Journal of Modern Applied Statistical Methods

Volume 10 | Issue 1 Article 6

5-1-2011

Estimation of Population Mean In Successive Sampling by Sub-Sampling Non-Respondents

Housila P. Singh
Vikram University, Ujjain, India, hpsujn@rediffmail.com

Sunil Kumar
University of Jammu, (J & K), India, sunilbhougal06@gmail.com

Sandeep Bhougal
Shri Mata Vaishno Devi University, Kakryal, Jammu (J & K), India, sandeep.bhougal@smvdu.ac.in

Recommended Citation

Singh, Housila P.; Kumar, Sunil; and Bhougal, Sandeep (2011) "Estimation of Population Mean In Successive Sampling by Sub-Sampling Non-Respondents," *Journal of Modern Applied Statistical Methods*: Vol. 10 : Iss. 1 , Article 6.
DOI: 10.22237/jmasm/1304222700
Available at: http://digitalcommons.wayne.edu/jmasm/vol10/iss1/6

Journal of Modern Applied Statistical Methods
May 2011, Vol. 10, No. 1, 51-60

1538 – 9472/11/$95.00

Estimation of Population Mean In Successive Sampling by Sub-Sampling Non-Respondents

Housila P. Singh
Vikram University,
Ujjain, India

Sunil Kumar
University of Jammu,
(J & K), India

Sandeep Bhougal
Shri Mata Vaishno Devi University,
Kakryal, Jammu (J & K), India

The estimation of the population mean in mail surveys is investigated in the context of sampling on two occasions where the population mean of the auxiliary variable is available in the presence of non-response only for the current occasion in two occasion successive sampling. The behavior of the proposed estimator is compared with the estimator for the same situation but in the absence of non-response. An empirical illustration demonstrates the performance of the proposed estimator.

Key words: Variance, study variable, auxiliary variable, non-response, successive sampling.

Introduction

A very important problem for many countries is the management and conservation of food resources. However, it commonly occurs that the classical theory of sampling cannot be directly applied in situations calling for quantification of environmental resources. If a population is subject to change, a survey carried out on a single occasion cannot of itself give any information of the nature or rate of such change (Miranda, 2007, p. 385).

The problem of sampling on two successive occasions was first considered by Jessen (1942) and has also been discussed by Patterson (1950), Narain (1953), Eckler (1955), Adhvaryu (1978), Sen (1979), Gorden (1983) and Arnab and Okafor (1992). In addition to the information from previous research, Singh, et al. (1991), Artes and Garcia (2001), Singh and Singh (2001), Garcia and Artes (2002), Singh (2003) and Singh and Vishwakarma (2007), used auxiliary information on current occasion for estimating the current population mean in two-occasion successive sampling.

It is common experience in sample surveys that a proportion of people among those invited to participate in a non-compulsory interview survey, or other study, choose not to take part or are unobtainable for other reasons. Non-response covers all causes of non-participation including, direct refusals, people who are away temporarily on holiday and non-contacts for other reasons. Those who are found to be outside the scope of the survey are classified as ineligible and excluded altogether. Ineligibles include people who had died or moved to an area outside the survey area, businesses that had closed down and changed addresses.

Hansen and Hurwitz (1946) were the first to suggest a technique of handling non-response in mail surveys. Cochran (1977), Okafor and Lee (2000) extended the Hansen and Hurwitz technique to the case when along with the information on character under study, information is also available on an auxiliary character. More recently Choudhary, et al. (2004), Okafor (2005) and Singh and Priyanka (2007) used the Hansen and Hurwitz technique for estimating the population mean on current occasion in the context of sampling on two occasions. This article investigates successive sampling theory in the presence of non-response and examines the efficiency over the estimate

Housila P. Singh is a Professor in the School of Studies in Statistics. Email him at: hpsujn@rediffmail.com. Sunil Kumar is in the Department of Statistics, University of Jammu. Email him at: sunilbhougal06@gmail.com. Sandeep Bhougal is in the School of Mathematics, SMVDU. Email him at: sandeep.bhougal@smvdu.ac.in.

defined for the same situation with complete response.

