



Methods for Estimating the Mean with Imputation in Survey Sampling

Housila P. Singh and Anita Yadav

School of Studies in Statistics Vikram University Ujjain- 456010, M.P., India.

ARTICLE INFO

Article history:

Received: 26 March 2021;

Received in revised form:

10 May 2021;

Accepted: 20 May 2021;

Keywords

Imputation Method,
Study Variate,
Auxiliary Variate,
Generalized Family of
Estimators,
Bias and Mean Squared Errors.

ABSTRACT

In this paper, various generalized imputation procedures have been suggested and their corresponding families of estimators of population mean have been considered. Asymptotic expressions for biases and mean squared errors have been obtained. Optimum conditions are investigated at which the proposed family of estimators has minimum mean squared errors. It has been shown that the proposed imputation procedures and the resulting families of estimators due to Singh and Horn (2000), Ahmed et al. (2006), Shukla and Thakur (2008) and Pandey et al (2015) are members of the suggested procedures and estimators. The results of this paper are general and of immense useful.

© 2021 Elixir All rights reserved.

1. Introduction

In survey sampling for estimating the parameters it is commonly assumed that all the observations on selected units in the sample are available. This may not hold true in many practical situations encountered in sample surveys and some observations may be missing for various reasons, for instance see Toutenburg and Srivastava (1998). Imputation procedure is used to substitute values for missing data. In literature, various imputation procedures are available, some of them are better over others. Statisticians have recognized that for some time that failure to account for the stochastic nature of incompleteness can spoil inference. A natural question arises what one needs to assume to establish ignoring the incomplete mechanism. Rubin(1976) addressed three concepts: missing at random (MAR), observed at random (OAR) and parameter distribution (PD). Heitjan and Basu (1996) have distinguished the meaning of missing at random (MAR) and missing completely at random (MCAR) in a very acceptable way. In what follows MCAR is used in the present investigation.

Let $U = (U_1, U_2, \dots, U_N)$ be a finite population of N identifiable units taking values (y_1, y_2, \dots, y_N) on study variable y . Let x be an auxiliary variable taking the corresponding values (x_1, x_2, \dots, x_N) for the units (U_1, U_2, \dots, U_N) . We wish to estimate the population mean $\bar{y} = \sum_{i=1}^N y_i / N$ of the study variable y . A simple random sample without replacement (SRSWOR), s of size n is drawn from the population U for estimating the population mean. Let r be the number of responding units out of sampled n units. Let the set of responding units be denoted by

D and that of non-responding units be denoting by D^c . For every unit $i \in D$, the value y_i is observed. However for the $i \in D^c$, the y_i values are missing and imputed values are to be derived. We assume that imputation is carried out with the aid of quantitative auxiliary variable x , such that x_i , the value of x for unit i , is known and positive for every $i \in s$. In other words, the data $s_s = (x_i : i \in s)$ are known and $s = D \cup D^c$.

In many situations the values of the auxiliary variable x may be available for each unit in the population. In such a case knowledge on population mean \bar{X} and variance σ_x^2 (or population mean square S_x^2) and possibly on some other parameters may be utilized simultaneously, for instance see Das and Tripathi (1978, 1981), Srivastava and Jhajj (1980, 1981), Jhajj et al (2005) and Singh and Agnihotri (2008). Thus it is worth to mention that the imputation may be carried out with the aid of quantitative auxiliary variable for which the data $s_x = (x_i : i \in s)$ and the population mean \bar{X} or variance σ_x^2 (or population mean square S_x^2) or coefficient of variation C_x or both \bar{X} and σ_x^2 (or S_x^2) are known.

It is to be mentioned that assuming MACR Singh and Horn (2000), Ahmed et al (2006) and Shukla and Thakur (2008), Pandey et al (2015) have given some methods of imputation and considered their corresponding estimators for population mean \bar{y} of the variable y under investigation. In this paper following the same procedure as adopted by Srivastava (1971), Srivastava and Jhajj (1981) and Singh et al (2001) we have given some general procedure of imputation and derived their corresponding

families of estimators of the population mean \bar{Y} . It is interesting to note that this study unifies the imputation procedures of Singh and Horn (2000), Singh and Deo (2003), Ahmed et al (2006), Shukla and Thakur (2008) and Pandey et al (2015).

2. Notations and Useful Results

For simplicity we assume that population size N is very large as compared to sample size r and n so that finite population correction (fpc) terms are ignored. In what follows we shall use the following notations:

$$\bar{y}_r = \sum_{i=1}^r \frac{y_i}{r}, \bar{x}_r = \sum_{i=1}^r \frac{x_i}{r}, \bar{x}_n = \sum_{i=1}^n \frac{x_i}{n}, s_{x(r)}^2 = \sum_{i=1}^r \frac{(x_i - \bar{x}_r)^2}{(r-1)}, s_{x(n)}^2 = \sum_{i=1}^n \frac{(x_i - \bar{x}_n)^2}{(n-1)}, \bar{Y} = \sum_{i=1}^N \frac{y_i}{N}, \bar{X} = \sum_{i=1}^N \frac{x_i}{N},$$

$$S_x^2 = \sum_{i=1}^N (x_i - \bar{X})^2 / (N-1), \mu_{pq} = \sum_{i=1}^N (y_i - \bar{Y})^p (x_i - \bar{X})^q / N, (p, q) \text{ being non-negative integers.}$$

$$C_y^2 = S_y^2 / \bar{Y}^2 = \mu_{20} / \bar{Y}^2, C_x^2 = S_x^2 / \bar{X}^2 = \mu_{02} / \bar{X}^2, \rho_{yx} = \mu_{11} / (S_y S_x), \lambda_0 = \mu_{12} / (\bar{Y} S_x^2), \gamma_1 = \mu_{03} / S_x^3, \beta_1 = \gamma_1^2, \beta_2(x) = \mu_{04} / S_x^4.$$

$$\text{Let } u = \bar{x}_n / \bar{X}, a = s_{x(n)}^2 / S_x^2, v = \bar{x}_r / \bar{x}_n, b = s_{x(r)}^2 / s_{x(n)}^2, w = \bar{x}_r / \bar{X}, d = s_{x(r)}^2 / S_x^2, \varepsilon_0 = \left(\frac{\bar{y}_r}{\bar{Y}} - 1 \right),$$

$$\varepsilon_1 = (u-1) = \left(\frac{\bar{x}_n}{\bar{X}} - 1 \right), \varepsilon_2 = (a-1) = \left(\frac{s_{x(n)}^2}{S_x^2} - 1 \right), \eta_1 = (w-1) = \left(\frac{\bar{x}_r}{\bar{X}} - 1 \right), \eta_2 = (d-1) = \left(\frac{s_{x(r)}^2}{S_x^2} - 1 \right), v = \frac{w}{u},$$

$$v = \frac{(1+\eta_1)}{(1+\varepsilon_1)} \Rightarrow (v-1) \cong (\eta_1 - \varepsilon_1 + \varepsilon_1^2 - \eta_1 \varepsilon_1), \varepsilon_0(v-1) \cong (\eta_1 \varepsilon_0 - \varepsilon_0 \varepsilon_1), b = \frac{d}{a} = \frac{(1+\eta_2)}{(1+\varepsilon_2)}, (b-1) \cong (\eta_2 - \varepsilon_2 + \varepsilon_2^2 - \eta_2 \varepsilon_2),$$

$$\varepsilon_0(b-1) \cong (\eta_2 \varepsilon_0 - \varepsilon_0 \varepsilon_2),$$

$$(v-1)(b-1) \cong (\eta_1 - \varepsilon_1)(\eta_2 - \varepsilon_2) = (\eta_1 \eta_2 - \eta_1 \varepsilon_2 - \eta_2 \varepsilon_1 + \varepsilon_1 \varepsilon_2)$$

Using the concept of two phase sampling by following Rao and Sitter (1995), for given r and n , we have

$$E(\varepsilon_0) = E(\varepsilon_1) = E(\varepsilon_2) = E(\eta_1) = E(\eta_2) = 0$$

and

$$E(\varepsilon_0^2) = (1/r)C_y^2, E(\varepsilon_1^2) = (1/n)C_x^2, E(\eta_1^2) = (1/r)C_x^2, E(\varepsilon_0 \varepsilon_1) = (1/n)K_{yx}, E(\varepsilon_0 \eta_1) = (1/r)K_{yx}C_x^2, E(\varepsilon_1 \eta_1) = (1/n)C_x^2,$$

where $K_{yx} = \rho_{yx} C_y / C_x$.

and to the first degree of approximation,

$$E(\varepsilon_2^2) = (1/n)(\beta_2(x)-1), E(\eta_2^2) = (1/r)(\beta_2(x)-1), E(\varepsilon_2 \eta_2) = (1/n)(\beta_2(x)-1), E(\varepsilon_0 \varepsilon_2) = (1/n)\lambda_0, E(\varepsilon_0 \eta_2) = (1/r)\lambda_0$$

$$, E(\varepsilon_1 \varepsilon_2) = (1/n)\gamma_1 C_x, E(\varepsilon_1 \eta_2) = (1/r)\gamma_1 C_x, E(\varepsilon_2 \eta_1) = (1/n)\gamma_1 C_x, E(\eta_1 \eta_2) = (1/r)\gamma_1 C_x.$$

Thus we have following Lemmas:

Lemma 2.1: To the first degree of approximation,

$$E(v-1) = 0 \text{ and } E(v-1)^2 = \left(\frac{1}{\gamma} - \frac{1}{n} \right) C_x^2$$

Proof: We have

$$\begin{aligned} v-1 &= \frac{\bar{x}_r}{\bar{x}_n} - 1 \\ &= \frac{\bar{X}(1+\eta_1)}{\bar{X}(1+\varepsilon_1)} - 1 \\ &= (1+\eta_1)(1+\varepsilon_1)^{-1} - 1 \end{aligned}$$

We assume that $|\varepsilon_1| < 1$ so that we may expand $(1+\varepsilon_1)^{-1}$ as a series in powers of ε_1 . Expanding, multiplying out and neglecting terms of η_1 and ε_1 having power greater than two we have

$$(v-1) = (\eta_1 - \varepsilon_1 + \varepsilon_1^2 - \eta_1 \varepsilon_1) \tag{2.1}$$

Taking expectation of both sides of (2.1) we have

$$E(v-1) = 0$$

Squaring both sides of (2.1) and neglecting terms of η_1 and ε_1 having power greater than two we have

$$(v-1)^2 = (\eta_1^2 + \varepsilon_1^2 - 2\eta_1 \varepsilon_1) \tag{2.2}$$

Taking expectation of both sides of (2.3) we get

$$E(v-1)^2 = \left(\frac{1}{\gamma} - \frac{1}{n} \right) C_x^2.$$

Hence the lemma.

