Jour. Ind. Soc. Ag. Statistics
56(2), 2003 : 155-165

Two Stage Successive Sampling with Partial Replacements of Units

R.C. Gola and Anil Rai
Indian Agricultural Statistics Research Institute, New Delhi-110012
(Received : February, 2001)

SUMMARY

Multistage sampling designs are practically inevitable in any large scale surveys. In the available literature the attention has, by and large been given to equal sized first stage units (fsu) with same replacement fraction within the fsu because of algebraic simplicity. However, in the practical situations often met with, the fsu are generally of unequal size. Therefore, in this article partial matching among fsu as well as ssu has been considered for estimation of mean on the second occasion under two stage sampling designs. Three estimators of mean have been proposed for this purpose (i) General linear estimator (ii) Linear composite estimator and (iii) Ratio type composite estimator. Further, the above three estimators were compared empirically with the help of real survey data. It has been observed that the gain in efficiency for general linear estimator was of the order of 200% and for the ratio type composite estimator, it was of about 50% while for linear composite estimator, the gain was only marginal. Thus it may be concluded that the substantial gains in the efficiency of the estimators may be achieved by making use of the technique of successive sampling coupled with an appropriate choice of the type of estimator to be used.

Key words : Two stage successive sampling, Unequal first stage units, Partial replacements.

1. Introduction

Consider a population consists of N first stage units (fsu) having M_i second stage units (ssu) in the i-th fsu $\{i = 1, 2, \ldots, N\}$. Further, n denotes the number of fsu in the simple random sample whereas m is the number of ssu in the sampled fsu. The character under study from the j-th ssu in the i-th fsu on the h-th occasion is denoted by y_{hij}. The fraction of fsu retained on the current occasion from the sample drawn on the preceding occasion is denoted by λ. The fraction of fsu drawn afresh in the sample on the current occasion is denoted by μ such that $(\lambda + \mu = 1)$. Further p is the fraction of ssu retained in each of the sampled fsu on the current occasion from the ones selected on the preceding occasion whereas q is the fraction of ssu drawn afresh in each of the above fsu. This means that p is the fraction of ssu retained in each of the sampled fsu which are selected on preceding occasion and retained for the second occasion.

Multistage sampling designs are practically inevitable in any large scale surveys. The applicability of multistage designs to successive sampling was studied by Kathuria ([1], [2]), Singh and Kathuria [3], Tikkiwal [7], Singh [4], Singh and Srivastava ([6], [5]).

In the available literature the attention has, by and large been given to equal sized fsu with same replacement fraction within the fsu because of algebraic simplicity. However, in the practical situations, the fsu are generally of unequal size. Therefore, there is a need to investigate these procedures for estimating the population parameter i.e. mean, total etc. under such situations. In this article partial matching among fsu as well as ssu has been considered for estimation of mean on the second occasion under two stage sampling designs. Three estimators of mean have been proposed for this purpose

1. General linear estimator
2. Linear composite estimator and
3. Ratio type composite estimator

Further, the above three estimators are compared empirically with the help of data collected under the project "Sample survey for methodological investigations into high yielding varieties programme" during V-th five year plan by Indian Agricultural Statistics Research Institute, New Delhi. The data pertains to Ambala district (administrative Unit) of Haryana state.

2. General Linear Estimator of the Mean

A general linear estimator of the population mean at the second occasion i.e. $\bar{y}_2$ under the above sampling scheme viz. partial matching among fsu as well as among ssu may be expressed as follows

$$\bar{T}_2' = a\bar{t}_1' + b\bar{t}_1'' - (a+b)\bar{t}_1''' + d\bar{t}_2' + e\bar{t}_2'' + (1-d-e)\bar{t}_2''' \tag{2.1}$$

where

$$\bar{t}_1' = \frac{1}{\lambda n}\sum_{i=1}^{\lambda n}\frac{M_i}{\bar{M}}\frac{1}{mp}\sum_{j=1}^{mp} y_{1ij}$$

