Estimation of Population Variance in Two-Phase Sampling in Presence of Random Non - Response

A. Bandyopadhyay
Department of Mathematics, Asansol Engineering College
Asansol–713305, India
arnabbandyopadhyay4@gmail.com

G. N. Singh
Department of Applied Mathematics
Indian School of Mines, Dhanbad-826004, India
gnsingh_ism@yahoo.com

Abstract
The present investigation deals with the problem of estimation of population variance in presence of random non-response in two-phase (double) sampling. Using information on two auxiliary variables, two general classes of estimators have been suggested in two different situations of random non-response and studied their properties under two different set up of two-phase sampling. It is shown that several estimators may be generated from our proposed classes of estimators. Proposed classes of estimators are empirically compared with some contemporary estimators of population variance under the similar realistic situations and their performances have been demonstrated through numerical illustration and graphical interpretation which are followed by suitable recommendations.

Mathematics Subject Classification: 62D05

Keywords: Variance estimation, Two-phase sampling, Random non-response, Study variable, Auxiliary variable, Bias, Mean square error.

1. Introduction
The problem of estimation of population variance arises in many practical situations. For example, a physician needs a full understanding of variations in the degree of human blood pressure, body temperature and pulse rate for adequate prescription. An agriculturist needs an adequate understanding of the variations in climatic factors especially from place to place (or time to time) to be able to plan on when, how and where to plant his crop. The variance estimation technique using auxiliary variable was first considered by Das and Tripathi (1978). Further this was extended by Srivastava and Jhaji (1980), Isaki (1983), Singh (1983), Upadhyay and Singh (1983), Tripathi et al. (1988), Singh and Joarder (1998) and Ahamed et al. (2003) among others. In many situations, information on an auxiliary variable may be readily available on all unit of the population; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation and number of beds in different hospitals may be known in hospital surveys.

However in some practical situations, it is common experience in sample surveys that data cannot always be collected from all the units selected in the sample. For example, the selected families may not be at home at the first attempt and some of them may refuse
to cooperate with the interviewer even if contacted. As many respondents do not reply, available sample of returns is incomplete. The resulting incompleteness is called non-response and is sometimes so large that can completely vitiate the results. Statisticians have long known that failure to account for the stochastic nature of incompleteness can damage the actual conclusion. An obvious problem, that one needs to justify, arises when ignoring the incomplete mechanism. Rubin (1976) advocated three concepts: missing at random (MAR), observed at random (OAR), and parameter distribution (PD). Rubin defined: “The data are MAR if the probability of the observed missingness pattern, given the observed and unobserved data, does not depend on the value of the unobserved data”. Singh and Joarder (1998) studied the properties of ratio type estimator of population variance suggested by Isaki (1983) under two different situations of random non-response (MAR) advocated by Tracy and Osahan (1994) when (i) random non-response on both the study and auxiliary variables and (ii) only on the study variable. Singh et al. (2012) revisited the family of estimators of population variance suggested by Srivastava and Jhajj (1980) under the above situations of random non-responses.

It is worth to be mentioned that all the above recent works of estimation of population variance in presence of random non-response are discussed on the assumption that either population mean or both population mean and variance of the auxiliary variable are known and even if they are unknown, it is assumed that no non-response situations occur on the auxiliary variable in the sampled unit. This may not often be the case. In such situations, it is more generously advisable to draw a large preliminary sample in which auxiliary variable alone is measured. This technique is known as double sampling or two-phase sampling. Two-phase sampling happens to be a powerful and cost effective (economical) technique for obtaining the reliable estimate in first-phase (preliminary) sample for the unknown population parameters of the auxiliary variables. Motivated with these arguments and using information on two auxiliary variables, we have proposed two general classes of estimators of population variance in two-phase sampling applicable for two different realistic situations of random non-response and studied their properties under two different set up of two-phase sampling. It is shown that several estimators may be generated as member of the proposed classes of estimators. The superiorities of the proposed classes of estimators over some contemporary estimators of population variance under the similar realistic conditions have been established through numerical illustration and graphical interpretation. Suitable recommendations have been put forward to the survey statistician.

2. Formulation of Estimators

2.1. Two-Phase Sampling Structure

Consider a finite population $U = (U_1, U_2, \ldots, U_N)$ of $N$ units, $y$, $x$ and $z$ are the variables under study, first auxiliary variable and second auxiliary variable respectively with population means $\bar{Y}$, $\bar{X}$ and $\bar{Z}$. Let $y_k$, $x_k$ and $z_k$ be the values of $y$, $x$ and $z$ for the $k$-th ($k = 1, 2, \ldots, N$) unit in the population. We wish to estimate the population variance
Estimation of Population Variance in Two-Phase Sampling in Presence of Random Non-Response

\[ S^2_y \left( S^2_y = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2 \right) \] of the study variable y in the presence of the auxiliary variables x and z, when the population variance \[ S^2_x \left( S^2_x = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2 \right) \] of x is unknown but the information on z is available for all the units of population. To estimate \( S^2_y \), a first phase sample \( S' \) of size \( n \) is drawn by simple random sampling without replacement scheme (SRSWOR) from the entire population U and observed for the auxiliary variables x and z to estimate \( S^2_x \). Again a second-phase sample \( S \) of size \( m \) (\( m < n \)) is drawn according to the following cases by SRSWOR scheme to observe the characteristic y and x.

**Case I:** Second phase sample is drawn as a subsample of the first phase sample (i.e. \( S \subset S' \)).

**Case II:** Second phase sample \( S \) is drawn independently of the first phase sample \( S' \).

Hence onwards, we use the following notations:

- \( \bar{y}_m, \bar{x}_m, \bar{z}_n \): Sample means of the respective variables based on the sample sizes shown in suffixes.

\[ S^2_z = (N - 1)^{-1} \sum_{i=1}^{N} (z_i - \bar{Z})^2 \] Population variance of the auxiliary variable z.

\[ s^2_{x_m} = (m - 1)^{-1} \sum_{i=1}^{m} (x_i - \bar{x}_m)^2 \] Sample variance of the auxiliary variable x based on sample of size \( m \).

\[ s^2_{y_m}, s^2_{x_n} \] and \( s^2_z \): Sample variances of the respective variables based on sample sizes shown in suffixes.