Lemma 2.2: To the first degree of approximation,

$$E(b-1) = 0 \text{ and } E(b-1)^2 = \left(\frac{1}{\gamma} - \frac{1}{n}\right) (\beta_2(x) - 1)$$

Proof: We have

$$\begin{aligned} b-1 &= \frac{s_{x(r)}^2}{s_{x(n)}^2} - 1 \\ &= \frac{S_x^2(1+\eta_2)}{S_x^2(1+\varepsilon_2)} - 1 \\ &= (1+\eta_2)(1+\varepsilon_2)^{-1} - 1 \end{aligned}$$

We assume that $|\varepsilon_2| < 1$ so that we may expand $(1+\varepsilon_2)^{-1}$ as a series in powers of ε_2 . Expanding, multiplying out and neglecting terms of η_2 and ε_2 having power greater than two we have

$$(b-1) = (\eta_2 - \varepsilon_2 + \varepsilon_2^2 - \eta_2\varepsilon_2) \quad (2.3)$$

Taking expectation of both sides of (2.3) we have

$$E(b-1) = 0$$

Squaring both sides of (2.3) and neglecting terms of η_2 and ε_2 having power greater than two we have

$$(b-1)^2 = (\eta_2^2 + \varepsilon_2^2 - 2\eta_2\varepsilon_2) \quad (2.4)$$

Taking expectation of both sides of (2.4) we get

$$E(b-1)^2 = \left(\frac{1}{\gamma} - \frac{1}{n}\right) (\beta_2(x) - 1).$$

Thus the lemma is proved.

Lemma 2.3: To the first degree of approximation,

$$E\{(b-1)(v-1)\} = \left(\frac{1}{\gamma} - \frac{1}{n}\right) \gamma_1 C_x$$

Proof: We have

$$\begin{aligned} (b-1)(v-1) &= \left(\frac{s_{x(r)}^2}{s_{x(n)}^2} - 1\right) \left(\frac{\bar{x}_r}{\bar{x}_n} - 1\right) \\ &= \left\{ \frac{S_x^2(1+\eta_2)}{S_x^2(1+\varepsilon_2)} - 1 \right\} \left\{ \frac{\bar{X}(1+\eta_1)}{\bar{X}(1+\varepsilon_1)} - 1 \right\} \\ &= \left\{ (1+\eta_2)(1+\varepsilon_2)^{-1} - 1 \right\} \left\{ (1+\eta_1)(1+\varepsilon_1)^{-1} - 1 \right\} \\ &= \left\{ (1+\eta_2)(1-\varepsilon_2 + \varepsilon_2^2 - \dots) - 1 \right\} \left\{ (1+\eta_1)(1-\varepsilon_1 + \varepsilon_1^2 - \dots) - 1 \right\} \\ &= \left\{ (1+\eta_2 - \varepsilon_2 - \varepsilon_2\eta_2 + \varepsilon_2^2 + \varepsilon_2^2\eta_2 + \dots)(1+\eta_1 - \varepsilon_1 - \varepsilon_1\eta_1 + \varepsilon_1^2 + \varepsilon_1^2\eta_1 + \dots) - 1 \right\} \\ &= \left[\eta_1\eta_2 - \eta_1\varepsilon_2 - \eta_2\varepsilon_1 + \varepsilon_1\varepsilon_2 + 0(\varepsilon^2) \right] \end{aligned} \quad (2.5)$$

Taking expectation of both sides of (2.5) we have

$$E\{(b-1)(v-1)\} = \left(\frac{1}{\gamma} - \frac{1}{n}\right) \gamma_1 C_x.$$

which proves the lemma.

Lemma 2.4: To the first degree of approximation,

$$E\{\varepsilon_0(v-1)\} = \left(\frac{1}{r} - \frac{1}{n}\right) K_{yx} C_x^2 \quad \text{and} \quad E\{\varepsilon_0(b-1)\} = \left(\frac{1}{r} - \frac{1}{n}\right) \lambda_0$$

Proof: We have

$$\begin{aligned} \varepsilon_0(v-1) &= \left(\frac{\bar{x}_r}{\bar{x}_n} - 1\right) \\ &= \varepsilon_0 \left\{ \frac{\bar{X}(1+\eta_1)}{\bar{X}(1+\varepsilon_1)} - 1 \right\} \\ &= \varepsilon_0 \left\{ (1+\eta_1)(1+\varepsilon_1)^{-1} - 1 \right\} \\ &= \varepsilon_0 \left\{ (1+\eta_1)(1-\varepsilon_1 + \varepsilon_1^2 - \dots) - 1 \right\} \\ &= \varepsilon_0 \left\{ 1 + \eta_1 - \varepsilon_1 - \varepsilon_1\eta_1 + \varepsilon_1^2 + \varepsilon_1^2\eta_1 + \dots - 1 \right\} \end{aligned}$$

$$= \left[\eta_1 - \varepsilon_2 - \eta_1 \varepsilon_1 + \varepsilon_1^2 + \eta_1 \varepsilon_1^2 + \dots \right]$$

which can be written as

$$E(v-1) = (\varepsilon_0 \eta_1 - \varepsilon_0 \varepsilon_1) + O(\varepsilon^2) \quad (2.6)$$

Taking expectation of both sides of (2.6) we get

$$E\{\varepsilon_0(v-1)\} = \left(\frac{1}{r} - \frac{1}{n} \right) K_{yx} C_x^2. \quad (2.7)$$

where $K_{yx} = \rho_{yx} C_y / C_x$

Similarly we can express

$$\begin{aligned} E\{\varepsilon_0(b-1)\} &= \varepsilon_0 \left\{ \frac{(1+\eta_2) S_x^2}{(1+\varepsilon_2) S_x^2} - 1 \right\} \\ &= \varepsilon_0 \left\{ (1+\eta_2) (1+\varepsilon_2)^{-1} - 1 \right\} \\ &= (\varepsilon_0 \eta_2 - \varepsilon_0 \varepsilon_2) + O(\varepsilon^2) \end{aligned} \quad (2.8)$$

Taking expectation of both sides of (2.8) we get

$$E\{\varepsilon_0(b-1)\} = \left(\frac{1}{r} - \frac{1}{n} \right) \lambda_0. \quad (2.9)$$

which proves the lemma.

3. Some known Procedures of Imputation and Estimators

In this section we give some classical methods of imputation which are available in the literature.

3.1 Ratio Method of Imputation

Following the notations of Lee et al (1994), in the case of single value imputation, if the i th unit requires imputation, the value \hat{b}_{xi} is imputed, where $\hat{b}_i = \sum_{i \in D} y_i / \sum_{i \in D} x_i$. The data after imputation becomes

$$y_{oi} = \begin{cases} y_i & \text{if } i \in D \\ \hat{b}_i x_i & \text{if } i \in D^C \end{cases} \quad (3.1)$$

where D and D^C denote the responding and non-responding units in the sample. This method of imputation is called the ratio method of imputation. Under this method of imputation, the point estimator of the population mean is given by

$$\begin{aligned} \bar{y}_s &= \frac{1}{n} \sum_{i=1}^n y_i \quad (3.2) \\ &= \frac{1}{n} \left[\sum_{i \in D} y_{oi} + \sum_{i \in D^C} y_{oi} \right] \\ &= \frac{1}{n} \left[\sum_{i=1}^r y_i + \sum_{i=r+1}^n \hat{b}_i x_i \right] \\ &= \frac{1}{n} \left[r \bar{y}_r + \hat{b} \sum_{i=r+1}^n x_i \right] \\ &= \frac{1}{n} \left[r \bar{y}_r + \left(\frac{\sum_{i=1}^n y_i}{\sum_{i=1}^r x_i} \right) (n \bar{x}_n - r \bar{x}_r) \right] \\ &= \frac{1}{n} \left[r \bar{y}_r + \frac{\bar{y}_r}{\bar{x}_r} (n \bar{x}_n - r \bar{x}_r) \right] \\ &= \frac{1}{n} \left[\frac{r}{n} \bar{y}_r + \frac{n}{n} \frac{\bar{y}_r}{\bar{x}_r} \bar{x}_n - \frac{r}{n} \bar{y}_r \right] \\ &= \bar{y}_r \frac{\bar{x}_r}{\bar{x}_n} = t_R \quad (\text{say}) \end{aligned} \quad (3.3)$$

$$\bar{y}_r = \frac{1}{r} \sum_{i=1}^r y_i, \quad \bar{x}_r = \frac{1}{r} \sum_{i=1}^r x_i \quad \text{and} \quad \bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i.$$

3.2 Mean Method of Imputation

Under mean method of imputation, the data after imputation take the form :

$$y_{.i} = \begin{cases} y_i & \text{if } i \in D \\ \bar{y}_r & \text{if } i \in D^C \end{cases} \quad (3.4)$$

and the point estimator (3.2) becomes

$$\bar{y}_{.m} = \frac{1}{r} \sum_{i \in D} y_i = \bar{y}_r \quad (3.5)$$

3.3 Compromised method of Imputation

Singh and Horn (2000) suggested the compromised imputation procedure, where the data take the form,

$$y_{oi} = \begin{cases} \frac{\alpha n}{r} y_i + (1-\alpha) \hat{b} x_i & \text{if } i \in D \\ (1-\alpha) \hat{b} x_i & \text{if } i \in D^C \end{cases} \quad (3.6)$$

where α is a chosen constant, such that the variance of the resultant estimator is minimum. It is to be mentioned that this procedure is also using information from imputed values for the responding units in addition to non-responding units. The point estimator (3.2) of the population mean \bar{Y} under compromised method of imputation becomes:

$$t_1 = \alpha \bar{y}_r + (1-\alpha) \bar{y}_r \frac{\bar{x}_n}{\bar{x}_r} \quad (3.7)$$

3.4 Power Transformation Method of Imputation

Singh and Deo (2003) proposed a power transformation imputation procedure, where the data take form:

$$y_{.i} = \begin{cases} y_i & \text{if } i \in D \\ \bar{y}_r \left[n \left(\frac{\bar{x}_n}{\bar{x}_r} \right)^{\alpha_1} - r \right] & \text{if } i \in D^C \end{cases} \quad (3.8)$$

where α_1 is a suitably chosen constant, such that the variance of the resultant estimator is minimum. The limitation of adjusting responding units in the methods proposed by Singh and Horn (2000) and Singh et al (2001) has been removed under this imputation procedure. The point estimator (2.2) of the population mean \bar{Y} under proposed method of imputation becomes

$$t_2 = \bar{y}_r \left(\frac{\bar{x}_n}{\bar{x}_r} \right)^{\alpha_1} \quad (3.9)$$

For $\alpha_1 = 0, 1$ and -1 , the point estimator t_2 respectively reduces to \bar{y}_r , $t_R = \bar{y}_r \left(\frac{\bar{x}_n}{\bar{x}_r} \right)$ and $t_p = \bar{y}_r \left(\frac{\bar{x}_r}{\bar{x}_n} \right)$. Thus the transformation method of imputation is some sort of compromise between mean, ratio and product methods of imputation.

