International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:11, 2016

Investigating the Efficiency of Stratified Double Median Ranked Set Sample for Estimating the Population Mean

Mahmoud I. Syam

***Abstract*—**Stratified double median ranked set sampling (SDMRSS) method is suggested for estimating the population mean. The SDMRSS is compared with the simple random sampling (SRS), stratified simple random sampling (SSRS), and stratified ranked set sampling (SRSS). It is shown that SDMRSS estimator is an unbiased of the population mean and more efficient than SRS, SSRS, and SRSS. Also, by SDMRSS, we can increase the efficiency of mean estimator for specific value of the sample size. SDMRSS is applied on real life examples, and the results of the example agreed the theoretical results.

***Keywords*—**Efficiency, double ranked set sampling, median ranked set sampling, ranked set sampling, stratified.

I. Introduction

DURING the last years, the ranked set sampling method which was proposed by McIntyre [6] to estimate mean pasture yields was developed and modified by many authors to estimate the mean of the population. Takahasi and Wakimoto [10] have established an accurate mathematical theory of ranked set sampling and they get the same results. Dell and Clutter [4] showed that the mean of the RSS is an unbiased estimator of the population mean, whatever or not there are errors in ranking. Samawi [9] introduced the stratified ranked set sample method for estimating the population mean. Muttlak [7] suggested using median ranked set sampling (MRSS) to estimate the population mean. Al-Saleh and Al-Kadiri [1] introduced double ranked set sampling for estimating the population mean. Al-Saleh and Al-Omari [2] suggested multistage ranked set sampling (MSRSS) that increases the efficiency of estimating the population mean for specific value of the sample size. Jemain and Al-Omari [5] suggested multistage median ranked set sampling (MMRSS) to estimate the population mean and they showed that the efficiency of the mean estimator using MMRSS can be increased for specific value of the sample size *m* by increasing the number of stages. For more about RSS, see [3], [8].

In this paper, the SDMRSS is suggested to estimate the population mean of symmetric and asymmetric distributions. The organization of this paper is as follows: In Section II, SDMRSS is presented. Estimation of the population mean is given in Section III. A simulation study is considered in Section IV. A real life example using SDMRSS is discussed in Section V. Finally, conclusions on the suggested estimator are introduced in Section VI.

Mahmoud. I. Syam is a lecturer in Qatar University, Foundation Program, Mathematics Department, P.O.Box 2713, Doha, Qatar (phone: 00974-55468238; e-mail: M.syam@qu.edu.qa).

II. SDMRSS

In stratified sampling, the population of N units is first divided into L subpopulations of $N_1, N_2, \cdots, N_L$ units, respectively. These subpopulations are no overlapping and together they comprise the whole population, so that $N_1 + N_2 + \cdots + N_L = N$. The subpopulations are called strata. To obtain the full benefit from stratification, the values of the N_h $(h = 1,2,\ldots,L)$ must be known. When the strata have been determined, a sample is drawn from each, and the drawings are made in different strata. The sample sizes within the strata are denoted by $n_1, n_2, \cdots, n_L$, respectively. The appropriate allocation of samples to different strata is very important in stratified sampling. In this article, the type of allocation method is proportional to the stratum size. If a simple random sample is taken in each stratum, the whole procedure is described as SSRS.

The double median ranked set sampling (DMRSS) is described as:

1) Identify n^3 elements from the target population and divide these elements randomly into n sets each of size n^2 elements.
2) If the sample size is even, select from the first $\frac{n^2}{2}$ sets the $\left(\frac{n}{2}\right)$th smallest rank, and from the second $\frac{n^2}{2}$ sets the $\left(\frac{n}{2}+1\right)$th smallest rank. If the sample size is odd, select from all sets the $\left(\frac{n+1}{2}\right)$th smallest rank which means the median of the sample. This step yields n sets of size n.
3) Apply the MRSS procedure again on the sets obtained from step (2) to obtain a DMRSS of size n.
4) The cycle can be repeated m times if needed to get a sample of size nm units.
5) If the DMRSS is used in each stratum, the whole procedure is described as SDMRSS.

