Ratio Cum Regression Estimator for Estimating a Population Mean with a Sub Sampling of Non Respondents

Sunil Kumar[1,a]

[a]Department of Statistics, University of Jammu

Abstract

In the present study, a combined ratio cum regression estimator is proposed to estimate the population mean of the study variable in the presence of a non-response using an auxiliary variable under double sampling. The expressions of bias and mean squared error(MSE) based on the proposed estimator is derived under double (or two stage) sampling to the first degree of approximation. Some estimators are also derived from the proposed class by allocating the suitable values of constants used. A comparison of the proposed estimator with the usual unbiased estimator and other derived estimators is carried out. An empirical study is carried out to demonstrate the performance of the suggested estimator and of others; it is endow that the empirical results backing the theoretical study.

Keywords: Study variable, auxiliary variable, ratio estimator, regression estimator, mean squared error, double sampling, non-response.

1. Introduction

Auxiliary information is often used to improve the precision or accuracy of the estimator of unknown population parameter of interest. In this direction, early work was done by Watson (1937), Cochran (1940), Robson (1957) and Murthy (1964) that introduced the concept of a usual regression estimator, ratio estimator and product estimator of the population mean $\bar{Y}$, respectively. All of them utilized auxiliary information at the estimation stage. Cochran (1940) developed the ratio estimator of a population mean or total of the study variable y using supplementary information on an auxiliary variable x, positively correlated with y. It was proven that if the relationship between y and x is a straight line passing through the neighborhood of the origin and the variance of y about this line is proportional to the auxiliary variable x, then the ratio estimator is as good as a regression variable. However, in many situations of practical importance, the regression line doesnot pass through the neighbourhood of the origin. In these situations, a ratio estimator does not perform equally as well as the regression estimator. This fact, motivated by various authors such as Singh (1965), Gupta (1978, 1979), Sahai (1979),Srivastava (1967), Reddy (1973, 1974), Walsh (1970), Srivenkataramana (1980), suggest some modified ratio type estimators in order to provide better alternatives.

It is common experience in surveys that data cannot always be collected for all units selected in the sample. Thus, the selected farmers or families may not be found at home at the first attempt and some may refuse to cooperate with the interviewer even if contacted. In general during surveys, it is observed that information in most cases is not obtained at the first attempt even after some callbacks. An estimate obtained from such incomplete data may be misleading because of the biased estimator.

---

[1] Assistant Professor, Department of Statistics, University of Jammu, J & K, India. E-mail: sunilbhougal06@gmail.com

This is the case of a non-response and the usual approach to face the non-response to recontact the non-respondents and obtain the information as much as possible see, Tripathi and Khare (1997) and Tabasum and Khan (2004, 2006). Hansen and Hurwitz (1946) considered the problem of non-response while estimating the population mean by taking a sub sample from the non-respondent group with the help of some extra effort. An unbiased estimator was suggested by combining the information available from response and non-response groups. Further, rectification in the non-response using an auxiliary variable was proposed by Cochran (1977), Rao (1986, 1987), Khare and Srivastava (1993, 1995, 1997), Okafor and Lee (2000), Tabasum and Khan (2004, 2006), Singh and Kumar (2008a, 2008b, 2009a, 2009b, 2010a, 2010b, 2011), Kumar and Singh (2010) and Singh *et al.* (2010) using Hansen and Hurwitz (1946) technique.

From a finite population $U = (U_1, U_2, \ldots, U_N)$ of size N, a large first phase sample of size n' is selected by simple random sampling without replacement(SRSWOR). A smaller second phase sample of size n is selected from n' by SRSWOR. Non-response occurs on the second phase sample of size n in which n_1 units respond and n_2 units do not. From the n_2 non-respondents, by SRSWOR, a sample of $r = n_2/k; k > 1$ units is selected where k is the inverse sampling rate. The second phase sample of size n. All the r units respond this time round. Here, we have $(n_1 + r)$ responding units on the study variable y and consequently the estimator for the population mean $\bar{Y}$ of the study variable y using the sub sampling scheme envisaged by Hansen and Hurwitz (1946) as defined as

$$\bar{y}^* = w_1 \bar{y}_1 + w_2 \bar{y}_{2r},$$

where $w_1 = n_1/n$; $w_2 = n_2/n$; $\bar{y}_1 = \sum_{i=1}^{n_1} y_i/n_1$ and $\bar{y}_{2r} = \sum_{i=1}^{r} y_i/r$. The estimator $\bar{y}^*$ is the unbiased estimator of the population mean $\bar{Y}$ of the study variable y and has the variance as given by