3.5 Ahmed et al (2006) Methods of Imputation:

When the population mean $\bar{X} = \sum_{i=1}^N x_i / N$ of an auxiliary variable x is known, Ahmed et al (2006) have given the following imputation procedures:

$$(A.1) \ y_{1i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \left[n \bar{y}_r \left(\frac{\bar{X}}{\bar{x}_n} \right)^{\delta_1} - r \bar{y}_r \right] & \text{if } i \in D^C \end{cases} \quad (3.10)$$

$$(A.2) \ y_{2i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \left[n \bar{y}_r \left(\frac{\bar{X}}{\bar{x}_r} \right)^{\delta_2} - r \bar{y}_r \right] & \text{if } i \in D^C \end{cases} \quad (3.11)$$

$$(A.3) \ y_{3i} = \begin{cases} y_i & \text{if } i \in D \\ \bar{y}_r + \frac{n \delta_3}{(n-r)} (\bar{X} - \bar{x}_n) & \text{if } i \in D^C \end{cases} \quad (3.12)$$

$$(A.4) \quad y_{4i} = \begin{cases} y_i & \text{if } i \in D \\ \bar{y}_r + \frac{n\delta_4}{(n-r)}(\bar{X} - \bar{x}_r) & \text{if } i \in D^C \end{cases} \quad (3.13)$$

$$(A.5) \quad y_{5i} = \begin{cases} y_i & \text{if } i \in D \\ \bar{y}_r + \frac{n\delta_5}{(n-r)}(\bar{X} + \bar{x}_n) + \delta_6(x_i - \bar{x}_r) & \text{if } i \in D^C \end{cases} \quad (3.14)$$

$$(A.6) \quad y_{6i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \left[\frac{n\bar{y}_r \bar{X}}{\delta_7 \bar{x}_n + (1-\delta_7) \bar{X}} - r\bar{y}_r \right] & \text{if } i \in D^C \end{cases} \quad (3.15)$$

$$(A.7) \quad y_{7i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \left[\frac{n\bar{y}_r \bar{X}}{\delta_8 \bar{x}_r + (1-\delta_8) \bar{X}} - r\bar{y}_r \right] & \text{if } i \in D^C \end{cases} \quad (3.16)$$

where y_{ji} denotes the i th available observation for the j th imputation method; and δ_i 's ($i=1$ to 7) are suitably chosen constant, such that the variance(s) of the resulting estimator(s) is minimum.

The point estimator (3.2) of the population mean \bar{Y} under proposed methods of imputation ((A.1) - (A.7)) respectively turn out to be:

$$t_3 = \bar{y}_r \left(\frac{\bar{X}}{\bar{x}_n} \right)^{\delta_1} \quad (3.17)$$

$$t_4 = \bar{y}_r \left(\frac{\bar{X}}{\bar{x}_r} \right)^{\delta_2} \quad (3.18)$$

$$t_5 = \bar{y}_r + \delta_3(\bar{X} - \bar{x}_n) \quad (3.19)$$

$$t_6 = \bar{y}_r + \delta_4(\bar{X} - \bar{x}_r) \quad (3.20)$$

$$t_7 = \bar{y}_r + \delta_5(\bar{X} - \bar{x}_n) + \delta_6(\bar{x}_n - \bar{x}_r) \quad (3.21)$$

$$t_8 = \frac{\bar{y}_r \bar{X}}{[\delta_7 \bar{x}_n + (1-\delta_7) \bar{X}]} \quad (3.22)$$

$$t_9 = \frac{\bar{y}_r \bar{X}}{[\delta_8 \bar{x}_r + (1-\delta_8) \bar{X}]} \quad (3.23)$$

It is to be mentioned that for $\delta_2 = 1$ and -1 , the class of estimators t_2 respectively reduces to:

$$t_{R1} = \bar{y}_r \left(\frac{\bar{X}}{\bar{x}_r} \right) \quad (\text{Ratio estimator}) \quad (3.24)$$

and

$$t_{p1} = \bar{y}_r \left(\frac{\bar{x}_r}{\bar{X}} \right) \quad (\text{Product estimator}) \quad (3.25)$$

When the population mean \bar{X} of the auxiliary variable x is not known, Ahmed et al (2006) suggested the following imputation procedures:

$$(A.8) \quad y_{8i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \left[n\bar{y}_r \left(\frac{\bar{x}_n}{\bar{x}_r} \right)^{\delta_8} - r\bar{y}_r \right] & \text{if } i \in D^C \end{cases} \quad (3.26)$$

$$(A.9) \quad y_{9i} = \begin{cases} y_i & \text{if } i \in D \\ \bar{y}_r + \delta_9(x_i - \bar{x}_r) & \text{if } i \in D^C \end{cases} \quad (3.27)$$

$$(A.10) \quad y_{10i} = \begin{cases} y_i & \text{if } i \in D \\ \left[\frac{\bar{y}_r \left(x_i + \frac{r}{n-r} \bar{x}_r \right)}{\delta_{10} \bar{x}_r + (1-\delta_{10}) \bar{x}_n} - \frac{r}{n-r} \bar{y}_r \right] & \text{if } i \in D^C \end{cases} \quad (3.28)$$

The point estimator (3.2) of the population mean \bar{Y} under suggested procedures of imputation ((A.8) - (A.10)) respectively becomes:

$$t_{10} = \bar{y}_r \left(\frac{\bar{x}_m}{\bar{x}_r} \right)^{\delta_8} \quad (3.29)$$

$$t_{11} = \bar{y}_r + \delta_9 (\bar{x}_n - \bar{x}_r) \quad (3.30)$$

$$t_{12} = \frac{\bar{y}_r \bar{x}_n}{\{\delta_{10} \bar{x}_r + (1 - \delta_{10}) \bar{x}_n\}} \quad (3.31)$$

Here we note that the estimator t_8 at (3.29) is same as obtained by Singh and Deo (2003).

3.6 Shukla and Thakur (2008) Methods of Imputation

When the population mean \bar{X} of the auxiliary variable x is known, Shukla and Thakur (2008) suggested the following method of imputation:

$$(S.1) \quad y_{11i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{\bar{y}_r}{(n-r)} [n\phi_1(k) - r] & \text{if } i \in D^C \end{cases} \quad (3.32)$$

$$(S.2) \quad y_{12i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{\bar{y}_r}{(n-r)} [n\phi_2(k) - r] & \text{if } i \in D^C \end{cases} \quad (3.33)$$

and when the population mean \bar{X} of the auxiliary variable x is not known, Shukla and Thakur (2008) proposed the following method of imputation:

$$(S.3) \quad y_{13i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{\bar{y}_r}{(n-r)} [n\phi_3(k) - r] & \text{if } i \in D^C \end{cases} \quad (3.34)$$

where

$$\phi_1(k) = \frac{\{(A+C)\bar{X} + fB\bar{x}_n\}}{\{(A+fB)\bar{X} + C\bar{x}_n\}},$$

$$\phi_2(k) = \frac{\{(A+C)\bar{X} + fB\bar{x}_r\}}{\{(A+fB)\bar{X} + C\bar{x}_r\}},$$

$$\phi_3(k) = \frac{\{(A+C)\bar{x}_n + fB\bar{x}_r\}}{\{(A+fB)\bar{x}_n + C\bar{x}_r\}}.$$

where $A = (k-1)(k-2)$, $B = (k-1)(k-4)$, $C = (k-2)(k-3)(k-4)$, $f = n/N$ and $0 < k < \infty$ is a constant. Under the imputation procedures ((S.1) – (S.3)), the point estimator (3.2) respectively becomes:

$$t_{13} = \bar{y}_r \phi_1(k) \quad (3.35)$$

$$t_{14} = \bar{y}_r \phi_2(k) \quad (3.36)$$

$$t_{15} = \bar{y}_r \phi_3(k) \quad (3.37)$$

3.7 Pandey, Thakur and Yadav (2015) Methods of Imputation

Pandey et al (2015) three imputation strategies using known population mean \bar{X} of the auxiliary variable x based on exponential ratio-type estimator envisaged by Bahl and Tuteja (1991). Let y_{ij} denote the i th available observations for the j th imputation.

$$\bullet \quad y_{1i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{n\bar{y}_r}{(n-r)} \left[\exp \left(\frac{\bar{X} - \bar{x}_n}{\bar{X} + \bar{x}_n} \right) - \frac{r}{n} \right] & \text{if } i \in D^C \end{cases} \quad (3.38)$$

Resulting point estimator of the population mean \bar{Y} is

$$t_{16} = \bar{y}_r \exp \left(\frac{\bar{X} - \bar{x}_n}{\bar{X} + \bar{x}_n} \right) \quad (3.39)$$

$$\bullet y_{2i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{n\bar{y}_r}{(n-r)} \left[\exp\left(\frac{\bar{x}_r - \bar{x}_n}{\bar{x}_r + \bar{x}_n}\right) - \frac{r}{n} \right] & \text{if } i \in D^c \end{cases}$$

Resulting point estimator of the population mean \bar{Y} is

$$t_{17} = \bar{y}_r \exp\left(\frac{\bar{x}_r - \bar{x}_n}{\bar{x}_r + \bar{x}_n}\right) \quad (3.40)$$

$$\bullet y_{3i} = \begin{cases} y_i & \text{if } i \in D \\ \frac{n}{(n-r)} \left[\exp\left(\frac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) - \frac{n}{(n-r)} \right] & \text{if } i \in D^c \end{cases} \quad (3.41)$$

Resulting point estimator of the population mean \bar{Y} is

$$t_{18} = \bar{y}_r \exp\left(\frac{\bar{X} - \bar{x}_r}{\bar{X} + \bar{x}_r}\right) \quad (3.42)$$

For further study in this context the reader are referred to Pandey and Yadav (2016, 2017, 2018) and Mohamed et al (2018).

4. Some Suggested General Methods of Imputation and their Estimators

In what follow, y_{ji} denotes the i th available observation for the j th imputation method. We suggest the following imputation procedures:

- When the population mean \bar{X} of the auxiliary variable x is known:

$$y_{fi} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \{nf(\bar{y}_r, u) - r\bar{y}_r\} & \text{if } i \in D^c \end{cases} \quad (4.1)$$

- When the population mean \bar{X} of the auxiliary variable x is not known:

$$y_{gi} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \{ng(\bar{y}_r, v) - r\bar{y}_r\} & \text{if } i \in D^c \end{cases} \quad (4.2)$$

- When the population mean \bar{X} of the auxiliary variable x is known:

$$y_{hi} = \begin{cases} y_i & \text{if } i \in D \\ \frac{1}{(n-r)} \{nh(\bar{y}_r, w) - r\bar{y}_r\} & \text{if } i \in D^c \end{cases} \quad (4.3)$$

where $l(\bar{y}_r, u)$, $g(\bar{y}_r, v)$ and $h(\bar{y}_r, w)$ are the functions of (\bar{y}_r, u) , (\bar{y}_r, v) and (\bar{y}_r, w) respectively. The functions $l(\bar{y}_r, u)$, $g(\bar{y}_r, v)$ and $h(\bar{y}_r, w)$ assume values in a bounded closed convex subset $Q \subset \mathcal{R}_2$, which contain the point $(\bar{Y}, 1)$ and are such that

$$l(\bar{Y}, 1) = \bar{Y} \Rightarrow l_1(\bar{Y}, 1) = \frac{\partial f(\cdot)}{\partial \bar{y}_r} \Big|_{(\bar{Y}, 1)} = 1, \quad g(\bar{Y}, 1) = \bar{Y} \Rightarrow g_1(\bar{Y}, 1) = \frac{\partial g(\cdot)}{\partial \bar{y}_r} \Big|_{(\bar{Y}, 1)} = 1 \quad \text{and}$$

$$h(\bar{Y}, 1) = \bar{Y} \Rightarrow h_1(\bar{Y}, 1) = \frac{\partial h(\cdot)}{\partial \bar{y}_r} \Big|_{(\bar{Y}, 1)} = 1$$

