Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 64(3) 2010 407-415

On the Use of Super-Population Model in Two-Occasion Rotation Patterns

G.N. Singh* and Jaishree Prabha Karna
Indian School of Mines, Dhanbad

Received 06 August 2009; Revised 30 April 2010; Accepted 24 June 2010

SUMMARY

The present work is an attempt to study the estimation of the population mean on the current occasion in two-occasion successive (rotation) sampling under a super-population model. An estimator of the population mean for the current occasion in two-occasion successive (rotation) sampling has been proposed. Optimum replacement policy and performance of the proposed estimator have been discussed. Results are interpreted via empirical studies.

Keywords: Successive sampling, Super-population, Linear model, Auxiliary variable, Variance, Optimum replacement policy.

INTRODUCTION

There are many problems of practical interest in different fields of applied and environmental sciences in which the various characters opt to change over time. For such situations one time surveys are not sufficient, there is a need of continuous monitoring. Surveys where sampling is done on successive occasions (over years or seasons or months) according to a specified rule, with partial replacement of units, is called successive (rotation) sampling. Successive (rotation) sampling provides a strong statistical tool for generating the reliable estimates on different occasions.

The data collected on the first occasion can be used for estimating population characteristics of the second occasion as done by Jessen (1942), Patterson (1950), Rao and Graham (1964) among others. Avadhani and Sukhatme (1970) used data collected on the first occasion for selection of sample on the second occasion while Arnab (1998) used the data collected on the first occasion for selection of sample as well as stratification for the second occasion. Sen (1971) applied this theory with success in designing the estimator for the population mean on the current occasion using information on two auxiliary variables available on previous occasion. Sen (1972, 1973) extended his work for several auxiliary variables. Singh *et al.* (1991) and Singh and Singh (2001) used the auxiliary information on current occasion and proposed estimators for the current population mean in two-occasion successive (rotation) sampling. Singh (2003) extended their work for h-occasions successive sampling.

In many situations, information on an auxiliary variable may be readily available on the first as well as on the second occasion; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation, number of beds in different hospitals may be known in hospital surveys, number of polluting industries and vehicles are known in environmental survey, nature of employment status, educational status, food availability and medical aids of a locality are well known in advance for estimating the various demographic parameters in demographic surveys. Utilizing the auxiliary information on both the occasions, Feng and Zou (1997), Biradar and Singh (2001), Singh (2005), Singh and Priyanka (2006, 2007a and 2008a) have proposed several estimators for

*Corresponding author : G.N. Singh
E-mail address : gnsingh_ism@yahoo.com

estimating the population mean on the current (second) occasion in two-occasion successive (rotation) sampling.

All the above quoted works are design based and consists of some known and unknown population parameters. Sometimes, it may be unrealistic to get the readymade information on these population parameters, in such situations; it is more realistic to assume a super-population model with unknown model parameters. Model parameters may be estimated from the available data. Such a model could efficiently link the study and auxiliary variable on different occasions. Motivated with these arguments, Singh and Priyanka (2007b and 2008b) proposed several estimators for estimating the population mean on the current (second) occasion in two-occasion successive (rotation) sampling and studied their behaviors under a super-population model.

There are various ways for utilizing the available information at estimation stage in successive sampling. Considering the best linear combination of all the available information for estimating the current population mean in two-occasion successive sampling is one of them. In the present work, we utilize the auxiliary information which is available on both the occasions through a super-population model. We consider that the auxiliary information available on both the occasions is stable over time. It is to be mentioned that under the assumed super-population model, errors are correlated over two-occasion and the auxiliary variable is gamma distributed. Following the works of Patterson (1950), Kulldorff (1963) and utilizing all the available information we propose a generalized estimator of population mean on the current occasion in two-occasion successive sampling. Unbiasedness and minimum expected variance criterions of the proposed estimator have been explored. Optimum replacement policy is discussed and results are demonstrated through empirical means of comparison.

2. NOTATIONS

Let $U = (U_1, U_2, ..., U_N)$ be the N-element finite population, which has been sampled over two-occasion, and the character under study be denoted by y_h (h = 1, 2) on the h^{th} occasion. It is assumed that information on a stable auxiliary variable z (with known population mean) is available on both the occasions. A simple random sample (without replacement) s_1 of size n is selected on the first occasion. A random sub sample s_{2m} by SRSWOR of size $m = n\lambda$ $(0 \le \lambda \le 1)$ from the sample s_1 on the first occasion is retained (matched) for its use on the second occasion. A fresh sample (un-matched sample) s_{2u} of size $u = (n - m) = n\mu$ $(\mu = 1 - \lambda)$ is selected on the second occasion from the entire population by SRSWOR method so that the sample size on the second occasion is also n. λ and μ $(\lambda + \mu = 1)$ are the fractions of matched and fresh samples on the current (second) occasion respectively. Hence onwards we consider the following notations for their further use:

$\bar{Y}_h$: The population mean of the study variable y_h on the h^{th} (h = 1, 2) occasion.

