Ann. Inst. Statist. Math.
Vol. 47, No. 3, 465–482 (1995)

PARAMETRIC RANKED SET SAMPLING*

LYNNE STOKES

Department of Management Science and Information Systems, University of Texas at Austin, CBA 5.202, Austin, TX 78712-1175, U.S.A.

(Received March 15, 1994; revised December 6, 1994)

Abstract. Ranked set sampling was introduced by McIntyre (1952, *Australian Journal of Agricultural Research*, **3**, 385–390) as a cost-effective method of selecting data if observations are much more cheaply ranked than measured. He proposed its use for estimating the population mean when the distribution of the data was unknown. In this paper, we examine the advantage, if any, that this method of sampling has if the distribution is known, for a specific family of distributions. Specifically, we consider estimation of μ and σ for the family of random variables with cdf's of the form $F(\frac{x-\mu}{\sigma})$. We find that the ranked set sample does provide more information about both μ and σ than a random sample of the same number of observations. We examine both maximum likelihood and best linear unbiased estimation of μ and σ, as well as methods for modifying the ranked set sampling procedure to provide even better estimation.

Key words and phrases: Order statistic, ranked set sample, maximum likelihood estimator, best linear unbiased estimator.

1. Introduction

Ranked set sampling was introduced and applied to the problem of estimating mean pasture yields by McIntyre (1952). Its function was to improve the efficiency of the sample mean as an estimator of the population mean μ in situations in which the characteristic of interest was difficult or expensive to measure, but could be cheaply ranked. There has been recent interest in ranked set sampling for environmental applications (Johnson *et al.* (1993), Patil and Taillie (1993), Patil *et al.* (1993*a*, 1993*b*, 1994), Gore *et al.* (1994)), for which measurement of the variable of interest for sample units may require expensive testing, but ranking of small sets of samples with respect to the characteristic can be done more cheaply.

The ranked set sampling process consists of drawing m random samples, each of size m, from the population. The m members of each sample are ordered

* This paper has been prepared with partial support from the United States Environmental Protection Agency under Cooperative Agreement Number CR821801-01-0. The contents have not been subjected to Agency review and therefore do not necessarily reflect the views or policies of the Agency and no official endorsement should be inferred.

among themselves by eye, or by some other inexpensive method. Then the smallest observation from the first sample is measured, as is the second smallest observation from the second sample. The process continues in this manner until the largest observation from the m-th sample is measured. This entire cycle is repeated n times until a total of nm^2 observations have been drawn from the population but only nm have been measured. These nm measured observations are referred to as the ranked set sample (RSS). Since accurate eye ordering for large m would be difficult in most practical situations, an increase in sample size is typically implemented by increasing n rather than m.

The attractive feature of a RSS is that it allows improved estimation of a variety of parameters, when judged against a random sample (RS) having the same number (nm) of measured observations. Let $X_{(r)i}$ denote the r-th order statistic in the i-th cycle; then $X_{(r)i}$, $r = 1, \ldots, m$; $i = 1, \ldots, n$ denotes the RSS. Let X_i, $i = 1, \ldots, nm$ denote the RS. Takahasi and Wakimoto (1968) showed that the relative precision of $\hat{\mu}^*$ to $\bar{X}$ always exceeds 1; i.e.,

$$1 \leq RP(\hat{\mu}^*, \bar{X}) = \frac{\mathrm{Var}(\bar{X})}{\mathrm{Var}(\hat{\mu}^*)}, \tag{1.1}$$

where $\bar{X} = \frac{1}{nm}\sum_{i=1}^{nm} X_i$ and

$$\hat{\mu}^* = \frac{1}{nm}\sum_{i=1}^{n}\sum_{r=1}^{m} X_{(r)i}. \tag{1.2}$$

Stokes (1980) showed that the sample variance does not enjoy the same advantage as an estimator of population variance for all values of m and n, but that

$$1 \leq \frac{mse(s^2)}{mse((\hat{\sigma}^*)^2)} \quad \text{for } nm \text{ sufficiently large,}$$

where

$$(nm)s^2 = \sum_{j=1}^{nm}(X_j - \bar{X})^2 \tag{1.3}$$

and

$$(nm)(\hat{\sigma}^*)^2 = \sum_{i=1}^{n}\sum_{r=1}^{m}(X_{(r)i} - \hat{\mu}^*)^2. \tag{1.4}$$

Stokes and Sager (1988) showed that estimation of the cumulative distribution function (cdf) can also be improved by ranked set sampling, since

$$1 \leq RP[\hat{F}^*(t), \hat{F}(t)],$$

for all t, where $\hat{F}$ and $\hat{F}^*$ are the empirical cdf's of the RS and RSS, respectively. In each of these cases, the advantage of the RSS estimator over the RS estimator remains regardless of the distribution of the population. A further remarkable

result is that the advantage for RSS is maintained in each of these three cases even if the ranking process is imperfect (Dell and Clutter (1972), Stokes (1980), Stokes and Sager (1988)).

There has also been some study of parametric versions of RSS estimators for parameters of specific distributions. For example, Sinha *et al.* (1994) examined estimation of parameters of normal and exponential distributions. The disadvantage of the parametric versions of RSS estimators is that they generally do not allow for imperfect ranking; that is, the parametric RSS estimators generally do not retain their unbiasedness when errors in ranking occur. Their potential advantage, however, is that use of knowledge of the distribution along with ranked set sampling might provide improvement in estimation over RS estimators, as well as over the comparable nonparametric RSS estimator, if one exists.

In this paper, the improvement that RSS will allow over RS is characterized for estimators of parameters of a class of random variables. The class of random variables studied is the location-scale family having cdf of the form $F(\frac{x-\mu}{\sigma})$, with F known. This class includes as special cases the distributions studied by Sinha *et al.* (1994). In Section 2, we examine maximum likelihood estimation and in Section 3 best linear unbiased estimation of μ and σ. In Section 4, we discuss the advantages of modified ranked set sampling, in which not all order statistics are measured in turn, but more general selection schemes are allowed. A discussion of robustness to ranking errors and conclusions follows in Section 5. Throughout this paper, we assume that perfect ranking is possible for at least a small number of observations and is very inexpensive when compared with measurement.

2. Maximum likelihood estimation

One reason for examining maximum likelihood estimators (MLE's) from RSS is that they serve as a standard against which other estimators can be measured, since under fairly general regularity conditions an MLE is asymptotically efficient. Thus, by comparing the performance of nonparametric estimators to MLE's, we can assess the value of knowledge of the distribution. Similarly, by comparing the performance of MLE's to that of more easily calculated parametric estimators, we can determine if the extra effort in computation is warranted.

Let $\boldsymbol{X} = (X_1, \ldots, X_{nm})'$ be a RS from a distribution having cdf

$$F\left(\frac{x-\mu}{\sigma}\right) \tag{2.1}$$

and pdf $\frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$. Let $\boldsymbol{X}^* = (X_{(1)1}, \ldots, X_{(m)1}, \ldots, X_{(1)n}, \ldots, X_{(m)n})'$ be a RSS from the same distribution. The loglikelihood function of the RS is

$$L = -nm \ln \sigma + \sum_{j=1}^{nm} \ln f(Z_j),$$

where $Z_j = \frac{X_j - \mu}{\sigma}$, and the loglikelihood function of the RSS is

$$L^* = K - nm \ln \sigma + \sum_{i=1}^{n} \sum_{r=1}^{m} \ln f(Z_{(r)i}) + \sum_{i=1}^{n} \sum_{r=1}^{m} (r-1) \ln F(Z_{(r)i})$$

$$+\sum_{i=1}^{n}\sum_{r=1}^{m}(m-r)\ln[1-F(Z_{(r)i})],$$

where K is a constant and $Z_{(r)i} = \frac{X_{(r)i}-\mu}{\sigma}$. Throughout this paper, we will consider separately the two special cases of the location-scale family mentioned above: (1) single parameter unknown, (2) both μ and σ unknown.

