

Assessment of Non-Response under Ratio Method of Imputation in Two-Occasion Successive Sampling

G. N. Singh¹, D. Majhi¹, S. Prasad¹ and F. Homa¹

¹Department of Applied Mathematics, Indian School of Mines, Dhanbad-826004, India and E.
mail: gnsingh_ism@yahoo.com

Received 11 August 2012

Accepted 16 August 2013

Abstract

This paper considers the problem of estimation of the population mean of the study character under ratio method of imputation when some observations in the sample data are missing at random and the information on an auxiliary variable is readily available on both the occasions in two-occasion successive sampling. Estimator for the population mean on the second (current) has been proposed and the expressions of bias and mean square error are derived. Special cases and optimum replacement policies are discussed. Empirical studies are carried out and recommendations are made.

Key words: Non-response, imputation, successive sampling, auxiliary variable, bias, mean square error, optimum replacement policy.

Mathematics subject classification: 62D05

1. Introduction

Non-response is one of the major problems encountered by the survey statisticians. Longitudinal surveys (surveys on successive occasions) are more prone to this problem than single occasion survey. For example, in agriculture yield surveys, it might be possible that crop on certain plots are destroyed due to some natural calamities or disease so that yield on these plots are impossible to be measured. Such non-response may have different patterns and causes. Statisticians have recognized for some time that if the suitable information about the nature of non-response in the population is unknown, the inference concerning population parameters may be spoiled. Many

methods are used to reduce the negative impact of non-response in sample surveys. Imputation is one which deals with the filling up method of incomplete data for adapting the standard analytical model in statistics. It is typically used when needed to substitute missing item values with certain fabricated values in the sample surveys. To deal with missing values effectively, Rubin (1976), Sande (1979) and Kalton *et al.* (1981) have suggested imputation methods which make incomplete data sets structurally complete and its analysis simple. Further useful and intelligible imputation methods were suggested by Kalton and Kasprzyk (1982), Singh and Singh (1991), Lee *et al.* (1994, 1995), Singh and Horn (2000), Singh and Deo (2003), Ahmed *et al.* (2006), Singh (2009) and Singh *et al.* (2010) to deal with the problems of non-response in sample surveys.

Motivated with the above works, Singh *et al.* (2008), Singh *et al.* (2009), Singh and Karna (2010), Singh and Priyanka (2010) and Singh *et al.* (2012) have suggested impressive imputation methods to deal with the problems of non-response in two-occasion successive sampling. There may be situations where the non-response may occur on both the occasions, following this argument; the objective of the present work is to study the effect of non-response on both the occasions in two-occasion successive sampling. Utilizing the information on a dynamic (changing over occasions) auxiliary variable, ratio method of imputation has been used to cope with the problems of non-response on both the occasions in two-occasion successive sampling. Estimator for the current occasion is derived as a particular case when there is non-response either on the first occasion or on the second occasion. The performance of the proposed estimator is compared in two different situations: with and without non-response. Empirical studies are carried out and suitable interpretations are made in the present work.

2. Notations and Sample Structures on Two-Occasion

Let $U = (U_1, U_2, \dots, U_N)$ be the finite population of size N , which has been sampled over two occasions. The character under study is denoted by x (y) on the first (second) occasion. It is assumed that the information on an auxiliary variable z_h ($h=1,2$) (with known population mean) is available on h^{th} ($h = 1, 2$) occasion. We assume that non-response occurs at both the occasions. A simple random sample (s_n) (without replacement) of n units is drawn on the first occasion. Let the number of responding units out of sampled n units, which are drawn at the first occasion, be

denoted by r_1 , the set of responding units by R_1 and that of non-responding units by R_1^c . A random sub-sample (s_m) of $m = n\lambda$ units is retained (matched) for its use on the second occasion from the units which responded at the first occasion and it is assumed that these matched units will respond at the second (current) occasion as well. A fresh simple random sample (s_u) (without replacement) of $u = (n - m) = n\mu$ units is drawn on the second occasion from the entire population so that the sample size on the second occasion remains n . Let the number of responding units out of sampled u units, which are drawn afresh on the current occasion, be denoted by r_2 , the set of responding units by R_2 and that of non-responding units by R_2^c . λ and μ ($\lambda + \mu = 1$) are the fractions of the matched and fresh samples, respectively, on the current occasion. For every unit $i \in R_j$ ($j = 1, 2$) the value x_i (y_i) is observed, but for the units $i \in R_j^c$ ($j = 1, 2$) the x_i (y_i) values are missing and instead the imputed values are derived.

The following notations are used hereafter:

$\bar{X}, \bar{Y}, \bar{Z}_1, \bar{Z}_2$: The population means of the variables x, y, z_1 and z_2 respectively.

$\bar{x}_n, \bar{z}_{1n}, \bar{y}_m, \bar{x}_m, \bar{z}_{2m}, \bar{x}_{r_1}, \bar{y}_{r_2}, \bar{y}_u, \bar{z}_{2u}, \bar{z}_{1r_1}, \bar{z}_{2r_2}$: The sample means of the respective variables based on the sample sizes shown in suffices.