$\bar{y}_{1n}$: The sample mean based on n units on the first occasion.

$\bar{y}_{hm}$: The sample mean on h^{th} (h = 1, 2) occasion, based on m units common to both the occasions.

$\bar{y}_{2u}$: The sample mean based on u units drawn afresh on the current (second) occasion.

$\bar{Z}$: The population mean of the auxiliary variable z.

$\bar{z}_n, \bar{z}_u, \bar{z}_m$: The sample means of the auxiliary variable z of the sample sizes shown in suffices.

$f\left(=\dfrac{n}{N}\right)$: The sampling fraction.

3. SUPER POPULATION MODEL

Here we assume that the study variable y_h (h = 1, 2) follows the following super-population model:

$$y_{hi} = \alpha_h + \beta_h z_i + e_{hi},\ h = 1, 2 \text{ and } i = 1, 2, ..., N \quad (1)$$

where α_h, β_h(h = 1, 2) are unknown real constants and e_{hi}'s are random errors (disturbances) over h^{th} (h = 1, 2) occasion, such that

$$E_c(e_{hi} \mid z_i) = 0 \quad (2)$$

$$E_c(e_{hi}e_{hj} \mid z_i, z_j) = 0\ \forall\ (i \ne j = 1, 2, ..., N) \quad (3)$$

$$E_c\left(e_{hi}^2 \mid z_i\right) = \delta_h z_i^{g_h};\ \delta_h > 0;\ 0 \le g_h \le 2 \qquad (4)$$

$$E_c(e_{hi} e_{h'j} \mid z_i, z_j) = 0 \ \forall\ (i \ne j = 1, 2, ..., N) \text{ and } h \ne h' = 1, 2 \qquad (5)$$

$$E_c\left(e_{hi} e_{h'i} \mid z_i\right) = \rho_{e_h e_{h'}} \sqrt{\delta_h \delta_{h'}}\, z_i^{\frac{g_h + g_{h'}}{2}};$$
$$\delta_{h'} > 0;\ 0 \le g_{h'} \le 2;\ h \ne h' = 1, 2 \qquad (6)$$

where $\rho_{e_h e_{h'}}$ is the correlation coefficient between the random errors (disturbances) over the two-occasion.

E_c denotes, the conditional expectation given z_i ($i = 1, 2, \ldots, N$). We assume that z_i's are independently and identically distributed gamma random variables with common density

$$f(z) = \frac{1}{\lceil\theta} e^{-z} z^{\theta - 1};\ z \ge 0,\ \theta > 1 \qquad (7)$$

Let us denote the expectation with respect to common distribution of z_i by E_z, model expectation by E_m ($= E_z E_c$) and the sampling design expectation by E_d.

4. PROPOSED ESTIMATOR

We wish to estimate the population mean $\bar{Y}_2$ on the current (second) occasion. Following Kulldorff (1963), we may propose an unbiased estimator of $\bar{Y}_2$, which is as follows:

$$T = \left\{a_1 \bar{y}_{2m} + a_2 \bar{y}_{2u}\right\} + \left\{b_1 \bar{y}_{1m} + b_2 \bar{y}_{1n}\right\} + \left\{c_1 \bar{z}_m + c_2 \bar{z}_u + c_3 \bar{z}_n + c_4 \bar{Z}\right\} \qquad (8)$$

where a_1, a_2, b_1, b_2, c_1, c_2, c_3 and c_4 are constants to be determined so that

(i) T becomes unbiased for $\bar{Y}_2$ i.e $E_d(T) = \bar{Y}_2$ and

(ii) the expected variance of $T = E_m\{V_d(T)\}$ attains a minimum.

For unbiasedness condition i.e $E_d(T) = \bar{Y}_2$, we must have

$$a_1 + a_2 = 1;\ b_1 + b_2 = 0 \text{ and } c_1 + c_2 + c_3 + c_4 = 0$$

Substituting $a_1 = \varphi$, $b_1 = \beta$ and $c_4 = -(c_1 + c_2 + c_3)$, the estimator T defined in equation (8) reduces to the following form:

$$T = \left\{\varphi \bar{y}_{2m} + (1-\varphi)\bar{y}_{2u}\right\} + \beta\left\{\bar{y}_{1m} - \bar{y}_{1n}\right\} + \left\{c_1\left(\bar{z}_m - \bar{Z}\right) + c_2\left(\bar{z}_u - \bar{Z}\right) + c_3\left(\bar{z}_n - \bar{Z}\right)\right\}$$