$$\text{Var}(\bar{y}^*) = \lambda S_y^2 + \theta S_{y(2)}^2,$$

where $\lambda = (1 - f)/n$; $f = n/N$; $\theta = W_2(k-1)/n$; $W_2 = N_2/N$; $S_y^2 = 1/(N-1) \sum_{i=1}^{N} (y_i - \bar{Y})^2$; $S_{y(2)}^2 = \sum_{i=1}^{N=N_1+N_2} (y_i - \bar{Y}_2)^2/(N_2 - 1)$; $\bar{Y} = \sum_{i=1}^{N} y_i/N$; $\bar{Y}_2 = 1/N_2 \sum_{i=1}^{N=N_1+N_2} y_i$; $\bar{y}_{2r} = \sum_{i=1}^{r} y_i/r$; N_1 and $N_2 (= N - N_1)$ are the sizes of the responding and non-responding units from the finite population N.

In estimating the population mean, sample survey experts sometimes use auxiliary information to improve the precision of the estimates. Let x denote an auxiliary variable with population mean $\bar{X} = \sum_{i=1}^{N} x_i/N$. Let $\bar{X}_1 = \sum_{i=1}^{N_1} x_i/N_1$ and $\bar{X}_2 = \sum_{i=1}^{N_2} x_i/N_2$ denote the population means of the response and non-response groups (or strata). Let $\bar{x} = \sum_{i=1}^{n} x_i/n$ denote the mean of all the n units. Let $\bar{x}_1 = \sum_{i=1}^{n_1} x_i/n_1$ and $\bar{x}_2 = \sum_{i=1}^{n_2} x_i/n_2$ denote the means of the n_1 responding units and n_2 non-responding units. Further, let $\bar{x}_{2r} = \sum_{i=1}^{r} x_i/r$ denote the means of the $r (= n_2/k); k > 1$ sub-sampled units. With this background, one can define an unbiased estimator of population mean $\bar{X}$ as

$$\bar{x}^* = w_1 \bar{x}_1 + w_2 \bar{x}_{2r}.$$

The variance of $\bar{x}^*$ is given by

$$\text{Var}(\bar{x}^*) = \lambda S_x^2 + \theta S_{x(2)}^2,$$

where $S_x^2 = 1/(N-1) \sum_{i=1}^{N} (x_i - \bar{X})^2$; $S_{x(2)}^2 = \sum_{i=1}^{N=N_1+N_2} (x_i - \bar{X}_2)^2/(N_2 - 1)$.

In the present study, a general combined ratio and regression estimator is suggested to estimate the population mean $\bar{Y}$ of the study variable y in the presence of a non-response. To the first degree of approximation, the expressions of bias and mean squared error(MSE) of the suggested estimator have been obtained. Some estimators are shown to be particular members of this family. A comparison of the proposed study is also presented to expound the performance of the proposed estimator. An empirical study is also carried to demonstrate the performance of the suggested estimator.

2. The Proposed Class of Estimators

The proposed class of estimators for estimating the population mean $\bar{Y}$ of the study variable y is given by

$$T_{CRR} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\frac{a\bar{x}^* + b}{a\bar{x}' + b}\right)^\alpha \left(\frac{a\bar{x}' + b}{a\bar{x} + b}\right)^\beta$$

or

$$T_{CRR} = T_{OL}\left(\frac{a\bar{x}^* + b}{a\bar{x}' + b}\right)^\alpha \left(\frac{a\bar{x}' + b}{a\bar{x} + b}\right)^\beta, \tag{2.1}$$

where $T_{OL} = \bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)$ is the regression estimator proposed by Okafor and Lee (2000), $\bar{x}'$ denote the sample mean of x based on functions of known parameters such as standard deviation (σ), etc. of the auxiliary variable x, and α, β are suitable chosen constants, $b_{yx}^* = s_{yx}^*/s_x^{*2}$ is an estimator of population regression coefficient; $\beta_{yx} = S_{yx}/S_x^2$; $s_{yx}^* = 1/(n-1)(\sum_{i=1}^n x_i y_i + r\sum_{i=1}^r x_i y_i - n\bar{x}\bar{y}^*)$; $s_x^{*2} = 1/(n-1)(\sum_{i=1}^n x_i^2 + r\sum_{i=1}^r x_i^2 - n\bar{x}\bar{x}^*)$; $S_{yx} = 1/(N-1)\sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y})$; $S_x^2 = 1/(N-1)\sum_{i=1}^N (x_i - \bar{X})^2$.

