Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 67(3) 2013 345-353

Variance Estimation using Jackknife Method in Ranked Set Sampling under Finite Population Framework

Ankur Biswas, Tauqueer Ahmad and Anil Rai
Indian Agricultural Statistics Research Institute, New Delhi

Received 21 October 2011; Revised 22 May 2013; Accepted 03 July 2013

SUMMARY

When measuring an observation is expensive, but ranking a small subset of observations is relatively easy, ranked set sampling (RSS) can be used to increase the precision of the estimators. Estimating the variance in case of RSS has been found to be cumbersome in the context of finite population. Therefore, in this paper, we propose two different variance estimation procedures using Jackknife method in RSS under finite population framework. We compare the efficiency of these proposed variance estimation procedures with each other through a simulation study.

Keywords: Variance estimation, Ranked set sampling, Jackknife method, Strata based approach, Cycle based approach.

1. INTRODUCTION

Ranked set sampling (RSS) is a method of sampling that provides a more precise estimator of population mean than simple random sampling (SRS) when actual measurements are either difficult, time consuming or expensive in terms of time, money or labour and ranking on the basis of visual inspection or any other rough method, not requiring actual measurement, is relatively easy. The method of RSS was first introduced by McIntyre (1952) to improve upon SRS for situations where some preliminary ranking of sampled units is feasible. In situations where visual inspection is not directly available, ranking can sometimes be done on the basis of a covariate that is more accessible requiring less costs than, but correlated with, the character of interest. Thus, if we are interested in the volumes of trees, we may use the ranking by diameter to approximate the ranking by volume. This procedure is called as ranking using concomitant variables. This was first discussed by Stokes (1977) and referred it as "ranked set sampling with concomitant variables". Patil *et al.* (1994) have discussed various aspects of RSS in detail. Chen *et al*. (2004) discussed RSS including the method of bootstrapping from ranked set samples.

1.1 RSS in the Context of Finite Population

The majority of research in RSS has been concerned with estimating the mean. Estimation of variance of an estimator from ranked set sample has received less attention. Most of the works which have been done in the area of RSS are in the context of infinite population. However, Patil *et al.* (1995) discussed the methods of RSS without replacement in the context of finite population sampling and provided unbiased estimator of population mean and its variance expression. The expression of variance of the estimator is not quite simple. Moreover, the expression of estimator of variance has not been provided. Krishna (2002), Sud and Mishra (2006, 2007) and Kankure and Rai (2008) made an attempt to extend the theory of RSS

Corresponding author: Tauqueer Ahmad
E-mail address: tauqueer@iasri.res.in

when the sampling is done from a finite population. Modarres *et al.* (2006) proposed two different bootstrap methods for ranked set samples.

The RSS procedure involves randomly drawing m^2 units by simple random sampling without replacement (SRSWOR) from a population of size N with mean μ and a finite variance σ^2. Then, m^2 units are randomly partitioned into m equal-sized sets with set size m. The units within each set are then ranked on the basis of some auxiliary character which may or may not be quantifiable. The unit receiving the smallest rank is accurately quantified from the first set, the unit receiving the 2^{nd} smallest rank is accurately quantified from the 2^{nd} set and so forth, until the unit with largest rank is accurately quantified from the m^{th} set. This constitutes one cycle. The entire cycle is replicated r times until altogether $n = mr$ observations are quantified out of m^2r originally selected units. These n quantified units constitute the ranked set sample.

Patil *et al.* (1995) showed that the probability of the q^{th} ranked unit in the population has the i^{th} rank in any of the disjoint SRSWOR subset of size m is given by

$$\mathrm{P}\left[\, y_{(i:m)} = \mathrm{Y_q} \,\right] = A_i^q = \binom{q-1}{i-1}\binom{N-q}{m-i}\Big/\binom{N}{m},$$

$$\forall\, i = 1, 2, ..., m \text{ and } q = 1, 2, ...\, N \quad (1)$$

and the probability of the q^{th} ranked unit in the population has the i^{th} rank in the disjoint subset 1 and the t^{th} ranked unit in the population has the j^{th} rank in the disjoint subset 2 as

$$\mathrm{P}\left[\, y_{(i:m)} = Y_q, y_{(j:m)} = Y_t \,\right] = B_{ij}^{qt}$$

$$= \sum_{\lambda=0}^{m-i} \frac{\binom{q-1}{i-1}\binom{t-q-1}{\lambda}\binom{N-t}{m-i-\lambda}\binom{t-1-i-\lambda}{j-1}\binom{N-t-m+i+\lambda}{m-j}}{\binom{N}{m,m}},$$

$$\forall\, i, j = 1, 2, ..., m \text{ and } \mathrm{q} < \mathrm{t} \quad (2)$$

where multinomial coefficient

$$\binom{N}{m,m} = \frac{N!}{m!m!(N-2m)!}.$$

If $q = t$ then $B_{ij}^{qt} = 0$ and if $q > t$ then $B_{ij}^{qt} = B_{ji}^{tq}$.

