CHAPTER – 3

AN EFFICIENT ESTIMATOR UNDER MISSING DATA, IMPUTATION AND MEASUREMENT ERROR

3.1 INTRODUCTION

We look for the alternative estimators in the setup of missing data and measurement error. In surveys, we generally receive the complex type of collected data. Information may be missing or may be under-reported, over-reported. The variable of main interest might suffer from these sources of incompleteness. This chapter presents some new estimators using imputation technique and measurement error modals.

* A part of this chapter has been published in Research & Reviews: A Journal of Statistics (STM Journals) in Volume 1, Issue 1, pp. 1-8. (Please see appendix for full length paper).

* A part of this chapter has been published in Journal of Current Engineering Research in Volume 2, Issue 1, pp. 50-55. (Please see appendix for full length paper).
3.2 ESTIMATION IN PRESENCE OF MISSING DATA

The problem of non-response is one of the enormous and most realistic problems for interviewers in any household survey. As solution to this problem there are so many methods in the literature of survey-sampling like Mean method, Ratio method, Product method, Compromised method and so on. To look up at the accuracy in sample surveys theory the use of helping variables for estimation of population mean of a variable under study has played an influential role. A number of estimators are accessible in the literature of sample surveys where supporting information is the contributor to improve the methodology. Out of all ratio and product estimators are high-quality examples to state this. The ratio estimation method is practical when the correlation coefficient between the study and auxiliary variable is positive [Cochran (1940, 42, 77)]. If the correlation coefficient between the study and auxiliary variable is negative then the use of product estimation will make the study effective in practice. [Robson (1957) and Murthy (1964)].

Abu-Dayyeh et al. (2003) used auxiliary variables to show estimators of finite population mean. Sahoo and Sahoo (2001) discussed about predictive estimation of finite population mean in two-phase sampling using two auxiliary variables. Singh and Shukla (1987) have a discussion on one parameter family of factor type ratio estimator. In a similar study Shukla et al. (1991) transformed the factor type estimator to make the estimation more effective. Shukla (2002) studied F-T estimator where sampling procedure undertaken was in two-phase sampling. Singh and Tailor (2003) described a procedure to cope up the variation in surveys by using correlation coefficient ($\rho$) as follows
\[ \hat{Y} = \bar{y} \left( \frac{\bar{X} + \rho}{\bar{x} + \rho} \right) \]


### 3.2.1 NOTATIONS AND ASSUMPTIONS

Notations for the study are:

\( \bar{Y}; \bar{X} \) : Population Parameters.

\( \bar{y}; \bar{x} \) : Mean per unit estimates for a simple random sample of size \( n \).

\( n \) : Sample size.

\( r \) : Units available as response \( (n - r) \) are missing units in the sample.

\( f \) : Sampling friction \( (f = n/N) \)

\( N \) : Population size

### 3.2.2 PROPOSED ESTIMATOR

Using the motivation due to ratio method of imputation, mean method of imputation and Singh and Tailor (2003) procedure we are suggesting some estimators as
\[ y_{1i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_h - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.1) \]

\[ y_{2i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_2 - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.3) \]

\[ y_{3i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_3 - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.5) \]

\[ y_{4i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_4 - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.7) \]

\[ y_{5i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_5 - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.9) \]

\[ y_{6i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_6 - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.11) \]

\[ y_{7i} = \begin{cases} 
  y_i & \text{if } i \in R \\
  \frac{\bar{y}_r}{(1-f_1)} [\phi_7 - f_1] & \text{if } i \in R^C 
\end{cases} \quad \ldots (3.13) \]

Then the proposed point estimates are
\[ \bar{y}_1^* = \bar{y}_r \phi_1 \]  \hspace{1cm} \text{(3.15)}

\[ \bar{y}_2^* = \bar{y}_r \phi_2 \]  \hspace{1cm} \text{(3.16)}

\[ \bar{y}_{SD1}^* = \bar{y}_r \phi_3 \]  \hspace{1cm} \text{(3.17)}

\[ \bar{y}_{SD2}^* = \bar{y}_r \phi_4 \]  \hspace{1cm} \text{(3.18)}

\[ \hat{y}_1 = \bar{y}_r \phi_5 \]  \hspace{1cm} \text{(3.19)}

\[ \hat{y}_2 = \bar{y}_r \phi_6 \]  \hspace{1cm} \text{(3.20)}

\[ \hat{y}_3 = \bar{y}_r \phi_7 \]  \hspace{1cm} \text{(3.21)}

### 3.2.3 PROPERTIES OF PROPOSED ESTIMATOR

**THEOREM 3.1** Up to first order of approximation

**[1]:** The estimator \( \bar{y}_1^* \) could be expressed as:

\[ \bar{y}_1^* = \bar{Y} [1 + e_0 - \psi (e_1 + e_0 e_1)] \]  \hspace{1cm} \text{(3.22)}

**[2]:** Bias of \( \bar{y}_1^* \) is:

\[ E(\bar{y}_1^* - \bar{Y}) = -\bar{Y} \psi M_1 \rho C_0 C_1 \]  \hspace{1cm} \text{(3.23)}

**[3]:** Mean squared error of \( \bar{y}_1^* \) is:

\[ E(\bar{y}_1^* - \bar{Y})^2 = \bar{Y}^2 \left[ M_1 C_0^2 + \psi^2 M_1 C_1^2 - 2 \psi M_1 \rho C_0 C_1 \right] \]  \hspace{1cm} \text{(3.24)}
THEOREM 3.2 Up to first order of approximation

[4]: The estimator $\bar{y}_2^*$ could be expressed as:

$$
\bar{y}_2^* = \bar{Y} \left[ 1 - \psi e_2 + e_0 - \psi e_0 e_2 \right]
$$

... (3.26)

[5]: Bias of $\bar{y}_2^*$ is:

$$
E(\bar{y}_2^* - \bar{Y}) = -\bar{Y} \psi M_2 \rho C_0 C_1
$$

... (3.27)

[6]: Mean squared error $\bar{y}_2^*$ of is:

$$
E(\bar{y}_2^* - \bar{Y})^2 = \bar{Y}^2 \left[ M_1 C_0^2 + \frac{f^2 \theta^2}{(1-f)^2} M_2 C_1^2 - \frac{2f \theta}{(1-f)} M_2 \rho C_0 C_1 \right]
$$

... (3.28)

THEOREM 3.3 Up to first order of approximation

[7]: The estimator $\bar{y}_{SD1}^*$ could be expressed as:

$$
\bar{y}_{SD1}^* = \bar{Y} \left[ 1 + \lambda e_1 - \beta \lambda e_1^2 + e_0 + \lambda e_0 e_1 \right]
$$

... (3.30)

[8]: Bias of $\bar{y}_{SD1}^*$ is:

$$
E\left[\bar{y}_{SD1}^* - \bar{Y}\right] = \bar{Y} \left[ \lambda M_1 \rho C_0 C_1 - \beta \lambda M_1 C_1^2 \right]
$$

