

# Efficient Classes of Modified Ratio and Product Estimators for Estimating Population Mean under Two-Phase Sampling

Alok Kumar Shukla

Department of Statistics,

D.A-V. College, Kanpur-208001, U.P., INDIA

**Abstract:-** In this Scripture, a ratio and product type exponential estimators have been suggested for improved estimation of population mean in two-phase random sampling. The biases and the mean squared errors (MSE) of the suggested estimators are derived up to the first order of approximation. The optimum values of the characterizing scalars are also obtained. The minimum values of the MSEs of the suggested estimators are obtained for these optimum values of the characterizing constants. A theoretical comparison has been made with the existing competing estimators and the efficiency conditions are verified using two natural populations.

**Keywords-** Ratio estimator, Product estimator, Efficiency, Mean squared errors

## 1. INTRODUCTION

The most commensurate estimator for estimating the population mean is the sample mean that is the corresponding statistic. Unfortunately the sampling distribution of sample mean is widely dispersed round the true population mean. Thus its sampling variance is reasonably large. Thus we search for even biased estimator but having sampling distribution close to population mean. This is achieved through the use of auxiliary variable, which is highly positively or negatively associated with the main variable under study. Ratio and product method of estimation are applied when main and auxiliary variables are highly positively and negatively correlated respectively. Many times we come across the problems when the

population mean of the auxiliary variable is not known. In such situation, two-phase sampling is used when a larger sample is drawn to estimate the population mean of the auxiliary variable and then a small sample of appropriate size is drawn to estimate the population mean of the study variable.

Let the population under consideration consists of  $N$  distinct and identifiable units. A sample of size  $n_1$  is drawn without replacement to estimate the population mean of the auxiliary variable and then a sub-sample of size  $n$  is drawn without replacement from  $n_1$  units in two ways:

**Case I:** When the second phase sample of size  $n$  is a subsample of the first phase of size  $n_1$ .

**Case II:** When the second phase sample of size  $n$  is drawn independently of the first phase sample of size  $n_1$ .

## 2. REVIEW OF EXISTING ESTIMATORS

Various authors used auxiliary information and suggested ratio and product type estimators for enhanced estimation of population mean in two-phase random sampling. The following Table-1 represents various estimators of population mean under two-phase random sampling for both the cases along with their mean squared error up to the first order of approximation.

Table-1: Various estimators and their MSE for two different cases

| S. No. | Estimator   | MSE for Case-I  | MSE for Case-II   |
|--------|---|---|---|
| 1.     | $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$<br>Sample mean estimator | $\bar{Y}^2 \frac{1-f}{n} C_y^2$   | $\bar{Y}^2 \frac{1-f}{n} C_y^2$   |
| 2.     | $\bar{y}_R^d$<br>Usual Ratio Estimator                            | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1-f^*}{n} C_x^2 (1-2C) \right]$                         | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + f^{**} C_x^2 - 2 \frac{1-f}{n} C C_x^2 \right]$               |
| 3.     | $\bar{y}_R^{*d}$<br>Kumar and Bahl (2006) Estimator               | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + g' \frac{1-f^*}{n} C_x^2 (g'-2C) \right]$                     | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + g' C_x^2 \left( g' f^{**} - 2C \frac{1-f}{n} \right) \right]$ |
| 4.     | $\bar{y}_{Re}^d$<br>Singh and Vishwakarma (2007) Estimator        | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1-f^*}{n} C_x^2 \left( \frac{1}{4} - C \right) \right]$ | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1}{4} f^{**} C_x^2 - \frac{1-f}{n} C C_x^2 \right]$     |

