Improved Classes of Estimators for Ratio of Two Means with Double Sampling the Non Respondents

Brij Behari Khare[1] | *Banaras Hindu University, Varanasi, India*
Raghaw Raman Sinha[2] | *Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, India*

Abstract

In this paper, we have considered the problem of estimation of ratio of two population means (R) using multivariate auxiliary characters with known population means under incomplete information. Following Tripathi (1970) and Tripathi and Chaturvedi (1979), general classes of estimators for estimating R using multi-variate auxiliary characters under incomplete information have been proposed and their properties are studied. The expressions of the conditions for attaining minimum mean square error of the proposed classes of estimators have been derived and the minimum values of their mean square errors are given. The justification for using the proposed classes of estimators has been given efficiently with the help of theoretical and empirical studies.

Keywords	***JEL code***
Ratio, auxiliary characters, bias, mean square error, incomplete information	*C83*

INTRODUCTION

The estimation of the ratio of two population means using multivariate auxiliary characters has been widely used in the different field of science and humanities. The problem of estimation of ratio of two population means using one and multi variate auxiliary characters with known population means have been studied by Hartley and Ross (1954), Singh (1965), Tripathi (1970), Tripathi and Chaturvedi (1979) and Khare (1991). But in most of the sample surveys based on mail questionnaire or related to human population, we often find incomplete information due to the occurrence of non-response. To reduce the effect of non-response in such situations, Hansen and Hurwitz (1946) first suggested the method of sub-sampling on the non-responding group and suggested an unbiased estimator for estimating the

1 Department of Statistics, Banaras Hindu University, Varanasi – 221005, India, e-mail: bbkhare56@yahoo.com.
2 Department of Mathematics, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar – 144011, India. Corresponding author: e-mail: raghawraman@gmail.com.

population mean by using the information available from responding and non-responding group. Later on, using the technique of Hansen and Hurwitz (1946), some estimators for estimating the population mean using auxiliary characters with known and unknown population means have been proposed by Rao (1986, 1990), Khare and Srivastava (1993, 1995, 1997, 2000), Khare and Sinha (2002, 2009) and Singh and Kumar (2009). Toutenberg and Srivastava (1998) have considered the problem of estimating the ratio of two population means in sample survey when some observations are missing due to random non-response while Khare and Sinha (2004) have proposed classes of estimators for the estimation of finite population ratio using two phase sampling scheme in presence of non-response.

In this paper, we have proposed two general classes of estimators using multi-auxiliary characters with known population means under different situations of non-response and studied their properties. The superiority of the proposed classes of estimators has been shown through theoretical and empirical comparisons.

1 THE PROPOSED CLASSES OF ESTIMATORS

Let $Y_{il}\ (i = 1,2)$ and $X_{jl}\ (j = 1,2,\dots,p)$ be the non-negative value of l^{th} unit of the study characters $y_i\ (i = 1,2)$ and the auxiliary characters $x_j\ (j = 1,2 \dots,p)$ for a population of size N with population means $\bar{Y}_i\ (i = 1,2)$ and $\bar{X}_j\ (j = 1,2,\dots,p)$. Let n be the size of the sample drawn from the population of size N using simple random sampling without replacement (SRSWOR) scheme of sampling and it has been observed that n_1 units respond and n_2 units do not respond in the sample of size 'n'. In this procedure, the whole population is supposed to be consisting of two non-overlapping strata of N_1 responding and $N_2\ (= N - N_1)$ non-responding units, though they are not known in advance. The stratum weights of responding and non-responding groups are given by $W_1 = \frac{N_1}{N}$ and $W_2 = \frac{N_2}{N}$ and their estimates are respectively given by $\widehat{W}_1 = \frac{n_1}{n}$ and $\widehat{W}_2 = \frac{n_2}{n}$. In this problem, we have considered that the responding and non-responding units are same for the study and auxiliary characters. Further by making extra effort, a subsample of size $r\ (= n_2 / k, k > 1)$ from n_2 non-responding units has been drawn by using SRSWOR method of sampling. Now, following Hansen and Hurwitz (1946) technique, the unbiased estimator for estimating the population mean using $(n_1 + r)$ observations on $y_i\ (i = 1,2)$ characters is given by:

$$\bar{y}_i^* = \frac{n_1}{n}\,\bar{y}_{i(1)} + \frac{n_2}{n}\bar{y}'_{i(2)}\,; \qquad i = 1,2, \tag{1}$$

and the variance of the estimator $\bar{y}_i^*$ upto the order (n^{-1}) is given by:

$$V(\bar{y}_i^*) = \frac{1-f}{n}S_i^{*2} + \frac{W_2(k-1)}{n}S_{i(2)}^{*2}\,; \qquad i = 1,2, \tag{2}$$

where $f = \frac{n}{N}$, $W_i = \frac{N_i}{N}$, $\bar{y}_{i(1)}$ and $\bar{y}'_{i(2)}\ (i = 1,2)$ are the sample means of characters y_i based on n_1 and r units and S_i^{*2} and $S_{i(2)}^{*2}$ are the population mean square errors of y_i for the entire population and non-responding part of the population.

