مجلهی پژوهشهای آماری ایران

J. Statist. Res. Iran **9** (2012): 87–102

دورهی ۹، شمارهی ۱، بهار و تابستان ۱۳۹۱، صص ۸۷–۱۰۲

پژوهشکدهی آمار

Stratified Median Ranked Set Sampling: Optimum and Proportional Allocations

Samineh Hajighorbani[†] and Roshanak Aliakbari Saba[‡,*]

[†] Allameh Tabataba'i University

[‡] Statistical Research and Training Centre

Abstract. In this paper, for the Stratified Median Ranked Set Sampling (SMRSS), proposed by Ibrahim et al. (2010), we examine the proportional and optimum sample allocations that are two well-known methods for sample allocation in stratified sampling. We show that the variances of the mean estimators of a symmetric population in SMRSS using optimum and proportional allocations to strata are smaller than the corresponding variances in Stratified Random Sampling (STRS). It is also shown that for a fixed value of sampling cost in strata, the variance of mean estimator with optimum allocation is less than or equal to the variance of mean estimator with proportional allocation in SMRSS. In addition, we develop the results obtained by Ibrahim et al. (2010) for proportional allocation in SMRSS for some symmetric and non-symmetric distributions when the parameters of distributions are varying.

Keywords. Ranked Set Sampling; Stratified Median Ranked Set Sampling; optimum allocation; proportional allocation.

MSC 2010: 62D05.

1 Introduction

Ranked Set Sampling (RSS) was first suggested by McIntyre (1952) to estimate mean pasture and forage yields. Takahasi and Wakimoto (1968) provided the necessary mathematical theory. They proved that the sample mean

[*] Corresponding author

of RSS is an unbiased estimator of the population mean with smaller variance than the sample mean of a Simple Random Sample (SRS) of the same sample size. Dell and Clutter (1972) showed that the mean of the RSS is an unbiased estimator of the population mean, whether or not there are errors in ranking, and is at least as efficient as the SRS estimator with the same number of quantifications. Muttlak (1997) suggested the Median Ranked Set Sampling (MRSS) which reduce the errors in ranking and increase the efficiency over RSS with perfect ranking for some probability distribution functions.

In recent years, RSS has been used in conjunction with other sampling designs, as the last stage of sampling in multi-stage designs, or in stratified sampling designs. For example, Samawi (1996) introduced the concept of Stratified Ranked Set Sampling (SRSS) and Nematollahi et al. (2008) employed Ranked Set Sampling in the second stage of a two-stage cluster sampling design to improve the precision of the population mean estimate.

Ibrahim et al. (2010) suggested Stratified Median Ranked Set Sampling (SMRSS) for estimating the population mean. They indicate that under symmetric distributions, the mean of SMRSS is an unbiased estimator of the population mean with greater efficiency than SRS and STRS estimators.

When stratified sampling designs are to be employed, one key question which has to be immediately addressed is how many observations should be taken in each stratum? The answer to this question is very important, because the sample size is an important feature for making inferences about the population based on sample data.

In this paper, for the SMRSS designs, we examine the proportional and optimum sample allocations. In Section 2, we review some sampling methods considered in this study as well as the estimation of the population mean and its properties. In Section 3, the results of using proportional and optimal sample allocation methods in SMRSS are described. In Section 4, we develop the Ibrahim et al. (2010) results for some symmetric and non-symmetric distributions by a simulation study when the parameters of distributions are varying. Some conclusions are given in Section 5.

2 Sampling Methods

In the following subsections we describe some sampling methods considered in the paper.

2.1 Ranked Set Sampling

To select a sample of size $n_0 = m \times n$ units from a population using RSS method, we proceed as follows:

Step 1. Randomly select a sample of size n^2 units from the population.

Step 2. Allocate the n^2 selected units as randomly as possible into n sets, each of size n.

Step 3. Without knowing any values for the variable of interest, rank the units within each set based on perception of relative values for this variable. This may be based on personal judgment or done with measurements of a covariate that is correlated with the variable of interest.

Step 4. Choose a sample for actual measurement including the smallest ranked unit in the first set, the second smallest ranked unit in the second set, continuing in this fashion until the largest ranked unit is selected in the last set.

Step 5. Repeat steps 1 to 4 for m cycles until desired sample size, $n_0 = m \times n$, is obtained for measurement.

Since in some situations ranking n units for a large sample size is difficult, we select a ranked set sample with small n and then repeat the sampling scheme m times (Step 5). So if the ranking wasn't difficult and would be done with high precision, we can consider $m = 1$ and eliminate the Step 5 from sampling scheme.