2.1 *One parameter families*

First suppose that σ in (2.1) is known, and we wish to estimate μ by maximum likelihood (ML). The ML estimator from a RS, denoted by $\hat{\mu}_{ML}$, is the solution of the equation

$$\sum_{j=1}^{nm}\frac{f'(Z_j)}{f(Z_j)} = 0,$$

while the ML estimator of μ from a RSS, denoted by $\hat{\mu}^*_{ML}$, is the solution of the equation

$$\sum_{i=1}^{n}\sum_{r=1}^{m}\frac{f'(Z_{(r)i})}{f(Z_{(r)i})} - \sum_{i=1}^{n}\sum_{r=1}^{m}(r-1)\frac{f(Z_{(r)i})}{F(Z_{(r)i})} + \sum_{i=1}^{n}\sum_{r=1}^{m}(m-r)\frac{f(Z_{(r)i})}{1-F(Z_{(r)i})} = 0. \tag{2.2}$$

We examine Fisher Information for μ under each sampling scheme. This will allow us to compare $RP(\hat{\mu}^*_{ML}, \hat{\mu}_{ML})$ and $RP(\hat{\mu}^*, \bar{X})$. First note that for distributions of type (2.1) with σ known, the Fisher Information from a RS is

$$I_{nm}(\mu) = \frac{nm}{\sigma^2}E\left\{\frac{f'(Z_j)}{f(Z_j)}\right\}^2, \tag{2.3}$$

provided $E[\frac{\partial \ln f}{\partial \mu}] = 0$. It is shown in the Appendix that the Fisher Information for μ from a RSS for distributions satisfying the usual regularity conditions is

$$I^*_{nm}(\mu) = \frac{nm}{\sigma^2}E\left\{\frac{f'(Z_j)}{f(Z_j)}\right\}^2 + \frac{nm(m-1)}{\sigma^2}E\left\{\frac{f^2(Z_j)}{F(Z_j)(1-F(Z_j))}\right\}. \tag{2.4}$$

A comparison of (2.3) and (2.4) shows that $I_{nm}(\mu) \leq I^*_{nm}(\mu)$, since the second term of (2.4) is non-negative. In fact, the magnitude of the second term shows that an efficient estimator of μ from a RSS has a smaller variance by an order of m than an efficient estimator from a RS. In particular, for distributions satisfying the regularity conditions specified, we have the asymptotic relative precision

$$\lim_{n\to\infty} RP(\hat{\mu}^*_{ML}, \hat{\mu}_{ML}) = \frac{I^*_{nm}(\mu)}{I_{nm}(\mu)} = 1 + (m-1)E\left\{\frac{f^2(Z_j)}{F(Z_j)(1-F(Z_j))}\right\} \Big/ E\left\{\frac{f'(Z_j)}{f(Z_j)}\right\}^2.$$

Example (a). Let $X_j \sim N(\mu, 1)$, $j = 1, \ldots, nm$, denote a RS. Let $X_{(r)i}$, $r = 1, \ldots, m$; $i = 1, \ldots, n$, denote a RSS from the same distribution. Then, of course, $\hat{\mu}_{ML} = \bar{X}$, but $\hat{\mu}^*_{ML}$ is the solution of the equation

$$\hat{\mu}^*_{ML} = \hat{\mu}^* + \frac{1}{m}\sum_{r=1}^{m}[(m-r)\bar{g}_r - (r-1)\bar{h}_r],$$

where $\hat{\mu}^*$ is defined as in (1.2), $\bar{g}_r = \frac{1}{n}\sum_{i=1}^{n}\frac{\phi(Z_{(r)i})}{1-\Phi(Z_{(r)i})}$, and $\bar{h}_r = \frac{1}{n}\sum_{i=1}^{N}\frac{\phi(Z_{(r)i})}{\Phi(Z_{(r)i})}$. The fact that $\hat{\mu}^*_{ML} \neq \hat{\mu}^*$ in this example shows that if the underlying distribution is known to be normal with σ known, then estimation of the mean can be improved over the nonparametric approach suggested by Dell and Clutter. To examine the amount of improvement, we compare $I_{nm}(\mu) = nm$ and

$$I^*_{nm}(\mu) = nm + nm(m-1)E\left\{\frac{\phi^2(Z_j)}{\Phi(Z_j)(1-\Phi(Z_j))}\right\},$$

where $Z_j \sim N(0,1)$. The expectation above can be evaluated numerically (see, for example, Abramowitz and Stegun (1968), p. 890), and is .4805. Thus

$$\lim_{n\to\infty} RP(\hat{\mu}^*_{ML}, \hat{\mu}_{ML}) = 1 + (m-1)(.4805).$$

Table 1(a) compares $\lim_{n\to\infty} RP(\hat{\mu}^*_{ML}, \hat{\mu}_{ML})$ and $RP(\hat{\mu}^*, \bar{X})$ (from Dell and Clutter (1972), Table 1) for some small m. It shows that large improvement over RS is available from either nonparametric or parametric estimation of μ from a RSS. However, the added improvement from parametric estimation is quite small; e.g., for $m = 5$, the increase in precision for large n is only about $(2.92 - 2.77)/2.77 \approx 5\%$. Thus most of the improvement to be had from RSS in estimation of μ does not require knowledge of the sample's normality.

Now suppose that μ in (2.1) is known, and we wish to estimate σ by maximum likelihood. The ML estimator of σ from a RS, denoted by $\hat{\sigma}_{ML}$, is the solution of the equation

$$nm + \sum_{j=1}^{nm}\frac{Z_j f'(Z_j)}{f(Z_j)} = 0, \tag{2.5}$$

while the ML estimator of σ from a RSS, denoted by $\hat{\sigma}^*_{ML}$, is the solution of the equation

$$\begin{aligned} nm + \sum_{i=1}^{n}\sum_{r=1}^{m}\frac{Z_{(r)i}f'(Z_{(r)i})}{f(Z_{(r)i})} - \sum_{i=1}^{n}\sum_{r=1}^{m}(r-1)\frac{Z_{(r)i}f(Z_{(r)i})}{F(Z_{(r)i})} \\ + \sum_{i=1}^{n}\sum_{r=1}^{m}(m-r)\frac{Z_{(r)i}f(Z_{(r)i})}{1-F(Z_{(r)i})} = 0. \end{aligned} \tag{2.6}$$

Table 1. Relative precision of RSS and RS estimators for three examples[1].

	m			
Relative precision	2	3	4	5
	(a) $N(\mu, 1)$			
$RP(\hat{\mu}^*, \bar{X})$	1.47	1.91	2.35	2.77
$\lim_{n\to\infty} RP(\hat{\mu}^*_{ML}, \hat{\mu}_{ML})$	1.48	1.96	2.44	2.92
	(b) $N(0, \sigma^2)$			
$\lim_{n\to\infty} RP(\hat{\sigma}^*, s)$	1.00	1.08	1.18	1.27
$\lim_{n\to\infty} RP(\hat{\sigma}^*_{ML}, \hat{\sigma}_{ML})$	1.14	1.27	1.41	1.81
	(c) Exp(σ)			
$RP(\hat{\mu}^*, \bar{X})$	1.33	1.64	1.92	2.19
$\lim_{n\to\infty} RP(\hat{\sigma}^*_{ML}, \hat{\sigma}_{ML})$	1.40	1.81	2.21	2.62

[1]Values of RP in cases (a) and (c) are from Table 1 of Dell and Clutter (1972).