Patil *et al.* (1995) showed that the ranked set estimator *i.e.* sample mean is unbiased estimator of the population mean, $\mu = \overline{Y}$, which is given by

$$\overline{y}_{RSS} = \frac{1}{mr}\sum_{k=1}^{r}\sum_{i=1}^{m} y_{(i:m)k} \quad (3)$$

Variance of the ranked set estimator, $\overline{y}_{RSS}$, is given by

$$V(\overline{y}_{RSS}) = \frac{1}{mr}\left\{\frac{N-1-mr}{N-1}\sigma^2 - \overline{\gamma}\right\} \quad (4)$$

where

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \overline{Y})^2, \quad \overline{\gamma} = \frac{m!(m-1)!}{N(N-1)...(N-2m+1)}\gamma,$$

$\gamma = (\underline{Y} - \overline{Y})'\underset{\sim}{\Gamma}(\underline{Y} - \overline{Y})$ and $\underset{\sim}{\Gamma} = \binom{N}{m,m}\sum_{i=1}^{m} B_{ii}$ is a symmetric matrix with zeroes on the diagonal and a function of N and m only and it does not depend on population values $\underline{Y}$.

Patil *et al.* (1995) obtained an unbiased estimator of population mean and its variance expression. But, estimating the variance of the estimator from an observed sample in case of RSS under finite population context has been found to be cumbersome as some π_i (probability of inclusion of i^{th} unit in the sample) and π_{ij} (inclusion probability for i^{th} and j^{th} units together in the sample) come out to be zero.

The Jackknife method of variance estimation is one of the popular resampling procedures available in the literature. Therefore, we propose two different variance estimation procedures using Jackknife method under ranked set sampling framework in the context of finite population in section 2.

1.2 Jackknife Method of Variance Estimation

The Jackknife method is also known as the Quenouille-Tukey Jackknife, since this tool was invented by Quenouille (1949) and later developed by Tukey (1958). The Jackknife and Bootstrap resampling procedures have been discussed in detail by Shao and Tu (1995). Wolter's book (1985) serves as a key source of reference for methods of variance estimation.

Let $\hat{\theta}$ be an estimator of θ defined for a sample $\underline{X} = (X_1, ..., X_n)$. The complete sample is to be partitioned into k groups of m observations such that $n = mk$. Let $\hat{\theta}_{(\alpha)}$ be the estimator of the same functional form as $\hat{\theta}$, but computed from the reduced sample of size $m(k-1)$ obtained by omitting the α^{th} group and define the 'pseudovalues' as

$$\hat{\theta}_\alpha = k\hat{\theta} - (k-1)\hat{\theta}_{(\alpha)} \tag{5}$$

Quenouille's unbiased estimator of θ is the mean of the pseudovalues, given by

$$\hat{\hat{\theta}} = \sum_{\alpha=1}^{k} \hat{\theta}_\alpha / k \tag{6}$$

The Jackknife estimator of variance is given by

$$v_1\left(\hat{\hat{\theta}}\right) = \frac{1}{k(k-1)} \sum_{\alpha=1}^{k} \left(\hat{\theta}_\alpha - \hat{\hat{\theta}}\right)^2$$

The Jackknife estimator of variance can alternatively be given by

$$v_1\left(\hat{\hat{\theta}}\right) = \frac{k-1}{k}\left[\sum_{\alpha=1}^{k} \hat{\theta}_{(\alpha)}^2 - k\hat{\theta}^2\right] \tag{7}$$

In the context of finite population, the Quenouille's estimator, $\hat{\hat{\theta}}$, has to satisfy two important Lemmas (Wolter 1985). These are as follows:

(a) $\hat{\hat{\theta}} = \hat{\theta}$

(b) $E\left(v_1(\hat{\hat{\theta}})\right) = V\{\hat{\hat{\theta}}\} = V\{\hat{\theta}\}$.

In the context of finite population, Lemma (a) ensures the unbiasedness of the Quenouille's estimator, $\hat{\hat{\theta}}$ for the parameter of interest θ and Lemma (b) ensures the unbiasedness of the Jackknife estimator of variance, $v_1(\hat{\hat{\theta}})$, for $V\{\hat{\theta}\}$.