Estimation of the Population Mean

Let $X_{ij}^{(i)}$ be the ith $(i = 1, 2, \ldots, n)$ order statistic of ith sample from jth cycle $(j = 1, 2, \ldots, m)$ in RSS. The estimator of population mean (μ) using RSS method and a sample of size $n_0 = m \times n$ with m cycles is given by:

$$\bar{X}_{RSS} = \frac{1}{mn} \sum_{j=1}^{m} \sum_{i=1}^{n} X_{ij}^{(i)}.$$

Variance of this estimator is given by

$$\begin{aligned}\mathrm{Var}(\bar{X}_{RSS}) &= \frac{\sigma^2}{mn} - \frac{1}{mn^2}\sum_{i=1}^{n}(\mu_{(i)} - \mu)^2 \\ &= \frac{1}{mn^2}\sum_{i=1}^{n}\sigma^2_{(i)}\end{aligned}$$

where $\mu_{(i)}$ and $\sigma^2_{(i)}$ are mean and variance of the ith order statistic of a random sample of size n, respectively. See Takahasi and Wakimoto (1968) for more details.

Now if $X_1, X_2, \ldots. X_{mn}$ is a sample of size $n_0 = m \times n$ units from the population selected by using SRS method, the population mean estimator would be $\bar{X}_{SRS} = \frac{1}{mn}\sum_{i=1}^{mn} X_i$ and $\mathrm{Var}(\bar{X}_{SRS}) = \frac{\sigma^2}{mn}$.

Comparing the variance of mean estimators of RSS and SRS methods shows that the precision of RSS is always more than SRS in estimating the population mean based on the same sample size. The preference of RSS is confirmed even if ranking were completely random and provided no information.

2.2 Median Ranked Set Sampling

Median Ranked Set Sampling (MRSS) proposed by Muttlak (1997) is an alternative RSS method that can be used to estimate the population mean instead of RSS. The MRSS can be performed with less error in ranking in practical applications. Since all we have to do is to find the element in the middle of the sample and measure it, the MRSS method can be easily employed in the field and will save some time in performing the ranking of the units with respect to the variable of interest. The MRSS method can be summarized as follows:

Select n random samples of size n units from the population and rank the units within each sample.

From each sample of size n, select the median of sample for measurement. If the sample size n, is odd, from each sample select for measurement the $\left(\frac{n+1}{2}\right)$th smallest rank i.e., the median of the sample. If the sample size is even, select for measurement from the first $\frac{n}{2}$ samples the $\left(\frac{n}{2}\right)$th smallest rank and from the second $\frac{n}{2}$ samples the $\left(\frac{n}{2}+1\right)$th smallest rank. The cycle may be repeated m times if needed to get a sample of size $n_0 = m \times n$ units.

Estimation of the Population Mean

Suppose we select a sample of size n using MRSS from the population of interest, with pdf $f(\cdot)$, mean μ and variance σ^2, repeated in m cycles. The estimator of population mean based on MRSS in the case of an odd sample size, denoted as $\bar{X}_{MRSS1}$ can be defined by

$$\bar{X}_{MRSS1} = \frac{1}{mn} \sum_{j=1}^{m} \sum_{i=1}^{n} X_{ij}^{\left(\frac{n+1}{2}\right)}$$

and in the case of an even sample size, denoted as $\bar{X}_{MRSS2}$, is defined as

$$\bar{X}_{MRSS2} = \frac{1}{mn} \left\{ \sum_{j=1}^{m} \sum_{i=1}^{\frac{n}{2}} X_{ij}^{\left(\frac{n}{2}\right)} + \sum_{j=1}^{m} \sum_{i=\frac{n}{2}+1}^{n} X_{ij}^{\left(\frac{n}{2}+1\right)} \right\}$$

where $X_{ij}^{(r)}$ is the rth order statistic of the ith sample from jth cycle. Muttlak (1997) showed that the variance of $\bar{X}_{MRSS1}$ is given by

$$\text{Var}\left(\bar{X}_{MRSS1}\right) = \frac{1}{mn} \sigma^2_{\left(\frac{n+1}{2}\right)}$$

and when n is even, the variance of the estimator of population mean $\bar{X}_{MRSS2}$ is given by

$$\text{Var}\left(\bar{X}_{MRSS2}\right) = \frac{1}{2mn} \left[\sigma^2_{\left(\frac{n}{2}\right)} + \sigma^2_{\left(\frac{n}{2}+1\right)} \right].$$

Also he showed that, if the population distribution were symmetric about μ the MRSS estimators of the population mean in two cases of an odd and an even sample size are unbiased and the variances of these estimators are less than the variance of SRS estimator. He also showed that using MRSS increases the efficiency of population mean estimator with respect to RSS method with perfect ranking for some probability distribution functions. Symmetric distributions are appearing in some situations in practice, for example consider the distribution of floor area of housing units in a known city that can be a variable with normal distribution if the floor areas of housing units are not so big or so small.

