THE EFFICIENT CLASSES OF ESTIMATORS FOR THE PRODUCT OF TWO POPULATION MEANS IN THE EXISTENCE OF NON-RESPONSE UNDER THE STRATIFIED POPULATION-A SIMULATION STUDY

Manish Mishra, B. B. Khare & Sachin Singh

•

School of Liberal Studies, UPES, Dehradun-248001, India.
Department of Statistics, Banaras Hindu University, Varanasi-221005, India.
Department of Mathematics, Sharda University, Greater Noida-201306, India.
manish.mishra@ddn.upes.ac.in, bbkhare56@yahoo.com, singhat619@gmail.com

Abstract

This paper focuses on estimating the product of two population means. Within this paper, we have introduced three distinct classes of estimators for product of two population means. These estimators take into account the known population mean of an auxiliary variable under the framework of stratified random sampling and the presence of non-response in the study variable. Basically, for case (I) we assume the non-response on the study variable and utilize the auxiliary information corresponding to the responding units of the study variable and in case (II), we utilize the complete dataset from the auxiliary variable while also accounting for non-response in the study variable. In case (III) we combined both the information of the auxiliary variable and assumed the non-response on the study variable. Expressions for bias and mean square error have been derived, extending up to the first-order derivative. We have also pinpointed some specific members of the proposed estimator. We have conducted a simulation study to evaluate the valuable insights into the performance of the suggested classes of estimators with the conventional estimator.

Keywords: Product of two population means, Auxiliary variable, Stratified Sampling, Non-response, Mean square error

1. Introduction

The product of two population means is a common parameter used in various areas, including agriculture, economics, social sciences, public health and other scientific investigations. The estimation problem of the product of two population means is very useful in practice. For instance, suppose we want to estimate the total production in a farm having N plants then we select some plants and observe the number of pods and the seeds in pods in a plant then from selected units we will calculate the average number of pods/plant and average number of seeds/plant. Multiplying the average number of seeds/plant and the number of pods/plant with the total number of plants we get the total number of seeds.

In the case of finite population utilizing the auxiliary information(s), the estimation problems related to the product of population means have been examined by several researchers such as Singh[14], Singh[15], Singh[16], Singh[17], Ray and Singh[11], Srivastava et al.[20] and Khare[3] Kumar and Srivastava [8].

Any researcher or statistician may typically encounter the phenomena of the non-response in a scientific investigation or in a sample survey. The reasons for the occurrence of the non-response are the respondent's unwillingness to respond to some questions that are sentimental, the person not at home, Lack of enthusiasm, etc.

To cope with the complications of non-response, Hansen and Hurwitz [2] have put forward a method of sub-sampling from non-responding units. Using information from responding and sub-sampling units picked from non-responding units in the sample, he suggested a population mean estimator.

In the existence of non-response, the estimation of the product of population means utilizing auxiliary information has been considered by Khare and Sinha[4, 5] and Khare et al.[6].

For stratified populations, Khare and Jha[7], Singh et al.[18] and Mishra et al.[9] have suggested precise estimators of population parameters in the situation of non-response on study character using known and unknown population mean of auxiliary variable.

Motivated by Khare and Jha[7], Three classes of estimators are proposed to estimate the product of two population means under the stratified random sampling with the known population mean of auxiliary character having the non-response on study character. The expressions for bias and mean square error (MSE) of the suggested estimators have been obtained. For the numerical study, the data generated by simulation using R-software has also been given to validate the supremacy of the suggested estimators.

2. Notations and Sampling procedure:

In the study of the population $\eta : (\eta_1, \eta_2, \eta_3 ... \eta_N)$ of the size N. The population is divided into L strata. We denote y_h $(h = 1, 2)$ as study characters of the population having population means $\overline{Y}_{hi}$ $(h = 1, 2)$ in the i^{th} stratum, $(i = 1, 2, 3..L)$ and for the auxiliary character (x), the population mean is denoted by $\overline{X}_i$ which is known for each stratum. For each stratum, the stratum proportion and size is also known. Let N_{i1} and $N_{i(2)}$ are the number of units belonging to the responding and non-responding part of the i^{th} stratum such that $N_{i1} + N_{i(2)} = N_i$. For the i^{th} stratum, $W_{i1} = \frac{N_{i1}}{N}$ and $W_{i2} = \frac{N_{i(2)}}{N}$ represent the response and non-response rates.