$$\bar{M} = \frac{1}{N}\sum_{i=1}^{N} M_i$$

$$\bar{t}_1'' = \frac{1}{\lambda n}\sum_{i=1}^{\lambda n} \frac{M_i}{\bar{\bar{M}}} \frac{1}{mq}\sum_{j=1}^{mq} y_{1ij}$$

$$\bar{t}_1''' = \frac{1}{\mu n}\sum_{i=1}^{\mu n} \frac{M_i}{\bar{\bar{M}}} \frac{1}{m}\sum_{j=1}^{m} y_{1ij}$$

$$\bar{t}_2' = \frac{1}{\lambda n}\sum_{i=1}^{\lambda n} \frac{M_i}{\bar{\bar{M}}} \frac{1}{mp}\sum_{j=1}^{mp} y_{2ij}$$

$$\bar{t}_2'' = \frac{1}{\lambda n}\sum_{i=1}^{\lambda n} \frac{M_i}{\bar{\bar{M}}} \frac{1}{mq}\sum_{j=1}^{mpq} y_{2ij}$$

$$\bar{t}_2''' = \frac{1}{\mu n}\sum_{i=1}^{\mu n} \frac{M_i}{\bar{\bar{M}}} \frac{1}{m}\sum_{j=1}^{m} y_{2ij}$$

Here a, b, d and e are constants to be evaluated so as to provide minimum variance for the proposed estimator. Following the zero function approach $\bar{\bar{T}}_2'$ will be minimum variance linear unbiased estimator (MVULE) if and only if

$$\begin{aligned}
&\text{Cov}\,(\bar{\bar{T}}_2', \bar{t}_1' - \bar{t}_1'') = 0 \\
&\text{Cov}\,(\bar{\bar{T}}_2', \bar{t}_1' - \bar{t}_1''') = 0 \\
&\text{Cov}\,(\bar{\bar{T}}_2', \bar{t}_2' - \bar{t}_2'') = 0 \\
&\text{Cov}\,(T_2', \bar{t}_2' - \bar{t}_2''') = 0
\end{aligned} \qquad (2.2)$$

Solving the above equations and assuming

$$S_{1b} = S_{2b} = S_b$$

$$S_{1i} = S_{2i} = S_i$$

$$S_{12i} = \rho_w S_i^2$$

$$S_{12b} = \rho_b\, S_b^2$$

where $$S_{hi}^2 = \frac{1}{M_i - 1}\sum_{j=1}^{M_i} (Y_{hij} - \bar{Y}_{hi})^2, \quad h = 1, 2$$

$$\bar{Y}_{hi} = \frac{1}{M_i}\sum_{j=1}^{M_i} Y_{hij}$$

$$S_{hb}^2 = \frac{1}{N-1}\sum_{i=1}^{N} \left(\frac{M_i}{\bar{\bar{M}}}\bar{Y}_{hi} - \bar{Y}_h\right)^2, \ h = 1, 2$$

$$\overline{Y}_h = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{M_i}\sum_{j=1}^{M_i} Y_{hij}$$

$$S_{hh'i} = \frac{1}{M_i - 1}\sum_{j=1}^{M_i} (Y_{hij} - \overline{Y}_{hi})(Y_{h'ij} - \overline{Y}_{h'i}), \quad h = 1, 2$$

$$\overline{S}_w^2 = \frac{1}{N}\sum_{i=1}^{N} S_i^2$$

$$\rho_w = \frac{\sum_{i=1}^{N} M_i^2 \sum_{j=1}^{M_i} (Y_{hij} - \overline{Y}_{hi})(Y_{h'ij} - \overline{Y}_{h'i})}{\sqrt{\sum_{i=1}^{N} M_i^2 \sum_{j=1}^{M_i} (Y_{hij} - \overline{Y}_{hi})^2}}$$

$$\rho_{hh'b} = \frac{\sum_{i=1}^{N}\left(\frac{M_i}{\overline{M}}\overline{Y}_{hi} - \overline{Y}_h\right)\left(\frac{M_i}{\overline{M}}\overline{Y}_{h'i} - \overline{Y}_{h'}\right)}{\sqrt{\sum_{i=1}^{N}\left(\frac{M_i}{\overline{M}}\overline{Y}_{hi} - \overline{Y}_h\right)^2 \sum_{i=1}^{N}\left(\frac{M_i}{\overline{M}}\overline{Y}_{h'i} - \overline{Y}_{h'}\right)^2}}$$

$$\rho_b = \frac{\sum_{i=1}^{N}\left(\frac{M_i}{\overline{M}}\overline{Y}_{hi} - \overline{Y}_h\right)\left(\frac{M_i}{\overline{M}}\overline{Y}_{hi'} - \overline{Y}_h\right)}{\sum_{i=1}^{N}\left(\frac{M_i}{\overline{M}}\overline{Y}_{hi} - \overline{Y}_h\right)^2}$$