Similarly, the estimator $\bar{x}_j^*\ (j = 1,2 \dots p)$ for estimating the population mean $\bar{X}_j$ is given by:

$$\bar{x}_j^* = \frac{n_1}{n}\,\bar{x}_{j(1)} + \frac{n_2}{n}\,\bar{x}'_{j(2)}. \tag{3}$$

Let $\hat{R}\left(= \frac{\bar{y}_1^*}{\bar{y}_2^*}\right)$ denotes a conventional estimator for estimating the ratio of two population means $R\left(= \frac{\bar{Y}_1}{\bar{Y}_2}\right)$. So we have proposed two different classes of estimators for estimating R utilizing the multi-auxiliary characters with known population means in two different situations.

1.1 $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p$ known and there are incomplete information on y_i $(i = 1, 2)$ and x_j $(j = 1, 2, \dots p)$

In this case, we observe that n_1 units respond for y_1, y_2 and $x_1, x_2, \dots, x_p$ in the sample of size n and $\bar{X}_j$'s $(j = 1,2, \dots, p)$ are known. We now propose a class of estimators t_r for estimating the ratio of two population means (R) using multi-auxiliary characters $x_1, x_2, \dots, x_p$ with their respective known population means in presence of non-response as:

$$t_r = g(m, \boldsymbol{u}'), \tag{4}$$

such that $g(R, \boldsymbol{e}') = R$; $\qquad g_1(R, \boldsymbol{e}') = \left(\frac{\partial}{\partial m} g(m, \boldsymbol{u}')\right)_{(R, \boldsymbol{e}')} = 1,$ (5)

where $m = \frac{\bar{y}_1^*}{\bar{y}_2^*}$, $u_j = \frac{\bar{x}_j^*}{\bar{X}_j}$; $(j = 1, 2, \dots p)$, $\boldsymbol{u}$ and $\boldsymbol{e}$ denote the column vectors $(u_1, u_2, \dots, u_p)'$ and $(1, 1, \dots, 1)'$ respectively.

1.2 $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p$ known and incomplete information on y_i $(i = 1, 2)$ but complete information on x_j $(j = 1, 2, \dots p)$ available in the sample

In this case we observe that n_1 units respond on y_1, y_2 but there is complete information on $x_1, x_2, \dots, x_p$ in the sample of size n (see Rao, 1986) and $\bar{X}_j$'s $(j = 1, 2, \dots, p)$ are known. In such case we propose a class of estimators t_r^* for estimating the ratio of two population means (R) using multi-auxiliary characters $x_1, x_2, \dots, x_p$ with their known population means in presence of non-response as:

$$t_r^* = h(m, \boldsymbol{\omega}'), \tag{6}$$

such that $h(R, \boldsymbol{e}') = R$; $\qquad h_1(R, \boldsymbol{e}') = \left(\frac{\partial}{\partial m} h(m, \boldsymbol{\omega}')\right)_{(R, \boldsymbol{e}')} = 1,$ (7)

where $\boldsymbol{\omega}$ denotes the column vector $(\omega_1, \omega_2, \dots, \omega_p)'$ and $\omega_j = \frac{\bar{x}_j}{\bar{X}_j}$, $(j = 1, 2, \dots p)$.

The functions $g(m, \boldsymbol{u}')$ and $h(m, \boldsymbol{\omega}')$ satisfy the following conditions:

(i) For any sampling design, whatever be the sample chosen, $(m, \boldsymbol{u}')$ [or $(m, \boldsymbol{\omega}')$] assumes value in a bounded, closed convex subset D_r [or D_r^*] of the $p + 1$ dimensional real space containing the point $(R, \boldsymbol{e}')$.

(ii) In D_r [or D_r^*], the function $g(m, \boldsymbol{u}')$ [or $h(m, \boldsymbol{\omega}')$] is continuous and bounded.

(iii) The first and second partial derivatives of $g(m, \boldsymbol{u}')$ [or $h(m, \boldsymbol{\omega}')$] exist and are continuous and bounded in D_r [or D_r^*].