The point estimators of the population mean \bar{Y} under (4.1), (4.2) and (4.3) methods of imputation are respectively given by

$$\bar{y}_l = l(\bar{y}_r, u), \quad (4.4)$$

$$\bar{y}_g = g(\bar{y}_r, v), \quad (4.5)$$

$$\text{and } \bar{y}_h = h(\bar{y}_r, w). \quad (4.6)$$

Remark 4.1

It is to be mentioned that the point estimators

$$\left. \begin{aligned}
 \bar{y}_l^{(1)} &= \bar{y}_r u^{-1}, \\
 \bar{y}_l^{(2)} &= \bar{y}_r u, \\
 \bar{y}_l^{(3)} &= \bar{y}_r u^{-\beta_1}, \\
 \bar{y}_l^{(4)} &= \bar{y}_r \{\beta_1 u + (1 - \beta_1)\}^{-1}, \\
 \bar{y}_l^{(5)} &= \bar{y}_r + b_1(1 - u) \bar{X}, \\
 \bar{y}_l^{(6)} &= \bar{y}_r \frac{\{(A + C) + lBu\}}{\{(A + fB) + Cu\}}, \\
 \bar{y}_l^{(7)} &= \bar{y}_r \exp\left\{\frac{\beta_1(1 - u)}{(1 + u)}\right\}, \\
 \bar{y}_l^{(8)} &= \beta_1 \bar{y}_r + (1 - \beta_1) \bar{y}_r u^{-a} \\
 \bar{y}_l^{(9)} &= t_{16} = \bar{y}_r \exp\left(\frac{1 - u}{1 + u}\right)
 \end{aligned} \right\} \quad (4.7)$$

etc. are members of the family of estimators \bar{y}_l given by (4.4), where $\alpha, \beta_1, b_1, A = (k - 1)(k - 2), B = (k - 1)(k - 4), C = (k - 2)(k - 3)(k - 4), 0 < k < \infty$ are constants and $f = n/N$. We note that the estimators $\bar{y}_l^{(3)}, \bar{y}_l^{(4)}$ and $\bar{y}_l^{(5)}$ are due to Ahmed et al (2006) estimators while the estimators $\bar{y}_l^{(6)}$ and $\bar{y}_l^{(10)}$ due to Shukla and Thakur (2008) and Pandey et al (2015) respectively.

Remark 4.2

The following point estimators

$$\left. \begin{aligned}
 \bar{y}_g^{(1)} &= \bar{y}_r v^{-1}, \text{ Singh and Horn(2000) estimator} \\
 \bar{y}_g^{(2)} &= \bar{y}_r v, \\
 \bar{y}_g^{(3)} &= \alpha_1 \bar{y}_r + (1 - \alpha_1) \bar{y}_r v^{-1}, \text{ Singh and Horn(2000) estimator} \\
 \bar{y}_g^{(4)} &= \bar{y}_r v^{-\alpha_1}, \bar{y}_g^{(5)} = \bar{y}_r \{\alpha_1 v + (1 - \alpha_1)\}^{-1}, \text{ Ahmed et al(2006) estimators} \\
 \bar{y}_g^{(6)} &= \bar{y}_r + b_2(1 - v), \\
 \bar{y}_g^{(7)} &= \bar{y}_r \frac{\{(A + C) + fBv\}}{\{(A + fB) + Cv\}} \text{ Shukla and Thakur (2008),} \\
 \bar{y}_g^{(8)} &= \bar{y}_r \exp\left\{\frac{\alpha_1(1 - v)}{(1 + v)}\right\}, \\
 \bar{y}_g^{(9)} &= \bar{y}_r \{\alpha_1 + (1 - \alpha_1) v^{-\beta_1}\} \\
 \bar{y}_g^{(10)} &= t_{17} = \bar{y}_r \exp\left(\frac{v - 1}{v + 1}\right) \text{ Pandey et al (2015)}
 \end{aligned} \right\} \quad (4.8)$$

etc. are members of the family of estimators \bar{y}_g given by (4.5), where α_1, β_1, b_2 are suitably chosen constants.

Remark 4.3

It is to be noted that the estimators

$$\left. \begin{aligned} \bar{y}_h^{(1)} &= \bar{y}_r w^{-1}, \bar{y}_h^{(2)} = \bar{y}_r w, \bar{y}_h^{(3)} = \bar{y}_r w^{a_3}, \\ \bar{y}_h^{(4)} &= \bar{y}_r + b_3(1-w)\bar{X}, \\ \bar{y}_h^{(5)} &= \bar{y}_r + \{\alpha_2 w + (1-w_2)\}^{-1} \\ \bar{y}_h^{(6)} &= \bar{y}_r \frac{\{(A+C) + fBw\}}{\{(A+fB) + Cw\}} \text{ Shukla and Thakur (2008),} \\ \bar{y}_h^{(7)} &= \bar{y}_r \exp\left\{\frac{\alpha_2(1-w)}{(1+w)}\right\}, \\ \bar{y}_h^{(8)} &= \bar{y}_r \{\alpha_2 + (1-\alpha_2)w^{-\beta_2}\} \\ \bar{y}_h^{(9)} &= t_{18} = \bar{y}_r \exp\left(\frac{1-w}{1+w}\right) \text{ Pandey et al (2015)} \end{aligned} \right\} \text{ Ahmed et al(2006) estimators} \quad (4.9)$$

etc. are members of the family of estimators \bar{y}_h given by (4.6), where α_1, β_1, b_3 are suitably chosen constants.

5. Biases and Mean Squared Errors of the Families of Estimators \bar{y}_l, \bar{y}_g and \bar{y}_h

To the first degree of approximation (ignoring finite population correction terms) the biases and mean squared errors of \bar{y}_l, \bar{y}_g and \bar{y}_h are respectively given by

$$B(\bar{y}_l) = (C_x^2/2n) \{l_{22}(\bar{Y}, 1) + 2\bar{Y}K_{yx}l_{12}(\bar{Y}, 1)\}, \quad (5.1)$$

$$B(\bar{y}_g) = \{(1/r) - (1/n)\} (C_x^2/2) \{g_{22}(\bar{Y}, 1) + 2\bar{Y}K_{yx}g_{12}(\bar{Y}, 1)\}, \quad (5.2)$$

$$B(\bar{y}_h) = (C_x^2/2r) \{h_{22}(\bar{Y}, 1) + 2\bar{Y}K_{yx}h_{12}(\bar{Y}, 1)\}, \quad (5.3)$$

$$MSE(\bar{y}_l) = \left[(S_y^2/r) + (C_x^2/n) l_2(\bar{Y}, 1) \{l_2(\bar{Y}, 1) + 2\bar{Y}K_{yx}\} \right], \quad (5.4)$$

$$MSE(\bar{y}_g) = \left[(S_y^2/r) + \{(1/r) - (1/n)\} C_x^2 g_2(\bar{Y}, 1) \{g_2(\bar{Y}, 1) + 2\bar{Y}K_{yx}\} \right], \quad (5.5)$$

$$MSE(\bar{y}_h) = (1/r) \left[S_y^2 + C_x^2 h_2(\bar{Y}, 1) \{h_2(\bar{Y}, 1) + 2\bar{Y}K_{yx}\} \right], \quad (5.6)$$

where $l_{ij}(\bar{Y}, 1)$, $g_{ij}(\bar{Y}, 1)$ and $h_{ij}(\bar{Y}, 1)$, ($i, j = 1, 2$) denote the second order derivatives of the functions $l(\cdot)$, $g(\cdot)$ and $h(\cdot)$ respectively about the point $(\bar{Y}, 1)$. It is to be mentioned that the biases and mean squared errors of the estimators $\bar{y}_l^{(i)}$, $i = 1$ to 9, $\bar{y}_g^{(j)}$, $i = 1$ to 10 and $\bar{y}_h^{(l)}$, $i = 1$ to 9 can be obtained from {(5.1),(5.4)}, {(5.2), (5.5)}, and {(5.3),(5.6)} respectively just by

putting the values of derivatives $\{l_2(\bar{Y}, 1), l_{12}(\bar{Y}, 1), l_{22}(\bar{Y}, 1)\}$, $\{g_2(\bar{Y}, 1), g_{12}(\bar{Y}, 1), g_{22}(\bar{Y}, 1)\}$ and $\{h_2(\bar{Y}, 1), h_{12}(\bar{Y}, 1), h_{22}(\bar{Y}, 1)\}$. The mean squared error of \bar{y}_l, \bar{y}_g and \bar{y}_h respectively in (5.4), (5.5) and (5.6) are minimized for

$$l_2(\bar{Y}, 1) = g_2(\bar{Y}, 1) = h_2(\bar{Y}, 1) = -\beta\bar{X} \quad (5.7)$$

$$\min.MSE(\bar{y}_l) = S_y^2 \left[(1/r) (1 - \rho_{yx}^2) + \{(1/r) - (1/n)\} \rho_{yx}^2 \right], \quad (5.8)$$

$$\min.MSE(\bar{y}_g) = S_y^2 \left[(1/r) (1 - \rho_{yx}^2) + (1/n) \rho_{yx}^2 \right], \quad (5.9)$$

$$\min.MSE(\bar{y}_h) = (S_y^2/r) (1 - \rho_{yx}^2). \quad (5.10)$$

Thus we state the following Theorem:

Theorem 5.1

To the first degree of approximation

$$(i) \text{ } MSE(\bar{y}_l) \geq S_y^2 \left[(1/r) (1 - \rho_{yx}^2) + \{(1/r) - (1/n)\} \rho_{yx}^2 \right]$$

with equality holding if $l_2(\bar{Y}, 1) = -\beta\bar{X}$.

$$(ii) \text{ } MSE(\bar{y}_g) \geq S_y^2 \left[(1/r) (1 - \rho_{yx}^2) + (1/n) \rho_{yx}^2 \right]$$

with equality holding if $g_2(\bar{Y}, 1) = -\beta\bar{X}$.

$$(iii) \text{ } MSE(\bar{y}_h) \geq (S_y^2/r) (1 - \rho_{yx}^2)$$

with equality holding if $h_2(\bar{Y}, 1) = -\beta\bar{X}$.

We also state the following theorem:

Theorem 5.2

The suggested family of estimators \bar{y}_g is more efficient than the family of estimators \bar{y}_l at its optimum condition if $r < n/2$.

Proof:

From (4.8) and (4.9) we have

$$\min.MSE(\bar{y}_g) - \min.MSE(\bar{y}_h) = S_y^2 \{(2/n) - (1/r)\} \rho_{yx}^2$$

which is less than 'zero' if

$$\{(2/n) - (1/r)\} < 0$$

i.e. if $(2/n) - (1/r)$

i.e. if $r < (n/2)$.

Thus the theorem is proved.

Theorem 5.3

The suggested family of estimators \bar{y}_h is the best among the families of estimators \bar{y}_l , \bar{y}_g and \bar{y}_h .

Proof:

From (5.8), (5.9) and (5.10) we have

$$\min.MSE(\bar{y}_g) - \min.MSE(\bar{y}_h) = S_y^2 \{(2/n) - (1/r)\} \rho_{yx}^2$$

which is less than 'zero' if

$$\{(2/n) - (1/r)\} < 0$$

i.e. if $(2/n) - (1/r)$

i.e. if $r < (n/2)$.

Thus the Theorem is proved.

$$\min.MSE(\bar{y}_l) - \min.MSE(\bar{y}_h) = \{(1/r) - (1/n)\} \rho_{yx}^2 S_y^2 \geq 0, \quad (5.11)$$

$$\min.MSE(\bar{y}_g) - \min.MSE(\bar{y}_h) = (1/n) \rho_{yx}^2 S_y^2 \geq 0. \quad (5.12)$$

From the above it is clear that the proposed family of estimators \bar{y}_h has least MSE than the families of estimators \bar{y}_l and \bar{y}_g at their optimum conditions. Hence the theorem.