To illustrate the method, let us consider the following example, which combines an even and an odd sample size in two strata.

Example 1: Consider a population with two strata $L = 2$ $(h = 1, 2)$ and in the first stratum there are 27 elements divided into 3 sets, 9 elements in each set, and in the second stratum there are 64 elements divided into 3 sets, 16 elements in each set, as the following:

Stratum (1): Assume the 27 elements are

$$X_{11}^{(1)}, X_{12}^{(1)}, \cdots, X_{33}^{(1)}, X_{11}^{(2)}, X_{12}^{(2)}, \cdots, X_{33}^{(2)}, X_{11}^{(3)}, X_{12}^{(3)}, \cdots, X_{33}^{(3)}$$

After ranking the elements in each set, the following sets are obtained

$$\begin{bmatrix} X_{(11)}^{(1)} & X_{(12)}^{(1)} & X_{(13)}^{(1)} \\ X_{(21)}^{(1)} & X_{(22)}^{(1)} & X_{(23)}^{(1)} \\ X_{(31)}^{(1)} & X_{(32)}^{(1)} & X_{(33)}^{(1)} \end{bmatrix}, \begin{bmatrix} X_{(11)}^{(2)} & X_{(12)}^{(2)} & X_{(13)}^{(2)} \\ X_{(21)}^{(2)} & X_{(22)}^{(2)} & X_{(23)}^{(2)} \\ X_{(31)}^{(2)} & X_{(32)}^{(2)} & X_{(33)}^{(2)} \end{bmatrix} \text{ and } \begin{bmatrix} X_{(11)}^{(3)} & X_{(12)}^{(3)} & X_{(13)}^{(3)} \\ X_{(21)}^{(3)} & X_{(22)}^{(3)} & X_{(23)}^{(3)} \\ X_{(31)}^{(3)} & X_{(32)}^{(3)} & X_{(33)}^{(3)} \end{bmatrix}$$

Apply MRSS on each of the 9 elements to get 3 sets as:

Set (1): $X_{(12)}^{(1)}, X_{(22)}^{(1)}, X_{(32)}^{(1)}$

Set (2): $X_{(12)}^{(2)}, X_{(22)}^{(2)}, X_{(32)}^{(2)}$

Set (3): $X_{(12)}^{(3)}, X_{(22)}^{(3)}, X_{(32)}^{(3)}$

Elements of the double median ranked set sample in the first stratum are

$$X_{(22)}^{(1)}, X_{(22)}^{(2)}, X_{(22)}^{(3)}$$

Stratum (2): Assume the 64 elements are

$$Y_{11}^{(1)}, Y_{12}^{(1)}, \cdots, Y_{44}^{(1)}, Y_{11}^{(2)}, Y_{12}^{(2)}, \cdots, Y_{44}^{(2)},$$
$$y_{11}^{(3)}, y_{12}^{(3)}, \cdots, Y_{44}^{(3)}, y_{11}^{(4)}, y_{12}^{(4)}, \cdots, Y_{44}^{(4)}$$

After ranking the elements in each set, the following sets are obtained

$$\begin{bmatrix} Y_{(11)}^{(1)} & Y_{(12)}^{(1)} & Y_{(13)}^{(1)} & Y_{(14)}^{(1)} \\ Y_{(21)}^{(1)} & Y_{(22)}^{(1)} & Y_{(23)}^{(1)} & Y_{(24)}^{(1)} \\ Y_{(31)}^{(1)} & Y_{(32)}^{(1)} & Y_{(33)}^{(1)} & Y_{(34)}^{(1)} \\ Y_{(41)}^{(1)} & Y_{(42)}^{(1)} & Y_{(43)}^{(1)} & Y_{(44)}^{(1)} \end{bmatrix}, \begin{bmatrix} Y_{(11)}^{(2)} & Y_{(12)}^{(2)} & Y_{(13)}^{(2)} & Y_{(14)}^{(2)} \\ Y_{(21)}^{(2)} & Y_{(22)}^{(2)} & Y_{(23)}^{(2)} & Y_{(24)}^{(2)} \\ Y_{(31)}^{(2)} & Y_{(32)}^{(2)} & Y_{(33)}^{(2)} & Y_{(34)}^{(2)} \\ Y_{(41)}^{(2)} & Y_{(42)}^{(2)} & Y_{(43)}^{(2)} & Y_{(44)}^{(2)} \end{bmatrix}$$

