A semi-parametric bootstrap-based best linear unbiased estimator of location under symmetry

Alan D. Hutson\textsuperscript{a,b}

\textsuperscript{a}Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, New York, USA

\textsuperscript{b}Biostatistics, University at Buffalo, Buffalo, New York, USA

Abstract

In this note we provide a novel semi-parametric best linear unbiased estimator (BLUE) of location and its corresponding variance estimator under the assumption the random variate is generated from a symmetric location-scale family of distributions. The approach follows in a two-stage fashion and is based on the exact bootstrap estimate of the covariance matrix of the order statistic. We generalize our approach to add a robustness component in order to derive a trimmed BLUE of location under a semi-parametric symmetry assumption.

Keywords

Bootstrap; Mid-range; Order statistics; Robust estimator

1. Introduction

There is a substantial literature on best linear unbiased estimation (BLUE) based on order statistics for both uncensored and type II censored data, both grouped and ungrouped; See Balakrishnan and Rao (1997) for an introduction to the topic and Balakrishnan and Rao (2003), and the references therewithin, for a more technical discussion.

First, as background we outline the classic BLUE notation, which is derived from a generalized least-squares approach. Toward this end let \( X_{(1)} < X_{(2)} < \cdots < X_{(n)} \) denote the order statistics from an i.i.d. sample from a continuous c.d.f. \( F \). The c.d.f. \( F \) is assumed to come from a location-scale family such that \( F(x) = F_0((x - \mu)/\sigma) \ (\sigma > 0) \), or similarly having a quantile function of the form \( F^{-1}(u) = \mu + \sigma F_0^{-1}(u) \). Also, denote \( Y_{(i)} = (X_{(i)} - \mu)/\sigma \), with known \( n \times 1 \) mean vector \( \alpha \), with elements \( \alpha_i = E(Y_{(i)}) \), and known \( n \times n \) variance-covariance matrix \( B \). Let \( A = (1, \alpha) \), where \( 1 \) is an \( n \times 1 \) column vector of 1’s. Denote \( \Omega = B^{-1} \) as the inverse of the variance-covariance matrix \( B \) corresponding to standardized order statistics within vector \( Y \). If we denote the \( n \times 1 \) vector of ordered
observations as \( X=(X_1, X_2, \ldots, X_n)' \) then it is well known, e.g., see David (1981), that
the BLUE estimators are given as
\[
\hat{\mu} = -\alpha^T X, \quad (1.1)
\]
\[
\hat{\sigma} = 1^T X, \quad (1.2)
\]
where
\[
\Gamma = \frac{\Omega(1\alpha' - \alpha'1)}}{\Delta},
\]
\[
\Delta = |A^T \Omega A|.
\]
If follows from least-squares theory that
\[
\text{Var}(\hat{\mu}) = \frac{\alpha^T \Omega \alpha \sigma^2}{\Delta}, \quad (1.4)
\]
\[
\text{Var}(\hat{\sigma}) = \frac{1^T \Omega \sigma^2}{\Delta}, \quad (1.5)
\]
where \( \sigma^2 \) is known and \( \Delta \) is defined above.

The limitation of the BLUE approach for location-scale families is that one has to assume a parametric form for \( F_0 \). Hence, it is generally not well-used when current statistical software can be used to estimate \( \mu \) and \( \sigma \) using more popular approaches such as maximum likelihood based on assuming the same \( F_0 \). However, if we make one straightforward assumption that \( F_0 \) represents the c.d.f. of a symmetric parent distribution then we can estimate \( \mu \) and \( \text{Var}(\hat{\mu}) \) using a completely novel two-stage BLUE semi-parametric approach. Unlike the parametric approach our method also has a trimmed location estimator counterpart. In general given that \( \mu \) is typically the parameter of interest in statistical inference procedures, e.g., a one-sample t-test or sign-test, this approach provides a robust and meaningful BLUE estimate of \( \mu \) based on the exact bootstrap estimate of the variance-covariance matrix for \( X_1, X_2, \ldots, X_n \) and given in Sec. 2. In Sec. 3 we define the new BLUE semi-parametric estimator of location along with a trimmed counterpart. This estimator could readily be used in exact permutation testing under the symmetry assumption.