6. Efficiency Comparisons When the Value of the Derivative Does Not Coincide With Its Optimum Value

6.I: To compare the family of estimators \bar{y}_l with the conventional unbiased estimator \bar{y}_r , the ratio estimator $\bar{y}_l^{(1)} = \bar{y}_r (\bar{X}/\bar{x}_n)$ and the product estimator $\bar{y}_l^{(2)} = \bar{y}_r (\bar{x}_n/\bar{X})$ we write the variance of \bar{y}_r and mean squared errors of $\bar{y}_l^{(1)}$ and $\bar{y}_l^{(2)}$ to the first degree of approximation (ignoring finite correction factor) are respectively given by

$$Var(\bar{y}_r) = S_y^2 / r, \quad (6.1)$$

$$MSE(\bar{y}_l^{(1)}) = \left\{ \left(S_y^2 / r \right) + \bar{Y}^2 \left(C_x^2 / n \right) \left(1 - 2K_{yx} \right) \right\} \quad (6.2)$$

$$MSE(\bar{y}_l^{(2)}) = \left\{ \left(S_y^2 / r \right) + \bar{Y}^2 \left(C_x^2 / n \right) \left(1 + 2K_{yx} \right) \right\} \quad (6.3)$$

From (5.4), (6.1), (6.2) and (6.3) it is observed that the suggested family of estimators \bar{y}_l is more efficient than

(i) the conventional unbiased estimator \bar{y}_r if

$$\min.(0, -\beta\bar{X}) < l_2(\bar{Y}, 1) < \max.(0, -\beta\bar{X}), \quad (6.4)$$

(ii) the usual ratio estimator $\bar{y}_l^{(1)}$ if

$$\min.\{-R\bar{X}, \bar{X}(R-2\beta)\} < l_2(\bar{Y}, 1) < \max.\{-R\bar{X}, \bar{X}(R-2\beta)\}, \quad (6.5)$$

(iii) the usual product estimator $\bar{y}_l^{(2)}$ if

$$\min.\{-R\bar{X}, -\bar{X}(R+2\beta)\} < l_2(\bar{Y}, 1) < \max.\{-R\bar{X}, -\bar{X}(R+2\beta)\} \quad (6.6)$$

6.II: To compare the family of estimators \bar{y}_g with usual unbiased estimator \bar{y}_r , usual ratio estimator $\bar{y}_g^{(1)} = \bar{y}_r (\bar{x}_n/\bar{x}_r)$ and the product estimator $\bar{y}_g^{(2)} = \bar{y}_r (\bar{x}_r/\bar{x}_n)$, we write the mean squared errors of $\bar{y}_g^{(1)}$ and $\bar{y}_g^{(2)}$ to the first degree of approximation, respectively as

$$MSE(\bar{y}_g^{(1)}) = \left[\left(S_y^2 / r \right) + \bar{Y} \{ (1/r) - (1/n) \} C_x^2 (1 - 2K_{yx}) \right], \quad (6.7)$$

$$MSE(\bar{y}_g^{(2)}) = \left[\left(S_y^2 / r \right) + \bar{Y} \{ (1/r) - (1/n) \} C_x^2 (1 + 2K_{yx}) \right]. \quad (6.8)$$

From (5.4) (6.1), (6.7) and (6.8) we note that the suggested family of estimators \bar{y}_l is more efficient than \bar{y}_r , $\bar{y}_g^{(1)}$ and $\bar{y}_g^{(2)}$ respectively if the following inequalities:

$$\min.(0, -\beta\bar{X}) < g_2(\bar{Y}, 1) < \max.(0, -\beta\bar{X}), \quad (6.9)$$

$$\min.\{-R\bar{X}, \bar{X}(R-2\beta)\} < g_2(\bar{Y}, 1) < \max.\{-R\bar{X}, \bar{X}(R-2\beta)\}, \quad (6.10)$$

$$\min.\{-R\bar{X}, -\bar{X}(R+2\beta)\} < g_2(\bar{Y}, 1) < \max.\{-R\bar{X}, -\bar{X}(R+2\beta)\}. \quad (6.11)$$

6.III: Further compare the family of estimators \bar{y}_h with usual unbiased estimator \bar{y}_r , usual ratio estimator $\bar{y}_h^{(1)} = \bar{y}_r(\bar{X}/\bar{x}_r)$ and the usual product estimator $\bar{y}_h^{(2)} = \bar{y}_r(\bar{x}_r/\bar{X})$, we write the

mean squared errors of $\bar{y}_h^{(1)}$ and $\bar{y}_h^{(2)}$ to the first degree of approximation, as $MSE(\bar{y}_h^{(1)}) = \left[(\bar{Y}^2/r) \{C_y^2 + C_x^2(1-2K_{yx})\} \right]$,

(6.12)

$$MSE(\bar{y}_h^{(2)}) = \left[(\bar{Y}^2/r) \{C_y^2 + C_x^2(1+2K_{yx})\} \right]. \quad (6.13)$$

From (6.1) (6.12) and (6.13) it is easily observed that the proposed family of estimators \bar{y}_h is better than

(i) the unbiased estimator \bar{y}_r if

$$\min.(0, -\beta\bar{X}) < h_2(\bar{Y}, 1) < \max.(0, -\beta\bar{X}), \quad (6.14)$$

(ii) usual ratio estimator $\bar{y}_h^{(1)}$ if

$$\min.\{-R\bar{X}, \bar{X}(R-2\beta)\} < h_2(\bar{Y}, 1) < \max.\{-R\bar{X}, \bar{X}(R-2\beta)\}, \quad (6.15)$$

(iii) usual product estimator $\bar{y}_h^{(2)}$ if

$$\min.\{-R\bar{X}, -\bar{X}(R+2\beta)\} < h_2(\bar{Y}, 1) < \max.\{-R\bar{X}, -\bar{X}(R+2\beta)\} \quad (6.16)$$

6.IV: Suppose the form of the estimators under (5.1), (5.2) and (5.3) are similar i.e. $l_2(\bar{Y}, 1) = g_2(\bar{Y}, 1) = h_2(\bar{Y}, 1)$. Then from (5.4), (5.5) and (5.6) we note that the family of estimators \bar{y}_h is better than the families of estimators \bar{y}_l and \bar{y}_g if

$$\min.(0, -2\beta\bar{X}) < h_2(\bar{Y}, 1) < \max.(0, -2\beta\bar{X}).$$

7. Estimators Based on Estimated optimum Values

It is to be mentioned that $z_2(\bar{Y}, 1) = -\beta\bar{X}$ with $z = l, g, h$ is seldom known in practice, hence it is worth advisable to replace this by its estimated optimum value based on the sample data at hand. Since the optimum value of $z_2(\bar{Y}, 1)$ is

$$z_2(\bar{Y}, 1) = -\beta\bar{X} = \delta \text{ (say)} \quad (7.1)$$

or $\delta = -S_{xy}\bar{X}/S_x^2$,

$$\text{where } S_{xy} = \sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y}) / (N-1).$$

Replacing S_{xy} and S_x^2 and \bar{X} by their unbiased estimators

$$S_{x(r)}^2 = \sum_{i=1}^r (x_i - \bar{x}_r)^2 / (r-1), S_{y(r)} = \sum_{i=1}^r (x_i - \bar{x}_r)(y_i - \bar{y}_r) / (r-1) \text{ and } \bar{x}_r \text{ respectively in (7.1) we define a consistent estimate}$$

of δ as

$$\hat{\delta} = -\hat{\beta}_{(r)}\bar{x}_r. \quad (7.2)$$

Thus the following the procedure adopted by Singh and Tracy (2001) and Singh and Horn (2000), we define the following imputation methods based on 'estimated optimum' value:

$$\text{I. } y_{l_1}^* = \begin{cases} y_i & \text{if } i \in D \\ \{nl^*(\bar{y}_r, u, \hat{\delta}) - r\bar{y}_r\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (7.3)$$

$$\text{II. } y_{g_1}^* = \begin{cases} y_i & \text{if } i \in D \\ \{ng^*(\bar{y}_r, v, \hat{\delta}) - r\bar{y}_r\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (7.4)$$

$$\text{III. } y_{h_1}^* = \begin{cases} y_i & \text{if } i \in D \\ \{nh^*(\bar{y}_r, w, \hat{\delta}) - r\bar{y}_r\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (7.5)$$

where $l^*(\bar{y}_r, u, \hat{\delta})$, $g^*(\bar{y}_r, v, \hat{\delta})$ and $h^*(\bar{y}_r, w, \hat{\delta})$ are the functions of $(\bar{y}_r, u, \hat{\delta})$, $(\bar{y}_r, v, \hat{\delta})$ and $(\bar{y}_r, w, \hat{\delta})$ respectively. The functions $l^*(\bar{y}_r, u, \hat{\delta})$, $g^*(\bar{y}_r, v, \hat{\delta})$ and $h^*(\bar{y}_r, w, \hat{\delta})$ assume values in bounded closed convex subset $Q^* \in R_3$, which contain the point

$M = (\bar{Y}, 1, \delta)$ and are such that

$$\left. \begin{aligned} l^*(M) = g^*(M) = h^*(M) = \bar{Y} \\ \Rightarrow l_1^*(M) = g_1^*(M) = h_1^*(M) = 1 \end{aligned} \right\} \quad (7.6)$$

$$l_2^*(M) = g_2^*(M) = h_2^*(M) = -\beta\bar{X} \quad (7.7)$$

$$\text{and } l_3^*(M) = g_3^*(M) = h_3^*(M) = 0 \quad (7.8)$$

where $z_j^*(M)$ with $z^* = l^*, g^*, h^*$ and $j = 1, 2, 3$; denote the first order partial derivative of the function $z_j^*(\cdot)$ about the point $M = (\bar{Y}, 1, \delta)$.

The point estimators of the population mean \bar{Y} under (7.3), (7.4) and (7.5) methods of imputation are respectively given by

$$\bar{y}_{l^*} = l^*(\bar{y}_r, u, \hat{\delta}). \quad (7.9)$$

$$\bar{y}_{g^*} = g^*(\bar{y}_r, v, \hat{\delta}). \quad (7.10)$$

$$\bar{y}_{h^*} = h^*(\bar{y}_r, w, \hat{\delta}). \quad (7.11)$$

Under the conditions (7.6) (7.7) and (7.8) it can be shown to the first degree of approximation that

$$MSE(\bar{y}_{l^*}) = S_y^2 \left[\left\{ \frac{1 - \rho_{yx}^2}{r} \right\} + \left\{ \frac{1}{r} - \frac{1}{n} \right\} \rho_{yx}^2 \right] = \min.MSE(\bar{y}_f), \quad (7.12)$$

$$MSE(\bar{y}_{g^*}) = S_y^2 \left[\left\{ \frac{1 - \rho_{yx}^2}{r} \right\} + \left\{ \frac{\rho_{yx}^2}{n} \right\} \right] = \min.MSE(\bar{y}_g), \quad (7.13)$$

$$MSE(\bar{y}_{h^*}) = \left\{ S_y^2 \frac{1 - \rho_{yx}^2}{r} \right\} = \min.MSE(\bar{y}_h). \quad (7.14)$$

It can easily be seen that the proposed family of estimators \bar{y}_{h^*} is better than \bar{y}_{l^*} and \bar{y}_{g^*} .