We assume that no non-response situations occur at the first phase sample \( S' \) while random non-response situations occur either on both the variables y and x or on the variable y alone in the second phase sample \( S \). We have considered that the occurrences of random non-response situation follow the discrete probability distribution as presented below.

**2.2. Non-Response Probability Model**

If random non-response situations occur at the second phase sample \( S \) of size \( m \) and \( r \) \( \{r = 0, 1, 2, \ldots, (m - 2)\} \) denotes the number of sampling units on which information could not be collected due to random non-response, then the observations of the respective variables on which random non-response occur can be taken from the remaining \( (m - r) \) units of the second phase sample. It is assumed that \( r \) is less than \( (m - 1) \), that is, \( 0 \leq r \leq (m - 2) \). We also assume that if \( p \) denotes the probability of...
a non-response among the \((m - 2)\) possible values of non-response, then \(r\) has the following discrete distribution

\[
P(r) = \frac{(m - r)^{m - 2}C_r}{mq + 2p} p^r q^{m - 2 - r}, \quad r = 0, 1, 2, \ldots, (m - 2)
\]

where \(q = 1 - p\) and \(m - 2C_r\) denote the total number of ways of obtaining \(r\) non-responses out of the \((m - 2)\) total possible non-responses, for instance, see Singh and Joarder (1998).

It is to be noted, the probability model, defined in equation (1), is free from actual data values; hence, can be considered as a model suitable for MAR situation.

We have defined following variables based on the responding part of the sample as

\[
\bar{x}_m = \frac{1}{m - r} \sum_{i=1}^{m-r} x_i, \quad \bar{y}_m = \frac{1}{m - r} \sum_{i=1}^{m-r} y_i: \text{ Sample means of the respective variables based on the responding part of the second phase sample } S.
\]

\[
s^2_{x_m} = (m - r - 1) \sum_{i=1}^{m-r} (x_i - \bar{x}_m)^2: \text{ Sample variance of the variable } x \text{ based on the responding part of the second phase sample } S.
\]

\[
s^2_{y_m}: \text{ Sample variance of the study variable } y \text{ based on the responding part of the second phase sample } S.
\]

### 2.3. Proposed Estimation Strategies

Utilizing information on an auxiliary variable \(x\) with unknown \(S^2_x\) and following the work of Isaki (1983), one may propose the ratio type estimator of population variance \(S^2_y\) in two-phase sampling as

\[
t_r = \frac{s^2_y}{s^2_{x_m}}
\]

(2)

Similarly, if \(\bar{X}\) and \(S^2_x\) both are unknown, then following the work of Srivastava and Jhajj (1980), one may define a general class of estimators of population variance \(S^2_y\) in two-phase sampling set up as

\[
t_g = g\left(s^2_{y_m}, u, v\right)
\]

(3)

where \(u = \frac{\bar{X}_m}{\bar{X}_n}\), \(v = \frac{s^2_{y_m}}{s^2_{x_m}}\) and \(g\left(s^2_{y_m}, u, v\right)\) is a parametric function that satisfies similar regularity conditions as given in Srivastava and Jhajj (1980) and is such that \(g\left(S^2_y, 1, 1\right) = S^2_y\).
Motivated with the work of Singh and Joarder (1998) and Singh et al. (2012), one may propose the estimators $t_R$ and $t_g$ for two different situations of a random non-response at the second phase sample $S$ as presented below.

(i) If random non-response occurs on both the variables $y$ and $x$, the estimators $t_R$ and $t_g$ may be considered as

$$t_R = \frac{s_{ym}^{s_y} s_m^{s_y}}{s_m^{s_y}}$$

(4)

and

$$t_g = g(s_y^{s_y}, u^*, v^*)$$

(5)

where $u^* = \frac{\bar{x_m}^*}{\bar{x}_n}$ and $v^* = \frac{\bar{s}_m^{s_x}}{s_s^{s_x}}$.

(ii) If random non-response occurs only on the study variable $y$, then the estimators $t_R$ and $t_g$ may be consider as

$$t_R^{**} = \frac{s_{ym}^{s_y} s_m^{s_y}}{s_m^{s_y}}$$

(6)

and

$$t_g^{**} = g(s_y^{s_y}, u, v)$$

(7)

Motivated with the above suggestions and assuming that the population variance $S_x^2$ of the auxiliary variable $x$ is unknown, we have proposed two general classes of estimators of population variance $S_y^2$ in two-phase sampling set up applicable for two different situations of random non-response and presented below.

**Situation I:** In this situation, we assume that random non-response conditions occur on both the study variable $y$ and the auxiliary variable $x$ at the second phase sample $S$ and also the population variance $S_z^2$ of the auxiliary variable $z$ is known. Accordingly, we have suggested the general class of estimators of population variance $S_y^2$ in two-phase sampling set up as

$$T_i = f \left( s_{ym}^{s_y}, s_m^{s_y}, h_1( s_x^{s_x}, s_z^{s_z} ) \right)$$

(8)

where $h_1( s_x^{s_x}, s_z^{s_z} )$ is a class of estimators of $S_x^2$ using information on $s_x^{s_x}$ and $s_z^{s_z}$, such that

$$h_1( S_x^2, S_z^2 ) = S_x^2.$$  

(9)
We consider the composite function \( f \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) as one-to-one function of \( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2 \) and \( s_{z_n}^2 \) denoted by \( T_1 = F \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) such that

\[
F \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) = S_y^2 \Rightarrow \frac{\partial F \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right)}{\partial s_{y_m}^2} = 1 \quad (10)
\]

with \( \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) and \( F \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) satisfy the following regularity conditions:

1. Whatever be the chosen samples, \( \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) assume values in a closed convex subspace, \( \mathbb{R}^4 \) of the four dimensional real space containing the point \( \left( S_y^2, S_x^2, S_x^2, S_z^2 \right) \).
2. The function \( F \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) is continuous and bounded in \( \mathbb{R}^4 \).
3. The first, second and third order partial derivatives of \( F \left( s_{y_m}^2, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2 \right) \) exist and are continuous and bounded in \( \mathbb{R}^4 \).