To assess the value of knowledge of F, we again compute Fisher Information for σ under each scheme. First, under RS, we have, for distributions satisfying $E[\frac{\partial \ln L}{\partial \sigma}] = 0$,

$$I_{nm}(\sigma) = \frac{nm}{\sigma^2} E\left\{\left[\frac{Z_j f'(Z_j)}{f(Z_j)}\right]^2 - 1\right\}, \tag{2.7}$$

since $E[\frac{Z_j f'(Z_j)}{f(Z_j)}] = -1$. By a method similar to that used in the Appendix, we can show

$$I^*_{nm}(\sigma) = \frac{nm}{\sigma^2}\left\{E\left[\frac{Z_j f'(Z_j)}{f(Z_j)}\right]^2 - 1\right\} + \frac{nm(m-1)}{\sigma^2} E\left\{\frac{[Z_j f(Z_j)]^2}{F(Z_j)(1-F(Z_j))}\right\}, \tag{2.8}$$

if f is such that $\lim_{z\to\infty} zf(z) = \lim_{z\to-\infty} zf(z) = \lim_{z\to\infty} z^2 f(z) = \lim_{z\to-\infty} z^2 f(z)$. A comparison of (2.7) and (2.8) shows that $I^*_{nm}(\sigma) \geq I_{nm}(\sigma)$. As before, for distributions satisfying the regularity conditions specified, we have

$$\begin{aligned}\lim_{n\to\infty} RP(\hat{\sigma}^*_{ML}, \hat{\sigma}_{ML}) &= \frac{I^*_{nm}(\sigma)}{I_{nm}(\sigma)} \\ &= 1 + (m-1)E\left\{\frac{[Z_j f(Z_j)]^2}{F(Z_j)(1-F(Z_j))}\right\} \Big/ \left\{E\left[\frac{Z_j f'(Z_j)}{f(Z_j)}\right]^2 - 1\right\}.\end{aligned}$$

Example (b). Let $X_j \sim N(0, \sigma^2)$, $j = 1, \ldots, nm$, denote a RS and $X_{(r)i}$, $r = 1, \ldots, m$; $i = 1, \ldots, n$, denote a RSS from the same distribution. The MLE's $[\frac{1}{n}\sum_{i=1}^{nm} X_i^2]^{1/2}$ and $\hat{\sigma}^*_{ML}$, obtained iteratively from (2.6), can be compared through their Fisher information, $I_{nm}(\sigma) = \frac{2nm}{\sigma^2}$ and

$$I^*_{nm}(\sigma) = \frac{2nm}{\sigma^2} + \frac{nm(m-1)}{\sigma^2} E\left\{\frac{[Z_j\phi(Z_j)]^2}{\Phi(Z_j)(1-\Phi(Z_j))}\right\}.$$

The expectation above can be evaluated numerically and is .2705. Thus

$$\lim_{n\to\infty} RP(\hat{\sigma}^*_{ML}, \hat{\sigma}_{ML}) = 1 + \frac{(m-1)}{2}(.2705).$$

If the distribution of X_j were unknown, σ could be estimated from a RSS by $\hat{\sigma}^*$, defined in (1.4) and from a RS by s, defined in (1.3). Using the delta method, one can show that $RP(\hat{\sigma}^*, s) \approx RP((\hat{\sigma}^*)^2, s^2)$ for large n. The limiting value of the latter expression can be obtained from Stokes ((1980), p. 39). Table 1(b) displays these relative precisions. From the table, we can observe that the gain from RSS over RS in estimation of σ is modest in either the parametric or nonparametric version, compared to its gain for estimating μ. However, unlike the findings from Example (a), we see that the relative improvement from knowledge of the distribution is substantial. For example, for $m = 5$, the increase in precision from using the knowledge of normality is $(1.81 - 1.27)/1.27 \approx 43\%$.

Example (c). Consider a RS and RSS from the exponential distribution (denoted Exp(σ)) having cdf $F(x) = 1 - e^{-(x/\sigma)}$. As before, we compare the ML estimator from the RS (which is $\hat{\sigma}_{ML} = \bar{X}$) with the ML estimator from the RSS (which must be obtained iteratively from (2.6)) by comparing $I_{nm}(\sigma) = \frac{nm}{\sigma^2}$ and $I^*_{nm}(\sigma) = \frac{nm}{\sigma^2}[1 + (m-1)\int_0^\infty \frac{z^2 e^{-2z}}{1-e^{-z}} dz]$, so that $\lim_{n\to\infty} RP(\hat{\sigma}^*_{ML}, \hat{\sigma}_{ML}) = 1 + (m-1)(.4041)$. If the distribution were unknown, the parameter σ could be estimated from the RSS nonparametrically by $\hat{\mu}^*$, since σ denotes the mean of this distribution. Table 1(c) compares the RSS and RS estimators by their relative precision. There is less benefit from RSS in estimation of the exponential mean as compared to the normal mean, but there is more benefit from knowledge of the distribution coupled with the use of RSS. For $m = 5$, the increase in precision from knowledge of the distribution is $(2.62 - 2.19)/2.62 \approx 20\%$.

2.2 *Two-parameter families*

To compare ML estimators of μ and σ estimated simultaneously, the Fisher Information matrices $I_{nm}(\mu, \sigma)$ and $I^*_{nm}(\mu, \sigma)$ must be compared. The diagonal elements of these matrices are given by (2.3), (2.7) and (2.4), (2.8). Using a method similar to that of the Appendix, it can be shown that the off-diagonal elements are

$$\frac{nm}{\sigma^2} E\left\{Z_j\left[\frac{f'(Z_j)}{f(Z_j)}\right]^2\right\} \tag{2.9}$$

and

$$\frac{nm}{\sigma^2}E\left\{Z_j\left[\frac{f'(Z_j)}{f(z_j)}\right]^2\right\}+\frac{nm(m-1)}{\sigma^2}E\left\{\frac{Z_j f^2(Z_j)}{F(Z_j)[1-F(Z_j)]}\right\}, \tag{2.10}$$

for the RS and RSS estimators, respectively. If the random variable Z_j is symmetric, then the off-diagonal elements (2.9) and (2.10) are 0, so that $|I_{nm}(\mu,\sigma)| < |I^*_{nm}(\mu,\sigma)|$ and $(\hat{\mu}^*_{ML},\hat{\sigma}^*_{ML})$ has an ellipse of concentration which lies completely within that of $(\hat{\mu}_{ML},\hat{\sigma}_{ML})$ for any sample size. However, it does not appear to be possible to determine that $|I_{nm}(\mu,\sigma)| < |I^*_{nm}(\mu,\sigma)|$ generally for non-symmetric distributions, unless m is sufficiently large.