We get following equations after ignoring finite correction factor

$$aq - bp + dq\,\rho_w = 0 \tag{2.3}$$

$$a\left[S_b^2 + \left(\frac{\mu}{p} + \lambda\right)\frac{\overline{S}_w^2}{m}\right] + b\left[S_b^2 + \lambda\frac{\overline{S}_w^2}{m}\right] + d\,\frac{\mu}{p}\left[p\rho_b\, S_b^2 + \rho_w\frac{\overline{S}_w^2}{m}\right] + e\,\mu\,\rho_b\, S_b^2 = 0 \tag{2.4}$$

$$aq\,\rho_w + dq - dq - ep = 0 \tag{2.5}$$

$$a\left[p\mu\rho_b S_b^2 + \mu\rho_w \frac{\overline{S}_w^2}{m}\right] + bp\mu\rho_b S_b^2 + d\left[pS_b^2 + (\mu+\lambda p)\frac{\overline{S}_w^2}{m}\right]$$

$$+ e\left[pS_b^2 + (\lambda p)\frac{\overline{S}_w^2}{m}\right] - \left[\lambda p S_b^2 + \frac{\overline{S}_w^2}{m}\right] = 0 \qquad (2.6)$$

Now by solving above equations for constants with the help of Cramer's rule and substituting it in the variance expression, after ignoring finite correction factors we obtain

$$V(\overline{T}_2') = \frac{\left(S_b^2 + \mu\frac{\overline{S}_w^2}{m}\right)}{\mu n}\left[1 - \frac{\lambda\left(S_b^2 + \frac{\overline{S}_w^2}{m}\right)\left[\left\{1 - q^2\rho_w^2\right\}S_b^2 + \left\{1 - q(\mu+\lambda q)\rho_w^2\right\}\frac{\overline{S}_w^2}{m}\right]}{\Delta''}\right]$$

where

$$\Delta'' = \left[\left\{(1+\rho_b\rho_w\,\mu q)S_b^2 + \frac{\overline{S}_w^2}{m}\right\}^2 - \left\{(\rho_b\mu + \rho_w q)S_b^2 + \rho_w(\mu+\lambda\, q)\frac{\overline{S}_w^2}{m}\right\}\right]$$

3. Linear Composite Estimator of the Mean

A linear composite estimator of the population mean for the second occasion i.e. $\overline{y}_2$ under the above sampling scheme viz. partial matching among fsu as well as ssu may be obtained in two steps. First by considering the part of the sample on both the occasions, where all fsu are retained with partial matching among ssu. In this case a linear composite estimator of the mean on the second occasion is worked out. In the second step, we consider the part of the sample, where fsu are unmatched. In this case a linear estimator of the mean on the second occasion is worked out. The two estimators so obtained are then suitably combined to get an unbiased composite estimator of $\overline{Y}_2$. A composite estimator for the second occasion based on first matching criterion can be written as

$$\overline{t}_{12} = Q(\overline{t}_{\lambda 1} + \overline{t}_2' - \overline{t}_1') + (1-Q)\overline{t}_2'' \qquad (3.1)$$

where

$$\overline{t}_{\lambda 1} = \frac{1}{\lambda n}\sum_{i=1}^{\lambda n}\frac{M_i}{\overline{M}\, m}\sum_{j=1}^{m} y_{1ij}$$

and Q is a constant to be evaluated so as to provide minimum variance for the estimator proposed.