Here, we are considering the problem in stratified population using non-response in each stratum, the estimation of P $(= \overline{Y}_1 \overline{Y}_2)$ for known $\overline{X}$ has been considered. The sampling procedure we use for the study is as:
we select n_i units from i^{th} stratum, n_{i1} units are selected from N_{i1} units using simple random sampling without replacement (SRSWOR) method of sampling, Here, we obtain n_{i1} units as respondent and $n_{i(2)}$ units as non-responding units out of n_i selected units. We select r_i units, $\left(r_i = \frac{n_{i(2)}}{k_i}, k_i > 1\right)$ from $n_{i(2)}$ units of the i^{th} stratum.
Using the information available on $(n_{i1} + r_i)$ units, the estimator $\overline{y}^*_{hi}$, $(h = 1, 2)$ is given by Hansen and Hurwitz[2] method as follows:

$$\overline{y}^*_{hi} = \frac{n_{i1}}{n_i} \overline{y}_{hi(1)} + \frac{n_{i(2)}}{n_i} \overline{y}_{hi(2)}, \tag{1}$$

where, $\overline{y}_{hi(1)}$ denotes the sample means of $\overline{y}_h$ for n_{i1} responding units in the i^{th} stratum and $\overline{y}_{hi(2)}$ is the sample mean of r_i units drawn from non-responding units $(n_{i(2)})$ in the i^{th} stratum. The estimator $\overline{y}^*_{hi}$ $(h = 1, 2)$ is unbiased and has variance given by:

$$V(\overline{y}^*_{hi}) = \frac{(1 - f_i)}{n_i} S^{*2}_{yhi} + \frac{W_i(k_i - 1)}{n_i} S^{*2}_{yhi(2)} \tag{2}$$

where, S^{*2}_{yhi} and $S^{*2}_{yhi(2)}$ are the population mean square for n_{i1} responding units and non-responding units of i^{th} stratum of the population.

The Hansen and Hurwitz [2] estimator for the auxiliary variable x, we have

$$\overline{x}^*_i = \frac{n_{i1}}{n_i} \overline{x}_{i(1)} + \frac{n_{i(2)}}{n_i} \overline{x}_{i(2)} \tag{3}$$

where, $\overline{x}_{i(1)}$ and $\overline{x}_{i(2)}$ are the sample means of x for n_{i1} responding units and based on r_i units sub-sampled from $n_{i(2)}$ non-responding units of the i^{th} stratum.

In the presence of non-response, the stratified sample means for $\overline{Y}_h$ $(h = 1, 2)$ and $\overline{X}$ are given as follows:

$$\overline{y}^*_{hst} = \sum_{i=1}^{L} W_i \overline{y}^*_{hi} \text{ and } \overline{x}^*_{st} = \sum_{i=1}^{L} W_i \overline{x}^*_i \tag{4}$$

In each stratum utilizing the information on n_i units, the stratified sample mean to estimate $\overline{X}$ is given by:

$$\overline{x}_{st} = \sum_{i=1}^{L} W_i \overline{x}_i \tag{5}$$

where, $W_i = \frac{N_i}{N}$.

3. Proposed Classes of Estimators:

Let $\hat{P}_{st} = \overline{y}^*_{1st}\overline{y}^*_{2st}$ denotes an estimator for P using stratified random sampling in the existence of non-response. We propose three different estimators for P under different situations of non-response which are given as follows:

Case (A): **Utilizing Incomplete information on y_h, $(h = 1, 2)$ and corresponding values of x when $\overline{X}$ is known**.

$$l_1 = d_{(1)}(\hat{P}_{st}, m_1) \tag{6}$$

Case (B): **Incomplete information on y_h, $(h = 1, 2)$ and complete information on x for the known $\overline{X}$**.

$$l_2 = d_{(2)}(\hat{P}_{st}, m_2) \tag{7}$$

Case (C): **Utilizing Incomplete information on y_h, $(h = 1, 2)$ and corresponding values of x and complete information on auxiliary variable.**

$$l_3 = h(\hat{P}_{st}, m_1, m_2) \tag{8}$$

such that,

$$\begin{aligned}
&d_{(j)}(P, 1) = P,\ d_{1(j)}(P, 1) = 1,\ h(P, 1, 1) = P \text{ and } h_1(P, 1, 1) = 1 \text{ where,}\\
&\qquad d_{1(j)}(P, 1) = \left[\frac{\delta}{\delta \hat{P}_{st}} d_{(j)}(\hat{P}_{st}, m_j)\right]_{(P,1)}, d_{2(j)}(P, 1) = \left[\frac{\delta}{\delta m_j} d_{(j)}(\hat{P}_{st}, m_j)\right]_{(P,1)},\\
&\qquad h_1(P, 1, 1) = \left[\frac{\delta}{\delta \hat{P}_{st}} h(\hat{P}_{st}, m_1, m_2)\right]_{(P,1,1)}, h_2(P, 1, 1) = \left[\frac{\delta}{\delta m_1} h(\hat{P}_{st}, m_1, m_2)\right]_{(P,1,1)}\\
&\qquad h_3(P, 1, 1) = \left[\frac{\delta}{\delta m_2} h(\hat{P}_{st}, m_1, m_2)\right]_{(P,1,1)} \text{ for j=1,2.}
\end{aligned} \tag{9}$$