Example (d). Let $X_j \sim N(\mu,\sigma^2)$, $j=1,\ldots,nm$, denote a RS and $X_{(r)i}$, $r=1,\ldots,m$; $i=1,\ldots,n$, denote a RSS from the same distribution. Because of the symmetry of the normal distribution, the efficiencies for $\hat{\mu}^*_{ML}$ and $\hat{\sigma}^*_{ML}$ reported in Table 1 hold for this case as well, although the ML estimates themselves will differ from the ML estimates in the cases in which one parameter is known. This is true in general for symmetric distributions.

3. Best Linear Unbiased Estimators (BLUE's)

Sinha *et al.* (1994) have suggested BLUE's as parametric estimators from ranked set samples for some specific distributions. In this section, we describe a general method for obtaining RSS BLUE's of μ and σ from location-scale distributions of the form (2.1). The resulting estimators are simple to compute and might therefore be preferable to the ML estimators derived in Section 2 if it could be shown that they are nearly efficient. We examine several example distributions and find that for some, the BLUE's are nearly efficient, while for others, they perform badly.

Recall that $Z_{(r)i}=\frac{X_{(r)i}-\mu}{\sigma}$. Following the method of Lloyd (1952), we define $\alpha_{(r:m)}=E[Z_{(r)i}]$ and $\nu_{(r:m)}=\mathrm{Var}[Z_{(r)i}]$. Then $E[X_{(r)i}]=\mu+\alpha_{(r:m)}\sigma$ and $\mathrm{Var}[X_{(r)i}]=\sigma^2\nu_{(r:m)}$. Thus when F is known, the vector of unknown parameters $\boldsymbol{\theta}'=(\mu,\sigma)$ can be estimated by

$$\hat{\boldsymbol{\theta}}^*_{BLU}=(\boldsymbol{A}'\boldsymbol{V}^{-1}\boldsymbol{A})^{-1}(\boldsymbol{A}'\boldsymbol{V}^{-1}\boldsymbol{X}^*) \tag{3.1}$$

where $\boldsymbol{A}'=\begin{bmatrix}\mathbf{1}' & \cdots & \mathbf{1}'\\ \boldsymbol{\alpha}' & \cdots & \boldsymbol{\alpha}'\end{bmatrix}$, $\mathbf{1}$ is an $m\times 1$ vector of 1's and $\boldsymbol{\alpha}'=(\alpha_{(1:m)},\ldots,\alpha_{(m:m)})$, $\boldsymbol{V}=\mathrm{Diag}(\nu,\ldots,\nu)$, where $\boldsymbol{\nu}'=(\nu_{(1:m)},\ldots,\nu_{(m:m)})$. Further

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}^*_{BLU})=\sigma^2(\boldsymbol{A}'\boldsymbol{V}^{-1}\boldsymbol{A})^{-1}. \tag{3.2}$$

We again consider separately the one- and two-parameter families of form (2.1). First, when σ is known, (3.1) yields

$$\hat{\mu}^*_{BLU}=\frac{\sum_{r=1}^m[(\bar{X}_{(r)}-\sigma\alpha_{(r:m)})/\nu_{(r:m)}]}{\sum_{r=1}^m[1/\nu_{(r:m)}]} \tag{3.3}$$

where $\bar{X}_{(r)} = \frac{1}{n}\sum_{i=1}^{n} X_{(r)i}$ and (3.2) yields

$$\text{Var}(\hat{\mu}^*_{BLU}) = \frac{\sigma^2}{n}\left[\sum_{r=1}^{m}(1/\nu_{(r:m)})\right]^{-1}. \tag{3.4}$$

Second, when μ is known,

$$\hat{\sigma}^*_{BLU} = \frac{\sum_{r=1}^{m}[\alpha_{(r:m)}(\bar{X}_{(r)} - \mu)/\nu_{(r:m)}]}{\sum_{r=1}^{m}[\alpha^2_{(r:m)}/\nu_{(r:m)}]} \tag{3.5}$$

and

$$\text{Var}(\hat{\sigma}^*_{BLU}) = \frac{\sigma^2}{n}\left[\sum_{r=1}^{m}(\alpha^2_{(r:m)}/\nu_{(r:m)})\right]^{-1}. \tag{3.6}$$

Thus, $\hat{\mu}^*_{BLU}$ and $\hat{\sigma}^*_{BLU}$ are weighted averages of sample means of order statistics, each adjusted with a bias correction term which is a function of the known parameter. Finally, if both μ and σ are unknown and the distribution is symmetric around μ, then both estimators can be written more simply since their bias correction terms are 0. Thus for symmetric distributions, knowledge of the unknown nuisance parameter is not required to compute $\hat{\mu}^*_{BLU}$ or $\hat{\sigma}^*_{BLU}$, and $\text{Var}(\hat{\mu}^*_{BLU})$ and $\text{Var}(\hat{\sigma}^*_{BLU})$ are still given by (3.4) and (3.6), since the off-diagonal elements of $\boldsymbol{A}'\boldsymbol{V}^{-1}\boldsymbol{A}$ are 0.

Now we reconsider the examples of Section 2. For each example, the efficiency of the BLU estimators for $m = 2, \ldots, 5$ is shown in Table 2, where efficiency is defined as the ratio of the minimum achievable variance to that of the BLUE from the RSS.

Table 2. Efficiency of BLU estimators from ranked set samples.

		m			
Distribution		2	3	4	5
(a) $N(\mu, 1)$	$eff(\hat{\mu}^*_{BLU})$	.99	.99	.99	.98
(b) $N(0, \sigma)$	$eff(\hat{\sigma}^*_{BLU})$	.21	.34	.44	.49
(c) Exp(σ)	$eff(\hat{\sigma}^*_{BLU})$	1.00	.99	.99	.99

Example (a)′. Let $F = \Phi$, the standard normal cdf, and suppose that μ is the parameter of interest. Note that knowledge of σ is unnecessary in calculating $\hat{\mu}^*_{BLU}$, because of the symmetry of the normal distribution. This estimator, also suggested by Sinha *et al.* (1994), has efficiency

$$eff(\hat{\mu}^*_{BLU}) = [\text{Var}(\hat{\mu}^*_{BLU})I^*_{nm}(\mu)]^{-1} = \frac{\sum_{r=1}^{m}(1/\nu_{(r:m)})}{m[1 + (m-1)(.4805)]},$$

where $\nu_{(r:m)}$ is the variance of the standard normal order statistic. Table 2 shows that the potential loss from using the BLU rather than the ML estimator in this case is negligible. As noted in Table 1, however, neither is a great improvement over $\hat{\mu}^*$.