... (3.31)

[9]: Mean squared error of $\bar{y}_{SD1}^*$ is:
$$E(\bar{y}_{SD1}^* - \bar{Y})^2 = \bar{Y}^2 M_1 \left[ C_0^2 + \lambda^2 C_1^2 + 2\lambda \rho C_0 C_1 \right] \quad \ldots (3.32)$$

THEOREM 3.4 Up to first order of approximation

[10]: The estimator $\bar{y}_{SD2}^*$ could be expressed as:

$$\bar{y}_{SD2}^* = \bar{Y} \left[ 1 + \lambda e_2 - \beta \lambda e_2^2 + e_0 + \lambda e_0 e_2 \right] \quad \ldots (3.34)$$

[11]: Bias of $\bar{y}_{SD2}^*$ is:

$$E[\bar{y}_{SD2}^* - \bar{Y}] = \bar{Y} \left[ \lambda M_2 \rho C_0 C_1 - \beta \lambda M_2 C_1^2 \right] \quad \ldots (3.35)$$

[12]: Mean squared error of $\bar{y}_{SD2}^*$ is:

$$E(\bar{y}_{SD2}^* - \bar{Y})^2 = \bar{Y}^2 \left[ M_1 C_0^2 + M_2 (\lambda^2 C_1^2 + 2\lambda \rho C_0 C_1) \right] \quad \ldots (3.36)$$

THEOREM 3.5 Up to first order of approximation

[13]: The estimator $\hat{y}_1$ could be expressed as:

$$\hat{y}_1 = \bar{Y} (1 + e_0 - \theta e_1 + \theta^2 e_1^2 - \theta e_0 e_1) \quad \ldots (3.38)$$

[14]: Bias of $\hat{y}_1$ is:

$$E(\hat{y}_1 - \bar{Y}) = \bar{Y} \theta M_1 (\theta C_1^2 - \rho C_0 C_1) \quad \ldots (3.39)$$
[15]: Mean squared error of $\hat{y}_1$ is:

$$E(\hat{y}_1 - \bar{Y})^2 = \bar{Y}^2 M_1 (C_0^2 + \theta^2 C_1^2 - 2\theta \rho C_0 C_1) \quad \ldots (3.40)$$

**THEOREM 3.6** Up to first order of approximation

[16]: The estimator $\hat{y}_2$ could be expressed as:

$$\hat{y}_2 = \bar{Y} (1 + e_0 - \theta e_2 + \theta^2 e_2^2 - \theta e_0 e_2) \quad \ldots (3.42)$$

[17]: Bias of $\hat{y}_2$ is:

$$E(\hat{y}_2 - \bar{Y}) = \bar{Y} \theta M_2 (\theta C_1^2 - \rho C_0 C_1) \quad \ldots (3.43)$$

[18]: Mean squared error of $\hat{y}_2$ is:

$$MSE(\hat{y}_2) = \bar{Y}^2 (M_1 C_0^2 + M_2 (\theta^2 C_1^2 - 2\theta \rho C_0 C_1)) \quad \ldots (3.44)$$

**THEOREM 3.7** Up to first order of approximation

[19]: The estimator $\hat{y}_3$ could be expressed as:

$$\hat{y}_3 = \bar{Y} (1 + e_0 - \theta e_1 + \theta e_2 + \theta^2 e_1^2 - \theta^2 e_1 e_2 - \theta e_0 e_1 + \theta e_0 e_2) \quad \ldots (3.46)$$

[20]: Bias of $\hat{y}_3$ is:

$$E(\hat{y}_3 - \bar{Y}) = \bar{Y} \theta M_3 (\theta C_1^2 - \rho C_0 C_1) \quad \ldots (3.47)$$
Mean squared error of $\hat{y}_3$ is:

$$E(\hat{y}_3 - \bar{Y})^2 = \bar{Y}^2 (M_1 C_0^2 + \theta M_3 (\theta C_1^2 - 2 \rho C_0 C_1))$$

... (3.48)

### 3.2.4 COMPARISON(S)

**LEMMA 3.1** Estimator $\bar{y}_1^*$ will be better than $\bar{y}_2^*$ if $n > r$

$$\left[ \text{MSE}(\bar{y}_1^*) \right]_{\text{Min}} - \left[ \text{MSE}(\bar{y}_2^*) \right]_{\text{Min}} < 0$$

$$\bar{Y}^2 C_0^2 \rho^2 [M_2 - M_1] < 0$$

or $n > r$  
This is always true.

**LEMMA 3.2** $\bar{y}_{SD1}^*$ will be better than $\bar{y}_{SD2}^*$ if $n > r$

$$\left[ \text{MSE}(\bar{y}_{SD1}^*) \right]_{\text{Min}} - \left[ \text{MSE}(\bar{y}_{SD2}^*) \right]_{\text{Min}} < 0$$

$\Rightarrow n > r$  
which is always true.

**LEMMA 3.3** $\text{MSE}(\hat{y}_1) < \text{MSE}(\hat{y}_2)$ if

$\therefore r < n$

which is always true.

Again $\text{MSE}(\hat{y}_1) < \text{MSE}(\hat{y}_3)$ if

$\therefore n < N$
which is always true.

Again $MSE(\hat{\mu}_2) < MSE(\hat{\mu}_3)$ if

$\therefore M_2 - M_3 > 0$

### 3.2.5 ILLUSTRATIVE EXAMPLE

The target in this section is to evaluate the gain in efficiencies (in terms of mean squared error) obtained by the proposed estimators. To see the performance of the various estimators discussed here, we are considering two different population data used earlier by other researchers. The empirical analysis is discussed below.

**POPULATION – I [Sources: Cochran (1977)]**

The required information is given in Table 3.1.

**TABLE 3.1 Population - I Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{Y}$</td>
<td>56.46</td>
<td>$n$</td>
<td>51</td>
<td>$C_0$</td>
<td>1.4177</td>
<td>$r$</td>
<td>45</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>44.5</td>
<td>$N$</td>
<td>257</td>
<td>$C_1$</td>
<td>1.4045</td>
<td>$\rho$</td>
<td>0.8870</td>
</tr>
</tbody>
</table>

**TABLE 3.2 Percent Relative Efficiency of various estimators with respect to mean per unit estimator for Population – I**

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Estimator(s)</th>
<th>% Relative efficiencies with respect to $\bar{y}_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$\bar{y}_r$ (Mean method)</td>
<td>100</td>
</tr>
</tbody>
</table>
2. $\hat{Y}_1 = \bar{y}_r \left( \frac{\bar{X} + \rho}{\bar{x}_r + \rho} \right)$ Singh and Tailor (2003) with NR 453.83