$\rho_{yx}, \rho_{xz_1}, \rho_{xz_2}, \rho_{yz_1}, \rho_{yz_2}, \rho_{z_1z_2}$: The correlation coefficients between the variables shown in suffices.

$S_x^2, S_y^2, S_{z_1}^2, S_{z_2}^2$: The population variance of the variables x, y, z_1 and z_2 respectively.

$f_1 \left(= \frac{r_1}{n} \right), f_2 \left(= \frac{r_2}{u} \right)$: The fractions of respondents in the samples of sizes n and u respectively.

$t_1 (= 1 - f_1), t_2 (= 1 - f_2)$: The fractions of non-respondents in the samples of sizes n and u respectively.

3. Formulation of the Estimator

We consider two independent estimators of the population mean \bar{Y} on the second (current) occasion that use information on an auxiliary variable $z_1(z_2)$. Estimators are proposed to cope with the problems of non-response on both the occasions. The first estimator is based on the sample of size u drawn afresh on the second (current) occasion. The missing values are replaced by the derived imputed values generated through the ratio method of imputation. The following ratio method of imputation has been considered for the missing data on the second (current) occasion:

$$y_{\cdot i} = \begin{cases} y_i \left(\frac{\bar{Z}_2}{\bar{Z}_{2u}} \right) & \text{if } i \in R_2 \\ \delta_2 \left(\frac{\bar{Z}_2}{\bar{Z}_{2u}} z_{2i} \right) & \text{if } i \in R_2^c \end{cases} \quad (1)$$

$$\text{where } \delta_2 = \frac{\sum_{i \in R_2} y_i}{\sum_{i \in R_2} z_{2i}} = \frac{\bar{y}_{r_2}}{\bar{z}_{2r_2}}, \bar{y}_{r_2} = \frac{1}{r_2} \sum_{i \in R_2} y_i \text{ and } \bar{z}_{2r_2} = \frac{1}{r_2} \sum_{i \in R_2} z_{2i}.$$

Following the above imputation method, the estimator for \bar{Y} based on the fresh sample of size u drawn on the second (current) occasion is given by

$$T_{1u} = \frac{1}{u} \sum_{i \in s_u} y_{\cdot i} = \frac{1}{u} \left[\sum_{i \in R_2} y_{\cdot i} + \sum_{i \in R_2^c} y_{\cdot i} \right] = \frac{\bar{y}_{r_2}}{\bar{z}_{2r_2}} \bar{Z}_2. \quad (2)$$

The second estimator based on the sample of size m , which is common to both the occasions and utilizes the information from the first occasion. Since, there is non-response on the first occasion as well; therefore, the missing values are replaced by the derived imputed values using the following ratio method of imputation. Hence, the data after imputation takes the following form:

$$x_{\cdot i} = \begin{cases} x_i \left(\frac{\bar{Z}_1}{\bar{Z}_{1n}} \right) & \text{if } i \in R_1 \\ \delta_1 \left(\frac{\bar{Z}_1}{\bar{Z}_{1n}} z_{1i} \right) & \text{if } i \in R_1^c \end{cases} \quad (3)$$

$$\text{where } \delta_1 = \frac{\sum_{i \in R_1} x_i}{\sum_{i \in R_1} z_{1i}} = \frac{\bar{x}_{r_1}}{\bar{z}_{1r_1}}, \bar{x}_{r_1} = \frac{1}{r_1} \sum_{i \in R_1} x_i \text{ and } \bar{z}_{1r_1} = \frac{1}{r_1} \sum_{i \in R_1} z_{1i}.$$

Under the above imputation method the estimator based on sample of size n on the first occasion is given by

$$\bar{x}_n^* = \frac{1}{n} \sum_{i \in s_n} x_{\cdot i} = \frac{1}{n} \left[\sum_{i \in R_1} x_{\cdot i} + \sum_{i \in R_1^c} x_{\cdot i} \right] = \left(\frac{\bar{x}_{r_1}}{\bar{z}_{1r_1}} \right) \bar{z}_1. \quad (4)$$

Therefore, the second estimator for estimating the population mean \bar{Y} on the second (current) occasion is a new ratio type estimator based on a sample of size $m (= n\lambda)$ common to both occasions and utilizes the ratio method of imputation for imputing the missing values on the first occasion and it is defined as

$$T_{1m} = \bar{y}_m \left(\frac{\bar{x}_n^*}{\bar{x}_m} \right) \left(\frac{\bar{z}_2}{\bar{z}_{2m}} \right). \quad (5)$$

Considering the convex linear combination of estimators T_{1u} and T_{1m} ; the final estimator T for estimating the population mean \bar{Y} at the second (current) occasion is proposed as:

$$T = \phi T_{1u} + (1 - \phi) T_{1m} \quad (6)$$

where ϕ is an unknown real constant to be determined so as to minimize the mean square error of the estimator T .