Example (b)′. Let $F = \Phi$, and suppose that σ is the parameter of interest. Note that knowledge of μ is unnecessary in calculating $\hat{\sigma}^*_{BLU}$ because of the symmetry of the normal distribution. Then

$$eff(\hat{\sigma}^*_{BLU}) = [\mathrm{Var}(\hat{\sigma}^*_{BLU}) I^*_{nm}(\sigma)]^{-1} = \frac{\sum_{r=1}^{m} (\alpha^2_{(r:m)}/\nu_{(r:m)})}{m[2 + (m-1)(.2705)]},$$

where $\alpha_{(r:m)}$ is the mean of the standard normal order statistic. Table 2 shows that the potential loss from using the BLU rather than the ML estimator in this case is extremely large. Because of the poor performance of $\hat{\sigma}^*_{BLU}$, an alternative estimator of σ is desirable. Sinha *et al.* (1994) suggested an alternative estimator of σ^2 for this estimation problem. Its counterpart for estimation of σ is

$$\tilde{\sigma}^*_m = \left[c_m \sum_{i=1}^{n} \sum_{r=1}^{m} (X_{(r)i} - \hat{\mu}^*)^2 \right]^{1/2}$$

where c_m is a constant chosen to make $(\tilde{\sigma}^*_m)^2$ unbiased for σ^2. They show that $\mathrm{Var}((\tilde{\sigma}^*_m)^2) = \frac{k_m \sigma^4}{n}$, where the constants k_m are provided in their Table 7 for $m = 2, \ldots, 5$. This estimator is not linear like the others we have considered in this section, but its unbiasedness and variance properties do arise from moments of the normal order statistics. From the delta method, we know that $\mathrm{Var}(\tilde{\sigma}^*_m) \approx \mathrm{Var}((\tilde{\sigma}^*_m)^2)/4\sigma^2$ for large n, so that $\mathrm{Var}(\tilde{\sigma}^*_m) \approx \frac{k_m \sigma^2}{4n}$. Table 3 compares $\tilde{\sigma}^*_m$ to both the BLU and ML estimators of σ by displaying

$$\begin{aligned} \lim_{n\to\infty} RP(\tilde{\sigma}^*_m, \hat{\sigma}^*_{BLU}) &= \lim_{n\to\infty} \frac{\mathrm{Var}(\hat{\sigma}^*_{BLU})}{\mathrm{Var}(\tilde{\sigma}^*_m)} \\ &= 4 \Big/ \left\{ k_m \sum_{r=1}^{m} [\alpha^2_{(r:m)}/\nu_{(r:m)}] \right\} \end{aligned}$$

and

$$\begin{aligned} \lim_{n\to\infty} RP(\tilde{\sigma}^*_m, \hat{\sigma}^*_{ML}) &= \lim_{n\to\infty} \frac{\mathrm{Var}(\hat{\sigma}^*_{ML})}{\mathrm{Var}(\tilde{\sigma}^*_m)} \\ &= 4/\{m k_m [2 + (m-1)(.2705)]\}. \end{aligned}$$

The table shows that $\tilde{\sigma}^*_m$ is a dramatic improvement over the BLU estimator, but is still far from efficient.

Example (c)′. Let $F(x) = 1 - e^{-(x/\sigma)}$. Then $\hat{\sigma}^*_{BLU}$, suggested in Sinha *et al.* (1994), has

$$eff(\hat{\sigma}^*_{BLU}) = \frac{\sum_{r=1}^{m} (\alpha^2_{(r:m)}/\nu_{(r:m)})}{m[1 + (m-1)(.4041)]},$$

Table 3. Comparison of $\tilde{\sigma}_m^*$ with BLU and ML estimators from ranked set samples.

	m			
Relative precision	2	3	4	5
$\lim_{n\to\infty} RP(\tilde{\sigma}_m^*, \hat{\sigma}_{BLU}^*)$	2.61	2.02	1.51	1.54
$\lim_{n\to\infty} RP(\tilde{\sigma}_m^*, \hat{\sigma}_{ML}^*)$	.54	.68	.66	.76

where $\alpha_{(r:m)}$ and $\nu_{(r:m)}$ are the mean and variance of the standard exponential order statistic. Table 2 shows that $\hat{\sigma}_{BLU}^*$ is nearly efficient, so that $\hat{\sigma}_{BLU}^*$ would probably be preferred to $\hat{\sigma}_{ML}^*$, since it is easier to compute and its small sample properties are known.

4. BLUE's from Modified Ranked Set Samples (MRSS)

Several authors, including Yanagawa and Shirahata (1976), Ridout and Cobby (1987), and Muttlak and McDonald (1990), have suggested modifications to the ranked set sampling procedure. In these procedures, not every order statistic is selected for measurement in each cycle. Instead, the sampler chooses which measurements would be most beneficial. In this section, we use that idea to find improved BLU estimators of μ and σ for distributions of type (2.1), as Ogawa (1951) did for RS. A general method for determining how to modify the RSS is developed for large m. We consider this situation, even though it is impractical in most applications. The reason is that the results provide theoretical insight not afforded by the numerical results for small m, about what characteristics of distributions allow them to benefit most from MRSS. Further, the large sample results are seen to work well even in small samples for the examples we have considered.

We use the following theorem (Mosteller (1946)), which describes the asymptotic distribution of selected quantiles.

THEOREM 4.1. *For k given real numbers $0 < \lambda_1 < \cdots < \lambda_k < 1$, let the λ_j-quantile of the population be x_j; i.e., $\int_{-\infty}^{x_j} g(t)dt = \lambda_j$, $j = 1, \ldots, k$, where $g(x)$ is the pdf of the population. Assume that $g(x)$ is differentiable in the neighborhoods of $x = x_j$ and that $g_j \equiv g(x_j) \neq 0$ for all j. Then the joint distribution of the k order statistics $X_{(n_1:N)}, \ldots, X_{(n_k:N)}$ where $n_j = [N\lambda_j] + 1$, $j = 1, \ldots, k$ tends to a k-dimensional normal distribution with means $x_1, \ldots, x_k$ and variance-covariance matrix with elements $[\frac{1}{N}\frac{\lambda_j(1-\lambda_l)}{g_j g_l}]$ as $N \to \infty$.*

Suppose we modify the RSS procedure so that in each of the n cycles of m samples, a set of independent order statistics $X_{(r_1:m)}, \ldots, X_{(r_m:m)}$ is measured. The theorem above allows us to observe that for the location-scale family of (2.1),

when $r_j = [m\lambda_j] + 1$, $j = 1, \ldots, m$,

$$X_{(r_j:m)} \to N\left(\mu + \sigma z_{\lambda_j}, \frac{\sigma^2}{m}\frac{\lambda_j(1-\lambda_j)}{f_j^2(z_{\lambda_j})}\right) \tag{4.1}$$

as $m \to \infty$, where $z_{\lambda_j} = F^{-1}(\lambda_j)$. Further, the random variables making up the MRSS are independent, so that their variance-covariance matrix is diagonal.

A BLU estimator of μ and/or σ from the MRSS can be constructed following the method used in Section 3. The questions we address in this section are (1) how should $r_1, \ldots, r_m$ be selected, and (2) how does this procedure compare with others? We examine the answers to these questions both for small m and as $m \to \infty$.