3. $\hat{Y}_2 = \bar{y}_r \left( \frac{\bar{X} + \rho}{\bar{x}_n + \rho} \right)$ Singh and Tailor (2003) with NR 301.62

4. $\hat{Y}_3 = \bar{y}_r \left( \frac{\bar{x}_n + \rho}{\bar{x}_r + \rho} \right)$ Singh and Tailor (2003) with NR 112.510524

5. $\bar{y}_{Rat1} = \bar{y}_r \frac{\bar{X}}{\bar{x}_r}$ Ratio method of imputation 446.46

6. $\bar{y}_{Rat2} = \bar{y}_r \frac{\bar{X}}{\bar{x}_n}$ Ratio method of imputation 298.81

7. $\bar{y}_1^*$

8. $\bar{y}_2^*$

9. $\bar{y}_{SD1}^*$

10. $\bar{y}_{SD2}^*$

11. $(\hat{y}_1)$

12. $(\hat{y}_2)$

13. $(\hat{y}_3)$

---

**Graph:**

- **Mean per unit estimator**
- **Proposed estimator $\bar{Y}_{SD1}^*$**

**Axes:**
- **X-axis:** $K$ values (0.8557, 1.6062, 2.9511, 4.0607)
- **Y-axis:** MSE (0 to 120)
3.3 ESTIMATION IN PRESENCE OF MEASUREMENT ERROR

In sample survey we get the efficiency of the estimators on the basis of collected or simulated data. Data for the analysis may originate from various kind of sampling sources such as simple random sampling, stratified sampling, systematic sampling etc. Various methods of estimation are generally analyzed under the assumption that observations collected are true and without any error. In real life, this kind situations are not tenable. In fact real life data contains observational error due to many reasons like memory failure, over reporting etc. These are also called measurement error [See Cochran (2005) and Sukhatme et al (1984)].
There may be two possibilities in survey, either incorrect response or non-response. Several methods available in the literature of survey sampling to handle non-response but if the informer provides the incorrect information then additional methodologies are required. There may be some observational error due to surveyor also.


This Chapter presents an estimation strategy under measurement error model in light of Shalabh (1997), Manisha and Singh (2001). The utilization of dual to ratio estimator by Srivenkatramana and Tracy (1981) is the main focus in the content.
3.3.1 NOTATIONS FOR THE STUDY

Assume given a set of information obtained through simple random sampling procedure on two characteristics $X$ and $Y$. Suppose $(x_i, y_i)$ be the pair of observational values and $(X_i, Y_i)$ are corresponding true values on the characteristic $(X, Y)$ respectively.

For the $i^{th}$ unit ($i=1, 2...n$) in the same sample suppose the measurement error are as follow.

$$U_i = (y_i - Y_i) \quad \ldots \quad (3.50)$$

$$V_i = (x_i - X_i) \quad \ldots \quad (3.51)$$

Notations for the study are:

$\bar{Y}, \bar{X}$ : Population Parameters

$\bar{y}$ and $\bar{x}$ : Mean per unit estimates for a simple random sample of size $n$.

$n$ : Sample size

$N$ : Population size

$\sigma^2_U$ and $\sigma^2_V$ : Variances for measurement error of $Y$ and $X$ respectively

$\sigma^2_Y$ and $\sigma^2_X$ : Variances of variable $Y$ and $X$ respectively

$\rho$ : Correlation coefficient between variables

$f$ : Sampling friction
\[ C_Y = \frac{\sigma_Y}{Y} \quad : \text{Coefficient of variation for variable } Y \]

\[ C_X = \frac{\sigma_X}{X} \quad : \text{Coefficient of variation for variable } X \]

Let these measurement error be stochastic in nature and are uncorrelated \([i.e. \rho(U_i, V_i) = 0]\) along with sum of measurement error is zero and variances are \(\sigma_u^2\) and \(\sigma_v^2\) respectively. The population means are \(\bar{X}\) and \(\bar{Y}\) for the true values with population variances are \(\sigma_X^2\) and \(\sigma_Y^2\) respectively. The \(\bar{Y}\) is unknown and mean of auxiliary information \(\bar{X}\) is known which is used as a source to estimate \(\bar{Y}\). From (3.50), we write

\[ \bar{y} = \frac{1}{n} \sum_i (U_i + Y_i) \quad \ldots (3.52) \]

### 3.3.2 SOME EXISTING ESTIMATORS

**1) MEAN PER UNIT ESTIMATOR**

Mean per unit (or Mean) estimator is a well known estimator in the literature and in setup of measurement error \(\bar{y} = n^{-1} \sum_i [U_i + Y_i]\) as shown in (3.52).

Bias for \(\bar{y}\) is zero \(i.e. \) \(E(\bar{y}) = \left[ \frac{1}{n} \sum_i (U_i + Y_i) \right] = (\bar{Y}) \quad \ldots (3.53a)\)

\[ Variance (\bar{y}) = \frac{\sigma_y^2}{n} \left[ 1 + \frac{\sigma_u^2}{\sigma_y^2} \right] \quad \ldots (3.53b) \]
For the estimation of $\bar{Y}$, we use the sample statistic $\bar{y}$ which provides an unbiased estimator. In mean per unit estimator $\bar{y}$ we are not using any additional information. To use the auxiliary $X$ characteristic, one can adopt several ways.

(2) SHALABH (1997) ESTIMATOR

Shalabh (1997) proposed estimator

$$t_R = \frac{\bar{y}}{X} \mu_X$$

Bias of $t_R$, $B(t_R) = \frac{\mu_Y}{n} \left[ C_X (C_X - \rho C_Y) + \frac{\sigma_Y^2}{\mu_X^2} \right]$ ... (3.54a)

MSE of $t_R$, $MSE(t_R) = \frac{\sigma_Y^2}{n} \left[ 1 - \frac{C_X}{C_Y} \left( 2 \rho - \frac{C_X}{C_Y} \right) \right] + \frac{1}{n} \left[ \sigma_u^2 + \left( \frac{\mu_Y}{\mu_X} \right)^2 \sigma_v^2 \right]$ ... (3.54b)

where $\mu_X$ denotes the population mean of $X$.

(3) MANISHA AND SINGH (2001) ESTIMATOR

Manisha and Singh (2001) proposed estimator

$$\tilde{y}_\theta = \theta t_R + (1-\theta) \bar{y}$$

Bias of $\tilde{y}_\theta$, $B(\tilde{y}_\theta) = \theta \left\{ \frac{\mu_Y}{n \mu_X^2} \left( \frac{\sigma_X^2}{\mu_X^2} + \sigma_v^2 \right) - \frac{1}{n \mu_X} \rho \sigma_X \sigma_Y \right\}$ ... (3.55a)
MSE of $\bar{y}_\theta, M(\bar{y}_\theta) = \frac{\sigma_Y^2}{n} \left[ 1 - \theta \frac{C_X}{C_Y} \left( 2\rho - \theta \frac{C_X}{C_Y} \right) \right] + \frac{1}{n} \left[ \theta^2 \frac{\mu_Y^2}{\mu_X^2} \sigma_Y^2 + \sigma_U^2 \right] \quad \ldots (3.55b)

where $\theta$ is characterizing scalar, $U$ and $V$ are measurement errors corresponding to $Y$ and $X$ respectively.