2.3 Stratified Median Ranked Set Sampling

Suppose there are N units in the population, divided into L strata of $N_1, N_2, \ldots, N_L$ units that are not overlapping, so that $N = \sum_{h=1}^{L} N_h$. In a stratified sampling method a sample of size n_h is drawn from hth stratum ($h =$

$1, 2, \ldots, L$) to have a sample of size $n = \sum_{h=1}^{L} n_h$ from the population. If the SRS method is used for sampling in each stratum, we have the Stratified Random Sampling (STRS). In this sampling method

$$\bar{X}_{STRS} = \sum_{h=1}^{L} \frac{W_h}{n_h} \sum_{j=1}^{n_h} X_{hj}$$

is an unbiased estimator of population mean where $W_h = N_h/N$ and X_{hj} is the jth sample unit in the hth stratum.

If the RSS or MRSS is performed in each stratum instead of SRS, the method is known as Stratified Ranked Set Sampling (SRSS) or Stratified Median Ranked Set Sampling (SMRSS) respectively.

Estimation of the Population Mean

Assume in each stratum h, a sample of size $n_{0h} = m_h \times n_h$ units is drawn using MRSS where m_h is the number of cycles and n_h is an odd sample size of each cycle. The estimator of the population mean is given by:

$$\bar{X}_{SMRSS1} = \sum_{h=1}^{L} \frac{W_h}{m_h n_h} \left\{ \sum_{j=1}^{m_h} \sum_{i=1}^{n_h} X_{hij}^{\left(\frac{n_h+1}{2}\right)} \right\}$$

where $X_{hij}^{(r)}$ is the measure of the variable of interest for rth order statistic of ith random sample in jth cycle of hth stratum. The variance of this estimator is

$$\operatorname{Var}\left(\bar{X}_{SMRSS1}\right) = \sum_{h=1}^{L} \frac{W_h^2}{m_h n_h} \sigma^2_{h\left(\frac{n_h+1}{2}\right)} \tag{1}$$

where $\sigma^2_{h(r)}$ denotes the variance of rth order statistic of a random sample of size n_h. If n_h is even, the estimator of population mean using SMRSS is given by

$$\bar{X}_{SMRSS2} = \sum_{h=1}^{L} \frac{W_h}{m_h n_h} \left\{ \sum_{j=1}^{m_h} \sum_{i=1}^{\frac{n_h}{2}} X_{hij}^{\left(\frac{n_h}{2}\right)} + \sum_{j=1}^{m_h} \sum_{i=\frac{n_h}{2}+1}^{n_h} X_{hij}^{\left(\frac{n_h}{2}+1\right)} \right\}$$

with variance given by

$$\operatorname{Var}\left(\bar{X}_{SMRSS2}\right) = \sum_{h=1}^{L} \frac{W_h^2}{2 m_h n_h} \left\{ \sigma^2_{h\left(\frac{n_h}{2}\right)} + \sigma^2_{h\left(\frac{n_h}{2}+1\right)} \right\}. \tag{2}$$

Ibrahim et al. (2010) showed that if the distribution of population is symmetric about μ, then

$$E\left(\bar{X}_{SMRSS1}\right) = E\left(\bar{X}_{SMRSS2}\right) = \mu$$

and

$$\text{Var}\left(\bar{X}_{SMRSS1}\right) \leqslant \text{Var}\left(\bar{X}_{STRS}\right) \leqslant \text{Var}\left(\bar{X}_{SRS}\right),$$
$$\text{Var}\left(\bar{X}_{SMRSS2}\right) \leqslant \text{Var}\left(\bar{X}_{STRS}\right) \leqslant \text{Var}\left(\bar{X}_{SRS}\right).$$

3 Sample Size Allocations

When designing a sample survey on a stratified population, one of the important considerations is how to allocate the total sample size n among the L identified strata. If unit variances or costs of sampling differ among the strata, sampling efficiency can be increased by over-sampling the more variable or cheaper strata. So, the sample designer may decide to minimize the variance of estimation for a specified cost of taking the sample or to minimize the cost for a specified value of variance of the estimation. This allocation method is described as optimum allocation.

If sampling fraction is the same in all strata, the stratification is described as stratification with proportional allocation. This allocation is used to ensure that the distribution of the sample in subpopulations (strata) is proportional to their size.

By this background in mind, optimum and proportional allocations in SMRSS for symmetric distributions are described in this section and a comparison is made between their variances.