It can be observed from equation (8) that the class of estimators \( T_1 \) is very wide in the sense for any parametric function, \( f \left( s_{y_m}^2, s_{x_m}^2, h \left( s_{x_n}^2, s_{z_n}^2 \right) \right) \) satisfying above regularity conditions with \( F \left( S_y^2, S_x^2, S_x^2, S_z^2 \right) = S_y^2 \) may generate an estimators of \( S_y^2 \). For examples, the following ratio, product, regression and exponential type estimators of \( S_y^2 \) are the members of the class \( T_1 \).

\[
\begin{align*}
t_1 &= s_{x_m}^2 - s_{x_n}^2, \quad t_2 = s_{x_m}^2 - s_{x_n}^2, \quad t_3 = s_{x_m}^2 + b_1 \left( s_{x_n}^2 - s_{x_m}^2 \right), \quad t_4 = s_{x_m}^2 \exp \left( \frac{s_{x_m}^2 - s_{x_n}^2}{s_{x_m}^2 + s_{x_n}^2} \right); \quad (i = 1, 2, \ldots, 4)
\end{align*}
\]

where

\[
\begin{align*}
s_{1x_m}^2 &= s_{x_m}^2 - s_{x_n}^2, \quad s_{2x_m}^2 = s_{x_m}^2 - s_{x_n}^2, \quad s_{3x_m}^2 = s_{x_m}^2 + b_2 \left( S_x^2 - s_{x_n}^2 \right), \quad s_{4x_m}^2 = s_{x_m}^2 \exp \left( \frac{S_x^2 - s_{x_n}^2}{S_x^2 + s_{x_n}^2} \right)
\end{align*}
\]

and \( b_1, b_2 \) are the real scalars.

**Situation II:** In this case, we assume that random non-response situation occurs only on the study variable \( y \) while the complete information on the auxiliary variable \( x \) is available at the second phase sample \( S \) and also the population variance \( S_x^2 \) is known. Considering this aspect, we have proposed the general class of estimators of population variance \( S_y^2 \) in two-phase sampling set up as

\[
T_2 = g \left( s_{y_m}^2, s_{x_m}^2, h \left( s_{x_n}^2, s_{z_n}^2 \right) \right). \quad (11)
\]
Estimation of Population Variance in Two-Phase Sampling in Presence of Random Non-Response

We consider the composite function $g(s_{ym}^2, s_{xn}^2, s_{xn}^2, s_{zn}^2)$ as one-to-one function of $s_{ym}^2, s_{xn}^2, s_{xn}^2$ and $s_{zn}^2$ denoted by $T_2 = G(s_{ym}^2, s_{xn}^2, s_{xn}^2, s_{zn}^2)$ such that

$$G(S_y^2, S_x^2, S_x^2, S_z^2) = S_y^2 \Rightarrow \frac{\partial G(s_{ym}^2, s_{xn}^2, s_{xn}^2, s_{zn}^2)}{\partial s_{ym}^2} = 1 \quad (12)$$

with $(S_y^2, S_x^2, S_x^2, S_z^2)$ and $G(s_{ym}^2, s_{xn}^2, s_{xn}^2, s_{zn}^2)$ satisfy the similar regularity conditions as given for $(S_y^2, S_x^2, S_x^2, S_z^2)$ and $F(s_{ym}^2, s_{xn}^2, s_{xn}^2, s_{zn}^2)$ in equation (10).

Proceeding as above, it may be found that the class of estimators $T_2$ is also very wide and we present below some estimators of $S_y^2$ which are members of the class $T_2$.

$$t_1' = s_{ym}^2 \frac{s_{zn}^2}{s_{xn}^2}, \quad t_2' = s_{ym}^2 \frac{s_{zn}^2}{s_{xn}^2}, \quad t_3' = s_{ym}^2 + b_3 (s_{xn}^2 - s_{xm}^2), \quad t_4' = s_{ym}^2 \exp \left( \frac{s_{zn}^2 - s_{xn}^2}{s_{zn}^2 + s_{xn}^2} \right); \quad (i = 1, 2, \ldots, 4)$$

where $b_3$ is a real scalar.

3. Bias and Mean Square Errors of the Proposed Classes of Estimators $T_1$ and $T_2$

The bias and mean square errors (M. S. E.s) of our proposed classes of estimators $T_1$ and $T_2$ are derived up to first order of approximations under large sample assumptions and using the following transformations:

$$s_{ym}^2 = S_y^2 (1 + e_0), \quad s_{xn}^2 = S_x^2 (1 + e_1), \quad s_{xn}^2 = S_x^2 (1 + e_2), \quad s_{zn}^2 = S_z^2 (1 + e_3), \quad s_{zn}^2 = S_z^2 (1 + e_4).$$

Such that $|e_i| < 1 \quad \forall \ (i = 0, 1, \ldots, 4)$.

We have derived the bias and mean square errors of the proposed classes of estimators $T_1$ and $T_2$ separately for the cases I and II of the two-phase sampling structure defined in section 2.1 and present them below.

3. 1 Bias and Mean Square Errors of the Proposed Classes of Estimators under Case I

In this section, we have considered that the second phase sample $S$ of size $m$ is drawn as a subsample from of the first phase sample $S'$ of size $n$ and we have the following results.

$$E(e_0^2) = f' C_0^2, \quad E(e_1^2) = f' C_1^2, \quad E(e_2^2) = f_2 C_2^2, \quad E(e_3^2) = f_2 C_1^2, \quad E(e_4^2) = f_2 C_1^2, \quad E(e_0 e_1) = f_{01} C_0 C_1, \quad E(e_0 e_2) = f_{02} C_0 C_2, \quad E(e_0 e_3) = f_{03} C_0 C_3, \quad E(e_0 e_4) = f_{04} C_0 C_4,$$

$$E(e_1 e_2) = f_{12} C_1 C_2, \quad E(e_1 e_3) = f_{13} C_1 C_3, \quad E(e_1 e_4) = f_{14} C_1 C_4, \quad E(e_2 e_3) = f_{23} C_2 C_3, \quad E(e_2 e_4) = f_{24} C_2 C_4, \quad E(e_3 e_4) = f_{34} C_3 C_4 \quad (13)$$