To obtain the bias and MSE of the class of estimator T_{CRR}, we have

$$\bar{y}^* = \bar{Y}(1 + \epsilon_0); \qquad \bar{x}^* = \bar{X}(1 + \epsilon_1^*); \qquad \bar{x} = \bar{X}(1 + \epsilon_1);$$
$$\bar{x}' = \bar{X}(1 + \epsilon_1'); \qquad s_{xy}^* = S_{xy}(1 + \epsilon_2); \qquad s_x^{*2} = S_x^2(1 + \epsilon_3),$$

such that

$$E(\epsilon_0) = E(\epsilon_1^*) = E(\epsilon_1) = E(\epsilon_1') = E(\epsilon_2) = E(\epsilon_3) = 0$$

and

$$E\left(\epsilon_0^2\right) = \frac{1}{\bar{Y}^2}\left(\lambda S_y^2 + \theta S_{y(2)}^2\right); \quad E\left(\epsilon_1^{*2}\right) = \frac{1}{\bar{X}^2}\left(\lambda S_x^2 + \theta S_{x(2)}^2\right); \quad E\left(\epsilon_1^2\right) = \frac{\lambda}{\bar{X}^2}S_x^2; \quad E\left(\epsilon_1'^2\right) = \frac{\lambda'}{\bar{X}^2}S_x^2;$$

$$E\left(\epsilon_0\epsilon_1^*\right) = \lambda\frac{S_{yx}}{\bar{Y}\bar{X}} + \theta\frac{S_{yx(2)}}{\bar{Y}\bar{X}}; \quad E\left(\epsilon_0\epsilon_1\right) = \lambda\frac{S_{yx}}{\bar{Y}\bar{X}}; \quad E\left(\epsilon_0\epsilon_1'\right) = \lambda'\frac{S_{yx}}{\bar{Y}\bar{X}}; \quad E\left(\epsilon_1^*\epsilon_1'\right) = \lambda'\frac{S_x^2}{\bar{X}^2};$$

$$E\left(\epsilon_1\epsilon_1^*\right) = \frac{\lambda}{\bar{X}^2}S_x^2; \quad E\left(\epsilon_1\epsilon_1'\right) = \lambda'\frac{S_x^2}{\bar{X}^2}; \quad E\left(\epsilon_1'\epsilon_2\right) = \frac{N(N-n')}{(N-1)(N-2)}\frac{\mu_{31}}{n'\bar{X}S_{yx}};$$

$$E\left(\epsilon_1^*\epsilon_3\right) = \frac{N(N-n)}{(N-1)(N-2)}\frac{\mu_{30}}{n\bar{X}S_x^2} + \frac{W_2(k-1)}{n}\frac{\mu_{30(2)}}{\bar{X}S_x^2}; \quad E\left(\epsilon_1'\epsilon_3\right) = \frac{N(N-n')}{(N-1)(N-2)}\frac{\mu_{30}}{n'\bar{X}S_x^2};$$

$$E\left(\epsilon_1^*\epsilon_2\right) = \frac{N(N-n)}{(N-1)(N-2)}\frac{\mu_{21}}{n\bar{X}S_{xy}} + \frac{W_2(k-1)}{n}\frac{\mu_{21(2)}}{\bar{X}S_{xy}},$$

where $\lambda' = (1 - f')/n$; $f' = n'/N$; $\mu_{rs} = 1/N\sum_{i=1}^{N=N_1+N_2}(y_i - \bar{Y})^s(x_i - \bar{X})^r$; $\mu_{rs(2)} = 1/N_2\sum_{i=1}^{N=N_1+N_2}(y_i - \bar{Y}_2)^s(x_i - \bar{X}_2)^r$; $\bar{X}_2 = 1/N_2\sum_{i=1}^{N=N_1+N_2} x_i$; $\bar{Y}_2 = 1/N_2\sum_{i=1}^{N=N_1+N_2} y_i$; (r, s) being non-negative integers.

Now expressing T_{CRR} in terms of $\epsilon's$, we have

$$T_{CRR} = \bar{Y}\left\{1 + \epsilon_0 + K_{yx}(1 + \epsilon_2)(1 + \epsilon_3)^{-1}(\epsilon_1' - \epsilon_1^*)\right\}(1 + \emptyset\epsilon_1^*)^\alpha(1 + \emptyset\epsilon_1^*)^{-\alpha}(1 + \emptyset\epsilon_1)^\beta(1 + \emptyset\epsilon_1')^{-\beta}, \tag{2.2}$$

where $\emptyset = (a\bar{X})/(a\bar{X} + b)$; $K_{yx} = \beta_{yx}/R$; $R = \bar{Y}/\bar{X}$.