4.1 *One-parameter families*

If σ were known, the BLU estimator of μ from the MRSS is

$$\hat{\mu}_\Delta^* = \frac{\sum_{i=1}^n \sum_{j=1}^m [(X_{(r_j)i} - \sigma\alpha_{(r_j)})/\nu_{(r_j)i}]}{n\sum_{j=1}^m (1/\nu_{(r_j)})}, \tag{4.2}$$

which has variance $\mathrm{Var}(\hat{\mu}_\Delta^*) = \frac{\sigma^2}{n}[\sum_{j=1}^m (1/\nu_{(r_j:m)})]^{-1}$. This variance is minimized by choosing the same order statistic from each sample, i.e., the one which minimizes $\nu_{(r_j:m)}$, say $r_j = r_\mu$. In the special case in which X is symmetric, one could, with equivalent precision, select for measurement the $(m - r_\mu + 1)$-th order statistics in place of any number of r_μ order statistics, since they have equal precision. For such a MRSS, we have

$$\mathrm{Var}(\hat{\mu}_\Delta^*) = \frac{\sigma^2}{nm}\nu_{(r_\mu:m)}. \tag{4.3}$$

If m were large, (4.1) shows that the optimal order statistic is $r_\mu = [m\lambda_\mu + 1]$ where λ_μ minimizes

$$h_1(\lambda) = \frac{\lambda(1-\lambda)}{f^2(z_\lambda)}. \tag{4.4}$$

For such a MRSS,

$$m^2 \,\mathrm{Var}(\hat{\mu}_\Delta^*) \to \frac{\sigma^2}{n}\frac{\lambda_\mu(1-\lambda_\mu)}{f^2(z_{\lambda_\mu})} \tag{4.5}$$

as $m \to \infty$.

Similarly, if μ were known, the BLU estimator from a MRSS is

$$\hat{\sigma}_\Delta^* = \frac{\sum_{i=1}^n \sum_{j=1}^m [\alpha_{(r_j:m)}(X_{(r_j:m)i} - \mu)/\nu_{(r_j:m)}]}{n\sum_{j=1}^m (\alpha_{(r_j:m)}^2/\nu_{(r_j:m)})}, \tag{4.6}$$

which has variance $\mathrm{Var}(\hat{\sigma}^*_\Delta) = \frac{\sigma^2}{n}[\sum_{j=1}^m (\alpha^2_{(r_j:m)}/\nu_{(r_j:m)})]^{-1}$. This variance is minimized by selecting from each sample the order statistic which maximizes $\alpha^2_{(r_j:m)}/\nu_{(r_j:m)}$, say $r_i = r_\sigma$. As in the previous case, when X is symmetric one could select any combination of r_σ and $(m - r_\sigma + 1)$ order statistics. Then

$$\mathrm{Var}(\hat{\sigma}^*_\Delta) = \frac{\sigma^2}{nm}\nu_{(r_\sigma:m)}/\alpha^2_{(r_\sigma:m)}. \tag{4.7}$$

As before we see that if m were large, $r_\sigma = [m\lambda_\sigma + 1]$ where λ_σ minimizes

$$h_2(\lambda) = \frac{\lambda(1-\lambda)}{f^2(z_\lambda)z^2_\lambda}. \tag{4.8}$$

For such a MRSS,

$$m^2\,\mathrm{Var}(\hat{\sigma}^*_\Delta) \to \frac{\sigma^2}{n}\frac{\lambda_\sigma(1-\lambda_\sigma)}{z^2_{\lambda_\mu} f^2(z_{\lambda_\sigma})} \tag{4.9}$$

as $m \to \infty$.

To examine whether the BLU estimator from a MRSS competes favorably with the best possible estimator from a RSS, we can compute

$$imp(\hat{\mu}^*_\Delta) = [\mathrm{Var}(\hat{\mu}^*_\Delta)I^*_{nm}(\mu)]^{-1} \tag{4.10}$$

and

$$imp(\hat{\sigma}^*_\Delta) = [\mathrm{Var}(\hat{\sigma}^*_\Delta)I^*_{nm}(\sigma)]^{-1}, \tag{4.11}$$

where $\mathrm{Var}(\hat{\mu}^*_\Delta)$ and $\mathrm{Var}(\hat{\sigma}^*_\Delta)$ are defined in (4.3) and (4.7). It is of theoretical interest to examine the improvement available from MRSS for large m, even though in practice, one is unlikely to have samples available with extremely large m. From (2.4), (2.8), (4.5), and (4.9), one can show that

$$\lim_{m\to\infty} imp(\hat{\mu}^*_\Delta) = \frac{f^2(z_{\lambda_\mu})}{\lambda_\mu(1-\lambda_\mu)} \Big/ E\left\{\frac{f^2(Z_j)}{F(Z_j)(1-F(Z_j))}\right\} \tag{4.12}$$

and

$$\lim_{m\to\infty} imp(\hat{\sigma}_\Delta) = \frac{z^2_{\lambda_\sigma} f^2(z_{\lambda_\sigma})}{\lambda_\sigma(1-\lambda_\sigma)} \Big/ E\left\{\frac{Z_j^2 f^2(Z_j)}{F(Z_j)(1-F(Z_j))}\right\}. \tag{4.13}$$

As m increases, the improvement available from $\hat{\mu}^*_\Delta$ and $\hat{\sigma}^*_\Delta$ converges to the ratio between the maximum and expected value of the random variables in brackets in (4.12) (for estimating μ) and (4.13) (for estimating σ). So modifying the RSS procedure is most beneficial for those distributions for which these random variables have sharp modes. We now return to the three examples of the previous sections.

Example (a)″. Let $F = \Phi$, and assume that μ is the parameter of interest. Then $h_1(\lambda)$ in (4.4) is minimized by $\lambda_\mu = .5$, so that for large m, the sample median is the optimal order statistic to select for the MRSS. (Obviously this

same choice is optimal for any symmetric unimodal random variable.) Sinha *et al.* (1994) proved in their Lemma 1 that the sample median is optimal for *all* m for this normal estimation problem. Thus for this example, the large m results lead us to the correct choice for small m as well. It is unnecessary that σ be known in this example (or for any other symmetric distribution). If m is odd, (4.2) shows that $\hat{\mu}^*_\Delta$'s bias correction term is 0. If m is even, we can, with equal precision, alternate selection of the $m/2$ and $(m/2+1)$-st order statistics in each of the mn samples, also yielding a bias correction term of 0. Table 4(a) displays values of $imp(\hat{\mu}^*_\Delta)$ (from (4.10)) for $m = 2, \ldots, 5$. We see that for $m > 2$, $\hat{\mu}^*_\Delta$ offers at least a 14% improvement over the best possible from RSS. From (4.12) it can be shown that $imp(\hat{\mu}^*_\Delta) \to 1.32$ as $m \to \infty$.

Table 4. Improvement available from BLUE's in modified RSS for three examples.

		m			
Distribution		2	3	4	5
(a) $N(\mu, 1)$	r_μ	1 or 2	2	3 or 4	3
	$imp(\hat{\mu}^*_\Delta)$	.99	1.14	1.14	1.19
(b) $N(0, \sigma)$	r_σ	1 or 2	1 or 3	1 or 4	1 or 5
	$imp(\hat{\sigma}^*_\Delta)$	.21	.50	.77	.98
(c) Exp(σ)	r_σ	2	3	4	5
	$imp(\hat{\sigma}^*_\Delta)$	1.28	1.37	1.38	1.36

Example (b)$''$. Let $F = \Phi$, and assume that σ is the parameter of interest. Then $h_2(\lambda)$ in (4.8) is minimized by $\lambda_\sigma = .942$, so that σ can best be estimated when m is large by choosing from each sample the order statistic $r_\sigma = ([.942m]+1)$. For small m, r_σ differs infrequently and by at most 1 from this value. (If μ were unknown, one could achieve equal precision and make the bias correction term 0 by alternate selection of the r_σ and $m - r_\sigma + 1$ order statistics.) Table 4(b) displays r_σ and $imp(\hat{\sigma}^*_\Delta)$ (from (4.11)) for $m = 2, \ldots, 5$. For these small m, even MRSS does not produce a BLU estimator preferable to $\hat{\sigma}^*_{ML}$. From (4.13), however, we see that $\lim_{m\to\infty} imp(\hat{\sigma}^*_\Delta) = 2.18$. So if one were able to accurately eyeorder a large enough number of observations, or at least accurately identify the $(1 - .942)m = .058m$ most extreme observations from each sample, then a great advantage in estimation of σ could be realized from MRSS.