Building an Estimator

Suppose that the two samples are of size n on both occasions and simple random sampling and the size of the population N is used which is sufficiently large for the correlation factor to be ignored.

Let $U = (U_1, U_2, ..., U_N)$ represent the total population of N identifiable units that have been sampled over two occasions. Let $x(y)$ be the character under study on the first (second) occasions respectively. It is deduced that information on an auxiliary variable x is available on both the occasions with known population mean. A simple random sample without replacement of n units is taken on the first occasion.

On the second occasion, a simple random sample without replacement of $m = n\lambda$ units is retained while an independent sample of $u = n\mu = n - m$ units is selected so that the sample size on both the occasions is the same, n units. It is assumed that there is non-response at the second (current) occasion, so that the population can be divided into two classes, those who will respond at the first attempt and those who will not: let the sizes of these two classes be N_1 and N_2 respectively. Assume that in the matched (unmatched) portion of the sample on two occasions $m_1(u_1)$ units respond and $m_2(u_2)$ units do not. Let $m_{h_2}(u_{h_2})$ units denote the size of the sub-sample drawn from the non-response class from the matched (unmatched) portion of the sample on the two occasions for collecting information through personal interview.

This study considers the same situation as outlined in Singh and Kumar (2010), where the information on the auxiliary variable is completely available for all the second phase sample of size n units while, out of n sample units on the current occasion, some units refused to respond on the study variable y. Hansen and Hurwitz (1946) technique to sub sampling from $m_2(u_2)$ non-respondents of size $m_{h_2}(u_{h_2})$ units selected at random and is enumerated by direct interview, such that, by $(m_{h_2} = m_2/k)$ $(u_{h_2} = u_2/k)$; $k > 1$, one will obtain the estimate

$$\left(\bar{y}_{2m_{h_2}} = \sum_{j=1}^{m_{h_2}} y_j / m_{h_2}\right)\left(\bar{y}_{2u_{h_2}} = \sum_{j=1}^{u_{h_2}} y_j / u_{h_2}\right).$$

Using $\bar{y}_{2m_2}(\bar{y}_{2u_2})$, an unbiased estimator $\bar{y}^*$ of the population mean $\bar{Y}$ of the study variable y on the current occasion will be constructed. For these $m_{h_2}(u_{h_2})$ units selected from $m_2(u_2)$ non-respondent units one can also obtain the estimate

$$\left(\bar{x}_{2m_{h_2}} = \sum_{j=1}^{m_{h_2}} x_j / m_{h_2}\right)\left(\bar{x}_{2u_{h_2}} = \sum_{j=1}^{u_{h_2}} x_j / u_{h_2}\right)$$

and using this estimate results in the unbiased estimate $\bar{x}^*$ on the current occasion.

Further, an estimator is constructed when there is non-response only on the second occasion as:

$$t_m = \bar{y}_m^* + \lambda_1\left(\bar{x}_n - \bar{x}_m^*\right) + \lambda_2\left(\bar{x}_m - \bar{x}_n\right), \tag{2.1}$$

where λ_1 and λ_2 are suitably chosen constants,

$$\bar{y}_m^* = \frac{m_1\bar{y}_{m_1} + m_2\bar{y}_{m_{h_2}}}{m}$$

is the Hansen and Hurwitz (1946) estimator for the population mean $\bar{Y}$ for matched portion of the sample on second occasion;

$$\bar{x}_m^* = \frac{m_1\bar{x}_{m_1} + m_2\bar{x}_{m_{h_2}}}{m}$$

is the Hansen and Hurwitz (1946) estimator for the population mean $\bar{X}$ for matched portion of the sample on second occasion;

$$\bar{x}_n = \sum_{i=1}^{n} x_i / n$$

is the estimate of the population mean $\bar{X}$ of the sample;

$$\bar{x}_m = \sum_{i=1}^{m} x_i / m$$

is the estimate of the population mean $\bar{X}$ on second occasion for the matched portion of the sample;

$$\bar{x}_{m_1} = \sum_{i=1}^{m_1} x_i / m_1$$

is the estimate of the population mean $\bar{X}_1$ on second occasion for the matched portion of the sample;

$$\bar{x}_{m_{h_2}} = \sum_{i=1}^{m_{h_2}} x_i / m_{h_2}$$

is the sub-sample mean of variable x based on m_{h_2} units on the second occasion;

$$\bar{y}_{m_1} = \sum_{i=1}^{m_1} y_i / m_1$$

is the estimate of the population mean $\bar{Y}_1$ on second occasion for the matched portion of the sample; and

$$\bar{y}_{m_{h_2}} = \sum_{i=1}^{m_{h_2}} y_i / m_{h_2}$$

is the sub-sample of variable y based on m_{h_2} units on the second occasion.