We assume that $|\epsilon_3| < 1$, $|Ø\epsilon_1^*| < 1$, $|Ø\epsilon_1| < 1$ and $|Ø\epsilon_1'| < 1$ so that the right hand side of (2.2) to the first degree of approximation, we have

$$
\begin{aligned}
\left(T_{CRR} - \bar{Y}\right) = \bar{Y}\Big[& \epsilon_0 + K_{yx}\left(\epsilon_1' - \epsilon_1^* + \epsilon_2\epsilon_1' - \epsilon_2\epsilon_1^* - \epsilon_1'\epsilon_3 + \epsilon_1^*\epsilon_3\right) \\
& + \alpha Ø\left\{\epsilon_1^* - \epsilon_1' + \epsilon_0\epsilon_1^* - \epsilon_0\epsilon_1' + K_{yx}\left(\epsilon_1^*\epsilon_1' - \epsilon_1^{*2} - \epsilon_1'^2 + \epsilon_1'\epsilon_1^*\right)\right\} \\
& + \beta Ø\left\{\epsilon_1 - \epsilon_1' + \epsilon_0\epsilon_1 - \epsilon_0\epsilon_1' + K_{yx}\left(\epsilon_1'\epsilon_1 - \epsilon_1'^2 - \epsilon_1\epsilon_1^* + \epsilon_1'\epsilon_1^*\right)\right\} \\
& + Ø^2\left(\alpha^2\epsilon_1'\epsilon_1^* - \beta^2\epsilon_1'\epsilon_1\right) + \frac{\alpha(\alpha+1)}{2}Ø^2\left(\epsilon_1^{*2} + \epsilon_1'^2\right) + \frac{\beta(\beta+1)}{2}Ø^2\left(\epsilon_1^2 + \epsilon_1'^2\right) \\
& + \alpha\beta Ø^2\left(\epsilon_1^*\epsilon_1 - \epsilon_1^*\epsilon_1' - \epsilon_1'\epsilon_1 + \epsilon_1'^2\right)\Big].
\end{aligned}
\tag{2.3}
$$

Taking expectations of both sides of (2.3), we get the bias of T_{CRR} to the first degree of approximation is given by

$$
\begin{aligned}
B\left(T_{CRR}\right) = \frac{1}{\bar{X}}\Big[& R\alpha Ø\theta\left(\beta_{yx(2)} + \beta_{yx}\right)S_{x(2)}^2 + RØ^2\left(\alpha^2 - \beta^2\right)\lambda'S_x^2 \\
& + \frac{\alpha(\alpha+1)}{2}RØ^2\left\{(\lambda + \lambda')S_x^2 + \theta S_{x(2)}^2\right\} + \frac{\beta(\beta+1)}{2}RØ^2(\lambda + \lambda')S_x^2 + R\alpha\beta Ø^2(\lambda - \lambda')S_x^2 \\
& + RK_{yx}\left\{\frac{N}{(N-1)(N-2)}(\lambda - \lambda')\left(\frac{\mu_{30}}{S_x^2} - \frac{\mu_{21}}{S_{yx}}\right) + \theta\left(\frac{\mu_{30(2)}}{S_x^2} - \frac{\mu_{21(2)}}{S_{yx}}\right)\right\}\Big],
\end{aligned}
\tag{2.4}
$$

where $\beta_{yx(2)} = S_{yx(2)}/S_{x(2)}^2$; $S_{yx(2)} = \sum_{i=1}^{N_2}(x_i - \bar{X}_2)(y_i - \bar{Y}_2)/(N_2 - 1)$.

Squaring both sides of (2.3) and neglecting terms of $\epsilon's$ involving power greater than two, we have

$$
\begin{aligned}
\left(T_{CRR} - \bar{Y}\right)^2 = \bar{Y}^2\Big[& \epsilon_0^2 + K_{yx}^2\left(\epsilon_1' - \epsilon_1^*\right)^2 + 2K_{yx}\left(\epsilon_0\epsilon_1' - \epsilon_0\epsilon_1^*\right) + \alpha^2 Ø^2\left(\epsilon_1^* - \epsilon_1'\right)^2 + \beta^2 Ø^2(\epsilon_1 - \epsilon_1')^2 \\
& + \alpha Ø\left\{2\left(\epsilon_0\epsilon_1^* - \epsilon_0\epsilon_1'\right) + 2K_{yx}\left(\epsilon_1'\epsilon_1^* - \epsilon_1'^2 - \epsilon_1^{*2} + \epsilon_1'\epsilon_1^*\right)\right\} + \beta\alpha Ø2\left\{\left(\epsilon_0\epsilon_1 - \epsilon_0\epsilon_1'\right)\right. \\
& \left. + 2K_{yx}\left(\epsilon_1'\epsilon_1 - \epsilon_1'^2 - \epsilon_1^*\epsilon_1 + \epsilon_1'\epsilon_1^*\right)\right\} + 2\beta\alpha Ø^2\left(\epsilon_1^*\epsilon_1 - \epsilon_1'\epsilon_1^* - \epsilon_1'\epsilon_1 + \epsilon_1'^2\right)\Big].
\end{aligned}
\tag{2.5}
$$