4. Properties of the Proposed Estimator

Since, T_{1u} and T_{1m} both are ratio and ratio type estimators, they are biased for population mean \bar{Y} , therefore, the resulting estimator T defined in equation (6) is also a biased estimator of \bar{Y} . Using the MCAR response mechanism for given r_1, r_2, u, m and n , the bias $B(\cdot)$ and mean square error $M(\cdot)$ of the estimator T are derived (ignoring fpc) up-to the first order of approximations and shown in the following theorems:

Theorem 4.1: The bias of the estimator T in estimating the population mean \bar{Y} , to the first order of approximations is

$$B(T) = \phi B(T_{lu}) + (1 - \phi) B(T_{lm}) \quad (7)$$

$$\text{where } B(T_{lu}) = \bar{Y} \left(\frac{1}{r_2} \right) [C_{z_2}^2 - \rho_{yz_2} C_y C_{z_2}] \quad (8)$$

$$\text{and } B(T_{lm}) = \bar{Y} \left[\left(\frac{1}{m} \right) (C_x^2 + C_{z_2}^2 + \rho_{xz_2} C_x C_{z_2} - \rho_{yz_2} C_y C_{z_2} - \rho_{yx} C_y C_x) \right] + \left[\left(\frac{1}{r_1} \right) (C_{z_1}^2 - C_x^2 + \rho_{z_1 z_2} C_{z_1} C_{z_2} + \rho_{yx} C_y C_x - \rho_{yz_1} C_y C_{z_1} - \rho_{xz_2} C_x C_{z_2}) \right] \quad (9)$$

Theorem 4.2: The mean square error of the estimator T of the population mean \bar{Y} , to the first order of approximations and ignoring finite population corrections is given by

$$M(T) = \phi^2 M(T_{lu}) + (1 - \phi)^2 M(T_{lm}) \quad (10)$$

$$\text{where } M(T_{lu}) = \frac{1}{r_2} \{ 2(1 - \rho_{yz_2}) \} S_y^2 \quad (11)$$

$$\text{and } M(T_{lm}) = \left[\frac{1}{m} \{ 3 + 2(\rho_{xz_2} - \rho_{yz_2} - \rho_{yx}) \} + \frac{1}{r_1} \{ 2(\rho_{z_1 z_2} + \rho_{yx} - \rho_{xz_2} - \rho_{yz_1}) \} \right] S_y^2 \quad (12)$$

Remark 4.1: Since the estimators T_{lu} and T_{lm} are based on two different non-overlapping samples, the covariance type term is of $o(N^{-1})$, hence it is ignored for large population size.

Remark 4.2: The expression of mean square error is derived under assumption that the coefficients of variation of the variables x , y , z_1 and z_2 are approximately equal.

4. 2. 1. Minimum Mean Square error of the Estimator T

Since the mean square error of the estimator T in equation (10) is the function of the unknown constant ϕ , therefore, it is minimized with respect to ϕ and subsequently the optimum value of ϕ (say $\phi_{opt.}$) is obtained as

$$\phi_{\text{opt.}} = \frac{M(T_{\text{lm}})}{M(T_{\text{lu}}) + M(T_{\text{lm}})}. \quad (13)$$

Now substituting the value of $\phi_{\text{opt.}}$ in equation (10), we have the optimum mean square error of the estimator T as

$$M(T)_{\text{opt.}} = \frac{M(T_{\text{lu}}) \cdot M(T_{\text{lm}})}{M(T_{\text{lu}}) + M(T_{\text{lm}})}. \quad (14)$$

Further, substituting the values from equations (11) and (12) in equation (14) the simplified value of $M(T)_{\text{opt.}}$ is shown in theorem 4.3.

Theorem 4.3: The $M(T)_{\text{opt.}}$ is derived as

$$M(T)_{\text{opt.}} = \frac{[A(f_1 B + C) - \mu A C] S_y^2}{[f_1 A + \mu D - f_2 C \mu^2] n} \quad (15)$$

$$\text{where } A = \{2(1 - \rho_{yz_2})\}, B = \{3 + 2(\rho_{xz_2} - \rho_{yz_2} - \rho_{yx})\}, C = \{2(\rho_{z_1 z_2} + \rho_{yx} - \rho_{xz_2} - \rho_{yz_1})\}$$

$$\text{and } D = f_1 f_2 B + f_2 C - f_1 A.$$

The $M(T)_{\text{opt.}}$ derived in equation (15) is the function of μ . To estimate the population mean on each occasion, a good choice for μ is 1 (the case of no matching) while for estimating the change from one occasion to the other, μ should be 0 (the case of complete matching). But to design a strategy that would be efficient for both the problems simultaneously, the optimum choice of μ is desired.