The variance of t_m (if fpc is ignored) to the first degree of approximation is given by

$$Var(t_m) = \left[\left(\frac{1}{m}-\frac{1}{n}\right)\left\{S_y^2 + (\lambda_1-\lambda_2)^2 S_x^2 - 2(\lambda_1-\lambda_2)\beta S_x^2\right\} + \frac{W_2(k-1)}{m}\left\{S_{y(2)}^2 + \lambda_1(\lambda_1 - 2\beta_{(2)})S_{x(2)}^2\right\} + \frac{1}{n}S_y^2\right] \quad (2.2)$$

where

$$W_2 = N_2 / N;$$

$$\beta = \rho(S_y / S_x);$$

$$\beta_{(2)} = \rho_{(2)}(S_{y(2)} / S_{x(2)});$$

$$k = (u_2 / u_{h_2}) = (m_2 / m_{h_2});$$

and ρ and $\rho_{(2)}$ are the correlation coefficient between the variables (y and x) and ($y_{(2)}$ and $x_{(2)}$);

$$S_y^2 = \sum_{i=1}^{N} (y_i - \bar{Y})^2 / (N-1)$$

denotes the population mean square of the variable y;

$$S_x^2 = \sum_{i=1}^{N} (x_i - \bar{X})^2 / (N-1)$$

denotes the population mean square of the variable x;

$$S_{y(2)}^2 = \sum_{i=1}^{N_2} (y_i - \bar{Y}_{(2)})^2 / (N_2-1)$$

denotes the population mean square pertaining to the non-response class of the variable y;

$$S_{x(2)}^2 = \sum_{i=1}^{N_2} (x_i - \bar{X}_{(2)})^2 / (N_2-1)$$

denotes the population mean square pertaining to the non-response class of the variable x.

Differentiating the variance of t_m, that is, $Var(t_m)$ at (2.2) with respect to λ_1 and λ_2, and equating to zero, results in the optimum values of λ_1 and λ_2 as

$$\lambda_1 = \rho_{(2)}\left(S_{y(2)}/S_{x(2)}\right) = \beta_{(2)}$$

and

$$\begin{aligned}\lambda_2 &= \left\{\rho_{(2)}\left(S_{y(2)}/S_{x(2)}\right)\right\} - \left\{\rho\left(S_y/S_x\right)\right\} \\ &= \left(\beta_{(2)} - \beta\right)\end{aligned}.$$

Substituting the optimum values of λ_1 and λ_2 in (2.1), results in the optimum estimate of the estimator t_m as

$$t_m^{(0)} = \bar{y}_m^* - \beta_{(2)}\left(\bar{x}_m^* - \bar{x}_m\right) - \beta\left(\bar{x}_m - \bar{x}_n\right), \tag{2.3}$$

with variance (ignoring fpc), the result is

$$Var\left(t_m^{(0)}\right) = \left[\left(\frac{1}{m} - \frac{1}{n}\right)\left(1-\rho^2\right)S_y^2 + \frac{W_2(k-1)}{m}\left(1-\rho_{(2)}^2\right)S_{y(2)}^2 + \frac{1}{n}S_y^2\right] \tag{2.4}$$