where
A. Bandyopadhyay, G. N. Singh

\[ f^* = \frac{1}{mq + 2p} \cdot \frac{1}{N}, \quad f_1 = \left( \frac{1}{m} - \frac{1}{N} \right), \quad f_2 = \left( \frac{1}{n} - \frac{1}{N} \right), \quad f_3 = \left( \frac{1}{m} - \frac{1}{n} - \frac{1}{N} \right), \quad f' = \left( \frac{1}{mq + 2p} - \frac{1}{n} \right), \]

\[ \mu_{abc} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{Y})^a (x_i - \bar{X})^b (z_i - \bar{Z})^c; \quad (a, b, c) \) being non negative integers,

\[ \lambda_{abc} = \mu_{abc} \left\{ \frac{\mu_{a20}^{m/2} \mu_{b20}^{m/2} \mu_{c20}^{m/2}}{\mu_{020}^{m/2}} \right\}, \quad C_0 = \sqrt{\lambda_{220} - 1}, \quad C_1 = \sqrt{\lambda_{040} - 1}, \quad C_2 = \sqrt{\lambda_{004} - 1}, \]

\[ \rho_{01} = (\lambda_{220} - 1) / (\lambda_{220} - 1)(\lambda_{040} - 1), \quad \rho_{02} = (\lambda_{202} - 1) / (\lambda_{202} - 1)(\lambda_{004} - 1), \]

\[ \rho_{12} = (\lambda_{022} - 1) / (\lambda_{022} - 1)(\lambda_{004} - 1). \]

From the above expectations, it is to be noted that:

(a) If \( p = 0 \) (there is no non-response), the above expected values of the sample statistics on which random non-responses occur coincide with the usual results.

(b) \( \rho_{01} \) is the correlation between \((y - \bar{Y})^2\) and \((x - \bar{X})^2\). Similarly \( \rho_{12} \) is the correlation between \((x - \bar{X})^2\) and \((z - \bar{Z})^2\) and \( \rho_{02} \) is the correlation between \((y - \bar{Y})^2\) and \((z - \bar{Z})^2\); see for instance Upadhyaya and Singh (2006).

Now, to express the class of estimators \( T_i \) in terms of \( e \)'s, we expand \( F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) \) about the point \((S_y, S_x, S_x, S_z)\) in a third order of Taylor's series expansions and we have

\[ F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) = F(S_y, S_x, S_x, S_z) + d_1(s_{y_n} - S_y) + d_2(s_{x_m} - S_x) + d_3(s_{x_n} - S_x) + d_4(s_{z_n} - S_z) \quad (14) \]

\[ + \frac{1}{2} \left\{ d_{11}(s_{y_n}^2 - S_y^2) + d_{22}(s_{x_m}^2 - S_x^2) + d_{33}(s_{x_n}^2 - S_x^2) + d_{44}(s_{z_n}^2 - S_z^2) \right\} \]

\[ + 2d_{12}(s_{y_n}^2 - S_y^2)(s_{x_m} - S_x) + 2d_{13}(s_{y_n}^2 - S_y^2)(s_{x_n} - S_x) + 2d_{14}(s_{y_n}^2 - S_y^2)(s_{z_n} - S_z) \]

\[ + 2d_{23}(s_{x_m}^2 - S_x^2)(s_{x_n} - S_x) + 2d_{24}(s_{x_m}^2 - S_x^2)(s_{z_n} - S_z) + 2d_{34}(s_{x_n}^2 - S_x^2)(s_{z_n} - S_z) \]

\[ + \frac{1}{6} \left\{ (s_{y_n}^2 - S_y^2) \frac{\partial}{\partial s_{y_n}} + (s_{x_m}^2 - S_x^2) \frac{\partial}{\partial s_{x_m}} + (s_{x_n}^2 - S_x^2) \frac{\partial}{\partial s_{x_n}} + (s_{z_n}^2 - S_z^2) \frac{\partial}{\partial s_{z_n}} \right\} ^3 F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) \]

where

\[ d_1 = \left. \frac{\partial}{\partial s_{y_n}} F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) \right|_{(s_{y_n}^2, s_{x_m}^2, s_{x_n}^2, s_{z}}} , \quad d_2 = \left. \frac{\partial}{\partial s_{x_m}} F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) \right|_{(s_{y_n}^2, s_{x_m}^2, s_{x_n}^2, s_{z}}} , \]

\[ d_3 = \left. \frac{\partial}{\partial s_{x_n}} F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) \right|_{(s_{y_n}^2, s_{x_m}^2, s_{x_n}^2, s_{z})} , \quad d_4 = \left. \frac{\partial}{\partial s_{z_n}} F(s_{y_n}, s_{x_m}, s_{x_n}, s_{z}) \right|_{(s_{y_n}^2, s_{x_m}^2, s_{x_n}^2, s_{z})} . \]
and \( (d_{11}, d_{12}, d_{13}, d_{14}, d_{21}, d_{22}, d_{23}, d_{24}, d_{34}) \) are the second order partial derivatives of \( F(s_{y_m}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) \) at the point \( (S_y^2, S_x^2, S_z^2, S_z^2) \) and \( s_{ym}^2 = S_y^2 + \theta (s_{ym}^2 - S_y^2) \), \( s_{x1}^2 = S_x^2 + \theta (s_{x1}^2 - S_x^2) \), \( s_{x2}^2 = S_x^2 + \theta (s_{x2}^2 - S_x^2) \), \( s_{x3}^2 = S_x^2 + \theta (s_{x3}^2 - S_x^2) \) for \( 0 < \theta < 1 \).