5. Optimum Replacement Policy

The $M(T)_{\text{opt.}}$ in equation (15) is a function of μ (fraction of fresh sample on the current occasion), which is an important factor in reducing the cost of the survey. It is also the functions of f_1 , f_2 and various correlations ρ_{yx} , ρ_{xz_1} , ρ_{xz_2} , ρ_{yz_1} , ρ_{yz_2} and $\rho_{z_1 z_2}$. To get the optimum value of μ , the $M(T)_{\text{opt.}}$ in equation (15) is minimized with respect to μ for the fixed values of f_1 , f_2 and correlations. The minimum value of μ is obtained as

$$\hat{\mu} = \frac{f_2 AC(f_1 B + C) \pm \sqrt{\{f_2 AC(f_1 B + C)\}^2 - f_2 AC^2 \{f_1 A^2 C + AD(f_1 B + C)\}}}{f_2 AC^2}. \quad (16)$$

From equation (16), it is clear that the real values of $\hat{\mu}$ exist, iff, the quantity under square root is greater than or equal to zero. For any combination of $\rho_{yx}, \rho_{xz_1}, \rho_{xz_2}, \rho_{yz_1}, \rho_{yz_2}$ and $\rho_{z_1 z_2}$, which satisfy the condition of real solution, two real values of $\hat{\mu}$ are possible. Hence, while choosing the value of $\hat{\mu}$, it should be remembered that $0 \leq \hat{\mu} \leq 1$, all other values of $\hat{\mu}$ are inadmissible. If both the values are admissible lowest one will be the best choice. Substituting the admissible value of $\hat{\mu}$ say $\mu^{(0)}$ from equation (16) into equation (15), we have the optimum value of mean square error of T , which is shown below

$$M(T)_{\text{opt.}^*} = \frac{[A(f_1 B + C) - AC\mu^{(0)}] S_y^2}{[f_1 A + D\mu^{(0)} - f_2 C\mu^{(0)^2}] n}. \quad (17)$$

6. Special Cases

Case 6.1: When there is Non-Response only at the First Occasion

The estimator for the population mean \bar{Y} at the current occasion can be obtained as

$$T^* = \phi^* T_{1u}^* + (1 - \phi^*) T_{1m}$$

where $T_{1u}^* = \frac{\bar{y}_u}{\bar{z}_{2u}} \bar{Z}_2$ and T_{1m} is defined in equation (5) and ϕ^* is an unknown real constant to be

determined so as to minimize the mean square error of the estimator T^* .

The optimum fraction of the fresh sample to be drawn in this case is obtained as

$$\hat{\mu} = \frac{AC(f_1 B + C) \pm \sqrt{\{AC(f_1 B + C)\}^2 - (AC^2) \{f_1 A^2 C + AE(f_1 B + C)\}}}{AC^2}, \text{ where } E = f_1 (B - A) + C$$

and the minimum mean square error of the estimator T^* at the admissible value of $\hat{\mu}$ (say $\mu_1^{(0)}$) is derived as

$$M(T^*)_{\text{opt.}^*} = \frac{[A(f_1 B + C) - AC\mu_1^{(0)}] S_y^2}{[f_1 A + E\mu_1^{(0)} - C\mu_1^{(0)2}] n} \quad (18)$$

Case 6.2: When there is Non-Response only at the Second (Current) Occasion

The estimator for the population mean \bar{Y} at the current occasion can be obtained as $T^{**} = \phi^{**} T_{lu} + (1 - \phi^{**}) T_{lm}^*$, where T_{lu} is defined in equation (2) and

$T_{lm}^* = \bar{y}_m \left(\frac{\bar{x}_n^{**}}{\bar{x}_m} \right) \left(\frac{\bar{z}_2}{\bar{z}_{2m}} \right)$; where $\bar{x}_n^{**} = \left(\frac{\bar{x}_n}{\bar{z}_{1n}} \right) \bar{z}_1$. ϕ^{**} is an unknown real constant to be determined

so as to minimize the mean square error of the estimator T^{**} .

The optimum fraction of the fresh sample to be drawn in this case is obtained as

$$\hat{\mu} = \frac{f_2 AC(B + C) \pm \sqrt{\{f_2 AC(B + C)\}^2 - f_2 AC^2 \{A^2 C + AF(B + C)\}}}{f_2 AC^2}, \text{ where } F = f_2 (B + C) - A$$

and the minimum mean square error of the estimator T^{**} at the admissible value of $\hat{\mu}$ (say $\mu_2^{(0)}$) is given as

$$M(T^{**})_{\text{opt.}^*} = \frac{[A(B + C) - AC\mu_2^{(0)}] S_y^2}{[A + F\mu_2^{(0)} - f_2 C\mu_2^{(0)2}] n} \quad (19)$$

7. Efficiency Comparison

The percent relative loss in efficiency of the estimator T with respect to the estimator under the same circumstances but under the complete information (with no missing data) has been obtained to judge the effect of non-response on the precision of estimates under the two-occasion successive (rotation) sampling.