Here $[g_1(m, \boldsymbol{u}'), g_2(m, \boldsymbol{u}')]$ and $[h_1(m, \boldsymbol{\omega}'), h_2(m, \boldsymbol{\omega}')]$ denote the first partial derivatives of $g(m, \boldsymbol{u}')$ and $h(m, \boldsymbol{\omega}')$ with respect to $[m, \boldsymbol{u}']$ and $[m, \boldsymbol{\omega}']$ respectively. The second partial derivatives of $g(m, \boldsymbol{u}'), h(m, \boldsymbol{\omega}')$ with respect to $\boldsymbol{u}'$ and $\boldsymbol{\omega}'$ are denoted by $g_{22}(m, \boldsymbol{u}'), h_{22}(m, \boldsymbol{\omega}')$ and first partial derivatives of $g_2(m, \boldsymbol{u}')$ and $h_2(m, \boldsymbol{\omega}')$ with respect to m are denoted by $g_{12}(m, \boldsymbol{u}')$ and $h_{12}(m, \boldsymbol{\omega}')$.

It may be seen that the bias and mean square error of the estimators t_r and t_r^* will always exist under the regularity conditions imposed on $g(m, \boldsymbol{u}')$ and $h(m, \boldsymbol{\omega}')$.

Now expanding $g(m, \boldsymbol{u}')$ and $h(m, \boldsymbol{\omega}')$ about the point $(R, \boldsymbol{e}')$ using Taylor's series upto second partial derivatives and using the condition (5) and (7) we have:

$$\begin{aligned} t_r = {} & R + (m - R) + (\boldsymbol{u} - \boldsymbol{e})' g_2(R, \boldsymbol{e}') + \frac{1}{2}\{(m - R)^2 g_{11}(m^*, \boldsymbol{u}^{*\prime}) \\ & + 2(m - R)(\boldsymbol{u} - \boldsymbol{e})' g_{12}(m^*, \boldsymbol{u}^{*\prime}) + (\boldsymbol{u} - \boldsymbol{e})' g_{22}(m^*, \boldsymbol{u}^{*\prime})(\boldsymbol{u} - \boldsymbol{e})\}, \end{aligned} \tag{8}$$

and

$$\begin{aligned} t_r^* = {} & R + (m - R) + (\boldsymbol{\omega} - \boldsymbol{e})' h_2(R, \boldsymbol{e}') + \frac{1}{2}\{(m - R)^2 h_{11}(m^*, \boldsymbol{\omega}^{*\prime}) \\ & + 2\,(m - R)(\boldsymbol{\omega} - \boldsymbol{e})' h_{12}(m^*, \boldsymbol{\omega}^{*\prime}) + (\boldsymbol{\omega} - \boldsymbol{e})' h_{22}(m^*, \boldsymbol{\omega}^{*\prime})(\boldsymbol{\omega} - \boldsymbol{e})\}, \end{aligned} \tag{9}$$

where $m^* = R + \phi_r(m - R)$, $\boldsymbol{u}^* = \boldsymbol{e} + \boldsymbol{\phi}_1(\boldsymbol{u} - \boldsymbol{e})$, $\boldsymbol{\omega}^* = \boldsymbol{e} + \boldsymbol{\phi}_2(\boldsymbol{\omega} - \boldsymbol{e})$ such that, $0 < \phi_r, \phi_{1j}, \phi_{2j} < 1$ and $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ are the $(p \times p)$ diagonal matrix having ϕ_{1j} and ϕ_{2j} as their j^{th} diagonal elements.

2 BIAS AND MEAN SQUARE ERROR (MSE) OF t_r AND t_r^*

From (8) and (9), the expressions for bias and mean square error of t_r and t_r^* for any sampling design upto the terms of order (n^{-1}) are given by:

$$Bias(t_r) = Bias(\hat{R}) + E(\hat{R} - R)(\boldsymbol{u} - \boldsymbol{e})' g_{12}(m^*, \boldsymbol{u}^{*'}) + \tfrac{1}{2} E(\boldsymbol{u} - \boldsymbol{e})' g_{22}(m^*, \boldsymbol{u}^{*'})(\boldsymbol{u} - \boldsymbol{e}), \quad (10)$$

$$MSE(t_r) = MSE(\hat{R}) + 2E(\hat{R} - R)(\boldsymbol{u} - \boldsymbol{e})' g_2(R, \boldsymbol{e}') + E(g_2(R, \boldsymbol{e}'))'(\boldsymbol{u} - \boldsymbol{e})(\boldsymbol{u} - \boldsymbol{e})' g_2(R, \boldsymbol{e}'), \quad (11)$$

$$Bias(t_r^*) = Bias(\hat{R}) + E(\hat{R} - R)(\boldsymbol{\omega} - \boldsymbol{e})' h_{12}(m^*, \boldsymbol{\omega}^{*'}) + \tfrac{1}{2} E(\boldsymbol{\omega} - \boldsymbol{e})' h_{22}(m^*, \boldsymbol{\omega}^{*'})(\boldsymbol{\omega} - \boldsymbol{e}), \quad (12)$$

$$MSE(t_r^*) = MSE(\hat{R}) + 2E(\hat{R} - R)(\boldsymbol{\omega} - \boldsymbol{e})' h_2(R, \boldsymbol{e}') + E(h_2(R, \boldsymbol{e}'))'(\boldsymbol{\omega} - \boldsymbol{e})(\boldsymbol{\omega} - \boldsymbol{e})' h_2(R, \boldsymbol{e}'). \quad (13)$$