Example (c)$''$. Let $F(x) = 1 - e^{-x/\sigma}$. The BLU estimator (4.6) is suggested by Sinha *et al.* (1994) for this example, and they identify r_σ in their Table 9 for $m = 2, \ldots, 5$. $h_2(\lambda)$ in (4.8) is minimized by $\lambda_\sigma = .797$, so that σ can best be estimated when m is large by choosing from each sample the order statistic $r_\sigma = ([.797m] + 1)$. For small m, r_σ differs infrequently and by at most 1 from this value. Table 4(c) displays r_σ and $imp(\hat{\sigma}^*_\Delta)$ (from (4.11)) for $m = 2, \ldots, 5$, and

shows that BLU estimation of μ from a MRSS is beneficial here, even for small m. Equation (4.13) shows that $\lim_{m\to\infty} imp(\hat{\sigma}^*_\Delta) = 1.61$, so that an even greater advantage could be realized if a sufficiently large number of observations could be accurately ranked.

4.2 *Two-parameter families*

Finally, suppose both μ and σ are unknown. Unlike the one-parameter cases just considered, it is impossible to estimate both μ and σ if the same order statistic is chosen from each sample, for then $\boldsymbol{A}$ has rank 1. Therefore, at least two different order statistics must be included in the MRSS. Suppose we include exactly two different order statistics in our sample; that is, suppose we measure $mn/2$ r_i-th and $mn/2$ r_j-th order statistics, each from samples of size m, with $r_i > r_j$. Depending upon the objective of estimation, one might be most interested in minimizing

$$\operatorname{Var}(\hat{\mu}^*_\Delta) \propto \frac{\alpha^2_{(r_i:m)}\sigma^2_{(r_i:m)} + \alpha^2_{(r_j:m)}\sigma^2_{(r_j:m)}}{\delta_{ij}},$$

$$\operatorname{Var}(\hat{\sigma}^*_\Delta) \propto \frac{\sigma^2_{(r_i:m)} + \sigma^2_{(r_j:m)}}{\delta_{ij}},$$

or the generalized variance

$$|\operatorname{Var}(\hat{\theta}^*_\Delta)| \propto \frac{\sigma^2_{(r_i:m)}\sigma^2_{(r_j:m)}}{\delta_{ij}}, \tag{4.14}$$

where $\delta_{ij} = (\alpha_{(r_i:m)} - \alpha_{(r_j:m)})^2$. As before, (4.1) can be used to determine the optimal pair of order statistics when m is large.

Now consider the special case in which X is symmetric. It has already been noted that for any symmetric distribution, a MRSS choice of any combination of r_μ (or r_σ) and $(m - r_\mu + 1)$ (or $(m - r_\sigma + 1)$) order statistics yield equally efficient BLU estimators. But only when $nm/2$ of each are chosen from among the nm samples of size m are the bias-correction terms $\hat{\mu}^*_\Delta$ and $\hat{\mu}^*_\Delta$ equal to zero. So this is the optimal choice for the MRSS if only μ (or σ) is to be estimated when X is symmetric, but both parameters are unknown. However, if both parameters are of interest, one might choose to minimize $|\operatorname{Var}(\hat{\theta}^*_\Delta)|$ in (4.14). To do this for large m, one should select the pair of order statistics $r_{\theta 1} = ([m\lambda_\theta] + 1)$ and $r_{\theta 2} = (m - r_{\theta 1} + 1)$, where λ_θ minimizes

$$h_3(\lambda) = \left|\frac{\lambda(1-\lambda)}{z_\lambda f^2(z_\lambda)}\right|.$$

Then

$$m^2|\operatorname{Var}(\hat{\theta}^*_\Delta)| \to \frac{\sigma^2}{n}\frac{\lambda_\theta(1-\lambda_\theta)}{z_{\lambda_\theta} f^2(z_{\lambda_\theta})}$$

as $m \to \infty$.

Example (d)$''$. Let $F = \Phi$ and assume both μ and σ are of interest. Then $h_3(\lambda)$ is minimized by $\lambda_\theta = .872$, so that θ can best be estimated when m is large by choosing $nm/2$ order statistics of rank $r_\sigma = ([.872m] + 1)$ and the remaining $nm/2$ of rank $(m - r_\sigma + 1)$.

5. Conclusions

RSS has previously proven advantageous for estimating mean, variance, and even the entire cdf of populations whose distributions are unknown. In this paper, we have investigated whether and how much additional improvement in estimation can be achieved for the location-scale family of distributions if one assumes that knowledge of the distribution is available. We have shown that improvement over RSS always yields improved estimation in large samples for both μ and σ via maximum likelihood estimation. Further, we saw that for some distributions, BLU estimation is a close competitor to maximum likelihood, but for others is much poorer. Finally, we saw that modifying the RSS procedure does in some, but not all cases, produce considerably improved precision.

The greatest practical usefulness of this investigation may be that it gives the researcher a method for determining when the quest for improved estimators can be halted. For example, when a parametric method is justified for normal (μ to be estimated) or exponential data, BLUE's make nearly full use of the information in the data. Thus the additional effort of computing MLE's is unnecessary. The same cannot be said if interest lies in estimating σ for normal data. In that case, the additional computing effort may be worthwhile. In those cases in which MLE's are deemed worthwhile, their standard errors can be computed by making use of the estimated Fisher information (from (2.4), (2.8), and/or (2.10)).

The obstacle to use of any of these parametric RSS procedures is their lack of robustness to errors in the ranking process. If such errors occur, they can bias the estimators. To assess the seriousness of the bias, consider the ranking error model used by Stokes (1977). Suppose that the moments of the judgment order statistic $X_{[r:m]}$ are

$$E(X_{[r:m]}) = \mu + \rho\sigma\alpha_{(r:m)} \tag{5.1}$$

and $\mathrm{Var}(X_{[r:m]}) = \sigma^2(1-\rho^2)+\rho^2\sigma^2\nu_{(r:m)}$, where $-1 \leq \rho \leq 1$. This model contains the two extremes of perfect ranking ($\rho = 1$) or perfectly worthless ranking ($\rho = 0$) as special cases.

Without specifying a more complete model for ranking errors, one cannot assess their effect on $\hat{\mu}^*_{ML}$ and $\hat{\sigma}^*_{ML}$. But it is easy to see their effect on the BLU estimators. Letting $\boldsymbol{X}^*$in (3.1) denote the vector of judgment order statistics $(X_{[1:m]1}, \ldots, X_{[m:m]n})'$ with means shown in (5.1), one can see that $E(\hat{\theta}^*_{BLU}) = (\mu, \rho\sigma)'$. Thus this estimator is still unbiased for μ, but can be badly biased for estimation of σ. The BLU estimators from modified ranked set samples fare worse. Even $\hat{\mu}^*_{\Delta}$ is biased for μ under this model for ranking errors, except in the special case of symmetric distributions. So the sampler must be especially careful when using these parametric estimators (except possibly in BLU estimation of μ) that m is chosen small enough that the ranking can be done without error.

