Improved Family of Estimators of Population Mean in Simple Random Sampling

Rajesh Singh, Mukesh Kumar and Manoj K. Chaudhary

Department of Statistics, Banaras Hindu University (U.P.), India

Abstract: In this paper we have suggested a general procedure for estimating the population mean through defining a class of estimators. Many of the existing estimators are shown particular members of the proposed class of estimators. It has been shown to the first degree of approximation that mean squared error (MSE) of the proposed class of estimators are better than the competing ratio, regression, exponential ratio-type and other ratio-type estimators considered here. Numerical illustrations are given in support of the present study.

Key words: Auxiliary variable, ratio-type and exponential ratio-type estimators, efficiency

INTRODUCTION

The problem of estimating the population mean in the presence of an auxiliary variable has been widely discussed in finite population sampling literature. Ratio, product and difference methods of estimation are good examples in this context. Ratio method of estimation is quite effective when there is high positive correlation between study and auxiliary variables. On the other hand, if correlation is negative (high), the product method of estimation can be employed efficiently.

In recent years, a number of research papers on ratio-type, exponential ratio-type and regression-type estimators have appeared, based on different types of transformations. Some important contributions in this area are due to Singh and Tailor [12], Shabbir and Gupta [8, 9], Kadilar and Cingi [4-6], Khosnevisan et al. [7], Singh et al. [10], Singh et al. [11] and Singh et al. [13]. We begin by introducing some terminology used in the article.

Let \( Y \) be a finite population consisting of \( N \) units from which a sample of size \( n \) is to be drawn by simple random sampling without replacement (SRSWOR). Let \( y \) and \( x \) denote the study and auxiliary variables having sample means \( \bar{y} \) and \( \bar{x} \) respectively. It is desired to estimate the population mean \( \bar{Y} \) of the study variate \( y \) using information on the population mean \( \bar{X} \) of the auxiliary variate \( x \).

Let \( e_0 = \frac{(\bar{y} - \bar{Y})}{\bar{y}} \) and \( e_1 = \frac{(\bar{X} - \bar{x})}{\bar{x}} \). This gives \( E(e_0) = E(e_1) = 0 \),

\[
E(e_0^2) = \frac{(\bar{y} - \bar{Y})^2}{\bar{y}^2}, \quad \text{and} \quad E(e_1^2) = \frac{(\bar{X} - \bar{x})^2}{\bar{x}^2},
\]

where

\[ f = \frac{\bar{Y}}{\bar{y}}, \quad c_y = \frac{\bar{y}}{\bar{y}^2}, \quad c_X = \frac{\bar{x}}{\bar{x}^2}, \quad \rho = \frac{\bar{Y}\bar{x}}{\bar{y}\bar{X}}. \]

We discuss below some of the estimators available in the literature.

The most common estimator of \( \bar{Y} \) is the sample mean estimator defined as

\[
\bar{Y}_B = \bar{y}
\]

The mean square error (MSE) of \( \bar{Y}_B \) is given by

\[
\text{MSE}(\bar{Y}_B) = \frac{1 - \rho^2}{\bar{y}^2} c_y^2
\]

The usual ratio estimator for \( \bar{Y} \) is

\[
\bar{Y}_R = \bar{y} \left( \frac{\bar{x}}{\bar{X}} \right)
\]

The MSE of \( \bar{Y}_R \), to first-order of approximation is given by

\[
\text{MSE}(\bar{Y}_R) - \frac{\bar{y}^2}{\bar{X}^2} \left[ c_y^2 + c_X^2 - 2\rho c_y c_X \right]
\]

Corresponding Author: Dr. Rajesh Singh, Department of Statistics, Banaras Hindu University (U.P.), India


