11-1-2002

Double Median Ranked Set Sample: Comparing To Other Double Ranked Samples For Mean And Ratio Estimators

Hani M. Samawi
Sultan Qaboos University, Sultanate of Oman

Eman M. Tawalbeh
Yarmouk University, Jordan

Double Median Ranked Set Sample: Comparing To Other Double Ranked Samples For Mean And Ratio Estimators

Hani M. Samawi
Department of Mathematics & Statistics
Sultan Qaboos University
Sultanate of Oman

Eman M. Tawalbeh
Department of Statistics
Yarmouk University
Jordan

Double median ranked set sample (DMRSS) and its properties for estimating the population mean, when the underlying distribution is assumed to be symmetric about its mean, are introduced. Also, the performance of DMRSS with respect to other ranked set samples and double ranked set samples, for estimating the population mean and ratio, is considered. Real data that consist of heights and diameters of 399 trees are used to illustrate the procedure. The analysis and simulation indicate that using DMRSS for estimating the population mean is more efficient than using the other ranked samples and double ranked samples schemes except in case of uniform distribution. Also, using double sampling schemes substantially increase the relative efficiency of ratio estimators relative to their counterpart schemes of one stage samples. Moreover, DMRSS is superior to other double sampling schemes for ratio estimation.

Key words: Double extreme ranked set sample; double median ranked set sample, ratio estimation.

Introduction

In many agricultural and environmental studies and recently in human populations, it is common for quantification of a sampling unit to be costly as compared with the physical acquisition of the unit. For example, level of bilirubin in the blood of infants can be ranked visually by observing: a) color of the face, b) color of the chest, c) color of lower part of the body, & d) color of terminal parts of the whole body. Then, as the yellowish goes from i to iv, the level of bilirubin in the blood goes higher (Samawi & Al-Sakeer, 2001). In such circumstances, considerable cost savings can be achieved if the number of quantification is only a small fraction of the number of available units but all units contribute to the information content of the quantification.

Hani Michel Samawi is an Associate Professor of Biostatistics. His areas of research are in bootstrap and resampling methods, ranked set sampling estimators, sampling method, testing hypothesis, estimation, and analysis of biostatistics data. E-mail: hsamawi@squ.edu.om.

Ranked set sampling (RSS) is considered to be a new method of sampling compared with other sampling methods that can achieve this goal. RSS was first introduced by McIntyre (1952). The use of RSS is highly powerful and much superior to the standard simple random sampling (SRS) for estimating some of the population parameters.

As a variation of RSS Samawi et al. (1996) and Muttlak (1997) investigated extreme ranked set sample (ERSS) and median ranked set sample (MRSS) respectively. Samawi and Muttlak (1996 & 2001) used RSS and MRSS to improve the performance of the ratio estimator. Also, Samawi (2001) suggested the double extreme ranked set sampling (DERSS). They showed that ERSS, MRSS and DERSS are more practical than RSS and more efficient at least than SRS for estimating the population mean. Moreover, Al-Saleh and Al-Kadiri (2000) showed that the efficiency of estimating the population mean could be improved even more by double ranked set sampling technique (DRSS). Also, they proved that ranking in the second stage is easier than in the first stage.

In this article, DMRSS is introduced. The properties of DMRSS for estimating the population mean, when the underlying distribution

is assumed to be symmetric about its mean, are discussed. Also, the performance of DMRSS with respect to the other ranked set samples and double ranked set samples, for estimating the population mean and ratio, is considered.

In Section 2 samples notations and definition and some basic results are introduced . DMRSS scheme and properties are introduced in Section 3. Also, its performance with other sampling schemes will be compared for estimating the population mean. In Section 4, the performance of different double ranked samples schemes will be compared with their counterpart one stage ranked samples for ratio estimation based on the relative efficiency. Illustration of the procedure using real data set with final comments and conclusions is discussed in Section 5.

Sample Notations And Definitions With Some Useful Results
One Stage Sampling

Univariate population
For any of RSS, ERSS and MRSS schemes, the procedure can be described by selecting r random sets each of size r from the target population. In the most practical situations, the size r will be 2, 3 or 4. Rank each set by a suitable method of ranking like prior information, visual inspection or by the experimenter. In sampling notation this implies:

$$\begin{bmatrix} X_{11}, & X_{12}, & \cdots, & X_{1r} \\ X_{21}, & X_{22}, & \cdots, & X_{2r} \\ \vdots & & \vdots & \\ X_{r1}, & X_{r2}, & \cdots, & X_{rr} \end{bmatrix} \xrightarrow{\text{after ranking}}$$

$$\begin{bmatrix} X_{1(1)}, & X_{1(2)}, & \cdots, & X_{1(r)} \\ X_{2(1)}, & X_{2(2)}, & \cdots, & X_{2(r)} \\ \vdots & & \vdots & \\ X_{r(1)}, & X_{r(2)}, & \cdots, & X_{r(r)} \end{bmatrix}$$

(2.1)

where X_{ji} denotes the i-th observation in the j-th set and $X_{j(i)}$ the i-th ordered statistic in the j-th set.

1) If only $X_{1(1)}, X_{2(2)},..., X_{r(r)}$, quantified by obtaining the element with smallest rank from the

first set, the second smallest from the second set, and so on until the largest unit from the r-th set is measured. Then, this represents one cycle of RSS. We can repeat the whole procedure m times to get a RSS of size n = mr. (See Takahasi and Wakimoto, 1968.)

2) Similarly, as in Samawi et al. (1996), we have two cases: In case of r is even, and if only RSS, $X_{1(1)k}, X_{2(r)k}, \cdots, X_{r-1(1)k}, X_{r(r)k}$, k=1,2,…,m, quantified, then this will denote the $ERSS_E$. In case of r is odd, and if only $X_{1(1)k}, X_{2(r)k}, \cdots, X_{r-1(r)k}, X_{r\left(\frac{r+1}{2}\right)k}$, k=1,2,…,m, quantified, then this will denote the $ERSS_O$.

3) Again, similar to Muttlak (1997), we have two cases: In case of r is odd, and if only $X_{1\left(\frac{r+1}{2}\right)k} ,…, X_{r\left(\frac{r+1}{2}\right)k}, k = 1, 2, ..., m$, quantified, then this will denote the $MRSS_O$. In case of r is even, select for measurement from the first $\frac{r}{2}$ samples the $\left(\frac{r}{2}\right)$-th smallest unit and from the last $\frac{r}{2}$ samples select the $\left(\frac{r}{2}+1\right)$-th smallest unit. This will be denoted by $MRSS_E$ (i.e.

$$X_{1\left(\frac{r}{2}\right)k} ,…, X_{\frac{r}{2}\left(\frac{r}{2}\right)k}, X_{\frac{r}{2}+1\left(\frac{r}{2}+1\right)k},$$
$$…, X_{r\left(\frac{r}{2}+1\right)k}, k = 1, 2,…, m \quad).$$

For bivariate population
Samawi and Muttluk (1996) modified the above procedure in case of bivariate distributions to estimate the population ratio. The procedure is described as follows:
First choose r^2 independent bivariate elements from a population, with bivariate distribution function F(x, y). Rank each set with respect to one of the variables Y or X. Suppose ranking is on variable X. Apply the same procedures as in case of univariate population but for each measured unit from the X's, the associated unit from the Y's is measured too. This may be repeated m times to get a bivariate sample of size n = rm.

In sample notation:

1) The sample $\{(X_{i(i)k}, Y_{i[i]k}),$ i=1,2,...,r, k=1,2,...,m} will denote the bivariate RSS.

2) The sample,

$\{(X_{1(1)k}, Y_{1[1]k}), (X_{2(r)k}, Y_{2[r]k}),$

$\dots, (X_{r-1(1)k}, Y_{r-1[1]k}), (X_{r(r)k}, Y_{r[r]k})\}'$

k=1,2,...,m, will denote the bivariate ERSS$_E$ and

$\{(X_{1(1)k}, Y_{1[1]k}), (X_{2(r)k}, Y_{2[r]k}),$

$\dots, (X_{r-1(r)k}, Y_{r-1[r]k}), (X_{r(\frac{r+1}{2})k}, Y_{r[\frac{r+1}{2}]k})\}'$

k=1,2,...,m, will denote the bivariate ERSS$_O$.