8. Some Generalized Procedures of Imputation and Their Estimators

Let $L(\bar{y}_r, u, a)$, $G(\bar{y}_r, v, b)$, $H(\bar{y}_r, w, d)$ be the functions of (\bar{y}_r, u, a) , (\bar{y}_r, v, b) and (\bar{y}_r, w, d) respectively such that

$L(Z) = \bar{Y} \Rightarrow L_1(Z) = 1$, $G(Z) = \bar{Y} \Rightarrow G_1(Z) = 1$ and $H(Z) = \bar{Y} \Rightarrow H_1(Z) = 1$ with $Z = (\bar{Y}, 1, 1)$ and such that these satisfy the following conditions:

1. Whatever be the sample chosen (\bar{y}_r, u, a) , (\bar{y}_r, v, b) and (\bar{y}_r, w, d) assume values in abounded closed convex subset, R^* , R^{**} and R^{***} respectively of the three dimensional real space containing the point $Z = (\bar{Y}, 1, 1)$.
2. In R^* , R^{**} and R^{***} respectively the functions $L(\bar{y}_r, u, a)$, $G(\bar{y}_r, v, b)$ and $H(\bar{y}_r, w, d)$ are continuous and bounded.
3. The first and second order partial derivatives of $L(\bar{y}_r, u, a)$, $G(\bar{y}_r, v, b)$ and $H(\bar{y}_r, w, d)$ exist and are continuous and bounded in R^* , R^{**} and R^{***} respectively.

8.I: When the population mean \bar{X} and mean square S_x^2 are known, we define the following imputation procedure as:

$$y_{L_1} = \begin{cases} y_i & \text{if } i \in D \\ \{nL(\bar{y}_r, u, a) - r\bar{y}_r\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (8.1)$$

8.II: When the population mean \bar{X} and mean square S_x^2 of the auxiliary variable x are not known, we define the following imputation procedure as:

$$y_{G_1} = \begin{cases} y_i & \text{if } i \in D \\ \{nG(\bar{y}_r, v, b) - r\bar{y}_r\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (8.2)$$

8.III: When the population mean \bar{X} and mean square S_x^2 are not known, we define another imputation procedure as:

$$y_{H_1} = \begin{cases} y_i & \text{if } i \in D \\ \{nH(\bar{y}_r, w, d) - r\bar{y}_r\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (8.3)$$

Under (8.1), (8.2) and (8.3) the point estimators of the population mean \bar{Y} are

$$\bar{y}_L = L(\bar{y}_r, u, a), \quad (8.4)$$

$$\bar{y}_G = G(\bar{y}_r, v, b), \quad (8.5)$$

$$\bar{y}_H = H(\bar{y}_r, w, d). \quad (8.6)$$

9. Properties of the Suggested Generalized Estimators \bar{y}_L , \bar{y}_G and \bar{y}_H

9.I: To obtain the biases of the estimators \bar{y}_L , \bar{y}_G and \bar{y}_H we further assume that third order partial derivatives of $L(\bar{y}_r, u, a)$, $G(\bar{y}_r, v, b)$ and $H(\bar{y}_r, w, d)$ exist and are continuous and bounded. Then expanding $L(\bar{y}_r, u, a)$, $G(\bar{y}_r, v, b)$ and $H(\bar{y}_r, w, d)$ about the point $Z = (\bar{Y}, 1, 1)$ in a third order Taylor's series we obtain:

$$\bar{y}_F = \left[\begin{aligned} & F(Z) + (\bar{y}_r - \bar{Y})F_1(Z) + (u-1)F_2(Z) + (a-1)F_3(Z) \\ & + (1/2) \left\{ (u-1)^2 F_{22}(Z) + (a-1)^2 F_{33}(Z) + 2(u-1)(a-1)F_{23}(Z) \right\} \\ & + (1/6) \left\{ (\bar{y}_r - \bar{Y}) \frac{\partial}{\partial \bar{y}_r} + (u-1) \frac{\partial}{\partial u} + (a-1) \frac{\partial}{\partial a} \right\}^3 F(\bar{y}_r^*, u^*, a^*) \end{aligned} \right] \quad (9.1)$$

$$\bar{y}_G = \left[\begin{array}{l} G(Z) + (\bar{y}_r - \bar{Y})G_1(Z) + (v-1)G_2(Z) + (b-1)G_3(Z) \\ + (1/2) \left\{ (v-1)^2 G_{22}(Z) + (b-1)^2 G_{33}(Z) + 2(v-1)(b-1)G_{23}(Z) \right\} \\ + 2(\bar{y}_r - \bar{Y})(v-1)G_{12}(Z) + 2(\bar{y}_r - \bar{Y})(b-1)G_{13}(Z) \\ + (1/6) \left\{ (\bar{y}_r - \bar{Y}) \frac{\partial}{\partial \bar{y}_r} + (v-1) \frac{\partial}{\partial v} + (b-1) \frac{\partial}{\partial b} \right\}^3 G(\bar{y}_r^*, v^*, b^*) \end{array} \right] \quad (9.2)$$

$$\bar{y}_H = \left[\begin{array}{l} H(Z) + (\bar{y}_r - \bar{Y})H_1(Z) + (w-1)H_2(Z) + (d-1)H_3(Z) \\ + (1/2) \left\{ (w-1)^2 H_{22}(Z) + (d-1)^2 H_{33}(Z) + 2(w-1)(d-1)H_{23}(Z) \right\} \\ + 2(\bar{y}_r - \bar{Y})(w-1)H_{12}(Z) + 2(\bar{y}_r - \bar{Y})(d-1)H_{13}(Z) \\ + (1/6) \left\{ (\bar{y}_r - \bar{Y}) \frac{\partial}{\partial \bar{y}_r} + (w-1) \frac{\partial}{\partial w} + (d-1) \frac{\partial}{\partial d} \right\}^3 G(\bar{y}_r^*, w^*, d^*) \end{array} \right] \quad (9.3)$$

where $\bar{y}_r^* = \bar{Y} + \psi^*(\bar{y}_r - \bar{Y})$, $u^* = 1 + \psi^*(u-1)$, $a^* = 1 + \psi^*(a-1)$, $v^* = 1 + \psi^*(v-1)$, $w^* = 1 + \psi^*(w-1)$, $b^* = 1 + \psi^*(b-1)$, $d^* = 1 + \psi^*(d-1)$, $0 < \psi^* < 1$.

$L_j(Z), G_j(Z), H_j(Z)$, $j = 1, 2, 3$ denote the first order partial derivatives of $L(\cdot)$, $G(\cdot)$ and $H(\cdot)$ about the point $Z = (\bar{Y}, 1, 1)$. Noting that $L(Z) = G(Z) = H(Z) = \bar{Y}$, $L_1(Z) = G_1(Z) = H_1(Z) = 1$ and expressing (8.7), (8.8) and (8.9) in terms of $\varepsilon_0, \varepsilon_1, \varepsilon_2, \eta_1, \eta_2$ we have

$$(\bar{y}_L - \bar{Y}) = \bar{Y}\varepsilon_0 + \varepsilon_1 L_2(Z) + \varepsilon_2 L_3(Z) + (1/2) \left\{ \varepsilon_1^2 L_{22}(Z) + \varepsilon_2^2 L_{33}(Z) + 2\varepsilon_1 \varepsilon_2 L_{23}(Z) \right. \\ \left. + 2\bar{Y}\varepsilon_0 \varepsilon_1 L_{12}(Z) + 2\bar{Y}\varepsilon_0 \varepsilon_2 L_{13}(Z) \right\} \quad (9.4)$$

$$(\bar{y}_G - \bar{Y}) = \bar{Y}\varepsilon_0 + (\eta_1 - \varepsilon_1 + \varepsilon_1^2 - \eta_1 \varepsilon_1)G_2(Z) + (\eta_2 - \varepsilon_2 + \varepsilon_2^2 - \eta_2 \varepsilon_2)G_3(Z) \\ + (1/2) \left\{ (\eta_1 - \varepsilon_1)^2 G_{22}(Z) + (\eta_2 - \varepsilon_2)^2 G_{33}(Z) + 2(\eta_1 - \varepsilon_1)(\eta_2 - \varepsilon_2)G_{23}(Z) \right. \\ \left. + 2\bar{Y}\varepsilon_0 (\eta_1 - \varepsilon_1) G_{12}(Z) + 2\bar{Y}\varepsilon_0 (\eta_2 - \varepsilon_2) G_{13}(Z) \right\} \quad (9.5)$$

$$(\bar{y}_H - \bar{Y}) = \bar{Y}\varepsilon_0 + \eta_1 H_2(Z) + \eta_2 H_3(Z) + (1/2) \left\{ \eta_1^2 H_{22}(Z) + \eta_2^2 H_{33}(Z) + 2\eta_1 \eta_2 \right. \\ \left. H_{23}(Z) + 2\bar{Y}\varepsilon_0 \eta_1 H_{12}(Z) + 2\bar{Y}\varepsilon_0 \eta_2 H_{13}(Z) \right\} \quad (9.6)$$

Taking expectation both sides of (9.4), (9.5) and (9.6) we obtain the biases of the estimators \bar{y}_L , \bar{y}_G and \bar{y}_H up to the terms of order n^{-1}

$$B(\bar{y}_L) = (1/2n) \left\{ C_x^2 L_{22}(Z) + (\beta_2(x) - 1)L_{33}(Z) + 2\gamma_1 C_x L_{23}(Z) \right. \\ \left. + 2\bar{Y}\rho_{yx} C_y C_x L_{12}(Z) + 2\bar{Y}\lambda_0 L_{13}(Z) \right\}, \quad (9.7)$$

$$B(\bar{y}_G) = (1/2n) \left\{ (1/r) - (1/n) \right\} \left\{ C_x^2 G_{22}(Z) + (\beta_2(x) - 1)G_{33}(Z) + 2\gamma_1 C_x G_{23}(Z) \right. \\ \left. + 2\bar{Y}\rho_{yx} C_y C_x G_{12}(Z) + 2\bar{Y}\lambda_0 G_{13}(Z) \right\}, \quad (9.8)$$

$$B(\bar{y}_H) = (1/2r) \left\{ C_x^2 H_{22}(Z) + (\beta_2(x) - 1)H_{33}(Z) + 2\gamma_1 C_x H_{23}(Z) \right. \\ \left. + 2\bar{Y}\rho_{yx} C_y C_x H_{12}(Z) + 2\bar{Y}\lambda_0 H_{13}(Z) \right\}, \quad (9.9)$$