$$\begin{bmatrix} Y_{(11)}^{(3)} & Y_{(12)}^{(3)} & Y_{(13)}^{(3)} & Y_{(14)}^{(3)} \\ Y_{(21)}^{(3)} & Y_{(22)}^{(3)} & Y_{(23)}^{(3)} & Y_{(24)}^{(3)} \\ Y_{(31)}^{(3)} & Y_{(32)}^{(3)} & Y_{(33)}^{(3)} & Y_{(34)}^{(3)} \\ Y_{(41)}^{(3)} & Y_{(42)}^{(3)} & Y_{(43)}^{(3)} & Y_{(44)}^{(3)} \end{bmatrix}, \begin{bmatrix} Y_{(11)}^{(4)} & Y_{(12)}^{(4)} & Y_{(13)}^{(4)} & Y_{(14)}^{(4)} \\ Y_{(21)}^{(4)} & Y_{(22)}^{(4)} & Y_{(23)}^{(4)} & Y_{(24)}^{(4)} \\ Y_{(31)}^{(4)} & Y_{(32)}^{(4)} & Y_{(33)}^{(4)} & Y_{(34)}^{(4)} \\ Y_{(41)}^{(4)} & Y_{(42)}^{(4)} & Y_{(43)}^{(4)} & Y_{(44)}^{(4)} \end{bmatrix}$$

Apply MRSS on each of the 16 elements to get four sets as:

Set (1): $Y_{(12)}^{(1)}, Y_{(22)}^{(1)}, Y_{(32)}^{(1)}, Y_{(42)}^{(1)}$

Set (2): $Y_{(12)}^{(2)}, Y_{(22)}^{(2)}, Y_{(32)}^{(2)}, Y_{(42)}^{(2)}$

Set (3): $Y_{(13)}^{(3)}, Y_{(23)}^{(3)}, Y_{(33)}^{(3)}, Y_{(43)}^{(3)}$

Set (4): $Y_{(13)}^{(4)}, Y_{(23)}^{(4)}, Y_{(33)}^{(4)}, Y_{(43)}^{(4)}$

Elements of the double median ranked set sample in the second stratum are $Y_{(22)}^{(1)}, Y_{(22)}^{(2)}, Y_{(33)}^{(3)}, Y_{(33)}^{(4)}$. Therefore, the SDMRSS units are,

$$X_{(22)}^{(1)}, X_{(22)}^{(2)}, X_{(22)}^{(3)}, Y_{(22)}^{(1)}, Y_{(22)}^{(2)}, Y_{(33)}^{(3)}, Y_{(33)}^{(4)}$$

III. Estimation of the Population Mean

Assume that the variable of interest *X* has density *f(x)* and cumulative distribution function *F(x)*, with mean μ and variance σ^2. Let $X_1, X_2, \ldots, X_n$ be *i.i.d. f(x).* Let $X_{11}, X_{12}, \ldots, X_{1n}; X_{21}, X_{22}, \ldots, X_{2n}; \ldots; X_{n1}, X_{n2}, \ldots, X_{nn}$ be *n* independent simple random samples each of size *n*. Let $X_{i(1:n)}, X_{i(2:n)}, \ldots, X_{i(n:n)}$ be the order statistics of the *i*th sample $X_{i1}, X_{i2}, \ldots, X_{in}$, $i = 1, 2, \ldots, n$. Therefore $X_{i(1:n)}, X_{i(2:n)}, \ldots, X_{i(n:n)}$ is the measured ranked set sample.