2. PROPOSED METHODOLOGY FOR VARIANCE ESTIMATION IN RSS

We propose the following two different approaches in order to develop the variance estimation procedures using Jackknife method in ranked set sampling:

1. Cycle based approach
2. Strata based approach

2.1 Cycle Based Approach

We propose cycle based approach for variance estimation using Jackknife procedure in the context of RSS. Under this approach, in the sample of mr observations, there are r cycles considered as r groups each of size m. For applying usual Jackknife procedure, it is proposed to drop one complete cycle and obtain the estimator from the reduced sample. Then, drop another cycle and proceed in the same fashion until all the cycles are dropped once. Here, dropping of m units at a time is repeated r times. In this way, one unit is dropped from each rank in the process of dropping one cycle.

Let $\hat{\mu}$ be the original estimator of population mean, μ. Again, let $\hat{\mu}_{(\alpha)}$ be the estimator of the same functional form as $\hat{\mu}$, but computed from the reduced sample of size $m(r-1)$ after deleting α^{th} group or cycle. The expressions of $\hat{\mu}$ and $\hat{\mu}_{(\alpha)}$ are given by

$$\hat{\mu} = \bar{y}_{RSS} = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{r}\sum_{k=1}^{r} y_{(i:m)k} \text{ and} \tag{8}$$

$$\hat{\mu}_{(\alpha)} = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{r-1}\sum_{\substack{k=1\\(k\neq\alpha)}}^{r} y_{(i:m)k}, \quad \forall \alpha = 1,2,...,r. \tag{9}$$

The pseudovalues are given by

$$\hat{\mu}_\alpha = r\hat{\mu} - (r-1)\hat{\mu}_{(\alpha)}, \tag{10}$$

The Quenouille's estimator of μ is given by

$$\hat{\hat{\mu}} = \frac{1}{r}\sum_{\alpha=1}^{r} \hat{\mu}_\alpha \tag{11}$$

and the Jackknife estimator of variance of the estimator $\hat{\hat{\mu}}$ is given by

$$v_1\left(\hat{\hat{\mu}}\right) = \frac{1}{r(r-1)}\sum_{\alpha=1}^{r}\left(\hat{\mu}_\alpha - \hat{\hat{\mu}}\right)^2. \tag{12}$$

By Lemma (a) of section 1.2, it follows that, $\hat{\hat{\mu}} = \hat{\mu} = \bar{y}_{RSS}$ and thus $\hat{\hat{\mu}}$ is unbiased estimator of population mean, μ.

The Jackknife estimator of variance can alternatively be given by

$$v_1(\hat{\hat{\mu}}) = v_1(\hat{\mu}) = \frac{r-1}{r}\left[\sum_{\alpha=1}^{r} \hat{\mu}_{(\alpha)}^2 - r\hat{\mu}^2\right] \tag{13}$$

Now

$$E\left[v_1(\hat{\bar{\mu}})\right] = \frac{r-1}{r}\left[\sum_{\alpha=1}^{r} E\left(\hat{\mu}^2_{(\alpha)}\right) - rE\left(\hat{\mu}^2\right)\right]$$

$$= (r-1)\left[V(\hat{\mu}_{(\alpha)}) + \{E(\hat{\mu}_{(\alpha)})\}^2 - V(\hat{\mu}) - \{E(\hat{\mu})\}^2\right]$$

$$= (r-1)\left[V(\hat{\mu}_{(\alpha)}) - V(\hat{\mu})\right] \quad (14)$$

For moderate to large population size N, an approximation to the expression of variance of RSS estimator (as in Equation 4) obtained by Patil *et al.* (1995) can be given as

$$V(\hat{\mu}) = V(\bar{y}_{RSS}) = \frac{1}{mr}\left\{\frac{N-1-mr}{N-1}\sigma^2 - \bar{\gamma}\right\}$$

$$\cong \frac{1}{mr}\frac{N-1-mr}{N-1}\left\{\sigma^2 - \bar{\gamma}\right\} \quad (15)$$

Now, it can be noted that after deletion of a complete cycle exactly r-1 units are there for each ranks in the sample. Then, using the similar notations of Patil *et al.* (1995), from Equation (9) we get

$$m^2(r-1)^2 V(\hat{\mu}_{(\alpha)})$$

$$= (r-1)\sum_{i=1}^{m}\sigma^2_{(i:m)} + (r-1)^2\sum_{i=1}^{m}\sum_{j=1}^{m}C_{ij} - (r-1)\sum_{i=1}^{m}C_{ii}$$