We denote, $\hat{P}_{st} = \overline{y}^*_{1st}\overline{y}^*_{2st}$, $m_1 = \frac{\overline{x}^*_{st}}{\overline{X}}$ and $m_2 = \frac{\overline{x}_{st}}{\overline{X}}$.

We assume that the function $d_{(1)}(\hat{P}_{st}, m_1)$, $d_{(2)}(\hat{P}_{st}, m_2)$ and $h(\hat{P}_{st}, m_1, m_2)$ are fulfilling the regularity conditions which are given as follows:

- For any sample chosen from any design, the selected functions $d_{(j)}(\hat{P}_{st}, m_j)$, and $h(\hat{P}_{st}, m_1, m_2)$ take the values in two-dimensional and three-dimensional real space G_1 and G_2 having the points $(P, 1)$ and $(P, 1, 1)$.
- The partial derivatives of the first-order and second-order with respect to $\hat{P}$, m_1 and m_2 for the functions $d_{(j)}(\hat{P}_{st}, m_j)$, and $h(\hat{P}_{st}, m_1, m_2)$ are continuous and bounded in G_1 and G_2.

(10)

Using the regularity conditions (10) and expanding the function $d_{(j)}(\hat{P}_{st}, m_j)$, for $j = 1, 2$ and $h(\hat{P}_{st}, m_1, m_2)$ about the point $(P, 1)$ and $(P, 1, 1)$ respectively by utilizing Taylor's series up to the second-order partial derivatives, we get

$$l_j = d_{(j)}(\theta_1) + (\hat{P}_{st} - P)d_{1(j)}(\theta_1) + (m_j - 1)d_{2(j)}(\theta_1) + \frac{1}{2}\Big[(\hat{P}_{st} - P)^2 d_{11(j)} (\theta_1^*) + (m_j - 1)^2 d_{22(j)}(\theta_1^*) + 2(\hat{P}_{st} - P)(m_j - 1)d_{12(j)}(\theta_1^*)\Big] \tag{11}$$

$$l_3 = h(\theta_2) + (\hat{P}_{st} - P)h_1(\theta_2) + (m_1 - 1)h_2(\theta_2) + (m_2 - 1)h_3(\theta_2) + \frac{1}{2}\Big[(\hat{P}_{st} - P)^2 h_{11}(\theta_2^*) + (m_1 - 1)^2 h_{22}(\theta_2^*) + (m_2 - 1)^2 h_{33}(\theta_2^*) + 2(\hat{P}_{st} - P)(m_1 - 1)h_{12}(\theta_2^*) + 2(\hat{P}_{st} - P)(m_2 - 1)h_{13}(\theta_2^*) + 2(m_1 - 1)(m_2 - 1)h_{23}(\theta_2^*)\Big] \tag{12}$$

Where, $\theta_1 = (P, 1)$, $\theta_1^* = (\hat{P}_{st}^*, m_j^*)$, $\theta_2 = (P, 1, 1)$, $\theta_2^* = (\hat{P}_{st}^*, m_1^*, m_1^*)$, $\hat{P}^* = P + \alpha_1(\hat{P}_{st} - P)$ and $m_j^* = 1 + \alpha_2(m_j - 1)$, $0 < \alpha_j > 1$ for $j = 1, 2$.

Here, $d_{1(j)}(\theta_1)$ and $d_{11(j)}(\theta_1)$ are the first and second-order partial derivatives with respect to $\hat{P}_{st}$ and $d_{2(j)}(\theta_1)$ and $d_{22(j)}(\theta_1)$ are the first and second-order partial derivatives with respect to m_j for the function $d_{(j)}(\theta_1)$. Similarly $h_1(\theta_2)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ are the first-order partial derivatives with respect to $\hat{P}_{st}$, m_1 and m_2 respectively and using the condition given in equation (9) and presume that the second-order derivatives are very small in equations (11) and (12), the expression for l_j and l_3 are given as follows:

$$l_j = \hat{P}_{st} + (m_j - 1)d_{2(j)}(\theta_1) + P^{-1}(\hat{P}_{st} - P)(m_j - 1)d_{2(j)}(\theta_1^*) \tag{13}$$

$$l_3 = \hat{P}_{st} + (m_1 - 1)h_2(\theta_2) + (m_2 - 1)h_3(\theta_2) \tag{14}$$

For the function $d_{(j)}(\hat{P}_{st}, m_j)$, for $j = 1, 2$ and $h(\hat{P}_{st}, m_1, m_2)$, under the conditions given in equation (10) the Bias(l_j), bias(l_3), MSE(l_j) and MSE$()l_3)$ are always exist.