Taking expectations of both sides of (2.5), we get the MSE of T_{CRR} to the first degree of approximation, we get

$$
\begin{aligned}
\mathrm{MSE}(T_{CRR}) = \Big[& (\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + R^2Ø^2(\alpha + \beta)^2 S_x^2\right\} \\
& + \theta\left\{S_{y(2)}^2 + \left(\beta_{yx} - \alpha ØR\right)\left(\beta_{yx} - \alpha ØR - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} + \lambda'S_y^2\Big]
\end{aligned}
\tag{2.6}
$$

which is minimum when

$$
\beta = -\alpha \quad \text{and} \quad \alpha = \frac{1}{ØR}\left(\beta_{yx} - \beta_{yx(2)}\right).
\tag{2.7}
$$

Substituting equation (2.7) in (2.6), we get the minimum MSE of T_{CRR}, is given by

$$
\min \mathrm{MSE}(T_{CRR}) = \left[(\lambda - \lambda')\left(1 - \rho_{yx}^2\right)S_y^2 + \theta\left(1 - \rho_{yx(2)}^2\right)S_{y(2)}^2 + \lambda'S_y^2\right].
\tag{2.8}
$$

Table 1: The estimators of the population mean that can be obtained by suitable choices of constants α, β, a and b

Estimator	α	β	a	b
$T_{CRR}^{(0)} = \bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)$ Okafor and Lee's (2000) regression estimator	0	0	a	b
$T_{CRR}^{(1)} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\dfrac{\bar{x}'}{\bar{x}^*}\right)$	-1	0	1	0
$T_{CRR}^{(2)} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\dfrac{\bar{x}'}{\bar{x}}\right)$	0	-1	1	0
$T_{CRR}^{(3)} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\dfrac{\bar{x}'}{\bar{x}^*}\right)\left(\dfrac{\bar{x}'}{\bar{x}}\right)$	-1	-1	1	0
$T_{CRR}^{(4)} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\dfrac{\bar{x}' + \rho_{yx}}{\bar{x}^* + \rho_{yx}}\right)$	-1	0	1	ρ_{yx}
$T_{CRR}^{(5)} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\dfrac{\bar{x}' + \rho_{yx}}{\bar{x} + \rho_{yx}}\right)$	0	-1	1	ρ_{yx}
$T_{CRR}^{(6)} = \left\{\bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)\right\}\left(\dfrac{\bar{x}' + \rho_{yx}}{\bar{x}^* + \rho_{yx}}\right)\left(\dfrac{\bar{x}' + \rho_{yx}}{\bar{x} + \rho_{yx}}\right)$	-1	-1	1	ρ_{yx}

ρ_{yx} is the correlation coefficient between the study variable y and the auxiliary variable x

3. Some Members of the Proposed Class of Estimators T_{CRR}

Table 1 is the estimators of the population mean that can be obtained by suitable choices of constants α, β, a and b.

More estimators can also be generated from the proposed estimator in (2.1) just by substituting different values of α, β, a and b, respectively. Expressions of MSE's of the above defined estimators can be obtained by substituting the values of α, β, a and b in (2.6). Thus, the expressions of MSE of different estimators to the first degree of approximation are as follows:

$$\text{MSE}\left(T_{CRR}^{(0)}\right) = \left[(\lambda - \lambda')\left(1 - \rho_{yx}^2\right)S_y^2 + \theta\left\{S_{y(2)}^2 + \beta_{yx}\left(\beta_{yx} - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} + \lambda' S_y^2\right], \tag{3.1}$$

$$\text{MSE}\left(T_{CRR}^{(1)}\right) = \left[(\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + R^2 S_x^2\right\} + \theta\left\{S_{y(2)}^2 + \left(\beta_{yx} + R\right)\left(\beta_{yx} + R - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} \right.$$
$$\left. + \lambda' S_y^2\right], \tag{3.2}$$

$$\text{MSE}\left(T_{CRR}^{(2)}\right) = \left[(\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + R^2 S_x^2\right\} + \theta\left\{S_{y(2)}^2 + \left(\beta_{yx} + R\right)\left(\beta_{yx} - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} \right.$$
$$\left. + \lambda' S_y^2\right], \tag{3.3}$$