### 3.3.3 PROPOSED ESTIMATOR

The estimator of Shalabh (1997) is ratio type whereas estimator of Manisha and Singh (2001) is a linear combination of ratio type and sample mean estimator.

Singh and Shukla (1987) have suggested using Factor-type estimator for estimating $\bar{Y}$ which incorporates the properties of ratio estimator, product estimator, Dual to ratio estimator and usual mean estimator depending upon the suitable choice of a parameter. A novel property it possesses is optimization of mean squared error at the multiple values of the estimator parameter. Another is the existence of almost unbiasedness property on the multiple choices reducing the mean squared error. The application of Factor-type (F-T) estimator is further encouraged due to Singh and Shukla (1991), Singh et al. (1993), Shukla (2002). The estimator is effectively utilized in imputation problem by Shukla and Thakur (2008), Shukla et al. (2009a & b), et al. (2011) etc. the measurement error extension of the same is the objective of this paper.

The basic structure of Factor-type (F-T) estimator, as suggested by Singh and Shukla (1987) is
\[
\bar{y}_{F-T} = \bar{y} \left[ \frac{(A + C)\bar{X} + f\bar{B}}{(A + fB)\bar{X} + C\bar{x}} \right]
\]

where \( A = (K-1) (K-2), B = (K-1) (K-4), C = (K-2) (K-3) (K-4); 0 < f < 1; 0 < K < \infty \)

In light of Manisha and Singh (2001), we suggest to replace the ratio estimator by dual
to ratio estimator and factor-type estimator. F-T is used with the sound logic that (a) F-T
estimator contains ratio estimator for \( K=1 \). (b) \( K \) is already a parameter which could be used in
linear combination also (c) F-T estimator contains some good properties not covered in ratio
estimator. A modified ratio estimator is also been used with some constant scalar.

We propose estimators as follows

[A]: \( \bar{y}_D = \left[ K\bar{y}^* + (1 - K)\bar{y} \right] \) \hspace{1cm} \ldots (3.56)

where \( \bar{y}^* = \bar{y} \frac{N\bar{X} - n\bar{x}}{(N - n)\bar{X}} \) \hspace{1cm} \ldots (3.56a)

[B]: \( \bar{y}_{FTME} = K\bar{y}_{F-T} + (1 - K)\bar{y} \) \hspace{1cm} \ldots (3.57)

where \( A = (K-1) (K-2) \)
\( B = (K-1) (K-4) \)
\( C = (K-2) (K-3) (K-4) \)
\( \bar{y}_{F-T} = \bar{y} \left[ \frac{(A + C)\bar{X} + f\bar{B}\bar{x}}{(A + fB)\bar{X} + C\bar{x}} \right] \) \hspace{1cm} \ldots (3.57a)

[C]: \( \hat{y} = \bar{y} \frac{\bar{X} + K}{\bar{x} + K} \) \hspace{1cm} \ldots (3.58)
where $K (\geq 0)$ is suitably chosen constant to make the proposed estimator more efficient.

### 3.3.4 PROPERTIES OF PROPOSED ESTIMATOR

**THEOREM 3.8** Up to first order of approximation

[22]: The estimator $(\bar{y}_D)$ could be expressed as:

$$
\bar{y}_D = \bar{Y} + \delta_0 \frac{\bar{Y}}{X} \delta_1 - \frac{\psi}{X} \delta_0 \delta_1 
$$ ... (3.60)

[23]: Bias of $(\bar{y}_D)$ is:

$$
Bias[\bar{y}_D] = -\frac{\bar{Y}}{n} \rho C_Y C_X 
$$ ... (3.61)

[24]: Mean squared error of $(\bar{y}_D)$ is:

$$
E[\bar{y}_D - \bar{Y}]^2 = \frac{1}{n} \left[ (\sigma_Y^2 + \sigma_U^2) + \bar{Y}^2 \psi^2 C_X^2 \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) - 2\bar{Y}^2 \psi \rho C_Y C_X \right] 
$$ ... (3.62)

**NOTE**

The mean squared error of $(\bar{y}_D)$ in equation (3.62) is a function with unknown $K$, so it is practical to get an optimum value of $K$ in such a way that the mean squared error of the resultant proposed estimator becomes minimum.
Differentiating (3.62) with respect to $\psi$ and equating to zero (assuming $K \neq 0$), we have

$$\psi = \left[ \rho \frac{C_Y}{C_X} - \frac{\sigma_X^2}{\sigma_Y^2 + \sigma_X^2} \right]$$ \hspace{1cm} \text{... (3.63)}

This value of $\psi$ given in (3.63) provides a minimum MSE to proposed estimator.(Second partial derivative is positive). Using this minimum mean squared error is

$$MSE(\bar{y}_D) = \frac{1}{n} \left( \frac{\sigma_Y^2 + \sigma_U^2}{n(\sigma_X^2 + \sigma_Y^2)} \right)$$ \hspace{1cm} \text{... (3.64)}

**THEOREM 3.9** Up to first order of approximation

[25]: The estimator ($\bar{y}_{FTME}$) could be expressed as:

$$\bar{y}_{FTME} = \bar{y} + \delta_0 + \frac{P\bar{y}}{X} \delta_1 - \frac{\rho P\bar{y}^2}{X^2} \delta_1^2 + \frac{P}{X} \delta_0 \delta_1$$ \hspace{1cm} \text{... (3.65)}

[26]: Bias of ($\bar{y}_{FTME}$) is:

$$E[\bar{y}_{FTME} - \bar{y}] = \frac{P}{n} \left[ \rho C_Y C_X - \beta C_X \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) \right]$$ \hspace{1cm} \text{... (3.66)}

[27]: Mean squared error of ($\bar{y}_{FTME}$) is:

$$E[\bar{y}_{FTME} - \bar{y}]^2 = \frac{1}{n} \left( \sigma_Y^2 + \sigma_U^2 \right) + \frac{P\bar{y}^2}{n} \left[ PC_{C_X} \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) + 2 \rho C_Y C_X \right]$$ \hspace{1cm} \text{... (3.67)}

**NOTE 3.9** (Bias Reduction and Minimum Mean Squared Error)
Bias for the proposed estimator is given

\[ P = 0 \text{ or } \beta = \frac{\rho C_Y \sigma^2_Y}{C_X \left( \sigma^2_X + \sigma^2_Y \right)} = \phi \quad \ldots (3.68) \]

At this value of \( \beta \) the proposed estimator will convert into an unbiased estimator.