4. Properties of the Proposed Classes of Estimators

Utilizing the large sample approximation, we assume that:

$$\overline{y}_{1st}^* = \overline{Y}_1(1 + \epsilon_0), \overline{y}_{2st}^* = \overline{Y}_2(1 + \epsilon_1), \overline{x}_{st}^* = \overline{X}(1 + \epsilon_2) \text{ and } \overline{x}_{st} = \overline{X}(1 + \epsilon_3) \tag{15}$$

Such that $|\epsilon_i| < 1$ and $E(\epsilon_i) = 0 \; \forall \; i = 0, 1, 2, 3$.

We also have,

$$E(\epsilon_0^2) = \frac{V(\overline{y}_{1st}^*)}{\overline{Y}_1^2} = \frac{1}{\overline{Y}_1^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{y_{1i}}^2 + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{1i}(2)}^2\right\},$$

$$E(\epsilon_1^2) = \frac{V(\overline{y}_{2st}^*)}{\overline{Y}_2^2} = \frac{1}{\overline{Y}_2^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{y_{2i}}^2 + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{2i}(2)}^2\right\},$$

$$E(\epsilon_3^2) = E(\epsilon_2\epsilon_3) = \frac{V(\overline{x}_{st})}{\overline{X}^2} = \frac{1}{\overline{X}^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{x_i}^2\right\},$$

$$E(\epsilon_2^2) = \frac{V(\overline{x}_{st}^*)}{\overline{X}^2} = \frac{1}{\overline{X}^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{x_i}^2 + \frac{(k_i-1)}{n_i}W_{i2}S_{x_i(2)}^2\right\},$$

$$E(\epsilon_0\epsilon_1) = \frac{Cov(\overline{y}_{1st}^*\overline{y}_{1st}^*)}{\overline{Y}_1\overline{Y}_2} = \frac{1}{\overline{Y}_1\overline{Y}_2}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{1i}y_{2i}} + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{1i}y_{2i}(2)}\right\},$$

$$E(\epsilon_0\epsilon_2) = \frac{Cov(\overline{y}_{1st}^*\overline{x}_{st}^*)}{\overline{Y}_1\overline{X}} = \frac{1}{\overline{Y}_1\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{1i}x_i} + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{1i}x_i(2)}\right\},$$

$$E(\epsilon_1\epsilon_2) = \frac{Cov(\overline{y}_{2st}^*\overline{x}_{st}^*)}{\overline{Y}_2\overline{X}} = \frac{1}{\overline{Y}_2\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{2i}x_i} + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{2i}x_i(2)}\right\},$$

$$E(\epsilon_0\epsilon_3) = \frac{Cov(\overline{y}_{1st}^*\overline{x}_{st})}{\overline{Y}_1\overline{X}} = \frac{1}{\overline{Y}_1\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{1i}x_i}\right\},$$

$$E(\epsilon_1\epsilon_3) = \frac{Cov(\overline{y}_{2st}^*\overline{x}_{st})}{\overline{Y}_2\overline{X}} = \frac{1}{\overline{Y}_2\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{2i}x_i}\right\} \tag{16}$$

Where, $(S_{y_{1i}}^2, S_{y_{2i}}^2)$ represent the population mean square error of (y_1, y_2) for i^{th} stratum and $(S_{y_{1i}(2)}^2, S_{y_{2i}(2)}^2)$ are the population mean square of r_i units sub-sampled from $n_{i(2)}$ units of the study variable for i^{th} stratum and $(S_{x_i}^2, S_{x_{i(2)}}^2)$ represent the population mean square error of x for i^{th} stratum and r_i units sub-sampled from $n_{i(2)}$ for i^{th} stratum respectively.