The mean square error of t_r and t_r^* will attain their minimum values if:

$$g_2(R, \boldsymbol{e}') = -\,(E(\boldsymbol{u} - \boldsymbol{e})(\boldsymbol{u} - \boldsymbol{e})')^{-1} E(\hat{R} - R)(\boldsymbol{u} - \boldsymbol{e}), \quad (14)$$

and $$h_2(R, \boldsymbol{e}') = -\,(E(\boldsymbol{\omega} - \boldsymbol{e})(\boldsymbol{\omega} - \boldsymbol{e})')^{-1} E(\hat{R} - R)(\boldsymbol{u} - \boldsymbol{e}), \quad (15)$$

respectively. By putting the value of $g_2(R, \boldsymbol{e}')$ from (14) in (11) and $h_2(R, \boldsymbol{e}')$ from (15) in (13), the minimum values of mean square error of t_r and t_r^* are given by:

$$MSE(t_r)_{min.} = MSE(\hat{R}) - E(\hat{R} - R)(\boldsymbol{u} - \boldsymbol{e})'(E(\boldsymbol{u} - \boldsymbol{e})(\boldsymbol{u} - \boldsymbol{e})')^{-1} E(\hat{R} - R)(\boldsymbol{u} - \boldsymbol{e}), \quad (16)$$

and

$$MSE(t_r^*)_{min.} = MSE(\hat{R}) - E(\hat{R} - R)(\boldsymbol{\omega} - \boldsymbol{e})'(E(\boldsymbol{\omega} - \boldsymbol{e})(\boldsymbol{\omega} - \boldsymbol{e})')^{-1} E(\hat{R} - R)(\boldsymbol{\omega} - \boldsymbol{e}). \quad (17)$$

To derive the expressions for bias mean square error of the proposed estimator t_r and t_r^* under SRSWOR upto the order (n^{-1}), we assume that:

$$\bar{y}_i^* = \bar{Y}_i + \bar{\epsilon}_i, \;\; \bar{x}_j^* = \bar{X}_j + \bar{\epsilon}_j' \text{ such that } E(\bar{\epsilon}_i) = E(\bar{\epsilon}_j') = 0; \; (i = 1,2\,;\; j = 1,2, \dots p).$$

Let $A = [a_{jj'}]$ and $A_0 = [a_{0jj'}]$ are two $p \times p$ positive definite matrix such that:

$$a_{jj'} = \frac{1-f}{n}\rho_{jj'}C_jC_{j'} + \frac{W_2(k-1)}{n}\rho_{jj'(2)}C_j'C_{j'}' \;\; \text{and} \;\; a_{0jj'} = \rho_{jj'}C_jC_{j'} \;\; \forall\; j \neq j' = 1,2, \dots, p.$$

Also let $\boldsymbol{q} = (q_1, q_2, \dots, q_p)'$ and $\boldsymbol{q}_{(2)} = (q_{1(2)}, q_{2(2)}, \dots, q_{p(2)})'$ are two column vectors such that:

$$q_j = C_j\{\rho_{1j}^* C_1^* - \rho_{2j}^* C_2^*\}, \qquad \text{and} \qquad q_{j(2)} = C_j'\{\rho_{1j(2)}^* C_1^{*\prime} - \rho_{2j(2)}^* C_2^{*\prime}\},$$

where $C_j^2 = \frac{S_j^2}{\bar{X}_j^2}$, $C_j'^2 = \frac{S_{j(2)}^2}{\bar{X}_j^2}$, $C_i^{*2} = \frac{S_j^{*2}}{\bar{Y}_i^2}$, $C_i^{*\prime 2} = \frac{S_{i(2)}^{*2}}{\bar{Y}_i^2}$, $\forall\; i = 1,2;\;\; j = 1,2, \dots p.$

Here S_j^2 and $S_{j(2)}^2$ denote the mean square error of x_j for the entire and non-responding part of the population. Let $\rho_{jj'}, \rho_{ij}^*$ are the correlation coefficients between $(x_j, x_{j'})$ and (y_i, x_j) respectively for the entire population and $\rho_{jj'(2)}, \rho_{ij(2)}^*$ are the correlation coefficients between $(x_j, x_{j'})$ and (y_i, x_j) for the non-responding group of the population.