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

$$-(r-1)\sum_{i=1}^{m}C_{ii}$$

$$= (r-1)\left[\frac{m\{N-1-m(r-1)\}}{N-1}\sigma^2 - \sum_{i=1}^{m}(\mu_{(i:m)} - \mu)^2 - \sum_{i=1}^{m}C_{ii}\right]$$

$$\Rightarrow V(\hat{\mu}_{(\alpha)}) = \frac{1}{m(r-1)}\left[\frac{N-1-m(r-1)}{N-1}\sigma^2 - \bar{\gamma}\right]$$

$$\cong \frac{1}{m(r-1)}\frac{N-1-m(r-1)}{N-1}\left[\sigma^2 - \bar{\gamma}\right] \quad (16)$$

where

$$V(y_{(i:m)k}) = \sigma^2_{(i:m)},\ \text{Cov}\left(y_{(i:m)k}, y_{(j:m)k'}\right) = C_{ij},$$

$$\sum_{i=1}^{m}\sigma^2_{(i:m)} = m\sigma^2 - \sum_{i=1}^{m}(\mu_{(i:m)} - \mu)^2, \sum_{i=1}^{m}\sum_{j=1}^{m}C_{ij} = -\frac{m^2}{N-1}\sigma^2$$

and $\bar{\gamma} = \frac{1}{m}\left[\sum_{i=1}^{m}(\mu_{(i:m)} - \mu)^2 + \sum_{i=1}^{m}C_{ii}\right].$

Then, using these results in Equation (14) we get

$$E\left[v_1(\hat{\bar{\mu}})\right] = \frac{r-1}{r}\left[\frac{N-1-m(r-1)}{m(r-1)(N-1)} - \frac{N-1-mr}{mr(N-1)}\right]\left[\sigma^2 - \bar{\gamma}\right]$$

$$= \frac{1}{mr}\left[\sigma^2 - \bar{\gamma}\right] \quad (17)$$

It can be seen that $v_1(\hat{\bar{\mu}})$ is no longer an unbiased estimator of the variance, $V(\hat{\mu})$. This problem of bias of the Jackknife estimator of variance can be eliminated by using

$$\hat{\mu}^*_{(\alpha)} = \hat{\mu} + (1-f)^{1/2}(\hat{\mu}_{(\alpha)} - \hat{\mu}), \text{ instead of } \hat{\mu}_{(\alpha)};$$

where $f = \frac{mr}{N-1}$.

Hence, in this case, the rescaling factor is $(1-f)^{1/2} = \left(1 - \frac{mr}{N-1}\right)^{1/2}$. This results in the following definitions:

The pseudovalues as $\hat{\mu}^*_{\alpha} = r\hat{\mu} - (r-1)\hat{\mu}^*_{(\alpha)}$, Quenouille's estimator as $\hat{\bar{\mu}}^* = \frac{1}{r}\sum_{\alpha=1}^{r}\hat{\mu}^*_{\alpha}$ and the Jackknife estimator of variance as

$$v_1(\hat{\bar{\mu}}^*) = \frac{r-1}{r}\sum_{\alpha-1}^{r}(\hat{\mu}^*_{(\alpha)} - \hat{\mu})^2. \quad (18)$$

Now

$$E\left[v_1(\hat{\bar{\mu}}^*)\right]$$

$$= E\left[\frac{r-1}{r}\sum_{\alpha=1}^{r}\left(\hat{\mu}^*_{(\alpha)} - \hat{\mu}\right)^2\right]$$

$$= E\left[\frac{r-1}{r}\sum_{\alpha=1}^{r}\left\{\hat{\mu} + (1-f)^{1/2}(\hat{\mu}_{(\alpha)} - \hat{\mu}) - \hat{\mu}\right\}^2\right]$$

$$= (1-f)E\left[v_1(\hat{\bar{\mu}})\right] \quad (19)$$

Substituting the expression of Equation (17) in Equation (19), we get

$$E\left[v_1\left(\hat{\hat{\mu}}^*\right)\right] = \left(1-\frac{mr}{N-1}\right)\frac{1}{mr}\left[\sigma^2-\overline{\gamma}\right]$$

$$= \frac{1}{mr}\frac{N-1-mr}{N-1}\left[\sigma^2-\overline{\gamma}\right]$$

$$\cong V(\overline{y}_{RSS}) = V(\hat{\mu}) \tag{20}$$

Hence, the proposed estimator of variance after using the obtained rescaling factor becomes approximately unbiased for the variance of the estimator.

