6.1 A Generalised Family of Estimators for Estimating Unknown Variance Using Two Auxiliary Variables

In theory of sample surveys, we are estimating the population variance $S_y^2$ of study variable $y$. It is well known that use of auxiliary information improves the precision of proposed estimator. If information on an auxiliary variable is readily available then it is a well-known fact that the ratio-type and regression-type estimators can be used for estimation of parameters of interest, due to increase in efficiency of these estimators. The problem of estimating the population variance of $S_y^2$ of study variable $y$ received a considerable attention of the statistician in survey sampling including Grover(2010), Isaki(1983), Jhajj et al.(2005), Kadilar and Cingi (2005,2006,2007), Singh and Singh(2001,2003), Singh and Solanki (2012), Singh et al. (2008a, 2008b) and Singh et al. (2011) have suggested improved estimators for estimation of $S_y^2$.

Let $\varphi_i$ (i=1,2,…, N) be the population having N units such that $y$ is highly correlated with the auxiliary variables $x$ and $z$. We assume that a simple random sample without replacement (SRSWOR) of size $n$ is drawn from the finite population of size N. Let $(s_x^2,s_z^2)$ and $(s_x^2,s_z^2,s_y^2)$ be the sample variances defined over $n'$ and $n$ and $S_x^2, S_z^2$ and $S_y^2$ be the population variances of variables $x$, $z$ and $y$ respectively.

To estimate the population variance $S_y^2$ of study variable $y$, consider two cases:

- The population variances $S_x^2$ and $S_z^2$ of auxiliary variable $x$ and $z$ are known.
- When both $S_x^2$ and $S_z^2$ are unknown.

Where, $S_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2$, $S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2$, $S_z^2 = \frac{1}{N-1} \sum_{i=1}^{N} (z_i - \bar{Z})^2$. 
6.1.1. Large Sample Approximations

Let, $s_y^2 = (1 + \varepsilon_0)S_y^2$, $s_x^2 = (1 + \varepsilon_1)S_x^2$, $s_{xy}^2 = (1 + \varepsilon')S_{xy}^2$, $s_z^2 = (1 + \varepsilon_2)S_z^2$, and $s_{xz}^2 = (1 + \varepsilon')S_{xz}^2$, such that $E(\varepsilon_i) = E(\varepsilon_i') = 0, (i = 0, 1, 2)$.

Also to the first order of approximation, we have

$E(\varepsilon_0^2) = \gamma_1 V_{400}$, $E(\varepsilon_1^2) = \gamma_1 V_{040}$, $E(\varepsilon_2^2) = \gamma_1 V_{220}$, $E(\varepsilon_0 \varepsilon_1) = \gamma_1 V_{022}$, $E(\varepsilon_0 \varepsilon_2) = \gamma_1 V_{020}$, $E(\varepsilon_1 \varepsilon_2) = \gamma_1 V_{200}$, $E(\varepsilon_0 \varepsilon_1) = \gamma_1 V_{202}$, $E(\varepsilon_0 \varepsilon_2) = \gamma_1 V_{202}$.

where $\gamma_1 = \frac{1}{n}$, $\gamma_2 = \frac{1}{n}$, $\gamma = \left(\frac{1}{n} - \frac{1}{n^2}\right)$, $\nabla_{pq}^* = (\nabla_{pq} - 1)$ and $\nabla_{pq} = (\mu_{pq}/\mu_{200}^{020})^{\alpha/2}$, $\mu_{pq}^{020}$, $\mu_{200}$, being non negative integers.

6.1.2. Estimation of $S_y^2$ when both $S_x^2$ and $S_z^2$ are Known

In case when $S_x^2$ and $S_z^2$ are known, the conventional unbiased estimator is given as

The conventional unbiased estimator, $\hat{S}_y^2 = s_y^2$

Usual chain-ratio type estimator, is defined by,

$t_R = s_y^2 \left(\frac{S_x^2}{s_x^2} + \frac{S_z^2}{s_z^2}\right)$

The expressions of the variances and mean square error (MSE) of $\hat{S}_y^2$ and $t_R$, up to first order of approximation, are given by

\[\text{var}(\hat{S}_y^2) = \gamma_1 S_y^2 V_{400}^*\]

\[\text{MSE}(t_R) = \gamma_1 S_y^2 \left[ V_{400}^* + V_{040}^* + V_{022}^* - 2V_{220}^* - 2V_{202}^* + 2V_{020}^*\right]\]

Motivated by Koyuncu (2013), we propose a generalised estimator for estimating the population variance $S_y^2$, as

\[T_{pr} = s_y^2 \left[k_1 \left(\frac{S_x^2}{s_x^2}\right)^{\gamma_1} \exp\left(\frac{a(s_x^2 - S_x^2)}{S_x^2 + b(s_x^2 - S_x^2)}\right) + k_2 \left(\frac{S_z^2}{s_z^2}\right)^{\gamma_2} \exp\left(\frac{c(s_z^2 - S_z^2)}{S_z^2 + d(s_z^2 - S_z^2)}\right)\right]\]

where, $k_1$, $k_2$ are weights, $g_1$ and $g_2$ are constants $a$, $b$, $c$ and $d$ are either real numbers or the functions of known parameters.
Note: Here we will use the notations $T_{pr(1)}$ and $T_{pr(2)}$, rather than $T_{pr}$ for two different cases (ie: $k_1 + k_2 = 1$ and $k_1 + k_2 \neq 1$).

Bias and MSE of the suggested estimators are derived under two different conditions:

- **Case 1: When $k_1 + k_2 = 1$, where $k_1$ and $k_2$ are weights.**

**Theorem 2.1:** Estimator $T_{pr(1)}$ [equation (6.3)] in terms of $\varepsilon_i; i = 0, 1$ could be expressed as:

$$
T_{pr(1)} = S^2\gamma [1 + \varepsilon_0 + k_1 \{ \varepsilon_1 \eta_1 + \varepsilon_1^2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \} + k_2 \{ \varepsilon_2 \eta_3 + \varepsilon_2^2 \eta_4 + \varepsilon_0 \varepsilon_2 \eta_3 \}]
$$

(6.4)

Where, $\eta_1 = (a - g_1), \eta_2 = \left\{ \frac{a^2}{2} - ab - ag_1 + \frac{g_1 (g_1 + 1)}{2} \right\}, \eta_3 = (c - g_2)$ and $\eta_4 = \left\{ \frac{c^2}{2} - cd - cg_2 + \frac{g_2 (g_2 + 1)}{2} \right\}$.

**Proof:**

$$
T_{pr(1)} = S^2\gamma \left[ k_1 \left( \frac{S^2_x}{S^2_x} \right)^{\varepsilon_1} \exp \left( \frac{a(S^2_x - S^2_x)}{S^2_x + b(S^2_x - S^2_x)} \right) + k_2 \left( \frac{S^2_x}{S^2_x} \right)^{\varepsilon_2} \exp \left( \frac{c(S^2_x - S^2_x)}{S^2_x + d(S^2_x - S^2_x)} \right) \right]
$$

$$
= (1 + \varepsilon_0)S^2\gamma \left[ k_1 (1 + \varepsilon_1)^{-\varepsilon_1} \exp \left\{ a \{ e \varepsilon_1 (1 + c \varepsilon_1) \}^{-1} \right\} + k_2 (1 + \varepsilon_2)^{-\varepsilon_2} \exp \left\{ c \{ e \varepsilon_2 (1 + d \varepsilon_2) \}^{-1} \right\} \right]
$$

(6.5)

Here we assume $|\varepsilon_1|, |b \varepsilon_1|, |c \varepsilon_2|$ and $|d \varepsilon_2| < 1$, so that the terms $(1 + \varepsilon_1)^{-\varepsilon_1}, (1 + b \varepsilon_1)^{-1}, (1 + \varepsilon_2)^{-\varepsilon_2}$ and $(1 + d \varepsilon_2)^{-1}$ are expandable. By expanding the right hand side of equation (6.5) and neglecting the terms of $\varepsilon_i$, having power greater than two, we have

$$
T_{pr(1)} = S^2\gamma \left[ k_1 \{ l + \varepsilon_0 \eta_1 + \varepsilon_1^2 \eta_2 \} + k_2 \{ l + \varepsilon_0 \eta_3 + \varepsilon_2^2 \eta_4 \} \right]
$$

$$
T_{pr(1)} = S^2\gamma \left[ l + \varepsilon_0 + k_1 \{ \varepsilon_1 \eta_1 + \varepsilon_1^2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \} + k_2 \{ \varepsilon_2 \eta_3 + \varepsilon_2^2 \eta_4 + \varepsilon_0 \varepsilon_2 \eta_3 \} \right]
$$

(6.6)

**Theorem 2.2:** Bias of $T_{pr(1)}$ is given as

$$
B[T_{pr(1)}] = S^2\gamma \left[ k_1 \{ \eta_2 \mathbf{V}_0^{040} + \eta_1 \mathbf{V}_2^{220} \} + k_2 \{ \eta_4 \mathbf{V}_0^{044} + \eta_3 \mathbf{V}_2^{022} \} \right]
$$

(6.7)

**Proof:**

$$
B[T_{pr(1)}] = E[T_{pr(1)}] - S^2\gamma
$$

$$
= S^2\gamma \left[ \varepsilon_0 + k_1 \{ \varepsilon_1 \eta_1 + \varepsilon_1^2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \} + k_2 \{ \varepsilon_2 \eta_3 + \varepsilon_2^2 \eta_4 + \varepsilon_0 \varepsilon_2 \eta_3 \} \right]
$$

$$
B[T_{pr(1)}] = S^2\gamma \left[ k_1 \{ \eta_2 \mathbf{V}_0^{040} + \eta_1 \mathbf{V}_2^{220} \} + k_2 \{ \eta_4 \mathbf{V}_0^{044} + \eta_3 \mathbf{V}_2^{022} \} \right]
$$
Theorem 2.3: Mean square error of $T_{pr(1)}$, up to the first order of approximation: $O\left(\frac{1}{n}\right)$ is

$$MSE[T_{pr(1)}] = \frac{S^4}{n} \left[ \nabla_{400}^* + k_1^2 A_1 + k_2^2 B_1 + 2k_1 C_1 + 2k_2 D_1 + 2k_1 k_2 E_1 \right]$$

(6.8)

Where, $A_1 = \eta_1^2 \nabla_{040}^*, B_1 = \eta_2^2 \nabla_{004}^*, C_1 = \eta_1 \nabla_{220}^*, D_1 = \eta_3 \nabla_{202}^*$ and $E_1 = \eta_1 \eta_3 \nabla_{022}^*$

Proof:

$$MSE[T_{pr(1)}] = E[T_{pr(1)} - S_{ij}^2]^2 = \frac{S^4}{n} \eta_1^{2} \eta_2^{2} \eta_1 \eta_2^{2} + 2k_1 \eta_1 \varepsilon_0 \varepsilon_1 + 2k_2 \eta_2 \varepsilon_0 \varepsilon_2 + 2k_1 k_2 \eta_1 \eta_3 \varepsilon_1 \varepsilon_2]$$

$$MSE[T_{pr(1)}] = \frac{S^4}{n} \left[ \nabla_{400}^* + k_1^2 A_1 + (1 + k_1^2 - 2k_1) B_1 + 2k_1 C_1 + 2(1 - k_1) D_1 + 2k_1 (1 - k_1) E_1 \right]$$

(6.9)

Differentiating equation (6.9) w.r.t. $k_1$ and equating it to zero, we get the optimum value of $k_1$ as

$$k_1^* = \frac{(B_1 + D_1 - C_1 - E_1)}{(A_1 + B_1 - 2E_1)}$$

Putting the optimum value of $k_1$ in equation (6.9), we get the minimum MSE of the suggested method $T_{pr(1)}$.