Consider the following estimator of \bar{Y}

$$\tau = \psi T_{lu}^* + (1 - \psi) T_{lm}^* \quad (20)$$

where ψ is an unknown real constant to be determined by the minimization of the mean square error of the estimator τ .

Since, the estimator τ is based on two different non-overlapping samples and the covariance type term is $o(N^{-1})$, hence for large population size covariance type term is ignored while deriving the mean square error of the estimator τ . Following Sukhatme *et al.* (1984), the optimum mean square error of τ is given by

$$M(\tau)_{\text{opt.}^*} = \frac{[A(B+C) - AC\mu_3^{(0)}] S_y^2}{[A + G\mu_3^{(0)} - C\mu_3^{(0)2}] n} \quad (21)$$

where $G = B + C - A$ and $\mu_3^{(0)}$ is the admissible value of $\hat{\mu}$ for the estimator τ ,

$$\text{where } \hat{\mu} = \frac{AC(B+C) \pm \sqrt{\{AC(B+C)\}^2 - (AC^2)(A^2C + AG(B+C))}}{AC^2}. \quad (22)$$

Remark 7.1: The admissible value of $\hat{\mu}$ (i.e. $\mu_3^{(0)}$) in equation (22) is obtained in the similar manner as $\mu^{(0)}$ is derived.

The percent relative loss in precision of the estimators T , T^* and T^{**} with respect to the estimator τ under their respective optimality conditions are given by

$$L_1 = \frac{M(T)_{\text{opt.}^*} - M(\tau)_{\text{opt.}^*}}{M(\tau)_{\text{opt.}^*}} \times 100, \quad L_2 = \frac{M(T^*)_{\text{opt.}^*} - M(\tau)_{\text{opt.}^*}}{M(T^*)_{\text{opt.}^*}} \times 100$$

and $L_3 = \frac{M(T^{**})_{\text{opt.}^*} - M(\tau)_{\text{opt.}^*}}{M(T^{**})_{\text{opt.}^*}} \times 100$ respectively.

7.1 Empirical Study

The expressions of the minimum μ and the percent relative losses L_1 , L_2 and L_3 are in terms of population correlation coefficients. There are six correlations ρ_{yx} , ρ_{xz_1} , ρ_{xz_2} , ρ_{yz_1} , ρ_{yz_2} and $\rho_{z_1z_2}$ are involved in the expressions of μ , L_1 , L_2 and L_3 . To present the empirical results in tabular form we assumed the assumption $\rho_{xz_1} = \rho_{xz_2} = \rho_{yz_1} = \rho_{yz_2} = \rho_0$. Hence under above assumption the optimum admissible values of μ , L_1 , L_2 and L_3 have been computed for different choices of positive correlations $(\rho_{yx}, \rho_0, \rho_{z_1z_2})$, t_1 and t_2 and shown in Tables 1-3.

Table-1

Percent relative loss L_1 in precision of T with respect to τ for $\rho_{z_1z_2} = 0.5$.

| ρ_0 | | | 0.7 | | | 0.8 | | | 0.9 | | |
|----------|-------|-------------|-------------|---------------|-------|-------------|---------------|-------|-------------|---------------|-------|
| t_1 | t_2 | ρ_{yx} | $\mu^{(0)}$ | $\mu_3^{(0)}$ | L_1 | $\mu^{(0)}$ | $\mu_3^{(0)}$ | L_1 | $\mu^{(0)}$ | $\mu_3^{(0)}$ | L_1 |
| 0.05 | 0.05 | 0.3 | 0.23 | 0.13 | 0.25 | 0.18 | 0.10 | -2.25 | 0.23 | 0.18 | -1.68 |
| | | 0.5 | 0.83 | 0.79 | 4.67 | 0.40 | 0.36 | 2.51 | 0.29 | 0.26 | 0.61 |
| | | 0.7 | * | - | - | 0.75 | 0.74 | 4.51 | 0.38 | 0.37 | 2.39 |
| | | 0.9 | * | - | - | - | - | - | 0.54 | 0.54 | 3.77 |
| | 0.10 | 0.3 | 0.30 | 0.13 | 3.92 | 0.22 | 0.10 | 1.20 | 0.24 | 0.18 | 1.95 |
| | | 0.5 | 0.89 | 0.79 | 9.64 | 0.44 | 0.36 | 6.52 | 0.30 | 0.26 | 4.36 |
| | | 0.7 | * | - | - | 0.79 | 0.74 | 9.28 | 0.40 | 0.37 | 6.36 |
| | | 0.9 | * | - | - | - | - | - | 0.56 | 0.54 | 8.12 |
| | 0.15 | 0.3 | 0.36 | 0.13 | 8.28 | 0.26 | 0.10 | 5.17 | 0.26 | 0.18 | 5.82 |
| | | 0.5 | 0.96 | 0.79 | 14.8 | 0.48 | 0.36 | 10.8 | 0.32 | 0.26 | 8.31 |
| | | 0.7 | * | - | - | 0.83 | 0.74 | 14.2 | 0.42 | 0.37 | 10.5 |
| | | 0.9 | * | - | - | * | - | - | 0.58 | 0.54 | 12.6 |
| 0.10 | 0.05 | 0.3 | 0.26 | 0.13 | -1.17 | 0.22 | 0.10 | -5.03 | 0.25 | 0.18 | -5.30 |
| | | 0.5 | 0.80 | 0.79 | 4.57 | 0.41 | 0.36 | 1.54 | 0.30 | 0.26 | -2.21 |
| | | 0.7 | * | - | - | 0.72 | 0.74 | 4.34 | 0.38 | 0.37 | 0.68 |
| | | 0.9 | * | - | - | * | - | - | 0.52 | 0.54 | 3.09 |
| | 0.10 | 0.3 | 0.32 | 0.13 | 3.01 | 0.25 | 0.10 | -0.78 | 0.27 | 0.18 | -1.18 |
| | | 0.5 | 0.86 | 0.79 | 9.50 | 0.44 | 0.36 | 5.73 | 0.31 | 0.26 | 1.82 |
| | | 0.7 | * | - | - | 0.76 | 0.74 | 9.07 | 0.39 | 0.37 | 4.77 |
| | | | | | | | | | | | |