|    |  |   |   |
|----|--|---|---|
| 5. | $\bar{y}_{Re}^{*d}$<br>Kalita and Singh<br>(2013) Estimator      | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + g' \frac{1-f^*}{n} C_x^2 \left( \frac{1}{4} g' - C \right) \right]$ | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1}{4} g'^2 f^{**} C_x^2 - g' \frac{1-f}{n} C C_x^2 \right]$ |
| 6. | $\bar{y}_P$<br>Usual Product<br>Estimator                        | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1-f^*}{n} C_x^2 (1+2C) \right]$                               | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + f^{**} C_x^2 + 2 \frac{1-f}{n} C C_x^2 \right]$                   |
| 7. | $\bar{y}_P^{*d}$<br>Singh and Chaudhury<br>(2012) Estimator      | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + g' \frac{1-f^*}{n} C_x^2 (g'+2C) \right]$                           | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + g' C_x^2 \left( g' f^{**} + 2C \frac{1-f}{n} \right) \right]$     |
| 8. | $\bar{y}_{Pe}^d$<br>Singh and<br>Vishwakarma (2007)<br>Estimator | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1-f^*}{n} C_x^2 \left( \frac{1}{4} + C \right) \right]$       | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1}{4} f^{**} C_x^2 + \frac{1-f}{n} C C_x^2 \right]$         |
| 9. | $\bar{y}_{Pe}^{*d}$<br>Kalita and Singh<br>(2013) Estimator      | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + g' \frac{1-f^*}{n} C_x^2 \left( \frac{1}{4} g' + C \right) \right]$ | $\bar{Y}^2 \left[ \frac{1-f}{n} C_y^2 + \frac{1}{4} g'^2 f^{**} C_x^2 + g' \frac{1-f}{n} C C_x^2 \right]$ |

Where,

$$S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2, \quad S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2, \quad S_{yx} = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})(x_i - \bar{X}), \quad \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i,$$

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \quad C_y = \frac{S_y}{\bar{Y}}, \quad C_x = \frac{S_x}{\bar{X}}, \quad C = \rho_{yx} \frac{C_y}{C_x}, \quad \rho_{yx} = \frac{S_{yx}}{S_y S_x}, \quad f = \frac{n}{N}, \quad f_1 = \frac{n_1}{N}, \quad f^* = \frac{n}{n_1},$$

$$f^{**} = \frac{1-f}{n} + \frac{1-f_1}{n_1}, \quad g' = \frac{n}{n_1 - n}.$$

### 3. PROPOSED ESTIMATORS

Under this section, an exponential dual to ratio and product-type estimators, respectively, are suggested as follows:

$$\tau_{Re} = \bar{y} \left\{ \alpha + (1-\alpha) \exp \left[ \frac{\bar{x}^{*d} - \bar{x}_1}{\bar{x}^{*d} + \bar{x}_1} \right] \right\} \quad (1)$$

$$\tau_{Pe} = \bar{y} \left\{ \alpha + (1-\alpha) \exp \left[ \frac{\bar{x}_1 - \bar{x}^{*d}}{\bar{x}_1 + \bar{x}^{*d}} \right] \right\} \quad (2)$$

Where,  $\alpha$  is a characterizing scalar and is obtained by minimizing the mean squared error of the proposed estimators. The Bias and MSE of the proposed estimators are obtained for the following two cases.

#### Case I:

To study the large sample properties of the proposed class of estimators, we define

$$\bar{y} = \bar{Y}(1+e_0), \quad \bar{x} = \bar{X}(1+e_1) \quad \text{and} \quad \bar{x}_1 = \bar{X}(1+e_2) \quad \text{such that} \quad E(e_0) = E(e_1) = E(e_2) = 0 \quad \text{and}$$

$$E(e_0^2) = \gamma C_y^2, \quad E(e_1^2) = \gamma C_x^2, \quad E(e_2^2) = \gamma^* C_x^2, \quad E(e_0 e_1) = \gamma C C_x^2, \quad E(e_0 e_2) = \gamma^* C C_x^2, \quad E(e_1 e_2) = \gamma^* C_x,$$

$$\gamma = \left( \frac{1}{n} - \frac{1}{N} \right), \quad \gamma^* = \left( \frac{1}{n_1} - \frac{1}{N} \right), \quad \gamma^{**} = \left( \frac{1}{n} - \frac{1}{n_1} \right), \quad C = \rho_{yx} \frac{C_y}{C_x} \quad \text{and} \quad g = \frac{n}{n_1 - n}$$