3.1 Optimum Allocation

Let c_h, $h = 1, 2, \ldots, L$ be the cost of sampling for a unit in stratum h. The costs can differ substantially between strata. Assume $n_{0h} = m_h n_h$ be the sample size from hth stratum in SMRSS with m_h cycles in hth stratum.

The total cost of the survey can be taken as $C = c_0 + \sum_{h=1}^{L} c_h m_h n_h$. This relation is a linear cost function and c_0 is a fixed cost of sampling. Suppose we want to minimize the variance of the mean estimator for a specified cost C and minimize the cost for a specified variance of the mean estimator. Let us consider the two cases, whether an odd or even number of sample units were selected from each stratum in each cycle.

At the first case, for an odd sample size n_h, we minimize $(C - c_0)$ $\text{Var}\left(\bar{X}_{SMRSS1}\right)$. The Cauchy-Schwarz inequality is used for this minimization. According to this inequality

$$\left(\sum_{h=1}^{L} a_h^2\right)\left(\sum_{h=1}^{L} b_h^2\right) \geqslant \left(\sum_{h=1}^{L} a_h b_h\right)^2$$

and equality holds if and only if $\frac{b_h}{a_h}$ is a constant for all h. This leads to

$$\frac{W_h\sqrt{\frac{\sigma^2_{h\left(\frac{n_h+1}{2}\right)}}{m_h n_h}}}{\sqrt{c_h m_h n_h}} = k$$

where k is a constant. From this equation we have:

$$n_h = \frac{W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{k m_h \sqrt{c_h}}. \tag{3}$$

So the total sample size n from the entire population is given by:

$$n = \sum_{h=1}^{L} m_h n_h = \sum_{h=1}^{L} \frac{W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{k\sqrt{c_h}}. \tag{4}$$

Equations (3) and (4) lead to the following equation.

$$n_h = \frac{\frac{W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{\sqrt{c_h}}}{m_h \sum_{h=1}^{L} \frac{W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{\sqrt{c_h}}} \times n. \tag{5}$$

By substituting the value of n_h of equation (5) in equation (1) we get the variance of optimum allocation of the mean estimator, $\bar{X}_{SMRSS1}$ as:

$$\text{Var}_{opt}\left(\bar{X}_{SMRSS1}\right) = \frac{1}{n}\left\{\sum_{h=1}^{L} W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}\sqrt{c_h}\right\}\left\{\sum_{h=1}^{L} \frac{W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{\sqrt{c_h}}\right\}.$$

In a special case, if the cost per unit is the same in all strata, i.e., $c_h = c$ for all h, then for a given fix sample size n, the n_h can be written as:

$$n_h = \frac{W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{m_h \sum_{h=1}^{L} W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}} \times n = \frac{N_h \sigma_{h\left(\frac{n_h+1}{2}\right)}}{m_h \sum_{h=1}^{L} N_h \sigma_{h\left(\frac{n_h+1}{2}\right)}} \times n$$

and the optimum variance can be summarized as:

$$\text{Var}_{opt}\left(\bar{X}_{SMRSS1}\right) = \frac{1}{n}\left\{\sum_{h=1}^{L} W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}\right\}^2. \tag{6}$$

Now for an even number of n_h, we minimize $(C - c_0)\text{Var}\left(\bar{X}_{SMRSS2}\right)$ by using the Cauchy-Schwarz inequality again, and define:

$$\sigma_h'^2 = \frac{1}{2}\left\{\sigma^2_{h\left(\frac{n_h}{2}\right)} + \sigma^2_{h\left(\frac{n_h}{2}+1\right)}\right\},$$

so, n_h can be written as:

$$n_h = \frac{\frac{W_h \sigma_h'}{\sqrt{c_h}}}{m_h \sum_{h=1}^{L} \frac{W_h \sigma_h'}{\sqrt{c_h}}} \times n. \tag{7}$$

Substituting the n_h of equation (7) in equation (2), the variance of optimum allocation of the mean estimator, $\bar{X}_{SMRSS2}$ is:

$$\text{Var}_{opt}\left(\bar{X}_{SMRSS2}\right) = \frac{1}{n}\left(\sum_{h=1}^{L} W_h \sigma_h' \sqrt{c_h}\right)\left(\sum_{h=1}^{L} \frac{W_h \sigma_h'}{\sqrt{c_h}}\right).$$

In a special case, when $c_h = c$ for all h, for a given fix sample size n, the n_h can be written as:

$$n_h = \frac{W_h \sigma_h'}{m_h \sum_{h=1}^{L} W_h \sigma_h'} \times n = \frac{N_h \sigma_h'}{m_h \sum_{h=1}^{L} N_h \sigma_h'} \times n$$

and the optimum variance can be summarized as:

$$\text{Var}_{opt}\left(\bar{X}_{SMRSS2}\right) = \frac{1}{n}\left(\sum_{h=1}^{L} W_h \sigma_h'\right)^2.$$