Case 2: When $k_1 + k_2 \neq 1$, where $k_1$ and $k_2$ are weights.

Theorem 2.4: The estimator $T_{pr(2)}$ (equation 6.3) in terms of $\varepsilon_1$'s, could be expressed as:

$$T_{pr(2)} = S_{ij}^2 \left[ k_1 \left[ l + \varepsilon_0 + \varepsilon_1 \eta_1 + \varepsilon_2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \right] + k_2 \left[ l + \varepsilon_0 + \varepsilon_2 \eta_2 + \varepsilon_0 \varepsilon_2 \eta_2 \right] \right]$$

Where, $\eta_1 = (a - g_{1})$, $\eta_2 = \frac{a - ab - ag_1 + \frac{g_1 (g_1 + 1)}{2}}{2}$, $\eta_4 = (c - g_{2})$ and $\eta_4 = \frac{c^2}{2} - cd - cg_2 + \frac{g_2 (g_2 + 1)}{2}$

Proof: $T_{pr(2)} = S_{ij}^2 \left[ k_1 \left[ S_{s}^2 \right] \exp \left[ \frac{a(s_1^2 - S_{s}^2)}{S_{s}^2 + b(s_1^2 - S_{s}^2)} \right] + k_2 \left[ S_{s}^2 \right] \exp \left[ c(s_2^2 - S_{s}^2) \right] \right]$ $\left[ l + \varepsilon_0 \right] S_{ij}^2 \exp \left[ a \varepsilon_1 (1 + b \varepsilon_1)^{-1} \right] + k_2 \left( l + \varepsilon_2 \right)^{-1} \exp \left[ c \varepsilon_2 (1 + d \varepsilon_2)^{-1} \right]$ $\left(6.10\right)$

Here we assume $|\varepsilon_1|, |b \varepsilon_1|, |\varepsilon_2|$ and $|d \varepsilon_2| < 1$, so that the terms $(l + \varepsilon_1)^{-1}, (l + b \varepsilon_1)^{-1}, (l + \varepsilon_2)^{-1}$.
and \((1 + \kappa)\) are expandable. By expanding the right hand side of equation (6.10) and neglecting the terms of \(\varepsilon\)'s having power greater than two, we have
\[
T_{pr(1)} = S^2 \left[ k_1 \left( 1 + \varepsilon_0 + \varepsilon_1 \eta_1 + \varepsilon_1^2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \right) + k_2 \left( 1 + \varepsilon_0 + \varepsilon_2 \eta_3 + \varepsilon_2^2 \eta_4 + \varepsilon_0 \varepsilon_2 \eta_3 \right) \right] \tag{6.11}
\]

**Theorem 2.5:** Bias of \(T_{pr(2)}\) is given as:
\[
B[T_{pr(2)}] = S^2 \left[ k_1 \left( 1 + \eta_2 \nabla^{*}_{040} + \eta_1 \nabla^{*}_{220} \right) + k_2 \left( 1 + \eta_4 \nabla^{*}_{004} + \eta_3 \nabla^{*}_{202} \right) - 1 \right] \tag{6.12}
\]

**Proof:**
\[
B[T_{pr(2)}] = E[T_{pr(2)} - S^2] = S^2 E \left[ k_1 \left( 1 + \varepsilon_0 + \varepsilon_1 \eta_1 + \varepsilon_1^2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \right) + k_2 \left( 1 + \varepsilon_0 + \varepsilon_2 \eta_3 + \varepsilon_2^2 \eta_4 + \varepsilon_0 \varepsilon_2 \eta_3 \right) - 1 \right]
\]

**Theorem 2.6:** Mean square error of \(T_{pr(2)}\), up to the first order of approximation is:
\[
\text{MSE}[T_{pr(2)}] = S^4 \left[ 1 + k_1^2 A_2 + k_2^2 B_2 - 2k_1 C_2 - 2k_2 D_2 + 2k_1 k_2 E_2 \right] \tag{6.13}
\]

Where,
\[
A_2 = 1 + \frac{1}{n} \left[ \nabla^{*}_{400} + (\eta_1^2 + 2\eta_2) \nabla^{*}_{040} + 4\eta_1 \nabla^{*}_{220} \right]
\]
\[
B_2 = 1 + \frac{1}{n} \left[ \nabla^{*}_{400} + (\eta_3^2 + 2\eta_4) \nabla^{*}_{004} + 4\eta_3 \nabla^{*}_{202} \right]
\]
\[
C_2 = 1 + \frac{1}{n} \left[ \eta_2 \nabla^{*}_{040} + \eta_1 \nabla^{*}_{220} \right], \quad D_2 = 1 + \frac{1}{n} \left[ \eta_4 \nabla^{*}_{004} + \eta_3 \nabla^{*}_{202} \right]
\]
\[
E_2 = 1 + \frac{1}{n} \left[ \nabla^{*}_{400} + \eta_2 \nabla^{*}_{040} + \eta_4 \nabla^{*}_{004} + 2\eta_1 \nabla^{*}_{220} + 2\eta_3 \nabla^{*}_{202} + \eta_1 \eta_3 \nabla^{*}_{022} \right]
\]

**Proof:** From equation (6.10), we have
\[
[T_{pr(2)} - S^2]^2 = S^2 \left[ k_1 \left( 1 + \varepsilon_0 + \varepsilon_1 \eta_1 + \varepsilon_1^2 \eta_2 + \varepsilon_0 \varepsilon_1 \eta_1 \right) + k_2 \left( 1 + \varepsilon_0 + \varepsilon_2 \eta_3 + \varepsilon_2^2 \eta_4 + \varepsilon_0 \varepsilon_2 \eta_3 \right) - 1 \right]^2
\]

After expanding the above equation up to the first order of approximation ie: \(O\left(\frac{1}{n}\right)\) and then taking expectations of both sides, we get
\[
\text{MSE}[T_{pr(2)}] = S^4 \left[ 1 + k_1^2 A_2 + k_2^2 B_2 - 2k_1 C_2 - 2k_2 D_2 + 2k_1 k_2 E_2 \right]
\]

Minimising equation (6.13) with respect to \(k_1\) and \(k_2\), we get the optimum values \(k_1^*\) and \(k_2^*\) respectively given as
\[
k_1^* = \frac{B_2 C_2 - D_2 E_2}{A_2 B_2 - E_2^2} \quad \text{and} \quad k_2^* = \frac{A_2 D_2 - C_2 E_2}{A_2 B_2 - E_2^2}
\]

98
Putting this value of $k_1$ and $k_2$ in equation (6.13), we get the minimum MSE of the suggested estimator $T_{pr(2)}$.

6.1.3 Estimation of $s^2_y$ when both $s^2_x$ and $s^2_z$ are Unknown

Consider a finite population with $N (< \infty)$ identifiable units. Let $y$ be the variable under study taking values $y_i$ $(i = 1, 2, \ldots, N)$ for $i^{th}$ unit of the population. To estimate the population variance $S^2_y$ of $y$ in the presence of two auxiliary variables, when the population variances of $x$ and $z$ are not known and SRWSOR scheme is used in selecting samples for both the phases. The following two phase sampling scheme may be recommended.

1) The first phase sample $s'$ of size $(n' < N)$ is drawn in order to observe $x$ and $z$.

2) The second phase sample $s$ of size $(n < n')$ is drawn in order to observe $y$, $x$ and $z$.

Let $s'^2_x$ and $s'^2_z$ be the unbiased estimators of $S^2_x$ and $S^2_z$ respectively, based on first phase sample $s'$; $s'^2_y$, $s'^2_x$ and $s'^2_z$ be the unbiased estimators $S^2_y$, $S^2_x$ and $S^2_z$, respectively, based on second phase sample $s$. When $S^2_x$ and $S^2_z$ unknown, then usual unbiased and ordinary ratio estimator are given as

Usual unbiased estimator $\hat{S}^2_y = s'^2_y$

Usual chain ratio-type estimator in two phase sampling, is defined as

$$t_R = s'^2_y \left( \frac{s'^2_x}{s^2_x} \right) \left( \frac{s'^2_z}{s^2_z} \right)$$

MSE's expressions for the estimators $\hat{S}^2_y$ and $t_R$, up to the first order of approximation are respectively, given by

$$\text{MSE}(\hat{S}^2_y) = S^4_y \gamma_1 \nabla_{00}$$

$$\text{MSE}(t_R) = S^4_y \nabla_{040} \gamma_1 + \gamma \nabla_{040}^* + \gamma \nabla_{004} - 2\gamma \nabla_{220}^* - 2\gamma \nabla_{202} + 2 \gamma \nabla_{022}^*$$

when $S^2_x$ and $S^2_z$ are not known a priori, then under double sampling technique, suggested estimator $T_{pr}$ takes the following form:

$$T_{pr} = s'^2_y \left[ k_1 \left( \frac{s'^2_x}{s^2_x} \right)^{g_1} \exp \left\{ \frac{a(s'^2_x - s'^2_z)}{s'^2_x + b(s'^2_x - s'^2_z)} \right\} + k_2 \left( \frac{s'^2_z}{s^2_z} \right)^{g_2} \exp \left\{ \frac{c(s'^2_z - s'^2_x)}{s'^2_z + d(s'^2_z - s'^2_x)} \right\} \right]$$

where, $k_1$, $k_2$ are weights, $g_1$ and $g_2$ are constants and $a$, $b$, $c$ and $d$ are either real numbers or the functions of known parameters.

99
Note: Here we have taken the notation $T_{pr(1)}'$ and $T_{pr(2)}'$, rather than $T_{pr}$ for two different cases (ie: $k_1 + k_2 = 1$ and $k_1 + k_2 \neq 1$).