$$\text{MSE}\left(T_{CRR}^{(3)}\right) = \left[(\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + 4R^2 S_x^2\right\} + \theta\left\{S_{y(2)}^2 + \left(\beta_{yx} + R\right)\left(\beta_{yx} + R - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} \right.$$
$$\left. + \lambda' S_y^2\right], \tag{3.4}$$

$$\text{MSE}\left(T_{CRR}^{(4)}\right) = \left[(\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + R^2 \emptyset^{*2} S_x^2\right\} \right.$$
$$\left. + \theta\left\{S_{y(2)}^2 + \left(\beta_{yx} + \emptyset^* R\right)\left(\beta_{yx} + \emptyset^* R - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} + \lambda' S_y^2\right], \tag{3.5}$$

$$\text{MSE}\left(T_{CRR}^{(5)}\right) = \left[(\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + R^2 \emptyset^{*2} S_x^2\right\} + \theta\left\{S_{y(2)}^2 + \beta_{yx}\left(\beta_{yx} - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} + \lambda' S_y^2\right], \tag{3.6}$$

$$\text{MSE}\left(T_{CRR}^{(6)}\right) = \left[(\lambda - \lambda')\left\{\left(1 - \rho_{yx}^2\right)S_y^2 + 4R^2 \emptyset^{*2} S_x^2\right\} \right.$$
$$\left. + \theta\left\{S_{y(2)}^2 + \left(\beta_{yx} + \emptyset^* R\right)\left(\beta_{yx} + \emptyset^* R - 2\beta_{yx(2)}\right)S_{x(2)}^2\right\} + \lambda' S_y^2\right], \tag{3.7}$$

where $\emptyset^* = \{\bar{X}/(\bar{X} + \rho_{yx})\}$.

4. Efficiency Comparison

The proposed class of estimator T_{CRR} is more efficient than

i) The usual estimator $\bar{y}^*$, if

$$\frac{-B - \sqrt{B^2 - AC}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC}}{\text{Ø}RA} \tag{4.1}$$

where $A = (\lambda - \lambda')S_x^2 + \theta S_{x(2)}^2$; $B = (\lambda - \lambda')R\text{Ø}\beta S_x^2 - \theta(\beta_{yx} - \beta_{yx(2)})S_{x(2)}^2$; $C = (\lambda - \lambda')(R^2\text{Ø}^2\beta^2 S_x^2 - \rho_{yx}^2 S_x^2) + \theta(\beta_{yx} - \beta_{yx(2)})\beta_{yx}S_{x(2)}^2$.

ii) The regression estimator $T_{CRR}^{(0)}$ if

$$\frac{-B - \sqrt{B^2 - AC^*}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC^*}}{\text{Ø}RA}, \tag{4.2}$$

where $C^* = (\lambda - \lambda')R^2\text{Ø}^2\beta^2 S_x^2$.

iii) The estimator $T_{CRR}^{(1)}$ if

$$\frac{-B - \sqrt{B^2 - AC'}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC'}}{\text{Ø}RA}, \tag{4.3}$$

where $C' = \{(\lambda - \lambda')R^2(\text{Ø}^2\beta^2 - 1)S_x^2 + 2\theta R(\beta_{yx(2)} - \beta_{yx})S_{x(2)}^2\}$.

iv) The estimator $T_{CRR}^{(2)}$ if

$$\frac{-B - \sqrt{B^2 - AC^{**}}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC^{**}}}{\text{Ø}RA}, \tag{4.4}$$

where $C^{**} = (\lambda - \lambda')R^2(\text{Ø}^2\beta^2 - 1)S_x^2$.

v) The estimator $T_{CRR}^{(3)}$ if

$$\frac{-B - \sqrt{B^2 - AC''}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC''}}{\text{Ø}RA} \tag{4.5}$$

where $C'' = (\lambda - \lambda')(\text{Ø}^2\beta^2 - 4)R^2 S_x^2 + \theta(2\beta_{yx(2)} - 2\beta_{yx} - R)RS_{x(2)}^2$.

vi) The estimator $T_{CRR}^{(4)}$ if

$$\frac{-B - \sqrt{B^2 - AC_1^*}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC_1^*}}{\text{Ø}RA}, \tag{4.6}$$

where $C_1^* = (\lambda - \lambda')R^2(\text{Ø}^2\beta^2 - \text{Ø}^{*2})S_x^2 + \theta(\text{Ø}^*R - 2\beta_{yx} + 2\beta_{yx(2)})\text{Ø}^*RS_{x(2)}^2$.

vii) The estimator $T_{CRR}^{(5)}$ if

$$\frac{-B - \sqrt{B^2 - AC_1'}}{\text{Ø}RA} < \alpha < \frac{-B + \sqrt{B^2 - AC_1'}}{\text{Ø}RA}, \tag{4.7}$$

Table 2: Percent relative efficiency(PRE) of the estimators with respect to $\bar{y}^*$.