2. Exact bootstrap variance estimators

The results in this section were first derived in Hutson and Ernst (2000) and are critical to the results presented in this paper. Let \( (\hat{\mu}_{1:n}, \hat{\mu}_{2:n}, \ldots, \hat{\mu}_{n:n})' \) denote the bootstrap mean vector of the order statistics, where the exact bootstrap estimate of \( \mu_{r:n} \), \( 1 \leq r \leq n \), may be written as
\[ \hat{\mu}_{r:n} = \sum_{j=1}^{n} w_{j(r)} X_{j:n}, \quad \text{(2.1)} \]

\[ w_{j(r)} = r \left( \frac{n}{j} \right) B \left( \frac{j}{n}; r, n - r + 1 \right) - B \left( \frac{j - 1}{n}; r, n - r + 1 \right). \quad \text{(2.2)} \]

and \( B(x; a, b) = \int_0^x t^{a-1} (1-t)^{b-1} \, dt \) is the incomplete beta function.

Furthermore, denote the \( n \times n \) bootstrap estimated covariance matrix of the order statistics elements of vector \( X \) as

\[ \tilde{\Sigma} = \begin{pmatrix} \hat{\sigma}_{11:n} & \hat{\sigma}_{12:n} & \cdots & \hat{\sigma}_{1n:n} \\ \hat{\sigma}_{21:n} & \hat{\sigma}_{22:n} & \cdots & \hat{\sigma}_{2n:n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{n1:n} & \hat{\sigma}_{n2:n} & \cdots & \hat{\sigma}_{nn:n} \end{pmatrix}, \quad \text{(2.3)} \]

where

\[ \hat{\sigma}_{r:n}^2 = \sum_{j=1}^{n} w_{j(r)} (X_{j:n} - \hat{\mu}_{r:n})^2, \quad \text{(2.4)} \]

\( w_{j(r)} \) is defined at (2.2),

\[ \hat{\sigma}_{rs:n} = \sum_{j=2}^{n} \sum_{i=1}^{j-1} w_{ij(rs)} (X_{i:n} - \hat{\mu}_{r:n}) (X_{j:n} - \hat{\mu}_{s:n}) \quad \text{(2.5)} \]

\[ w_{ij(rs)} = \int_{(j-1)/n}^{j/n} \int_{(i-1)/n}^{i/n} f_{rs}(u_r, u_s) \, du_r \, du_s \quad \text{(2.6)} \]

\[ v_{j(rs)} = \int_{(j-1)/n}^{j/n} \int_{(j-1)/n}^{j/n} f_{rs}(u_r, u_s) \, du_r \, du_s, \quad \text{(2.7)} \]

and

\[ f_{rs}(u_r, u_s) = n C_{rs} u_r^{s-r-1} (1-u_r)^{s-r-1} (1-u_s)^{n-s}, \quad \text{(2.8)} \]

is the joint distribution of two uniform order statistics \( U_{r:n} \) and \( U_{s:n} \) with \( n C_{rs} = n! / [(r-1)! (s-r+1)!] \).
The weights \(w_{i(rs)}\) and \(v_{j(rs)}\) in (2.6) and (2.7) can be easily calculated by evaluating the integrals and writing them in closed form. The key is to note that the binomial series expansion of \((u_s - u_r)^{s-r-1}\) in (2.8) is

\[
(u_s - u_r)^{s-r-1} = \sum_{k=0}^{s-r-1} \frac{(s-r-1)}{k!} (-1)^{s-r-1-k} u_r^k u_s^{s-r-1-k}.
\]