The mean squared error of \( \bar{y}_{FTME} \) is a function with unknown parameter \( P \), whereas \( P \) is a function of \( K \) solely. So it is practical to get an optimum value of \( K \) in such a way that the mean squared error of the resultant proposed estimator becomes minimum.

Differentiating \( MSE(\bar{y}_{FTME}) \) with respect to \( P \) and equating to zero (assuming \( P \neq 0 \)), we have

\[ P = -\phi = \Delta \text{(say)} \quad \ldots (3.69) \]

This value of \( P \) provides a minimum MSE to proposed estimator. (Second derivative is + ve). Minimum mean squared error is will be as

\[
MSE \left( \bar{y}_{FTME} \right) \bigg|_{\text{Minimum}} = \frac{1}{n} \left( \sigma^2_Y + \sigma^2_{e_1} \right) - \frac{\left( \rho \sigma_Y \sigma_X \right)^2}{n \left( \sigma^2_X + \sigma^2_Y \right)} \quad \ldots (3.70)
\]

**THEOREM 3.10** Up to first order of approximation

[28]: The estimator \( \hat{y} \) could be expressed as:

\[
\hat{y} = \bar{Y} + \delta_0 - \xi \delta_1 + \frac{\xi}{(\bar{X} + K)} \delta_1^2 - \frac{1}{\bar{X} + K} \delta_0 \delta_1 \quad \ldots (3.71)
\]

[29]: Bias of \( \hat{y} \) is:
\[ E[\hat{y} - \bar{Y}] = \frac{1}{n(\bar{X} + K)} \left[ \tilde{\xi}(\sigma^2_Y + \sigma^2_U) - \rho \sigma_Y \sigma_X \right] \] ... (3.72)

[30]: Mean squared error of \( \hat{y} \) is:

\[ E(\hat{y} - \bar{Y})^2 = \frac{1}{n} (\sigma^2_Y + \sigma^2_U) + \frac{\xi^2}{n} (\sigma^2_X + \sigma^2_Y) - \frac{2 \xi}{n} \rho \sigma_Y \sigma_X \] ... (3.73)

NOTE 3.10

\[ \text{MSE}(\hat{y})_{\text{Min}} = \frac{1}{n} \left[ (\sigma^2_Y + \sigma^2_U) - \left( \frac{\rho \sigma_Y \sigma_X}{(\sigma^2_X + \sigma^2_Y)} \right)^2 \right] \] ... (3.74)

3.3.5 COMPARISON(S)

In this section, we compare the proposed estimator \( \bar{y}_D \) with other existing estimators under measurement error. On comparison of mean squared error of proposed estimator given in (3.62) with Manisha and Singh (2001) is given in (3.55b), we found

\[ \frac{1}{n} \left[ (\sigma_Y^2 + \sigma_U^2) + \bar{Y}^2 \psi^2 C_X \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) - 2 \bar{Y}^2 \psi \rho C_Y C_X \right] < \\
\frac{\sigma_Y^2}{n} \left[ 1 - \theta \left( \frac{C_X}{C_Y} \left( 2 \rho - \theta \frac{C_X}{C_Y} \right) \right) + \frac{1}{n} \left[ \theta^2 \frac{\mu_Y^2 - \sigma_Y^2 + \sigma_U^2}{\mu_X^2 - \sigma_X^2} \right] \right] \]

Thus \( \rho > \frac{\alpha}{2} \frac{1}{(1-f)} \frac{C_X}{C_Y} \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) \) when \( \bar{X} \) and \( \bar{Y} \) have the same signs.

and \( \rho < -\frac{\alpha}{2} \frac{1}{(1-f)} \frac{C_X}{C_Y} \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) \) when \( \bar{X} \) and \( \bar{Y} \) have the opposite signs.
These are efficiency conditions with the assumption that characterizing scalars for both mentioned estimators are same and denoted by \( \alpha \) to pass over the mathematical complexity. In particular when \( C_X \) and \( C_Y \) are identical in magnitudes, then the efficiency conditions will be as follows

\[
\rho > \frac{\alpha}{2} \left(1 - f\right) \left[1 + \frac{\sigma_V^2}{\sigma_X^2}\right] \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the same signs}
\]

and

\[
\rho < - \frac{\alpha}{2} \left(1 - f\right) \left[1 + \frac{\sigma_V^2}{\sigma_X^2}\right] \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the opposite signs}
\]

To compare with Shalabh (1997) shown in (3.54b), we see that the proposed estimator will be better if (3.62) < (3.54b)

\[
\rho > \frac{1}{2} \frac{n^2 K^2 - N^2 + 2nN}{n(K+1) - N(N-n)} C_X \left\{1 + \frac{\sigma_V^2}{\sigma_X^2}\right\} \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the same signs.}
\]

\[
\rho < -\frac{1}{2} \frac{n^2 K^2 - N^2 + 2nN}{n(K+1) - N(N-n)} C_X \left\{1 + \frac{\sigma_V^2}{\sigma_X^2}\right\} \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the opposite signs.}
\]

For \( \psi = 0 \) the reduction becomes as follows

\[
\rho > \frac{1}{2} \left[1 + \frac{\sigma_V^2}{\sigma_X^2}\right] \frac{C_X}{C_Y} \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the same signs.}
\]

\[
\rho < -\frac{1}{2} \left[1 + \frac{\sigma_V^2}{\sigma_X^2}\right] \frac{C_X}{C_Y} \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the opposite signs.}
\]
As we know the mean squared error of $\overline{y}$ is same as variance of $\overline{y}$ given by (3.53b).

Hence from (3.62) and (3.53b), $\text{MSE}(\overline{y}_D) < \text{MSE}(\overline{y})$ if

$$\rho > \frac{1}{2} \frac{fK}{1-f} \frac{C_X}{C_Y} \left[1 + \frac{\sigma_Y^2}{\sigma_X^2}\right]$$

when $X$ and $Y$ have the same signs.

$$\rho < \frac{1}{2} \frac{fK}{1-f} \frac{C_X}{C_Y} \left[1 + \frac{\sigma_Y^2}{\sigma_X^2}\right]$$

when $X$ and $Y$ have the opposite signs.

