J. Ind. Soc. Agril. Statist.
62(3), 2008 : 203-207

Estimation of Finite Population Mean using Double Ranked Set Sampling

U.C. Sud and Dwijesh Chandra Mishra
Indian Agricultural Statistics Research Institute, New Delhi
(Received : October 2006)

SUMMARY

The Double Ranked Set Sampling (DRSS) procedure has been extended to the case of estimation of finite population mean under classical approach of survey sampling. In view of the complexity of the theoretical framework, the sample sizes are restricted to '2'. Using real data, it is empirically demonstrated that an estimator based on DRSS procedure performs better than estimators based on the Ranked Set Sampling (RSS) procedure and Simple Random Sampling (SRS) respectively.

Key words: Double Ranked Set Sampling, Ranked Set Sampling, Finite population sampling, Simple random sampling.

INTRODUCTION

In statistical settings where actual measurements of the sample observations are difficult or costly or time consuming or destructive etc. but acquisition and subsequent ranking of the potential sample data is relatively easy, improved methods of statistical inference can result from using Ranked Set Sampling (RSS) technique.

The method of RSS was first introduced by McIntyre (1952) to improve upon simple random sampling for situations where some preliminary ranking of sampled units is possible. The basic idea was to randomly partition the sampled units into small groups and each member of the individual group is ranked relative to other members of the group. Based on this ranking one member from each group is selected for quantification.

The method involves selecting 'm' samples, each of size m, and ordering each of the samples by eye or by some other relatively inexpensive means, without actual measurement on the individuals. Then the smallest ordered observation from the 1st sample is accurately measured, as is the 2nd smallest observation from the 2nd. The process is continued until the largest observation from the mth sample is measured. This constitutes one cycle. The entire cycle is repeated n times, (because accurate judgment ordering of a large number of observations would be difficult in most experimental situations, therefore an increase in sample size is typically implemented by increasing n, the number of replication of cycle, rather than m) until altogether N = mn observations have been quantified.

The method of Double Ranked Set Sampling for estimation of parameter of interest was developed by Al-Saleh and Al-Kadiri (2000). The method involves following steps

- Identify m^3 elements from the target population and divide these elements randomly into m sets each of size m^2.
- Use the usual RSS procedure on each set to obtain m ranked set samples of size m each. Apply the RSS procedure again on the m ranked set samples to obtain a DRSS of size m.

It was shown that an estimator of population mean based on DRSS was more efficient than an estimator based on RSS.

However, they developed the method under the assumption that the population under investigation follows some distribution. In this paper we propose to develop a DRSS based estimator under a finite population sampling frame work. Due to complexity of the theoretical frame work, we limit ourselves to the case of sample of size '2'.

2. THEORETICAL FRAMEWORK OF DRSS

Let there be a finite population of size $m^3 = 8$. Let the finite population unit values be given by $x_1, x_2, ..., x_8$. Without any loss of generality, let the following hold

$$x_1 \leq x_2 \leq ... \leq x_8$$

Let the finite population mean be defined by

$$\overline{X} = \frac{1}{8}\sum_{i=1}^{8} x_i$$

Let a sample of size '2' be selected from the population by the procedure of DRSS. This implies that the eight units are randomly divided into two groups of size '4' each. The units in the subgroups of size 4 are randomly arranged into 2 rows of 2 units each. The RSS procedure is applied independently on the two randomly formed groups. The RSS procedure is applied a second time on the resultant samples from the two groups which in turn gives a sample of size '2'. Let the units selected by DRSS be denoted by x_1 and x_2 respectively.

To estimate $\overline{X}$, we propose to develop a Horvitz Thomson type estimator. To be able to do so working out expressions for inclusion probabilities of units of the population is a prerequisite. In what follows, we develop expressions for the inclusion and joint inclusion probabilities of the population units under DRSS framework.

Inclusion Probability of a Unit

Let the s-th ranked unit in the population of 8 units has the i-th rank in any of the two groups of size '4'. We consider all possibilities where by the s-th ranked unit in the population has the i-th rank.