Then, \(f_{rs}(u_r, u_s)\) can be written as

\[
f_{rs}(u_r, u_s) = n C_{rs} \sum_{k=0}^{s-r-1} \frac{(s-r-1)}{k!} (-1)^{s-r-1-k} u_r^k u_s^{s-r-1-k} - \frac{s-r-1}{s-k-1} \int_{r}^{n} \frac{u_r}{n} du_r
\]

which is easily integrated. Using this expression in (2.6) results in

\[
w_{i(rs)} = n C_{rs} \sum_{k=0}^{s-r-1} \frac{(s-r-1)}{k!} (-1)^{s-r-1-k} u_r^k u_s^{s-r-1-k} - \frac{s-r-1}{s-k-1} \int_{r}^{n} \frac{u_r}{n} du_r
\]

which can be readily calculated using the incomplete beta function,

\[
B(x; a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt.
\]

Applying the same technique to (2.7) leads to

\[
v_{j(rs)} = n C_{rs} \sum_{k=0}^{s-r-1} \frac{(s-r-1)}{k!} (-1)^{s-r-1-k} u_r^k u_s^{s-r-1-k} - \frac{s-r-1}{s-k-1} \int_{r}^{n} \frac{u_r}{n} du_r
\]

3. Two-stage semi-parametric BLUE of \(\mu\)

First note that if \(F_0\) represents the c.d.f. of a symmetric parent distribution and \(E\left(\frac{X_i + X_n - i + 1}{2}\right) < \infty, i = 1, 2, \ldots, \left[\frac{n+1}{2}\right]\) then we have the following basic result:

\[
E\left(\frac{X_i + X_n - i + 1}{2}\right) = \mu, i = 1, 2, \ldots, \left[\frac{n+1}{2}\right],
\]

where \([·]\) denotes the integer part. For the specific case when \(n\) is odd we have \(E(X_{(n+1)/2}) = \mu\) for the expectation of the sample median.
The first step in our estimation process is to obtain the exact bootstrap variance-covariance matrix for the vector of all possible midpoint estimators. Toward that end for $n$ even denote the $n/2 \times n$ matrix of weights as

$$
H = \begin{pmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & 1/2 & 0 & 0 & 0 & 0 & 1/2 & \vdots \\
\vdots & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & \vdots \\
\vdots & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & \vdots \\
\vdots & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & \vdots \\
\end{pmatrix}
$$

(3.2)

and for $n$ odd denote the $(n + 1)/2 \times n$ matrix of weights as

$$
H = \begin{pmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & 1/2 & 0 & 0 & 0 & 0 & 1/2 & \vdots \\
\vdots & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & \vdots \\
\vdots & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & \vdots \\
\vdots & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \vdots \\
\end{pmatrix}
$$

(3.3)

As above denote the $n \times 1$ vector of ordered observations as $X = (X_1, X_2, ..., X_n)'$. We then have the $n/2 \times 1$ (even) or $(n + 1)/2 \times 1$ (odd) vector of midpoint estimators given as $M = HX$ with variance-covariance matrix $\Gamma = H\Sigma H'$, where $n/2 \times 1$ (even) or $(n + 1)/2 \times 1$ (odd) vector of expectations is given as

$$
E(M) = E(HX) = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix}
$$

(3.4)

and

$$
\Sigma = \begin{pmatrix}
\sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\
\sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \\
\end{pmatrix}
$$

(3.5)

is the variance-covariance matrix of the set of order statistics.