$$= \varphi\left[\bar{y}_{2m} + \frac{\beta}{\varphi}\left\{\bar{y}_{1m} - \bar{y}_{1n}\right\} + \frac{c_1}{\varphi}\left\{\bar{z}_m - \bar{Z}\right\} + \frac{c_3}{\varphi}\left\{\bar{z}_n - \bar{Z}\right\}\right] + (1-\varphi)\left[\bar{y}_{2u} + \frac{c_2}{(1-\varphi)}\left\{\bar{z}_u - \bar{Z}\right\}\right]$$

$$= \varphi \hat{\hat{\bar{Y}}}_{2m} + (1-\varphi)\hat{\hat{\bar{Y}}}_{2u} \qquad (9)$$

where $\hat{\hat{\bar{Y}}}_{2m} = \bar{y}_{2m} + k_1\left\{\bar{y}_{1m} - \bar{y}_{1n}\right\} + k_2\left\{\bar{z}_m - \bar{Z}\right\} + k_3\left\{\bar{z}_n - \bar{Z}\right\}$ i.e. estimator based on the matched sample and $\hat{\hat{\bar{Y}}}_{2u} = \bar{y}_{2u} + k_4\left\{\bar{z}_u - \bar{Z}\right\}$ i.e. estimator based on the un-matched sample, where $k_1 = \frac{\beta}{\varphi}$, $k_2 = \frac{c_1}{\varphi}$, $k_3 = \frac{c_3}{\varphi}$ and $k_4 = \frac{c_2}{(1-\varphi)}$.

Remark 4.1: It is obvious from equation (9) that for estimating the population mean on the current (second) occasion, ignoring findings from the first occasion, the estimator $\hat{\hat{\bar{Y}}}_{2u}$ is suitable, and it would be appropriate to choose φ to be 0, while for estimating the change from one occasion to the next, the estimator $\hat{\hat{\bar{Y}}}_{2m}$ may be emphasized, so choosing φ as 1. For asserting both the problems simultaneously, the suitable (optimum) choice of φ is desired.

5. PROPERTIES OF THE ESTIMATOR T

Theorem 5.1: The expected variance of the estimator T is obtained as

$$E_m V_d(T) = \varphi^2 E_m V_d\left(\hat{\hat{\bar{Y}}}_{2m}\right) + (1-\varphi)^2 E_m V_d\left(\hat{\hat{\bar{Y}}}_{2u}\right) + 2\varphi(1-\varphi) E_m \mathrm{Cov}_d\left(\hat{\hat{\bar{Y}}}_{2m}, \hat{\hat{\bar{Y}}}_{2u}\right) \qquad (10)$$

where E_m denotes the expectation with respect to the model, V_d and Cov_d respectively denote the variance and the covariance with respect to the sampling design d.

$$E_m V_d\left(\hat{\hat{Y}}_{2m}\right) = \left(\frac{1}{m}-\frac{1}{N}\right)\left\{\left(\beta_2+k_2\right)^2\theta+\delta_2\frac{\left\lceil g_2+\theta\right.}{\left\lceil\theta\right.}\right\}$$
$$+\left(\frac{1}{n}-\frac{1}{N}\right)\left\{k_3^2\theta+2k_3\beta_2\theta+2k_2k_3\theta\right\}$$
$$+\left(\frac{1}{m}-\frac{1}{n}\right)\left[k_1^2\left(\beta_1^2\theta+\delta_1\frac{\left\lceil g_1+\theta\right.}{\left\lceil\theta\right.}\right)\right.$$
$$+2k_1\beta_1\beta_2\theta+2k_1k_2\beta_1\theta$$
$$\left.+2k_1\rho_{e1e2}\sqrt{\delta_1\delta_2}\frac{\left\lceil g+\theta\right.}{\left\lceil\theta\right.}\right] \quad (11)$$

$$E_m V_d\left(\hat{\hat{Y}}_{2u}\right) = \left(\frac{1}{u}-\frac{1}{N}\right)\left[\left(\beta_2+k_4\right)^2\theta+\delta_2\frac{\left\lceil g_2+\theta\right.}{\left\lceil\theta\right.}\right] \quad (12)$$

and

$$E_m Cov_d\left(\hat{\hat{Y}}_{2m},\hat{\hat{Y}}_{2u}\right) = -\frac{1}{N}\left[\left(\beta_2+k_2+k_3\right)\right.$$
$$\left.\left(\beta_2+k_4\right)\theta+\delta_2\frac{\left\lceil g_2+\theta\right.}{\left\lceil\theta\right.}\right] \quad (13)$$

where $g = \dfrac{g_1+g_2}{2}$.