2.2 Strata Based Approach

In this section, we propose another approach of variance estimation in RSS under finite population context named as "Strata based approach". Consider the ranks as strata and observations of each rank as units within a stratum. In the sample of *mr* observations, there appears a natural grouping where there are *r* cycles considered as *r* groups each of size *m*. Here, there are *m* strata and each stratum consisting of *r* units. For applying usual Jackknife procedure, it is proposed to drop one unit randomly from each rank (i.e. stratum) and obtain the estimator from the reduced sample. Here, *m* units are dropped at a time and these *m* units create a group. Then, other *m* units are dropped (one from each stratum) in such a way that these *m* units do not contain those units which were dropped in first dropping. In this manner, exhaustively *mr* sample units (i.e. *r* groups each of size *m*) are dropped in the process of *r* times of dropping procedure. Hence, one unit is dropped from each rank in the process of adopting the proposed strata based approach.

Let $\hat{\mu}$ be the original estimator of population mean, μ. Again let $\hat{\mu}_{(\alpha)}$ be the estimator of the same functional form as $\hat{\mu}$, but computed from the reduced sample of size *m(r-1)* after deleting α^{th} group. The expressions of $\hat{\mu}$ and $\hat{\mu}_{(\alpha)}$ are given by

$$\hat{\mu} = \overline{y}_{RSS} = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{r}\sum_{k=1}^{r} y_{(i:m)k} \quad \text{and} \tag{21}$$

$$\hat{\mu}_{(\alpha)} = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{r-1}\sum_{\substack{k=1\\(k\neq\alpha)}}^{r} y_{(i:m)k}, \quad \forall\, \alpha = 1, 2, ..., r. \tag{22}$$

The pseudovalues are given by

$$\hat{\mu}_{\alpha} = r\hat{\mu}-(r-1)\hat{\mu}_{(\alpha)}, \tag{23}$$

The Quenouille's estimator of μ is given by

$$\hat{\hat{\mu}} = \frac{1}{r}\sum_{\alpha=1}^{r}\hat{\mu}_{\alpha} \tag{24}$$

and the Jackknife estimator of variance of the estimator is given by

$$v_1\left(\hat{\hat{\mu}}\right) = \frac{1}{r(r-1)}\sum_{\alpha=1}^{r}(\hat{\mu}_{\alpha}-\hat{\hat{\mu}})^2. \tag{25}$$

By Lemma (a) of section 1.2, it follows that, $\hat{\hat{\mu}} = \hat{\mu} = \overline{y}_{RSS}$ and thus $\hat{\hat{\mu}}$ is unbiased estimator of population mean, μ.

The Jackknife estimator of variance can alternatively be given by

$$v_1(\hat{\hat{\mu}}) = v_1(\hat{\mu}) = \frac{r-1}{r}\left[\sum_{\alpha=1}^{r}\hat{\mu}_{(\alpha)}^2 - r\hat{\mu}^2\right] \tag{26}$$

Now, in the similar way, from Equation (14) we get

$$E\left[v_1(\hat{\hat{\mu}})\right] = (r-1)\left[V\left(\hat{\mu}_{(\alpha)}\right)-V(\hat{\mu})\right] \tag{27}$$

For moderate to large N,

$$V(\hat{\mu}) = V(\overline{y}_{RSS}) = \frac{1}{mr}\left\{\frac{N-1-mr}{N-1}\sigma^2-\overline{\gamma}\right\}$$

$$\cong \frac{1}{mr}\frac{N-1-mr}{N-1}\left\{\sigma^2-\overline{\gamma}\right\} \tag{28}$$

In the similar way as in previous section, we get

$$V(\hat{\mu}_{(\alpha)}) \cong \frac{1}{m(r-1)}\frac{N-1-m(r-1)}{N-1}\left\{\sigma^2-\overline{\gamma}\right\}, \tag{29}$$

since, after deletion of units by this approach exactly *r*-1 units are there for each ranks in the sample.

Then, using these results in Equation (27) we get

$$E\left[v_1(\hat{\hat{\mu}})\right] = \frac{1}{mr}\left[\sigma^2-\overline{\gamma}\right] \tag{30}$$

It can be seen that $v_1(\hat{\hat{\mu}})$ is no longer an unbiased estimator of the variance, $V(\hat{\mu})$. This problem of bias of the Jackknife estimator of variance can be eliminated by using

$$\hat{\mu}^*_{(\alpha)} = \hat{\mu} + (1-f)^{1/2}(\hat{\mu}_{(\alpha)} - \hat{\mu}),$$ instead of $\hat{\mu}_{(\alpha)}$; where $f = \dfrac{mr}{N-1}$.