1. s-th ranked unit in the population having i-th rank is selected in 1st row of the first group as first rank. This probability is given by

$$\frac{\left[\binom{s-1}{i-1}\binom{m^3-s}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m} + \binom{s-1}{i-1}\binom{m^3-s}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\right]}{\binom{m^3}{m^2}\binom{m^2}{m}} \tag{2.1}$$

2. s-th ranked unit in the population is selected in the second row of first group and i selected as first rank. This probability is given by

$$\frac{\binom{s-1}{i-1}\binom{m^3-s}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}}{\binom{m^3}{m^2}\binom{m^2}{m}} \tag{2.2}$$

3. s-th ranked unit in the population is selected in the second group as second rank but i selected in the first row, the probability is given by

$$\frac{\binom{s-1}{i-1}\binom{m^3-s}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{2}\binom{m^2-m-i+1}{m-2}}{\binom{m^3}{m^2}\binom{m^2}{m}} \tag{2.3}$$

4. s-th ranked unit in the population is selected in the second group as second rank but i-th ranked unit is selected in the second row, the probability is given by

$$\frac{\left[\binom{s-1}{i-1}\binom{m^3-s}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m-1}\binom{i-2}{1}\binom{m^2-m-i+1}{m-2} + \binom{s-1}{i-1}\binom{m^3-s}{m^2-i}\binom{i-1}{2}\binom{m^2-i}{m-2}\binom{i-3}{1}\binom{m^2-m-i+2}{m-2}\right]}{\binom{m^3}{m^2}\binom{m^2}{m}} \tag{2.4}$$

Since the four cases are mutually exclusive the inclusion probability is the sum of probabilities of four cases.

Using hypothetical data, it was seen that the inclusion probability of a given unit is '1/4'. This was also verified by substituting the values m^3, m^2, s and i (i = 1, 2, 3, 4) in the expression for inclusion probability of a unit developed above.

Thus, it can be concluded that $\pi_1 = \pi_2 = ... = \pi_8 = \frac{1}{4}$ = constant.

Joint Inclusion Probabilities of Units

To develop expression for π_{st} (t > s) we proceed as follows. Let the s-th and t-th ranked units in the population have i-th and j-th ranks respectively in the sub-groups of size '4' formed randomly. The sample space here comprises $\binom{8}{4}\frac{4!}{2!2!}\times\frac{4!}{2!2!}=2520$ samples.

This is due to the fact that 4 units out of 8 can be selected in $\binom{8}{4}$ ways and the remaining 4 units can be selected in $\binom{4}{4}$ ways. Now each of the 4 units can be arranged in 2 rows, each row can have '2' units in $\frac{4!}{2!2!}$ ways. This way the total number of samples work out to $\binom{8}{4}\binom{4}{4}\times\frac{4!}{2!2!}\times\frac{4!}{2!2!}$.

We consider all the possible ways in which s-th and t-th ranked units have i-th and j-th ranks respectively in the sub-groups of size '4' each.

Case 1. s selected in the first group as i-th rank and t selected in the second group as j-th rank, s-th ranked units in the population having i-th rank is selected in 1st row of the first group and t-th ranked unit in the population having j-th rank is selected in 1st row of the second group. The number of ways this can happen is

$$\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m}\times$$
$$\binom{t-1-i-\lambda}{j-1}\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{0}\binom{m^2-j}{m-1}\binom{j-1}{2}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\times$$
$$\binom{t-1-i-\lambda}{j-1}\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{0}\binom{m^2-j}{m-1}\binom{j-1}{2}\binom{m^2-m-j+1}{m-2}$$

s-th ranked unit in the population having i-th rank is selected in 2nd row of the first group and t-th ranked unit in the population having j-th rank is selected in 1st row of the second group. The number of ways this can happen is

$$\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}\binom{t-1-i-\lambda}{j-1}\times$$
$$\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{0}\binom{m^2-j}{m-1}\binom{j-1}{2}\binom{m^2-m-j+1}{m-2}$$