Hence, the expressions of bias and mean square error of t_r and t_r^* upto the terms of order (n^{-1}) under SRSWOR method of sampling are given by:

$$Bias(t_r) = Bias(\hat{R}) + R\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)' g_{12}(m^*, \boldsymbol{u}^{*'}) + \frac{1}{2}\, trace\, \boldsymbol{A}\, g_{22}(m^*, \boldsymbol{u}^{*'}), \quad (18)$$

$$MSE(t_r) = MSE(\hat{R}) + (g_2(R, \boldsymbol{e}'))' \boldsymbol{A} g_2(R, \boldsymbol{e}') + 2R\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)' g_2(R, \boldsymbol{e}')\,, \quad (19)$$

$$Bias(t_r^*) = Bias(\hat{R}) + \left(\frac{1-f}{n}\right)\left[R\boldsymbol{q}' h_{12}(m^*, \boldsymbol{\omega}^{*\prime}) + \frac{1}{2}\ trace\ \boldsymbol{A}_0\ h_{22}(m^*, \boldsymbol{\omega}^{*\prime})\right], \tag{20}$$

$$MSE(t_r^*) = MSE(\hat{R}) + \left(\frac{1-f}{n}\right)\left[\left(h_2(R, \boldsymbol{e}')\right)' \boldsymbol{A}_0\ h_2(R, \boldsymbol{e}') + 2R\boldsymbol{q}' h_2(R, \boldsymbol{e}')\right], \tag{21}$$

where:

$$Bias(\hat{R}) = R\left[\frac{1-f}{n}\{C_2^{*2} - \rho C_1^* C_2^*\} + \frac{W_2(k-1)}{n}\{C_2^{*\prime 2} - \rho_{(2)} C_1^{*\prime} C_2^{*\prime}\}\right], \tag{22}$$

and $$MSE(\hat{R}) = R^2\left[\frac{1-f}{n}\{C_1^{*2} + C_2^{*2} - 2\rho C_1^* C_2^*\} + \frac{W_2(k-1)}{n}\{C_1^{*\prime 2} + C_2^{*\prime 2} - 2\rho_{(2)} C_1^{*\prime} C_2^{*\prime}\}\right]. \tag{23}$$

The conditions for which $MSE(t_r)$ and $MSE(t_r^*)$ will attain minimum values are given by:

$$g_2(R, \boldsymbol{e}') = -R\boldsymbol{A}^{-1}\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right), \tag{24}$$

and $$h_2(R, \boldsymbol{e}') = -R\ \boldsymbol{A}_0^{-1}\boldsymbol{q} \tag{25}$$

respectively. Substituting the values of $g_2(R, \boldsymbol{e}')$ and $h_2(R, \boldsymbol{e}')$ from (24) and (25) in (19) and (21), we obtain the expressions of minimum mean square error of t_r and t_r^* as:

$$MSE(t_r)_{min.} = MSE(\hat{R}) - R^2\left\{\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)' \boldsymbol{A}^{-1}\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)\right\}, \tag{26}$$

and $$MSE(t_r^*)_{min.} = MSE(\hat{R}) - R^2\frac{1-f}{n}\boldsymbol{q}'\ \boldsymbol{A}_0^{-1}\boldsymbol{q}. \tag{27}$$

3 SOME MEMBERS OF THE PROPOSED CLASSES OF ESTIMATORS

Since so many members of the proposed classes of estimators t_r and t_r^* may be possible. So following the lines of Khare and Sinha (2009), we have given some members of t_r and t_r^* which are denoted by $[T_{r1}, T_{r2}, T_{r3}]$ and $[T_{r1}^*, T_{r2}^*, T_{r3}^*]$ as:

$$T_{r1} = m\ exp\left[\sum_{j=1}^{p}\theta_{1j} log u_j\right], \tag{28}$$

$$T_{r2} = m\sum_{j=1}^{p} W_j\, u_j^{\theta_{2j}/W_j}, \qquad \sum_{i=1}^{p} W_j = 1, \tag{29}$$

$$T_{r3} = \sum_{j=1}^{p}\left[W_j u_j^{\theta_{3j}/W_j}\right]\left[m + \beta_{1j}^*(u_j - 1)\right], \tag{30}$$

$$T_{r1}^* = m\ exp\left[\sum_{j=1}^{p}\alpha_{1j} log\omega_j\right], \tag{31}$$

$$T_{r2}^* = m\sum_{j=1}^{p} W_j\omega_j^{\alpha_{2j}/W_j}, \qquad \sum_{i=1}^{p} W_j = 1, \tag{32}$$

and $$T_{r3}^* = \sum_{j=1}^{p}\left[W_j\omega_j^{\alpha_{3j}/W_j}\right]\left[m + \beta_{2j}^*(\omega_j - 1)\right]. \tag{33}$$

Here all the estimators discussed from (28) to (33) satisfy the conditions given in (5) and (7) accordingly. Hence the estimators $[T_{r1}, T_{r2}, T_{r3}]$ and $[T_{r1}^*, T_{r2}^*, T_{r3}^*]$ will attain the minimum mean square errors equal to the expressions given in (26) and (27) if their optimum values of the constants are calculated by (24) and (25) respectively. Sometimes the values of parameters in the optimum values of the constants are not known then one may estimate them on the basis of the sample values or may use past data. Reddy (1978) has shown that such values are not only stable overtime and region but also don't affect the mean square error of the estimators upto the terms of order n^{-1} (Srivastava and Jhajj, 1983).