Hence, in this case, the rescaling factor is $(1-f)^{1/2} = \left(1 - \dfrac{mr}{N-1}\right)^{1/2}$. This results in the following definitions:

The pseudovalues as $\hat{\mu}^*_\alpha = r\hat{\mu} - (r-1)\hat{\mu}^*_{(\alpha)}$, Quenouille's estimator as $\hat{\hat{\mu}}^* = \dfrac{1}{r}\sum_{\alpha=1}^{r}\hat{\mu}^*_\alpha$ and the Jackknife estimator of variance as

$$v_1(\hat{\hat{\mu}}^*) = \frac{r-1}{r}\sum_{\alpha=1}^{r}(\hat{\mu}^*_{(\alpha)} - \hat{\mu})^2. \quad (31)$$

Now,

$$E\left[v_1(\hat{\hat{\mu}}^*)\right] = E\left[\frac{r-1}{r}\sum_{\alpha=1}^{r}(\hat{\mu}^*_{(\alpha)} - \hat{\mu})^2\right]$$

$$= (1-f)E\left[v_1(\hat{\hat{\mu}})\right] \quad (32)$$

Substituting the value of (30) in (32), we get

$$E\left[v_1(\hat{\hat{\mu}}^*)\right] = \frac{1}{mr}\frac{N-1-mr}{N-1}\left[\sigma^2 - \overline{\gamma}\right]$$

$$\cong V(\overline{y}_{RSS}) = V(\hat{\mu}) \quad (33)$$

Hence, the proposed estimator of variance after using the obtained rescaling factor becomes approximately unbiased for the variance of the estimator.

3. COMPARISON OF PROPOSED VARIANCE ESTIMATION PROCEDURES

In order to study the performance of the proposed variance estimation procedures using Jackknife method in RSS under finite population framework, we did a simulation study. Under simulation study, we generated a bivariate normal population using SAS (Statistical Analysis System) software of size 1000. Here, two variables X and Y were taken into consideration where Y was treated as variable of interest and X was treated as auxiliary variable. The auxiliary variable was used for ranking as required in the procedure of RSS. The parameters of the generated bivariate normal population are given below:

Mean for $X = \overline{X} = 30$, mean for $Y = \overline{Y} = 35$, standard deviation for $X = \sigma_X = 8$, standard deviation for $Y = \sigma_Y = 7$ and the correlation between X and $Y = \rho = 0.7$.

Further, 500 samples of different sample sizes with different combination of number of cycles (r) and number of ranks (m) were drawn using RSS scheme from this simulated population. Then, the estimates of RSS estimator as well as its variance, % CV, skewness and kurtosis were obtained on the basis of estimates from these 500 samples for each sample size separately. Percentage relative bias was obtained using following expression given by

$$\%Bias = \left[\frac{\overline{y}_{RSS} - \overline{Y}}{\overline{Y}}\right] \times 100$$

where $\overline{Y}$ and $\overline{y}_{RSS}$ are the population mean and the estimate of population mean based on RSS estimator respectively.

At the same time, 500 SRSWOR samples were generated to compare the RSS scheme with usual SRSWOR scheme for each RSS sample size. Further, percentage gain in efficiency of the RSS estimator with respect to SRSWOR estimator of population mean was obtained using the following expression

$$GE = \left[\frac{V(\overline{y}_{SRS}) - V(\overline{y}_{RSS})}{V(\overline{y}_{RSS})}\right] \times 100$$

where, $V(\overline{y}_{RSS})$ and $V(\overline{y}_{SRS})$ are the variance obtained based on 500 samples for the RSS estimator and usual SRSWOR estimator respectively.

Further, in order to study the performance of developed variance estimation procedures using Jackknife method, these procedures were applied on each selected RSS sample for different combination of number of cycles (r) and number of ranks (m). For this, Relative Bias (R.B.) and Relative Stability (R.S.) of the estimates of variance of RSS estimator of population mean, $v_1(\hat{\hat{\mu}})$, were computed for the proposed approaches. The formula for R.B. and R.S. are given by

$$R.B. = \left[\frac{\frac{1}{s}\sum_{s} v_{1s}(\hat{\hat{\mu}}^*) - V(\overline{y}_{RSS})}{V(\overline{y}_{RSS})}\right] \times 100$$

and

$$R.S. = \frac{\sqrt{MSE\{v_1(\hat{\bar{\mu}}^*)\}}}{MSE(\bar{y}_{RSS})}$$

$$= \frac{\left[\frac{1}{s}\sum_{s}\left\{v_{1s}(\hat{\bar{\mu}}^*) - V(\bar{y}_{RSS})\right\}^2\right]^{1/2}}{V(\bar{y}_{RSS})}$$

where MSE denotes the mean square error and s denotes the number of samples selected for variance estimation.

SAS codes were written for selection of ranked set samples and for obtaining variance of the RSS estimator of population mean, estimates of variance, Relative Bias and Relative Stability for both the approaches in the context of finite population.