Now compare the proposed estimator with other existing estimators under measurement error. The proposed estimator will be better as compare to other estimators if this provides minimum MSE. (i.e. $(\text{MSE})_{\text{FMTE}} < (\text{MSE})_{\text{Manisha&Singh}}$)

$$\frac{1}{n}(\sigma_Y^2 + \sigma_U^2) + \frac{P\overline{y}^2}{n} + PC_X^2 \left\{1 + \frac{\sigma_Y^2}{\sigma_X^2}\right\} + 2\rho C_Y C_X < \frac{\sigma_Y^2}{n} + \frac{\sigma_U^2}{n} - 2\theta \rho \frac{\sigma_Y^2}{n} \frac{C_X}{C_Y} + \theta^2 \frac{C_X^2}{C_Y} \left[\frac{n}{\overline{y}^2} \frac{\sigma_Y^2}{n} + \frac{n}{\overline{y}^2} \frac{\sigma_U^2}{n} \frac{C_X^2}{C_Y} \right]$$

Where $\theta$ is characterizing scalar and $0 \leq \theta \leq 1$. For $\theta = 1$, Manisha and Singh (2001) estimator provide the estimator proposed by Shalabh (1997). Then

$$\rho > \frac{(\theta - P)}{2} \frac{C_X}{C_Y} \left[1 + \frac{\sigma_Y^2}{\sigma_X^2}\right]$$

when $X$ and $Y$ have the same sign.

and

$$\rho < \frac{(\theta - P)}{2} \frac{C_X}{C_Y} \left[1 + \frac{\sigma_Y^2}{\sigma_X^2}\right]$$

when $X$ and $Y$ have the opposite sign.
For $\theta = 1$ all above expressions provides the information regarding to Shalabh (1997) estimator. Proposed estimator will be more effective when

$$(MSE)_{FTME} < (MSE)_{Shalabh(97)}$$

$$\frac{1}{n} (\sigma_y^2 + \sigma_u^2) + \frac{P \bar{Y}^2}{n} \left[ PC_X^2 \left\{ 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right\} + 2\rho C_Y C_X \right] <$$

$$\left[ \frac{\sigma_y^2}{n} + \frac{\sigma_u^2}{n} - 2 \rho \frac{\sigma_Y^2}{n} \frac{C_X}{C_Y} + \frac{C_X^2 \sigma_Y^2}{C_Y^2} \frac{n}{n} + \frac{\sigma_Y^2 \bar{Y}^2}{n} \right]$$

$$\rho > \frac{(1-P) C_X}{2 \ C_Y} \left[ 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right]$$

when $\bar{X}$ and $\bar{Y}$ have the same sign.

And $\rho < -\frac{(1-P) C_X}{2 \ C_Y} \left[ 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right]$ when $\bar{X}$ and $\bar{Y}$ have the opposite sign.

Further the mean squared error of $\bar{y}$ is same as variance of $\bar{y}$. Hence from (3.67) and (3.73b), we see that $MSE(\bar{y}_{FTME}) < Variance(\bar{y})$ if

$$\frac{1}{n} (\sigma_y^2 + \sigma_u^2) + \frac{P \bar{Y}^2}{n} \left[ PC_X^2 \left\{ 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right\} + 2\rho C_Y C_X \right] < \frac{\sigma_y^2}{n} \left\{ 1 + \frac{\sigma_U^2}{\sigma_Y^2} \right\}$$

$$\rho > \frac{P \ C_X}{2 \ C_Y} \left[ 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right]$$

when $\bar{X}$ and $\bar{Y}$ have the same sign.

$$\rho < -\frac{P \ C_X}{2 \ C_Y} \left[ 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right]$$

when $\bar{X}$ and $\bar{Y}$ have the opposite sign.

Similarly while comparison of $\hat{y}$ with $\bar{y}$
\[
\frac{1}{n} \left( \sigma_y^2 + \sigma_U^2 \right) + \frac{\bar{z}}{n} \left( \sigma_X^2 + \sigma_V^2 \right) - \frac{2\bar{z}}{n} \rho \sigma_Y \sigma_X < \frac{1}{n} \left( \sigma_y^2 + \sigma_U^2 \right)
\]

\[\rho > \frac{1}{2} \frac{\sigma_X}{\sigma_Y} \left[ 1 + \frac{\sigma_V^2}{\sigma_X^2} \right] \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the same sign.}\]

\[\text{and } \rho < -\frac{1}{2} \frac{\sigma_X}{\sigma_Y} \left[ 1 + \frac{\sigma_V^2}{\sigma_X^2} \right] \quad \text{when } \bar{X} \text{ and } \bar{Y} \text{ have the opposite sign.}\]

### 3.3.6 NUMERICAL ILLUSTRATION

Population of size 250 (N=250) is given in appendix-1. The population parameters are displayed in Table 3.3.

**TABLE -3.3 Population Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\bar{Y})</td>
<td>63.396</td>
<td>(n)</td>
<td>50</td>
<td>(C_0)</td>
<td>0.2899</td>
<td>(\rho_{01})</td>
<td>0.8544</td>
</tr>
<tr>
<td>(\bar{X}_1)</td>
<td>48.136</td>
<td>(\sigma_y)</td>
<td>18.3819</td>
<td>(C_1)</td>
<td>0.4637</td>
<td>(\rho_{02})</td>
<td>0.8249</td>
</tr>
<tr>
<td>(\bar{X}_2)</td>
<td>56.364</td>
<td>(\sigma_x)</td>
<td>23.03</td>
<td>(C_2)</td>
<td>0.4085</td>
<td>(\rho_{12})</td>
<td>0.8289</td>
</tr>
</tbody>
</table>

**TABLE -3.4 Mean Squared Error**

<table>
<thead>
<tr>
<th>Estimator</th>
<th>% RE (in terms of MSE) with respect to mean per unit estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{With var } (y, x_1))</td>
<td>(\text{With var } (y, x_2))</td>
</tr>
<tr>
<td></td>
<td>(\bar{y})</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Shalab (1997)</td>
<td>36.29266</td>
</tr>
<tr>
<td>Manisha and Singh (2001)</td>
<td>100 ((\theta = 0))</td>
</tr>
<tr>
<td>(\bar{y}_D) at ((\psi = 0.1694))</td>
<td>36.29266 ((\theta = 1))</td>
</tr>
<tr>
<td>(\bar{y}_D) at ((\psi = 0.1802))</td>
<td>108.2484</td>
</tr>
<tr>
<td>(\bar{y}_{FTME}) at ((P = -0.1694))</td>
<td>108.2484</td>
</tr>
<tr>
<td>(\bar{y}_{FTME}) at ((P = -0.1802))</td>
<td>108.2126</td>
</tr>
<tr>
<td>(\hat{y})</td>
<td>108.2484</td>
</tr>
</tbody>
</table>

\(\xi = 0.22321\) and \(0.20268\)

**Fig. 3.4** MSE comparison at opt. \(K\) for \(\bar{y}_{FTME}\) with \(X_1\) variable
3.4 ESTIMATION IN PRESENCE OF MISSING DATA AND MEASUREMENT ERROR TOGETHER

A general problem in sampling surveys which encountered by investigators is the problem of non-response. Non-response is an error occurs due to unsuccessful attempt to obtain the desired information from an eligible individual. To solve the problem of non-response the method of imputation is usual method to cope up this.

Kreuter et al. (2010) has a unique opportunity to examine non-response and measurement error jointly and provide estimates for bias, variance, and mean square error by linking data from a survey on labor market participation to administrative data. They found that increased contact attempts resulted in significant reductions in non-response bias. An exilient discussion on non-response error, measurement error, and mode of data collection:
tradeoffs in a multi-mode survey of sensitive and non-sensitive item is made by Sakshaug et al. (2010).