Let $\bar{t}_{2m} = \bar{t}_{\lambda 1} + \bar{t}'_2 - \bar{t}'_1$

so $\bar{t}_{12} = Q\bar{t}_{2m} + (1-Q)\,\bar{t}''_2$

its variance can be written as

$$V(\bar{t}_{12}) = Q^2 V(\bar{t}_{2m}) + (1-Q)^2\, V(\bar{t}''_2) + 2Q(1-Q)\,\mathrm{Cov}(\bar{t}_{2m}, \bar{t}''_2) \qquad (3.2)$$

The optimum value of Q is obtained by minimizing the above variance expression with respect to Q . This gives

$$Q_{opt} = \frac{V(\bar{t}''_2) - \mathrm{Cov}(\bar{t}_{2m'}, \bar{t}''_2)}{V(\bar{t}_{2m}) + V(\bar{t}''_2) - 2\mathrm{Cov}\,(\bar{t}_{2m'}, \bar{t}''_2)} \qquad (3.3)$$

Substituting different variances and covariances in the variance expression and ignoring finite population correction we get

$$V(\bar{t}_{12}) = \frac{1}{\lambda n} S^2_{2b} + \frac{(1-Q)}{\lambda n m_2} \overline{S}^2_{2w} \qquad (3.4)$$

where $m_2 = mq$

$$\overline{S}^2_{2w} = \frac{1}{N} \sum_{i=1}^{N} S^2_{2i}$$

Now, a linear composite estimator of $\overline{Y}_2$ based on all the sampling units may be written as

$$\hat{T}_{12} = k(\bar{t}_1 + \bar{t}_{12} - \bar{t}_{\lambda 1}) + (1-k)\bar{t}'''_2$$

where $\bar{t}_1 = \frac{1}{n} \sum_{i=1}^{n} \frac{M_i}{\overline{M}} \frac{1}{m} \sum_{j=1}^{m} (y_{1ij})$ and k is a constant which is derived such that the variance of the $\hat{T}_{i2}$ minimizes.

Let

$$\bar{u}_2 = (\bar{t}_1 + \bar{t}_{12} - \bar{t}_{\lambda 1})$$

Therefore

$$\hat{T}_{12} = k\bar{u}_2 + (1-k)\bar{t}'''_2 \qquad (3.5)$$

The variance expression and the optimum value of k can be written as

$$V(\hat{T}_{12}) = k^2 V(\bar{u}_2) + (1-k)^2 V(\bar{t}'''_2) + 2k(1-k)\,\mathrm{Cov}(\bar{u}_2, t'''_2)$$

and

$$K_{opt} = \frac{V(\bar{t}'''_2) - \mathrm{Cov}(\bar{u}_2, \bar{t}'''_2)}{V(\bar{u}_2) + V(\bar{t}'''_2) - 2\mathrm{Cov}(\bar{u}_2, \bar{t}'''_2)} \qquad (3.6)$$

Ignoring finite population correction and under the assumptions

$$S^2_{1b} = S^2_{2b} = S^2_b$$

$\overline{S}_{1w}^2 = \overline{S}_{2w}^2 = \overline{S}_w^2$

The K_{opt} can be written as

$$K_{opt} = \frac{\dfrac{1}{\mu n}\left[S_b^2 + \dfrac{\overline{S}_w^2}{m}\right]}{\left[\dfrac{1}{\mu n}\left[S_b^2 + \dfrac{\overline{S}_w^2}{m}\right] + \dfrac{1}{\lambda n}S_b^2 - \dfrac{2\mu}{\lambda n}\rho_b S_b^2 - \dfrac{1}{n}\dfrac{\overline{S}_w^2}{m} + \left[\dfrac{1-Q}{mn}\right]\overline{S}_w^2 + \dfrac{1}{\lambda n}\dfrac{\overline{S}_w^2}{m} - \dfrac{2Q\mu}{\lambda n}\rho_w \dfrac{\overline{S}_w^2}{m}\right]}$$

The expression of the variance of $\hat{T}_{12}$ can be written as

$$V(\hat{T}_{12}) = (1 - K_{opt})\left[\frac{S_b^2}{\mu n} + \frac{1}{\mu n}\frac{\overline{S}_w^2}{m}\right] \tag{3.7}$$