$(S_{y_{1i}y_{2i}}, S_{y_{1i}x_i}, S_{y_{2i}x_i})$ and $(S_{y_{1i}y_{2i}(2)}, S_{y_{1i}x_i(2)}, S_{y_{2i}x_i(2)})$ are the covariance between $(S_{y_{1i}y_{2i}}, S_{y_{1i}x_i}, S_{y_{2i}x_i})$ and $(S_{y_{1i}y_{2i}(2)}, S_{y_{1i}x_i(2)}, S_{y_{2i}x_i(2)})$stratum of the population and r_i units sub sampled from $n_{i(2)}$ units of the population for the i^{th} stratum respectively and $f_i = \left(\frac{1}{n_i} - \frac{1}{N_i}\right)$.

The Bias and MSE of l_j and l_3 by using the equation (81) and (82) respectively up to the n^{-1} terms of order are given as below:

$$\begin{aligned} Bias(l_j) = E(l_j - P) &= E\Big[\hat{P}_{st} + (m_j-1)d_{2(j)}(\theta_1) \\ &+ P^{-1}(\hat{P}_{st} - P)(m_j-1)d_{2(j)}(\theta_1^*) - P\Big] \\ &= E(\hat{P}_{st} - P) + P^{-1}E(\hat{P}_{st} - P)(m_j-1)d_{2(j)}(\theta_1) \end{aligned} \tag{17}$$

$$\begin{aligned} MSE(l_j) = E(l_j - P)^2 &= E(\hat{P}_{st} - P)^2 + E(m_j-1)^2 d_{2(j)}^2(\theta_1) \\ &+ 2E(\hat{P}_{st} - P)(m_j-1)d_{2(j)}(\theta_1) \end{aligned} \tag{18}$$

$$\begin{aligned} Bias(l_3) = E(l_3 - P) &= E\Big[\hat{P}_{st} + (m_1-1)h_2(\theta_2) + (m_2-1)h_3(\theta_2) - P\Big] \\ &= E(\hat{P}_{st} - P) \end{aligned} \tag{19}$$

$$\begin{aligned} MSE(l_3) = E(l_3 - P)^2 &= E(\hat{P}_{st} - P)^2 + E(m_1-1)^2 h_2^2(\theta_2) + E(m_2-1)^2 h_3^2(\theta_2) \\ &- 2E(\hat{P}_{st} - P)(m_1-1)h_2(\theta_2) + 2E(\hat{P}_{st} - P)(m_2-1)h_3(\theta_2) \\ &+ 2E(m_1-1)(m_2-1)h_2(\theta_2)h_3(\theta_2) \end{aligned} \tag{20}$$

The equation (86) can be written as:

$$MSE(l_3) = E(l_3 - P)^2 = MSE(\hat{P}_{st}) + Ah_2(\theta_2) + Bh_3(\theta_2) + 2Ch_2(\theta_2) + 2Dh_3(\theta_2) + 2Eh_2(\theta_2)h_3(\theta_2) \tag{21}$$

Where, $MSE(\hat{P}_{st}) = E(\hat{P}_{st} - P)^2$, $A = E(m_1 - 1)^2$, $B = E(m_2 - 1)^2$, $C = E(\hat{P}_{st} - P)(m_1 - 1)$, $D = E(\hat{P}_{st} - P)(m_2 - 1)$ and $E = E(m_1 - 1)(m_2 - 1)$.

Now, to get the optimum value of $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ We partially differentiate the expression (18) w.r.t. $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$ and equation (21) w.r.t $h_2(\theta_2)$ and $h_3(\theta_2)$ and equating to zero. Assuming that the partial derivatives of order second are positive, we get

$$d_{2(1)}(\theta_1)_{(opt)} = -\frac{\overline{X}}{V(\overline{x}^*_{st})}\left[\overline{Y}_2 Cov(\overline{y}^*_{1st}, \overline{x}^*_{st}) + \overline{Y}_1 Cov(\overline{y}^*_{2st}, \overline{x}^*_{st})\right] \tag{22}$$

$$d_{2(2)}(\theta_1)_{(opt)} = -\frac{\overline{X}}{V(\overline{x}_{st})}\left[\overline{Y}_2 Cov(\overline{y}^*_{1st}, \overline{x}_{st}) + \overline{Y}_1 Cov(\overline{y}^*_{2st}, \overline{x}_{st})\right] \tag{23}$$

$$h_2(\theta_2)_{opt} = \frac{(DE - BC)}{(AB - E^2)} \tag{24}$$

$$h_3(\theta_2)_{opt} = \frac{(CE - AD)}{(AB - E^2)} \tag{25}$$

The minimum mean square errors after putting the optimum values of $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ in equation (18) and (21) are given by:

$$MSE(l_1)_{min} = MSE(\hat{p}_{st}) - \frac{P^2}{V(\overline{x}^*_{st})}\left[\frac{Cov(\overline{y}^*_{1st}, \overline{x}^*_{st})}{\overline{Y}_1} + \frac{Cov(\overline{y}^*_{2st}, \overline{x}^*_{st})}{\overline{Y}_2}\right]^2 \tag{26}$$

$$MSE(l_2)_{min} = MSE(\hat{p}_{st}) - \frac{P^2}{V(\overline{x}_{st})}\left[\frac{Cov(\overline{y}^*_{1st}, \overline{x}_{st})}{\overline{Y}_1} + \frac{Cov(\overline{y}^*_{2st}, \overline{x}_{st})}{\overline{Y}_2}\right]^2 \tag{27}$$

$$MSE(l_3)_{min} = MSE(\hat{p}_{st}) - \left[\frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right] \tag{28}$$

Members of the proposed classes l_1, l_2 and l_3:

For the given condition in (9), any parametric function $d_{(1)}(\hat{P}_{st}, m_1)$, $d_{(2)}(\hat{P}_{st}, m_2)$ and $h(\hat{P}_{st}, m_2)$ can produce a class of asymptotic estimators. Such types of estimators have a very large number of classes. Some of the members are given below for the proposed classes of estimators:

$$\begin{aligned} l_{j1} &= \hat{P}_{st}(\lambda_1 + (1-\lambda_1)m_j), l_{j2} = (\hat{P}_{st} + a_1(m_j - 1))m_j^{\beta} \\ l_{j3} &= \hat{P}_{st}(m_j)^{\alpha_2}, l_{j4} = \hat{P}_{st}(2 - m_j^{\alpha_3}), l_{j5} = \hat{P}_{st}\exp\left[\frac{m_j - 1}{m_j + 1}\right] \\ l_{31} &= \hat{P}_{st}(\gamma_1 + m_1\gamma_2 + (1-\gamma_1-\gamma_2)m_2), l_{32} = \hat{P}_{st}\exp[\alpha_4(m_1 - 1) \\ &+ \alpha_5(m_2 - 1)], l_{33} = \hat{P}_{st} m_1^{\alpha_6} m_2^{\alpha_7}, l_{32} = \hat{P}_{st}\frac{1}{2}\left[1 + \alpha_8 m_1^{\gamma_3} + (1-\alpha_8)m_2^{\gamma_4}\right] \end{aligned} \tag{29}$$

The MSEs of the suggested classes of the estimators will attain the minimum value of the MSE given in equation (26), (27), (28) for the optimum value $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ given in equation (22), (23), (24) and (25). The member of the proposed classes of the estimators l_1, l_2 and l_3 given in equation (29) will also attain the same minimum MSE. The optimum values are occasionally in the form of some unknown parameters and occasionally in the form of the value of unknown constants these values can be obtained from past data (Reddy [13]) or can be estimated by sample values that do not affect the minimum MSE of the estimator up to the term of order $(\frac{1}{n})$ (Srivastva and Jhajj [19]).

5. Comparison of (l_1, l_2 and l_3) with Pertinent Estimators [$\hat{P}_{st}$]

In stratified random sampling, the estimator for P in the case of non-response is defined as:

$$\hat{P}_{st} = \overline{y}^*_{1st}\overline{y}^*_{2st} \tag{30}$$

The MSE of the $\hat{P}_{st}$ is given as:

$$MSE(\hat{P}) = P^2\left[\frac{V(\overline{y}^*_{1st})}{\overline{Y}^2_1} + \frac{V(\overline{y}^*_{2st})}{\overline{Y}^2_2} + 2\frac{Cov(\overline{y}^*_{1st}, \overline{y}^*_{2st})}{\overline{Y}_1\overline{Y}_2}\right] \tag{31}$$

6. Simulation Study:

To evaluate the characteristics of suggested classes of estimators, we perform a simulation study to artificially generate the population using normal distribution. For this study, we generate a population of size 4500. The following table (1) shows the distributions of the population generated to perform the study.

Table 1: *Sample Size and Distribution*

Strata No.	Stratum Size(N_i)	Sample Size (n_i)	Distribution of $y_{1h_i} \sim N(\mu, \sigma^2)$	Distribution of $y_{2h_i} \sim N(\mu, \sigma^2)$	Distribution of $x_{h_i} \sim N(\mu, \sigma^2)$
1	1800	600	N(500,81)	N(900,121)	N(600,100)
2	1200	400	N(300,64)	N(500,100)	N(300,121)
3	900	300	N(400,100)	N(400,64)	N(500,81)
4	600	200	N(500,121)	N(300,81)	N(300,64)

Here, we employ some transformations suggested by Reddy et al.[12]. To generate the population for study variables and auxiliary variable with some association. The following table (2) shows the required transformation and correlation to generate the random variables. Now, to estimate the approximate mean square error of the suggested classes of the estimators and estimator $\hat{P}_{st}$. We average the outcomes after 6000 iterations of the loop.