3.2 Proportional Allocation

According to the description of proportional sample allocation in SMRSS, the number of sample units in the hth stratum $(h = 1, 2, \ldots, L)$ is proportional to the stratum size, that is

$$\frac{m_h n_h}{n} = \frac{N_h}{N} \Rightarrow n_h = \frac{N_h n}{N m_h}.$$

It can be shown that the variance of proportional allocation for an odd sample size n_h in each cycle from hth stratum is given by:

$$\mathrm{Var}_{prop}\left(\bar{X}_{SMRSS1}\right) = \frac{1}{n}\left\{\sum_{h=1}^{L} W_h \sigma^2_{h\left(\frac{n_h+1}{2}\right)}\right\} \tag{8}$$

and for an even sample size in each cycle from hth stratum, the variance of the proportional allocation is given by:

$$\mathrm{Var}_{prop}\left(\bar{X}_{SMRSS2}\right) = \frac{1}{n}\left(\sum_{h=1}^{L} W_h \sigma'^2_h\right).$$

The following theorem for symmetric distributions compares the variances of mean estimators under STRS and SMRSS both with the optimum and proportional allocations for the same sample size.

Theorem 1. *If distribution of the population is symmetric and the cost per unit c_h is the same in all strata, then the variance of the mean estimator in SMRSS with sample size $n_{0h} = m_h \times n_h$ from hth stratum using optimum and proportional allocations in strata is smaller than the corresponding variance of the mean estimator in STRS.*

See Ibrahim et al. (2010) for a proof of Theorem 1.

The next theorem helps us to compare the efficiency of SMRSS mean estimator under optimum allocation and proportional allocation.

Theorem 2. *In SMRSS, if $c_h = c$ $(h = 1, 2, \ldots, L)$ for all strata, then the variance of mean estimator with optimum allocation is less than or equal to the variance of mean estimator with proportional allocation.*

Proof. We prove the theorem in the case of odd sample size. For even sample size the proof is similar.

Using equations (6) and (8) we have

$$\begin{aligned}\mathrm{Var}_{prop}\left(\bar{X}_{SMRSS1}\right) - \mathrm{Var}_{opt}\left(\bar{X}_{SMRSS1}\right) = \frac{1}{n}\Bigg[&\sum_{h=1}^{L} W_h \sigma^2_{h\left(\frac{n_h+1}{2}\right)} \\ &- \left\{\sum_{h=1}^{L} W_h \sigma_{h\left(\frac{n_h+1}{2}\right)}\right\}^2\Bigg].\end{aligned}$$

Now if we define:

$$\bar{\sigma}^2 = \left\{ \sum_{h=1}^{L} W_h \sigma_{h\left(\frac{n_h+1}{2}\right)} \right\}^2 ,$$

then

$$\begin{aligned}\mathrm{Var}_{prop}\left(\bar{X}_{SMRSS1}\right) - \mathrm{Var}_{opt}\left(\bar{X}_{SMRSS1}\right) &= \frac{1}{n}\left\{ \sum_{h=1}^{L} W_h \sigma^2_{h\left(\frac{n_h+1}{2}\right)} - \bar{\sigma}^2 \right\} \\ &= \frac{1}{n}\sum_{h=1}^{L} W_h \left\{ \sigma_{h\left(\frac{n_h+1}{2}\right)} - \bar{\sigma}^2 \right\}^2 \geqslant 0,\end{aligned}$$

therefore

$$\mathrm{Var}_{opt}\left(\bar{X}_{SMRSS1}\right) \leqslant \mathrm{Var}_{prop}\left(\bar{X}_{SMRSS1}\right),$$

that completes the proof. □

4 Simulation Study

In this section, the efficiency of population mean estimator using SMRSS method is evaluated with respect to corresponding estimators in SRSS and STRS methods by a simulation study. We consider the proportional allocation method for sample allocation, so W_h is omitted from the estimation formulas and with assuming a fixed number of cycles, results can be extended to any population with same number of strata. For this study, three samples are selected with the mentioned sampling methods for 100,000 times and the estimators of population mean are calculated for each selected sample. Then the efficiency of SMRSS is estimated. Sample selection and estimations are done with SAS software. Results are obtained for the Normal and Beta distributions with different parameters and samples with size 12 ($n_1 = 5$ and $n_2 = 7$), 14 ($n_1 = 8$ and $n_2 = 6$) and 18 ($n_1 = 10$ and $n_2 = 8$) for a population with two strata and again 18 ($n_1 = 4$, $n_2 = 6$ and $n_3 = 8$) for a population with three strata, in each replication. Since the true value of population mean, μ is known in each distribution, the efficiency of population mean estimator using SMRSS method is calculated with respect to corresponding estimators in SRSS and STRS methods by