To obtain the bias and mean square error of the suggested estimator we have the following two conditions given as-

**Case 1: When $k_1 + k_2 = 1$, where $k_1$ and $k_2$ are defined weights.**

**Theorem 3.1:** The suggested estimator $T_{pr(1)}'$ (see equation 6.16) in terms of $\varepsilon_i$s, could be expressed as:

$$T_{pr(1)}' = S_y^2 \left[ 1 + \varepsilon_0 + k_1 \left( a - g_1 \right) \left( \varepsilon_i - \varepsilon'_i \right) + \varepsilon_i^2 Q_1 + \varepsilon'_i Q_2 + \varepsilon_1 \varepsilon'_i Q_3 + \left( a - g_1 \right) \left( \varepsilon_i \varepsilon_1 - \varepsilon_0 \varepsilon'_i \right) \right]$$

where,

$$Q_1 = \frac{a^2}{2} - ab + g_1 \left( g_1 + 1 \right) / 2, \quad Q_2 = \frac{a^2}{2} + a - ab - ag_1 + \frac{g_1 \left( g_1 - 1 \right)}{2},$$

$$Q_3 = 2ab - a - a^2 + 2ag_1 - g_1^2, \quad Q_4 = \frac{c^2}{2} - cd - g_2 c + \frac{g_2 \left( g_2 + 1 \right)}{2},$$

$$Q_5 = c + \frac{c^2}{2} - cd - g_2 c + \frac{g_2 \left( g_2 - 1 \right)}{2}, \quad Q_6 = 2cd - c - c^2 + 2g_2 c - g_2^2$$

**Proof:**

$$T_{pr(1)}' = S_y^2 \left[ k_1 \left( \frac{s_x^2}{s_x^2} \right)^{\gamma_1} \exp \left( \frac{a \left( s_x^2 - s_x^2 \right)}{s_x^2 + b \left( s_x^2 - s_x^2 \right)} \right) + k_2 \left( \frac{s_x^2}{s_x^2} \right)^{\gamma_2} \exp \left( \frac{c \left( s_x^2 - s_x^2 \right)}{s_x^2 + d \left( s_x^2 - s_x^2 \right)} \right) \right]$$

After expanding and arranging the above equation up to the first order of approximation, we have:

$$T_{pr(1)}' = S_y^2 \left[ 1 + \varepsilon_0 + k_1 \left( a - g_1 \right) \left( \varepsilon_i - \varepsilon'_i \right) + \varepsilon_i^2 Q_1 + \varepsilon'_i Q_2 + \varepsilon_1 \varepsilon'_i Q_3 + \left( a - g_1 \right) \left( \varepsilon_i \varepsilon_1 - \varepsilon_0 \varepsilon'_i \right) \right]$$

where, coefficient term of $\varepsilon_i$'s are already defined.

**Theorem 3.2:** Bias of $T_{pr(1)}'$ is given as

$$B(T_{pr(1)}') = S_y^2 \left[ k_1 \left( Q_1 + \gamma_1 V_{040}^* + \gamma_2 V_{040}^* \left( Q_2 + Q_3 \right) + \left( a - g_1 \right) \gamma V_{220}^* \right) \right]$$

**Proof:** Subtracting $S_y^2$ from both sides and taking expectations of equation (6.17), we have
\[ B(T'_{pr(1)}) = S_y^2 \left[ k_1 \left\{ Q_1 \gamma_1 \text{V}_{o40}^* + \gamma_2 \text{V}_{o40}^* (Q_2 + Q_3) + (a - g_1) \gamma \text{V}_{220}^* \right\} + k_2 \left\{ Q_4 \gamma_1 \text{V}_{o40}^* + \gamma_2 \text{V}_{o40}^* (Q_5 + Q_6) + (c - g_2) \gamma \text{V}_{220}^* \right\} \right] \] (6.18)

**Theorem 3.3:** Mean square error of \( T'_{pr(1)} \), up to the first order of approximation is given as:

\[
\text{MSE}(T'_{pr(1)}) = S_y^4 \left[ \frac{\text{V}_{400}^*}{n} + k_1^{-2} A + k_2^{-2} B + 2k_1' C + 2k_2' D + 2k_1' k_2' E \right] 
\]

Or

\[
\text{MSE}(T'_{pr(1)}) = S_y^4 \left[ \frac{\text{V}_{400}^*}{n} + k_1^{-2} A_3 + (1 - k_1')^2 B_3 + 2k_1' C_3 + 2(1 - k_1') D_3 + 2k_1' (1 - k_1') E_3 \right] 
\]

where, \( A_3 = \eta_1 \gamma \text{V}_{040}^*, B_3 = \eta_2 \gamma \text{V}_{040}^*, C_3 = \eta_1 \gamma \text{V}_{220}^*, D_3 = \eta_3 \gamma \text{V}_{202}^*, E_3 = \eta_1 \eta_3 \gamma \text{V}_{022}^* \)

such that \( \eta_1 = (a - g_1) \) and \( \eta_3 = (c - g_2) \)

**Proof:** from equation (6.18), we have

\[
(T'_{pr(1)} - S_y^{-2})^2 = S_y^4 [\tilde{e}_o + k_1' \eta_1 (\epsilon_1 - \epsilon_1') + k_2' \eta_3 (\epsilon_2 - \epsilon_2')]^2 
\]

After squaring and taking expectations of both sides of the above equation, up to the first order of approximation, we get

\[
\text{MSE}(T'_{pr(1)}) = S_y^4 \left[ \frac{\text{V}_{400}^*}{n} + k_1^{-2} A_3 + (1 - k_1')^2 B_3 + 2k_1' C_3 + 2(1 - k_1') D_3 + 2k_1' (1 - k_1') E_3 \right] 
\] (6.19)

Differentiating equation (6.19) partially w.r.t. to \( k_1' \), we get the optimum value of \( k_1' \) for minimum MSE as:

\[
k_1'^* = \left( \frac{B_3 + D_3 - C_3 - E_3}{A_3 + B_3 - 2E_3} \right) 
\]

Putting the optimum value of \( k_1'^* \) in (6.19), we get the minimum MSE of the suggested estimator \( T'_{pr(1)} \).

**Case 2:** When \( k_1' + k_2' \neq 1 \), where \( k_1' \) and \( k_2' \) are defined weights.

\[
T'_{pr(2)} = S_y^2 \left[ k_1 \left\{ \frac{s_x^2}{s_x^2} \right\}^{s_1} \exp \left( \frac{a(s_x^2 - s_x^2)}{s_x^2 + b(s_x^2 - s_x^2)} \right) + k_2 \left\{ \frac{s_x^2}{s_x^2} \right\}^{s_2} \exp \left( \frac{c(s_x^2 - s_x^2)}{s_x^2 + d(s_x^2 - s_x^2)} \right) \right] 
\] (6.20)
Theorem 3.4: The suggested estimator $T'_{pr(2)}$ in terms of $\varepsilon_i$’s, is written as
\[
T'_{pr(2)} = S_y^2 \left[ k_1 \left\{ l + \varepsilon_0 + \eta_1 (e_i - \varepsilon_i) + \varepsilon_1^2 Q_1 + \varepsilon_i^2 Q_2 + \varepsilon_i \varepsilon_i' Q_3 + \eta_1 (e_0 e_i - e_0 \varepsilon_i) \right\} \\
+ k_2 \left\{ l + \varepsilon_0 + \eta_3 (e_2 - \varepsilon_2) + \varepsilon_2^2 Q_4 + \varepsilon_2 \varepsilon_2' Q_5 + \varepsilon_2 \varepsilon_2' Q_6 + \eta_3 (e_0 e_2 - e_0 \varepsilon_2) \right\} \right]
\]
where, $\eta_1, \eta_3, Q_1, Q_2, Q_3, Q_4, Q_5$, and $Q_6$ are defined in theorem 3.1 and 3.2.

Proof:
\[
T'_{pr(2)} = S_y^2 \left[ k_1 \left\{ l + \varepsilon_0 + \eta_1 (e_i - \varepsilon_i) + \varepsilon_1^2 Q_1 + \varepsilon_i^2 Q_2 + \varepsilon_i \varepsilon_i' Q_3 + \eta_1 (e_0 e_i - e_0 \varepsilon_i) \right\} \\
+ k_2 \left\{ l + \varepsilon_0 + \eta_3 (e_2 - \varepsilon_2) + \varepsilon_2^2 Q_4 + \varepsilon_2 \varepsilon_2' Q_5 + \varepsilon_2 \varepsilon_2' Q_6 + \eta_3 (e_0 e_2 - e_0 \varepsilon_2) \right\} \right]
\] (6.21)
After expanding and arranging the above equation (6.21) up to the first order of approximation, we have
\[
T'_{pr(2)} = S_y^2 \left[ k_1 \left\{ l + \varepsilon_0 + \eta_1 (e_i - \varepsilon_i) + \varepsilon_1^2 Q_1 + \varepsilon_i^2 Q_2 + \varepsilon_i \varepsilon_i' Q_3 + \eta_1 (e_0 e_i - e_0 \varepsilon_i) \right\} \\
+ k_2 \left\{ l + \varepsilon_0 + \eta_3 (e_2 - \varepsilon_2) + \varepsilon_2^2 Q_4 + \varepsilon_2 \varepsilon_2' Q_5 + \varepsilon_2 \varepsilon_2' Q_6 + \eta_3 (e_0 e_2 - e_0 \varepsilon_2) \right\} \right]
\] (6.22)
where, coefficient term of $\varepsilon_i$’s are already defined.