4. Ratio Type Composite Estimator of Mean

The estimator based on partial matching among the fsu as well as among ssu for the current occasion i.e. $\overline{y}_2$ under the above sampling pattern could be built up by taking a suitable combination of the following ratio estimators based on matched sub sample is as follows

$$\overline{Z}_3^* = Q_1\left(\frac{\overline{t}_2'}{\overline{t}_1'}\overline{t}_1\right) + (1 - Q_1)\left(\frac{\overline{t}_2''}{\overline{t}_1''}\overline{t}_1\right) \tag{4.1}$$

where

$$t_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{M_i}{\overline{M}}\frac{1}{n}\sum_{j=1}^{m}(y_{1ij})$$

An overall estimator, by taking the unmatched units also into account, may be written as

$$T_3^* = Q_2\overline{Z}_3^* + (1 - Q_2)\overline{t}_2'' \tag{4.2}$$

where Q_1 and Q_2 are the constants, obtained such as to provide minimum variance of $\overline{Z}_3^*$ and T_3^* respectively.

After ignoring finite population correction and under the following assumptions

$$S_{1b}^2 = S_{2b}^2 = S_b^2$$
$$\bar{S}_{1w}^2 = \bar{S}_{2w}^2 = \bar{S}_w^2$$
$$S_{12w} = \rho_b S_b^2$$
$$\bar{S}_{12w} = \rho_w \bar{S}_w^2$$

we get

$$E(\bar{Z}_3^*) = Q_1 \bar{Y}_2 \left[\frac{1}{nm} \rho_w \frac{\bar{S}_w^2}{\bar{Y}_1 \bar{Y}_2} + \frac{1}{\lambda nmp} \frac{\bar{S}_w^2}{\bar{Y}_1^2} - \frac{1}{\lambda nmp} \rho_w \frac{\bar{S}_w^2}{\bar{Y}_1 \bar{Y}_2} - \frac{1}{nm} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right.$$
$$\left. - \frac{1}{\lambda nmp} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right] + \bar{Y}_2 \left[1 + \left(\frac{1}{n} - \frac{1}{\lambda n} \right) \frac{\rho S_b^2}{\bar{Y}_1 \bar{Y}_2} - \left(\frac{1}{n} - \frac{1}{\lambda n} \right) \frac{S_b^2}{\bar{Y}_1^2} + \frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right] \tag{4.3}$$

where

$$Q_{1opt} = \frac{\frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} + \frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_2^2} + \frac{1}{nm} \frac{\bar{S}_w^2}{\bar{Y}_1^2} - \frac{1}{nm} \frac{\rho_w}{\bar{Y}_1} \frac{\bar{S}_w^2}{\bar{Y}_2}}{\frac{1}{\lambda nmpq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} + \frac{1}{\lambda nmpq} \frac{\bar{S}_w^2}{\bar{Y}_2^2} - \frac{2}{\lambda nmp} \frac{\rho_w \bar{S}_w^2}{\bar{Y}_1 \bar{Y}_2}} \tag{4.4}$$

The relative bias of T_3^* can be obtained as

$$R.B.(T_3^*) = Q_1 Q_2 \left[\frac{1}{nm} \frac{\rho_w \bar{S}_w^2}{\bar{Y}_1 \bar{Y}_2} + \frac{1}{\lambda nmp} \frac{\bar{S}_w^2}{\bar{Y}_1^2} - \frac{1}{\lambda nmp} \frac{\rho_w \bar{S}_w^2}{\bar{Y}_1 \bar{Y}_2} \right.$$
$$\left. - \frac{1}{nm} \frac{\bar{S}_w^2}{\bar{Y}_1^2} - \frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right] + Q_2 \left[\frac{-\mu \rho_b S_b^2}{\lambda n \bar{Y}_1 \bar{Y}_2} + \frac{\mu S_b^2}{\lambda n \bar{Y}_1^2} + \frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right] \tag{4.5}$$