| | | | | | | | | | | | |
|------|------|-----|------|------|-------|------|------|-------|------|------|-------|
| | 0.15 | 0.9 | * | - | - | - | - | - | 0.54 | 0.55 | 7.43 |
| | | 0.3 | 0.38 | 0.13 | 7.71 | 0.29 | 0.10 | 3.77 | 0.28 | 0.18 | 3.10 |
| | | 0.5 | 0.92 | 0.79 | 14.6 | 0.48 | 0.36 | 10.1 | 0.33 | 0.26 | 6.03 |
| | | 0.7 | * | - | - | 0.80 | 0.74 | 13.9 | 0.41 | 0.37 | 9.02 |
| | | 0.9 | * | - | - | * | - | - | 0.55 | 0.54 | 11.8 |
| 0.15 | 0.05 | 0.3 | 0.29 | 0.13 | -2.41 | 0.26 | 0.10 | -6.97 | 0.28 | 0.18 | -8.46 |
| | | 0.5 | 0.78 | 0.79 | 4.44 | 0.42 | 0.36 | 0.46 | 0.31 | 0.26 | -5.52 |
| | | 0.7 | * | - | - | 0.70 | 0.74 | 4.09 | 0.37 | 0.38 | -1.62 |
| | | 0.9 | * | - | - | * | - | - | 0.50 | 0.54 | 2.07 |
| | | 0.3 | 0.35 | 0.13 | 2.19 | 0.29 | 0.10 | -2.23 | 0.29 | 0.18 | -4.07 |
| | 0.10 | 0.5 | 0.83 | 0.79 | 9.34 | 0.45 | 0.36 | 4.83 | 0.32 | 0.26 | -1.20 |
| | | 0.7 | * | - | - | 0.73 | 0.74 | 8.79 | 0.38 | 0.37 | 2.61 |
| | | 0.9 | * | - | - | - | - | - | 0.51 | 0.54 | 6.42 |
| | 0.15 | 0.3 | 0.40 | 0.13 | 7.16 | 0.32 | 0.10 | 2.70 | 0.30 | 0.18 | 0.44 |
| | | 0.5 | 0.89 | 0.79 | 14.4 | 0.48 | 0.36 | 9.41 | 0.33 | 0.26 | 3.24 |
| | | 0.7 | * | - | - | 0.76 | 0.74 | 13.6 | 0.40 | 0.37 | 6.99 |
| | | 0.9 | * | - | - | * | - | - | 0.53 | 0.54 | 10.8 |

Note: “*” indicate $\mu^{(0)}$ do not exist.

Table-2

Percent relative loss L_2 in precision of T^* with respect to τ for $\rho_{z_1 z_2} = 0.5$.