4. SIMULATION RESULTS AND DISCUSSIONS

The statistical properties of the RSS estimator (sample mean) of population mean such as variance, % Bias, % CV, skewness, kurtosis and percentage gain in efficiency (GE) of RSS estimator with respect to usual SRSWOR estimator of population mean were obtained for different sample sizes (n) *i.e.* 60, 120 and 180, using different combination of set size (m) and number of cycles (r) for 500 different samples and are presented in Table 1.

The graphical presentation of comparison of empirical variance of the RSS estimator was made for different sample sizes with different combination of set size(m) and number of cycles (r) and is presented in Fig. 1.

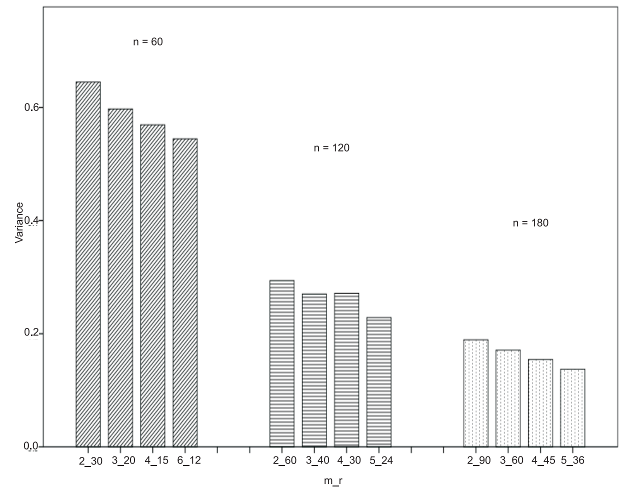

Fig. 1. Comparison of variance of the RSS estimator for different sample sizes with different combination of m and r.

Table 1. RSS estimator along with its variance, % Bias, % CV, Skewness, Kurtosis and GE with respect to SRSWOR estimator for different sample sizes (n) with different combination of set size (m) and number of cycles (r).

n	m	r	RSS Mean	Variance	% Bias	% CV	Skewness	Kurtosis	GE
60	2	30	35.1033	0.6449	–0.0936	2.2877	0.0737	–0.3026	28.9795
120	2	60	35.1789	0.2939	0.1215	1.5411	0.0304	–0.0283	28.9131
180	2	90	35.1262	0.1893	-0.0286	1.2385	0.0578	–0.1280	22.0715
60	3	20	35.2013	0.5974	0.1853	2.1956	–0.1879	–0.0844	39.2459
120	3	40	35.1856	0.2703	0.1406	1.4775	0.1689	0.2009	40.1996
180	3	60	35.1091	0.1709	–0.0772	1.1773	0.0806	–0.1904	35.2174
60	4	15	35.1052	0.5693	–0.0882	2.1494	0.1123	0.0643	46.1001
120	4	30	35.1405	0.2714	0.0122	1.4826	0.0759	–0.1794	39.5902
180	4	45	35.1395	0.1544	0.0093	1.1183	-0.0073	–0.1528	49.6255
60	5	12	35.1876	0.5445	0.1463	2.097	–0.0122	–0.1604	52.7666
120	5	24	35.1603	0.2287	0.0685	1.360	0.0893	–0.0916	65.7128
180	5	36	35.1480	0.1373	0.0335	1.054	0.0305	–0.1488	68.3294

It can be observed from the Table 1 that the estimates of RSS mean are almost unbiased in all the cases considered here. While considering the variance of RSS mean as shown in Table 1 and Fig. 1, it can be seen that the obtained variance is decreasing with the increase of sample size (n) as well as with the increase of set size (m) for a fixed sample size (n). Similar kind of trend is found for percentage gain in efficiency also. With the increase of sample size (n) as well as with the increase of set size (m) for a fixed sample size (n), the RSS estimator becomes more stable in terms of % CV. The RSS estimator is symmetric and mesokurtic in nature. These results ensure the superiority of RSS over SRS in case of finite population, as shown by several authors previously.

The variance of RSS estimator, estimates of variance, relative bias (R.B.) and relative stability (R.S.) following the proposed variance estimation approaches were obtained for different sample sizes (n) with different combination of set size (m) and number of cycles (r) for 500 different samples in case of RSS and are presented in Table 2.