Theorem 3.5: Bias of suggested method $T'_{pr(2)}$ is given as
\[
B(T'_{pr(2)}) = S_y^2 \left[ k_1 \left\{ l + Q_1 \gamma_{11} \gamma_{11}^{*} v_{11}^{0} + \gamma_2 \gamma_{11}^{*} + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_2 + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_3 + \eta_1 \gamma \gamma v_{11}^{*} \right\} \\
+ k_2 \left\{ l + Q_1 \gamma_{11} \gamma_{11}^{*} v_{11}^{0} + \gamma_2 \gamma_{11}^{*} + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_2 + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_3 + \eta_1 \gamma \gamma v_{11}^{*} \right\} \right]
\]
Proof: Subtracting $S_y$ both side of equation (6.22), we have
\[
T'_{pr(2)} \left[ k_1 \left\{ l + \varepsilon_0 + \eta_1 (e_i - \varepsilon_i) + \varepsilon_1^2 Q_1 + \varepsilon_i^2 Q_2 + \varepsilon_i \varepsilon_i' Q_3 + \eta_1 (e_0 e_i - e_0 \varepsilon_i) \right\} \\
+ k_2 \left\{ l + \varepsilon_0 + \eta_3 (e_2 - \varepsilon_2) + \varepsilon_2^2 Q_4 + \varepsilon_2 \varepsilon_2' Q_5 + \varepsilon_2 \varepsilon_2' Q_6 + \eta_3 (e_0 e_2 - e_0 \varepsilon_2) \right\} \right] = 1 - 1
\] (6.23)
Taking expectations of both sides, we get the required bias of the suggested method as
\[
B(T'_{pr(2)}) = S_y^2 \left[ k_1 \left\{ l + Q_1 \gamma_{11} \gamma_{11}^{*} v_{11}^{0} + \gamma_2 \gamma_{11}^{*} + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_2 + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_3 + \eta_1 \gamma \gamma v_{11}^{*} \right\} \\
+ k_2 \left\{ l + Q_1 \gamma_{11} \gamma_{11}^{*} v_{11}^{0} + \gamma_2 \gamma_{11}^{*} + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_2 + \gamma_2 \gamma_{11}^{*} v_{11}^{0} Q_3 + \eta_1 \gamma \gamma v_{11}^{*} \right\} \right]
\]

Theorem 3.6: Mean square error of the suggested estimator $T'_{pr(2)}$, is given by
\[
MSE(T'_{pr(2)}) = S_y^4 \left[ 1 + k_1^2 A + k_2^2 B - 2k_1 C - 2k_2 D + 2k_1 k_2 E \right]
\]
where, $A = \left[ 1 + \frac{\gamma_{11}^{*}}{n} + \gamma_{11}^{*} \left\{ \eta_1 \gamma + \frac{2Q_1}{n} + \frac{2Q_2}{n} + \frac{2Q_3}{n^2} \right\} + 4n_1 \gamma \gamma v_{11}^{*} \right]$, 
$B = \left[ 1 + \frac{\gamma_{11}^{*}}{n} + \gamma_{11}^{*} \left\{ \eta_1 \gamma + \frac{2Q_1}{n} + \frac{2Q_2}{n} + \frac{2Q_3}{n^2} \right\} + 4n_1 \gamma \gamma v_{11}^{*} \right]$. 

102
\[ C = \left[ I + \nabla_{y}^{*} \left( \frac{Q_{1}}{n} + \frac{Q_{2}}{n'} + \frac{Q_{3}}{n''} \right) + \eta_{1} \gamma \nabla_{220}^{*} \right], D = \left[ I + \nabla_{004}^{*} \left( \frac{Q_{4}}{n} + \frac{Q_{5}}{n'} + \frac{Q_{6}}{n''} \right) + \eta_{1} \gamma \nabla_{202}^{*} \right] \]

\[ E = \left[ I + \nabla_{400}^{*} + \nabla_{y}^{*} \left( \eta_{1} \gamma + \frac{Q_{1}}{n} + \frac{Q_{2}}{n'} + \frac{Q_{3}}{n''} \right) + \nabla_{004}^{*} \left( \eta_{1} \gamma + \frac{Q_{4}}{n} + \frac{Q_{5}}{n'} + \frac{Q_{6}}{n''} \right) \right] \]

\[ + 2 \eta_{1} \gamma \nabla_{202}^{*} + 2 \eta_{1} \gamma \nabla_{220}^{*} + \eta_{1} \eta_{1} \gamma \nabla_{022}^{*} \]

**Proof:** we have

\[ T_{pr(2)}^{*} - S_{y}^{2} = S_{y}^{2} \left[ k_{1}^{*} \left( 1 + e_{0} + \eta_{1} (e_{1} - e_{1}^{*}) + e_{1}^{*} Q_{1} + e_{1}^{*} Q_{2} + e_{1}^{*} Q_{3} + \eta_{1} (e_{0} e_{1} - e_{0} e_{1}^{*}) \right) \right] \]

After squaring and taking expectations of both sides of the above equation up to the first order of approximation, MSE of the estimator \( T_{pr(2)}^{*} \)

\[ \text{MSE}(T_{pr(2)}^{*}) = S_{y}^{2} \left[ 1 + k_{1}^{*2} A + k_{2}^{*2} B - 2 k_{1}^{*} C - 2 k_{2}^{*} D + 2 k_{1}^{*} k_{2}^{*} E \right] \]  

(6.24)

The optimum value of \( k_{1}^{*} \) and \( k_{2}^{*} \) is obtained by partially differentiating equation (6.24) w.r.t. \( k_{1}^{*} \) and \( k_{2}^{*} \) for minimum MSE, as

\[ k_{1}^{*} = \frac{B C - D E}{A B - E^{2}} \quad \text{and} \quad k_{2}^{*} = \frac{A D - C E}{A B - E^{2}} \]

Putting the optimum values of \( k_{1}^{*} \) and \( k_{2}^{*} \) in (6.19), we get the minimum MSE(\( T_{pr(2)}^{*} \)).

**6.1.4 Empirical Study**

**Population 1:** For empirical study we have taken the data of Murthy (1967), which contain the data of 34 villages. The variables are

- y- Area under wheat in 1964.
- x- Area under wheat in 1963.
- z- Cultivated area in 1961.

Also,

\[ V_{400} = 3.726, V_{040} = 2.912, V_{004} = 2.808, V_{220} = 3.105, V_{202} = 2.979, V_{022} = 2.738 \]

\[ S_{y}^{2} = 22564.56, S_{x}^{2} = 197095.3, S_{z}^{2} = 2652.05, S_{xy} = 60304.01, S_{yx} = 22158.05 \]

\[ n = 7, n' = 15, N = 34. \]

**Population 2:** The second data for the empirical study is taken from Ahmed (1995). The population consists of 340 villages. The variables are:

- y- Number of literate persons.
- x- Number of household.
To estimate the variance of $y$, we have used $x$ and $z$ as the prior information. From the data set, we have

$$N = 340, n' = 120, n = 50, S^2_y = 71379.47, S^2_x = 11838.85, S^2_z = 691820.23$$

$$\nabla^*_0 = 9.90334289, \nabla^*_0 = 7.05448224, \nabla^*_0 = 8.2552346, \nabla^*_0 = 6.31398563$$

$$\nabla^*_2 = 8.12904924, \nabla^*_2 = 6.13646859.$$ 

The percentage relative efficiency (PRE) of estimator is defined as

$$\text{PRE} = \frac{\text{VAR}(\hat{S}^*_y)}{\text{MSE}(\hat{S}^*_y)} \times 100$$

### Table 6.1: Percentage Relative Efficiency (PRE) for population-1.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>g_1</th>
<th>g_2</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S^2_y$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>$T_R$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>135.29</td>
</tr>
<tr>
<td>$T_{pr(1)}$</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>658.78</td>
</tr>
<tr>
<td>$T_{pr(1)}$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>897.38</td>
</tr>
<tr>
<td>$T_{pr(1)}$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4417.39</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5256.36</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1365.21</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>182.60</td>
<td></td>
</tr>
<tr>
<td>$T_{pr(1)}$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>190.08</td>
</tr>
<tr>
<td>$T_{pr(1)}$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>186.51</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>404.38</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>347.81</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>342.17</td>
</tr>
</tbody>
</table>

**Note:** Here the notations $T_{pr(1)}$, $T_{pr(2)}$ stands for suggested estimators in single phase and respectively $T^*_r$ $T^*_r$ is for two phase sampling.
### Table 6.2: Percentage Relative Efficiency (PRE) for Population-2.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>g₁</th>
<th>g₂</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_y^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>$t_R$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>115.15</td>
</tr>
<tr>
<td>$T_{pr(1)}$</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>521.62</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>342.12</td>
</tr>
<tr>
<td>$T_{pr(2)}$</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>868.25</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>709.37</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>189.21</td>
</tr>
<tr>
<td>$T_{pr(1)}'</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>170.35</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>169.32</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>315.53</td>
</tr>
<tr>
<td>$T_{pr(2)}'$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>342.16</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>404.38</td>
</tr>
</tbody>
</table>

**Note:** Here the notations $T_{pr(1)}$ , $T_{pr(2)}$ stands for suggested estimators in single phase and $T_{pr(1)}'$ , $T_{pr(2)}'$ stands for suggested estimators in two-phase sampling.

### 6.1.5. Conclusion

From the above Table 6.1 and Table 6.2, we observed that, the suggested estimators $T_{pr(1)}$ , $T_{pr(2)}$ (in single phase sampling) and $T_{pr(1)}'$ , $T_{pr(2)}'$ (in two phase sampling) perform much better than the usual unbiased and ratio estimator.
6.2 Some Generalised Class of Estimators for the Estimation of Coefficient of Variation of Finite Population in Two-Phase Sampling

In survey sampling, the use of auxiliary information has resulted in extensive gain in performance over the estimators which do not take such information. The auxiliary information has been effectively used in double sampling to estimate the coefficient of variation $C_y$ of study variable $y$. The problem of estimating finite population variance using multi-auxiliary information has attracted attention of many authors including Das and Tripathi (1978), Prasad and Singh (1990, 1992), Srivastava and Jhajj (1983) and Isaki (1983). However, the problem of estimation of population coefficient of variation (i.e. dispersion per unit mean in the population) has much importance, as it measures variability in relation to the mean (or average) and is used to compare the relative dispersion in one type of data with the relative dispersion in another type of data. For instance if we want to evaluate the relative dispersion of grades for two classes of students: Class A and Class B, the coefficient of variation can be used to compare these two groups and determine how the grade dispersion in class A compares to the grade dispersion to the class B. In literature coefficient of variation has not attracted much attention except in Das and Tripathi (1992, 93) and Singh and Singh (2002).

Let $V = (V_1, V_2, \ldots, V_N)$ be a finite population of $N$ identifiable units; $(y, x, z)$ be real non-negative variables taking values $(y_i, x_i, z_i)$ for the $i^{th}$ unit $i=1,2,\ldots,N$, where $y$ and $x, z$ are the variable under study and auxiliary variable respectively. In a certain situations where population mean $\bar{X}$, population variance, $\sigma_x^2$, population coefficient of variation $C_x$, population mean $\bar{X}$ and variance $\sigma_z^2$ of the auxiliary variable $x$ is/are known. Das and Tripathi (1992-93) have suggested various classes of estimators for estimating the population coefficient of variation $C_y$ with their properties. However, in other way when this auxiliary information indicated above is/are not known in advance, the technique of two phase sampling is used. In two phase sampling scheme:

1) First we draw a sample $s'$ of size $n_1$ from $N$ identifiable units to observed $x$ and $z$ and then furnish the estimates of $(\bar{X}, \sigma_x^2)$, then
2) Draw a preliminary subsample \( s \subset s' \) of size \( n \) from the first sample \( n_1 \) to observed \( y \) only.