Estimator	$N = 100; n' = 50; n = 30$			
	(1/k)			
	(1/5)	(1/4)	(1/3)	(1/2)
$\bar{y}^*$	100.00	100.00	100.00	100.00
$T_{CRR}^{(0)}$	432.37	380.54	328.12	275.08
$T_{CRR}^{(1)}$	91.75	92.51	93.54	95.02
$T_{CRR}^{(2)}$	192.52	168.90	145.17	121.31
$T_{CRR}^{(3)}$	51.17	48.34	45.02	41.07
$T_{CRR}^{(4)}$	92.29	93.03	94.04	95.49
$T_{CRR}^{(5)}$	193.33	169.62	145.78	121.83
$T_{CRR}^{(6)}$	51.51	48.66	45.31	41.34
$T_{CRR}^{(opt)}$	432.92	380.91	328.33	275.17

where $C_1' = (\lambda - \lambda')R^2(\emptyset^2\beta^2 - \emptyset^{*^2})S_x^2$.

viii) The estimator $T_{CRR}^{(6)}$ if

$$\frac{-B - \sqrt{B^2 - AC_2^*}}{\emptyset RA} < \alpha < \frac{-B + \sqrt{B^2 - AC_2^*}}{\emptyset RA}, \tag{4.8}$$

where $C_2^* = (\lambda - \lambda')R^2(\emptyset^2\beta^2 - 4\emptyset^{*^2})S_x^2 + \theta(2\beta_{yx(2)} - 2\beta_{yx} - \emptyset^*R)\emptyset^*RS_{x(2)}^2$.

5. Empirical Study

To illustrate the properties of the proposed estimators of the population mean $\bar{Y}$, we consider a real data set considered before by Srivastava (1993). The description of the sample is given below:

The sample of 100 consecutive trips (after omitting 20 outlier values) measured by two fuel meters for a small family car in normal usage given by Lewis *et al.* (1991) has been taken into consideration. The measurement of displacement meter (in cm^3) is considered as auxiliary variable x. We treat the last 25 percent of the values as non-response values. The values of the parameters are as follows:

$\bar{Y} = 3500.12; \quad \bar{X} = 260.84; \quad S_y = 2079.30; \quad S_x = 156.40; \quad \bar{Y}_2 = 3401.08; \quad \bar{X}_2 = 259.96;$

$S_{y(2)} = 1726.02; \quad S_{x(2)} = 134.36; \quad \rho_{yx} = 0.985; \quad \rho_{yx(2)} = 0.995; \quad N = 100; \quad n' = 50;$

$n = 30; \quad W_2 = 0.25.$

Here we have computed the percent relative efficiencies(PRE's) of different suggested estimators with respect to the usual unbiased estimator $\bar{y}^*$, for different values of k.

From the Table 2, it is envisaged that the proposed class of estimators under optimum condition is more desirable over all the considered estimators. It is observed that the percent relative efficiency of the estimators $T_{CRR}^{(1)}$ and $T_{CRR}^{(4)}$ increases as $(1/k)$ increases, but for the estimators $T_{CRR}^{(0)}$, $T_{CRR}^{(2)}$, $T_{CRR}^{(3)}$, $T_{CRR}^{(5)}$, $T_{CRR}^{(6)}$ and $T_{CRR}^{(opt)}$ decreases with the increase in the value of $(1/k)$. It is further observed that the estimator $T_{CRR}^{(0)}$ (regression estimator) seems to be a more appropriate estimator in comparison to others as it has appreciable efficiency (close to the efficiency of the optimum estimator $T_{CRR}^{(0pt)}$.

Finally, we conclude that the proposed estimator T_{CRR} under optimum conditions is recommended for future research.

Acknowledgement

References

Cochran, W. G. (1940). *Sampling Techniques*, 3rd edition, John Wiley & Sons, New York.

Cochran, W. G. (1977). *Sampling Techniques*, 3rd edition, John Wiley & Sons, New York.

Gupta, P. C. (1978). On some quadratic and higher degree ratio and product estimators, *Journal of the Indian Society of Agricultural Statistics*, **30**, 71–80.

Gupta, P. C. (1979). *Some Estimation Problems in Sampling Using Auxiliary Information*, Ph. D. thesis IARS, New Delhi.

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

Khare, B. B. and Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response, *National Academy Science Letters (India)*, **16**, 111–114.