The notation for the DMRSS will be used by replacing the sets of ordered statistics $X_{i(1:n)}, X_{i(2:n)}, \ldots, X_{i(n:n)}$, $i = 1, 2, \ldots, n$ which are obtained from an *i.i.d.* sample $X_{i1}, X_{i2}, \ldots, X_{in}$, $i = 1, 2, \ldots, n$. If the sample size *n* is even, the ordered statistics will be replaced by $X^*_{1(\frac{n}{2}:n)}, \ldots, X^*_{\frac{n}{2}(\frac{n}{2}:n)}, X^*_{\frac{n}{2}+1(\frac{n+2}{2}:n)}, \ldots, X^*_{n(\frac{n+2}{2}:n)}$, and if the sample size *n* is odd, the ordered statistics will be replaced by

$$X^*_{1(\frac{n+1}{2}:n)}, \ldots, X^*_{\frac{n+1}{2}-1(\frac{n+1}{2}:n)}, X^*_{\frac{n+1}{2}(\frac{n+1}{2}:n)}, X^*_{\frac{n+1}{2}+1(\frac{n+1}{2}:n)}, \ldots, X^*_{n(\frac{n+1}{2}:n)}$$

which are independent but not identically distributed random variables, where $X^*_{i(\frac{n+1}{2}:n)}$ is the median of the *i*th sample.

The notation for the SDMRSS will be by replacing the sets of ordered statistics $X_{hi(1:n_h)}, X_{hi(2:n_h)}, \ldots, X_{hi(n_h:n_h)}$, $i = 1, 2, \ldots, n$ and $h = 1, 2, \ldots, n$. If the sample size n_h is even, the ordered statistics will be replaced by $X^*_{h1(\frac{n_h}{2}:n_h)}, \ldots, X^*_{h\frac{n_h}{2}(\frac{n_h}{2}:n_h)}, X^*_{h\frac{n_h}{2}+1(\frac{n_h+2}{2}:n_h)}, \ldots, X^*_{hn_h(\frac{n_h+2}{2}:n_h)}$, and if the sample size n_h is odd, the ordered statistics will be replaced by

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:11, 2016

$$X^*_{h1(\frac{n_h+1}{2}:n_h)},\ldots,X^*_{h\frac{n_h+1}{2}-1(\frac{n_h+1}{2}:n_h)},X^*_{h\frac{n_h+1}{2}(\frac{n_h+1}{2}:n_h)},$$
$$X^*_{h\frac{n_h+1}{2}+1(\frac{n_h+1}{2}:n_h)},\ldots,X^*_{hn_h(\frac{n_h+1}{2}:n_h)}$$

which are independent but not identically distributed random variables, where $X^*_{hi(\frac{n_h+1}{2}:n_h)}$ is the median of the i^{th} sample in h^{th} stratum.

In the case of SDMRSS, when n_h is even, the estimator of the population mean is defined as

$$\overline{X}^*_{SDMRSS1} = \sum_{h=1}^{L} W_h \cdot \overline{X}_{DMRSS1h}$$
$$= \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{\frac{n_h}{2}} X^*_{hi(\frac{n_h}{2}:n_h)} + \sum_{i=\frac{n_h}{2}+1}^{n_h} X^*_{hi(\frac{n_h+2}{2}:n_h)}\right) \quad (1)$$

In the case of SDMRSS, when n_h is odd, the estimator of the population mean is defined as

$$\overline{X}^*_{SDMRSS2} = \sum_{h=1}^{L} W_h \cdot \overline{X}_{DMRSS2h}$$
$$= \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{n_h} X^*_{hi(\frac{n_h+1}{2}:n_h)}\right) \quad (2)$$

where $W_h = \frac{N_h}{N}$, N_h is the stratum size and N is the total population size.