Sometimes, even if \( (\bar{X}, \sigma^2_X, C_X) \) are unknown, information on a second auxiliary variable \( z \) related to \( y \) and \( x \) is readily available. Let \( (\bar{Z}, \sigma^2_Z) \) be the known population mean and variance respectively of variable \( z \). It may be mentioned that if the \( \bar{X} \) is not known, but the population mean \( \bar{Z} \) of auxiliary variable \( z \) closely related to \( x \) but remotely related to \( y \) (i.e., \( \rho_{xy} > \rho_{yz} \)) is available, Singh et al. (1994), Singh (1993), Sahoo and Sahoo (1993), Srivastava et al. (1990), Kiregyera (1980,84), Chand (1975) among others have suggested chain ratio type estimators for estimating population mean \( \bar{Y} \) of \( y \). Recently, Gupta et al. (1992-1993) have suggested three chain ratio type estimators for estimating population variance \( \sigma^2_Y \) of \( y \) in different situations and studied their properties.

### 6.2.1 Notation Used

Let,

- \( N \) - population size.
- \( n_1 \) - first phase sample.
- \( n \) - preliminary second phase sample taken from \( n_1 \).

\[ S^2_x = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2, \quad S^2_y = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2, \quad S^2_z = \frac{1}{N-1} \sum_{i=1}^{N} (z_i - \bar{Z})^2 \] - Population mean square.

\[ C_y = \frac{S^2_y}{\bar{Y}}, \quad C_x = \frac{S^2_x}{\bar{X}}, \quad C_z = \frac{S^2_z}{\bar{Z}} \] - Coefficient of variations of \( y, x \) and \( z \) respectively.

- \( \rho_{xy} \), \( \rho_{yz} \) - population correlation coefficient between \( (x, y) \) and \( (z, y) \), respectively.

- \( B(.) \) - Bias of estimator

- \( \text{MSE}(.) \) - Mean Squared Error
6.2.2 Large Sample Approximation

Here, we assume that simple random sampling without replacement (SRSWOR) scheme is adopted for selecting the samples at both the phases. For simplicity, we assume that $N$ is large as compared to $(n, n')$, so that finite population correction terms are ignored. We write,

$$y = \frac{1}{n} \sum_{i=1}^{n} y_i, \; x = \frac{1}{n} \sum_{i=1}^{n} x_i, \; \bar{y} = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i, \; \bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i$$

$$s_y^2 = \frac{1}{n - 1} \sum_{i=1}^{n} (y_i - \bar{y})^2, \; s_x^2 = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{x})^2, \; s_{yz}^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (z_i - \bar{z})^2$$

Then we have,

$$\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i, \; \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i, \; \bar{z} = \frac{1}{N} \sum_{i=1}^{N} z_i$$

$$\bar{y}' = y_1, \; \bar{x}' = x_1, \; \bar{z}' = z_1$$

Let

$$\bar{y} = \bar{y}(1 + e_{01}), \; \bar{x} = \bar{x}(1 + e_{1}), \; \bar{z} = \bar{z}(1 + e_{2})$$

$$s_y^2 = \sigma_y^2(1 + e_{01}), \; s_x^2 = \sigma_x^2(1 + e_{1}), \; s_{yz}^2 = \sigma_{yz}^2(1 + e_{2})$$

Then we have,
\[
E(e_{0i}) = E(e_0) = E(e_i) = E(e_z) = 0 \quad \text{and} \quad E(e'_i) = 0, \quad \text{for all } i=1,2,3,4.
\]

\[
E(e_{0i}^2) = C_x^2 / n, \quad E(e_i^2) = C_x^2 / n, \quad E(e_{z1}^2) = C_z^2 / n_1, \quad E(e_{z2}^2) = C_z^2 / n_1
\]

\[
E(e_{0i}e_i) = C_x^2 / n, \quad E(e_{0i}e_z^*) = C_x^2 / n_1, \quad E(e_{0i}e_{z2}^*) = C_z^2 / n_1
\]

\[
E(e_i'e_i^*) = C_x^2 / n_1, \quad E(e_i'e_{z2}^*) = \rho_{xz} C_x C_z / n_1, \quad E(e_i'e_{z2}^*) = \rho_{xz} C_x C_z / n_1 \quad \text{and} \quad f = \left( \frac{1}{n} - \frac{1}{n_1} \right)
\]

And to the first degree of approximation (or upto terms of order \(n^{-1}\))

\[
E(e_0^2) = \nabla^{000}_4 / n, \quad E(e_i^2) = \nabla^{000}_4 / n, \quad E(e_{z1}^2) = \nabla^{000}_4 / n_1, \quad E(e_{z2}^2) = \nabla^{000}_4 / n_1, \quad E(e_{0i}e_i) = \nabla^{000}_4 / n
\]

\[
E(e_{0i}e_{z2}^*) = \nabla^{000}_4 / n_1, \quad E(e_{0i}e_{z2}^*) = \nabla^{000}_4 / n_1, \quad E(e_{0i}e_{z2}^*) = \nabla^{000}_4 / n_1
\]

\[
E(e_i'e_i^*) = \nabla^{000}_4 / n_1, \quad E(e_i'e_{z2}^*) = \nabla^{000}_4 / n_1, \quad E(e_i'e_{z2}^*) = \nabla^{000}_4 / n_1
\]

\[
E(e_i'e_{z2}^*) = \nabla^{000}_4 / n_1, \quad E(e_i'e_{z2}^*) = \nabla^{000}_4 / n_1, \quad E(e_i'e_{z2}^*) = \nabla^{000}_4 / n_1
\]

The present study deals with the problem of estimating the population coefficient of variation \(C_y\) of study variable \(y\) using auxiliary information in two phase sampling and have suggested three different classes of estimators.

### 6.2.3 Existing Estimators and their Properties

When population mean \(\overline{Z}\) and variance \(\sigma_z^2\) are known, consider usual ratio type estimators defined by

\[
\tau_{RM} = \hat{C}_y \left( \overline{X} / \overline{Z} \right) \hat{Y} \left( \overline{Z} / \overline{Z} \right)^{\hat{Y}} \quad \text{(6.25)}
\]

109
Usual exponential ratio-type estimator using two auxiliary variates, when population mean $\bar{Z}$ and variance $\sigma^2_z$ are known is given by

$$\tau_{RV} = \hat{C}_y^2 \left( \frac{s^2_x}{\sigma^2_x} \right)^{\phi_1} \left( \frac{s^2_x}{\sigma^2_x} \right)^{\phi_4}$$

(6.26)

When population mean $\bar{Z}$ and variance $\sigma^2_z$ are known, Singh and Singh (2002) suggested a class of chain ratio-type estimators of $\hat{C}_y$, as

$$\tau_S = \hat{C}_y \left( \frac{\bar{X}}{\bar{X}} \right)^{\alpha_1} \left( \frac{s^2_x}{s^2_x} \right)^{\alpha_2} \left( \frac{\bar{Z}}{\bar{Z}} \right)^{\alpha_3} \left( \frac{s^2_z}{\sigma^2_z} \right)^{\alpha_4}$$

(6.29)

where $\alpha_i$'s (i =1, 2, 3, 4) are suitably chosen constants.

The mean square error (MSE) of $\tau_{RM}$, $\tau_{RV}$, $\tau_{ERM}$, and $\tau_{ERV}$, up to the first order of approximation, are respectively given by

$$\text{MSE}(\tau_{RM}) = \text{MSE}(\hat{C}_y^2) + C_y \left[ \phi_1^2 C_x^2 f_1 + \frac{\phi_2^2 C_x^2}{n_1} - \phi_1 f_1 \left( 2 CC_x^2 - \nabla_{210} C_x \right) - \frac{\phi_2}{n_1} \left( 2 C^* C_x^2 - \nabla_{201} C_x \right) \right]$$

(6.30)

where,

$$\phi_1 (\text{opt}) = \frac{2CC_x - \nabla_{210}}{C_x} \quad \text{and} \quad \phi_2 (\text{opt}) = \frac{2C^* C_x - \nabla_{201}}{C_x}$$

$$\text{MSE}(\tau_{RV}) = \text{MSE}(\hat{C}_y^2) + C_y \left[ \phi_3^2 \nabla_{040}^* f_1 + \frac{\phi_4^2 \nabla_{004}^*}{n_1} - \phi_3 f_1 \left( 2 \nabla_{120} C_y - \nabla_{220}^* \right) \right]$$

$$- \frac{\phi_4}{n_1} \left( 2 \nabla_{302} C_y - \nabla_{202}^* \right)$$

(6.31)
where,
\[ \phi_3(\text{opt}) = \frac{2V_{120} C_y - \nabla_{220}^*}{2V_{040}^*} \quad \text{and} \quad \phi_4(\text{opt}) = \frac{2V_{102} C_y - \nabla_{202}^*}{2V_{004}^*} \]

\[
\text{MSE}(\tau_{\text{ERM}}) = \text{MSE}(\hat{C}_y^2) + C_y^2 \left[ \phi_3^2 \frac{C_y^2}{4} f_i + \phi_2^2 - \phi_4 f_i \left( \frac{\nabla_{210} C_x}{2} - CC_x \right) \right.
\]
\[
\left. - \frac{\phi_2}{n_1} \left( \nabla_{201} C_z - C^* C_z^2 \right) \right] \quad (6.32)
\]

where, \( \phi_1(\text{opt}) = \frac{2(\nabla_{210} - 2CC_x)}{C_x} \) and \( \phi_2(\text{opt}) = \frac{2(2\nabla_{120} C_z - C^* C_z)}{C_z} \)

\[
\text{MSE}(\tau_{\text{ERV}}) = \text{MSE}(\hat{C}_y^2) + C_y^2 \left[ \phi_3^2 \frac{C_y^2}{4} f_i + \phi_4^2 \frac{C_y^2}{4n_1} - \phi_4 f_i \left( \frac{\nabla_{220}^*}{2} - \nabla_{120} C_y \right) \right.
\]
\[
\left. - \frac{\phi_4}{n_1} \left( \nabla_{202}^* - \nabla_{102} C_y \right) \right] \quad (6.33)
\]

The bias and MSE of the estimator \( \tau_s \), up to the first order of approximation, are respectively given by

\[
B(\tau_s) = B(\hat{C}_y^2) + C_y \left[ f_i \left( \alpha_1 \frac{C_y}{2} \left( (\alpha_1 - 1)C_x + \nabla_{210} - 2CC_x \right) + \alpha_1 \alpha_2 \nabla_{030} C_x \right) \right.
\]
\[
+ \frac{\alpha_2}{2} \left( (\alpha_2 - 1) - \nabla_{040}^* + \nabla_{220} - 2\nabla_{120} C_y - 1 \right) \left( \nabla_{120} C_y \right) - \frac{1}{n_1} \left( (\alpha_3 - 1)C_x + \nabla_{201} - 2CC_z \right) \]
\[
+ \frac{\alpha_4}{2} \left( (\alpha_4 - 1)\nabla_{004}^* + \nabla_{202} - 2\nabla_{102} C_y - 1 \right) \left( \nabla_{102} C_y \right) \right]
\]

where, \( B(\hat{C}_y^2) = \frac{C_y}{n} \left[ C_y^2 - \nabla_{300} C_y - \frac{n_1}{8} \right] \)

and

\[
\text{MSE}(\tau_s) = \text{MSE}(\hat{C}_y^2) + C_y^2 f_i \left[ \alpha_1 C_x^2 \left( \alpha_1 - 2 \left( C - \frac{\nabla_{210}}{2C_x} \right) \right) + \alpha_2 \left( \nabla_{040} - 1 \right) \right.
\]
\[
- \left( 2\nabla_{120} C_y - \nabla_{220} + 1 \right) + \frac{2\alpha_1 \alpha_2 \nabla_{030}}{C_x^2} \left[ \frac{C_y}{n_1} \left( \alpha_3 - 2 \left( C^* - \frac{\nabla_{201}}{2C_z} \right) \right) \right]
\]

\[111\]
\[ + \alpha_4 \left( \alpha_4 (V_{004} - 1) - (2V_{102} C_y - V_{102} + 1) \right) + \frac{2\alpha_3 \alpha_4}{C_z} \] (6.34)

where,

\[
\text{MSE}(\hat{C}_y) = \frac{C_y^2}{n} \left[ C_y^2 - \frac{V_{300} C_y + V_{400}^4}{4} \right] (6.35)
\]