In the light of the conditions mentioned for \( F(s_{y_m}^2, s_{x_i}^2, s_{x_j}^2, s_{x_k}^2) \) in equation (10), it is noted that
\[
F(S_y^2, S_x^2, S_z^2, S_z^2) = S_y^2 \Rightarrow d_1 = 1 \text{ and } d_{11} = \frac{\partial^2}{\partial (s_{ym}^2)^2} F(s_{ym}^2, s_{x1}^2, s_{x2}^2, s_{x3}^2) = 0. \quad (15)
\]

Since the population variance \( S_x^2 \) of the auxiliary variable \( x \) is unknown, therefore, we have to impose the constraint as
\[
d = d_3. \quad (16)
\]

Thus, expressing \( F(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) \) in terms of \( e \)'s and neglecting the terms of \( e \)'s having power greater than two we get
\[
T_1 = F(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) = S_y^2 (1+e_o) + d_3 S_y^2 (e_1 - e_1) + d_3 S_y^2 e_2 + \frac{1}{2} \left( S_y^4 (d_{12} e_1^2 + d_{13} e_1^2 + 2d_{13} e_1 e_2) + d_{14} S_y^4 e_2^2 + 2d_{15} S_y^4 e_1 e_2 + 2d_{14} S_y^4 e_2 e_2 + 2d_{15} S_y^4 e_1 e_2 + 2d_{14} S_y^4 e_2 e_2 \right) \quad (17)
\]

Similarly, expressing \( T_2 \) in terms of \( e \)'s we have
\[
T_2 = G(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) = S_y^2 (1+e_o) + c_2 S_x^2 e_1 + c_3 S_x^2 e_2 + \frac{1}{2} \left( c_4 S_x^4 e_1^2 + c_3 S_x^4 e_1^2 + 2c_2 S_x^4 e_1 e_2 + c_4 S_x^4 e_2^2 + 2c_3 S_x^4 e_1 e_2 + 2c_2 S_x^4 e_2 e_2 \right) \quad (18)
\]

where
\[
c_2 = \frac{\partial}{\partial s_{ym}^2} G(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) \bigg|_{(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2)} = - \frac{\partial}{\partial s_{ym}^2} G(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) \bigg|_{(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2)} \quad \text{(as \( S_x^2 \) is unknown),}
\]
\[
c_3 = \frac{\partial}{\partial s_{ym}^2} G(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) \bigg|_{(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2)} \quad \text{and} \quad \left( c_{22}, c_{33}, c_{44}, c_{12}, c_{13}, c_{14}, c_{23}, c_{24}, c_{34} \right) \text{ are the second order partial derivatives of } G(s_{ym}^2, s_{x_1}^2, s_{x_2}^2, s_{x_3}^2) \text{ at the point } (S_y^2, S_x^2, S_z^2, S_z^2).
\]

Taking expectations on both sides of the equations (17), (18) and using the results in equation (13), we obtain the expressions for bias \( B(.) \) and mean square errors \( M(.) \) of the proposed classes of estimators \( T_1 \) and \( T_2 \) to the first order of approximations as
\[
B(T_1) = E(T_1 - \bar{Y}) = \frac{1}{2} \left[ S_y^4 C_1 \left( d_{12} f_1 + d_{13} f_2 + 2d_{13} f_2 \right) + 2S_y^4 S_x^2 \left( d_{12} f_1 + d_{13} f_2 \right) \rho_{10} C_0 C_1 + d_{14} f_S S_x^2 + 2d_{14} f_S S_x^2 \rho_{10} C_0 C_2 + 2f_2 \left( d_{14} + d_{14} \right) S_x^2 S_y^2 \rho_{12} C_1 C_2 \right], \quad (19)
\]
\[ B(T_2) = E(T_2 - \bar{Y}) = \frac{1}{2} \left[ S_1 C_1 \left( c_{22} f_1 + c_{33} f_2 + 2c_{23} f_2 \right) + 2S_2^2 S_4 \left( c_{12} f_1 + c_{13} f_2 \right) \rho_{00} C_0 \right] \]
\[ + c_{44} f_2 S_4 C_2^2 + 2c_{14} f_2 S_3^2 \rho_{02} C_0 C_2 + 2f_2 \left( e_{24} + c_{34} \right) S_3^2 S_4^2 \rho_{12} C_1 C_2, \quad (20) \]

\[ M(T_1) = E(T_1 - \bar{Y})^2 = f^* S_3 C_0^2 + d_2^* S_4 C_2^2 + d_4^* f_2 S_3^2 C_2 + 2d_2^* f^* S_4^2 \rho_{00} C_0 C_1 + 2d_4^* f_2 S_3^2 \rho_{02} C_0 C_2 \quad (21) \]

\[ M(T_2) = E(T_2 - \bar{Y})^2 = f^* S_3 C_0^2 + c_2^* f_2 S_4^2 C_2^2 + f^* C_1 C_2 + f^* S_4^2 \rho_{12} C_1 C_2 \quad (22) \]

3.2 Bias and Mean Square Errors of the Proposed Classes of Estimators under Case II

If the second phase sample \( S \) is drawn independently of the first phase sample \( S' \), then we have the following results.

\[ E(e_0) = f^* C_0^2, \quad E(e_1^2) = f C_1^2, \quad E(e_2^2) = f_2 C_2^2, \quad E(e_3^2) = f_3 C_3^2, \quad E(e_4^2) = f_4 C_4^2, \]

\[ E(e_0, e_1) = f^* \rho_{00} C_0 C_1, \quad E(e_0, e_2) = f \rho_{00} C_0 C_1, \quad E(e_1, e_4) = f \rho_{00} C_0 C_1, \quad E(e_2, e_3) = f \rho_{00} C_0 C_1, \]

\[ E(e_0, e_2) = E(e_0, e_1) = E(e_1, e_4) = E(e_2, e_3) = E(e_3, e_4) = 0 \quad (23) \]

Proceeding as section 3.1 and using the results in equation (23), we have derived the expressions for bias \( B(.) \) and mean square errors \( M(.) \) of the proposed classes of estimators \( T_1 \) and \( T_2 \) to the first order of approximations as

\[ B(T_1) = E(T_1 - \bar{Y}) = \frac{1}{2} \left[ f^* S_3 C_0^2 \left( d_{22} f^* + d_{33} f_2 \right) + d_{44} f_2 S_4 C_2^2 + 2d_4^* f^* S_3^2 \rho_{00} C_0 C_1 + 2d_4^* f_2 S_3^2 \rho_{02} C_0 C_2 \right], \quad (24) \]

\[ B(T_2) = E(T_2 - \bar{Y}) = \frac{1}{2} \left[ f^* S_3 C_1^2 \left( c_{22} f_1 + c_{33} f_2 \right) + c_{44} f_2 S_4 C_2^2 + 2c_{34} f_2 S_3^2 \rho_{02} C_0 C_2 \right], \quad (25) \]

\[ M(T_1) = E(T_1 - \bar{Y})^2 = f^* S_3 C_0^2 + c_2^* f_2 S_4^2 C_2^2 + f^* S_3 C_2 + f^* S_4^2 \rho_{12} C_1 C_2, \quad (26) \]

and

\[ M(T_2) = E(T_2 - \bar{Y})^2 = f^* S_3 C_0^2 + c_2^* f_2 S_4^2 C_2^2 + f^* S_4^2 \rho_{12} C_1 C_2 \quad (27) \]

Remark 3.1.