The variance of SDMRSS1 (when n_h is even) is given by

$$Var(\overline{X}^*_{SDMRSS1}) = \sum_{h=1}^{L} \frac{W_h^2}{n_h^2}\left(\sum_{i=1}^{\frac{n_h}{2}} \sigma^{2*}_{hi(\frac{n_h}{2}:n_h)} + \sum_{i=\frac{n_h}{2}+1}^{n_h} \sigma^{2*}_{hi(\frac{n_h+2}{2}:n_h)}\right) \quad (3)$$

The variance of SDMRSS2 (when n_h is odd) is given by

$$Var(\overline{X}^*_{SDMRSS2}) = \sum_{h=1}^{L} \frac{W_h^2}{n_h^2}\left(\sum_{i=1}^{n_h} \sigma^2_{hi(\frac{n_h+1}{2}:n_h)}\right) \quad (4)$$

Property 1. $\overline{X}_{SDMRSS}$ is an unbiased estimator of a symmetric population mean.

Proof. Two cases are considered:

First: If the sample sizes in the strata $n_h, h = 1,\cdots,L$ are even,

$$\overline{X}^*_{SDMRSS1} = \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{\frac{n_h}{2}} X^*_{hi(\frac{n_h}{2}:n_h)} + \sum_{i=\frac{n_h}{2}+1}^{n_h} X^*_{hi(\frac{n_h+2}{2}:n_h)}\right)$$

$$E\left(\overline{X}^*_{SDMRSS1}\right) = E\left(\sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{\frac{n_h}{2}} X^*_{hi(\frac{n_h}{2}:n_h)} + \sum_{i=\frac{n_h}{2}+1}^{n_h} X^*_{hi(\frac{n_h+2}{2}:n_h)}\right)\right)$$
$$= \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{\frac{n_h}{2}} E(X^*_{hi(\frac{n_h}{2}:n_h)}) + \sum_{i=\frac{n_h}{2}+1}^{n_h} E(X^*_{hi(\frac{n_h+2}{2}:n_h)})\right)$$
$$= \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{\frac{n_h}{2}} \mu^*_{h(\frac{n_h}{2})} + \sum_{i=\frac{n_h}{2}+1}^{n_h} \mu^*_{h(\frac{n_h+2}{2})}\right)$$

Since the distribution is symmetric about μ, then $\mu^*_{h(\frac{n_h}{2})} + \mu^*_{h(\frac{n_h+2}{2})} = 2\mu^*_h$. Therefore,

$$E\left(\overline{X}^*_{SDMRSS1}\right) = \left(\sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{n_h} \mu^*_{hi}\right)\right)$$

It was shown by Al-Saleh and Al-Kadiri [1] that $\mu_h = \frac{1}{n_h}\left(\sum_{i=1}^{n_h} \mu^*_{hi}\right)$, therefore $E\left(\overline{X}^*_{SDMRSS1}\right) = \left(\sum_{h=1}^{L} W_h \cdot \mu_h\right) = \mu$

Second: If the sample sizes in the strata $n_h, h = 1,\cdots,L$ are odd,

$$E\left(\overline{X}^*_{SDMRSS2}\right) = E\left(\sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{n_h} X^*_{hi(\frac{n_h+1}{2}:n_h)}\right)\right)$$
$$= \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{n_h} E(X^*_{hi(\frac{n_h+1}{2}:n_h)})\right)$$
$$= \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{n_h} \mu^*_{h(\frac{n_h+1}{2})}\right)$$

Since the distribution is symmetric about the μ, then mean = median, which implies $\mu^*_{h(\frac{n_h+1}{2})} = \mu_{h(\frac{n_h+1}{2})}$. Therefore,

$$E\left(\overline{X}^*_{SDMRSS2}\right) = \sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left(\sum_{i=1}^{n_h} \mu_{h\left(\frac{n_h+1}{2}\right)}\right)$$

$$=\sum_{h=1}^{L} W_h \cdot \frac{1}{n_h}\left((n_h)(\mu_h)\right)=\sum_{h=1}^{L} W_h \cdot \mu_h \ =\mu$$

Property 2. If the distribution is symmetric about μ, then $Var\left(\overline{X}_{SDMRSS}\right) < Var\left(\overline{X}_{SRS}\right)$.