The MSE of \( \tau_s \) in (6.34) is minimised for

\[
\alpha_1 = \frac{\left[ \frac{V_{040}^0 (2\rho_{yx} C_y - V_{210}) - \frac{V_{040}^0 (2V_{120} C_y - V_{200} + 1)}{2} \right]}{(V_{040} - V_{030} - 1)C_x} \]
\[
\alpha_2 = \frac{\left[ \frac{V_{040}^0 (2V_{120} C_y - V_{220} + 1) - \frac{V_{030}^0 (2\rho_{yx} C_y - V_{210})}{2} \right]}{(V_{040} - V_{030} - 1)C_x} \]
\[
\alpha_3 = \frac{\left[ \frac{V_{040}^0 (2\rho_{yx} C_y - V_{201}) - \frac{V_{030}^0 (2V_{120} C_y - V_{200} + 1)}{2} \right]}{(V_{040} - V_{030} - 1)C_x} \]
\[
\alpha_4 = \frac{\left[ \frac{(2V_{120} C_y - V_{202} + 1) - \frac{V_{030}^0 (2\rho_{yx} C_y - V_{201})}{2} \right]}{(V_{040} - V_{030} - 1)C_x} \]

6.2.4 Suggested estimators

6.2.4 (a) A generalised class of exponential ratio-type estimator

Motivated by Singh and Vishwakarma (2007), we have suggested a modified class of estimator, when population mean \( \bar{Z} \) and variance \( \sigma_z^2 \) are known, is given as:

\[
\tau_1 = \hat{C}_y w_0 + \hat{C}_y f_1(x, \bar{Z}, \bar{Z}^*, \bar{Z}^*) w_1 + \hat{C}_y f_2(s_x^2, s_z^2, s_z^2, s_z^2) w_2 \] (6.36)

where

\[
f_1(x, \bar{Z}, \bar{Z}^*, \bar{Z}^*) = \hat{C}_y^2 \exp \left[ \frac{\bar{X}^*}{a\bar{Z} + b} (a\bar{Z} + b) - \bar{X} \right] \] (6.37)
Expanding and arranging equations (6.39) and (6.40) and putting them in (6.36), we optimised for minimum variance under the condition

Here ‘\(a\)’ and ‘\(b\)’ are constants and \(w_0, w_1\) and \(w_2\) are suitably chosen scalars, which are to be determined.

\(f_1\) and \(f_2\) in terms of \(e_i\)’s are written as

\[
f_1 = \frac{S_y(1 + e_0)^{1/2}}{\overline{Y}(1 + e_{01})} \exp \left[ \frac{(1 + e_i)\overline{X}}{a(1 + e_i) + b} (a\overline{Z} + b) - (1 + e_i)\overline{X} \right] \frac{(1 + e_i)\overline{X}}{a(1 + e_i) + b} (a\overline{Z} + b) + (1 + e_i)\overline{X} \right] \tag{6.39}
\]

\[
f_2 = \frac{S_y(1 + e_0)^{1/2}}{\overline{Y}(1 + e_{01})} \exp \left[ \frac{(1 + e_i)\sigma_x^2}{a(1 + e_i)\sigma_x^2 + b} (a\overline{Z} + b) - (1 + e_i)\sigma_x^2 \right] \frac{(1 + e_i)\sigma_x^2}{a(1 + e_i)\sigma_x^2 + b} (a\overline{Z} + b) + (1 + e_i)\sigma_x^2 \right] \tag{6.40}
\]

Expanding and arranging equations (6.39) and (6.40) and putting them in (6.36), we have

\[
\tau_i = C_y \left[ 1 - e_0 + e_0^2 + \frac{e_0^2}{8} - \frac{e_{01}e_0}{2} \right] + w_i C_y \left[ \frac{e_1}{2} - \frac{e_1}{2} - \frac{\theta e_1}{2} + \frac{\theta e_1}{2} + \frac{\theta e_1}{2} + \frac{\theta e_1}{2} + \frac{\theta e_1}{2} - \frac{\theta e_1 e_2}{2} \right] \\
+ w_2 C_y \left[ \frac{e_2}{2} - \frac{e_3}{2} - \frac{\theta e_4}{2} + \frac{\theta e_4}{2} + \frac{\theta e_4}{2} + \frac{\theta e_4}{2} + \frac{\theta e_4}{2} - \frac{\theta e_4 e_3}{2} \right] + \overline{Y}(1 + e_{01}) \left[ \frac{e_1}{2} - \frac{e_1}{2} - \frac{\theta e_1}{2} + \frac{\theta e_1}{2} + \frac{\theta e_1}{2} + \frac{\theta e_1}{2} + \frac{\theta e_1}{2} - \frac{\theta e_1 e_2}{2} \right] \\
\] 

Subtracting \(C_y\) from both sides of equation (6.41) and then taking expectations, the bias of \(\tau_i\) is given as

\[
B(\tau_i) = B(\hat{C}_y) + C_y w_i \left[ \frac{3C_y^2 f}{8} + \frac{3C_y^2}{8n_1} + \frac{C_y^2 f}{2} - \frac{\nu_{201} C_y f}{2} + \frac{\theta C'}{2n_1} - \frac{\theta \nu_{201} C_y}{2n_1} \right] 
\]

113
Differentiating (6.44) with respect to \( \theta \), equating it to zero, we get

\[
\text{MSE}(\tau_1) = \text{MSE}(\hat{C}_y) + C_y^2 \left[ w_1^2 \left( C_y^2 + \frac{\theta^2 C_z^2}{4n_1} \right) + w_2^2 \left( \frac{\nabla_{004}f}{2} + \frac{\theta \nabla_{004}f}{4n_1} \right) \right] + 2w_1 \left( \frac{CC_z^2 f}{2} \right) + \frac{\nabla_{210}C_x f}{4n_1} - \frac{\theta \nabla_{20}C_z}{4n_1} \left[ \frac{C_z^2}{2n_1} \right] + 2w_1w_2 \left( \frac{\nabla_{003}C_x f}{4n_1} + \theta \nabla_{003}C_x \right)
\]

Or

\[
\text{MSE}(\tau_1) = \text{MSE}(\hat{C}_y) + C_y^2 \left[ w_1^2 N_{11} + w_2^2 N_{12} + 2w_1 N_{13} + 2w_2 N_{14} + 2w_1w_2 N_{15} \right]
\]  

(6.44)

Differentiating (6.44) with respect to \( (w_1, w_2) \) and equating it to zero, we get the optimum value of \( w_1 \) and \( w_2 \) as

\[
w_{1\text{(opt)}} = \frac{N_{14}N_{15} - N_{12}N_{13}}{N_{11}N_{12} - N_{15}^2} \quad \text{and} \quad w_{2\text{(opt)}} = \frac{N_{13}N_{15} - N_{11}N_{14}}{N_{11}N_{12} - N_{15}^2}
\]

Substituting these optimum values in equation (6.44), we get the minimum MSE of the proposed estimator \( \tau_1 \).

where,

\[
N_{11} = \frac{C_y^2 f}{4} + \frac{\theta^2 C_z^2}{4n_1}, \quad N_{12} = \frac{\nabla_{004}f}{4} + \frac{\theta^2 \nabla_{004}f}{4n_1}, \quad N_{13} = \frac{CC_z^2 f}{2} + \frac{\nabla_{210}C_x f}{4n_1} - \frac{\theta \nabla_{20}C_z}{4n_1} + \frac{C_z^2}{2n_1}
\]
\[ N_{14} = \frac{V_{120} C_y}{2} f - \frac{V_{220} C_y}{4} f + \frac{\theta_1 V_{102} C_y}{2n_1} - \frac{\theta_1 V_{202} C_y}{4n_1} \], \[ N_{15} = \frac{V_{010} C_x}{4} f + \theta \theta_1 V_{000} C_z \]

### 6.2.4 (b) A generalised class of difference ratio cum exponential ratio-type estimator

We suggest a modified class of estimator, when population mean \( \bar{Z} \) and variance \( \sigma^2_z \) are known, is given as:

\[
\tau_2 = \hat{C} \left[ h_0 + h_1 \left( \frac{x^*}{x} \right)^{m_1} \left( \frac{Z^*}{Z} \right)^{m_2} + h_2 \exp \left( \frac{s_x^2 - \sigma_x^2}{s_x^2 + \sigma_x^2} \right)^{l_1} \exp \left( \frac{s_z^2 - \sigma_z^2}{s_z^2 + \sigma_z^2} \right)^{l_2} \right] \tag{6.45}
\]

where \( m_1, m_2, l_1, l_2 \) are suitably chosen constants and \( h_0, h_1, h_2 \) are suitably chosen scalars, which are to be optimised for minimum variance under the condition

\[
h_0 + h_1 + h_2 = 1 \tag{6.46}
\]

Equation (6.45) in terms of \( e_i \)'s is written as

\[
= \frac{S_z}{Y} \frac{(1 + e_0)^{l/2}}{(1 + e_0)} \left[ h_0 + h_1 \left( 1 + e_i \right)^{m_1} (1 + e_1)^{-m_1} (1 + e_2)^{-m_2} \right] + h_2 \exp \left\{ l_1 \frac{e_1^* - e_1}{2} + \left( 1 + \frac{e_3^* + e_3}{2} \right)^{-1} \right\} \exp \left\{ l_2 \frac{e_4^* - e_4}{2} + \left( 1 + \frac{e_5^* + e_5}{2} \right)^{-1} \right\} \tag{6.47}
\]