Using above up to the first order of approximation, the proposed estimator may be written as,

$$\tau_{Re} = \bar{Y}(1+e_0) \left\{ \alpha + (1-\alpha) \left( 1 + \frac{1}{2} g (e_2 - e_1 - e_2^2 + e_2 e_1) + \frac{1}{4} g^2 (e_1 e_2 - e_2^2 - e_1^2) + \frac{1}{8} g^2 (e_2^2 + e_1^2) \right) \right\} \quad (3)$$

Thus we have bias of the proposed estimator as

$$B(\tau_{Re}) = \bar{Y} \left[ \frac{1}{8} g^2 \gamma^* C_x^2 - \frac{1}{8} g^2 \gamma C_x^2 - \frac{1}{2} g \gamma^{**} C C_x^2 \right] (1-\alpha) \quad (4)$$

And MSE of the proposed estimator as,

$$MSE(\tau_{Re}) = \bar{Y}^2 \left\{ \begin{aligned} &\gamma C_y^2 + g \gamma^* C_x^2 \left( \frac{1}{4} g - C \right) + \frac{\alpha^2}{4} g^2 (\gamma C_x^2 - \gamma^* C_x^2) \\ &\quad - \frac{\alpha}{2} g^2 (\gamma C_x^2 - \gamma^* C_x^2) - \alpha g (\gamma^* C C_x^2 - \gamma C C_x^2) \end{aligned} \right\} \quad (5)$$

The optimum value of the characterizing scalar  $\alpha$  is obtained by minimizing MSE in (5) using the method of maxima-minima as,

$$\begin{aligned} \alpha g (\gamma C_x^2 - \gamma^* C_x^2) - g (\gamma C_x^2 - \gamma^* C_x^2) - 2(\gamma^* C C_x^2 - \gamma C C_x^2) &= 0 \\ \alpha &= \frac{g (\gamma C_x^2 - \gamma^* C_x^2) + 2(\gamma^* C C_x^2 - \gamma C C_x^2)}{g (\gamma C_x^2 - \gamma^* C_x^2)} = \frac{A}{gB} \end{aligned} \quad (6)$$

Where  $A = g (\gamma C_x^2 - \gamma^* C_x^2) + 2(\gamma^* C C_x^2 - \gamma C C_x^2)$  and  $B = (\gamma C_x^2 - \gamma^* C_x^2)$

The minimum value of bias of the proposed estimator is obtained by putting optimum value of  $\alpha$  in (4) as,

$$B(\tau_{Re}) = \bar{Y} \left[ \frac{1}{8} g^2 \gamma^* C_x^2 - \frac{1}{8} g^2 \gamma C_x^2 - \frac{1}{2} g \gamma^* C C_x^2 \right] \left( 1 - \frac{A}{gB} \right) \quad (7)$$

Minimum value of the MSE of the proposed estimator is obtained by putting the optimum value of  $\alpha$  in (5) and thus the

$$\text{minimum MSE is given as, } MSE(\tau_{Re}) = \bar{Y}^2 \left[ \gamma C_y^2 + g \gamma^* C_x^2 \left( \frac{1}{4} g - C \right) - \frac{A^2}{4B} \right] \quad (8)$$

Similarly, Bias and MSE of proposed product estimator in equation (2)

The minimum value of bias of the proposed estimator is obtained by putting optimum value of  $\alpha$  as,

$$B(\tau_{Pe}) = \bar{Y} \left[ \frac{3}{8} g^2 \gamma C_x^2 - \frac{1}{8} g^2 \gamma^* C_x^2 + \frac{1}{2} g \gamma^* C C_x^2 \right] \left( 1 - \frac{A}{gB} \right) \quad (9)$$