1) Similarly,

$$\left\{ \left(X_{i(\frac{r+1}{2})k}, Y_{i[\frac{r+1}{2}]k} \right) : i = 1,2,\dots, \atop r \text{ and } k = 1,2,\dots, m \right\}$$

2) will denote the bivariate MRSS$_O$ and

$$\left(X_{1(\frac{r}{2})k}, Y_{1[\frac{r}{2}]k} \right), \dots, \left(X_{\frac{r}{2}(\frac{r}{2})k}, Y_{\frac{r}{2}[\frac{r}{2}]k} \right)$$

$$, \left(X_{\frac{r}{2}+1(\frac{r}{2}+1)k}, Y_{\frac{r}{2}+1[\frac{r}{2}+1]} \right)$$

$$, \dots, \left(X_{r(\frac{r}{2}+1)}, Y_{r[\frac{r}{2}+1]} \right),$$ k=1,2, ..., m will denote the bivariate MRSS$_E$.

Double Ranked Samples (Two Stage Sampling)

1) Al-Saleh and Al-Kadiri (2000) introduced DRSS procedure as follows:

> 1. Identify r^3 elements from the target population and divide these elements randomly into r sets each of size r^2 elements.
> 2. Use the usual RSS procedure on each set to obtain r RSS each of size r.
> 3. Employ again the RSS procedure in Step 2, to obtain the DRSS of size r.
> 4. We may repeat steps 1-3 m times to obtain a sample of size n = rm.

In sampling notations, after ranking each sample separately in each subset, we get:

$$\begin{bmatrix} X_{1(1)k}^{(1)} X_{1(2)k}^{(1)} \dots X_{1(r)k}^{(1)} \\ X_{2(1)k}^{(1)} X_{2(2)k}^{(1)} \dots X_{2(r)k}^{(1)} \\ \vdots \\ X_{r(1)k}^{(1)} X_{r(2)k}^{(1)} \dots X_{r(r)k}^{(1)} \end{bmatrix}, \dots,$$

$$\begin{bmatrix} X_{1(1)k}^{(r)} X_{1(2)k}^{(r)} \dots X_{1(r)k}^{(r)} \\ X_{2(1)k}^{(r)} X_{2(2)k}^{(r)} \dots X_{2(r)k}^{(r)} \\ \vdots \\ X_{r(1)k}^{(r)} X_{r(2)k}^{(r)} \dots X_{r(r)k}^{(r)} \end{bmatrix}, \qquad (2.2)$$

k=1,2,...,m , where $X_{i(i)k}^{(l)}$ is the i-th ordered observation in the l-th set in the i-th sample in the k-th cycle. Use RSS scheme on each subset separately, we get

$$A_{1k} = \left\{ X_{1(1)k}^{(1)}, X_{2(2)k}^{(1)}, \dots X_{r(r)k}^{(1)} \right\}, \dots,$$

$$A_{rk} = \left\{ X_{1(1)k}^{(r)}, X_{2(2)k}^{(r)}, \dots X_{r(r)k}^{(r)} \right\}$$

Then in the second stage, let $W_{i(i)k}$=i-th smallest observation in A_{ik}, then $\{W_{i(i)k},$ i=1,2,...,r, k=1,2,...,m} will denote the DRSS. Now let $W_{(1)k},\dots,W_{(r)k}$, k=1, 2, ..., m, be a DRSS, with mean and variance of $W_{(i)k}$ are $\mu_{(i)}^{**}$ and $\sigma_{(i)}^{**2}$, respectively. Al-Saleh and Al-Kadiri (2000) also showed that:

$$\mu = \frac{1}{r} \sum_{i=1}^{r} \mu_{(i)}^{**} \qquad \text{and}$$

$$\sigma^2 = \frac{1}{r} \left[\sum_{i=1}^{r} \sigma_{(i)}^{**2} + \sum_{i=1}^{r} (\mu_{(i)}^{**} - \mu)^2 \right]$$

where μ and σ^2 are the mean and the variance of the population, respectively. Also, it was shown that ranking in the second stage is easier than in the first stage.

2) DERSS is an extension to ERSS procedure by Samawi (2001). The procedure is just similar to that for DRSS, but taking ERSS instead of RSS in the first and in the second stage. Implies that

$\{W_{1(1)k}, W_{2(r)k}, W_{3(1)k}, ..., W_{r(r)k}, k = 1, 2, ..., m\}$ d

enotes DERSS$_E$. The case when r is odd is similar. For more about RSS see for example Kaur *et al.,* (1995) and Patil *et al.* (1999).

Double Median Ranked Set Sample

     In this Section a modification to MRSS, namely double median ranked set sample (DMRSS) is introduced. The properties of this scheme for estimating the population mean, which is considered to be finite, is discussed when the underlying distribution function is assumed to be symmetric. Also, some numerical and theoretical comparisons with SRS, RSS, MRSS, ERSS, DERSS and DRSS are included.

Sample Notation and Definitions

     For each cycle k=1,2,…,m (m= number of cycles), assume a simple random sample, of size r^3, is selected from a target population with c.d.f. F(x) and p.d.f. f(x), where F(x) is assumed to be symmetric and absolutely continuous, with mean μ and variance σ². Suppose we divided the sample independently into r sets of data where each set contains r samples, each of size r. Two cases are considered:

Case 1: From (2.2) and when r is odd, for the k-th cycle, get r^2 ranked samples as in (2.2):

Take the median $X^{(j)}_{i\left(\frac{r+1}{2}\right)k}$ from each sample in

each set, then the following sets are resulted: A$_{1k}$=

$\{X^{(1)}_{1\left(\frac{r+1}{2}\right)k}$, $X^{(1)}_{2\left(\frac{r+1}{2}\right)k}$, ..., $X^{(1)}_{r\left(\frac{r+1}{2}\right)k}$ $\}$, A$_{2k}$=

$\{X^{(2)}_{1\left(\frac{r+1}{2}\right)k}$, $X^{(2)}_{2\left(\frac{r+1}{2}\right)k}$, ..., $X^{(2)}_{r\left(\frac{r+1}{2}\right)k}$ $\}$, ..., A$_{rk}$=

$\{X^{(r)}_{1\left(\frac{r+1}{2}\right)k}$, $X^{(r)}_{2\left(\frac{r+1}{2}\right)k}$, ..., $X^{(r)}_{r\left(\frac{r+1}{2}\right)k}$ $\}$.

These sets are the first stage MRSS samples. The second stage MRSS or double MRSS is the set of medians of A$_{1k}$, A$_{2k}$, …, A$_{rk}$. Define $W_{1\left(\frac{r+1}{2}\right)k}$ =

med(A$_{1k}$), $W_{2\left(\frac{r+1}{2}\right)k}$ = med(A$_{2k}$), …, $W_{r\left(\frac{r+1}{2}\right)k}$ =

med(A$_{rk}$), then the sample $\{W_{1\left(\frac{r+1}{2}\right)k}$, $W_{2\left(\frac{r+1}{2}\right)k}$, …,

$W_{r\left(\frac{r+1}{2}\right)k}$ $\}$, k=1,2,…,m is denoted by DMRSS$_O$. The sample mean using DMRSS$_O$ is given by

$$\overline{W}_{DMRSS_O} = \frac{1}{rm}\sum_{k=1}^{m}\sum_{i=1}^{r}W_{i\left(\frac{r+1}{2}\right)k}.$$
(3.1)

Case 2: When r is even, for the k-th cycle, after ranking each sample in each set, as in Case 1, divide the r sets in (2.2) in half to two independent sets. From the first $\frac{r}{2}$ sets take the ($\frac{r}{2}$)-th smallest unit from each sample and from the last $\frac{r}{2}$ sets take the ($\frac{r}{2}$+1)-th smallest unit from each sample, that is, we will get the following sets:

$A_{1k} = \left\{X^{(1)}_{1\left(\frac{r}{2}\right)k}, X^{(1)}_{2\left(\frac{r}{2}\right)k}, ..., X^{(1)}_{r\left(\frac{r}{2}\right)k}\right\}$,

$A_{2k} = \left\{X^{(2)}_{1\left(\frac{r}{2}\right)k}, X^{(2)}_{2\left(\frac{r}{2}\right)k}, ..., X^{(2)}_{r\left(\frac{r}{2}\right)k}\right\}$,...,

$A_{\frac{r}{2}k} = \left\{X^{\left(\frac{r}{2}\right)}_{1\left(\frac{r}{2}\right)k}, X^{\left(\frac{r}{2}\right)}_{2\left(\frac{r}{2}\right)k}, ..., X^{\left(\frac{r}{2}\right)}_{r\left(\frac{r}{2}\right)k}\right\}$,

$B_{1k} = \left\{X^{\left(\frac{r}{2}+1\right)}_{1\left(\frac{r}{2}+1\right)k}, X^{\left(\frac{r}{2}+1\right)}_{2\left(\frac{r}{2}+1\right)k}, ..., X^{\left(\frac{r}{2}+1\right)}_{r\left(\frac{r}{2}+1\right)k}\right\}$,

$B_{2k} = \left\{X^{\left(\frac{r}{2}+2\right)}_{1\left(\frac{r}{2}+1\right)k}, X^{\left(\frac{r}{2}+2\right)}_{2\left(\frac{r}{2}+1\right)k}, ..., X^{\left(\frac{r}{2}+2\right)}_{r\left(\frac{r}{2}+1\right)k}\right\}$,...

$B_{\frac{r}{2}k} = \left\{X^{(r)}_{1\left(\frac{r}{2}+1\right)k}, X^{(r)}_{2\left(\frac{r}{2}+1\right)k}, ..., X^{(r)}_{r\left(\frac{r}{2}+1\right)k}\right\}$,

k=1,2,…, m. This is the first stage. Again from each A$_{ik}$ take the ($\frac{r}{2}$)-th smallest units, while from each B$_{ik}$ take the ($\frac{r}{2}$+1)-th smallest unit as follows: $W_{i\left(\frac{r}{2}\right)k}$ = the$\left(\frac{r}{2}\right)$- th ordered statistic from A$_{ik}$, i=1, 2,… , $\frac{r}{2}$, k= 1, 2, …, m and $W_{i\left(\frac{r}{2}+1\right)k}$ = the$\left(\frac{r}{2}+1\right)$- th ordered statistic from B$_{ik}$, i=1, 2, … , $\frac{r}{2}$, k= 1,2,…,m. Then the resulted

sample
$$W_{1\left(\frac{r}{2}\right)k}, W_{2\left(\frac{r}{2}\right)k}, \ldots, W_{\frac{r}{2}\left(\frac{r}{2}\right)k},$$
$$W_{1\left(\frac{r}{2}+1\right)k}, W_{2\left(\frac{r}{2}+1\right)k}, \ldots, W_{\frac{r}{2}\left(\frac{r}{2}+1\right)k},$$

k=1,2,…,m denotes DMRSS$_E$. The sample mean using DMRSS$_E$ is given by

$$\overline{W}_{DMRSS\;E} = \frac{1}{rm}\sum_{k=1}^{m}\left(\sum_{i=1}^{r/2}W_{i\left(\frac{r}{2}\right)k} + \sum_{i=1}^{r/2}W_{i\left(\frac{r}{2}+1\right)k}\right).$$
(3.2)

To study the properties of DMRSS$_O$ and DMRSS$_E$, next we derive the distribution functions of $W_{\left(\frac{r+1}{2}\right)}$, $W_{\left(\frac{r}{2}\right)}$ and $W_{\left(\frac{r}{2}+1\right)}$ respectively and some of their properties.

Distribution Function and Properties of DMRSS
Case 1: When r is odd. To find the distribution of $W_{i\left(\frac{r+1}{2}\right)k}$, i=1,2, …, r say $G_{\left(\frac{r+1}{2}\right)}(w)$, first the distribution of $X^{(j)}_{i\left(\frac{r+1}{2}\right)k}$ say $F_{\left(\frac{r+1}{2}\right)}(x)$ is given by

$$F_{\left(\frac{r+1}{2}\right)}(x) = \int_{-\infty}^{x}\frac{r!}{\left(\frac{r-1}{2}\right)!\left(\frac{r-1}{2}\right)!}f(t)$$
$$\left(F(t)\right)^{\left(\frac{r-1}{2}\right)}\left(1-F(t)\right)^{\left(\frac{r-1}{2}\right)}dt$$
(3.3)

(see Arnold, et al. 1992). Let u = F(t), then

$$F_{\left(\frac{r+1}{2}\right)}(x) = \int_{0}^{F(x)}\frac{r!}{\left(\frac{r-1}{2}\right)!\left(\frac{r-1}{2}\right)!}(u)^{\left(\frac{r-1}{2}\right)}(1-u)^{\left(\frac{r-1}{2}\right)}du$$
$$= I_{F(x)}\left(\frac{r+1}{2},\frac{r+1}{2}\right),$$
(3.4)

which is the usual incomplete beta function.
Hence, $X^{(j)}_{1\left(\frac{r+1}{2}\right)k}$, $X^{(j)}_{2\left(\frac{r+1}{2}\right)k}$, …, $X^{(j)}_{r\left(\frac{r+1}{2}\right)k}$, are independent and identically distributed (i.i.d.) with incomplete beta $I_{F(x)}\left(\frac{r+1}{2},\frac{r+1}{2}\right)$ distribution. Now from the definition of DMRSS$_O$, the p.d.f. of $W_{i\left(\frac{r+1}{2}\right)k}$ will be

$$g_{\left(\frac{r+1}{2}\right)}(w) = \frac{r!}{\left(\frac{r-1}{2}\right)!\left(\frac{r-1}{2}\right)!}\left(F_{\left(\frac{r+1}{2}\right)}(w)\right)^{\left(\frac{r-1}{2}\right)}$$
$$\left(1-F_{\left(\frac{r+1}{2}\right)}(w)\right)^{\left(\frac{r-1}{2}\right)}f_{\left(\frac{r+1}{2}\right)}(w) =$$
$$\left(\frac{r!}{\left(\frac{r-1}{2}\right)!\left(\frac{r-1}{2}\right)!}\right)^{2}\left(\left(I_{F(w)}\left(\frac{r+1}{2},\frac{r+1}{2}\right)\right)\left(1-I_{F(w)}\left(\frac{r+1}{2},\frac{r+1}{2}\right)\right)\right)^{\left(\frac{r-1}{2}\right)}$$
$$\times\left(F(w)(1-F(w))\right)^{\left(\frac{r-1}{2}\right)}\times f(w).$$
(3.5)

Let $s = I_{F(t)}\left(\frac{r+1}{2},\frac{r+1}{2}\right)$, then the c.d.f. of $W_{i\left(\frac{r+1}{2}\right)k}$ is

$$G_{\left(\frac{r+1}{2}\right)}(w) = \int_{0}^{I_{F(w)}}\frac{r!}{\left(\frac{r-1}{2}\right)!\left(\frac{r-1}{2}\right)!}(s)^{\left(\frac{r-1}{2}\right)}(1-s)^{\left(\frac{r-1}{2}\right)}ds$$
$$= I_{I_{F(w)}\left(\frac{r+1}{2},\frac{r+1}{2}\right)}\left(\frac{r+1}{2},\frac{r+1}{2}\right).$$
(3.6)

Note that, $W_{1\left(\frac{r+1}{2}\right)k}, W_{2\left(\frac{r+1}{2}\right)k}, \ldots, W_{r\left(\frac{r+1}{2}\right)k}$, k=1 ,2 ,…, m are i.i.d. with the (3.6) distribution function.