In practice $\beta_{(2)}$ and β are usually unknown, it lacks the practical utility of the optimum estimator $t_m^{(0)}$, thus it is advisable to replace $\beta_{(2)}$ and β by their consistent estimates $\hat{\beta}_{(2)}^*$ and $\hat{\beta}^*$ respectively in (2.3) to calculate an estimate of the population mean $\bar{Y}$ based on matched portion on second occasion as

$$\hat{t}_m^{(0)} = \bar{y}_m^* - \hat{\beta}_{(2)}^*\left(\bar{x}_m^* - \bar{x}_m\right) - \hat{\beta}^*\left(\bar{x}_m - \bar{x}_n\right), \tag{2.5}$$

where

$$\hat{\beta}_{(2)}^* = s_{xy(2)}^* \big/ s_{x(2)}^{*2},$$

$$\hat{\beta}^* = s_{xy}^* \big/ s_x^{*2},$$

$$s_{xy}^* = \frac{1}{(m-1)}\left\{\sum_{j=1}^{m_1} x_j y_j + k\sum_{j=1}^{m_{h2}} x_j y_j - m\bar{x}_m\bar{y}_m^*\right\},$$

$$s_x^{*2} = \frac{1}{(m-1)}\left\{\sum_{j=1}^{m_1} x_j^2 + k\sum_{j=1}^{m_{h2}} x_j^2 - m\bar{x}_m\bar{x}_m^*\right\},$$

$$s_{xy(2)}^* = \frac{1}{\left(m_{h_2}-1\right)}\left[\sum_{j=1}^{m_{h2}}\left(x_j - \bar{x}_{m_{h_2}}\right)\left(y_j - \bar{y}_{m_{h_2}}\right)\right],$$

$$s_{x(2)}^{*2} = \frac{1}{\left(m_{h_2}-1\right)}\sum_{j=1}^{m_{h2}}\left(x_j - \bar{x}_{m_{h_2}}\right)^2,$$

$$\bar{x}_{m_{h2}} = \frac{1}{m_{h_2}}\sum_{j=1}^{m_{h2}} x_j,$$

$$\bar{y}_{m_{h2}} = \frac{1}{m_{h_2}}\sum_{j=1}^{m_{h2}} y_j,$$

and

$$\bar{x}_m = \frac{1}{m}\sum_{j=1}^{m} x_j.$$

It can be shown to the first degree of approximation that

$$\begin{aligned}Var\left(\hat{t}_m^{(0)}\right) &= \\ &= Var\left(t_m^{(0)}\right) \\ &= \left[\begin{aligned}&\left(\frac{1}{m} - \frac{1}{n}\right)\left(1-\rho^2\right)S_y^2 + \\ &\frac{W_2(k-1)}{m}\left(1-\rho_{(2)}^2\right)S_{y(2)}^2 + \frac{1}{n}S_y^2\end{aligned}\right]\end{aligned} \tag{2.6}$$

where $Var\left(t_m^{(0)}\right)$ is given by (2.4) (Singh & Kumar, 2008).

Hence, an estimate of the population mean $\bar{Y}$ of the study variable y is constructed in the presence of non-response on the current occasion by combining the two independent estimates $\bar{y}_u^*$ and $\hat{t}_m^{(0)}$ with α an unknown constant as

$$T_{21} = \alpha \bar{y}_u^* + (1-\alpha)\hat{t}_m^{(0)}, \qquad (2.7)$$

where

$$\bar{y}_u^* = \frac{u_1 \bar{y}_{u_1} + u_2 \bar{y}_{u_{h_2}}}{u},$$

$$\bar{y}_{u_1} = \sum_{i=1}^{u_1} y_i / u_1,$$

and

$$\bar{y}_{u_{h_2}} = \sum_{i=1}^{u_{h_2}} y_i / u_{h_2}.$$

The variance of $\bar{y}_u^*$, the Hansen and Hurwitz (1946) estimator is

$$Var(\bar{y}_u^*) = \left(\frac{1}{u} - \frac{1}{N}\right) S_y^2 + \frac{W_2(k-1)}{u} S_{y(2)}^2. \qquad (2.8)$$

The variance of T_{21} at (2.7) to the first degree of approximation is given by

$$Var(T_{21}) = \alpha^2 Var(\bar{y}_u^*) + (1-\alpha)^2 Var(\hat{t}_m^{(0)}). \qquad (2.9)$$

Because, the variance of T_{21} in equation (2.9) is a function of unknown constant α, it is minimized with respect to α and subsequently the optimum value of α is obtained as

$$\alpha_{opt} = \frac{Var(\hat{t}_m^{(0)})}{Var(\bar{y}_u^*) + Var(\hat{t}_m^{(0)})}. \qquad (2.10)$$