Minimum value of the MSE of the proposed estimator is obtained by putting the optimum value of  $\alpha$  and thus the minimum

$$\text{MSE is given as, } MSE(\tau_{Pe}) = \bar{Y}^2 \left[ \gamma C_y^2 + g \gamma^* C_x^2 \left( \frac{1}{4} g + C \right) - \frac{A^2}{4B} \right] \quad (10)$$

Where  $A = g (\gamma C_x^2 - \gamma^* C_x^2) + 2(\gamma C C_x^2 - \gamma^* C C_x^2)$  and  $B = (\gamma C_x^2 - \gamma^* C_x^2)$

### Case II

To study the large sample properties of the proposed class of estimator, we define

$$\bar{y} = \bar{Y}(1 + e_0), \quad \bar{x} = \bar{X}(1 + e_1) \quad \text{and} \quad \bar{x}_1 = \bar{X}(1 + e_2) \quad \text{such that} \quad E(e_0) = E(e_1) = E(e_2) = 0 \quad \text{and}$$

$$E(e_0^2) = \gamma C_y^2, E(e_1^2) = \gamma C_x^2, E(e_2^2) = \gamma^* C_x^2, E(e_0 e_1) = \gamma C C_x^2, E(e_0 e_2) = 0, E(e_1 e_2) = 0, \gamma = \left( \frac{1}{n} - \frac{1}{N} \right),$$

$$\gamma^* = \left( \frac{1}{n_1} - \frac{1}{N} \right), \gamma^{**} = \left( \frac{1}{n} - \frac{1}{n_1} \right), \gamma^{***} = (\gamma + \gamma^*) C = \rho_{yx} \frac{C_y}{C_x} \quad \text{and} \quad g = \frac{n}{n_1 - n}$$

Similarly, Bias and MSE of proposed estimators in equation (1) and (2)

The minimum value of bias and MSE of the proposed estimator is obtained by putting optimum value of  $\alpha$  as,

$$B(\tau_{Re}) = \bar{Y} \left[ \frac{1}{8} g^2 \gamma^{***} C_x^2 + \frac{1}{2} g C_x^2 (\gamma^* + \gamma C) \right] \left( 1 - \frac{A}{gB} \right) \quad (11)$$

$$MSE(\tau_{Re}) = \bar{Y}^2 \left[ \gamma C_y^2 + \frac{1}{4} g^2 \gamma^{***} C_x^2 - g \gamma C C_x^2 - \frac{A^2}{4B} \right] \quad (12)$$

Where  $A = g (\gamma C_x^2 + \gamma^* C_x^2) - 2\gamma C C_x^2$  and  $B = (\gamma C_x^2 + \gamma^* C_x^2)$

Minimum value of bias and MSE of the proposed estimator is obtained by putting the optimum value of  $\alpha$  and thus the minimum MSE is given as,

$$B(\tau_{Pe}) = \bar{Y} \left[ \frac{3}{8} g^2 \gamma^{***} + \frac{1}{2} g C_x^2 (\gamma^* + \gamma C) \right] \left( 1 - \frac{A}{gB} \right) \tag{13}$$

$$MSE(\tau_{Pe}) = \bar{Y}^2 \left[ \gamma C_y^2 + \frac{1}{4} g^2 \gamma^{***} C_x^2 + g \gamma C C_x^2 - \frac{A^2}{4B} \right] \tag{14}$$

Where  $A = g(\gamma C_x^2 + \gamma^* C_x^2) + 2\gamma C C_x^2$  and  $B = (\gamma C_x^2 + \gamma^* C_x^2)$

#### 4. EFFICIENCY COMPARISON

A theoretical comparison of various estimators with the suggested estimator has been made and presented in Table-2. The efficiency conditions for both the cases under which the suggested estimator performs better than the competing one are also presented in this table.

