_2 p_1 (K_1 u_2 - K_3) \tag{8}$$

and

$$f_2(1 - p_3)(K_1 u_2 - K_3) \leq f_3(1 - p_2)(K_1 u_3 - K_3). \tag{9}$$

Proof: As in the finite case, the asymptotic unbiasedness of $\hat{\delta} = \sum_{i=1}^{3} b_i X_{n_i}$ guarantees that $b_1 + b_2 + b_3 = 0$. Thus, by the Lemma, we will have $\Pr(\hat{\delta} > 0) = 1$ if both $b_1 \leq 0$ and $b_1 + b_2 \leq 0$. But, $b_3 = -(b_1 + b_2)$, so we need $b_1 \leq 0$ and $b_3 \geq 0$.

By (7), we have $b_1 \leq 0$ precisely when

$$\frac{f_1 K_1}{\Delta}\left(\frac{f_1 u_1}{p_1} - \frac{f_2 u_2 - f_1 u_1}{p_2 - p_1}\right) \leq \frac{f_1 K_3}{\Delta}\left(\frac{f_1}{p_1} - \frac{f_2 - f_1}{p_2 - p_1}\right),$$

which can be simplified to (8), as was to be shown. In a similar fashion, one can show that $b_3 \geq 0$ is equivalent to (9).

4 Discussion

The positivity of the BLUE and ABLUE for the scale parameter is still an open question. It has been considered by a number of authors. In particular, Bai, Sarkar, and Wang (1997) showed that the best unbiased L-estimator of scale is positive with probability 1 if either 1) the generating density function is log-concave, or 2) $n = 2$, or 3) $n = 3$ and the generating density function is symmetric about the origin and the censored sample is symmetric about the origin. In particular, they note that the spacings are negatively correlated for log-concave distributions, which is consistent with the theorem since

$$\frac{\sigma_1\,\mu_2}{\mu_1\,\sigma_2} \geq \rho \quad\text{and}\quad \frac{\sigma_2\,\mu_1}{\mu_2\,\sigma_1} \geq \rho$$

are trivially satisfied when $\rho \leq 0$.

The theorems present necessary and sufficient conditions for positivity. Perhaps these equivalent conditions will provide a means for establishing the positivity of the BLUE and the ABLUE for scale in general.

References

1. Bai, Z. D., S. K. Sarkar, and W. Wang. (1997). Positivity of the best unbiased L-estimator of the scale parameter with complete or selected order statistics from location-scale distribution. *Statistics & Probability Letters,* 32, 181-188.

2. Mosteller, F. (1946). On some useful "inefficient" statistics. *Annals of Mathematical Statistics,* 17, 377-408.

3. Sarhan, A. E. and Greenberg, B. G. (eds.) (1962). *Contributions to Order Statistics.* John Wiley and Sons:New York.

4. Sarkar, S. K. and W. Wang. (1998). Estimation of scale parameter based on a fixed set of order statistics. In *Handbook of Statistics V17: Order Statistics and Their Applications.* N. Balakrishnan and C. R. Rao (eds.) Elsevier Science Publishers B. V.: Amsterdam, 159-181.

5. Umbach, D. (2002). On the positivity of linear combinations of order statistics as estimators of scale. *Journal of Statistical Studies,* October, 231-237.