ESTIMATORS PROPOSED
BY OTHER AUTHORS

Singh and Tailor [12] estimator

By using the known correlation coefficient ($\rho$), Singh and Tailor [12] introduced the following ratio type estimator

$$\tilde{Y}_{ST} = \bar{Y} \left( \frac{2\pi^2}{\eta^2 + \rho^2} \right)$$

(5)

with MSE

$$\text{MSE}(\tilde{Y}_{ST}) = \frac{1 - \frac{1}{n}}{\eta} \bar{Y}^2 \left[ \frac{C_y^2 + \eta^2 C_x^2}{(n - 1)\eta \rho C_y C_x} \right]$$

(6)

where

$$\eta = \frac{X}{(X + \rho)}$$

Khosnevisan et al. [7] suggested following modified estimator

$$\tilde{Y}_{KH} = \bar{Y} \left[ \frac{aX + b}{a(\alpha X + b) + (1 - \alpha)(\alpha X + b)} \right]$$

(7)

where $a$ and $\alpha$ are suitable constants, $a$ and $b$ are either real numbers or the functions of the known parameters of the auxiliary variable, $x$, such as coefficient of variation ($C_x$), kurtosis ($\beta_2(x)$) and correlation coefficient ($\rho$).

The minimum MSE of $\tilde{Y}_{KH}$, to the first order of approximation, is given by

$$\text{MSE}(\tilde{Y}_{KH})_{\text{min}} = \frac{1 - \frac{1}{n}}{\eta} \bar{Y}^2 C_y^2 (1 - \rho^2)$$

(8)

Bahl and Tuteja [1] suggested an exponential ratio-type estimator defined as

$$\tilde{Y}_{BT} = \bar{Y} \exp \left[ \frac{X - \alpha \bar{X}}{X - \alpha \bar{X}} \right]$$

(9)

The MSE of $\tilde{Y}_{BT}$, to the first order of approximation, is given by

$$\text{MSE}(\tilde{Y}_{BT}) = \frac{1 - \frac{1}{n}}{\eta} \bar{Y}^2 \left[ \frac{C_y^2 + \frac{\rho}{4} C_y C_x}{(n - 1)\eta \rho C_y C_x} \right]$$

(10)

Singh et al. [10] suggested a modified form of Bahl and Tuteja [1] estimator

$$\tilde{Y}_{ST} = \bar{Y} \exp \left[ \frac{X - \alpha \bar{X}}{X - \alpha \bar{X}} \right]$$

(11)

The minimum MSE of $\tilde{Y}_{ST}$ to the first order of approximation is given by

$$\text{MSE}(\tilde{Y}_{ST})_{\text{min}} = \frac{1 - \frac{1}{n}}{\eta} \bar{Y}^2 C_y^2 (1 - \rho^2)$$

(12)

Note that MSE in (7) and (11) is equal to the MSE of the linear regression estimator

$$\tilde{Y}_{LR} = \bar{Y} + b(X - \bar{X})$$

Kadilar and Cingi [3] suggested following class of estimators for estimating $\bar{Y}$ -

$$\tilde{Y}_{KH} = \left[ \bar{Y} + b(X - \bar{X}) \right] Y_{i, j = 1 \ldots 5}$$

(13)

where $Y_{i, j}$ are defined as

$$\gamma^1_1 = \frac{X}{X}, \quad \gamma^2_2 = \frac{X + C_y}{X + C_y}, \quad \gamma^3_2 = \frac{X + C_y}{X + C_y}, \quad \gamma^4_2 = \frac{X + C_y}{X + C_y}$$

(14)

The MSE of $\tilde{Y}_{KH}$ (i=1, 2..5), to first order of approximation, is given by

$$\text{MSE}(\tilde{Y}_{KH}) = \frac{1 - \frac{1}{n}}{\eta} \bar{Y}^2 \left[ \gamma^2_1 C_y^2 + C_y^2 (1 - \rho^2) \right]$$

(15)

and

$$\gamma^1_2 = \frac{X}{X + C_y}, \quad \gamma^2_2 = \frac{X + C_y}{X + C_y}, \quad \gamma^3_2 = \frac{X + C_y}{X + C_y}$$