Case 2: Distribution function of $W_{i\left(\frac{r}{2}\right)}$ and $W_{i\left(\frac{r}{2}+1\right)}$, i=1, 2, …, $\frac{r}{2}$ when r is even.

Recall the assumption that the MRSS is based on a simple random sample of size r^3 with the symmetric and i.i.d. distribution function F(x).

Distribution function of $W_{i\left(\frac{r}{2}\right)}$

Using the same steps as in case 1, the p.d.f. and c.d.f. of $W_{i\left(\frac{r}{2}\right)k}$, i =1, 2, …, $\frac{r}{2}$, will be respectively

$$g_{\left(\frac{r}{2}\right)}(w) = \frac{r!}{\left(\frac{r}{2}\right)!\left(\frac{r}{2}-1\right)!}\left(F_{\left(\frac{r}{2}\right)}(w)\right)^{\left(\frac{r}{2}-1\right)}$$

$$\left(1-F_{\left(\frac{r}{2}\right)}(w)\right)^{\left(\frac{r}{2}\right)} f_{\left(\frac{r}{2}\right)}(w)$$

$$= \left(\frac{r!}{\left(\frac{r}{2}\right)!\left(\frac{r}{2}-1\right)!}\right)^2 \left(I_{F(w)}\left(\frac{r}{2},\frac{r}{2}+1\right)\right)^{\left(\frac{r}{2}-1\right)}$$

$$\left(1-I_{F(w)}\left(\frac{r}{2},\frac{r}{2}+1\right)\right)^{\left(\frac{r}{2}\right)}$$

$$\times (F(w))^{\left(\frac{r}{2}-1\right)}(1-F(w))^{\left(\frac{r}{2}\right)} \times f(w)$$

(3.7)

and

$$G_{\left(\frac{r}{2}\right)}(w) =$$

$$\int_0^{I_{F(w)}\left(\frac{r}{2},\frac{r}{2}+1\right)} \frac{r!}{\left(\frac{r}{2}\right)!\left(\frac{r}{2}-1\right)!}(s)^{\left(\frac{r}{2}-1\right)}(1-s)^{\left(\frac{r}{2}\right)}ds$$

$$= I_{F(w)\left(\frac{r}{2},\frac{r}{2}+1\right)}\left(\frac{r}{2},\frac{r}{2}+1\right) .$$

(3.8)

Note that, the $W_{i\left(\frac{r}{2}\right)k}$, i =1, 2, …, $\frac{r}{2}$, k=1, 2, …, m are i.i.d. with (3.7) distribution function. Similarly, the p.d.f. and c.d.f. of $W_{i\left(\frac{r}{2}\right)k}$, i=1, 2, …, $\frac{r}{2}$, k=1, 2, …, m, say, $g_{\left(\frac{r}{2}+1\right)}(w)$ and $G_{\left(\frac{r}{2}+1\right)}(w)$ respectively are,

$$g_{\left(\frac{r}{2}+1\right)}(w) = \frac{r!}{\left(\frac{r}{2}\right)!\left(\frac{r}{2}-1\right)!}\left(F_{\left(\frac{r}{2}+1\right)}(w)\right)^{\left(\frac{r}{2}\right)}$$

$$\left(1-F_{\left(\frac{r}{2}+1\right)}(w)\right)^{\left(\frac{r}{2}-1\right)} f_{\left(\frac{r}{2}+1\right)}(w)$$

$$= \left(\frac{r!}{\left(\frac{r}{2}\right)!\left(\frac{r}{2}-1\right)!}\right)^2 \left(I_{F(w)}\left(\frac{r}{2}+1,\frac{r}{2}\right)\right)^{\left(\frac{r}{2}\right)}$$

$$\times \left(1-I_{F(w)}\left(\frac{r}{2}+1,\frac{r}{2}\right)\right)^{\left(\frac{r}{2}-1\right)}$$

$$\times (F(w))^{\left(\frac{r}{2}\right)}(1-F(w))^{\left(\frac{r}{2}-1\right)} f(w)$$

(3.9)

and,

$$G_{\left(\frac{r}{2}+1\right)}(w) =$$

$$\int_0^{I_{F(w)}\left(\frac{r}{2}+1,\frac{r}{2}\right)} \frac{r!}{\left(\frac{r}{2}\right)!\left(\frac{r}{2}-1\right)!}(u)^{\left(\frac{r}{2}\right)}(1-u)^{\left(\frac{r}{2}-1\right)}du$$

$$= I_{F(w)\left(\frac{r}{2}+1,\frac{r}{2}\right)}\left(\frac{r}{2}+1,\frac{r}{2}\right).$$

(3.10)

Hence, $W_{1\left(\frac{r}{2}+1\right)k}$, $W_{2\left(\frac{r}{2}+1\right)k}$, …, $W_{\frac{r}{2}\left(\frac{r}{2}+1\right)k}$, are i.i.d. with the distribution function as in (3.10). However, $W_{1\left(\frac{r}{2}\right)k}$, $W_{2\left(\frac{r}{2}\right)k}$, …, $W_{\frac{r}{2}\left(\frac{r}{2}\right)k}$, $W_{1\left(\frac{r}{2}+1\right)k}$, $W_{2\left(\frac{r}{2}+1\right)k}$, …, $W_{\frac{r}{2}\left(\frac{r}{2}+1\right)k}$, are independent but not identically distributed.

DMRSS for Estimating the Population Mean

The following results are stated and proved in the Appendix. Using DMRSS when the underlying distribution is assumed to be symmetric. $\overline{W}_{DMRSS_O}$ and $\overline{W}_{DMRSS_E}$ are unbiased estimators for μ, and $Var\left(\overline{W}_{DMRSS}\right) \leq Var\left(\overline{X}_{MRSS}\right) \leq \frac{\sigma^2}{n}$, where

$$\overline{X}_{MRSS} = \begin{bmatrix} \dfrac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r/2}(X_{i(\frac{r}{2})k}+X_{i(\frac{r}{2}+1)k}), \text{if r is even} \\ \dfrac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r}X_{i(\frac{r+1}{2})k}, \qquad\qquad \text{if r is odd} \end{bmatrix}$$

(see the Appendix for Theorem 1, Lemma 1 and Theorem 2.

Simulation Study

Based on 5000 replication, a computer simulation is conducted to study the behavior of the efficiency of the sample mean using SRS, MRSS, RSS, ERSS, DERSS and DRSS with respect to DMRSS. Random observations are generated from (1) standard normal distribution (2) Logistic distribution with α =2, β=1 and (3) uniform distribution with θ_1 =0, θ_2=4. The performance of the samples means for r=4,5,6 and7 and m=4 and 6 are investigated.

Results of simulation study

The results of these simulations are summarized by the relative efficiency (the ratio of the variances) of the estimators of the mean. The simulation results are given in Table 3.1.

Table 3.1 shows that estimating the population mean using DMRSS is substantially more efficient than SRS, MRSS, ERSS and RSS. Comparing the sample mean using MRSS with the sample mean using DMRSS, our simulation confirms the results of Theorem 3.2 for the three distributions. Comparing the efficiency for estimating the population mean using DMRSS relative to DERSS, there is a notable difference between them according to the distributions. The best performance was in case of logistic distribution. In normal distribution, the relative efficiency was slightly lower than in logistic distribution.

Clearly in case of uniform distribution, estimating the population mean using DERSS is more efficient than using DMRSS. Also, the population mean estimator using DMRSS is more efficient than the population mean estimator using DRSS, when the underlying distribution is assumed to be symmetric.