Now denote a $n/2 \times 1$ (even) or $(n + 1)/2 \times 1$ (odd) column vector of 1’s. It follows from standard generalized least-squares theory that the BLUE estimator for $\mu$ based on the midpoints is given by

$$
\hat{\mu} = (1' \Gamma^{-1} 1)^{-1} 1' \Gamma^{-1} M
$$

(3.6)

$$
= LM
$$

(3.7)
with expected value

$$E(\hat{\mu}) = (1' \Gamma^{-1} 1)^{-1} 1' \Gamma^{-1} \mu$$  \hspace{1cm} (3.8)

$$= \mu,$$  \hspace{1cm} (3.9)

given (3.4) and variance

$$\text{Var}(\hat{\mu}) = L\Gamma L'.$$  \hspace{1cm} (3.10)

If we assume $F_0$ represents the c.d.f. of a symmetric parent distribution then the semi-parametric exact bootstrap estimator for $\mu$ is then given by “plugging in” $\hat{\Sigma}$ for $\Sigma$ in (3.7) to arrive at following:

$$\hat{\mu}^* = (1' \hat{\Gamma}^{-1} 1)^{-1} 1' \hat{\Gamma}^{-1} M$$  \hspace{1cm} (3.11)

and

$$\text{Var}(\hat{\mu}^*) = L\hat{\Gamma} L',$$  \hspace{1cm} (3.12)

where $\hat{\Gamma} = H\hat{\Sigma}H'$ and $\hat{\Sigma}$ is given at (2.3).

**Theorem 3.1** If $F_0$ is absolutely continuous and $E(X_{ij}^2 | X_{ij}) < \infty$, $\forall i, j = 1, 2, \ldots, n$ then

$$\hat{\mu}^* \rightarrow \hat{\mu}$$  \hspace{1cm} (3.13)

in distribution as $n \rightarrow \infty$.

**Proof.** A sketch of the proof follows from a direct application of Slutsky’s Theorem by noting that $\hat{\Sigma} \rightarrow \Sigma$ in probability, elementwise. The respective elements of $\hat{\Sigma}$, namely $\hat{\mu}_{r,n}$ at (2.1), $\hat{\sigma}_{r,n}^2$ at (2.4) and $\hat{\sigma}_{rs,n}$ at (2.5) are all well-behaved statistical functionals of the form $T(F_n) = \int h(x) dF_n(x) = \sum w_i h(X_i)$, where $F_n$ denotes the empirical distribution function.

Hence, $\hat{\mu}^* = (1' \hat{\Gamma}^{-1} 1)^{-1} 1' \hat{\Gamma}^{-1} M \rightarrow (1' \Gamma^{-1} 1)^{-1} 1' \Gamma^{-1} M$ in distribution. For a more technical description for when $T(F_n) \rightarrow T(F)$ in probability, see Theorem 6.4.1A in Serfling (1980).

### 3.1. Two-stage BLUE trimmed mean estimator of $\mu$

The $\alpha$ trimmed mean estimator of $T(F) = (1 / (1 - 2\alpha)) \int_{\alpha}^{1-\alpha} x dF$ is given as

$$T(F_n) = \frac{1}{n - 2\lfloor n\alpha \rfloor} \sum_{i = \lfloor n\alpha \rfloor + 1}^{n - \lfloor n\alpha \rfloor} X_{(i)}.$$  \hspace{1cm} (3.14)

where $E(T(F_n)) = \mu$. The thing to note is that there is an equal weight given to each $X_{(i)}$ involved in the summand at (3.14). We can rewrite the trimmed mean statistic at (3.14) in terms of the midpoint estimators for $n$ even as
\[ T(F_n) = \frac{2}{n - 2[n\alpha]} \sum_{i = [n\alpha] + 1}^{n/2} \frac{X_{(n-i+1)} + X_{(i)}}{2} \]  
(3.15)

and for \( n \) odd

\[ T(F_n) = \frac{2}{n - 2[n\alpha]} \sum_{i = [n\alpha] + 1}^{(n+1)/2} \frac{X_{(n-i+1)} + X_{(i)}}{2}. \]  
(3.16)

We can in fact use the framework above to develop a new BLUE trimmed mean estimator that does not weight all of the midpoints in (3.15) and (3.16) equally and given the desired breakdown point as a function of \( \alpha \).