Appendix

Show that

$$I^*_{nm}(\mu) = \frac{nm}{\sigma^2} E\left[\frac{f'(Z_j)}{f(Z_j)}\right]^2 + \frac{nm(m-1)}{\sigma^2} E\left[\frac{f^2(Z_j)}{F(Z_j)(1-F(Z_j))}\right]. \tag{A.1}$$

PROOF. For distributions satisfying the condition $\frac{\partial}{\partial\mu}\int_{-\infty}^{\infty} f(\frac{x-\mu}{\sigma})dx = \int_{-\infty}^{\infty}\frac{\partial}{\partial\mu}f(\frac{x-\mu}{\sigma})dx$, we know that $I^*_{nm}(\mu) = -E[\frac{\partial^2 \ln L^*}{\partial\mu^*}]$. Thus we can differentiate the right hand side of (2.2) to obtain

$$
\begin{aligned}
(A.2)\qquad I^*_{nm}(\mu) = \frac{n}{\sigma^2}\sum_{r=1}^{m} E\left\{-\frac{f''(Z_{(r)i})}{f(Z_{(r)i})} + \left[\frac{f'(z_{(r)i})}{f(Z_{(r)i})}\right]^2\right\} \\
+\sum_{r=1}^{m}(r-1)E\left\{-\frac{f'(Z_{(r)i})}{F(Z_{(r)i})} + \left[\frac{f(Z_{(r)i})}{F(Z_{(r)i})}\right]^2\right\} \\
+\sum_{r-1}^{m}(m-r)E\left\{\frac{f'(Z_{(r)i})}{1-F(Z_{(r)i})} + \left[\frac{f(Z_{(r)i})}{1-F(Z_{(r)i})}\right]^2\right\}.
\end{aligned}
$$

Letting $h(z) = \frac{f'(z)}{f(z)} - \frac{f''(z)}{f(z)}$, we see that the first term of the right hand side of (A.1) is

$$
\begin{aligned}
(A.3)\qquad &\frac{n}{\sigma^2}\sum_{r=1}^{m}\int_{-\infty}^{\infty} h(z)m\binom{m-1}{r-1}f(z)F^{r-1}(z)[1-F(z)]^{m-r}dz \\
&= \frac{n}{\sigma^2}\int_{-\infty}^{\infty} mh(z)f(z)\sum_{r=1}^{m}\binom{m-1}{r-1}F^{r-1}(z)[1-F(z)]^{m-r}dz \\
&= \frac{nm}{\sigma^2}\int_{-\infty}^{\infty} h(z)f(z)dz = \frac{nm}{\sigma^2}E\left\{\left[\frac{f'(Z_j)}{f(Z_j)}\right]^2 - \frac{f''(Z_j)}{f(Z_j)}\right\}.
\end{aligned}
$$

Now redefine $h(z) = [\frac{f(z)}{F(z)}]^2 - \frac{f'(z)}{F(z)}$. Then the second term of the right hand side of (A.1) is

$$
\begin{aligned}
(A.4)\qquad &\frac{n}{\sigma^2}\sum_{r=1}^{m}\int_{-\infty}^{\infty}(r-1)h(z)m\binom{m-1}{r-1}f(z)F^{r-1}(z)[1-F(z)]^{m-r}dz \\
&= \frac{nm}{\sigma^2}\int_{-\infty}^{\infty} h(z)f(z)\sum_{r=1}^{m}(r-1)\binom{m-1}{r-1}F^{r-1}(z)[1-F(z)]^{m-r}dz \\
&= \frac{nm(m-1)}{\sigma^2}\int_{-\infty}^{\infty} h(z)f(z)F(z)dz \\
&= \frac{nm}{\sigma^2}E\left\{\left[\frac{f^2(Z_j)}{F(Z_j)}\right]^2 - f'(Z_j)\right\}.
\end{aligned}
$$

Similarly, the third term of the right hand side of (A.1) can be shown to be

$$
(A.5)\qquad \frac{nm(m-1)}{\sigma^2}E\left\{\frac{f^2(Z_j)}{[1-F(Z_j)]} + f'(Z_j)\right\}.
$$

Combining terms (A.3) through (A.5) yields (A.1).

REFERENCES

Abramowitz, M. A. and Stegun, I. A. (1968). *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D.C.

Dell, T. T. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background, *Biometrics*, **28**, 545–555.

Gore, S. D., Patil, G. P. and Sinha, A. K. (1994). Environmental chemistry, statistical modeling, and observational economy, *Environmental Statistics, Assessment, and Forecasting* (eds. C. R. Cothern and N. P. Ross), 57–97, Lewis Publishing/CRC Press, Boca Raton.

Johnson, G. D., Patil, G. P. and Sinha, A. K. (1993). Ranked set sampling for vegetation research, *Abstracta Botanica*, **17**, 87–102.

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

McIntyre, G. A. (1952). A method for unbiased selective sampling using ranked sets, *Australian Journal of Agricultural Research*, **3**, 385–390.

Mosteller, F. (1946). On some useful inefficient statistics, *Ann. Math. Statist.*, **17**, 377–407.

Muttlak, H. A. and McDonald, L. L. (1990). Ranked set sampling with respect to concomitant variables and with size biased probability of selection, *Comm. Statist. Theory Methods*, **19**, 205–219.

Ogawa, J. (1951). Contributions to the theory of systematic statistics, *Osaka Mathematical Journal*, **3**, 175–213.

Patil, G. P. and Taillie, C. (1993). Environmental sampling, observational economy, and statistical inference with emphasis on ranked set sampling, encounter sampling, and composite sampling, *Bulletin ISI Proceedings of 49th Session*, 295–312, Firenze.

Patil, G. P., Sinha, A. K. and Taillie, C. (1993*a*). Ranked set sampling from a finite population in the presence of a trend on a site, *Journal of Applied Statistical Science*, **1**, 51–65.

Patil, G. P., Sinha, A. K. and Taillie, C. (1993*b*). Observational economy of ranked set sampling: comparison with the regression estimator, *Environmetrics*, **4**, 399–412.

Patil, G. P., Sinha, A. K. and Taillie, C. (1994). Ranked set sampling, *Handbook of Statistics Volume 12: Environmental Statistics* (eds. G. P. Patil and C. R. Rao), 167–200, North Holland/Elsevier Science Publishers, New York and Amsterdam.

Ridout, M. S. and Cobby, J. M. (1987). Ranked set sampling with non-random selection of sets and errors in ranking, *Applied Statistics*, **36**, 145–152.

Sinha, Bimal K., Sinha, Bikas K. and Purkayastha, S. (1994). On some aspects of ranked set sampling for estimation of normal and exponential parameters, *Statist. Decisions* (in press).

Stokes, S. L. (1977). Ranked set sampling with concomitant variables, *Comm. Statist. Theory Methods*, **6**, 1207–1211.

Stokes, S. L. (1980). Estimation of variance using judgment ordered ranked set samples, *Biometrics*, **36**, 35–42.

Stokes, S. L. and Sager, T. W. (1988). Characterization of a ranked set sample with application to estimating distribution functions, *J. Amer. Statist. Assoc.*, **83**, 374–381.

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering, *Ann. Inst. Statist. Math.*, **20**, 1–31.

Yanagawa, T. and Shirahata, S. (1976). Ranked set sampling theory with selective probability matrix, *Austral. J. Statist.*, **18**, 45–52.