### 3.4.1 PROPOSED ESTIMATORS

\[
\hat{y}_1^* = K \bar{y}_r \frac{N\bar{X} - n\bar{x}_r}{(N-n)\bar{X}} + (1-K)\bar{y}_r \quad \ldots (3.75)
\]

\[
\hat{y}_2^* = K \bar{y}_r \frac{N\bar{X} - n\bar{x}_n}{(N-n)\bar{X}} + (1-K)\bar{y}_r \quad \ldots (3.76)
\]

\[
\bar{y}_{SD1} = [K, \bar{y}_{FT1} + (1-K)\bar{y}_r] \quad \ldots (3.77)
\]

\[
\bar{y}_{SD2} = [K, \bar{y}_{FT2} + (1-K)\bar{y}_r] \quad \ldots (3.78)
\]

\[
\hat{y}_1^* = \bar{y}_r \frac{\bar{X} + K}{\bar{x}_r + K} \quad \ldots (3.79)
\]

\[
\hat{y}_2^* = \bar{y}_r \frac{\bar{X} + K}{\bar{x}_n + K} \quad \ldots (3.80)
\]

### 3.4.2 PROPERTIES OF PROPOSED ESTIMATOR

**THEOREM 3.11** Up to first order of approximation

[31]: The estimator could be expressed as:

\[
\bar{y}_1^* = \bar{Y} + \delta_0^j - \frac{\bar{Y}}{\bar{X}} \psi' \delta_1^j - \frac{1}{\bar{X}} \psi' \delta_0^j \delta_1^j \quad \ldots (3.81)
\]

[32]: Bias of is:
\[
E(\bar{y}_1^* - \bar{Y}) = -\frac{\bar{Y} \psi'}{r} \rho C_Y C_X
\] ...

(3.82)

[33]: Mean squared error of is:

\[
E(\bar{y}_1^* - \bar{Y})^2 = \frac{1}{r} \left( \sigma^2_Y + \sigma^2_U \right) + \frac{\bar{Y}^2}{r} \psi'^2 C_Y^2 \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) - \frac{2\bar{Y}^2}{r} \psi' \rho C_Y C_X
\] ...

(3.83)

THEOREM 3.12 Up to first order of approximation

[34]: The estimator could be expressed as:

\[
\bar{y}_2^* = \bar{Y} + \delta_0' - \frac{\bar{Y}}{X} \psi' \delta_2' - \frac{1}{X} \psi' \delta_0' \delta_2'
\] ...

(3.84)

[35]: Bias of is:

\[
E(\bar{y}_2^* - \bar{Y}) = -\frac{\bar{Y}}{\sqrt{nr}} \psi' \rho C_Y C_X
\] ...

(3.85)

[36]: Mean squared error of is:

\[
E(\bar{y}_2^* - \bar{Y})^2 = \frac{1}{r} \left( \sigma^2_Y + \sigma^2_U \right) + \frac{\bar{Y}^2}{X} \frac{1}{n} \psi'^2 \left( \sigma^2_X + \sigma^2_Y \right) - \frac{1}{\sqrt{m} X} \frac{2\bar{Y}^2}{X} \psi' \rho \sigma_Y \sigma_X
\] ...

(3.86)

THEOREM 3.13 Up to first order of approximation

[37]: The estimator \( \bar{y}_{SDI} \) could be expressed as:

\[
\bar{y}_{SDI} = \bar{Y} + \delta_0' + \frac{\bar{Y} \beta P_1}{X} \delta_1' - \frac{\bar{Y} \beta P_1}{X^2} \delta_1'^2 + \frac{P_1}{X} \delta_0' \delta_1'
\] ...

(3.87)
[38]: Bias of \( \bar{Y}_{SDI} \) is:

\[
E[\bar{Y}_{SDI} - \bar{Y}] = \frac{\bar{Y}_P}{r} \left[ \rho C_Y C_X - \beta C_X^2 \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) \right]
\]  

... (3.88)

[39]: Mean squared error of \( \bar{Y}_{SDI} \) is:

\[
E[\bar{Y}_{SDI} - \bar{Y}]^2 = \frac{1}{r} (\sigma_Y^2 + \sigma_Y^2) + \frac{\bar{Y}_P^2}{r} P_1 C_X \left[ P_1 C_X \left( 1 + \frac{\sigma_Y^2}{\sigma_X^2} \right) + 2 \rho \sigma_Y \right]
\]  

... (3.89)

NOTE

The mean squared error of \( \bar{Y}_{SDI} \) is a function with unknown \( K \), so it is practical to get an optimum value of \( K \) in such a way that the mean squared error of the resultant proposed estimator becomes minimum.

Differentiating \( MSE(\bar{Y}_{SDI}) \) with respect to \( P_1 \) and equating to zero (assuming \( P_1 \neq 0 \)), we have

\[
\frac{\partial}{\partial P_1} MSE(\bar{Y}_{SDI}) = 0
\]

\[
\Rightarrow \hat{P}_1 = -\beta = \Delta_1 \text{ (Say)}
\]  

... (3.90)

This value of \( P_1 \) given above provides a minimum MSE to proposed \( \bar{Y}_{SDI} \) estimator.

THEOREM 3.14 Up to first order of approximation
The estimator $\bar{y}_{SD2}$ could be expressed as:

$$
\bar{y}_{SD2} = \bar{Y} + \delta'_0 + \bar{Y}P_2 \frac{1}{\bar{X}} \delta'_2 - \bar{Y}P_2 \frac{1}{\bar{X}^2} \delta'^2_2 + P_2 \frac{1}{\bar{X}} \delta'_0 \delta'_2
$$

... (3.92)

Bias of $\bar{y}_{SD2}$ is:

$$
E[\bar{y}_{SD2} - \bar{Y}] = \bar{Y}P_2 \frac{1}{\sqrt{n}} \rho C_Y C_X - \beta' \frac{1}{\sqrt{n}} C_X^2 \left\{1 + \frac{\sigma_Y^2}{\sigma_X^2}\right\}
$$

... (3.93)

Mean squared error of $\bar{y}_{SD2}$ is:

$$
E[\bar{y}_{SD2} - \bar{Y}]^2 = \frac{1}{r} (\sigma_Y^2 + \sigma_U^2) + \bar{Y}^2 P_2 \frac{1}{\sqrt{n}} C_X \left[P_2 \frac{1}{\sqrt{n}} C_X \left\{1 + \frac{\sigma_Y^2}{\sigma_X^2}\right\} + 2 \frac{1}{\sqrt{r}} \rho C_Y \right]
$$

... (3.94)