Table-2: Efficiency comparison of proposed estimators with competing estimators

| S. No. | Estimator           | Efficiency condition for Case-I               | Efficiency condition for Case-II                 |
|--------|---------------------|---|--|
| 1.     | $\bar{y}$           | $V(\bar{y}) - MSE(\tau_{Re})_I > 0$           | $V(\bar{y}) - MSE(\tau_{Re})_{II} > 0$           |
| 2.     | $\bar{y}_R^d$       | $V(\bar{y}_R^d) - MSE(\tau_{Re})_I > 0$       | $V(\bar{y}_R^d) - MSE(\tau_{Re})_{II} > 0$       |
| 3.     | $\bar{y}_R^{*d}$    | $V(\bar{y}_R^{*d}) - MSE(\tau_{Re})_I > 0$    | $V(\bar{y}_R^{*d}) - MSE(\tau_{Re})_{II} > 0$    |
| 4.     | $\bar{y}_{Re}^d$    | $V(\bar{y}_{Re}^d) - MSE(\tau_{Re})_I > 0$    | $V(\bar{y}_{Re}^d) - MSE(\tau_{Re})_{II} > 0$    |
| 5.     | $\bar{y}_{Re}^{*d}$ | $V(\bar{y}_{Re}^{*d}) - MSE(\tau_{Re})_I > 0$ | $V(\bar{y}_{Re}^{*d}) - MSE(\tau_{Re})_{II} > 0$ |
| 6.     | $\bar{y}_P^d$       | $V(\bar{y}_P^d) - MSE(\tau_{Pe})_I > 0$       | $V(\bar{y}_P^d) - MSE(\tau_{Pe})_{II} > 0$       |
| 7.     | $\bar{y}_P^{*d}$    | $V(\bar{y}_P^{*d}) - MSE(\tau_{Pe})_I > 0$    | $V(\bar{y}_P^{*d}) - MSE(\tau_{Pe})_{II} > 0$    |
| 8.     | $\bar{y}_{Pe}^d$    | $V(\bar{y}_{Pe}^d) - MSE(\tau_{Pe})_I > 0$    | $V(\bar{y}_{Pe}^d) - MSE(\tau_{Pe})_{II} > 0$    |
| 9.     | $\bar{y}_{Pe}^{*d}$ | $V(\bar{y}_{Pe}^{*d}) - MSE(\tau_{Pe})_I > 0$ | $V(\bar{y}_{Pe}^{*d}) - MSE(\tau_{Pe})_{II} > 0$ |

#### 5. EMPIRICAL STUDY

Under this section, the theoretical efficiency conditions are verified using following two natural populations. The sources, descriptions and the parameters of these two natural populations are given below.

**Population I: Source :Murthy (1967)**

$Y$  = Output,  $X$  = Number of workers

$N = 80, n = 16, n_1 = 30, \bar{Y} = 5182.64, \rho_{yx} = 0.9150, C_y = 0.3542, C_x = 0.9484$

**Population II: Source: Kadilar and Cingi, 2006c**

$N = 200, n = 50, n_1 = 175, \bar{Y} = 500, \rho_{yx} = 0.90, C_y = 25, C_x = 2$

**Population III:- Source: Johnston 1972**

$Y$  = Mean January temperature,  $X$  = Date of flowering of a particular summer species (number of days from January 1)

$N = 10, n = 2, n_1 = 5, \bar{Y} = 42, \rho_{yx} = -0.73, C_y = 0.1303, C_x = 0.0458$

**Population IV:- Source: Johnston 1972**

$Y$  = Percentage of hives affected by disease,  $X$  = Date of flowering of a particular summer species (number of days from January 1)

$N = 10, n = 2, n_1 = 5, \bar{Y} = 52, \rho_{yx} = -0.94, C_y = 0.1562, C_x = 0.0458$

The MSE of suggested ratio estimator and various competing ratio estimators for both the cases are presented in Table-3.