The variance of T_3^* under the above assumptions can be written as

$$V(T_3^*) = (1 - Q_{2opt}) \left[\frac{1}{\mu_n} S_b^2 + \frac{1}{\mu n} \frac{\bar{S}_w^2}{m} \right]$$
$$+ Q_{2opt} \bar{Y}_2^2 \left[\frac{\mu}{\lambda n} \frac{\rho_b S_b^2}{\bar{Y}_1 \bar{Y}_2} - \frac{\mu S_b^2}{\lambda n} \frac{S_b^2}{\bar{Y}_1^2} - \frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right]$$
$$- Q_{1opt} Q_{2opt} \bar{Y}_2^2 \left[\frac{1}{mn} \frac{\rho_w}{\bar{Y}_1} \frac{\bar{S}_w^2}{\bar{Y}_2} + \frac{1}{\lambda nmp} \frac{\bar{S}_w^2}{\bar{Y}_1^2} - \frac{1}{\lambda nmq} \frac{\bar{S}_w^2}{\bar{Y}_1^2} \right.$$

$$-\frac{1}{\lambda mnp}\frac{\rho_w \overline{S}_w^2}{\overline{Y}_1\overline{Y}_2}-\frac{1}{nm}\frac{\overline{S}_w^2}{\overline{Y}_1^2}\Bigg] \tag{4.6}$$

where the optimum value of Q_2 is as follows

$$Q_{2opt}=\frac{V(\bar{t}_2^{**})-\mathrm{Cov}\left(\overline{Z}_3^*,\bar{t}_2^{**}\right)}{V(\overline{Z}_3^*)+V(\bar{t}_2^{**})-2\mathrm{Cov}\left(\overline{Z}_3^*,\bar{t}_2^{**}\right)} \tag{4.7}$$

By substituting the various variance-covariance terms in the above equation we can get the expression of the Q_{2opt} as

$$Q_{opt}=\frac{\left[\begin{aligned}&\frac{1}{\mu n}S_b^2+\frac{1}{\mu n}\frac{\overline{S}_w^2}{m}+\overline{Y}_2^2\left[-\frac{\mu}{\lambda n}\frac{\rho_b S_b^2}{Y_1Y_2}+\frac{\mu}{\lambda n}\frac{S_b^2}{Y_1^2}+\frac{1}{\lambda mnq}\frac{\overline{S}_w^2}{Y_1^2}\right]\\&+Q_1\overline{Y}_2^2\left[\frac{1}{mn}\frac{\rho_w\overline{S}_w^2}{\overline{Y}_1\overline{Y}_2}+\frac{1}{\lambda mnp}\frac{\overline{S}_w^2}{\overline{Y}_1^2}-\frac{1}{\lambda mnp}\frac{\rho_w\overline{S}_w^2}{Y_1Y_2}\right.\\&\left.-\frac{1}{mn}\frac{\overline{S}_w^2}{\overline{Y}_1^2}-\frac{1}{\lambda mnq}\frac{\overline{S}_w^2}{\overline{Y}_1^2}\right]\end{aligned}\right]}{\left[\begin{aligned}&\frac{1}{\mu n}S_b^2+\frac{1}{\mu n}\frac{\overline{S}_w^2}{m}+\overline{Y}_2^2\left[\frac{1}{\lambda n}\frac{S_b^2}{\overline{Y}_2^2}+\frac{3\mu}{\lambda n}\frac{S_b^2}{\overline{Y}_1^2}-\frac{4\mu}{n}\frac{\rho_b S_b^2}{\overline{Y}_1\overline{Y}_2}\right.\\&\left.+\frac{1}{\lambda mnq}+\frac{\overline{S}_w^2}{\overline{Y}_2^2}+\frac{1}{mn}\frac{\overline{S}_w^2}{\overline{Y}_1^2}+\frac{3}{\lambda mnq}\frac{\overline{S}_w^2}{\overline{Y}_1^2}\right]+Q_1\overline{Y}_2^2\left[-\frac{1}{\lambda mnq}\frac{\overline{S}_w^2}{\overline{Y}_2^2}-\frac{3\overline{S}_w^2}{mn\overline{Y}_1^2}\right.\\&\left.-\frac{3}{\lambda mnq}\frac{\overline{S}_w^2}{\overline{Y}_1^2}+\frac{3}{mn}\frac{\rho_w\overline{S}_w^2}{\overline{Y}_1\overline{Y}_2}+\frac{2}{\lambda knp}\frac{\overline{S}_w^2}{\overline{Y}_1^2}-\frac{2}{\lambda mnp}\frac{\rho_w\overline{S}_w^2}{\overline{Y}_1\overline{Y}_2}\right]\end{aligned}\right]}$$