To calculate the approximate mean square error (AMSE) the formula is given as follows:

$$AMSE(m^*_1) = \frac{1}{6000}\sum_{t=1}^{6000}(m^*_1 - P)^2$$

where, $m^*_1 = \hat{P}_{st},\ l_1\ l_2\ and\ l_3$.

Table 2: *Transformation and correlation:*

Strata no.	cor(y_{1h_i}, x_{h_i})	cor(y_{1h_i}, x_{h_i})	Transformed auxiliary variable x_{h_i} using y_{1h_i}	Transformed auxiliary variable x_{h_i} using y_{2h_i}
I	−0.58	−0.60	$x_{11i} = r_{y_1x_1}y_{11i} + x_{1i}\sqrt{1 - r^2_{y_1x_1}}$	$x_{21i} = r_{y_2x_1}y_{21i} + x_{1i}\sqrt{1 - r^2_{y_2x_1}}$
II	−0.62	−0.72	$x_{12i} = r_{y_1x_2}y_{12i} + x_{2i}\sqrt{1 - r^2_{y_1x_2}}$	$x_{22i} = r_{y_2x_2}y_{22i} + x_{2i}\sqrt{1 - r^2_{y_2x_2}}$
III	−0.70	−0.58	$x_{13i} = r_{y_1x_3}y_{13i} + x_{3i}\sqrt{1 - r^2_{y_1x_3}}$	$x_{23i} = r_{y_2x_3}y_{23i} + x_{3i}\sqrt{1 - r^2_{y_2x_3}}$
IV	−0.65	−0.65	$x_{14i} = r_{y_1x_4}y_{14i} + x_{4i}\sqrt{1 - r^2_{y_1x_4}}$	$x_{24i} = r_{y_2x_4}y_{24i} + x_{4i}\sqrt{1 - r^2_{y_2x_4}}$

Table 3: *Percentage relative efficiency (PRE) of the proposed estimator with respect to relevant estimators.*

Rate of Non-response	Estimators	1/k		
		1/2	1/3	1/4
10%	$\hat{P}_{st}$	100(3991.83)	100(4484.76)	100(5022.54)
	l_1	117.97(3383.59)	118.25(3792.38)	118.49(4238.51)
	l_2	115.26(3463.26)	113.39(3954.87)	111.83(4491.21)
	l_3	118.22(3376.57)	118.80(3774.79)	119.45(4204.63)
20%	$\hat{P}_{st}$	100(4509.60)	100(5508.47)	100(6616.83)
	l_1	118.06(3819.70)	118.10(4664.08)	118.08(5603.45)
	l_2	113.21(3983.25)	110.62(4979.58)	108.69(6087.46)
	l_3	118.26(3813.21)	118.40(4652.04)	115.57(5580.08)
30%	$\hat{P}_{st}$	100(5034.18)	100(6578.39)	100(8146.68)
	l_1	117.98(4266.66)	118.24(5563.51)	118.54(6872.50)
	l_2	111.72(4505.72)	108.73(6049.82)	106.93(7618.18)
	l_3	118.14(4261.11)	118.47(5552.53)	118.85(6854.30)

(Note: The AMSE of the estimators are shown in parenthesis.)

Table (3) depicts the $AMSE$ of the estimators $\hat{P}_{st}$, l_1, l_2 and l_3 for different values of k and non-response rates.

From the above Table (3) we can see that as we increase the value of k and non-response rates from 10% to 30% the approximate means square error of the estimators $\hat{P}_{st}$, l_1, l_2 and l_3 increases.

7. Conclusion

The proposed estimators l_1, l_2 and l_3 are found to be more efficient than the estimator $\hat{P}_{st}$. And it is also observed that from Table (3) the proposed classes of estimators l_1 and l_3 are found to be almost equally efficient for the different choices of sub-sampling fraction and non-response rates.

The $AMSE(l_1)$ less than or greater than $AMSE(l_2)$ depending upon the situation of correlation between study variables y_1, y_2 and auxiliary variable x. Rao[10] has shown the situation when the conventional estimator using $\overline{x}^*$ and alternate estimator using $\overline{x}$ will be less than each other depending upon the situations given by him. This theory also works here in the case of l_1, l_2 and l_3.