4 COMPARISONS OF EFFICIENCY

(i) From (26) and (27), we get:

$$MSE(\hat{R}) - MSE(t_r) = R^2\left\{\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)' \boldsymbol{A}^{-1}\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)\right\} \geq 0$$

and $MSE(\hat{R}) - MSE(t_r^*) = R^2\frac{1-f}{n}\boldsymbol{q}'\,\boldsymbol{A}_0^{-1}\boldsymbol{q} \geq 0.$

(ii) Whatever be the estimator belonging to the class of estimators $t_r = g(m, \boldsymbol{u})$, the minimum mean square error will be same as given in (26). Similarly the estimator belonging to the class of estimators $t_r^* = h(m, \boldsymbol{\omega}')$ will also have minimum mean square error as given in (27).

(iii) On comparing the estimator t_r with $\hat{R}$ in terms of precision, we find that

$MSE(t_r) < MSE(\hat{R})$ iff:

$$-MSE(\hat{R}) < (g_2(R, \boldsymbol{e}'))'\boldsymbol{A}g_2(R, \boldsymbol{e}') + 2R\left(\frac{1-f}{n}\boldsymbol{q} + \frac{W_2(k-1)}{n}\boldsymbol{q}_{(2)}\right)' g_2(R, \boldsymbol{e}') < 0. \qquad (34)$$

(iv) Similarly by comparing t_r^* with respect to $\hat{R}$ in terms of precision, we see that

$MSE(t_r^*) < MSE(\hat{R})$ iff:

$$-MSE(\hat{R}) < \left(\frac{1-f}{n}\right)[(h_2(R, \boldsymbol{e}'))'\boldsymbol{A}_0 + 2R\boldsymbol{q}']h_2(R, \boldsymbol{e}') < 0. \qquad (35)$$

(v) The applicable range for the values of the constants involved in t_r and t_r^* for the better efficiency of t_r and t_r^* with respect to $\hat{R}$ can be obtained by (34) and (35).

(vi) For $W_2 = 0$, i.e. when we have complete information on study characters as well as on the auxiliary characters, then under the optimum conditions, the estimators t_r and t_r^* are equally efficient to the class of estimators proposed by Khare (1991) for R. It shows that all the members of t_r and t_r^* attain minimum mean square error for one, two or p-auxiliary characters as described in (26) and (27) if the conditions (24) and (25) are satisfied respectively.

(vii) However it is very difficult to observe the nature of relative efficiency (R. E.) of t_r with respect to t_r^* for p-auxiliary characters due to the involvement of various parameters in it. But in case of one auxiliary character (say x_j) we find that R. E.(t_r) with respect to t_r^* increases for the higher values of $\frac{\rho_{2j}^*}{\rho_{1j}^*}$, $\frac{\rho_{1j(2)}^*}{\rho_{1j}^*}$ and for the lower value of $\frac{\rho_{2j(2)}^*}{\rho_{1j(2)}^*}$, failing which t_r^* will be more efficient than t_r. So one can have a choice for using t_r or t_r^* under the different situations.

5 AN EMPIRICAL STUDY

109 Village / Town / ward wise population of urban area under Police-station – Baria, Tahasil – Champua, Orissa, India has been taken under consideration from District Census Handbook, 1981, Orissa, published by Govt. of India. The last 25% villages (i.e. 27 villages) have been considered as non-response group of the population. Here we have taken the study characters and auxiliary characters as follows:

y_1 : Number of literate persons in the village,
y_2 : Number of main workers in the village,
x_1 : Number of non-workers in the village,
x_2 : Total population of the village and
x_3 : Number of cultivators in the village.

For this population, we have:

$\bar{Y}_1$= 145.3028	$\bar{Y}_2$= 165.2661	$\bar{X}_1$= 259.0826	$\bar{X}_2$= 485.9174	$\bar{X}_3$= 100.5505	
C_1^*= 0.7666	C_2^* = 0.6828	C_1= 0.7645	C_2 = 0.6590	C_3 = 0.7314	
$C_1^{*\prime}$ = 0.6899	$C_2^{*\prime}$ = 0.5769	C_1' = 0.5429	C_2' = 0.4877	C_3' = 0.5678	
ρ_{11}^*= 0.905	ρ_{12}^* = 0.905	ρ_{13}^* = 0.648	ρ_{21}^* = 0.819	ρ_{22}^* = 0.908	ρ_{23}^* = 0.841
$\rho_{11(2)}^*$ = 0.875	$\rho_{12(2)}^*$ = 0.871	$\rho_{13(2)}^*$ = 0.382	$\rho_{21(2)}^*$ = 0.746	$\rho_{22(2)}^*$ = 0.907	$\rho_{23(2)}^*$ = 0.785
ρ_{12}= 0.946	ρ_{13} = 0.732	ρ_{23} = 0.801	$\rho_{12(2)}$ = 0.905	$\rho_{13(2)}$ = 0.488	$\rho_{23(2)}$ = 0.654
		ρ = 0.816	$\rho_{(2)}$ = 0.787		