Toward this end we simply need to modify the weight matrix \( H \) at (3.2) and (3.3) according to the desired trimming proportion and whether or not the sample size is even or odd. In other words replace the first \([n\alpha]\) rows of \( H \) at (3.2) and (3.3) with elements all consisting of 0’s, i.e., we eliminate the midpoints in the sequence \( \frac{X_{(n)} + X_{(1)}}{2}, \frac{X_{(n-1)} + X_{(2)}}{2}, \ldots \) depending upon the desired trimming proportion or breakdown point.

For \( n \) even denote the \( \alpha \) trimmed \( n/2 \times n \) matrix of weights as

\[
H_\alpha = \begin{bmatrix}
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\cdots & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \\
\cdots & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\
\cdots & 0 & 0 & 1/2 & 0 & 0 & 0 & \cdots \\
\cdots & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\
\cdots & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots
\end{bmatrix},
\]  
(3.17)

and for \( n \) odd denote the \( \alpha \) trimmed \( (n+1)/2 \times n \) matrix of weights as

\[
H_\alpha = \begin{bmatrix}
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\cdots & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \\
\cdots & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\
\cdots & 0 & 0 & 1/2 & 0 & 0 & 0 & \cdots \\
\cdots & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\
\cdots & 0 & 0 & 0 & 1 & 0 & 0 & \cdots
\end{bmatrix},
\]  
(3.18)

where for both weight matrices the first \([n\alpha]\) rows consist of 0’s.

Now let the vector of trimmed midpoint estimators be given as \( \mathbf{M}_\alpha = H_\alpha \mathbf{X} \) with variance-covariance matrix \( \Gamma_\alpha = H_\alpha \Sigma H_\alpha^T \) where \( n/2 \times 1 \) (even) or \( (n+1)/2 \times 1 \) (odd) vector of expectations under the same symmetry assumptions as above is given as
\[
E(M_\alpha) = E(H_\alpha X) =
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0 \\
\mu \\
\mu \\
\vdots \\
\mu
\end{pmatrix}
\]

(3.19)

and $\Sigma$ is defined at (3.5).

Now denote a $n/2 \times 1$ (even) or $(n + 1)/2$ (odd) vector of 0’s and 1’s depending upon the trimming proportion as $1_\alpha = (0, \ldots, 0, 1, 1, \ldots, 1)'$, where the first $\lfloor n\alpha \rfloor$ elements are 0’s. As above for the untrimmed case it follows from standard generalized least-squares theory that the BLUE estimator for $\mu$ is given by

\[
\hat{\mu}_\alpha = (1_\alpha' \Gamma_\alpha^{-1} 1_\alpha)^{-1} 1_\alpha' \Gamma_\alpha^{-1} M_\alpha
\]

(3.20)

\[
= L_\alpha M_\alpha
\]

(3.21)

with expected value

\[
E(\hat{\mu}_\alpha) = (1_\alpha' \Gamma_\alpha^{-1} 1_\alpha)^{-1} 1_\alpha' \Gamma_\alpha^{-1} 1_\alpha \mu
\]

(3.22)

\[
= \mu,
\]

(3.23)

given (3.19) and variance

\[
\text{Var}(\hat{\mu}_\alpha) = L_\alpha \Gamma_\alpha L_\alpha'.
\]

(3.24)

As above assume $F_0$ represents the c.d.f. of a symmetric parent distribution then the semi-parametric exact bootstrap trimmed estimator for $\mu_\alpha$ is then given by “plugging in” $\hat{\Sigma}$ for $\Sigma$ in (3.7)

\[
\hat{\mu}_\alpha^* = (1_\alpha' \hat{\Gamma}_\alpha^{-1} 1_\alpha)^{-1} 1_\alpha' \hat{\Gamma}_\alpha^{-1} M_\alpha
\]

(3.25)

and

\[
\text{Var}(\hat{\mu}_\alpha^*) = L_\alpha \hat{\Gamma}_\alpha L_\alpha'.
\]

(3.26)

where $\hat{\Sigma}$ is given at (2.3).