Regarding the sample size r, the relative efficiency of the population mean estimators, using DMRSS with respect to any of the other

Table 3.1: The efficiency of the mean estimators using DMRSS relative to the others

m	r	SRS	MRSS	ERSS	RSS	DERSS	DRSS
Normal(2,1)							
4	4	7.51	2.74	3.56	3.13	2.71	1.99
	5	11.83	3.49	5.20	4.53	3.69	2.81
	6	16.41	4.16	6.73	5.21	4.59	3.01
	7	23.36	4.88	8.32	6.34	5.98	3.59
6	4	7.28	2.69	3.67	3.18	2.71	2.01
	5	12.42	3.66	5.26	4.67	3.83	2.81
	6	15.85	3.82	6.86	4.83	4.45	2.81
	7	22.96	4.93	8.54	6.65	6.13	3.80
Logistic (2,1)							
4	4	8.53	2.57	5.20	3.79	4.54	2.80
	5	14.63	3.61	7.70	6.06	7.00	4.08
	6	19.38	4.11	10.74	6.54	10.25	4.35
	7	29.47	4.81	14.74	8.79	13.95	5.43
6	4	8.50	2.77	5.11	3.89	4.61	2.79
	5	15.03	3.77	7.28	6.13	6.90	4.07
	6	19.85	3.83	11.68	6.29	9.84	4.11
	7	29.86	4.95	14.06	8.88	13.26	5.62
Uniform(0,4)							
	4	4.35	2.17	1.44	1.83	0.52	1.00
	5	6.82	2.97	1.89	2.29	0.81	1.21
	6	8.95	3.34	1.70	2.58	0.26	1.27
	7	11.75	4.00	2.10	3.10	0.68	1.35
	4	4.46	2.20	1.45	1.78	0.50	1.04
	5	7.45	3.11	1.94	2.39	0.87	1.25
	6	8.99	3.28	1.60	2.50	0.27	1.27
	7	18.19	4.40	1.75	3.42	0.16	1.46

previous sampling techniques, increases as r increases. While considering the cycle size m, the relative efficiency for the sample mean using DMRSS relative to the other sampling schemes is not affected by the value of m.

Ratio Estimators

Frequently the quantity that is to be estimated from a bivariate random sample is the ratio of two means of two correlated variables, say X and Y, which both vary from unit to unit. For example, in a household survey, the average expenditure on cosmetics per adult female, and the average number of hours per week spent watching television for child aged 10 to 15.

Examples of this kind occur frequently when the sampling unit (the household) comprises a group or cluster of elements and our interest is in the population mean per element. Also, the ratio estimation method is used to obtain increased precision of estimating the population mean or total by taking advantage of the correlation between an auxiliary variable X and the variable of interest Y. In this paper, we assume that the

bivariate random variable (X,Y) has symmetric marginal distributions.

Ratio Estimator Using SRS

Let the bivariate random variable (X,Y) has c.d.f. F(x,y) with means μ_x and μ_y, variances σ_x^2 and σ_y^2, and correlation coefficient ρ, then $R = \dfrac{\mu_y}{\mu_x}$ will denote the population ratio. Using a simple bivariate random sample from F(x, y), the estimator of R is given by:

$$\widehat{R}_{SRS} = \frac{\overline{Y}}{\overline{X}}$$

(4.1)

where $\overline{X}$ and $\overline{Y}$ are the means of X and Y respectively.

Hansen et al.(1953) showed that the variance of $\widehat{R}_{SRS}$ can be approximated by

$$Var(R_{SRS}) \cong \frac{R^2}{n}\left(V_x^2 + V_y^2 - 2\rho\ V_x V_y\right),$$

(4.2)

where $V_x = \dfrac{\sigma_x}{\mu_x}$, $V_y = \dfrac{\sigma_y}{\mu_y}$,

$\rho = \dfrac{E\big((X - \mu_x)(Y - \mu_y)\big)}{\sigma_x \sigma_y}$ and n= rm.

Ratio Estimator Using RSS

Samawi and Muttlak (1996) showed that the ratio estimator using RSS when ranking is on the variable X, is $\widehat{R}_{RSS2} = \dfrac{\overline{Y}_{[r]}}{\overline{X}_{(r)}}$, where

$\overline{Y}_{[r]} = \dfrac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r} Y_{i[i]k}$, $\overline{X}_{(r)} = \dfrac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r} X_{i(i)k}$ and the variance is given by

$$Var\left(\widehat{R}_{RSS}\right) \cong \frac{R^2}{n}$$

$$\times \left\{ \begin{array}{c} V_x^2 + V_y^2 - 2\rho V_x V_y \\[2mm] -m\left[\dfrac{\sum\limits_{i=1}^{r} T_{x(i)}^2}{n\mu_x^2} + \dfrac{\sum\limits_{i=1}^{r} T_{y[i]}^2}{n\mu_y^2} - 2\dfrac{\sum\limits_{i=1}^{r} T_{x(i)y[i]}}{n\mu_x \mu_y}\right] \end{array} \right\},$$

(4.3)

where $T_{x(i)} = \mu_{(i)} - \mu_x$, $T_{y[i]} = \mu_{y[i]} - \mu_y$ and $T_{x(i)y[i]} = (\mu_{x(i)} - \mu_x)(\mu_{y[i]} - \mu_y)$.

As demonstrated by Samawi and Muttluk (1996), that ranking on X is more efficient than ranking on Y in ratio estimation in terms of variance, therefore we introduce only the case where ranking on the variable X is assumed to be without errors. In the next subsections, we will introduced and study the performance of ratio estimators using the double ranked samples discussed in the pervious sections.

Ratio Estimation Using DRSS

Using the notation of Section 2.3, the second stage a subsample of size n=rm, $\{W_{i(i)k}, i = 1,2,\ldots,r, k = 1,2,\ldots,m\}$ is selected. Also, in the second stage, for each $W_{i(i)k}$ measure (quantify) the associated value of the random variable Y. The bivariate DRSS $\{(W_{i(i)k}, Y_{i[i]k}) : i = 1,2,\ldots,r, k = 1,2,\ldots,m\}$ is measured, where $W_{i(i)k}$ as defined above, and $Y_{i[i]k}$ is the corresponding value of Y obtained from the i-th RSS sample in the k-th cycle.

Now, let $\widehat{\mu}_x^{**} = \overline{W}$ and $\widehat{\mu}_y^{**} = \overline{Y}$,

$$\overline{Y} = \frac{1}{rm}\sum_i \sum_k Y_{[i]k}$$

where $\overline{W} = \dfrac{1}{rm}\sum_i \sum_k W_{(i)k}$, and then the estimate of population ratio R using DRSS is given by

$$\widehat{R}_{DRSS} = \frac{\overline{Y}}{\overline{W}}.$$

(4.4)

By using Taylor expansion and assuming large population size, it is easy to show that

$$E\left(\widehat{R}_{DRSS}\right) = \frac{\mu_y}{\mu_x} + O\left(\frac{1}{n}\right),$$

and the variance of $\widehat{R}_{DRSS}$ will be approximated by

$$Var\left(\widehat{R}_{DRSS}\right) \cong \frac{R^2}{n}$$

$$\left(V_x^2 + V_y^2 - 2\rho V_x V_y - m \left(\frac{\sum_{i=1}^{r} T_{w(i)}^{**2}}{\frac{n\mu_x^2}{2}} + \frac{\sum_{i=1}^{r} T_{y[i]}^{**2}}{\frac{n\mu_y^2}{2}} - 2 \frac{\sum_{i=1}^{r} T_{w(i)y[i]}^{**}}{n\mu_x\mu_y} \right) \right)$$

(4.5)

where V_x^2 and V_y^2 as in equation (4.2).

Ratio Estimation Using DMRSS

Using similar modification for bivariate case, and assuming that ranking is on variable X in the two stages. Then as in section 2, $(W_{1(s)k}, Y_{1[s]k})$, $(W_{2(s)k}, Y_{2[s]k})$, …,$(W_{r(s)k}, Y_{r[s]k})$ k=1, 2, …,m will denote the bivariate DMRSS where s is ($\frac{r}{2}$) for the first $\frac{r}{2}$ units and ($\frac{r}{2}+1$) for the last $\frac{r}{2}$ units in case when r is even and ($\frac{r+1}{2}$) when r is odd.