5. An Empirical Comparison

In view of complicated expressions for the variances of the estimators, an attempt has been made to have such comparisons empirically with the help of the real data. In this situation, the gain in precision of these estimators over the usual simple estimator of the mean of the second occasion without making use of the data collected on the first occasion have been investigated. Such comparisons, although would not be rigorous one, but would give an idea about the gain in efficiency of these estimators over the simple estimator.

In case of the present empirical study the data collected by Indian Agricultural Statistics Research Institute, New Delhi, under the research project entitled "Sampling investigations into high yielding varieties programme" conducted during IV-th and V-th five year plan periods have been used. The data of the study pertained to Ambala block of Ambala district (administrative unit) in Haryana state for the seasons of 1977-78 and 1978-79. The data were collected through two stage design with villages growing wheat crop as the first stage sampling units and the cultivators growing wheat crop as the second stage units. The character under study was area brought under cultivation of high yielding varieties (HYV) of wheat crop. The sample size at the first stage was 12 villages selected out of the totality of villages growing wheat crop in Ambala block. At the second stage 10 cultivators were selected per sampled village out of totality of cultivators growing wheat crop. The units at both the stages were selected by simple random sampling without replacement. In the second year viz. 1978-79, 6 villages out of 12 sampled villages in 1977-78 were retained and remaining 6 villages were freshly drawn out of the villages not sampled in 1977-78. Likewise in each of the retained 6 villages, 5 cultivators were retained in 1978-79 out of the 10 cultivators drawn in 1977-78 and the remaining 5 cultivators were drawn freshly from cultivators not sampled in 1977-78. Table 1 presents the estimated relative variance of different estimators and their efficiency compared to simple estimator based on second occasion data for estimation of total area under high yielding varieties of wheat in Ambala district.

Table 1

Sr. No.	Estimator	Estimated relative variance for total area under HYV of wheat	Efficiency (%)
1.	General linear estimator	7.25	290.49
2.	Ratio type composite estimator	10.34	152.27
3.	Linear composite estimator	12.74	109.35
4.	Simple estimator	13.07	–

From the above results it is observed that the gain in efficiency for general linear estimator is of the order of 200% and for the ratio type composite estimator, it is about 50% while for linear composite estimator the gain is only marginal. Thus it may be summarised that the substantial gains in the efficiency of the estimators may be achieved by making use of the technique of successive sampling coupled with an appropriate choice of the type of estimator to be used.

ACKNOWLEDGEMENT

Authors are thankful to the referee for his valuable comments for the improvement of the quality of paper.

REFERENCES

[1] Kathuria, O.P. (1959). Some aspects of successive sampling in multistage designs. Unpublished Diploma thesis, ICAR, New Delhi.

[2] Kathuria, O.P. (1973). On alternative replacement procedure in sampling on successive occasions with a two stage design and on use of multi auxiliary information in such designs. Unpublished Ph. D. thesis, IARI, New Delhi.

[3] Singh, D. and Kathuria, O.P. (1969). On two-stage successive sampling. *Australian Jour. Statist.*, **11(2).**

[4] Singh, S. (1970). Some studies in two-stage successive sampling. Unpublished Diploma thesis, ICAR, New Delhi.

[5] Singh, D., Singh, S. and Srivastava, A.K. (1976). On repeat surveys in two stage sampling designs. *Jour. Ind. Soc. Agril. Stat.*, **26 (1)**.

[6] Srivastava, A.K. and Singh, S. (1974). A note on two-stage successive sampling. *Jour. Ind. Soc. Agril. Stat.*, **26(1)**.

[7] Tikkiwal, B.D. (1965). Theory of two-stage sampling on successive occasions. *Jour. Ind. Statist. Assoc.*