$$\widehat{RE}_1 = \widehat{RE}\left(\hat{\mu}_{SMRSS}|\hat{\mu}_{SRSS}\right) = \frac{\widehat{MSE}\left(\hat{\mu}_{SRSS}\right)}{\widehat{MSE}\left(\hat{\mu}_{SMRSS}\right)} = \frac{\widehat{\mathrm{Var}}\left(\hat{\mu}_{SRSS}\right)}{\widehat{MSE}\left(\hat{\mu}_{SMRSS}\right)},$$

$$\widehat{RE}_2 = \widehat{RE}\left(\hat{\mu}_{SMRSS}|\hat{\mu}_{STRS}\right) = \frac{\widehat{MSE}\left(\hat{\mu}_{STRS}\right)}{\widehat{MSE}\left(\hat{\mu}_{SMRSS}\right)} = \frac{\widehat{\mathrm{Var}}\left(\hat{\mu}_{STRS}\right)}{\widehat{MSE}\left(\hat{\mu}_{SMRSS}\right)}.$$

In the above relations, for each mean estimator, $\hat{\mu}$:

$$\widehat{MSE}(\hat{\mu}) = \widehat{\mathrm{Var}}(\hat{\mu}) + (\bar{\hat{\mu}} - \mu)^2$$

where

$$\widehat{\mathrm{Var}}(\hat{\mu}) = \frac{1}{100000}\sum_{i=1}^{100000}\left(\hat{\mu}_i - \bar{\hat{\mu}}\right)^2,$$

$\hat{\mu}_i$ is the estimate of population mean in the ith replication ($i = 1, 2, \ldots,$ 100000) and $\bar{\hat{\mu}} = \frac{1}{100000}\sum_{i=1}^{100000}\hat{\mu}_i$.

Normal Distribution

The efficiency of population mean estimator using SMRSS method for the Normal distributions with different parameters and different sample sizes are presented in Table 1. As this table shows, the efficiency of population mean estimator using SMRSS is more than corresponding estimators using SRSS and STRS, but the efficiency is a constant value when the mean or variance of Normal distribution is varying. In other words, the efficiency of population mean estimator in Normal distribution is stable when mean or variance of distribution is varying.

Beta Distribution

As it known, the density function of Beta distribution takes different shapes depending on the different values of the parameters. The probability density function of Beta distribution is symmetric when $\alpha = \beta \geqslant 1$, skewed unimodal when $\alpha \neq \beta,\ \alpha > 1,\ \beta > 1$, U-shaped function when $\alpha < 1,\ \beta < 1$, strictly increasing when $\alpha > 1, \beta = 1$, and strictly decreasing function when $\alpha = 1, \beta > 1$. The efficiency of population mean estimator using SMRSS method for the Beta distributions with different parameters and different sample sizes are presented in Tables 2 to 4.

Table 1. Efficiency of population mean estimator using SMRSS with respect to SRSS and STRS in Normal distribution.

Sample Size	Distribution	$\widehat{RE}_1$	$\widehat{RE}_2$
$n=12,\ n_1=5,\ n_2=7$	$N(0,1)$	1.2891	4.1267
	$N(0,4)$	1.2891	4.1267
	$N(2,1)$	1.2891	4.1267
	$N(20,1)$	1.2891	4.1267
$n=14,\ n_1=8,\ n_2=6$	$N(0,1)$	1.2966	4.6791
	$N(0,4)$	1.2966	4.6791
	$N(2,1)$	1.2966	4.6791
	$N(20,1)$	1.2966	4.6791
$n=18,\ n_1=10,\ n_2=8$	$N(0,1)$	1.3625	6.0130
	$N(0,4)$	1.3625	6.0130
	$N(2,1)$	1.3625	6.0130
	$N(20,1)$	1.3625	6.0130
$n=18,\ n_1=4,\ n_2=6,\ n_3=8$	$N(0,1)$	1.2717	4.1085
	$N(0,4)$	1.2717	4.1085
	$N(2,1)$	1.2717	4.1085
	$N(20,1)$	1.2717	4.1085

Table 2. Efficiency of population mean estimator using SMRSS with respect to SRSS and STRS in Beta distribution when $\alpha=\beta\geqslant 1$ or $\alpha\neq\beta,\ \alpha>1,\ \beta>1$.