Using the optimum value of α from (2.10) in (2.9), results in the optimum variance of T_{21} as

$$Var(T_{21})_{opt} = \frac{Var(\bar{y}_u^*) Var(\hat{t}_m^{(0)})}{Var(\bar{y}_u^*) + Var(\hat{t}_m^{(0)})}. \qquad (2.11)$$

Further, substituting the values from (2.4) and (2.8) in (2.11), the optimum variance of T_{21} is simplified as

$$Var(T_{21})_{opt} = \left(\frac{S_y^2 + W_2(k-1)S_{y(2)}^2}{n}\right) \left[\frac{\left\{ \left(1-q\rho^2\right) S_y^2 + A \right\}}{\left\{ \begin{array}{l} \left(1-q^2\rho^2\right) S_y^2 + \\ W_2(k-1)\left(1-q\rho_{(2)}^2\right) S_{y(2)}^2 \end{array} \right\}} \right], \qquad (2.12)$$

where

$$A = W_2(k-1)\left(1-\rho_{(2)}^2\right) S_{y(2)}^2.$$

To determine the optimum value of q so that population mean $\bar{Y}$ of study variable y may be estimated with maximum precision, minimize $Var(T_{21})_{opt}$ in (2.12) with respect to q and the optimum value of q is obtained as

$$q = \frac{\left(A + S_y^2\right) \pm \sqrt{\left(1-\rho^2\right) S_y^4 + W_2(k-1) \left\{ \begin{array}{l} A + \\ \left(2-\rho^2-\rho_{(2)}^2\right) S_y^2 \end{array} \right\} S_{y(2)}^2}}{\rho^2 S_y^2} = q_0 \qquad (2.13)$$

The real value of q_0 exists if

$$\left[\left(1-\rho^2\right) S_y^4 + W_2(k-1) \left\{ A + \left(2-\rho^2-\rho_{(2)}^2\right) S_y^2 \right\} S_{y(2)}^2 \right] \geq 0.$$

For certain situations, there might be two values of q_0 satisfying the above condition, hence when selecting a value of q_0, it should be remembered that the existence of q_0 depends on the limit $0 \leq q_0 \leq 1$; all other values of q_0 are inadmissible. In the case where both the values of q_0 are admissible, choose the minimum as q_0.

Further, substituting the value of q_0 from (2.13) in (2.12),

$$Var(T_{21})_{opt^*} = \left(\frac{S_y^2 + W_2(k-1)S_{y(2)}^2}{n}\right)\left[\frac{\begin{Bmatrix}(1-q_0\rho^2)S_y^2 + \\ A\end{Bmatrix}}{\begin{Bmatrix}(1-q_0^2\rho^2)S_y^2 + \\ W_2(k-1)(1-q_0\rho_{(2)}^2)S_{y(2)}^2\end{Bmatrix}}\right], \tag{2.14}$$

where $Var(T_{21})_{opt^*}$ is the optimum variance of T_{21} with respect to both α and q.

Efficiency Comparison

To determine the effect of non-response in successive sampling, calculate the percent relative loss in efficiency of T_{21} with respect to the estimator under the same circumstances but in absence of non-response. The estimator is defined as

$$T_{21}^* = \varphi\bar{y}_u + (1-\varphi)t_{ld};$$

where

$$\bar{y}_u = \sum_{i=1}^{u} y_i/u \ ; \ t_{ld} = \bar{y}_m + \hat{\beta}(\bar{x}_n - \bar{x}_m),$$

$$\hat{\beta} = \left\{\sum_{i=1}^{m}(x_i - \bar{x}_m)(y_i - \bar{y}_m)\right\}\Big/\sum_{i=1}^{m}(x_i - \bar{x}_m)^2$$

and φ is an unknown constant to be determined under certain criterion. Because T_{21}^* is an unbiased estimator of $\bar{Y}$ and is based on two independent samples the covariance terms vanishes, therefore following the procedure of Sukhatme, et al. (1984), the optimum variance of T_{21}^* can be obtained as