| ρ_0 | | 0.5 | | | 0.7 | | | 0.9 | | |
|----------|-------------|---------------|---------------|-------|---------------|---------------|-------|---------------|---------------|-------|
| t_1 | ρ_{yx} | $\mu_1^{(0)}$ | $\mu_3^{(0)}$ | L_2 | $\mu_1^{(0)}$ | $\mu_3^{(0)}$ | L_2 | $\mu_1^{(0)}$ | $\mu_3^{(0)}$ | L_2 |
| 0.05 | 0.3 | 0.17 | 0.13 | -2.59 | 0.14 | 0.10 | -5.17 | 0.21 | 0.18 | -5.08 |
| | 0.5 | 0.77 | 0.79 | -0.04 | 0.37 | 0.36 | -1.15 | 0.27 | 0.26 | -2.91 |
| | 0.7 | * | - | - | 0.71 | 0.74 | -0.08 | 0.37 | 0.38 | -1.41 |
| | 0.9 | * | - | - | * | - | - | 0.52 | 0.54 | -0.44 |
| 0.10 | 0.3 | 0.20 | 0.13 | -4.78 | 0.19 | 0.10 | -8.95 | 0.24 | 0.18 | -9.27 |
| | 0.5 | 0.75 | 0.79 | -0.12 | 0.38 | 0.36 | -2.36 | 0.29 | 0.26 | -6.10 |
| | 0.7 | * | - | - | 0.69 | 0.74 | -0.23 | 0.36 | 0.37 | -3.25 |
| | 0.9 | * | - | - | * | - | - | 0.50 | 0.54 | -1.12 |
| 0.15 | 0.3 | 0.24 | 0.13 | -6.62 | 0.23 | 0.10 | -11.5 | 0.27 | 0.18 | -12.7 |
| | 0.5 | 0.73 | 0.79 | -0.23 | 0.39 | 0.36 | -3.66 | 0.30 | 0.26 | -9.70 |
| | 0.7 | * | - | - | 0.67 | 0.74 | -0.46 | 0.36 | 0.37 | -5.72 |
| | 0.9 | * | - | - | * | - | - | 0.48 | 0.54 | -2.15 |

Note: “*” indicate $\mu_1^{(0)}$ do not exist

Table-3

Percent relative loss L_3 in precision of T^{**} with respect to τ for $\rho_{z1z2} = 0.5$.

| ρ_0 | | 0.5 | | | 0.7 | | | 0.9 | | |
|----------|-------------|---------------|---------------|-------|---------------|---------------|-------|---------------|---------------|-------|
| t_2 | ρ_{yx} | $\mu_2^{(0)}$ | $\mu_3^{(0)}$ | L_3 | $\mu_2^{(0)}$ | $\mu_3^{(0)}$ | L_3 | $\mu_2^{(0)}$ | $\mu_3^{(0)}$ | L_3 |
| 0.05 | 0.3 | 0.20 | 0.13 | 1.91 | 0.14 | 0.10 | 1.49 | 0.20 | 0.18 | 2.53 |
| | 0.5 | 0.85 | 0.79 | 4.74 | 0.39 | 0.36 | 3.40 | 0.28 | 0.26 | 3.14 |
| | 0.7 | * | - | - | 0.78 | 0.74 | 4.63 | 0.39 | 0.37 | 3.68 |
| | 0.9 | * | - | - | * | - | - | 0.56 | 0.54 | 4.24 |
| 0.10 | 0.3 | 0.27 | 0.13 | 4.93 | 0.18 | 0.10 | 3.82 | 0.22 | 0.18 | 5.44 |
| | 0.5 | 0.92 | 0.79 | 9.75 | 0.44 | 0.36 | 7.22 | 0.30 | 0.26 | 6.55 |
| | 0.7 | * | - | - | 0.82 | 0.74 | 9.44 | 0.41 | 0.37 | 7.56 |
| | 0.9 | * | - | - | * | - | - | 0.58 | 0.54 | 8.61 |
| 0.15 | 0.3 | 0.35 | 0.13 | 8.89 | 0.22 | 0.10 | 6.96 | 0.22 | 0.18 | 5.44 |
| | 0.5 | 0.99 | 0.79 | 14.9 | 0.48 | 0.36 | 11.4 | 0.30 | 0.26 | 6.55 |
| | 0.7 | * | - | - | 0.86 | 0.74 | 14.4 | 0.41 | 0.37 | 7.56 |
| | 0.9 | * | - | - | * | - | - | 0.58 | 0.54 | 8.61 |

Note: “*” indicate $\mu_2^{(0)}$ do not exist

8. Interpretation of Empirical Results

The following interpretations can be read out from the Tables 1-3.

(1) From Table-1, it is observed that

(a) For the fixed values of t_2, ρ_{yx} and ρ_0 , the values of $\mu^{(0)}$ increase while the values of L_1 decrease with the increasing values of t_1 . Thus, the higher the non response rate on the first occasion, the larger the fresh sample is desirable on the current occasion which leads to the enhancement in the precision of the estimates.

(b) For fixed values of t_1, ρ_{yx} and ρ_0 the values of $\mu^{(0)}$ and L_1 are increasing with the increasing values of t_2 .

(c) For fixed values of t_1, t_2 and ρ_{yx} , the values of L_1 and $\mu^{(0)}$ do not follow any definite pattern with the increasing values of ρ_0 .

(2) From Table-2, it is clear that

(a) For the fixed values of ρ_{yx} and ρ_0 , the values of $\mu_1^{(0)}$ increase while the values of L_2 decrease with the increasing values of t_1 . Thus, the higher the non response rate on the first occasion, the larger the fresh sample is required on the current occasion along with the enhance precision of the estimates.

(b) For the fixed values of t_1 and ρ_{yx} , the values of $\mu_1^{(0)}$ and L_2 do not follow any pattern with the increasing values of ρ_0 .