\textbf{Corollary 3.1.1} As $n \to \infty$, $E(X_i^2 X_j^2) < \infty$, $\forall i, j = 1, 2, \ldots, n$ and $F_0$ absolutely continuous
\[ \hat{\mu}_a \to \hat{\mu} \quad (3.27) \]

in distribution.

**Proof.** Follows the same as Theorem 3.1.

**Comment.**—Under the symmetry assumption used throughout this note it will be straightforward to create an exact alpha-level permutation test or large sample z-test about the hypothesis \( H_0 : \mu = \mu_0 \) either using all weighted midpoints or the trimmed estimator. This will be a subject for future consideration.

4. Simulation study

We carried out a simulation study comparing the BLUE estimator of \( \mu \) defined at (1.1) versus the semi-parametric exact bootstrap estimator defined at (3.11). It should be noted that both estimators are unbiased of \( \mu \) under the symmetry assumption and given a location-scale family of distributions. Hence, we compared the relative standard deviations of each estimator under various parametric assumptions. The results are provided in Table 1.

The semi-parametric exact bootstrap estimator for \( \mu \) requires only the assumption of symmetry to be valid while the tradional BLUE estimator requires precise parametric assumptions. We simulated data from a standard normal distribution, Student’s t distribution with 3 degrees of freedom, standard logistic distribution, standard Laplace distribution and a \( \beta (1/2, 1/2) \) distribution centered at \( \mu = 0 \) for sample sizes \( n = 10, 20, 30 \). We utilized 1,000 Monte Carlo simulations for each scenario. For the parametric BLUE estimator we simulated values under the correctly assumed distribution and under the incorrectly assumed distribution from our five parametric distributions under consideration. The \( \beta (1/2, 1/2) \) distribution centered at \( \mu = 0 \) provides an example of generating weights given an assumed U-shaped distribution. This distribution represents a more extreme case used in order to illustrate how mispecification of the parent distribution can dramatically impact the properties of the BLUE estimator.

In Table 1 the true underlying distributions are represented by the columns while the assumed distributions relative to the BLUE estimator are given by the rows. We bold-faced the standard deviation estimate when the correct distribution is assumed as the optimal case. Let us start by examining the first column of results under the normality. For \( n = 10, 20, 30 \) the standard deviations for the BLUE estimator under correctly assuming normality are in bold and are 0.33, 0.22 and 0.18, respectively as compared to the standard deviations for the semi-parametric exact bootstrap estimator of 0.36, 0.25, and 0.21, respectively. It is not surprising that under the optimal assumption of normality that the BLUE estimator outperforms the semi-parametric exact bootstrap estimator in terms of efficiency. However, one may be reminded of the provocative assertion by R. C. Geary (1947): “Normality is a myth; there never was, and never will be, a normal distribution.” In other words, model assumptions are rarely met. If the true underlying distribution is normal, but we incorrectly assumed a Laplace distribution the standard deviations are higher for the BLUE estimator as compared to the semi-parametric exact bootstrap estimator. This is a common pattern for...
all of the distributions examined, i.e., the standard deviation of the semi-parametric exact bootstrap estimator falls within the range of standard deviations of the BLUE estimator given correct to incorrect parametric assumptions.

In certain instances assuming the wrong parametric distribution relative to the BLUE estimator can lead to dramatically inflated standard deviations. Even subtle differences between the true and assumed distributions may lead to standard deviations 10% to 20% higher than optimal. For example if the true underlying distribution is a Student’s $t$ distribution with 3 degrees of freedom and we assume the $\beta(1/2, 1/2)$ distribution the standard deviation of the estimator is roughly four to five times larger over the range of sample sizes examined. This simulation study provides numerical evidence for the utility of the semi-parametric exact bootstrap estimator as a robust estimator of $\mu$ under simple symmetry assumptions when compared to the BLUE parametric estimator, which is very sensitive to underlying parametric assumptions.