$$Var(T_{21}^*)_{opt^*} = \frac{(1-q_1\rho^2)S_y^2}{n(1-q_1^2\rho^2)}, \tag{3.1}$$

where

$$q_1 = \frac{1 \pm \sqrt{(1-\rho^2)}}{\rho^2}.$$

To select the optimum value of q_1, it is important to remember that $0 \le q_1 \le 1$, however, if both values of q_1 are admissible, then the least of two values of q_1 should be chosen. Thus, the percentage loss in precision of T_{21}^* with respect to T_{21} both at optimality condition is given by

$$L = \frac{Var(T_{21})_{opt^*} - Var(T_{21}^*)_{opt^*}}{Var(T_{21})_{opt^*}} \times 100. \tag{3.2}$$

Results

Table 1 shows the percentage loss in precision observed wherever the optimum value of q exists when non-response is taken into account at current occasion. For fixed values of ρ, $\rho_{(2)}$, $(k-1)$ and W_2, for $S_y < S_{y(2)}$, the loss in precision decreases with the increase in the value of S_y; for $S_y > S_{y(2)}$, the loss in precision shows negative values with the decrease in the value of $S_{y(2)}$; and for $S_y = S_{y(2)}$, the loss in precision remains constant. For fixed values of S_y, $S_{y(2)}$, $(k-1)$ and W_2, the loss in precision shows negative values for $\rho < \rho_{(2)}$ with the decrease in the value of ρ and for $\rho > \rho_{(2)}$, the loss in precision decreases with increase in the value of $\rho_{(2)}$ while it remains constant for $\rho = \rho_{(2)}$. Table 2 shows that, for the increased values of W_2, the percentage loss in precision increases and it decreases with the decreases in the value of $(k-1)$.

Table 1: Percentage Loss in Precision of T_{21}^{*} over T_{21}
for Different Values of ρ , $\rho_{(2)}$, S_y and $S_{y(2)}$

$\rho = 0.8,\ \rho_{(2)} = 0.2,\ (k-1) = 1, W_2 = 0.8$

$S_y < S_{y(2)}$		L	$S_y > S_{y(2)}$		L	$S_y = S_{y(2)}$		L
S_y	$S_{y(2)}$		S_y	$S_{y(2)}$		S_y	$S_{y(2)}$	
0.2	0.8	90.59	0.8	0.7	*	0.8	0.8	22.07
0.3	0.8	80.43	0.8	0.6	*	0.8	0.8	22.07
0.4	0.8	68.44	0.8	0.5	*	0.8	0.8	22.07

$(k-1) = 0.5, W_2 = 0.7, S_y = 0.4, S_{y(2)} = 0.8$

$\rho < \rho_{(2)}$		L	$\rho > \rho_{(2)}$		L	$\rho = \rho_{(2)}$		L
ρ	$\rho_{(2)}$		ρ	$\rho_{(2)}$		ρ	$\rho_{(2)}$	
0.7	0.8	*	0.7	0.3	39.01	0.8	0.8	16.67
0.6	0.8	*	0.7	0.4	37.04	0.8	0.8	16.67
0.5	0.8	*	0.7	0.5	33.94	0.8	0.8	16.67

Table 2: Percentage Loss in Precision of T_{21}^{*} over T_{21}
for Different Values of W_2 and $(k-1)$

$\rho = 0.8,\ \rho_{(2)} = 0.2,\ (k-1) = 0.5,\ S_y = 0.2, S_{y(2)} = 0.7$		$\rho = 0.8, \rho_{(2)} = 0.4,\ W_2 = 0.3,\ S_y = 0.3, S_{y(2)} = 0.8$	
W_2	L	$(k-1)$	L
0.3	19.32	1.5	65.95
0.4	37.83	1.0	54.38
0.5	49.40	0.5	30.03