Here we assume that \( |e_0|, |e_1|, |e_2|, \ldots, |e_i|, \ldots, |e_4|, |e_5| < 1 \), so that the terms which is in the form \( (1 + e_i)^k \) and \( (1 + e_i^*)^k \) for all \( i = 0, 1, 2, 3, 4 \) and \( k = 1/2, 1, -1, m_1, m_2 \) are expandable. By expanding the right hand side of equation (6.47) and neglecting the terms of \( e_i \)'s having power greater than two, we have

\[
\tau_2 = C \left[ 1 - e_0 + e_0 + e_0^2 \frac{e_0}{2} - \frac{e_0}{2} e_0 + h_1 \left( m_1 e_1^* - m_1 e_1 - m_2 e_2 + m_2 e_1 e_2 - m_1 m_2 e_1 e_2 - m_1 e_1^* e_1 - m_2 e_2^* e_2 \right) - m_1^2 e_1 e_1^* + m_2 (m_2 + l) e_2^* - m_1 (m_1 + l) e_1^* + m_1 (m_1 - l) e_1^* + m_2 e_1^* e_2 + m_1 e_1 e_1 - m_1 e_0^* e_1 - m_2 e_0^* e_2 \right] + h_2 \left[ \frac{1}{4} \left( l_1 e_4^* - e_4 \right)^2 + \frac{1}{8} \left( l_2 e_4^* - e_4 \right)^2 + \frac{1}{4} \left( l_1 e_3^* - e_3 \right) + \frac{1}{2} l_2 \left( e_3^* - e_3 \right) \right] \tag{6.47}
\]

115
Subtract $C_y$ from both sides of equation (6.48) and taking expectation, the bias of $\tau_2$ is respectively given as

$$B(\tau_2) = B(\hat{\tau}_2) + C_y \left[ h_1 \left( \frac{m_1^2 C_x^2}{2 n_1} + \frac{m_1^2 C_x^2 f}{2 n_1} + \frac{m_2^2 C_x^2}{2 n_1} + \frac{m^2 C_x^2 f}{n_1} \right) - \frac{m_2 V_{200} C_x}{2 n_1} - \frac{m_1 V_{210} C_x}{2 n_1} \right] + h_2 \left[ \frac{l_2^2 V_{004}^0}{8 n_1} + \frac{l_2^2 V_{004}^0}{8 n_1} + \frac{l_2^2 f V_{004}^0}{8} - \frac{l_2^2 V_{100} C_y}{2 n_1} \right. \\
\left. + \frac{l_1 V_{120} f C_y}{2} + \frac{l_2 V_{320}^0}{4 n_1} - \frac{l_1 V_{220}^0}{4 n_1} \right)$$

(6.49)

Subtracting $C_y$ and squaring both sides of equation (6.48), we have

$$\left( \tau_2 - C_y \right)^2 = C_y^2 \left[ e_{01}^2 \frac{e_0^2}{4} - e_0 e_0 + h_1 \left( m_1^2 e_1^2 + m_1^2 e_1^2 + m_1^2 e_1^2 - 2m_1 e_1 e_1 - 2m_1 m_2 e_1 e_1 + 2m_1 m_2 e_1 e_1 \right) \right. \\
\left. + h_2 \left[ \frac{l_2^2 e_4^2}{4} + \frac{l_2^2 e_2^2}{4} + \frac{l_2^2 e_2^2}{4} + 2 l_2 \frac{(e_1 e_1 - e_1 e_1)}{4} - 2 l_2 \frac{(e_1 e_1 - e_1 e_1)}{4} \right) \right. \\
\left. + m_1 e_1 e_1 + m_2 e_1 e_1 - m_2 e_1 e_1 + m_2 e_0 e_0 \right] + 2 h_1 \left[ \frac{l_1 e_0 e_4}{2} + \frac{l_1 e_0 e_3}{2} - \frac{l_1 e_0 e_3}{2} - \frac{l_1 e_0 e_3}{2} - \frac{l_1 e_0 e_3}{2} - \frac{l_1 e_0 e_3}{2} \right. \\
\left. + 2 h_2 \left[ \frac{m_1 e_4 e_4}{2} + \frac{m_1 e_4 e_4}{2} + \frac{m_1 e_4 e_4}{2} - \frac{m_1 e_4 e_4}{2} - \frac{m_1 e_4 e_4}{2} - \frac{m_1 e_4 e_4}{2} \\
\left. - m_2 e_0 e_1 + m_2 e_0 e_1 \right] \right) \right)$$

(6.50)

Taking expectation on both sides of equation (6.50), MSE of $\tau_2$ is given as

$$\text{MSE}(\tau_2) = \text{MSE}(\hat{\tau}_2) + C_y^2 \left[ h_1 \left( \frac{m_1^2 C_x^2 f}{2 n_1} + \frac{m_2^2 C_x^2}{2 n_1} \right) + h_2 \left[ \frac{l_1^2 f V_{004}^0}{4 n_1} + \frac{l_2^2 V_{004}^0}{4 n_1} \right] + 2 h_1 \left( m_1 C_x^2 f \right) \right. \\
\left. - \frac{m_2 V_{200} C_x}{2 n_1} - \frac{m_1 V_{210} C_x}{2 n_1} \right] + h_2 \left[ \frac{l_1 V_{120} f C_y}{2} + \frac{l_2 V_{202}^0}{4 n_1} - \frac{l_1 V_{220}^0}{4 n_1} \right. \right. \\
\left. \left. + 2 h_2 \left[ \frac{l_1 m_1 V_{004} C_x}{2 n_1} - \frac{l_2 m_2 V_{004} C_x}{2 n_1} \right] \right] \right)$$

(6.51)

Or

$$\text{MSE}(\tau_2) = \text{MSE}(\hat{\tau}_2) + C_y^2 \left[ h_1^2 N_{21} + \frac{h_1 N_{22}}{2} + 2 h_1 N_{23} + 2 h_2 N_{24} + 2 h_1 h_2 N_{25} \right]$$

(6.51)
Differentiating equation (6.51) with respect to \((h_1, h_2)\) and equating it to zero, we get the optimum values of \(h_1\) and \(h_2\) as

\[
h_{1\text{opt}} = \frac{N_{24}N_{25} - N_{22}N_{23}}{N_{21}N_{22} - N_{25}^2} \quad \text{and} \quad h_{2\text{opt}} = \frac{N_{23}N_{25} - N_{21}N_{24}}{N_{21}N_{22} - N_{25}^2}
\]

where,

\[
N_{21} = m_1C_x^2f + \frac{m_2C_x^2}{n_1}, \quad N_{22} = \frac{l_1^2\nu_{04}^O}{4} + \frac{l_2^2\nu_{04}^O}{4n_4},
\]

\[
N_{23} = m_iCC_x^2f - \frac{m_2\nu_{201}^O}{2n_1} - \frac{m_1\nu_{20}^O C_x f}{n_1} + \frac{m_2C_C^2}{n_1},
\]

\[
N_{24} = \frac{l_1\nu_{210}^O C_y}{2} + \frac{l_2\nu_{202}^O}{4n_1} - \frac{l_1\nu_{120}^O}{2n_1} - \frac{l_2\nu_{120}^O}{2n_1}, \quad N_{25} = \frac{l_1m_i\nu_{030}^O C_x f}{2} - \frac{l_2m_i\nu_{003}^O C_z}{2n_1}.
\]

Substituting these optimum values in equation (6.49) and (6.51), we get the minimum bias and MSE of suggested estimator \(\tau_2\).

### 6.2.4 (c) A generalised chain exponential ratio-type estimator

When population mean \(\bar{Z}\) and variance \(\sigma_z^2\) are known, motivated by Singh and Singh (2002), we have suggested a generalised chain exponential ratio-type estimator given as

\[
\tau_3 = \hat{C}_y \exp \left[ \frac{x_i - \bar{x}}{\bar{x} + \bar{x}} \right] \exp \left[ \frac{s_x^2 - s_x^2}{s_x^2 + s_x^2} \right] \exp \left[ \frac{\bar{Z} - Z^*}{\bar{Z} + Z^*} \right] \exp \left[ \frac{\sigma_x^2 - s_x^2}{\sigma_x^2 + s_x^2} \right]
\]

where, \(\lambda_i\)'s (i=1,2,3,4) are suitably chosen constants.

Equation (6.52) in terms of \(e_i\)'s is written as

\[
\tau_3 = \frac{S_y (1 + e_{01})}{2} \exp \left[ \frac{e_1^* - e_1}{2 + e_1^* + e_1} \right] \exp \left[ \frac{e_2^* - e_2}{2 + e_2^* + e_2} \right] \exp \left[ \frac{e_3^* - e_3}{2 + e_3^* + e_3} \right] \exp \left[ \frac{e_4^* - e_4}{2 + e_4^* + e_4} \right]
\]

Expanding and arranging equation (6.53), we have

\[
\tau_3 = \hat{C}_y \left[ 1 - \frac{\lambda_3 e_1^*}{2} - \frac{\lambda_4 e_1^*}{2} + \frac{\lambda_1(e_1^* - e_1)}{2} + \frac{\lambda_2(e_1^* - e_1)}{2} + \frac{\lambda_3 e_2^*}{2} + \frac{\lambda_4 e_2^*}{2} + \frac{\lambda_3 \lambda_4 e_1^* e_2^*}{8} + \frac{\lambda_3 \lambda_4 e_1^* e_2^*}{4} - \frac{\lambda_3 \lambda_4 (e_1^* e_1^* - e_1 e_1^*)}{4} \right]
\]

117
Subtracting \( C_y \) from both sides of equation (6.54) and taking expectation, the bias of \( \tau_3 \) is given as

\[
B(\tau_3) = B(\hat{C}_y) + C_y \left[ \frac{\lambda_1 C_x f}{2} \left\{ \lambda_1 C_x + \frac{C C_x - \frac{V_{210}}{2}}{4} \right\} + \frac{\lambda_2 f}{2} \left\{ \frac{\lambda_2 V_{040}^O}{4} + \frac{V_{120} C_y - \frac{V_{202}^O}{2}}{2} \right\} + \frac{\lambda_4}{4n_1} \right]
\]

Subtracting \( C_y \) from both sides of equation (6.54) and squaring, we have

\[
(\tau_3 - C_y)^2 = C_y^2 \left[ \frac{e_0}{2} - e_0 + \frac{\lambda_1 (e_1^* - e_1)}{2} + \frac{\lambda_2 (e_3^* - e_3)}{2} - \frac{\lambda_4 e_4^*}{2} \right]^2
\]

Taking expectation and arranging both sides of equation (6.56), MSE of \( \tau_3 \) is given as

\[
\text{MSE}(\tau_3) = \text{MSE}(\hat{C}_y) + C_y^2 \left[ \frac{\lambda_1^2 C_x^2 f}{4} + \frac{\lambda_1}{4} \left( CC_x^2 f - \frac{V_{210}^O}{2} C_x^2 f \right) + \frac{\lambda_2^2 C_x^2 f}{4n_1} + \frac{\lambda_4}{4n_1} \right]
\]