s-th ranked unit in the population having i-th rank is selected in 1st row of the first group and t-th ranked unit in the population having j-th rank is selected in the 2nd row of the second group. The number of ways this can be possible is

$$\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m}\binom{t-1-i-\lambda}{j-1}\times$$
$$\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{1}\binom{m^2-j}{m-1}\binom{j-2}{1}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m}\binom{t-1-i-\lambda}{j-1}\times$$
$$\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{2}\binom{m^2-j}{m-2}\binom{j-3}{1}\binom{m^2-m-j+2}{m-2}$$
$$+\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\times$$
$$\binom{t-1-i-\lambda}{j-1}\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{1}\binom{m^2-j}{m-1}\binom{j-2}{1}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i-1}\binom{t-1-s}{\lambda}\binom{m^3-t}{m^2-i-\lambda}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\times$$
$$\binom{t-1-i-\lambda}{j-1}\binom{m^3-m^2-t+i+\lambda}{m^2-j}\binom{j-1}{2}\binom{m^2-j}{m-2}\binom{j-3}{1}\binom{m^2-m-j+2}{m-2}$$

s-th ranked unit in the population having i-th rank is selected in 2nd row of the first group and t-th ranked unit in the population having j-th rank is selected in 2nd row of the second group. The number of ways this can happen is

$$\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{1}\binom{m^2-j}{m-1}\binom{j-2}{1}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{2}\binom{m^2-j}{m-2}\binom{j-3}{1}\binom{m^2-m-j+2}{m-2}$$

Case 2. s-th selected unit in the population is selected in the second group as i-th rank and t selected in the first group as j-th rank s-th ranked unit in the population having i-th rank is selected in 1st row of the second group and t-th ranked unit in the population having j-th rank is selected in 1st row of the first group. The number of ways this can happen is

$$\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{0}\binom{m^2-j}{m-1}\binom{j-1}{2}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\times$$
$$\binom{s-i_1-1}{j-1}\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{0}\binom{m^2-j}{m-1}\binom{j-1}{2}\binom{m^2-m-j+1}{m-2}$$

s-th ranked unit in the population having i-th rank is selected in 2nd row of the second group and t-th ranked unit in the population having j-th rank is selected in 1st row of the first group. This number can happen in following number of ways

$$\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{0}\binom{m^2-j}{m-1}\binom{j-1}{2}\binom{m^2-m-j+1}{m-2}$$

s-th ranked unit in the population having i-th rank is selected in 1st row of the second group and t-th ranked unit in the population having j-th rank is selected in 2nd row of the first group. The number of ways this is possible is

$$\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{1}\binom{m^2-j}{m-1}\binom{j-2}{1}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{m^2-m-i+1}{m}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{2}\binom{m^2-j}{m-2}\binom{j-3}{1}\binom{m^2-m-j+2}{m-2}$$
$$+\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\times$$
$$\binom{s-i_1-1}{j-1}\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{1}\binom{m^2-j}{m-1}\binom{j-2}{1}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{0}\binom{m^2-i}{m-1}\binom{i-1}{1}\binom{m^2-m-i+1}{m-1}\times$$
$$\binom{s-i_1-1}{j-1}\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{2}\binom{m^2-j}{m-2}\binom{j-3}{1}\binom{m^2-m-j+2}{m-2}$$

s-th ranked unit in the population having i-th rank is selected in 2nd row of the second group and t-th ranked unit in the population having j-th rank is selected in 2nd row of the first group. The number of ways this can happen is

$$\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{1}\binom{m^2-j}{m-1}\binom{j-2}{1}\binom{m^2-m-j+1}{m-2}$$
$$+\binom{s-1}{i_1}\binom{t-1-s}{i-i_1-1}\binom{m^3-t}{m^2-i}\binom{i-1}{1}\binom{m^2-i}{m}\binom{m^2-m-i}{m-2}\binom{s-i_1-1}{j-1}\times$$
$$\binom{m^3-m^2-s+i_1}{m^2-j}\binom{j-1}{2}\binom{m^2-j}{m-2}\binom{j-3}{1}\binom{m^2-m-j+2}{m-2}$$

Estimator of Population Mean

The proposed estimator of population mean is given by

$$\hat{\bar{X}}_{HT} = \sum_{i=1}^{2}\frac{x_i}{N\pi_i} \qquad (2.5)$$

Where the expression for π_i has been developed earlier and it is shown to be equal to $\frac{1}{4}$.