A tangible idea regarding obtaining cost saving through mail surveys in the context of successive sampling on two occasions for different assumed values of ρ , $\rho_{(2)}$, S_y , $S_{y(2)}$, W_2 and k is shown in Tables 3 and 4. Also, let $N = 500$ and $n = 50$ and $c_1/c_0 = 4$, $c_2 = 45$, where c_0 is the cost per unit for mailing a questionnaire (Rs. 1.00), c_1 is the cost per unit of processing the results from the first attempt respondents (Rs. 4.00), c_2 is the cost per unit for collecting data through personal interview (Rs. 45.00). Denote C= total cost incurred in collecting the data by personal interview from the whole sample, that is, when there is no non-response. Assuming that the cost incurred on data collection for the matched and unmatched portion of the sample are same and also cost incurred on data collection on both the occasions is same, the cost function in this case is given by:

$$C = 2nc_2 . \tag{3.3}$$

Setting the values of n and c_2 in (3.3), the total cost work out to be Rs. 4500.00.

Further, let n_1 denote number of units which respond at the first attempt and n_2 denote number of units which do not respond. The cost function for the case when there is non-response on both occasions is given by

$$C_1 = 2\{c_0 n + c_1 n_1 + (c_2 n_2 / k)\}.$$

The expected cost is given by

$$E(C_1) = 2n_0 \{c_0 + c_1 W_1 + (c_2 W_2 / k)\} = C_1^* ,$$

where

$$W_1 = N_1 / N$$

and

$$W_2 = N_2 / N ,$$

such that

$$W_1 + W_2 = 1 ,$$

$$n_0 = \frac{n(1 - q_1^2 \rho^2)(1 - B)(S_y^2 + W_2 (k-1) S_{y(2)}^2)}{(1 - q_1 \rho^2)(1 - q_0 B) S_y^2} ,$$

and

$$B = \frac{\{q_0 \rho^2 S_y^2 + W_2 (k-1) \rho_{(2)}^2 S_{y(2)}^2\}}{\{S_y^2 + W_2 (k-1) S_{y(2)}^2\}} .$$

From Table 3 it is noted that for fixed values of ρ , $\rho_{(2)}$, $(k-1)$ and W_2, for the case $S_y < S_{y(2)}$, the cost savings increases with decreases in the value of S_y. For the case $S_y > S_{y(2)}$, the cost savings decreases with the decreases in the value of $S_{y(2)}$, and for the case $S_y = S_{y(2)}$, it remains constant. Further, for the fixed values of $(k-1)$, W_2, S_y and $S_{y(2)}$, for the case $\rho < \rho_{(2)}$, the cost savings decreases with the decreases in the value of ρ and for the case $\rho > \rho_{(2)}$, it also decreases with the increase in the value of $\rho_{(2)}$ but it remains constant for the case $\rho = \rho_{(2)}$. It is to be observed from Table 4 that increases in the values of W_2 and decreases in the value of $(k-1)$, the cost savings increase respectively.

References

Adhvaryu, D. (1978). Successive sampling using multi auxiliary information. *Sankhya*, *C*, *2*, 167-173.

Arnab, R. and Okafor, F. C. (1992). A note on double sampling over two occasions. *Pakistan Journal of Statistics*, *B*, 8(3), 9-18.

Artes, E. and Garcia, A. (2001). Estimation of current population ratio in successive sampling. *Journal of the Indian Society of Agricultural Statistics*, *54*(*3*), 342-354.

Choudhary, R. K., Bathla, H. V. L., & Sud, U. C. (2004). On non-response in sampling on two occasions. *Journal of the Indian Society of Agricultural Statistics*, *58*, 331-343.

Cochran, W. G. (1963). *Sampling Techniques*. *2nd Ed*. New York: Wiley.

Eckler, A. R. (1955). Rotation sampling. *Annals of Mathematical Statistics*, *26*, 664-685.

Garcia, A., & Artes, E. (2002). Improvement on estimating current population ratio in successive sampling. *Brazilian Journal of Probability Statistics*, *16*, 107-122.

Gorden, L. (1983). Successive sampling in finite populations. *The Annals of Statistics*, *11*, 702-706.