$W_{i(s)k}$ is the s-th smallest X unit in the k-th cycle of the i-th bivariate MRSS in the first stage and $Y_{i[s]k}$ is the corresponding Y observation in the k-th cycle of the i-th bivariate MRSS.
Two cases are considered here:

Case(1): When r is odd, the estimate of the population ratio R using DMRSS$_O$ is defined by

$$\widehat{R}_{DMRSS_O} = \frac{\overline{Y}_O}{\overline{W}_O},\qquad \text{where}\qquad ,$$

$$\overline{W}_O = \frac{1}{mr}\sum_k\sum_i W_{i(\frac{r+1}{2})k}, \quad \overline{Y}_O = \frac{1}{mr}\sum_k\sum_i Y_{i[\frac{r+1}{2}]k}$$

(4.6)

Again, by using Taylor expansion we have

$$E\left(\widehat{R}_{MDRSS_o}\right) = \frac{\mu_y}{\mu_x} + O\left(\frac{1}{n}\right) \text{ and}$$

$$Var(\widehat{R}_{DMRSSo}) \cong$$

$$\frac{R^2}{mr}\left(V_{w(s)}^{**2} + V_{y[s]}^{**2} - 2\rho_{W(s)Y[s]} V_{w(s)}^{**} V_{y[s]}^{**}\right),$$

$$\text{where } s = \frac{r+1}{2}$$

(4.7)

$$V_{w(s)}^{**2} = \frac{\sigma_{w(s)}^{**2}}{\mu_x^2}, \text{and } V_{y[s]}^{**2} = \frac{\sigma_{y[s]}^{**2}}{\mu_y^2}.$$

Case(2): When r is even, the estimate of the population ratio R using DMRSS$_E$ is given by

$$\widehat{R}_{DMRSS_E} = \frac{\overline{Y}_E}{\overline{W}_E},$$

(4.8)

$$\overline{W}_E = \frac{1}{mr}\sum_k\sum_{i=1}^{r} W_{i(s)k}$$

where,

$$= \frac{1}{mr}\sum_k\left(\sum_{i=1}^{r/2} W_{i\left(\frac{r}{2}\right)k} + \sum_{i=1}^{r/2} W_{i\left(\frac{r}{2}+1\right)k}\right)$$

$$\overline{Y}_E = \frac{1}{mr}\sum_k\sum_{i=1}^{r} Y_{i[s]k}$$

and

$$= \frac{1}{mr}\sum_k\left(\sum_{i=1}^{r/2} Y_{i\left[\frac{r}{2}\right]k} + \sum_{i=1}^{r/2} Y_{i\left[\frac{r}{2}+1\right]k}\right).$$

Again by using Taylor expansion, and assuming symmetric underlying distributions then

$$E\left(\widehat{R}_{MDRSS_E}\right) = \frac{\mu_y}{\mu_x} + O\left(\frac{1}{n}\right) \text{ and}$$

$$Var\left(\widehat{R}_{DMRSS_E}\right) \cong \frac{\mu_y^2}{\mu_x^2}$$

$$\left(\frac{\sigma_{w(s)}^{**2}}{mr\mu_x^2} + \frac{\sigma_{y[s]}^{**2}}{mr\mu_y^2} - \frac{\sigma_{w(\frac{r}{2})y[\frac{r}{2}]}^{**} + \sigma_{w(\frac{r}{2}+1)y[\frac{r}{2}+1]}^{**}}{mr\mu_x\mu_y}\right)$$

$$= \frac{R^2}{mr}\left(V_{w(s)}^{**2} + V_{y[s]}^{**2} - 2\rho_{W(s)Y[s]}V_{w(s)}^{**}V_{y[s]}^{**}\right)$$

(4.9)

where $V_{w(s)}^{**2} = \dfrac{\sigma_{w(s)}^{**2}}{\mu_x^2}$, $V_{y[s]}^{**2} = \dfrac{\sigma_{y[s]}^{**2}}{\mu_y^2}$ and s can

be   either $\dfrac{r}{2}$ or $\left(\dfrac{r}{2}+1\right)$ since

$\sigma_{w(\frac{r}{2})}^{**} = \sigma_{w(\frac{r}{2}+1)}^{**}$ and $\sigma_{y[\frac{r}{2}]}^{**} = \sigma_{y[\frac{r}{2}+1]}^{**}$.

Ratio Estimation Using DERSS
        Assume without loss of generality that r is even. The case when r is odd is similar and it will be indicated in the numerical results only. Also assume ranking is on variable X. Let

$$\left(W_{1(1)k}, Y_{1[1]k}\right), \left(W_{2(r)k}, Y_{2[r]k}\right),$$
$$\left(W_{3(1)k}, Y_{3[1]k}\right), ..., \left(W_{r(r)k}, Y_{r[r]k}\right)$$
be the bivariate

DERSS (see Samawi, 2001). This set of bivariate observations is independent but not identically distributed.
The estimate of the population ratio R using *DERSS* is given by

$$\hat{R}_{DERSS} = \frac{\overline{Y}_{DERSS}}{\overline{W}_{DERSS}},$$

(4.10)

where

$$\overline{W}_{DERSS} = \frac{1}{mr}\sum_k\left(\sum_{i\,odd}W_{i(1)k} + \sum_{i\,even}W_{i(r)k}\right),$$

$$= \frac{1}{mr}\sum_k\left(\sum_{i=1}^{r/2}W_{2i-1(1)k} + \sum_{i=1}^{r/2}W_{2i(r)k}\right)$$

$$\overline{Y}_{DERSS} = \frac{\overline{Y}_{[1]} + \overline{Y}_{[r]}}{2},$$

$$\overline{Y}_{[1]} = \frac{1}{m}\sum_k\sum_{i=1}^{r/2}\left(\frac{Y_{2i-1[1]k}}{r/2}\right),$$.

$$and\ \overline{Y}_{[r]} = \frac{1}{m}\sum_k\sum_{i=1}^{r/2}\left(\frac{Y_{2i[1]k}}{r/2}\right)$$

Once again, by using Taylor expansion we have

$$E\left(\widehat{R}_{EDRSS}\right) = \frac{\mu_y}{\mu_x} + O\left(\frac{1}{n}\right)$$ and

$$Var\left(\widehat{R}_{DERSS}\right)$$

$$\cong \frac{R^2}{mr}\left(V_{w(1)}^{**2} + V_{y[1]}^{**2} - 2\rho_{W(1)Y[1]}V_{w(1)}^{**}V_{y[1]}^{**}\right)$$

$$= \frac{R^2}{mr}\left(V_{w(r)}^{**2} + V_{y[r]}^{**2} - 2\rho_{W(r)Y[r]}V_{w(r)}^{**}V_{y[r]}^{**}\right)$$

(4.11)

where

$$V_{w(j)}^{**} = \sigma_{w(j)}^{**}/\mu_x,\ V_{y[j]}^{**} =$$
$$\sigma_{y[j]}^{**}/\mu_y\ and\ \sigma_{w(j)y[j]}^{**} = \rho_{W(j)Y[j]}\sigma_{w(j)}^{**}\sigma_{y[j]}^{**},\ j=1$$

or r.

Simulation Study
        A computer simulation is conducted to study the efficiency of estimating R when ranking is performed on the variable X. Using SRS, RSS, MRSS, ERSS, DRSS, DERSS and DMRSS, bivariate random samples where generated from a bivariate normal distribution with $\mu_x=2$, $\mu_y=4$, $\sigma_x=1$, $\sigma_y=1$ and $\rho=\pm0.9, \pm0.8, \pm0.5$.

        The performance of the ratio estimate will be investigated for r=4, 5, 6 and 7 and m=4 and 6. The ratios of the population means are estimated from SRS, RSS, MRSS, ERSS, DRSS, DERSS and DMRSS data sets. Using 5000 replications,

estimates of the means, the mean square errors and the ratio of the mean squrare errors (relative efficiency) for the ratio were computed.