Proof. Two cases are considered:

First: If the sample sizes in the strata n_h , $h=1,\cdots,L$ are even, The variance of $\overline{X}_{SDMRSS}$ is

$$Var\left(\overline{X}^{*}_{SDMRSS1}\right)=\sum_{h=1}^{L}\frac{W_h^2}{n_h^2}\left(\sum_{i=1}^{\frac{n_h}{2}}\sigma^{2*}_{hi(\frac{n_h}{2}:n_h)}+\sum_{i=\frac{n_h}{2}+1}^{n_h}\sigma^{2*}_{hi(\frac{n_h+2}{2}:n_h)}\right)$$

Nevertheless, $\sigma^{2*}_{hi} < \sigma^2_h$ for each stratum $h=1,2,\cdots,L$, this implies

$$Var\left(\overline{X}^{*}_{SDMRSS1}\right)=\sum_{h=1}^{L} W_h^2 \cdot \frac{1}{n_h}\sigma^{2*}_{hi} < \sum_{h=1}^{L} W_h^2 \cdot \frac{1}{n_h}\sigma^{2}_{h}$$
$$=Var(\overline{X}_{SSRS}) < Var(\overline{X}_{SRS})$$

Second: The proof in case of odd sample sizes is similar.

IV. Simulation Study based on SDMRSS

In this section, a simulation study is conducted to investigate the performance of SDMRSS for estimating the population mean. Symmetric and asymmetric distributions have been considered for samples of sizes $n=9,12,14,15,18$ assuming that the population is partitioned into two or three strata and proportional allocation is considered. The simulation was performed for the SRSS, SSRS, and SRS data sets from different distributions symmetric and asymmetric. The symmetric distributions are uniform and normal, and the asymmetric distributions are exponential, gamma and Weibull. Using 100000 replications, estimates of the means, variances and mean square errors were computed.

If the distribution is symmetric, then the efficiency of SDMRSS relative to T is defined by

$$eff\left(\overline{X}_{SDMRSS}, \overline{X}_{T}\right)=\frac{Var\left(\overline{X}_{T}\right)}{Var\left(\overline{X}_{SDMRSS}\right)}, \tag{5}$$

where *T = SRS, SSRS SRSS.*

While if the distribution is asymmetric, then the efficiency of SDMRSS relative to T is defined by

$$eff\left(\overline{X}_{SDMRSS}, \overline{X}_{T}\right)=\frac{MSE\left(\overline{X}_{T}\right)}{MSE\left(\overline{X}_{SDMRSS}\right)}, \tag{6}$$

where T = SRS, SSRS SRSS.

The values of the relative efficiency found under different distributional assumptions are provided in Tables I and II.

TABLE I
THE RELATIVE EFFICIENCY FOR ESTIMATING THE POPULATION MEAN USING SDERSS WITH RESPECT TO SRSS, SSRS AND SRS WITH SAMPLE SIZE N = 9 AND N = 12

Distribution	n = 9: $n_1=4$, $n_2=5$			**n = 12:** $n_1=5$, $n_2=7$		
	SRSS	*SSRS*	*SRS*	***SRSS***	***SSRS***	***SRS***
Uniform (0,1)	24.9	29.0	29.3	38.2	39.4	40.8
Normal (0,1)	22.5	24.4	24.6	34.7	36.4	35.6
Exponential (1)	19.2	17.7	17.7	23.4	24.6	22.9
Gamma (1,2)	17.4	17.9	17.8	23.5	24.2	23.5
Weibull (1,2)	17.3	17.0	16.9	22.8	22.8	22.4