(c) For fixed values of t_1 and ρ_0 , the values of $\mu_1^{(0)}$ and L_2 are increasing with the increasing values of ρ_{yx} .

(3) From Table-3, it is read out that

(a) For the fixed values of ρ_{yx} and ρ_0 , the values of $\mu_2^{(0)}$ and L_3 are increasing with the increasing values of t_2 . This phenomenon is obvious since the higher the non-response rate on the current occasion, the larger the fresh sample is required and higher loss in precision occurs

(b) For the fixed values of t_2 and ρ_{yx} , the values of $\mu_3^{(0)}$ and L_3 do not follow any pattern with the increasing values of ρ_0 .

(c) For fixed values of t_2 and ρ_0 , the values of $\mu_3^{(0)}$ and L_3 are increasing with the increasing values of ρ_{yx} .

9. Conclusions and Recommendations

In the above empirical results, it may be seen that for all studied cases the loss in precisions were observed, but the losses were appreciably low, even in many cases negative losses (gains) were observed, which indicate the effectiveness of the rotation patterns developed with the help of suggested imputation methods. Hence, it may be concluded that the imputation methods used in this work, may be utilized effectively to handle the problems of non-response in search of good rotation patterns in two-occasion successive sampling if suitable auxiliary information is available on both the occasions. Looking on the novel behaviors of the proposed estimators, we may recommend the survey statisticians to utilize it for their practical uses.

Acknowledgements

Authors are thankful to the editor Prof. M. Ahsanullah and the learned referee for their inspiring and fruitful suggestions. Authors are also thankful to the University Grants Commission, New Delhi and Indian School of Mines, Dhanbad for providing financial assistance and necessary infrastructure to carry out the present research work.

References

1. Ahmed, M. S., Al-Titi, O., Al-Rawi, Z. and Abu-Dayyeh, W. (2006). Estimation of a population mean using different imputation methods. *Statistics in Transition* 7, 6: 1247-1264.
2. Kalton, G., Kasprzyk, D. and Santos, R. (1981). Issues of non-response and imputation of income and program participation. In: kreovski, D., Platek, R., Rao, J. N. K., eds. *Currents Topics in Survey Sampling*, New York, Acad. Press, 455-480.
3. Kalton, G. and Kasprzyk, D. (1982). Imputation for missing surveys responses. *Proc. Sec. Surv. Res. Meth. Amer. Statist. Assoc.* 22-31.
4. Lee, H., Rancourt, E. and Särndal, C. E. (1994). Experiments with variance estimation from survey data with imputed values. *Journal of Official Statistics*, 10, 3: 231-243.
5. Lee, H., Rancourt, E. and Särndal, C. E. (1995). Variance estimation in the presence of imputed data for the generalized estimation system. *Proceeding of the American Statistical Association (Social Survey Research Methods Section)*, 384-389.
6. Rubin, R. B. (1979). Inference and missing data. *Biometrika*, 63, 3:581-592.
7. Sande, I. G. (1979). A personal view of hot-deck imputation procedures. *Survey Methodology*, *Statistics Canada*, 238-247.
8. Singh, G. N. and Karna J. P. (2010). Some imputation methods to minimize the effect of non-response in two-occasion rotation patterns. *Communication in Statistics-Theory & Methods*, 39, 3264-3281.

9. Singh, G. N., Karna, J. P. and Sahoo, L. N. (2009). Some imputation methods for non-response at current occasion in two-occasion rotation patterns. *Journal Statistical Research*, 43, 2, 37-54.
10. Singh, G. N., Priyanka, K. and Kozak, M. (2008). Use of imputation methods at current occasion in two-occasion rotation patterns. *Model Assisted Statistics and Applications*, 3, 2: 99-112.
11. Singh, G. N., Priyanka, K., Kim, J. M and Singh, S. (2010). Estimation of population mean using imputation techniques in sample surveys. *Journal of Korean Statistical Society*, 39, 67-74.
12. Singh, G. N. and Priyanka, K. (2010). Use of imputation methods in two-occasion successive sampling. *Journal of Indian Society of Agricultural Statistics*, 64, 3: 417-432.
13. Singh, R and Singh, N. (1991). Imputation methods in two-dimensional surveys. In: Recent Advances in Agriculture Statistics Research. New Delhi: Wiley Eastern Ltd.
14. Singh, S. (2009): A new method of imputation in survey sampling. *Statistics*, 43, 5: 499-511.
15. Singh, S and Deo, B. (2003): Imputation by power transformation. *Statistical Papers*, 44, 555-579.
16. Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. *Metrika*, 51: 267-276
17. Singh, G. N., Singh, V. K., Priyanka, K., Prasad, S and Karna J. P. (2012). Rotation patterns under imputation of missing data over two-occasion . *Communication in Statistics-Theory & Methods*, 41, 1857-1874.
18. 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*.