The bias and mean square errors of the various estimators (indicated in section 2.3) belonging to the classes of estimators \( T_i \) and \( T_2 \) can be easily obtained by substituting the suitable values of the derivatives in equations (19)-(22) and (24)-(27) as suggested by Singh et al. (2007) and Singh and Vishwakarma (2007).

4. Minimum M. S. E.s of the Proposed Classes of Estimators \( T_i \) and \( T_2 \)

It is obvious from the equations (21), (22), (26), (27) and remark 3.1 that the mean square errors of the proposed classes of estimators \( T_i \) (\( i = 1, 2 \)) depend on the different values of the derivatives \( d_2, d_4, c_2 \) and \( c_4 \). Therefore, we desire to minimize the mean square errors of the proposed classes of estimators \( T_i \) separately for two different cases of two-phase sampling set up considered in this work and shown below:
**Estimation of Population Variance in Two-Phase Sampling in Presence of Random Non - Response**

**Case I**

When second phase sample $S$ is drawn as a sub sample of the first phase sample $S'$, the optimality conditions under which proposed classes of estimators $T_i (i = 1, 2)$ have minimum M. S. E.s are obtained as

$$d_2 = c_2 = - \rho_{01} \frac{S^2 C_0}{S^2 C_1}, \quad d_4 = c_4 = - \rho_{02} \frac{S^2 C_0}{S^2 C_2}$$

(28)

Substituting these optimum values of the derivatives in equations (21) and (22), we have the minimum M. S. E.s of the classes of estimators $T_i (i = 1, 2)$ as

$$\text{Min. } M(T_1) = \left( f^* - f'_0 \rho_{01}^2 - f'_2 \rho_{02}^2 \right) C^2_0 S^4_y$$

(29)

and

$$\text{Min. } M(T_2) = \left( f^* - f'_0 \rho_{01}^2 - f'_2 \rho_{02}^2 \right) C^2_0 S^4_y$$

(30)

**Case II**

When second phase sample $S$ is selected independently of the first phase sample $S'$, the optimality conditions which minimize the mean square errors of the proposed classes of estimators $T_1$ and $T_2$ are obtained as

$$d_2 = - \frac{f^* \rho_{01} S^2 C_0}{(f^* + f'_2) S^2 C_1}, \quad d_4 = - \frac{f^* \rho_{01} \rho_{12} S^2 C_0}{(f^* + f'_2) S^2 C_2}, \quad c_2 = - \frac{f_0 \rho_{01} S^2 C_0}{(f^* + f'_2) S^2 C_1}, \quad c_4 = - \frac{f_0 \rho_{01} \rho_{12} S^2 C_0}{(f^* + f'_2) S^2 C_2}$$

(31)

Substituting these optimum values of the derivatives $d_2, d_4, c_2$ and $c_4$ in equations (26) and (27), we have the expressions of minimum M. S. E. of the classes of estimators $T_i (i = 1, 2)$ as

$$\text{Min. } M(T_1) = \left\{ f^* - \left( f^* \frac{\rho_{01}^2}{f^* + f'_2} \right)^2 - f'_2 \left( f^* \frac{\rho_{01} \rho_{12}}{f^* + f'_2} \right)^2 \right\} C^2_0 S^4_y$$

(32)

and

$$\text{Min. } M(T_2) = \left\{ f^* - \left( f^* \frac{\rho_{01}^2}{f^* + f'_2} \right)^2 - f'_2 \left( f^* \frac{\rho_{01} \rho_{12}}{f^* + f'_2} \right)^2 \right\} C^2_0 S^4_y$$

(33)

**Remark 4.1:** It is to be noted from optimality conditions in equations (28) and (31) that the optimum values of derivatives of the proposed classes of estimators $T_i (i = 1, 2)$ depend on unknown population parameters such as $C_0, C_1, C_2, \rho_{01}, \rho_{12}, \rho_{02}, S^2_y$ and $S^2_x$. Thus, to use such estimators one has to use guessed or estimated values of them. Guessed values of population parameters can be obtained either from past data or experience gathered over time; for instance see Murthy (1967) and Tracy et al. (1996). If the guessed values are not known then it is advisable to use sample data to estimate these parameters as suggested by Singh et al. (2007) and Gupta and Shabbir (2008). It could be seen that the minimum mean square errors of the classes of estimators remain same up to the first
order of approximations, even if population parameters are replaced by their respective sample estimates.

5. Efficiency Comparisons of the Proposed Classes of Estimators $T_1$ and $T_2$

To examine the performances of the proposed classes of estimators under two different cases of two-phase sampling set up as suggested in this paper, we have compared their efficiencies with some other estimators of population variance such as $s^2_{ys}$ (sample variance estimator in presence of random non-response), $t^*_R$, $t^*_g$, $t^*_R^*$ and $t^*_g^*$. Proceeding as sections 3 and 4, the M. S. E.s/ minimum M. S. E.s of these estimators are derived up to the first order of approximations under the Cases I and II of the two phase-sampling set up and presented below.

Case I:

$$M(t^*_R) = S_{y}^4 \left[ \frac{f^* C_0^2 + v C_1^2}{f^* + f_1} - 2f^* \rho_{01} C_0 C_1 \right]$$  \hspace{1cm} (34)

Min. $M(t^*_g) = S_{y}^4 \left[ \frac{f^* C_0^2 + V C_1^2}{f^* + V} - 2f^* \rho_{01} C_0 C_1 \right]$  \hspace{1cm} (35)

Min. $M(t^*_R^*) = S_{y}^4 \left[ \frac{f^* C_0^2 + f_1^* C_1^2}{f^* + f_1} - 2f^* \rho_{01} C_0 C_1 \right]$  \hspace{1cm} (36)

Min. $M(t^*_g^*) = S_{y}^4 \left[ \frac{f^* C_0^2 + W_2^* f_1 C_1^2 + f_1 W_2^* C_1^2}{f^* + f_1} - 2f^* \rho_{01} C_0 C_1 + 2W_1 W_2^* f^* \lambda_{210} C_0 \right]$  \hspace{1cm} (37)