Proof: The expected variance of the estimator T is given by

$$E_m V_d(T) = E_m E_d\left[T-\bar{Y}_2\right]^2 = \varphi^2 E_m V_d\left(\hat{\hat{Y}}_{2m}\right)$$
$$+\left(1-\varphi\right)^2 E_m V_d\left(\hat{\hat{Y}}_{2u}\right)$$
$$+2\varphi\left(1-\varphi\right)E_m \mathrm{Cov}_d\left(\hat{\hat{Y}}_{2m},\hat{\hat{Y}}_{2u}\right) \quad (14)$$

Under the assumed linear super-population model $y_{hi} = \alpha_h + \beta_h z_i + e_{hi}$, ($i$ = 1, 2, ..., N; h = 1, 2) we can write

$$\left.\begin{aligned}\bar{y}_{2u} &= \alpha_2+\beta_2\bar{z}_u+\bar{e}_{2u}\\ \bar{y}_{1n} &= \alpha_1+\beta_1\bar{z}_n+\bar{e}_{1n}\\ \bar{y}_{2m} &= \alpha_2+\beta_2\bar{z}_m+\bar{e}_{2m}\\ \bar{y}_{1m} &= \alpha_1+\beta_1\bar{z}_m+\bar{e}_{1m}\\ \bar{Y}_2 &= \alpha_2+\beta_2\bar{Z}+\bar{e}_{2N}\end{aligned}\right\} \quad (15)$$

where

$$\bar{e}_{2u} = \frac{1}{u}\sum_{i=1}^{u}e_{2i},\ \bar{e}_{1n} = \frac{1}{n}\sum_{i=1}^{n}e_{1i},\ \bar{e}_{2m} = \frac{1}{m}\sum_{i=1}^{m}e_{2i},$$
$$\bar{e}_{1m} = \frac{1}{m}\sum_{i=1}^{m}e_{1i} \text{ and } \bar{e}_{2N} = \frac{1}{N}\sum_{i=1}^{N}e_{2i}.$$

Following Rao (1968), the expected variances and covariance of the estimators $\hat{\hat{Y}}_{2u}$ and $\hat{\hat{Y}}_{2m}$ have been derived in three steps. First, we operate the design expectation E_d, secondly the conditional expectation E_c and finally the expectation E_z with respect to the common distribution of z_i. The design expectation can be evaluated using the results given in Sukhatme *et al.* (1984).

Under the super-population model defined in equations (1) – (7) and using results of the equation (15), the expected variance of $\hat{\hat{Y}}_{2u}$ comes out as

$$E_m V_d\left(\hat{\hat{Y}}_{2u}\right) = E_m E_d\left[\hat{\hat{Y}}_{2u} - \bar{Y}_2\right]^2$$
$$= E_m E_d\left[\left(\beta_2+k_4\right)\left(\bar{z}_u-\bar{Z}\right)+\left(\bar{e}_{2u}-\bar{e}_{2N}\right)\right]^2$$
$$= E_m E_d\left[\left(\beta_2+k_4\right)^2\left(\bar{z}_u-\bar{Z}\right)^2+\left(\bar{e}_{2u}-\bar{e}_{2N}\right)^2+2\left(\beta_2+k_4\right)\left(\bar{z}_u-\bar{Z}\right)\left(\bar{e}_{2u}-\bar{e}_{2N}\right)\right]$$
$$= E_m\left[\left(\beta_2+k_4\right)^2 V_d\left(\bar{z}_u\right)+V_d\left(\bar{e}_{2u}\right)+2\left(\beta_2+k_4\right)\mathrm{Cov}_d\left(\bar{z}_u,\bar{e}_{2u}\right)\right]$$
$$= E_z E_c\left[\left(\beta_2+k_4\right)^2\left(\frac{1}{u}-\frac{1}{\mathrm{N}}\right)S_z^2+\left(\frac{1}{u}-\frac{1}{N}\right)S_{e_2}^2+2\left(\beta_2+k_4\right)\left(\frac{1}{u}-\frac{1}{N}\right)S_{ze_2}\right]$$

$$= E_z E_c \left[\frac{1}{N-1}\left(\frac{1}{u} - \frac{1}{N}\right) \left\{ \left(\beta_2 + k_4\right)^2 \left(\sum_{i=1}^{N} z_i^2 - N\,\bar{Z}^2\right) + \left(\sum_{i=1}^{N} e_{2i}^2 - N\,\bar{e}_{2N}^2\right) + 2\left(\beta_2 + k_4\right)\sum_{i=1}^{N}\left(z_i - \bar{Z}\right)\left(e_{2i} - \bar{e}_{2N}\right) \right\} \right] \quad (16)$$

Using the conditions in equations (2) – (7) we evaluate E_c, which gives

$$E_m V_d\left(\hat{\bar{Y}}_{2u}\right) = \left(\frac{1}{u} - \frac{1}{N}\right)\frac{1}{N-1} E_z \left[\left(\beta_2 + k_4\right)^2 \left(\sum_{i=1}^{N} z_i^2 - N\bar{Z}^2\right) + \left(\sum_{i=1}^{N} \delta_2 z_i^{g_2} - \frac{1}{N}\sum_{i=1}^{N} \delta_2 z_i^{g_2}\right) \right] \quad (17)$$