Khare, B. B. and Srivastava, S. (1995). Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response, In *Proceeding Indian National Science Academy Indian*, **65**, 195–203.

Khare, B. B. and Srivastava, S. (1997). Transformed ratio type estimators for the population mean in the presence of nonresponse, *Communications in Statistics - Theory and Methods*, **26**, 1779–1791.

Kumar, S. and Singh, H. P. (2010). Estimation of mean using multi auxiliary information in presence of non-response, *Communications of the Korean Statistical Society*, **17**, 391–411.

Lewis, P. A., Jones, P. W., Polak, J. W. and Tillotson, H. T. (1991). The problem of conversion in method comparison studies, *Journal of Applied Statistics*, **40**, 105–112.

Murthy, M. N. (1964). Product method of estimation, *Sankhya: The Indian Journal of Statistics, Series A*, **26**, 294–307.

Okafor, F. C. and Lee, H. (2000). Double sampling for ratio and regression estimation with subsampling the non-respondents, *Survey Methodology*, **26**, 183–188.

Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the non-respondents, *Survey Methodology*, **12**, 217–230.

Rao, P. S. R. S. (1987). Ratio and regression estimates with sub sampling the non respondents, *Paper presented at a special contributed session of the International Statistical Association Meeting*, **2-16**, Tokyo, Japan.

Reddy, V. N. (1973). On ratio and product methods of estimation, *Sankhya: The Indian Journal of Statistics, Series B*, **35**, 307–316.

Reddy, V. N. (1974). On a transformed ratio method of estimation, *Sankhya: The Indian Journal of Statistics, Series C*, **36**, 59–70.

Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio type estimators, *Journal of the American Statistical Association*, **52**, 511–522.

Sahai, A. (1979). An efficient variant of the product and ratio estimator, *Statistica Neerlandica*, **33**, 27–35.

Singh, H. P. and Kumar, S. (2008a). A regression approach to the estimation of finite population mean in presence of non-response, *Australian and New Zealand Journal of Statistics*, **50**, 395–408.

Singh, H. P. and Kumar, S. (2008b). A general family of estimators of finite population ratio, product and mean using two phase sampling scheme in the presence of non-response, *Journal of Statistical Theory and Practice*, **2**, 677–692.

Singh, H. P. and Kumar, S. (2009a). A general class of estimators of the population mean in survey sampling using auxiliary information with sub sampling the non-respondents, *The Korean Journal of Applied Statistics*, **22**, 387–402.

Singh, H. P. and Kumar, S. (2009b). A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information, *SORT*, **33**, 71–84.

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

Singh, H. P. and Kumar, S. (2010b). Improved estimation of population mean under two phase sampling with sub sampling the non-respondents, *Journal of Statistical Planning and Inference*, **140**, 2536–2550.

Singh, H. P. and Kumar, S. (2011). Combination of ratio and regression estimators in presence of non-response, *Brazilian journal of Probability and Statistics*, **25**, 205–217.

Singh, H. P., Kumar, S. and Kozak, M. (2010). Improved estimation of finite-population mean when sub-sampling is employed to deal with non-response, *Communication in Statistics - Theory and Methods*, **39**, 791–802.

Singh, M. P. (1965). On the estimation of ratio and product of the population parameters, *Sankhya: The Indian Journal of Statistics, Series B*, **27**, 321–328.

Srivastava, S. (1967). An estimator using auxiliary information in sample surveys, *Calcutta Statistical Association Bulletin*, **16**, 121–132.

Srivastava, S. (1993). *Some Problems on the Estimation of Population Mean Using Auxiliary Character in Presence of Non-Response in Sample Surveys*, Thesis submitted to Banaras Hindu University, Varanasi, India.

Srivenkataramana, T. (1980). A dual to ratio estimator in sample surveys, *Biometrika*, **67**, 199–204.

Tabasum, R. and Khan, I. A. (2004). Double sampling for ratio estimation with non-response, *Journal of the Indian Society of Agricultural Statistics*, **58**, 300–306.

Tabasum, R. and Khan, I. A. (2006). Double sampling ratio estimator for the population mean in presence of non-response, *Assam Statistical Review*, **20**, 73–83.

Tripathi, T. P. and Khare, B. B. (1997). Estimation of mean vector in presence of non-response, *Communications in Statistics - Theory and Methods*, **26**, 2255–2269.

Walsh, J. E. (1970). Generalization of ratio estimate for population total, *Sankhya: The Indian Journal of Statistics, Series A*, **32**, 99–106.

Watson, D. J. (1937). The estimation of leaf areas, *Journal of Agricultural Science*, **27**, 474.