Using Singh and Tailor [12] transformation, Kadilar and Cingi [5] introduced another class of estimators, given by

$$\tilde{Y}_{KH} = \left[ \bar{Y} + b(X - \bar{X}) \right] Y_{i, j = 1 \ldots 5}$$
The MSE of $\bar{y}_{KGI}$, to first order of approximation, is given by

$$\text{MSE}(\bar{y}_{KGI}) = \frac{1-f}{n} \left( \left[ \bar{y}_1^2 - \bar{y}_2^2 \right]^{\frac{1}{2}} + \bar{y}^2 \right)$$  \hspace{1cm} (16)$$

where

$$\bar{y}_1 = \frac{\bar{x}}{\overline{X+P}}, \quad \bar{y}_2 = \frac{\bar{X}C_X + P}{\overline{X+P}}, \quad \bar{y}_3 = \frac{\overline{X+P}}{\overline{X+P}}.$$

and

$$\bar{y}_4 = \frac{\bar{x}P}{\overline{X+P}}, \quad \bar{y}_5 = \frac{\bar{x}C_X}{\overline{X+P}}.$$

Gupta and Shabbir [2] introduced the following general class of ratio-type estimators:

$$\bar{y}_{GS} = \left[ \omega_1 \bar{p} + \omega_2 (\bar{x} - \bar{r}) \right] \left( \frac{aX + b}{aX + b} \right)$$ \hspace{1cm} (17)

where $\omega_1$ and $\omega_2$ are weights.

The minimum MSE of $\bar{y}_{GS}$, to first order of approximation is given by

$$\text{MSE}(\bar{y}_{GS})_{\text{min}} = \frac{\text{MSE}(\bar{y}_{GS})}{1 + (1 - \frac{f}{n}) \bar{C}_X^2 (1 - \bar{p}^2)}$$ \hspace{1cm} (18)

**PROPOSED ESTIMATOR**

Motivated by Gupta and Shabbir [2], we propose the following general class of ratio type estimators:

$$\bar{y}_\lambda = \left[ q_1 \bar{p} + q_2 (\bar{x} - \bar{r}) \right] \left[ \lambda \left( \frac{aX + b}{aX + b} \right) + (1 - \lambda) \exp \left\{ \frac{aX - aR}{aX + aR + 2b} \right\} \right]$$ \hspace{1cm} (19)

where $q_1$ and $q_2$ are weights whose values are to be determined later and $\lambda$ is a constant.

From (19), we have

$$\bar{y}_\lambda = \left[ q_1 \bar{p} - q_2 (\bar{x} - \bar{r}) \right] \left[ \frac{(1 + q_0) - (\frac{q_0}{2})}{1 + q_0} \bar{r} (1 + \lambda) + 1 \bar{X} \bar{C}_X (3 + 5 \bar{p}) - \frac{q_0}{2} e_0 (1 + \lambda) \right]$$ \hspace{1cm} (20)

where

$$\theta = \frac{aX}{aX + b}.$$

From (20), we have

$$\left( \bar{y}_\lambda - \bar{y} \right) = (q_1 - 1) \bar{p} + \omega_1 \left[ q_1 - \frac{\theta e_0 (1 + \lambda)}{2} + \frac{\theta^2 e_0^2}{8} (3 + 5 \lambda) - \frac{\theta}{2} e_0 e_1 (1 + \lambda) \right]$$ \hspace{1cm} (21)