Now, evaluating E_z using the probability distribution of z_i, equation (17) reduces to

$$E_m V_d\left(\hat{\bar{Y}}_{2u}\right) = \left(\frac{1}{u} - \frac{1}{N}\right)\left[\left(\beta_2 + k_4\right)^2 \theta + \delta_2 \frac{\overline{|g_2 + \theta}}{\overline{|\theta}}\right] \quad (18)$$

Similarly, we can derive the expressions of $E_m V_d\left(\hat{\bar{Y}}_{2m}\right)$ and $E_m Cov_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)$ given in equations (11) and (13) respectively.

6. OPTIMUM VALUES OF φ, k_1, k_2, k_3, k_4

Since, the expected variance of the estimator T in equation (10) is a function of unknown constants φ, k_1, k_2, k_3 and k_4, therefore, it is minimized with respect to φ, k_1, k_2, k_3 and k_4 respectively and subsequently the optimum values of φ, k_1, k_2, k_3 and k_4 are obtained as

$$\varphi_{\text{opt}} = \frac{E_m V_d\left(\hat{\bar{Y}}_{2u}\right) - E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)}{E_m V_d\left(\hat{\bar{Y}}_{2m}\right) + E_m V_d\left(\hat{\bar{Y}}_{2u}\right) - 2E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)} \quad (19)$$

$$k_1^* = k_{1\,\text{opt}} = -\frac{\rho_{\text{e1e2}}\sqrt{\delta_1\delta_2}\,\frac{\overline{|g + \theta}}{\overline{|\theta}}}{\delta_1 \frac{\overline{|g_1 + \theta}}{\overline{|\theta}}} \quad (20)$$

$$k_2^* = k_{2\,\text{opt}} = -\left(\beta_2 + k_1^*\beta_1\right) \quad (21)$$

$$k_3^* = k_{3\,\text{opt}} = k_1^*\beta_1 \quad (22)$$

and

$$k_4^* = k_{4\,\text{opt}} = -\beta_2 \quad (23)$$

Substituting the values of k_1^*, k_2^*, k_3^* and k_4^* from equations (20), (21), (22) and (23) into the equation (10), we get the optimum expected variance of T, say $E_m V_d(T)^*$ with respect to k_1, k_2, k_3 and k_4 as

$$E_m V_d(T)^* = \varphi^2 E_m V_d\left(\hat{\bar{Y}}_{2m}\right)^* + (1-\varphi)^2 E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^* + 2\varphi(1-\varphi) E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^* \quad (24)$$

where

$$E_m V_d\left(\hat{\bar{Y}}_{2m}\right)^* = \left(\frac{1}{m} - \frac{1}{N}\right)\left(\delta_2 \frac{\overline{|g_2 + \theta}}{\overline{|\theta}}\right) + \left(\frac{1}{m} - \frac{1}{n}\right)\left[k_1^{*2}\delta_1 \frac{\overline{|g_1 + \theta}}{\overline{|\theta}} + 2k_1^*\rho_{\text{e1e2}}\sqrt{\delta_1\delta_2}\,\frac{\overline{|g + \theta}}{\overline{|\theta}}\right] \quad (25)$$

$$E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^* = \left(\frac{1}{u} - \frac{1}{N}\right)\left(\delta_2 \frac{\overline{|g_2 + \theta}}{\overline{|\theta}}\right) \quad (26)$$

and

$$E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^* = -\frac{1}{\text{N}}\left(\delta_2 \frac{\overline{|g_2 + \theta}}{\overline{|\theta}}\right) \quad (27)$$

In the light of equations (25)-(27), the optimum value of φ shown in equation (19) becomes

$$\varphi_{\text{opt}} = \frac{E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^* - E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^*}{E_m V_d\left(\hat{\bar{Y}}_{2m}\right)^* + E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^* - 2E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^*} \quad (28)$$

7. MINIMUM EXPECTED VARIANCE OF T

Substituting the value of φ_{opt} given in equation (28) in equation (24), we get the minimum expected variance of T as

$$E_m V_d\,(T)^*_{\text{opt}} = \frac{\left[E_m V_d\left(\hat{\bar{Y}}_{2m}\right)^* . E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^* - \left\{E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^*\right\}^2\right]}{\left[E_m V_d\left(\hat{\bar{Y}}_{2m}\right)^* + E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^* - 2E_m \text{Cov}_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^*\right]} \tag{29}$$