**THEOREM 3.15** Up to first order approximation

The estimator $\hat{y}_1^*$ could be expressed as:

$$
\hat{y}_1^* = \bar{Y} + \delta'_0 - \frac{\bar{Y}}{\bar{X} + K} \delta' + \frac{\bar{Y}}{(\bar{X} + K)^2} \delta'^2_1 - \frac{1}{\bar{X} + K} \delta'_0 \delta'_1
$$

... (3.95)

Bias of $\hat{y}_1^*$ is:

$$
E(\hat{y}_1^* - \bar{Y}) = \frac{\bar{Y}}{(\bar{X} + K)^2} \frac{1}{r} (\sigma_Y^2 + \sigma_U^2) - \frac{1}{\bar{X} + K} \frac{1}{r} \rho \sigma_Y \sigma_X
$$

... (3.96)

Mean squared error of $\hat{y}_1^*$ is:
\[ E(\hat{y}_1^\bullet - \bar{Y})^2 = \frac{1}{r} \left( \sigma_Y^2 + \sigma_U^2 \right) + \frac{\bar{z}_1^2}{r} \left( \sigma_X^2 + \sigma_Y^2 \right) - \frac{2\bar{z}_1}{r} \rho \sigma_y \sigma_X \]  \quad (3.97)

**THEOREM 3.16** Up to first order approximation

[46]: The estimator \( \hat{y}_2^\bullet \) could be expressed as:

\[ \hat{y}_2^\bullet = \bar{Y} + \delta_0' - \frac{\bar{Y}}{\bar{X} + K} \delta_2' + \frac{\bar{Y}}{(\bar{X} + K)^2} \delta_2'^2 - \frac{1}{\bar{X} + K} \delta_0' \delta_2' \]  \quad (3.99)

[47]: Bias of \( \hat{y}_2^\bullet \) is:

\[ E(\hat{y}_2^\bullet - \bar{Y}) = \frac{\bar{Y}}{(\bar{X} + K)^2} \frac{1}{n} \left( \sigma_Y^2 + \sigma_U^2 \right) - \frac{1}{\bar{X} + K} \frac{1}{r \sqrt{nr}} \rho \sigma_y \sigma_X \]  \quad (3.100)

[48]: Mean squared error of \( \hat{y}_2^\bullet \) is:

\[ E(\hat{y}_2^\bullet - \bar{Y})^2 = \frac{1}{r} \left( \sigma_Y^2 + \sigma_U^2 \right) + \frac{\bar{z}_2^2}{n} \left( \sigma_X^2 + \sigma_Y^2 \right) - \frac{2\bar{z}_2}{\sqrt{nr}} \rho \sigma_y \sigma_X \]  \quad (3.101)

**3.4.3 NUMERICAL EXAMPLE**

For numerical illustration considering the population of appendix-1 as shown in the section 3.3.6. The population parameters are displayed in Table 3.3.
TABLE 3.6 %RE of $\hat{y}^*_1$ and $\hat{y}^*_2$ with respect to mean per unit estimator

<table>
<thead>
<tr>
<th>At opt $\xi_1$ and $\xi_2$</th>
<th>PRE with respect to mean per unit estimator for $\hat{y}^*_1$</th>
<th>PRE with respect to mean per unit estimator for $\hat{y}^*_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>With $(Y$ and $X_1)$</td>
<td>108.256446</td>
<td>108.2329134</td>
</tr>
<tr>
<td>variables</td>
<td></td>
<td>107.3888</td>
</tr>
<tr>
<td>With $(Y$ and $X_2)$</td>
<td>107.4098</td>
<td>107.3888</td>
</tr>
</tbody>
</table>

TABLE 3.7 %RE of $\bar{y}^*_1$ and $\bar{y}^*_2$ with respect to mean per unit estimator

<table>
<thead>
<tr>
<th>Opt K values</th>
<th>PRE with respect to mean per unit estimator for $\bar{y}^*_1$</th>
<th>PRE with respect to mean per unit estimator for $\bar{y}^*_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With $(Y$ and $X_1)$</td>
<td>With $(Y$ and $X_2)$</td>
</tr>
<tr>
<td></td>
<td>variables</td>
<td>variables</td>
</tr>
<tr>
<td>0.6779</td>
<td>108.2565</td>
<td>107.381614</td>
</tr>
<tr>
<td></td>
<td></td>
<td>108.2329</td>
</tr>
<tr>
<td>0.7466</td>
<td>108.1648</td>
<td>107.3995967</td>
</tr>
<tr>
<td></td>
<td></td>
<td>108.2385</td>
</tr>
<tr>
<td>0.7146</td>
<td>108.2303</td>
<td>107.4091963</td>
</tr>
<tr>
<td></td>
<td></td>
<td>108.2565</td>
</tr>
<tr>
<td>0.787</td>
<td>108.0256</td>
<td>107.3427357</td>
</tr>
<tr>
<td></td>
<td></td>
<td>108.1648</td>
</tr>
</tbody>
</table>

Whereas the gain by ratio estimator is **36.29** and **42.39** respectively first and second auxiliary variable.
3.5 DISCUSSION AND CONCLUSIONS

Many estimators are compared with mean per unit estimator (sample mean estimator) $\bar{y}$, in the setup of imputation for missing data problem. The relative efficiencies of $\bar{y}_1^*$, $\bar{y}_{SD1}^*$ and $\hat{y}_1$ are highest than all other similar estimators. All these three are the suggested estimators at their optimal level. The suggested estimator $\hat{Y}_1$ of Singh and Tailor (2003) is next better to these in terms of efficiency parameter with imputation setup. The other variants of Singh and Tailor (2003) are not good enough in performance. One similar estimator $\bar{Y}_{Rat1}$ looks a good choice than many other estimators but lesser efficient to suggested estimators. Some more suggested estimators $\bar{y}_2^*$, $\bar{y}_{SD2}^*$, $\hat{y}_2$ and $\hat{y}_3$ do not have impressive performance.

We conclude that suggested estimators $\bar{y}_1^*$, $\bar{y}_{SD1}^*$ and $\hat{y}_1$ are efficient in the setup of imputation for estimation of population mean.

While looking to the setup of measurement error the content of this chapter has suggested three different estimators as $\bar{Y}_D$, $\bar{Y}_{FM1}$ and $\hat{Y}$. The proposals of Shalabh (1997) and Manisha and Singh (2001) are lesser optimally efficient in terms of mean squared error. Therefore, we conclude that the suggested estimators under measurement error model are more efficient than Shalabh (1997) and Manisha and Singh (2001) at their optimal level. The fact is found true in computation for different two auxiliary variables by taking one at a time.
In view of a composite setup of imputation and measurement error we found that the suggested estimators $\overline{Y}_{SD_1}$, $\overline{Y}_{SD_2}$, $\hat{Y}_1$, $\hat{Y}_2$, $\overline{Y}_1$ and $\overline{Y}_2$ are effective and efficient for computation of population mean.