Using (21), we get the MSE of $\bar{y}_\lambda$, to first order of approximation, as given by
\[
MSE(\bar{y}_n) = E(\bar{y}_n - \bar{y})^2 \\
= E \left[ (q_1 - 1)^2 \bar{y}^2 + q_1 \bar{y} \left( e_0 - \frac{q_1}{2} \left( 1 + \lambda \right) \right) - q_2 X e_1 \right]^2 \\
= E \left[ \left( q_1 - 1 \right)^2 \bar{y}^2 + q_1 \bar{y} e_0 - \left( \frac{q_1}{2} q_1 \bar{y} (1 + \lambda) + q_2 X \right) e_1 \right]^2 \\
= \left( (q_1 - 1)^2 \bar{y} \right)^2 + (q_1 \bar{y})^2 E(e_0^2) + \left( \frac{q_1}{2} q_1 \bar{y} (1 + \lambda) + q_2 X \right)^2 E(e_1^2) \\
- 2 q_1 \bar{y} \left( \frac{q_1}{2} q_1 \bar{y} (1 + \lambda) + q_2 X \right) E(e_0 e_1) \\
= \left( (q_1 - 1)^2 \bar{y} \right)^2 + \left( q_1 \bar{y} \right)^2 \frac{1 - f}{n} c_y^2 \left( (1 + \lambda)^2 - \left( 1 + \frac{q_1}{2} \rho c_y c_h \right) \right) \\
- 2 q_1 \bar{y} \left( \frac{q_1}{2} q_1 \bar{y} (1 + \lambda) + q_2 X \right) \frac{1 - f}{n} \rho c_y c_h \\
\]

On rearranging the terms we get

\[
MSE(\bar{y}_n) = \left( \left( q_1 - 1 \right)^2 \bar{y} \right)^2 + \frac{1 - f}{n} \left[ q_1^2 \bar{y} \left( c_y^2 + \frac{q_1}{2} \lambda c_y c_h \right) \left( 1 - \rho^2 \right) \right] \\
- 2 q_1 q_2 \bar{y} \left( \frac{q_1}{2} q_1 \bar{y} (1 + \lambda) + q_2 X \right) \frac{1 - f}{n} \rho c_y c_h \\
\]

Minimization of (22) with respect to \( q_i \) (i = 1,2) yields optimum values of \( q_1 \) and \( q_2 \) as:

\[
q_{10} = \frac{1}{1 + \left( \frac{1 - f}{n} \right) c_y^2 \left( 1 - \rho^2 \right)} \\
\]

and

\[
q_{20} = q_{10} \left( \frac{\bar{y}}{5c_h} \right) \left( \rho c_y - \frac{1 - f}{n} \rho c_h \right) \\
\]

(23)

We assume that the unknown parameters \( c_y \) and \( \rho \) are easily estimable from the preliminary data as in Tracy et al. [15] and Singh et al. [10]. Substituting the optimum values \( q_{10} \) and \( q_{20} \) in (22), the minimum MSE of \( \bar{y}_p \), to first order of approximation, is given by

\[
MSE(\bar{y}_n)_{\text{min}} = \frac{\text{MSE}(\bar{y}_n)}{1 + \left( \frac{1 - f}{n} \right) c_y^2 \left( 1 - \rho^2 \right)} \\
\]

EFFICIENCY COMPARISONS

First we compare the efficiency of proposed estimator with mean per unit estimator

\[
MSE(\bar{y}_n)_{\text{min}} - MSE(\bar{y}_m)_{\text{min}} \geq 0 \\
\]

\[
f_1 \frac{\bar{y}^2}{c_y^2} \left( 1 - \rho^2 \right) + \rho^2 \geq 0 \\
\]

(25)

where

\[
f_1 = \left( \frac{1 - f}{n} \right) \\
\]

Next we compare the proposed estimator with usual ratio estimator

\[
MSE(\bar{y}_n)_{\text{min}} - MSE(\bar{y}_r)_{\text{min}} \geq 0 \\
\]

\[
f_1 \frac{\bar{y}^2}{c_y^2} \left( c_y^2 + \frac{c_y}{4} - \rho c_y c_h \right) - \frac{f_1 \bar{y}^2 c_y^2 (1 - \rho^2)}{1 + f_1 c_y^2 (1 - \rho^2)} \geq 0 \\
\]

(26)

Similarly, we can compare the proposed estimator with exponential ratio type estimator (with \( a=1 \) and \( b=0 \)).

\[
MSE(\bar{y}_n)_{\text{min}} - MSE(\bar{y}_e)_{\text{min}} \geq 0 \\
\]

\[
f_1 \frac{\bar{y}^2}{c_y^2} \left( c_y^2 + \frac{c_y}{4} - \rho c_y c_h \right) - \frac{f_1 \bar{y}^2 c_y^2 (1 - \rho^2)}{1 + f_1 c_y^2 (1 - \rho^2)} \geq 0 \\
\]

(27)
We see, conditions (26) and (27) are always satisfied.