Results of the simulation study

The values obtained by the simulation study are given in Table 4.1. In all cases the simulation showed that the efficiency of estimating R is not affected by the cycle size m, an explanation for this is that m is canceled in the numerator and dominator when relative efficiency is used. The values in the tables vary from a value of m to another because of the simulation variation. When the underlying distribution is $N_2(2,4,1,1,\rho)$, Table 4.1 shows that estimating the population ratio using DMRSS is more efficient

than using SRS, RSS, and MRSS. Also, using DRSS to estimate the population ratio is more efficient than using SRS and RSS, and using DERSS is more efficient than using SRS, RSS and ERSS.

Moreover, using the definition of relative efficiency, the double sampling schemes can be compared with each other. Our simulation indicates that, estimating the population ratio using DMRSS is more efficient than using DRSS and DERSS. Also, whenever $|\rho|$ increases the efficiency increases in all cases. Note that negative values of ρ give higher efficiency than the positive values.

Table 4.1 Efficiency of the estimators of R when ranking on X and (X,Y) has $N_2(4,2,1,1,\rho)$

M	r	DMRSS relative to			DRSS relative to		DERSS relative to		
		SRS	RSS	MRSS	SRS	RSS	SRS	RSS	ERSS
					$\rho= 0.9$				
4	4	4.15	1.95	1.67	2.79	1.31	2.19	1.03	1.15
	5	5.47	1.14	1.88	3.33	1.30	2.62	1.02	1.16
	6	5.77	2.27	1.89	3.61	1.42	2.77	1.09	1.38
	7	6.05	2.15	1.82	3.78	1.34	2.94	1.04	1.26
6	4	4.24	2.07	1.83	2.85	1.39	2.36	1.15	1.27
	5	4.87	2.14	1.81	3.02	1.33	2.48	1.09	1.20
	6	5.67	2.33	1.87	3.49	1.44	2.75	1.13	1.37
	7	5.91	2.17	1.85	3.72	1.36	2.76	1.01	1.24
					$\rho= 0.8$				
4	4	3.74	1.81	1.62	2.69	1.30	2.24	1.08	1.23
	5	4.28	1.88	1.66	2.99	1.31	2.37	1.04	1.14
	6	4.44	1.77	1.58	3.27	1.31	2.56	1.02	1.24
	7	4.41	1.79	1.61	3.31	1.35	2.54	1.03	1.20
6	4	3.55	1.71	1.53	2.52	1.21	2.14	1.03	1.21
	5	4.09	1.91	1.70	2.72	2.37	2.38	1.11	1.24
	6	4.10	1.91	1.65	2.83	1.32	2.30	1.07	1.19
	7	4.22	1.80	1.59	2.96	1.26	2.43	1.04	1.18
					$\rho= 0.5$				
4	4	3.11	1.69	1.52	2.37	1.29	2.04	1.11	1.19
	5	3.60	1.84	1.63	2.67	1.36	2.20	1.12	1.14
	6	3.90	1.80	1.59	2.88	1.33	2.43	1.12	1.30
	7	3.52	1.56	1.40	2.68	1.19	2.33	1.03	1.24

6	4	2.99	1.68	1.52	2.36	1.33	2.01	1.13	1.16
	5	3.45	1.67	1.51	2.53	1.23	2.15	1.04	1.10
	6	3.66	1.66	1.47	2.85	1.29	2.26	1.02	1.20
	7	3.46	1.65	1.46	2.63	1.26	2.29	1.09	1.22
					$\rho = -0.5$				
4	4	5.12	2.22	1.99	3.17	1.37	2.55	1.11	1.24
	5	6.08	2.40	2.01	3.51	1.38	2.91	1.15	1.34
	6	7.67	2.55	2.01	4.14	1.51	2.94	1.07	1.37
	7	7.23	2.23	1.96	4.22	1.41	3.23	1.08	1.33
6	4	4.52	2.08	1.78	2.96	1.36	2.39	1.10	1.27
	5	5.63	2.39	2.01	3.32	1.41	2.74	1.16	1.34
	6	6.93	2.44	2.00	4.17	1.88	3.11	1.09	1.40
	7	6.91	2.23	1.95	4.21	1.72	3.09	1.00	1.26
					$\rho = -0.8$				
4	4	6.73	2.73	2.37	3.83	1.56	2.89	1.17	1.35
	5	8.77	3.15	2.58	4.19	1.51	3.06	1.10	1.29
	6	10.31	3.52	2.72	4.97	1.70	3.41	1.17	1.42
	7	12.95	3.64	3.03	5.72	1.61	3.83	1.08	1.36
6	4	5.90	2.54	2.19	3.49	1.50	2.71	1.17	1.32
	5	9.33	3.12	2.62	4.24	1.42	3.25	1.09	1.27
	6	9.07	3.30	2.71	4.42	1.61	3.13	1.14	1.43
	7	11.84	3.60	2.79	5.62	1.71	3.72	1.13	1.39
					$\rho = -0.9$				
4	4	6.76	2.18	2.43	3.50	1.45	2.72	1.13	1.34
	5	11.03	3.84	3.05	4.54	1.58	3.52	1.23	1.36
	6	12.91	3.83	3.13	5.20	1.54	3.66	1.08	1.45
	7	17.19	4.66	3.56	6.36	1.72	3.99	1.08	1.42
6	4	6.84	2.90	2.46	3.56	1.51	2.78	1.18	1.40
	5	10.63	3.69	2.92	4.41	1.53	3.40	1.18	1.36
	6	12.82	4.04	3.19	5.17	1.63	3.43	1.08	1.41
	7	16.33	4.50	3.50	5.80	1.60	3.77	1.04	1.37

Application To Real Data Set And Conclusions

We illustrate the double ranked sample mean estimation procedure using a real data set which consists of the height (Y) and the diameter (X) at breast height of 399 trees. See Platt et al. (1988) for a detailed description of the data set. The summary statistics for the data are reported in Table 5.1. Note that the correlation coefficient is $\rho = 0.908$.

Table 5.1. Summary Statistics of trees data.

Variable	Mean	Variance
Height (Y) in feet	52.36	325.14
Diameter (X) in cm	20.84	310.11

In this article, ranking is performed on the variable X exactly measured. However, in practice ranking is done before any actual quantification. Using a set size r = 3 and the cycle size m = 3, we draw bivariate SRS, DRSS, and DMRSS of size 9,

however DERSS is the same as DRSS in this case. Table 5.2 contains all the above proposed estimators and their estimated variances using the drawn samples.

Table 5.2. Results from the drawn samples.

Sample	Naïve Estimator of the Diameter (X)	9(Estimated Variance)[*]	Ratio Estimator	9(Estimated Variance)[*]
SRS	13.57	168.60	2.50	1.036
DRSS	19.39	148.37	2.29	0.633
DMRSS	15.89	131.35	2.15	0.297

Table 5.2 confirms the simulation results. However, this example is just to illustrate the application using the proposed estimators.

Finally, the theortical and simulation results showed that the population mean estimator using DMRSS is an unbiased estimator for the population mean whenever the underlying distribution is assumed to be symmetric. Also, it was shown theoretically that the variance of this estimator is less than the variance of the sample mean using MRSS (the first stage). Although using numerical simulation it was noticed that the sample mean based on DMRSS is more efficient than using other sampling methods (see Table 3.1) with respect to there variances.

Note there are difficulties in selecting the DMRSS because of the similarity of the subjects from the first stage. However, in practice this is not a problem because the number of units we rank in the second stage will not exceed 5.

In ratio estimation using the two stage sampling for different schemes, the estimator of the population ratio of two variables was introduced and the variance in each case was derived. Our numerical study indicated that the two stage sampling is more efficient than the first stage sampling considering the same sampling scheme with respect to their variances.

Comparing the two stage sampling schemes, namely DRSS, DMRSS and DERSS, superiority in efficiency depends on the distribution of the bivariate variable. However, DMRSS was more efficient than DRSS and DERSS when the underlying distribution is the bivariate normal. Moreover, those efficiencies depend on the set size r and the strength of the correlation between X and Y.