The variance of this estimator is given by

$$V(\hat{\bar{X}}_{HT}) = \sum_{i=1}^{8}\sum_{j>i}^{8}(\pi_i\pi_j-\pi_{ij})\left(\frac{x_i}{N\pi_i}-\frac{x_j}{N\pi_j}\right)^2 \qquad (2.6)$$

3. EMPIRICAL STUDY

For the purpose of comparing the DRSS and RSS/SRS based estimators, an empirical study was carried out wherein a part the data of wheat crop for an experimental station as given in Singh *et al*. (1979) was taken for the estimation of population mean. The data comprised of 2 fields, each field having 4 plots. For DRSS protocol, plots in each field were ranked according to the perceived weight of wheat yield and '2' plots were selected by the method of DRSS. Three different estimators based on a sample of size '2' were considered i.e. DRSS based estimator, RSS and SRS based estimators respectively. The results of analysis of data are presented in Table 3.1.

Table 3.1. Relative efficiency of estimators of population mean based on DRSS and RSS over an estimator based on SRS

Case	Design	SE	Relative Efficiency
1	DRSS	0.192	235.99
2	RSS	0.205	205.16
3	SRS	0.295	

From Table 3.1 it can be seen that an estimator of population mean based on DRSS is quite precise as compared to an RSS or SRS based estimators. The relative efficiency of DRSS based estimator over SRS based estimator is to the tune of 236 percent while the corresponding figure over RSS based estimator is to the tune of 205.

Based on the results of this study it can be concluded that even under finite population framework the DRSS estimator scores over the RSS or SRS based estimator.

Although general results concerning ranked set sampling (RSS) in the context of finite population sampling are available, it is not the case in the context of DRSS. This is attributable to the fact that the DRSS procedure is far more complicated than the RSS. Due to this the DRSS procedure results in the context of finite population sampling are developed for the case of sample of size ‘2’. Although restriction to a sample of size ‘2’ limits the utility of the results in some ways, there are still situations where the results obtained in this paper can be gainfully used. Thus, the results obtained in this paper can be utilized for multistage sampling designs where units at the first stage can be selected by DRSS while at the subsequent stages the units can be selected by either RSS or SRS. This can result in considerably increased precision as the psu’s contribution to the total variance of the estimator is substantial. On the same analogy in a stratified sampling scenario, units in different strata can be selected by DRSS. Further, if the population size is not a multiple of m^3, a sample of size m^3, say by SRS, can be drawn from the population and the DRSS procedure can be applied on the resultant sample. However, the properties of the estimator in this cases will be required to be studied under all possible samples, both DRSS and SRS.

Implementing DRSS involves ranking of units. Ranking will become cumbersome as the number of units increase. When the population size is large, DRSS can be implemented by dividing the population into smaller groups and drawing samples independently from different groups. The estimates from different groups can then be combined to get overall estimates. It is also worthwhile to point out that when the samples are drawn by DRSS, π_{ij}’s will be ‘0’ for some of the pairs of units. Thus, it is not possible to unbiasedly estimate the variance of the estimator. There is therefore a need to develop suitable approximately unbiased estimators. Efforts are being made in this direction.

REFERENCES

Al-Saleh, M.F. and Al-Kadiri, M.A. (2000). Double ranked set sampling. *Statist. Probab. Letters*, **48,** 205-212.

McIntyre, G.A. (1952). A method of unbiased selective sampling using ranked sets. *Austr. J. Agric. Res.*, **3,** 385-390.

Horvitz, D.G. and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.*, **47**.

Singh, D., Singh, P. and Kumar, P. (1979). *Handbook on Sampling Methods*. Indian Agricultural Statistics Research Institute, New Delhi.