It can be observed from the Table 2 that the values of RS of both the proposed procedures are almost comparable for different cases considered here. Both the procedures show quite low amount of relative bias for estimation of the variance of the RSS estimator. In most of the cases, Cycle based approach shows less RB than the Strata based approach. Both the methods are at par with each other as far as RS is concerned. But with the increase of set size (m) for a fixed sample size, the estimator of the variance obtained following both the approaches becomes less stable in most of the cases. Therefore, it can be concluded that estimator of the variance obtained following both the proposed approaches are almost comparable with respect to RS for different sample sizes with different combination of set size (m) and number of cycles (r). But Cycle based approach is preferable than Strata based approach in terms of RB, simplicity and convenience of the approach.

5. CONCLUSION

In this paper, two different variance estimation procedures namely, Cycle based approach and Strata

Table 2. Comparison of proposed procedures for different sample sizes (n) with different combination of set size (m) and number of cycles (r).

m	r	Variance of the Estimator	Cycle Based Approach			Strata Based Approach		
			Estimate of Variance	RB	RS	Estimate of Variance	RB	RS
2	30	0.6449	0.6557	1.6694	0.2773	0.6504	0.8468	0.2909
	60	0.2939	0.3078	4.7089	0.1971	0.3076	4.6468	0.1945
	90	0.1893	0.1879	–0.7460	0.1441	0.1929	1.9077	0.1492
3	20	0.5974	0.5966	–0.1196	0.3255	0.5766	– 3.4766	0.3154
	40	0.2703	0.2771	2.5341	0.4469	0.2835	4.8806	0.4547
	60	0.1709	0.1727	1.0691	0.1876	0.1740	1.8309	0.1801
4	15	0.5693	0.5656	–0.6518	0.4023	0.5571	–2.1445	0.3588
	30	0.2714	0.2621	–3.4464	0.2478	0.2594	–4.4278	0.2609
	45	0.1544	0.1586	2.7458	0.2239	0.1632	5.7075	0.2306
5	12	0.5445	0.5343	–1.8759	1.0709	0.5265	–3.3098	1.0530
	24	0.2287	0.2510	9.7776	0.3327	0.2475	8.2593	0.3216
	36	0.1373	0.1533	11.6912	0.2976	0.1538	12.0695	0.2926

based approach using Jackknife method in ranked set sampling under finite population framework are proposed. Under the proposed approaches, rescaling factor is obtained in RSS in the context of finite population and it is shown theoretically that the proposed estimator of variance becomes approximately unbiased for variance of the RSS estimator using the obtained rescaling factor. The comparison of the proposed variance estimation procedures is done through simulation study. The variance estimation procedure following Cycle based approach and Strata based approach performs at par with respect to relative stability for different sample sizes with different combination of set size (m) and number of cycles (r). But Cycle based approach is preferable than Strata based approach in terms of relative bias, simplicity and convenience of the approach.

ACKNOWLEDGEMENTS

The authors are highly grateful to an anonymous referee for the valuable comments and suggestions which led to improvement of this research paper.

REFERENCES

Chen, Z., Bai, Z. and Sinha, B. (2004). *Ranked Set Sampling: Theory and Applications*. Springer, New York.

Kankure, A.K. and Rai, A. (2008). Spatial ranked set sampling from spatially correlated population. *J. Ind. Soc. Agril. Statist.*, **62(3)**, 221-230.

Krishna, P. (2002). Some aspects of ranked set sampling from finite population. M.Sc. Thesis, I.A.R.I., New Delhi.

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

Modarres, R., Hui, T.P. and Zheng, G. (2006). Resampling methods for ranked set samples. *Comput. Statist. Data Anal.*, **51**, 1039-1050.

Patil, G.P., Sinha, A.K. and Taillie, C. (1994). Ranked set sampling. In: *Handbook of Statistics.* **12**, (eds. Patil, G.P. and Rao, C.R.), 167-198, North-Holland, Amsterdam.

Patil, G.P., Sinha, A.K. and Taillie, C. (1995). Finite population corrections for ranked set sampling. *Ann. Inst. Statist. Math.*, **47**, 621-636.

Quenouille, M. (1949). Approximate tests of correlation in time series. *Jour. Roy. Statist. Soc.,* **B11**, 18-84.

Shao., J. and Tu, D. (1995). *The Jackknife and Bootstrap*. Springer, New York.

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

Sud, U.C. and Mishra, D.C. (2006). Estimation of finite population mean using ranked set two stage sampling designs. *J. Ind. Soc. Agril. Statist.*, **60(2)**, 108-117.

Sud, U.C. and Mishra, D.C. (2007). Estimation of finite population mean using double ranked set sampling. *J. Ind. Soc. Agril. Statist.*, **61(3)**, 389-399.

Tukey, J.W. (1958). Bias and confidence in not quite large samples. *Ann. Math. Statist*., **29**, 614.

Wolter, K.M. (1985). *Introduction to Variance Estimation*. Springer, New York.