Case II:

$$M(t^*_R) = S_{y}^4 \left[ \frac{f^* C_0^2 + (f^* + f_2) C_1^2}{f^* + f_2} - 2f^* \rho_{01} C_0 C_1 \right]$$  \hspace{1cm} (38)

Min. $M(t^*_g) = S_{y}^4 \left[ \frac{f^* C_0^2 + W_2^* (f^* + f_2) C_1^2}{f^* + f_2} - 2f^* \rho_{01} C_0 C_1 \right]$  \hspace{1cm} (39)

Min. $M(t^*_R^*) = S_{y}^4 \left[ \frac{f^* C_0^2 + (f_1 + f_2) C_1^2}{f^* + f_1} - 2f_1 \rho_{01} C_0 C_1 \right]$  \hspace{1cm} (40)

Min. $M(t^*_g^*) = S_{y}^4 \left[ \frac{f^* C_0^2 + W_2^* (f_1 + f_2) C_1^2 + W_1 W_2^* f^* \lambda_{210} C_0}{f^* + f_1} - 2f_1 \rho_{01} C_0 C_1 + 2W_1 W_2^* f^* \lambda_{210} C_0 \right]$  \hspace{1cm} (41)

where

$$C_x = \frac{S_x}{\bar{X}}, \hspace{1cm} w_1 = \frac{\lambda_{030} \rho_{01} C_0 - \lambda_{210} C_1}{C_x (C_x^2 - \lambda_{030}^2)}, \hspace{1cm} w_2 = \frac{\lambda_{030} \lambda_{210} - \rho_{01} C_0 C_1}{C_x^2 - \lambda_{030}^2},$$

$$w'_1 = \frac{f^* \{ \lambda_{030} \rho_{01} C_0 - \lambda_{210} C_1 \} C_1}{(f^* + f_1) C_x \left( C_x^2 - \lambda_{030}^2 \right)}, \hspace{1cm} w'_2 = \frac{f^* \{ \lambda_{030} \lambda_{210} - \rho_{01} C_0 C_1 \} C_1}{(f^* + f_1) C_x \left( C_x^2 - \lambda_{030}^2 \right)}, \hspace{1cm} v_1 = \frac{f_1 \{ \lambda_{030} \rho_{01} C_0 - \lambda_{210} C_1 \} C_1}{(f_1 + f_2) C_x \left( C_x^2 - \lambda_{030}^2 \right)}, \hspace{1cm} v_2 = \frac{f_1 \{ \lambda_{030} \lambda_{210} - \rho_{01} C_0 C_1 \} C_1}{(f_1 + f_2) C_x \left( C_x^2 - \lambda_{030}^2 \right)}.$$

The variance of $s^2_{ys}$ can be obtained to the first order of approximation as

$$V \left( s^2_{ys} \right) = f^* C_0^2 S_{y}^4.$$

(42)
The performances of the proposed classes of estimators $T_1$ and $T_2$ under their respective optimality conditions are compared with the other estimators considered in this paper and their dominance have been shown by empirical and graphical means of comparisons.

5.1. Numerical Illustration
We have chosen four natural population data sets to illustrate the efficacious performances of the proposed classes of estimators $T_1$ and $T_2$. The source of the populations, the nature of the variables $y$, $x$, $z$ and the values of the various parameters are given as follows.

**Population I-Source: Cochran (1977, Page-182)**

$y$: Number of ‘placebo’ children.
$x$: Number of paralytic polio cases in the placebo group.
$z$: Number of paralytic polio cases in the ‘not inoculated’ group.

$N= 34$, $n = 20$, $m=12$, $C_0=2.32188$, $C_1=1.82685$, $C_x=1.2333$, $\rho_{01}=0.6661$, $\rho_{02}=0.5657$, $\rho_{12}=0.6005$, $\lambda_{030}=1.5224$ and $\lambda_{210}=1.4083$.

**Population II-Source:Murthy (1967, Page-399)**

$y$: Area under wheat in 1964.
$x$: Area under wheat in 1963.
$z$: Cultivated area in 1961.

$N= 34$, $n = 20$, $m=12$, $C_0=1.6510$, $C_1=1.3828$, $C_x=0.7205$, $\rho_{01}=0.9218$, $\rho_{02}=0.8914$, $\rho_{12}=0.9346$, $\lambda_{030}=0.9345$ and $\lambda_{210}=1.0196$.


$y$: Area under wheat in 1937.
$x$: Area under wheat in 1936.
$z$: Total cultivated area in 1931.

$N= 34$, $n = 20$, $m=12$, $C_0=1.5959$, $C_1=1.5105$, $C_x=0.7678$, $\rho_{01}=0.6251$, $\rho_{02}=0.8007$, $\rho_{12}=0.5342$, $\lambda_{030}=1.0982$ and $\lambda_{210}=0.8886$.


$y$: Output.
$x$: Fixed Capital
$z$: Number of workers.

$N= 80$, $n = 60$, $m=40$, $C_0=1.1255$, $C_1=1.6065$, $C_x=0.9485$, $\rho_{01}=0.7319$, $\rho_{02}=0.7940$, $\rho_{12}=0.9716$, $\lambda_{030}=1.2761$ and $\lambda_{210}=0.5461$. 
For different choices of non-response rate \( p \), the performances of the proposed classes of estimators \( T_i \) \((i = 1, 2)\) under their respective optimality conditions are compared with the other estimators considered in this work. The performances of the proposed classes of estimators \( T_i \) \((i = 1, 2)\) have been shown in terms of the percent relative efficiencies and presented in Tables 1-2. The percent relative efficiencies of the proposed classes of estimators \( T_i \) with respect to an estimator \( t \) is defined as

\[
PRE = \frac{\text{M}(t)}{\text{Min. M}(T_i)} \times 100
\]

where \( \text{M}(t) \) denotes the M. S. E./ minimum M. S. E. of an estimator \( t \).