Up to the terms of order n^{-1} , the mean squared errors of \bar{y}_L , \bar{y}_G and \bar{y}_H respectively are

$$MSE(\bar{y}_L) = E(\bar{y}_L - \bar{Y})^2 \\ = E\left\{ \bar{Y}^2 \varepsilon_0^2 + \varepsilon_1^2 L_2^2(Z) + \varepsilon_2^2 L_3^2(Z) + 2\bar{Y}\varepsilon_0 \varepsilon_1 L_2(Z) + 2\bar{Y}\varepsilon_0 \varepsilon_2 L_3(Z) + 2\varepsilon_1 \varepsilon_2 L_2(Z) L_3(Z) \right\} \\ = \left[(1/r) S_y^2 + (1/n) \left\{ C_x^2 L_2^2(Z) + (\beta_2(x) - 1)L_3^2(Z) + 2\gamma_1 C_x L_2(Z) L_3(Z) \right. \right. \\ \left. \left. + 2\bar{Y}\rho_{yx} C_y C_x L_2(Z) + 2\bar{Y}\lambda_0 L_3(Z) \right\} \right] \quad (9.10)$$

$$MSE(\bar{y}_G) = E(\bar{y}_G - \bar{Y})^2 \\ = E\left\{ \bar{Y}^2 \varepsilon_0^2 + (\eta_1 - \varepsilon_1)^2 G_2^2(Z) + (\eta_2 - \varepsilon_2)^2 G_3^2(Z) + 2(\eta_1 - \varepsilon_1)(\eta_2 - \varepsilon_2)G_2(Z)G_3(Z) \right. \\ \left. + 2\bar{Y}\varepsilon_0 (\eta_1 - \varepsilon_1)G_2(Z) + 2\bar{Y}\varepsilon_0 (\eta_2 - \varepsilon_2)G_3(Z) \right\}$$

$$= \left[(1/r) S_y^2 + \left\{ (1/r) - (1/n) \right\} \left\{ C_x^2 G_{22}(Z) + (\beta_2(x) - 1) G_3^2(Z) + 2\gamma_1 C_x G_2(Z) G_3(Z) + 2\bar{Y} \rho_{yx} C_y C_x G_2(Z) + 2\bar{Y} \lambda_0 G_3(Z) \right\} \right] \quad (9.11)$$

$$MSE(\bar{y}_H) = E(\bar{y}_H - \bar{Y})^2 = \left[(1/r) S_y^2 + C_x^2 H_2^2(Z) + (\beta_2(x) - 1) H_3^2(Z) + 2\gamma_1 C_x H_2(Z) H_3(Z) + 2\bar{Y} \rho_{yx} C_y C_x H_2(Z) + 2\bar{Y} \lambda_0 H_3(Z) \right] \quad (9.12)$$

Minimization of (9.10), (9.11) and (9.12) yield the optimum values of $\{L_2(Z), L_3(Z)\}$, $\{G_2(Z), G_3(Z)\}$ and $\{H_2(Z), H_3(Z)\}$ respectively as

$$L_2(Z) = \left[\frac{R\bar{X}/C_x}{\left\{ \lambda_0 \gamma_1 - \rho_{yx} C_y (\beta_2(x) - 1) \right\} / (\beta_2(x) - \beta_1 - 1)} \right] = \psi^* \text{ (say)} \quad (9.13)$$

$$L_3(Z) = \left[\frac{R\bar{X} (\rho_{yx} C_y \gamma_1 - \lambda_0)}{(\beta_2(x) - \beta_1 - 1)} \right] = \theta \text{ (say)} \quad (9.14)$$

$$\left. \begin{aligned} G_2(Z) &= \psi \\ G_3(Z) &= \theta \end{aligned} \right\} \text{ and } \left. \begin{aligned} H_2(Z) &= \psi \\ H_3(Z) &= \theta \end{aligned} \right\} \quad (9.15)$$

Substitution of (ψ, θ) in (9.10), (9.11) and (9.12) respectively yield the minimum mean squared errors of \bar{y}_L , \bar{y}_G and \bar{y}_H respectively as

$$\min. MSE(\bar{y}_L) = S_y^2 \left\{ (1/r) - (1/n) (\rho_{yx}^2 + \Delta) \right\} \quad (9.16)$$

$$\min. MSE(\bar{y}_G) = \left[S_y^2 \left\{ (1/r) - (1/r) - (1/n) \right\} (\rho_{yx}^2 + \Delta) \right] \quad (9.17)$$

$$\min. MSE(\bar{y}_H) = \left(S_y^2 / r \right) \left\{ 1 - (\rho_{yx}^2 + \Delta) \right\} \quad (9.18)$$

Thus we state the following Theorem:

Theorem 9.1

To the first degree of approximation,

$$\mathbf{9. I} \quad MSE(\bar{y}_L) \geq S_y^2 \left\{ (1/r) - (1/n) (\rho_{yx}^2 + \Delta) \right\}$$

with equality holding if

$$L_2(Z) = \psi \text{ and } L_3(Z) = \theta.$$

$$\mathbf{9. II} \quad MSE(\bar{y}_G) \geq \left[S_y^2 \left\{ (1/r) - (1/r) - (1/n) \right\} (\rho_{yx}^2 + \Delta) \right]$$

with equality holding if

$$G_2(Z) = \psi \text{ and } G_3(Z) = \theta.$$

$$\mathbf{9. III} \quad MSE(\bar{y}_H) \geq \left(S_y^2 / r \right) \left\{ 1 - (\rho_{yx}^2 + \Delta) \right\}$$

with equality holding if

$$H_2(Z) = \psi \text{ and } H_3(Z) = \theta.$$

where ψ and θ are same as defined by (7.13).

Remark 9.1

The following estimators:

$$\bar{y}_{L(1)} = \bar{y}_r u^\alpha a^\beta, \bar{y}_{L(2)} = \bar{y}_r \left[\frac{1 + \alpha(u-1)}{1 - \beta(a-1)} \right], \bar{y}_{L(3)} = \bar{y}_r \{1 + \alpha(u-1) + \beta(a-1)\},$$

$$\bar{y}_{L(4)} = \bar{y}_r \{1 - \alpha(u-1) - \beta(a-1)\}^{-1}, \bar{y}_{L(5)} = \{\bar{y}_r + \alpha(u-1) + \beta(a-1)\},$$

$$\bar{y}_{L(6)} = \bar{y}_r \exp\{\alpha(u-1) + \beta(a-1)\}, \bar{y}_{L(7)} = \bar{y}_r \exp\{\alpha \log u + \beta \log a\}, \text{ etc are members of the proposed family of estimators } \bar{y}_F,$$

where α and β are suitably chosen constants. The biases and mean squared errors of the estimators $\bar{y}_{L(j)}$, $j = 1$ to 7 can be easily obtained from (8.7) and (8.10) just by putting the suitable values of the derivatives $L_2(Z), L_3(Z), L_{22}(Z), L_{23}(Z), L_{33}(Z)$.

Remark 9.2

The following estimators:

$$\bar{y}_{G(1)} = \bar{y}_r v^\alpha b^\beta, \bar{y}_{G(2)} = \bar{y}_r \left[\frac{1 + \alpha(v-1)}{1 - \beta(b-1)} \right], \bar{y}_{G(3)} = \bar{y}_r \{1 + \alpha(v-1) + \beta(b-1)\},$$

$$\bar{y}_{G(4)} = \bar{y}_r \{1 - \alpha(v-1) - \beta(b-1)\}^{-1}, \bar{y}_{G(5)} = \{\bar{y}_r + \alpha(v-1) + \beta(b-1)\},$$

$$\bar{y}_{G(6)} = \bar{y}_r \exp\{\alpha(v-1) + \beta(b-1)\}, \bar{y}_{G(7)} = \bar{y}_r \exp\{\alpha \log v + \beta \log b\}, \text{ etc are members of the family of estimators } \bar{y}_G.$$

The biases and mean squared errors of the estimators $\bar{y}_{G(j)}$, $j = 1$ to 7 can be easily obtained from (9.8) and (9.11) just by putting the values of the derivatives $G_2(Z), G_3(Z), G_{22}(Z), G_{23}(Z), G_{33}(Z)$.

Remark 9.3

The following estimators:

$$\bar{y}_{H(1)} = \bar{y}_r w^\alpha d^\beta, \bar{y}_{H(2)} = \bar{y}_r \left[\frac{1 + \alpha(w-1)}{1 - \beta(d-1)} \right], \bar{y}_{H(3)} = \bar{y}_r \{1 + \alpha(w-1) + \beta(d-1)\},$$

$$\bar{y}_{H(4)} = \bar{y}_r \{1 - \alpha(w-1) - \beta(d-1)\}^{-1}, \bar{y}_{H(5)} = \{\bar{y}_r + \alpha(w-1) + \beta(d-1)\},$$

$\bar{y}_{H(6)} = \bar{y}_r \exp\{\alpha(w-1) + \beta(d-1)\}$, $\bar{y}_{H(7)} = \bar{y}_r \exp\{\alpha \log w + \beta \log d\}$, etc are members of the family of estimators \bar{y}_H . The biases and mean squared errors of the estimators $\bar{y}_{H(j)}$, $j = 1$ to 7 can be easily obtained from (9.9) and (9.12) just by putting the values of the derivatives $H_2(Z), H_3(Z), H_{22}(Z), H_{23}(Z), H_{33}(Z)$.

It is to be mentioned that the families of estimators \bar{y}_F, \bar{y}_G and \bar{y}_H are very large. If the parameters in $F(\bar{y}_r, u, a), G(\bar{y}_r, v, b)$ and $H(\bar{y}_r, w, d)$ are so chosen that they satisfy (9.13), (9.14) and (9.15) respectively, the resulting estimators will have mean squared errors given by (9.16), (9.17) and (9.18).

10. Estimators Based on Estimated Optimum Values

The optimum values of $L_2(Z), L_3(Z)$ given by (9.13) can be expressed as

$$\left. \begin{aligned} \psi &= \left(\bar{X} / S_{x(r)}^2 \right) \left\{ \left[\mu_{12} \mu_{03} - \mu_{11} (\mu_{04} - S_{x(r)}^4) \right] / \left[(\mu_{04} - S_{x(r)} \mu_{03} - S_{x(r)}^4) \right] \right\} \\ \hat{\theta} &= \left\{ \left[\mu_{11} \mu_{03} - S_{x(r)}^2 \mu_{12} \right] / \left[(\mu_{04} - S_{x(r)} \mu_{03} - S_{x(r)}^4) \right] \right\} \end{aligned} \right\} \quad (10.1)$$

In practice the exact values of $\hat{\psi}$ and $\hat{\theta}$ or the guessed values of ψ and θ closer to exact values of ψ and θ may be rarely known in practice, hence it is advisable to replace them by their estimates from sample values. The consistent estimates of ψ and θ are respectively given by

$$\left. \begin{aligned} \hat{\psi} &= \left(\bar{x}_r / s_{x(r)}^2 \right) \left\{ \left[\hat{\mu}_{12} \hat{\mu}_{03} - \hat{\mu}_{11} (\hat{\mu}_{04} - s_{x(r)}^4) \right] / \left[(\hat{\mu}_{04} - s_{x(r)} \hat{\mu}_{03} - s_{x(r)}^4) \right] \right\} \\ \hat{\theta} &= \left\{ \left[\hat{\mu}_{11} \hat{\mu}_{03} - s_{x(r)}^2 \hat{\mu}_{12} \right] / \left[(\hat{\mu}_{04} - s_{x(r)} \hat{\mu}_{03} - s_{x(r)}^4) \right] \right\} \end{aligned} \right\} \quad (10.2)$$

where $\hat{\mu}_{pq} = (1/r) \sum_{j=1}^r (y_j - \bar{y}_r)^p (x_j - \bar{x}_r)^q$, (p, q) being non-negative integers.