Sample Size	$\alpha=\beta\geqslant 1$			$\alpha\neq\beta,\ \alpha>1,\ \beta>1$		
	Distribution	$\widehat{RE}_1$	$\widehat{RE}_2$	Distribution	$\widehat{RE}_1$	$\widehat{RE}_2$
$n=12$ $n_1=5,\ n_2=7$	$Beta(1,1)$	0.7637	2.6703	$Beta(5,2)$	0.8039	2.6068
	$Beta(2,2)$	0.9327	3.2279	$Beta(7,4)$	1.0514	3.4710
	$Beta(4,4)$	1.0725	3.6182	$Beta(9,6)$	1.1453	3.7506
	$Beta(6,6)$	1.3245	3.7964	$Beta(11,8)$	1.1744	3.8389
$n=14$ $n_1=8,\ n_2=6$	$Beta(1,1)$	0.7674	3.0455	$Beta(5,2)$	0.7434	2.7629
	$Beta(2,2)$	0.9346	3.6829	$Beta(7,4)$	1.0238	3.8236
	$Beta(4,4)$	1.0889	4.1182	$Beta(9,6)$	1.1378	4.2335
	$Beta(6,6)$	1.1417	4.2894	$Beta(11,8)$	1.1757	4.3554
$n=18$ $n_1=10,\ n_2=8$	$Beta(1,1)$	0.7446	3.7389	$Beta(5,2)$	0.5735	2.5663
	$Beta(2,2)$	0.9442	4.5683	$Beta(7,4)$	0.9608	4.3716
	$Beta(4,4)$	1.1035	5.1965	$Beta(9,6)$	1.1112	5.0695
	$Beta(6,6)$	1.1790	5.4110	$Beta(11,8)$	1.1932	5.3766
$n=18$ $n_1=4,\ n_2=6,\ n_3=8$	$Beta(1,1)$	0.7827	2.7838	$Beta(5,2)$	0.7282	2.3599
	$Beta(2,2)$	0.9428	3.2760	$Beta(7,4)$	1.0092	3.3359
	$Beta(4,4)$	1.0768	3.6341	$Beta(9,6)$	1.1118	3.7157
	$Beta(6,6)$	1.1255	3.7591	$Beta(11,8)$	1.1504	3.7740

Table 3. Efficiency of population mean estimator using SMRSS with respect to SRSS and STRS in Beta distribution when $\alpha = 1,\ \beta > 1$ or $\alpha > 1,\ \beta = 1$.

Sample Size	$\alpha = 1,\ \beta > 1$			$\alpha > 1,\ \beta = 1$		
	Distribution	$\widehat{RE}_1$	$\widehat{RE}_2$	Distribution	$\widehat{RE}_1$	$\widehat{RE}_2$
	$Beta(1,2)$	0.6241	2.0632	$Beta(2,1)$	0.6217	2.0632
$n = 12$	$Beta(1,4)$	0.5170	1.5486	$Beta(4,1)$	0.5157	1.5485
$n_1 = 5,\ n_2 = 7$	$Beta(1,6)$	0.4972	1.4144	$Beta(6,1)$	0.4965	1.4144
	$Beta(1,10)$	0.4900	1.3257	$Beta(10,1)$	0.4899	1.3257
	$Beta(1,2)$	0.5554	2.1061	$Beta(2,1)$	0.5648	2.1084
$n = 14$	$Beta(1,4)$	0.4344	1.4713	$Beta(4,1)$	0.4415	1.4733
$n_1 = 8,\ n_2 = 6$	$Beta(1,6)$	0.4131	1.3187	$Beta(6,1)$	0.4197	1.3204
	$Beta(1,10)$	0.4046	1.2197	$Beta(10,1)$	0.4112	1.2212
	$Beta(1,2)$	0.4040	1.9130	$Beta(2,1)$	0.4058	1.9131
$n = 18$	$Beta(1,4)$	0.2787	1.1630	$Beta(4,1)$	0.2803	1.1628
$n_1 = 10,\ n_2 = 8$	$Beta(1,6)$	0.2596	1.0123	$Beta(6,1)$	0.2612	1.0120
	$Beta(1,10)$	0.2520	0.9193	$Beta(10,1)$	0.2536	0.9190
	$Beta(1,2)$	0.5620	1.8999	$Beta(2,1)$	0.5627	1.8964
$n = 18$	$Beta(1,4)$	0.4282	1.3106	$Beta(4,1)$	0.4293	1.3079
$n_1 = 4,\ n_2 = 6,\ n_3 = 8$	$Beta(1,6)$	0.4026	1.1689	$Beta(6,1)$	0.4039	1.1666
	$Beta(1,10)$	0.3903	1.0768	$Beta(10,1)$	0.3918	1.0747

As Table 2 shows, the efficiency of population mean estimator using SMRSS is more than corresponding estimator using SRSS for large values of α and β. This superiority is increasing when α and β is increased, the case that leads to lower amount of variance. The efficiency of population mean estimator using SMRSS always is more than corresponding estimator using STRS.