Here, some times $AMSE(l_1) < AMSE(l_2)$ then $AMSE(l_3)$ will be less than $AMSE(l_2)$ and almost equal or less than $AMSE(l_1)$. But if $AMSE(l_1) > AMSE(l_2)$ then $AMSE(l_3)$ will be less than $AMSE(l_1)$ and almost equal to the $AMSE(l_2)$. So the use of $AMSE(l_3)$ is advisable if the condition when $AMSE(l_1) < AMSE(l_2)$ or $AMSE(l_1) > AMSE(l_2)$ is unknown.

Hence, for the estimation of P under the stratified random sampling in the existence of non-response for the Known $\overline{X}$. We suggest to use the estimators l_1, l_2 and l_3 depending on the situations discussed in the results.

References

[1] El-Badry, M. A. (1956). A sampling procedure for mailed questionnaires. *Journal of the American Statistical Association* , 51(274), 209-227.

[2] Hansen, M. H., Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, 41(236), 517-529.

[3] Khare, B. B. (1992). On class of estimators for product of two population means using multi-auxiliary characters with known and unknown means, *.Jour. Appl. Stat*, 1, 56-67.

[4] Khare, B. B., and R. R. Sinha(2010). On class of estimators for the product of two population means using auxiliary character in presence of non-response."*Inter. Trans. Appl. Sci* 2.4 841-846.

[5] Khare, B. B., & Sinha, R. R. (2012). Combined class of estimators for ratio and product of two population means in presence of non-response. *International Journal of Statistics and Economics*, 8(12), 12-20.

[6] Khare, B. B., Jha, P. S., & Kumar, K. (2014). Improved generalized chain estimators for ratio and product of two population means using two auxiliary characters in the presence of non-response. *International J. Stats & Economics*, 13(1), 108-121.

[7] Khare, B. B., Jha, P. S. (2017). Classes of estimators for population mean using auxiliary variable in stratified population in the presence of non-response. *Communications in Statistics-Theory and Methods*, 46(13), 6579-6589.

[8] Kumar, K. Srivastava, U. (2018). Estimation of ratio and product of two population means using exponential type estimators in sample surveys. *International Journal of Mathematics and Statistics*, 19"3, pp. 102"109.

[9] Mishra, M., Khare, B. B., Singh, S. (2023). TWO-CLASSES FOR REGRESSION TYPE OF ESTIMATORS FOR THE RATIO OF TWO POPULATION MEANS IN TWO-PHASE SAMPLING IN THE PRESENCE OF NON-RESPONSE FOR STRATIFIED POPULATION. *Reliability: Theory Applications*, 18(3 (74)), 610-620.

[10] Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the non-respondents. *Survey methodology*, 12(2), 217-230.

[11] Ray, S. K., & Singh, R. K. (1985). Some estimators for the ratio and product of population parameters. *Journal of the Indian Society of Agricultural Statistics*, 37(1), 1-10.

[12] Reddy, M. K., Rao, K. R., & Boiroju, N. K. (2010). Comparison of ratio estimators using Monte Carlo simulation. *International Journal of Agriculture and Statistical Sciences*, 6(2), 517-527.

[13] Reddy, V. N. (1978). A study on the use of prior knowledge on certain population parameters in estimation. *Sankhya C*, 40, 29-37.

[14] Singh, M. P. (1965). On the estimation of ratio and product of the population parameters. *Sankhy´: The Indian Journal of Statistics*, Series B, 321-328.

[15] Singh, M. P. (1967). cum product method of estimation. *Metrika*, 12(1), 34-42.

[16] Singh, M. P. (1969). of some ratio-cum-product estimators. *Sankhy´: The Indian Journal of Statistics, Series B*, 375-378.

[17] Singh, R.K. (1982). the estimating ratio and product of population parameters. . *Stat. Assoc. Bull*, 31, 69-76.

[18] Singh, R. , Khare, S. , Khare, B.B. Jha , P.S. (2020). The general classes of estimators for population mean under stratified two phase random sampling in the presence of non-response. *Int. J. Agricult. Stat. Sci.*, Vol. 16, No. 2, pp. 557-565, 2020.

[19] Srivastava, S. K., Jhajj, H. S. (1983). A class of estimators of the population mean using multiauxiliary information. *Calcutta Statistical Association Bulletin*, 32(1-2), 47-56.

[20] Srivastava, S. R., Khare, B. B., & Srivastava, S. R. (1988). On generalized chain estimator for ratio and product of two population means using auxiliary characters. *Assam Statistical Review*, 2(1), 21-29.