Table-3: Mean squared error of proposed and competing ratio estimators

| Estimators          | Case I         |                 | Case II        |                 |
|---------------------|----------------|-----------------|----------------|-----------------|
|                     | Population I   | Population II   | Population I   | Population II   |
| $\bar{y}$           | 168488.1       | 2343750         | 168488.1       | 2343750         |
| $\bar{y}_R^d$       | 391542.8       | 2036607         | 1054187        | 2021946         |
| $\bar{y}_R^{*d}$    | 538452.2       | 2217464         | 1460110        | 2211264         |
| $\bar{y}_{Re}^d$    | 103853.5       | 2186607         | 183515.9       | 2178929         |
| $\bar{y}_{Re}^{*d}$ | 123381.2       | 2280036         | 255511.2       | 2276879         |
| $\tau_{Re}$         | <b>86201.6</b> | <b>535714.3</b> | <b>68914.6</b> | <b>531605.1</b> |

The MSE of suggested product estimator and various competing product estimators for both the cases are presented in Table-4.

Table-4: Mean squared error of proposed and competing product estimators

| Estimators          | Case I         |               | Case II        |               |
|---------------------|----------------|---------------|----------------|---------------|
|                     | Population III | Population IV | Population III | Population IV |
| $\bar{y}$           | 11.9794        | 26.3893       | 11.9794        | 26.3893       |
| $\bar{y}_P^d$       | 8.4789         | 17.1808       | 7.6820         | 14.6784       |
| $\bar{y}_P^{*d}$    | 9.3991         | 19.8722       | 8.7034         | 17.9519       |
| $\bar{y}_{Pe}^d$    | 9.9518         | 21.3597       | 9.3683         | 19.8249       |
| $\bar{y}_{Pe}^{*d}$ | 10.5661        | 22.9417       | 10.1360        | 21.8555       |
| $\tau_{Pe}$         | <b>7.1917</b>  | <b>8.9011</b> | <b>6.8725</b>  | <b>7.7352</b> |

## 6. RESULTS AND CONCLUSION

In this manuscript, an exponential ratio and product type estimators have been suggested for improved estimation of population mean under two-phase random sampling scheme. The biases and the MSEs of the suggested estimators are derived up to the first order of approximation. The optimum values of the characterizing scalars are also obtained and the minimum MSEs of the proposed estimators for these characterizing scalars are obtained. A theoretical comparison has been made with the existing competing estimators and the conditions under which suggested estimators are better than the competing estimators are obtained. These efficiency conditions are verified using two natural populations. Thus it is recommended that the suggested estimators should be used for improved estimation of population mean under double sampling scheme.

## REFERENCES

- [1] Johnston, J. (1972): Econometric methods, (2<sup>nd</sup> ed), McGraw-Hill, Tokyo.
- [2] Kadilar, C., Cingi, H. (2006c). New ratio estimators using correlation coefficient. *Interstatistics* 4:1-11.
- [3] Kalita, D., Singh, B.K. (2013). Exponential dual to ratio and dual to product-type estimators for finite population mean in double sampling. *Elixir Statistics* 59: 15458-15470.
- [4] Kumar, M., and Bhal, S. (2006). Class of dual to ratio estimators for double sampling, *Statistical Papers*, 47, 319-326
- [5] Murthy, M. N. (1967). *Sampling theory and methods*. Statistical Publishing Society: Calcutta, India.
- [6] Singh, B.K., Choudhury, S. (2012). Dual to product estimator for estimating population mean in double sampling, *Int. J. Stat. Syst.* 7, 31-39.
- [7] Singh, H. P., Vishwakarma, G. K. (2007). Modified exponential ratio and product estimators for finite population mean in double sampling. *Austrian Journal of Statistics*, 36, 3, 217-225.