References

Al-Saleh, M. F., & Al-Kadiri, M. A. (2000). Double ranked set sampling. *Statistics and probability letters*, *48*(2), 205-212.

Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A first course in order statistics*. New York: John Wiley and Sons, Inc.

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). *Sampling survey methods and theorey. Vol. 2*. John Wiley and Sons.

Kaur, A., Patil, G. P., Sinha, A. K., & Tailie, C. (1995). Ranked set sampling: an annotated bibliography. *Environmental and Ecological Statistics*, *2*, 25-54.

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

Muttlak, H. A. (1997). Median ranked set sampling. *Journal of Applied Statistical Science*, *6*(4), 245-255.

Patil, G. P., Sinha, A. K., & Taillie, C. (1999). Ranked set sampling: A bibliography. *Environmental and Ecological Statistics*, 6, 91-98.

Platt, W. J., Evans, G. W., & Rathbun, S.L. (1988). The population dynamics of a long-lived conifer. *The American Naturalist*, *131*, 391-525.

Samawi, H. M. (2001, in press). On double extreme ranked set sample with application to regression estimator.

Samawi, H. M. Ahmed, M. S., & Abu Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. *Biometrical Journal*, *38*, (5) 577-586.

Samawi, H. M., & Al-Sageer, O. A. (2001). On the estimation of the distribution function using extreme and median ranked set sampling. *Biometrical Journal, 43*(3), 357- 373.

Samawi, H. M., & Muttlak, H. A. (1996). Estimation of ratio using ranked set sampling. *Biometrical Journal*, *38* (6), 753-764.

Samawi, H. M., & Muttlak, H. A.(2001). On ratio estimation using median ranked set sampling. *Journal of Applied Statistical Science, 10*(2), 89-98.

Takahasi, K., & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the stratified sampling by means of ordering. *Annals of the Institute of Statistical Mathematics*, *20*, 1-31.

Yang, H.(1980). On the Variance of median and some other statistics. *Bulletin of Institute Mathematics Academic Simulation*, *10*, 197-204.

Appendix

Theorem 1: Let X be a random variable with symmetric distribution function F (x) and mean μ, then

    (a) $g_{\left(\frac{r+1}{2}\right)}(w)$ is symmetric about μ , if r is odd.

    (b)   Without loss of generality, assume that $\mu = 0$, then $W_{\left(\frac{r}{2}\right)} \underset{=}{d} - W_{\left(\frac{r}{2}+1\right)}$ and

hence $\dfrac{\mu_{\frac{r}{2}}^{**} + \mu_{\frac{r}{2}+1}^{**}}{2} = 0$ if r is even, where

$\mu_{\frac{r}{2}}^{**} = E\left(W_{\left(\frac{r}{2}\right)}\right)$ and $\mu_{\frac{r}{2}+1}^{**} = E\left(W_{\left(\frac{r}{2}+1\right)}\right)$.

Proof:

    (a) Because $f_{\left(\frac{r+1}{2}\right)}(x)$ is symmetric about 0 (see Arnold, et al. 1992), then by using (3.5)

$g_{\left(\frac{r+1}{2}\right)}(-w) = g_{\left(\frac{r+1}{2}\right)}(w)$. Therefore, $g_{\left(\frac{r+1}{2}\right)}(w)$ is symmetric about 0.

    (b) Using the fact that in case of symmetry $X_{\left(\frac{r}{2}\right)} \underset{=}{d} - X_{\left(\frac{r}{2}+1\right)}$, when r is even, and from (3.7) and (3.9), then it is clear that

$g_{\left(\frac{r}{2}+1\right)}(-w) = g_{\left(\frac{r}{2}\right)}(w)$. Therefore, $W_{\left(\frac{r}{2}\right)} \underset{=}{d} - W_{\left(\frac{r}{2}+1\right)}$. Also note that

$E\left(W_{\left(\frac{r}{2}\right)}\right) = E\left(-W_{\left(\frac{r}{2}+1\right)}\right) \Rightarrow \mu_{\left(\frac{r}{2}\right)}^{**} = -\mu_{\left(\frac{r}{2}+1\right)}^{**}$

and hence $\dfrac{\mu_{\left(\frac{r}{2}\right)}^{**} + \mu_{\left(\frac{r}{2}+1\right)}^{**}}{2} = 0$.

Lemma 1: Let $W_{1\left(\frac{r+1}{2}\right)k}, W_{2\left(\frac{r+1}{2}\right)k}, \ldots, W_{r\left(\frac{r+1}{2}\right)k}$,

k=1,2,…, m be the $DMRSS_O$ when r is odd, and

$W_{1\left(\frac{r}{2}\right)k}, \quad W_{2\left(\frac{r}{2}\right)k}, \quad \cdots, \quad W_{\frac{r}{2}\left(\frac{r}{2}\right)k}, \quad W_{1\left(\frac{r}{2}+1\right)k},$

$W_{2\left(\frac{r}{2}+1\right)k}, \quad \cdots, \quad W_{\frac{r}{2}\left(\frac{r}{2}+1\right)k}$, k=1, 2, …, m be the

$DMRSS_E$ when r is even. If the c.d.f. F(x) is symmetric about its mean μ, then $\overline{W}_{DMRSS_O}$ and $\overline{W}_{DMRSS_E}$ are unbiased estimators for μ.

Proof: The proof is a consequence of Theorem 1.

Theorem 2: If the random variable X has a symmetric distribution function F(x) about μ, then

$$Var\left(\overline{W}_{DMRSS}\right) \le Var\left(\overline{X}_{MRSS}\right) \le \frac{\sigma^2}{n}, \text{ where}$$

$$\overline{X}_{MRSS} = \begin{bmatrix} \dfrac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r/2}(X_{i\left(\frac{r}{2}\right)k} + X_{i\left(\frac{r}{2}+1\right)k}), \text{if r is even} \\ \dfrac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r} X_{i\left(\frac{r+1}{2}\right)k}, \qquad \text{if r is odd} \end{bmatrix}$$

.

Proof: Because Yang (1982) showed that $Var\left(X_{(med)}\right) \le \sigma^2$,

where $X_{(med)}$ is the sample median of i.i.d sample of size r, then we need to prove only that $Var\left(\overline{W}_{DMRSS}\right) \le Var\left(\overline{X}_{MRSS}\right)$.

Case 1: When r is odd. Because $W_{1\left(\frac{r+1}{2}\right)k}, W_{2\left(\frac{r+1}{2}\right)k},$

…, $W_{r\left(\frac{r+1}{2}\right)k}$, k=1,2,…,m, are i.i.d. from $G_{\left(\frac{r+1}{2}\right)}(w)$ then the prove is similar to that by Yang (1982).

Case 2: When r is even, $W_{i\left(\frac{r}{2}\right)k}$, i =1, 2, …, $\frac{r}{2}$, k=1, 2, …, m are i.i.d. with (3.8) distribution function and $W_{1\left(\frac{r}{2}+1\right)k}$, $W_{2\left(\frac{r}{2}+1\right)k}$, …, $W_{\frac{r}{2}\left(\frac{r}{2}+1\right)k}$, are i.i.d. with the distribution function as in (3.10). The two samples are independent. Also, assuming that $\mu = 0$ we have that $W_{\left(\frac{r}{2}\right)}\underline{\underline{d}} - W_{\left(\frac{r}{2}+1\right)}$, then by using Yang (1982)

$$Var\left(\frac{W_{\left(\frac{r}{2}\right)} + W_{\left(\frac{r}{2}+1\right)}}{2}\right) \leq \frac{\sigma^2_{\left(\frac{r}{2}\right)} + \sigma^2_{\left(\frac{r}{2}+1\right)}}{2} = \sigma^2_{\left(\frac{r}{2}\right)}$$

,

and hence $Var\left(\overline{W}_{DMRSS_E}\right) \leq Var\left(\overline{X}_{MRSS_E}\right)$.