Table 3 shows that the efficiency of population mean estimator using SMRSS is more than corresponding estimator using STRS, but this estimator is not efficient than mean estimator using SRSS. It should be noted that in both cases, the relative efficiencies are similar because the shape of distributions are the same, of course in opposite directions.

The efficiency of mean estimator using SMRSS is less than corresponding estimator using SRSS in Beta distribution when both α and β is less than 1. This has been shown in Table 4. This table also shows that the efficiency of mean estimator using SMRSS with respect to STRS corresponding estimator is less than 1 when both α and β is very small, the case that is not very pleasant in practice.

Table 4. Efficiency of population mean estimator using SMRSS with respect to SRSS and STRS in Beta distribution when $\alpha < 1,\ \beta < 1$.

Sample Size	Distribution	$\widehat{RE}_1$	$\widehat{RE}_2$
$n = 12,\ n_1 = 5,\ n_2 = 7$	$Beta(0.2, 0.1)$	0.2348	0.6260
	$Beta(0.4, 0.3)$	0.4597	1.4972
	$Beta(0.6, 0.5)$	0.5863	2.0161
	$Beta(0.8, 0.7)$	0.6749	2.3857
$n = 14,\ n_1 = 8,\ n_2 = 6$	$Beta(0.2, 0.1)$	0.1978	0.5801
	$Beta(0.4, 0.3)$	0.4299	1.5740
	$Beta(0.6, 0.5)$	0.5735	2.2467
	$Beta(0.8, 0.7)$	0.6698	2.6622
$n = 18,\ n_1 = 10,\ n_2 = 8$	$Beta(0.2, 0.1)$	0.1068	0.3758
	$Beta(0.4, 0.3)$	0.3174	1.4366
	$Beta(0.6, 0.5)$	0.4996	2.4045
	$Beta(0.8, 0.7)$	0.6173	3.0456
$n = 18,\ n_1 = 4,\ n_2 = 6,\ n_3 = 8$	$Beta(0.2, 0.1)$	0.2088	0.5597
	$Beta(0.4, 0.3)$	0.4578	1.4837
	$Beta(0.6, 0.5)$	0.6021	2.0732
	$Beta(0.8, 0.7)$	0.6996	2.4223

5 Conclusions

In this paper, we used proportional and optimum allocations to determine sample size in each stratum for SMRSS under symmetric or nonsymmetric distribution of population. Then, some results were obtained as follows:

1. If the distribution of population is symmetric about μ, then the variance of the mean estimator in SMRSS with sample size $n_{0h} = m_h \times n_h$ from hth stratum, using optimum and proportional allocations in strata, is smaller than the corresponding variance of the mean estimator in STRS. This is an especial case of the study done by Ibrahim et al. (2010).

2. If $c_h = c (h = 1, 2, \ldots, L)$ for all strata, the variance of mean estimator with optimum allocation is less than or equal to the variance of mean estimator with proportional allocation in SMRSS.

3. For Normal distribution, by allocating samples to strata proportionally, the efficiency of mean estimator in SMRSS with respect to corresponding estimators in SRSS and STRS is a constant value when the mean

or variance of Normal distribution is varying.

4. For Beta distribution, by allocating samples to strata proportionally, the efficiency of mean estimator in SMRSS with respect to corresponding estimators in SRSS and STRS is a function of variance of distribution. If the density of distribution is a strictly monotone function, efficiency of SMRSS is an increasing function of variance and if the density of distribution is symmetric, skewed unimodal or U-shaped function, efficiency of SMRSS is a decreasing function of variance.

References

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

Ibrahim, K., Syam, M. and Ibrahim Al-omari, A. (2010). Estimating the population mean using stratified median ranked set sampling. *Applied Mathematical Science*, **47**, 2341-2354.

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

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

Nematollahi, N., Salehi, M.M. and Saba A.R. (2008). Two-stage cluster sampling with ranked set sampling in the secondary sampling frame. *Commun. Statist. Theory Meth.*, **37**, 2404-2415.

Samawi, H.M. (1996). Stratified ranked set sample. *Pakistan Journal of Statistics*, **12**, 9-16.

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*, **21**, 249-255.

Samineh Hajighorbani
Department of Statistics,
Allameh Tabataba'i University,
Tehran, Iran.

Roshanak Aliakbari Saba
Statistical Research and
Training Centre,
Tehran, Iran.
email: *r_saba@srtc.ac.ir*