Thus the imputation methods based on estimated optimum values $(\hat{\psi}, \hat{\theta})$ of (ψ, θ) are given by

$$y_{L_1^*} = \begin{cases} y_i & \text{if } i \in D \\ \left\{ nL^*(\bar{y}_r, u, a, \hat{\psi}, \hat{\theta}) - r\bar{y}_r \right\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (10.3)$$

$$y_{G_1^*} = \begin{cases} y_i & \text{if } i \in D \\ \left\{ nG^*(\bar{y}_r, v, b, \hat{\psi}, \hat{\theta}) - r\bar{y}_r \right\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (10.4)$$

$$y_{H_1^*} = \begin{cases} y_i & \text{if } i \in D \\ \left\{ nH^*(\bar{y}_r, w, d, \hat{\psi}, \hat{\theta}) - r\bar{y}_r \right\} / (n-r) & \text{if } i \in D^c \end{cases} \quad (10.5)$$

where $L^*(\bar{y}_r, u, a, \hat{\psi}, \hat{\theta})$, $G^*(\bar{y}_r, v, b, \hat{\psi}, \hat{\theta})$ and $H^*(\bar{y}_r, w, d, \hat{\psi}, \hat{\theta})$ are the functions of $(\bar{y}_r, u, a, \hat{\psi}, \hat{\theta})$, $(\bar{y}_r, v, b, \hat{\psi}, \hat{\theta})$ and $(\bar{y}_r, w, d, \hat{\psi}, \hat{\theta})$ respectively such that

$$\left. \begin{aligned} L^*(J) = G^*(J) = H^*(J) = \bar{Y} &\Rightarrow L_1^*(J) = G_1^*(J) = H_1^*(J) = 1 \\ L_2^*(J) = G_2^*(J) = H_2^*(J) = \psi, L_3^*(J) = G_3^*(J) = H_3^*(J) = \theta \\ L_4^*(J) = G_4^*(J) = H_4^*(J) = 0, L_5^*(J) = G_5^*(J) = H_5^*(J) = 0 \end{aligned} \right\} \quad (10.6)$$

where $J = (\bar{y}_r, 1, 1, \psi, \theta)$,

$$L_1^*(J) = \left. \frac{\partial L^*(.)}{\partial \bar{y}_r} \right|_J, \quad L_2^*(J) = \left. \frac{\partial L^*(.)}{\partial u} \right|_J, \quad L_3^*(J) = \left. \frac{\partial L^*(.)}{\partial a} \right|_J, \quad L_4^*(J) = \left. \frac{\partial L^*(.)}{\partial \hat{\psi}} \right|_J, \quad L_5^*(J) = \left. \frac{\partial L^*(.)}{\partial \hat{\theta}} \right|_J,$$

$$G_1^*(J) = \left. \frac{\partial G^*(.)}{\partial \bar{y}_r} \right|_J, \quad G_2^*(J) = \left. \frac{\partial G^*(.)}{\partial v} \right|_J, \quad G_3^*(J) = \left. \frac{\partial G^*(.)}{\partial b} \right|_J, \quad G_4^*(J) = \left. \frac{\partial G^*(.)}{\partial \hat{\psi}} \right|_J, \quad G_5^*(J) = \left. \frac{\partial G^*(.)}{\partial \hat{\theta}} \right|_J,$$

$$H_1^*(J) = \left. \frac{\partial H^*(.)}{\partial \bar{y}_r} \right|_J, \quad H_2^*(J) = \left. \frac{\partial H^*(.)}{\partial w} \right|_J, \quad H_3^*(J) = \left. \frac{\partial H^*(.)}{\partial d} \right|_J, \quad H_4^*(J) = \left. \frac{\partial H^*(.)}{\partial \hat{\psi}} \right|_J, \quad H_5^*(J) = \left. \frac{\partial H^*(.)}{\partial \hat{\theta}} \right|_J.$$

Under the imputation methods (10.3), (10.4) and (10.5) the point estimators of the population mean

$$\bar{Y} \text{ as } \bar{y}_{L^*} = L^*(\bar{y}_r, u, a, \hat{\psi}, \hat{\theta}) \quad (10.7)$$

$$\bar{y}_{G^*} = G^*(\bar{y}_r, v, b, \hat{\psi}, \hat{\theta}) \quad (10.8)$$

$$\bar{y}_{H^*} = H^*(\bar{y}_r, w, d, \hat{\psi}, \hat{\theta}) \quad (10.9)$$

It can be shown to the first degree of approximation that

$$MSE(\bar{y}_{L^*}) = S_y^2 \left\{ (1/r) - (1/n) (\rho_{yx}^2 + \Delta) \right\} = \min.MSE(\bar{y}_L), \quad (10.10)$$

$$MSE(\bar{y}_{G^*}) = \left[S_y^2 \left\{ (1/r) - (1/n) (\rho_{yx}^2 + \Delta) \right\} \right] = \min.MSE(\bar{y}_G), \quad (10.11)$$

$$MSE(\bar{y}_{H^*}) = \left(S_y^2 / r \right) \left\{ 1 - (\rho_{yx}^2 + \Delta) \right\} = \min.MSE(\bar{y}_H). \quad (10.12)$$

11. Efficiency Comparisons

From (10.10) and (10.11) we have

$$MSE(\bar{y}_{G^*}) - MSE(\bar{y}_{L^*}) = \min.MSE(\bar{y}_G) - \min.MSE(\bar{y}_L),$$

which is less than zero if

$$r < (n/2). \quad (11.1)$$

From (10.10), (10.11) and (10.12) we have

$$\begin{aligned} MSE(\bar{y}_{L^*}) - MSE(\bar{y}_{H^*}) &= \{ \min.MSE(\bar{y}_L) - \min.MSE(\bar{y}_H) \} \\ &= \{ (1/r) - (1/n) \} (\rho_{yx}^2 + \Delta) S_y^2 \geq 0, \end{aligned} \quad (11.2)$$

$$\begin{aligned} MSE(\bar{y}_{G^*}) - MSE(\bar{y}_{H^*}) &= \{ \min.MSE(\bar{y}_G) - \min.MSE(\bar{y}_H) \} \\ &= \left(S_y^2 / n \right) (\rho_{yx}^2 + \Delta) \geq 0. \end{aligned} \quad (11.3)$$

From (11.2) and (11.3) it follows that the proposed family of estimators \bar{y}_{H^*} (or \bar{y}_H) is better than \bar{y}_{L^*} (or \bar{y}_L). Thus the family of estimators \bar{y}_{H^*} (or \bar{y}_H) is the best estimator among the families of estimators \bar{y}_{L^*} (or \bar{y}_L), \bar{y}_{G^*} (or \bar{y}_G), \bar{y}_{H^*} (or \bar{y}_H), \bar{y}_{L^*} (or \bar{y}_L), \bar{y}_{G^*} (or \bar{y}_G) and \bar{y}_{H^*} (or \bar{y}_H).

12. Conclusion

The proposed generalized procedure of imputation methods are theoretically sound and of notable importance. It covers imputation methods reported by Singh and Horn (2000), Ahmed et al (2006), Shukla and Thakur (2008) and Pandey et al (2015) and their point estimators of the population mean. The suggested imputation methods in (8.1) (or (10.3)), (8.2) (or (10.4)) and (8.3) (or (10.5)) would be worth using when the relationship between study variate y and the auxiliary variate x is markedly non-linear and $(\beta_2(x) - \beta_1 - 1)$ is small. From the biases and MSEs of the resulting

families of estimators \bar{y}_l (or \bar{y}_{l^*}), \bar{y}_g (or \bar{y}_{g^*}), \bar{y}_h (or \bar{y}_{h^*}), \bar{y}_L (or \bar{y}_{L^*}), \bar{y}_G (or \bar{y}_{G^*}) and \bar{y}_H (or \bar{y}_{H^*}) of the proposed imputation methods, the biases and MSEs of any estimator belonging to these families of estimators can be easily obtained just by inserting the suitable values of the derivatives. Thus the envisaged procedure of imputation unifies several results. It is further noted that the suggested imputation methods and the resulting families of estimators \bar{y}_L (or \bar{y}_{L^*}), \bar{y}_G (or \bar{y}_{G^*}) and \bar{y}_H (or \bar{y}_{H^*}) are more efficient than those considered by Singh and Horn (2000), Ahmed et al (2006), Shukla and Thakur (2008) and Pandey et al (2015). So the envisaged imputation procedures \bar{y}_l (or \bar{y}_{l^*}), \bar{y}_g (or \bar{y}_{g^*}), \bar{y}_h (or \bar{y}_{h^*}), \bar{y}_L (or \bar{y}_{L^*}), \bar{y}_G (or \bar{y}_{G^*}) and \bar{y}_H (or \bar{y}_{H^*}) are put forward for their use in practice.

References

- Ahmed, M.S., Al-Titi, O, Al-Rawi, Z. and Abu-Dayyeh, W. (2006): Estimation of a population mean using different imputation methods. *Statistics in Transition*, 7, (6), 1247-1264.
- Heitjan, D.F. and Basu, S. (1996): Distinguishing 'Missing at Random' and 'Missing Completely at Random'. *The American Statistician*, 50, 207-213.
- Mohamed, C., Sedory, S.A., Singh, S. (2018): Improved Mean Methods of Imputation, *Statistics, Optimization and Information Computing*, vol. 6, 526-535.
- Pandey, R. and Yadav, K. (2016): Mean estimation under imputation based on two-phase sampling design using an auxiliary variable. *PJSOR*, 12(4), 639-658.
- Pandey, R. and Yadav, K. (2017): Population variance estimation using factor type imputation method. *Statistics in Transition, New Series*, 18(3), 375-392.
- Pandey, R. and Yadav, K. (2018): On variance estimation under factor type imputation using auxiliary attribute, *International Journal of agricultural statistical sciences*, 4(1), 95-101.
- Pandey, R., Singh, N. and Yadav, K. (2015): Separate regression type imputation methods to estimate population mean. *International Journal of Computer & Mathematical Sciences*, 4, special issue, 2347-8527.
- Rao, J.,N.,K. and Sitter, R.R. (1995): Variance estimation under two-phase sampling with application to imputation for missing data. *Biometrika*, 82: 453-460.
- Rubin, R.B. (1967): Interference and missing data. *Biometrika*, 63(6), 581-592.
- Shukla, D. and Thakur, N.S. (2008): Estimation of mean with imputation of missing data using factor type estimator. *Statistics in Transition*, 9, (1), 33-48.
- Singh, H.P. and Tracy, D.S. (2001): Estimation of population mean in presence of random non-response in sample surveys. *Statistica, anno LXI*, n.2, 231-248.
- Singh, S. and Horn, S. (2000): Compromized imputation in survey sampling. *Metrika*, 51:267-276.
- Singh, S., Joarder, A.H. and Tracy, D.S. (2001): Median estimation using double sampling. *Australian and New Zealand Journal of Statistics*, 43, (1), 33-46.

Srivastava, S.K. (1971): A generalized estimator for the mean of a finite population using multi-auxiliary information. *Journal of American Statistical Association*, 66, 404-407.

Srivastava, S.K. and Jhajj, H.S. (1981): A class of estimators of the population mean in survey sampling using auxiliary information. *Biometrika*, 68, 341-343.

Toutenburg, H. and Srivastava, V.K. (1998): Estimation of ratio of population means in survey sampling when some observations are missing. *Metrika*, 48, 177-187.