Table 1: PREs of the class of estimators \( T_i \) with respect to other estimators when non-response situations occur on the study variable \( y \) as well as on the auxiliary variable \( x \) at the second phase sample

<table>
<thead>
<tr>
<th>Population I</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimators</td>
<td>( s^2_{ym} )</td>
<td>( t_R^* )</td>
</tr>
<tr>
<td>( p = 0.05 )</td>
<td>166.5081</td>
<td>120.6259</td>
</tr>
<tr>
<td>( p = 0.10 )</td>
<td>167.2997</td>
<td>119.5488</td>
</tr>
<tr>
<td>( p = 0.15 )</td>
<td>168.0647</td>
<td>118.5078</td>
</tr>
<tr>
<td>( p = 0.20 )</td>
<td>168.8045</td>
<td>117.5012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Population II</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimators</td>
<td>( s^2_{ym} )</td>
<td>( t_R^* )</td>
</tr>
<tr>
<td>( p = 0.05 )</td>
<td>588.3309</td>
<td>269.9376</td>
</tr>
<tr>
<td>( p = 0.10 )</td>
<td>592.7544</td>
<td>260.4811</td>
</tr>
<tr>
<td>( p = 0.15 )</td>
<td>597.0525</td>
<td>251.2929</td>
</tr>
<tr>
<td>( p = 0.20 )</td>
<td>601.2304</td>
<td>242.3616</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Population III</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimators</td>
<td>( s^2_{ym} )</td>
<td>( t_R^* )</td>
</tr>
<tr>
<td>( p = 0.05 )</td>
<td>192.4331</td>
<td>156.9049</td>
</tr>
<tr>
<td>( p = 0.10 )</td>
<td>190.3244</td>
<td>153.9273</td>
</tr>
<tr>
<td>( p = 0.15 )</td>
<td>188.3483</td>
<td>151.1370</td>
</tr>
<tr>
<td>( p = 0.20 )</td>
<td>186.4927</td>
<td>148.5168</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Population IV</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimators</td>
<td>( s^2_{ym} )</td>
<td>( t_R^* )</td>
</tr>
<tr>
<td>( p = 0.05 )</td>
<td>229.5861</td>
<td>221.2544</td>
</tr>
<tr>
<td>( p = 0.10 )</td>
<td>228.2167</td>
<td>219.6065</td>
</tr>
<tr>
<td>( p = 0.15 )</td>
<td>226.9754</td>
<td>218.1129</td>
</tr>
<tr>
<td>( p = 0.20 )</td>
<td>225.8451</td>
<td>216.7528</td>
</tr>
</tbody>
</table>
Table 2: PREs of the class of estimators $T_2$ with respect to other estimators when non-response situation occurs only on the study variable $y$ at the second phase sample

<table>
<thead>
<tr>
<th>Population I</th>
<th>Estimators</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$s_{\bar{y}_m}^2$</td>
<td>$t^*_R$</td>
<td>$t^*_g$</td>
</tr>
<tr>
<td>$p = 0.05$</td>
<td>159.1079</td>
<td>119.5632</td>
<td>117.4200</td>
</tr>
<tr>
<td>$P = 0.10$</td>
<td>153.2846</td>
<td>117.6359</td>
<td>115.7038</td>
</tr>
<tr>
<td>$P = 0.15$</td>
<td>148.0951</td>
<td>115.9182</td>
<td>114.1743</td>
</tr>
<tr>
<td>$P = 0.20$</td>
<td>143.4411</td>
<td>114.3779</td>
<td>112.8027</td>
</tr>
</tbody>
</table>

5.2. Graphical Interpretation

For different choices of correlations $\rho_{01}$, $\rho_{02}$ and $\rho_{12}$, we have demonstrated the performances of our proposed classes of estimators by pictorial representation. This could not only improve the readability of the results but also allow the comparison of a much denser grid for different correlation values. For different values of $\rho_{01}$, $\rho_{02}$, $\rho_{12}$, $N = 500$, $n = 200$, $m= 100$ and $p = 0.10$, the PREs of the classes of estimators $T_1$ and $T_2$ are derived with respect to $s_{\bar{y}_m}^2$ and shown in Figures 1-4.
Figure 1: PRE of $T_1$ under Case I

Figure 2: PRE of $T_2$ under Case I

Figure 3: PRE of $T_1$ under Case II

Figure 4: PRE of $T_2$ under Case II

Note: $r_{01}, r_{02}$ and $r_{12}$ denote $\rho_{01}, \rho_{02}$ and $\rho_{12}$ respectively in the Figures 1-4.

6. Conclusions

The following conclusions can be read-out from the present study.

1. From Tables 1 and 2, it is observed that

(a) For high positive values of the correlation coefficients $\rho_{01}, \rho_{02}$ and $\rho_{12}$ (specially for the populations II and IV), the proposed class of estimators $T_i$ yields impressive gains in efficiency over the other estimators $s^{2*}_{\gamma_2}, t^*_R$ and $t^*_g$ and this behavior is visible from both the cases of the two-phase sampling set up as
suggested in this work. Similar situations are also observed for the class of estimators $T_2$ when it dominates the estimators $s^2_{y_m}, t^{**}_{R}$ and $t^{**}_{g}$.

(b) For the different choices of non-response rate $p$, proposed classes of estimators $T_i$ ($i = 1, 2$) are more efficient than the other estimators considered in this work.

2. From Figures 1-4, it is noticed that

(a) The percent relative efficiencies of $T_i$ ($i = 1, 2$) are increasing with the increasing values of the correlation coefficients $\rho_{01}, \rho_{02}$ and $\rho_{12}$. This phenomenon indicates that the proposed classes of estimators perform more precisely, if information on high positively correlated auxiliary variables is available.

Thus it is clear that the uses of auxiliary variables are highly rewarding in terms of the proposed classes of estimators. Hence, the propositions of the classes of estimators in the present study are highly justified as they unify several results. Therefore, the suggested classes of estimators are more attractive in comparison with the previous work of similar nature.

7. Recommendations of the Proposed Work for Real life Applications

In real life survey it may be found that the character of interest is sensitive or stigmatizing such as drinking alcohol, gambling habit, drug addiction, tax evasion, history of induced abortions etc. Hence, a direct survey is likely to yield unreliable responses because presence of random non – response situations in the sampled units. The suggested estimation strategies for estimating the character of interest are recommended to the survey statisticians to handle these realistic situations.

Acknowledgements: Authors are thankful to the reviewer for his valuable and constructive suggestions.

References