Acknowledgments

Funding

This work was supported by Roswell Park Cancer Institute and National Cancer Institute (NCI) grant P50CA016056, NRG Oncology Statistical and Data Management Center grant U10CA180822, Ovarian Cancer SPORE grant 50CA159981-06 and IOTN Moonshot grant U24CA232979-01. We wish to thank the two reviewers and associate editor for their comments, which led to a substantially improved presentation of our results.

References

Balakrishnan N, and Rao CR. 1997. A note on the best linear unbiased estimation based on order statistics. The American Statistician 51 (2):181–5. doi:10.2307/2685416.

Balakrishnan N, and Rao CR. 2003. Some efficiency properties of best linear unbiased estimators. Journal of Statistical Planning and Inference 113 (2):551–5. doi:10.1016/S0378-3758(02)00107-6.

David HA 1981. Order statistics, 2nd ed. New York: Wiley.

Geary RC 1947. Testing for normality. Biometrika 34 (Pt 3–4):209–42. doi:10.2307/2332434. [PubMed: 18918691]

Hutson AD, and Ernst MD. 2000. The exact bootstrap mean and variance of an $L$-estimator. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 62 (1):89–94. doi:10.1111/1467-9868.00221.

Serfling RJ 1980. Approximation theorems of mathematical statistics. New York: John Wiley & Sons.
Table 1.

Monte Carlo standard deviation estimates for the parametric BLUE estimator (B) versus the semi-parametric BLUE estimator (SB).

| Assumed Distribution | True Underlying Distribution | N(0, 1) | t3 | Logistic(0,1) | Laplace(0,1) | β(1/2, 1/2) |
|----------------------|-------------------------------|---------|----|---------------|--------------|-------------|
| n = 10               |                               | B       | SB | B             | SB           | B           | SB          |
| N(0, 1)              |                              | 0.33    | 0.36| 0.57          | 0.57         | 0.45        | 0.11        |
| t3                   |                              | 0.32    | 0.41| 0.52          | 0.56         | 0.39        | 0.15        |
| Logistic(0,1)        |                              | 0.32    | 0.49| 0.57          | 0.65         | 0.42        | 0.13        |
| Laplace(0,1)         |                              | 0.38    | 0.45| 0.62          | 0.37         | 0.48        | 0.20        |
| β(1/2, 1/2)          |                              | 0.47    | 2.15| 1.08          | 0.99         | 0.05        | 0.07        |
| n = 20               |                               | B       | SB | B             | SB           | B           | SB          |
| N(0, 1)              |                              | 0.22    | 0.25| 0.43          | 0.42         | 0.30        | 0.08        |
| t3                   |                              | 0.24    | 0.28| 0.34          | 0.39         | 0.25        | 0.12        |
| Logistic(0,1)        |                              | 0.23    | 0.30| 0.38          | 0.43         | 0.28        | 0.10        |
| Laplace(0,1)         |                              | 0.27    | 0.30| 0.45          | 0.28         | 0.32        | 0.15        |
| β(1/2, 1/2)          |                              | 0.42    | 2.41| 1.00          | 1.00         | 0.01        | 0.02        |
| n = 30               |                               | B       | SB | B             | SB           | B           | SB          |
| N(0, 1)              |                              | 0.18    | 0.21| 0.37          | 0.34         | 0.26        | 0.06        |
| t3                   |                              | 0.20    | 0.23| 0.27          | 0.32         | 0.21        | 0.09        |
| Logistic(0,1)        |                              | 0.19    | 0.25| 0.31          | 0.37         | 0.22        | 0.08        |
| Laplace(0,1)         |                              | 0.24    | 0.25| 0.37          | 0.20         | 0.24        | 0.14        |
| β(1/2, 1/2)          |                              | 0.38    | 2.18| 1.01          | 1.01         | 0.01        | 0.01        |