Further substituting the values of $E_m V_d\left(\hat{\bar{Y}}_{2m}\right)^*$, $E_m V_d\left(\hat{\bar{Y}}_{2u}\right)^*$ and $E_m Cov_d\left(\hat{\bar{Y}}_{2m}, \hat{\bar{Y}}_{2u}\right)^*$ from equations (25) – (27) in equation (29) the simplified value of $E_m V_d\,(T)^*_{\text{opt}}$ is obtained as

$$E_m V_d\,(T)^*_{\text{opt}} = \frac{A_1}{n}\left[\frac{A_1 + \mu A_2}{A_1 + \mu^2 A_2} - f\right] \tag{30}$$

where $A_1 = \delta_2 \dfrac{\lceil g_2 + \theta}{\lceil \theta}$,

$$A_2 = k_1^{*2}\delta_1 \frac{\lceil g_1 + \theta}{\lceil \theta} + 2k_1^* \rho_{\text{e1e2}} \sqrt{\delta_1 \delta_2}\, \frac{\lceil g + \theta}{\lceil \theta}$$

and $f = \dfrac{n}{N}$.

8. OPTIMUM REPLACEMENT POLICY

To determine the optimum value of μ (fraction of samples to be taken afresh on the current (second) occasion) so that population mean $\bar{Y}_2$ may be estimated with the maximum precision, we minimize $E_m V_d\,(T)^*_{\text{opt}}$ given in equation (30) with respect to μ, which yield

$$\hat{\mu} = \frac{-A_1 \pm \sqrt{A_1^2 + A_1 A_2}}{A_2} \tag{31}$$

From equation (31) it is obvious that real value of $\hat{\mu}$ exist, iff, the quantity under square root is greater than or equal to zero. For any combination of parameters that satisfy the condition of real solution; two real values of $\hat{\mu}$ is possible. Hence, while choosing the value of $\hat{\mu}$, it should be remembered that $0 \le \hat{\mu} \le 1$. All the other values of $\hat{\mu}$ is said to be inadmissible. Substituting the admissible value of $\hat{\mu}$ say $\mu^{(0)}$ from equation (31) into equation (30), we have the following optimum value of the expected variance of T with respect to μ

$$E_m V_d\,(T^0)^*_{\text{opt}} = \frac{A_1}{n}\left[\frac{A_1 + \mu^{(0)} A_2}{A_1 + \mu^{(0)2} A_2} - f\right] \tag{32}$$

9. EFFICIENCY COMPARISON

The percent relative efficiency of T with respect to the simple unbiased sample mean estimator $\bar{y}_{2n}$ of the population mean $\bar{Y}_2$ on the current (second) occasion has been obtained. The sample mean estimator $\bar{y}_{2n}$ is exclusively based on a sample of size n on the current (second) occasion, which uses no prior information. Its expected variance under the assumed super-population model is given by

$$E_m V_d\left(\bar{y}_{2n}\right) = \frac{(1-f)}{n}\left(\beta_2^2\theta + \delta_2 \frac{\lceil g_2 + \theta}{\lceil \theta}\right) \tag{33}$$

For different choices of β_2, g_1, g_2, ρ_{e1e2}, δ_1, δ_2 and θ, Table 1 shows the optimum values of $\mu^{(0)}$ and percent relative efficiencies E of T over the estimator $\bar{y}_{2n}$ under optimal condition, where $E = \dfrac{E_m V_d\left(\bar{y}_{2n}\right)}{E_m V_d\,(T^0)^*_{\text{opt}}} \times 100$.

Table 1. Optimum values of $\mu^{(0)}$ and percent relative efficiency E of the estimator T over the estimator $\bar{y}_{2n}$ for (f = 0.1, δ_1 = 2.0, δ_2 = 2.0, θ = 8.0)