The present problem is to estimate the ratio of the two population means i.e. $R(=\bar{Y}_1\ /\ \bar{Y}_2)$ using x_1, x_2 and x_3. The estimators:

$$T_{r1} = m\ exp\left[\sum_{j=1}^{p}\theta_{1j}logu_j\right],$$

and

$$T_{r1}^* = m\ exp\left[\sum_{j=1}^{p}\alpha_{1j}log\omega_j\right],$$

which are the member of the classes of estimator t_r and t_r^* respectively have been considered for comparing their relative efficiency with respect to $\hat{R}$.

The optimum values of the constants [θ_{1j} and α_{1j}] and relative efficiency of T_{r1} and T_{r1}^* with respect to $\hat{R}$ for the different values of sub-sampling fraction $1\ /\ k$ have been given in Table 1.

DISCUSSION AND CONCLUSION

It has been observed from Table 1 that the estimators T_{r1} and T_{r1}^* are more efficient than $\hat{R}$ for all the different values of the sub-sampling fraction $(1\ /\ k)$. The mean square error of T_{r1} and T_{r1}^* decreases as the sub-sampling fraction and the numbers of auxiliary characters increase. It has also been observed from the Table 1 that the relative efficiency of T_{r1} and T_{r1}^* increases when numbers of auxiliary characters increase. On comparing the relative efficiency of T_{r1} and T_{r1}^* with respect to $\hat{R}$, we observe that R. E.(T_{r1}^*) with respect to $\hat{R}$ increases as sub-sampling fraction increases but R. E.(T_{r1}) with respect to $\hat{R}$ decreases as sub-sampling fraction increases. This is due to the fact that $MSE(\hat{R})$ decreases at a faster rate than $MSE(T_{r1})$ as sub-sampling fraction increases. Since T_{r1} and T_{r1}^* are the particular members of the proposed classes of estimators t_r and t_r^* respectively, so on the basis of theoretical and empirical discussions we may recommend the use of t_r and t_r^* in the case of large sample surveys.

ACKNOWLEDGEMENT

The authors are thankful to the editor and two anonymous referees for their valuable comments and suggestions that helped to improve our manuscript.

Table 1 Optimum values of the constants and relative efficiency [R. E. (.)] in % of T_{r1} and $T_r{}^*{}_1$ with respect to R [Figures in the parenthesis give MSE(–) in 10^{-5}]

Estimators	Auxiliary character(s)	$N = 109,\ n = 30$											
		$n = 1/k$											
		1 / 4				1 / 3				1 / 2			
		Optimum values of constants			R. E.	Optimum values of constants			R. E.	Optimum values of constants			R. E.
$\hat{R}$	–				100.00 (722)				100.00 (605)				100.00 (489)
	x_1	$\theta_{11}= -0.2260$			113.17 (638)	$\theta_{11} = -0.2136$			112.45 (538)	$\theta_{11}= -0.1976$			111.39 (439)
T_{r1}	x_1 , x_2	$\theta_{11}= -0.9389$	$\theta_{12}= 0.8677$		140.47 (514)	$\theta_{11}= -0.9063$	$\theta_{12}= 0.8446$		137.50 (440)	$\theta_{11}= -0.8617$	$\theta_{12}= 0.8117$		137.97 (365)
	x_1, x_2, x_3	$\theta_{11}=$ −0.7205	$\theta_{12}=$ 0.1958	$\theta_{13}=$ 0.5264	224.22 (322)	$\theta_{11}=$ −0.7184	$\theta_{12}=$ 0.2059	$\theta_{13}=$ 0.5162	213.03 (284)	$\theta_{11}=$ −0.7146	$\theta_{12}=$ 0.2203	$\theta_{13}=$ 0.5135	199.59 (245)
	x_1	$\alpha_{11}= -0.1761$			104.94 (688)	$\alpha_{11}= -0.1761$			105.96 (571)	$\alpha_{11}= -0.1761$			107.47 (455)
T^*_{r1}	x_1, x_2	$\alpha_{11}= -0.7963$	$\alpha_{12}= 0.7604$		113.17 (638)	$\alpha_{11}= -0.7963$	$\alpha_{12}= 0.7604$		116.12 (521)	$\alpha_{11}= -0.7963$	$\alpha_{12}= 0.7604$		120.74 (405)
	x_1, x_2, x_3	$\alpha_{11}=$ −0.7064	$\alpha_{12}=$ 0.2427	$\alpha_{13}=$ 0.4734	129.39 (558)	$\alpha_{11}=$ −0.7064	$\alpha_{12}=$ 0.2427	$\alpha_{13}=$ 0.4734	137.19 (441)	$\alpha_{11}=$ −0.7064	$\alpha_{12}=$ 0.2427	$\alpha_{13}=$ 0.4734	150.46 (325)