Table 3: Sample Sizes and Corresponding Expected Cost of Survey, which Give Equal Precision of Proposed Estimate T_{21}^* over T_{21} for Different Values of ρ , $\rho_{(2)}$, S_y and $S_{y(2)}$

$\rho = 0.4, \rho_{(2)} = 0.8, (k-1) = 0.5, W_2 = 0.2,$

$S_y < S_{y(2)}$		C_1^*	$S_y > S_{y(2)}$		C_1^*	$S_y = S_{y(2)}$		C_1^*
S_y	$S_{y(2)}$		S_y	$S_{y(2)}$		S_y	$S_{y(2)}$	
0.7	0.8	2354.88	0.8	0.7	2288.02	0.6	0.6	2317.93
0.6	0.8	2407.35	0.8	0.6	2260.93	0.6	0.6	2317.93
0.5	0.8	2483.36	0.8	0.5	2237.14	0.6	0.6	2317.93

$(k-1) = 0.5, W_2 = 0.5, S_y = 0.7, S_{y(2)} = 0.8$

$\rho < \rho_{(2)}$		C_1^*	$\rho > \rho_{(2)}$		C_1^*	$\rho = \rho_{(2)}$		C_1^*
ρ	$\rho_{(2)}$		ρ	$\rho_{(2)}$		ρ	$\rho_{(2)}$	
0.5	0.6	3496.39	0.7	0.3	3762.69	0.6	0.6	3612.65
0.4	0.6	3476.17	0.7	0.4	3732.99	0.6	0.6	3612.65
0.3	0.6	3450.64	0.7	0.5	3693.69	0.6	0.6	3612.65

Table 4: Sample Sizes and Corresponding Expected Cost of Survey, which Give Equal Precision of Proposed Estimate T_{21}^* over T_{21} for Different Values of W_2 and $(k-1)$

$\rho = 0.2, \rho_{(2)} = 0.7,$ $(k-1) = 0.5,\ S_y = 0.8, S_{y(2)} = 0.4$		$\rho = 0.4, \rho_{(2)} = 0.5,$ $W_2 = 0.3,\ S_y = 0.7, S_{y(2)} = 0.8$	
W_2	C_1^*	$(k-1)$	C_1^*
0.2	2260.11	1.5	1802.75
0.3	3161.32	1.0	2213.82
0.4	4075.08	0.5	3523.25

Hansen, M. H., & Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, *41*,517-529.

Jessen, R. J. (1942). Statistical investigation of a sample survey for obtaining farm facts. *Iowa Agricultural Experiment Statistical Research Bulletin*, *304*.

Miranda, M. D. M., Garcia, M. R., Cebrian, A. A., Montoya, Y. R., & Aguilera, S. G. (2007). Quantile estimation under successive sampling. *Computational Statistics*, *20*(*3*), 385-399.

Narain, R. D. (1953. On the recurrence formula in sampling on successive occasions. *Journal of the Indian Society of Agricultural Statistics*, *5*, 96-99.

Okafor, F. C. (2005). Sub-sampling the non-respondents in two-stage sampling over two successive occasions. *Journal of the Indian Statistical Association*, *43*(1), 33-49.

Okafor, F. C., & Lee, H. (2000): Double sampling for ratio and regression estimation with sub sampling the non-respondent. *Survey Methodology*, *26*(*2*), 183-188.

Patterson, H. D. (1950). Sampling on successive occasions with partial replacement of units. *Journal of the Royal Statistical Society*, *Sec. B*, *12*, 241-255.

Sen, A. R. (1979). Sampling theory on repeated occasions with ecological applications. *Sampling Biological Populations*, 315-328.

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

Singh, G. N., & Priyanka, K. (2007). Estimation of population mean at current occasion in successive sampling under a super-population model. *Model Assisted Statistics and Applications*, *2*, 189-200.

Singh, G. N., & Singh, V. K. (2001). On the use of auxiliary information in successive sampling. *Journal of the Indian Society for Agricultural Statistics*, *54*(*1*), 1-12.

Singh, H. P., & Vishwakarma, G., K. (2007). A general class of estimators in successive sampling. *METRON*, LXV, *2*, 201-227.

Singh, H. P., & Kumar, S. (2010). Estimation of mean in presence of non-response using two phase sampling scheme. *Statistical Papers*, 50, 559-582.

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

Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S., & Asok, C. (1984). *Sampling theory of surveys with applications*. Ames, IA, USA: Iowa State University Press and New Delphi, India: Indian Society of Agricultural Statistics.