ρ_{e1e2}			0.3		0.5		0.7		0.9	
β_2	g_1	g_2	$\mu^{(0)}$	E	$\mu^{(0)}$	E	$\mu^{(0)}$	E	$\mu^{(0)}$	E
0.5	0.5	0.5	0.5118	139.48	0.5359	146.84	0.5834	161.57	0.6964	197.94
		1.0	0.5117	115.43	0.5356	121.46	0.5825	133.51	0.6931	162.93
		1.5	0.5114	106.87	0.5347	112.32	0.5800	123.09	0.6840	148.65
		2.0	0.5110	103.87	0.5333	108.95	0.5763	118.87	0.6710	141.43
	1.0	0.5	0.5117	139.45	0.5356	146.74	0.5825	161.29	0.6931	196.84
		1.0	0.5118	115.45	0.5359	121.54	0.5834	133.73	0.6964	163.84
		1.5	0.5117	106.94	0.5356	112.53	0.5826	123.70	0.6933	150.99
		2.0	0.5114	103.97	0.5348	109.28	0.5802	119.79	0.6846	144.76
	1.5	0.5	0.5114	139.37	0.5347	146.47	0.5800	160.52	0.6840	193.84
		1.0	0.5117	115.43	0.5356	121.47	0.5826	133.52	0.6933	162.98
		1.5	0.5118	106.96	0.5359	112.60	0.5834	123.90	0.6964	151.79
		2.0	0.5117	104.03	0.5356	109.47	0.5826	120.34	0.6934	146.92
	2.0	0.5	0.5110	139.24	0.5333	146.05	0.5763	159.35	0.6710	189.59
		1.0	0.5114	115.36	0.5348	121.25	0.5802	132.91	0.6846	160.63
		1.5	0.5117	106.94	0.5356	112.54	0.5826	123.71	0.6934	151.03
		2.0	0.5118	104.05	0.5356	109.54	0.5834	120.53	0.6964	147.66
1.0	0.5	0.5	0.5118	250.04	0.5359	263.24	0.5834	289.64	0.6964	354.85
		1.0	0.5117	153.90	0.5356	161.95	0.5825	178.01	0.6931	217.24
		1.5	0.5114	119.87	0.5347	125.98	0.5800	138.06	0.6840	166.73
		2.0	0.5110	108.14	0.5333	113.43	0.5763	123.76	0.6710	147.24
	1.0	0.5	0.5117	249.99	0.5356	263.06	0.5825	289.15	0.6931	352.87
		1.0	0.5118	153.93	0.5359	162.06	0.5834	178.31	0.6964	218.46
		1.5	0.5117	119.94	0.5356	126.22	0.5826	138.74	0.6933	169.35
		2.0	0.5114	108.24	0.5348	113.77	0.5802	124.71	0.6846	150.71
	1.5	0.5	0.5114	249.85	0.5347	262.58	0.5800	287.76	0.6840	347.50
		1.0	0.5117	153.90	0.5356	161.96	0.5826	178.03	0.6933	217.31
		1.5	0.5118	119.96	0.5359	126.30	0.5834	138.96	0.6964	170.25
		2.0	0.5117	108.30	0.5356	113.97	0.5826	125.29	0.6934	152.96
	2.0	0.5	0.5110	249.62	0.5333	261.83	0.5763	285.67	0.6710	339.88
		1.0	0.5114	153.82	0.5348	161.67	0.5802	177.22	0.6846	214.17
		1.5	0.5117	119.94	0.5356	126.22	0.5826	138.75	0.6934	169.40
		2.0	0.5118	108.32	0.5359	114.04	0.5834	125.48	0.6964	153.73

ρ_{e1e2}			0.3		0.5		0.7		0.9	
β_2	g_1	g_2	$\mu^{(0)}$	E	$\mu^{(0)}$	E	$\mu^{(0)}$	E	$\mu^{(0)}$	E
1.5	0.5	0.5	0.5118	434.31	0.5359	457.23	0.5834	503.10	0.6964	616.37
		1.0	0.5117	218.03	0.5356	229.43	0.5825	252.18	0.6931	307.76
		1.5	0.5114	141.53	0.5347	148.75	0.5800	163.01	0.6840	196.86
		2.0	0.5110	115.26	0.5333	120.89	0.5763	131.90	0.6710	156.93
	1.0	0.5	0.5117	434.23	0.5356	456.93	0.5825	502.24	0.6931	612.92
		1.0	0.5118	218.08	0.5359	229.58	0.5834	252.61	0.6964	309.49
		1.5	0.5117	141.62	0.5359	149.03	0.5826	163.81	0.6933	199.96
		2.0	0.5114	115.36	0.5348	121.25	0.5802	132.91	0.6846	160.63
	1.5	0.5	0.5114	433.98	0.5347	456.09	0.5800	499.84	0.6840	603.61
		1.0	0.5117	218.03	0.5356	229.44	0.5826	252.21	0.6933	307.86
		1.5	0.5118	141.64	0.5359	149.12	0.5834	164.08	0.6964	201.02
		2.0	0.5117	115.43	0.5356	121.47	0.5826	133.53	0.6934	163.03
	2.0	0.5	0.5110	433.59	0.5333	454.79	0.5763	496.20	0.6710	590.36
		1.0	0.5114	217.92	0.5348	229.04	0.5802	251.07	0.6846	303.41
		1.5	0.5117	141.62	0.5356	149.03	0.5826	163.83	0.6934	200.01
		2.0	0.5118	115.45	0.5359	121.54	0.5834	133.73	0.6964	163.84

10. CONCLUSION

The following points can be read out from Table 1:

(a) For the fixed values of β_2, g_1, and g_2 the values of $\mu^{(0)}$ and E are increasing with the increasing values of $\rho_{e_1e_2}$.