Source: Own construction

References

HANSEN, M. H., HURWITZ, W. N. The Problem of Nonresponse in Sample Surveys. *Jour. Amer. Statist. Assoc.,* 1946, 41, pp. 517–529.

HARTLEY, H. O., ROSS, A. Unbiased Ratio Estimators. *Nature,* 1954, 174, pp. 270–271.

KHARE, B. B. On Generalized Class of Estimators for Ratio of Two Population Means Using Multi-Auxiliary Characters. *Aligarh J. Stat.*, 1991, 11, pp. 81–90.

KHARE, B. B., SINHA, R. R. Estimation of Finite Population Ratio Using Two Phase Sampling Scheme in the Presence of Non-Response. *Aligarh J. Stat.*, 2004, 24, pp. 43–56.

KHARE, B. B., SINHA, R. R. Estimation of the Ratio of Two Population Means Using Auxiliary Character with Unknown Population Mean in Presence of Non-response. *Progress of Mathematics*, 2002, 36, pp. 337–348.

KHARE, B. B., SINHA, R. R. On Class of Estimators for Population Mean Using Multi-Auxiliary Character in Presence of Non-Response. *Statistics in Transition – New Series*, 2009, 10, pp. 3–14.

KHARE, B. B., SRIVASTAVA, S. Estimation of Population Mean Using Auxiliary Character in Presence of Nonresponse. *Nat. Acad. Sc. Letters*, 1993, 16(3), pp.111–114.

KHARE, B. B., SRIVASTAVA, S. Generalized Estimators for Population Mean in Presence of Non-Response. *Internal. J. Math & Statist. Sc.*, 2000. 9(1), pp. 75–87.

KHARE, B. B., SRIVASTAVA, S. Study of Conventional and Alternative Two Phase Sampling Ratio, Product and Regression Estimators in Presence of Non-Response. *Proc. Nat. Acad. Sci.*, India, 1995, 65(a) ii, pp. 195–203.

KHARE, B. B., SRIVASTAVA, S. Transformed Ratio Type Estimators for the Population Mean in the Presence of Non-Response. *Commun. Statis.-Theory Math.*, 1997, 26(7), pp. 1779–1791.

RAO, P. S. R. S. Ratio Estimation with Subsampling the Nonrespondents. *Survey Methodology*, 1986, 12(2), pp. 217–230.

RAO, P. S. R. S. Regression Estimators with Subsampling of Nonrespondents. In GUNAR LIEPINS, E., UPPULURI, V. R. R., eds. *Data Quality Control, Theory and Pragmatics.* New York: Marcel Dekker Inc., 1990, pp. 191–208.

REDDY, V. N. A Study of Use of Prior Knowledge on Certain Population Parameters in Estimation. *Sankhaya*, C, 1978, 40, pp. 29–37.

SINGH, H. P., KUMAR, S. A General Class of Estimators of Population Mean in Survey Sampling Using Auxiliary Information with Sub Sampling the Non-Respondents. *Korean J. Appl. Stat.*, 2009, 22, pp. 387–402.

SINGH, M. P. On the Estimation of Ratio and Product of Population Parameters. *Sankhya*, Ser. C, 1965, 27, pp. 321–328.

SRIVASTAVA, S. K., JHAJJ, H. S. A Class of Estimators of the Population Mean Using Multi-Auxiliary Information. *Cal. Stat. Assoc. Bull.*, 1983, 32, pp. 47–56.

TOUTENBERG, H., SRIVASTAVA, V. K. Estimation of Ratio of Population Means in Survey Sampling when Some Observations Are Missing. *Metrika*, 1988, 48, pp. 177–187.

TRIPATHI, T. P. *Contribution to the Sampling Theory Using Multivariate Information.* Ph.D. Thesis submitted to Punjabi University, Patiala, India, 1970.

TRIPATHI, T. P., CHATURVEDI, D. K. Use of Multivariate Auxiliary Information in Estimating the Population Ratio. Stat-Math. Tech. Report No. 24 / 79, I. S. I., Calcutta, India, Abs. *J. Indian Soc. Agricultural Statist.*, 1979, bf. 31.