TABLE II
THE RELATIVE EFFICIENCY FOR ESTIMATING THE POPULATION MEAN USING SDMRSS WITH RESPECT TO SRSS, SSRS AND SRS WITH SAMPLE SIZE N =14 AND N =18

Distribution	***n* =14:** $n_1=8$, $n_2=6$,			***n*=18:** $n_1=8$, $n_2=10$,		
	SRSS	***SSRS***	***SRS***	***SRSS***	***SSRS***	***SRS***
Uniform (0,1)	43.6	47.9	45.6	60.9	66.2	63.3
Normal (0,1)	39.5	41.8	38.8	45.3	49.0	49.3
Exponential (1)	21.9	23.2	21.6	22.6	21.4	21.9
Gamma (1,2)	23.3	22.9	22.3	22.5	22.3	22.4
Weibull (1,2)	21.2	21.4	21.2	23.7	22.9	29.8

From the tables, we can notice that greater efficiency is attained using SDMRSS method as opposed to the other contending methods that have been discussed to estimate the population mean of the variable of interest.

When the performance of SDMRSS is compared to either SRSS, SSRS, or SRS, it is found that SDMRSS is more efficient, as shown by all the values of relative efficiency which are greater than 1.

TABLE III
THE VALUES OF BIAS OF SDERSS FOR DIFFERENT DISTRIBUTIONS AND DIFFERENT NUMBERS OF STRATA

Exponential (1)	Gamma (1,2)	Weibull (1,2)
When $n=9$ and two strata $n_1=4$ and $n_2=5$		
0.06	0.18	0.
When $n=12$ and two strata $n_1=5$ and $n_2=7$		
0.03	0.19	0.3
When $n=15$ and three strata $n_1=3$, $n_2=5$ and $n_3=7$		
0.01	0.08	0.00
When $n=18$ and three strata $n_1=4$, $n_2=6$ and $n_3=8$.		
0.03	0.10	0.02

When the performances of the suggested estimators are compared, the efficiency of the suggested estimator is found to be more superior when the underlying distributions are symmetric as compared to asymmetric.

The relative efficiency of SDMRSS estimator with respect to those estimators based on SRS, SSRS and SRSS are increasing as the sample size increases.

V. Real Life Example Using SDMRSS

The marks of 787 students from scientific majors in Foundation Program at Qatar University during academic semester Fall 2011 are collected as a population to calculate the mean and variance of them, then to choose 100,000 samples using each of SDMRSS, simple random sample, stratified ranked set sample and stratified simple random sample methods with sample size $n = 7,12,14$ using *Matlab 7*. The mean and variance are obtained for each method and compared to evaluate the performance of SDMRSS to estimate the population mean for real data. Stratification is done according to the gender (males and females) and proportional allocation is considered.

Let $x_{1i}, \quad i = 1,2,\cdots,531$ be the mark of i^{th} female student in the population.

Let $x_{2j}, \quad j = 1,2,\cdots,256$ be the mark of j^{th} male student in the population

The mean μ and the variance σ^2 of the population, the mean μ_1 and the variance σ_1^2 of the female population and the mean μ_2 and the variance σ_2^2 of the male population respectively are

$$\mu = \frac{1}{787}\sum_{i=1}^{787} x_i = 61.98 \text{ and}$$

$$\sigma^2 = \frac{1}{787}\sum_{i=1}^{787}\left(x_i - \mu\right)^2 = 1060.86$$

$$\mu_1 = \frac{1}{531}\sum_{i=1}^{531} x_i = 65.08 \text{ and}$$

$$\sigma_1^2 = \frac{1}{531}\sum_{i=1}^{531}\left(x_i - \mu\right)^2 = 1025.85$$

$$\mu_2 = \frac{1}{256}\sum_{i=1}^{256} x_i = 55.55 \text{ and}$$

$$\sigma_2^2 = \frac{1}{256}\sum_{i=1}^{256}\left(x_i - \mu\right)^2 = 1076.29$$

Skewness and median of 787 students are -0.64 and 71.00. Skewness and median of 531 female students are -0.74 and 75.00. Skewness and median of 256 male students are -0.48 and 65.00.