(b) For the fixed values of β_2, g_2 and ρ_{e1e2} the values of $\mu^{(0)}$ and E decrease when the value of g_1 is increased.

(c) For the fixed values of β_2, g_1 and ρ_{e1e2} the values of $\mu^{(0)}$ and E decrease when the value of g_2 is increased.

(d) For the fixed values of g_1, g_2 and ρ_{e1e2} the values of $\mu^{(0)}$ are stable while the efficiency E increases with the increasing values of β_2.

(e) The percent relative efficiency E and $\mu^{(0)}$ is independent of β_1.

From the above discussions, it may be concluded that the proposed estimator T is appreciably preferable over sample mean estimator under the proposed super-population model. Hence, the use of auxiliary variable through a super-population model is highly beneficial and it may be recommended to the survey statisticians/practitioners for its further use.

ACKNOWLEDGEMENT

Authors are thankful to the honorable referee for his valuable and inspiring suggestions. Authors are also thankful to the University Grants Commission, New Delhi and Indian School of Mines, Dhanbad for providing the financial assistance and necessary infrastructure to carry out the present work.

REFERENCES

Arnab, R. (1998). Sampling on two occasions: Estimation of population total. *Survey Methodology*, **24**, 185-192.

Avadhani, M.S. and Sukhatme, B.V. (1970). A comparison of two sampling procedures with applications to successive sampling. *Appl. Statist.*, **19**, 251-259.

Biradar, R.S. and Singh, H.P. (2001). Successive sampling using auxiliary information on both occasions. *Cal. Stat. Assoc. Bull.*, **51**, 243-251.

Feng, S. and Zou, G. (1997). Sample rotation method with auxiliary variable. *Comm. Statist.-Theory Methods*, **26(6),** 1497-1509.

Jessen, R.J. (1942). Statistical Investigation of a Sample Survey for obtaining Farm Facts. *Iowa Agricultural Experiment Station Research Bulletin No. 304, Ames, Iowa, U.S.A.,* 1-104.

Kulldorff, G. (1963). Some problems of optimum allocation for sampling on two occasions. *Rev. Int. Statist. Inst.*, **31**, 24-57.

Patterson, H.D. (1950). Sampling on successive occasions with partial replacement of units. *J. Roy. Statist. Soc.,* **B12**, 241-255.

Rao, J.N.K. and Graham, J.E. (1964). Rotation design for sampling on repeated occasions. *J. Amer. Statist. Assoc.,* **59**, 492-509.

Rao, P.S.R.S. (1968). On three procedures of sampling from finite populations. *Biometrika*, **55**, 438-441.

Sen, A.R. (1971). Successive sampling with two auxiliary variables. *Sankhya*, **B33**, 371-378.

Sen, A.R. (1972). Successive sampling with p ($p \geq 1$) auxiliary variables. *Ann. Math. Statist.*, **43**, 2031-2034.

Sen, A.R. (1973). Theory and application of sampling on repeated occasions with several auxiliary variables. *Biometrics,* **29,** 381-385.

Singh, V.K., Singh, G.N. and Shukla, D. (1991). An efficient family of ratio-cum-difference type estimators in successive sampling over two occasions. *Jour. Sci. Res.* **41C**, 149-159.

Singh, G.N. and Singh, V.K. (2001). On the use of auxiliary information in successive sampling. *J. Ind. Soc. Agril. Statist*., **54(1),** 1-12.

Singh, G.N. (2003). Estimation of population mean using auxiliary information on recent occasion in h-occasion successive sampling. *Statist. Trans.,* **6**, 523-532.

Singh, G.N. (2005). On the use of chain-type ratio estimator in successive sampling. *Statist. Trans.*, **7**, 21-26.

Singh, G.N. and Priyanka, K. (2006). On the use of chain-type ratio to difference estimator in successive sampling. *Int. Jour. Appl. Math. Statist.*, **5 (S06),** 41-49.

Singh, G.N. and Priyanka, K. (2007a). On the use of auxiliary information in search of good rotation patterns on successive occasions. *Bull. Statist. Eco.,* **1 (A07),** 42 – 60.

Singh, G.N. and Priyanka, K. (2007b). Estimation of population mean at current occasion in successive sampling under a super – population model. *Model Assist. Statist. Appl.*, **2**, 189 – 200.

Singh, G.N. and Priyanka, K. (2008a). Search of good rotation patterns to improve the precision of estimates at current occasion. *Comm. Statist.- Theory Methods,* **37(3)**, 337-348.

Singh, G.N. and Priyanka, K. (2008b). Use of super – population model in search of good rotation patterns on successive occasions. *J. Statist. Res.,* **42(1),** 127-141.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). *Sampling Theory of Surveys with Applications*. Iowa State University Press, Ames, Iowa (USA) and Indian Society of Agricultural Statistics, New Delhi (India).