Or

\[
\text{MSE}(\tau_3) = \text{MSE}(\hat{C}_y) + C_y^2 \left[ \lambda_1 N_{31} + \lambda_1 N_{32} + \lambda_2 N_{33} + \lambda_2 N_{34} + \lambda_4 N_{35} + \lambda_3 N_{36} - \lambda_3 N_{37} \right]
\]

118
Differentiating (6.57) w.r.t \((\lambda_1, \lambda_2, \lambda_3, \lambda_4)\) and equating them to zero, we get the optimum values of \(\lambda_1, \lambda_2, \lambda_3\) and \(\lambda_4\) as

\[
\lambda_{1(\text{opt})} = \frac{N_{34}N_{35} - 2N_{32}N_{33}}{4N_{31}N_{33} - N_{35}^2}, \quad \lambda_{2(\text{opt})} = \frac{N_{32}N_{35} - 2N_{31}N_{34}}{4N_{31}N_{33} - N_{35}^2}
\]

\[
\lambda_{3(\text{opt})} = \frac{2N_{37}N_{38} - 2N_{39}N_{310}}{4N_{36}N_{38} - N_{310}^2} \quad \text{and} \quad \lambda_{4(\text{opt})} = \frac{2N_{36}N_{39} - 2N_{37}N_{310}}{4N_{36}N_{38} - N_{310}^2}
\]

where, \(N_{31} = \frac{C_x^2f}{4}, \quad N_{32} = CC_x^2f - \frac{z_{210}^4}{2}C_x^2 f, \quad N_{33} = \frac{\nabla_{210}^4f}{4}, \quad N_{34} = \nabla_{120}C_yf - \frac{\nabla_{220}^4f}{2}f, \quad N_{35} = \frac{C_xz_{204}f}{2}, \quad N_{36} = \frac{C_z^2}{4n_1}, \quad N_{37} = \left\{ \frac{\nabla_{201}^{201}C_y}{2} - C^*C_z^2 \right\}, \quad N_{38} = \frac{\nabla_{210}^{201}C_y}{4n_1}, \quad N_{39} = \frac{\nabla_{204}}{2n_1}C_y \frac{\nabla_{001}}{2n_1}.
\]

Substituting these optimum values of \(\lambda_1, \lambda_2, \lambda_3\) and \(\lambda_4\) in (6.55) and (6.57), we get the minimum bias and MSE of proposed estimator \(\tau_1\).

### 6.2.5 Determination of \(n, n_1\) and \(\lambda\) for fixed cost and fixed variance

#### 6.2.5 (a) The minimum MSE of proposed estimator \(\tau_1\) for the fixed cost \(C\)

Suppose \(C_0\) is the overhead cost, \(C_1\) is the cost of selection and processing of a single unit in the first phase and \(C_2\) is the cost of a single unit in the second phase, then the total cost function \(C\) for selecting \(n_1\) units in the first phase and \(n\) units in the second phase is given by

\[
C = C_0 + n_1C_1 + nC_2
\]  

(6.58)

The minimum MSE of proposed estimator \(\tau_1\), is given as

\[
\min \text{MSE}(\tau_1) = \frac{\psi_1}{n} + \left( \frac{1}{n} - \frac{1}{n_1} \right)\psi_2 + \frac{\psi_3}{n_1} = V_0
\]

(6.59)

where, \(\psi_1 = C_y^2 \left[ C_y^2 - \nabla_{3000} C_y + \frac{\nabla_{400}^{400}}{4} \right]\)
The Lagrange function is

\[
\psi_2 = C_y^2 \left[ \lambda_{\lambda_1(\text{opt})} C_z^2 + \lambda_{\lambda_2(\text{opt})} \left( C^2_z f - \frac{\nabla_{210}}{2} C_z^2 \right) + \frac{\lambda_{\lambda_1(\text{opt})} \lambda_{\lambda_2(\text{opt})} C_x \nabla_{030}}{2} + \frac{\lambda_2^2 \nabla_{040} f}{4} \right] \\
+ \lambda_2 \left( \nabla_{120} C_y f - \frac{\nabla_{020}^0 f}{2} \right)
\]

\[
\psi_3 = C_y^2 \left[ \frac{\lambda_{\lambda_3(\text{opt})} C_z^2}{4} - \lambda_{\lambda_3(\text{opt})} \left( \frac{\nabla_{201} C_z}{2} - C^2_z \right) + \frac{\lambda_3^2 \nabla_{040}^0}{4n_1} - \lambda_{\lambda_4(\text{opt})} \left\{ \frac{\nabla_{020}^0}{2} - \nabla_{102} C_y \right\} \right] \\
+ \frac{\lambda_3(\text{opt}) \lambda_{\lambda_4(\text{opt})}}{2} \nabla_{003} C_z
\]

The Lagrange function is

\[
L = \frac{\psi_1}{n} + \left( \frac{1}{n} - \frac{1}{n_1} \right) \psi_2 + \frac{\psi_3}{n_1} - \lambda [C - C_0 - n_1 C_1 - n C_2] \quad (6.60)
\]

On differentiating (6.60) with respect to \( n_1 \) and equating to zero, we have

\[
\frac{\partial L}{\partial n_1} = \frac{\psi_2}{n_1^2} - \frac{\psi_3}{n_1^2} + \lambda C_1 = 0
\]

which implies that

\[
n_1 = \frac{\sqrt{\psi_3 - \psi_2}}{\sqrt{C_1 \lambda}} \quad (6.61)
\]

On differentiating (6.60) with respect to \( n \) and equating to zero, we have

\[
\frac{\partial L}{\partial n} = -\frac{\psi_1}{n^2} - \frac{\psi_2}{n^2} + \lambda C_2 = 0
\]

which implies that

\[
n = \frac{\sqrt{\psi_1 + \psi_2}}{\sqrt{C_2 \lambda}} \quad (6.62)
\]

On substituting these values of \( n_1 \) and \( n \) in equation (6.60), we have

\[
\lambda = \frac{\sqrt{(\psi_3 - \psi_2) C_1} + \sqrt{(\psi_1 + \psi_2) C_2}}{(C - C_0)} = \frac{D_0}{(C - C_0)} \quad (6.63)
\]

On substituting this value of \( \lambda \) in (6.61) and (6.62), we obtain the optimum sample sizes as

\[
n_1 = \frac{(C - C_0) \sqrt{\psi_3 - \psi_2}}{D_0 \sqrt{\lambda}} \quad (6.64)
\]

and

120
\[ n = \frac{(C - C_0)\sqrt{\psi_1 + \psi_2}}{D_0 \sqrt{\lambda}} \]  

(6.65)

On substituting (6.64), (6.65) in (6.59), we obtain the minimum MSE of proposed estimator \( \tau_3 \) for fixed cost \( C \).

6.2.5 (b) The minimum cost \( C \) for the fixed MSE of proposed estimator \( \tau_3 \)

Consider Lagrange function

\[ L = C - C_0 - n_1C_1 - n_2C_2 - \lambda \left[ \frac{\psi_1}{n} + \left( \frac{1}{n} - \frac{1}{n_1} \right)\psi_2 + \frac{\psi_3}{n_1} \right] \]  

(6.66)

On differentiating (6.66) with respect to \( n_1 \) and equating to zero, we have

\[ \frac{\partial L}{\partial n_1} = -C_1 - \lambda \frac{\psi_2}{n_1} + \frac{\lambda \psi_3}{n_1^2} = 0 \]

which implies that

\[ n_1 = \frac{\sqrt{\lambda (\psi_3 - \psi_2)}}{\sqrt{C_1}} \]  

(6.67)

On differentiating (6.66) with respect to \( n \) and equating to zero, we have

\[ \frac{\partial L}{\partial n} = -C_2 + \frac{\lambda \psi_1}{n^2} + \frac{\lambda \psi_2}{n^2} = 0 \]

which implies that

\[ n = \frac{\sqrt{\lambda (\psi_1 + \psi_2)}}{\sqrt{C_2}} \]  

(6.68)

On substituting these values of \( n_1 \) and \( n \) in equation (6.66), we have

\[ \lambda = \frac{\sqrt{(\psi_3 - \psi_2)C_1 + \sqrt{(\psi_1 + \psi_2)C_2}}}{V_0} = \frac{D_1}{V_0} \]  

(6.69)

On substituting this value of \( \lambda \) in (6.67) and (6.68), we obtain the optimum sample sizes as

\[ n_1 = \frac{D_1 \sqrt{\psi_3 - \psi_2}}{V_0 \sqrt{C_1}} \]  

(6.70)

and

\[ n = \frac{D_1 \sqrt{\psi_1 + \psi_2}}{V_0 \sqrt{C_2}} \]  

(6.71)

On substituting (6.70) and (6.71) in \( C = C_0 + n_1C_1 + nC_2 \), we get the minimum cost \( C \) for the fixed MSE of proposed estimator \( \tau_3 \).

Note: If we have information about money spent in different phases, we can have the minimum cost and minimum MSE for \( \tau_3 \).
6.2.6 Efficiency Comparison

To illustrate the performance of various estimators of $C_y$, we consider the data given in (Murthy, 1967, p-226). The variables are

- $y$: output
- $x$: Numbers of workers
- $z$: Fixed capital.

The values of required parameters of the population are:

<table>
<thead>
<tr>
<th>Table 6.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N=80$</td>
</tr>
<tr>
<td>$\bar{Z}=1126$</td>
</tr>
<tr>
<td>$\bar{\nu}_{030}=1.295$</td>
</tr>
<tr>
<td>$\bar{\nu}_{210}=0.5475$</td>
</tr>
</tbody>
</table>

The percentage relative estimator efficiency (PRE) of an estimator is defined as

$$\text{PRE}(\cdot) = \frac{\text{MSE}(\hat{\theta})}{\nu \text{MSE}(\cdot)} \times 100$$

The percent relative efficiencies of different estimators are given in Table 6.4

<table>
<thead>
<tr>
<th>Table 6.4 MSE and PRE of the estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimator</td>
</tr>
<tr>
<td>$\tau_{RM}$</td>
</tr>
<tr>
<td>$\tau_{RV}$</td>
</tr>
<tr>
<td>$\tau_{ERM}$</td>
</tr>
<tr>
<td>$\tau_{ERV}$</td>
</tr>
<tr>
<td>$\tau_{S}$</td>
</tr>
<tr>
<td>$\tau_1$</td>
</tr>
<tr>
<td>$\tau_2$</td>
</tr>
<tr>
<td>$\tau_3$</td>
</tr>
</tbody>
</table>
6.2.7 Conclusion

From Table 6.4 the efficiency of the proposed class $\tau_3$ has been investigated by comparing different estimators with the Singh and Singh (2002) estimator and we conclude that PRE of $\tau_3$ is approximately equal to the PRE of Singh and Singh estimator and surpasses all other.