Since skewness for all students, female students and male students are -0.64, -0.74, -0.48, then the marks are asymmetrically distributed, which means that SDMRSS estimator is biased and hence the mean square error of the estimators will be calculated. The efficiency of SSRS, SRSS and SDMRSS with respect to SRS is computed and summarized in Table IV.

TABLE IV
THE EFFICIENCY SSRS, SRSS AND SDMRSS RELATIVE TO SRS

Sampling Method		$n=7$, $n_1=4$, $n_2=3$	$n=12$, $n_1=5$, $n_2=7$	$n=14$, $n_1=8$, $n_2=6$
SRS	Mean	61.7	61.7	61.5
	Variance	57.3	42.6	38.7
SSRS	Mean	61.5	61.8	62.0
	Variance	32.4	28.7	22.8
	Efficiency	3.76	3.82	3.65
SRSS	Mean	62.3	62.1	61.9
	Variance	19.5	16.9	14.6
	Efficiency	3.98	4.53	4.33
SDMRSS	Mean	61.8	61.7	61.6
	Variance	13.4	11.04	10.8
	Efficiency	4.81	4.97	4.88

From Table IV, the following are noticed:

1. The values of estimated mean using SDMRSS is very close to the population mean.
2. The small differences between the population mean and the estimated mean happened because the marks are asymmetrically distributed bout its mean.
3. The variance of SDMRSS estimator is less than the variances in case of SRS, SSRS and SRSS.
4. The efficiency values using SDMRSS are greater than those obtained using SRS, SSRS and SRSS.
5. Results in this real life example agrees with the theoretical results.

VI. Conclusions

In this paper, we have suggested an estimator of the population mean using SDMRSS. The performance of the estimator based on SDMRSS is compared with those found using SRSS, SSRS and SRS for the same number of measured units. It is found that SDMRSS produces an unbiased estimator of the population mean and more efficient than SRSS, SSRS and SRS. Thus, SDMRSS should be more preferred than SRSS, SSRS and SRS for both symmetric and asymmetric distributions.

Acknowledgment

This publication was made possible because of the support of Qatar University. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Qatar University.

References

[1] Al-Saleh, M.F. and Al-Kadiri, M. (2000). Double ranked set sampling. Statistics Probability Letters, 48, 205 - 212.
[2] Al-Saleh, M.F. and Al-Omari, A.I. (2002). Multistage ranked set sampling. Journal of Statistical Planning and Inference, 102, 273-286.
[3] Bouza, C.N. (2002) Ranked set subsampling the non-response strata for estimating the difference of means. Biometrical Journal, 44:903915.
[4] Dell, T.R., and Clutter, J.L. (1972). Ranked set sampling theory with order statistics background. Biometrika, 28, 545-555.
[5] Jemain, A.A. & Al-Omari, A.I. (2006). Double quartile ranked set samples. Pakistan Journal of Statistics. 22(3): 217-228.
[6] McIntyre, G. A. (1952). A method for unbiased selective sampling using ranked sets. Australian Journal of Agricultural Research. 3, 385–390.
[7] Muttlak, H.A. (1997). Median ranked set sampling, J. App. Statist. Sci. 6(4), 577-586.

[8] Ohyama, T., Doi, J., and Yanagawa, T. (2008). Estimating population characteristics by incorporating prior values in stratified random sampling/ranked set sampling. Journal of Statistical Planning and Inference, 138, 4021-4032.
[9] Samawi, H. M. (1996). Stratified ranked set sample. Vol. 12(1) pp. 9 – 16.
[10] Takahasi K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. Annals of the Institute of Statistical Mathematics, 20, 1–31.