Next we compare the efficiency of proposed estimator with Kadilar and Cingi [3] estimator $\bar{y}_{KCi}$.

Similarly we compare the proposed estimator with Kadilar and Cingi [5] estimator $\hat{y}_{KCI}$.

The only change in expression is to replace $\hat{y}_{IK}^2$ by $\hat{y}_{IK}^2$, we get

$$f_1\frac{\hat{y}_{IK}^2}{2}(1 - \rho^2) - f_2\frac{\hat{y}_{IK}^2}{2}(1 - \rho^2) \geq 0$$

We see, conditions (26) and (27) are always satisfied.

Next we compare the efficiency of proposed estimator with Kadilar and Cingi [3] estimator $\bar{y}_{KCi}$.

$$\text{MSE}(\overline{y}_{KCi}) - \text{MSE}(\overline{y}_{IK})_{\text{min}} \geq 0$$

On solving, we get final expressions as

$$f_1\frac{\hat{y}_{IK}^2}{2}(1 - \rho^2) - f_2\frac{\hat{y}_{IK}^2}{2}(1 - \rho^2) \geq 0$$

This is always true.
This is true always.

Finally we compare the efficiency of proposed estimator with Gupta and Shabbir [2] estimator.

\[
\text{MSE} (\delta_{g}) - \text{MSE} (\delta_{h})_{\min} \geq 0
\]

as we see both the MSE(s) are equal.

**EMPIRICAL STUDY**

For empirical study we use two data sets earlier used by Singh and Tailor [12] and Singh and Agnihotri [14].

When we examine Table 2, we observe that proposed class of estimators performs better than the other estimators discussed in this paper and the minimum MSE of the proposed estimator is same as that of Gupta and Shabbir [2] estimator.

**CONCLUSION**

In this article, we suggested some families of estimators in simple random sampling using known values of some population parameters. It has been shown that many existing estimators are particular members of the proposed family of estimator. We studied the effect of various transformations of auxiliary information on the families of estimators.

**REFERENCES**

1. Bahl, S. and R.K. Tuteja, 1991. Ratio and product type exponential estimator. Infrm. and Optim. Sci., 12 (1): 159-163.
2. Gupta, S. and J. Shabbir, 2008. On improvement in estimating the population mean in simple random sampling. Jour. of Appl. Statis., 35 (5): 559-566.
3. Kadilar, C. and H. Cingi, 2004. Ratio Estimators in Simple Random Sampling. Appl. Math. and Comp., 151: 893-902.
4. Kadilar, C. and H. Cingi, 2006. New ratio estimators using correlation coefficient. InterStat, pp: 1-11.
5. Kadilar, C. and H. Cingi, 2006. Improvement in estimating population mean using correlation coefficient. Hacet. J. Math. Stat., 35 (1): 103-109.
6. Kadilar, C. and H. Cingi, 2006. Improvement in estimating the population mean in simple random sampling. Appl. Math. Lett., 19 (1): 75-79.
7. Khoshnevisan, M., R. Singh, P. Chauhan, N. Sawan and F. Smarandache, 2007. A general family of estimators for estimating population mean using known value of some population parameter(s). Far east Jour. of Theor. Statis., 22 (2): 181-191.
8. Shabbir, J. and S. Gupta, 2005. Improved ratio estimators in stratified sampling. Am. J. Math. Manage. Sci. 25 (3-4): 293-311.
9. Shabbir, J. and S. Gupta, 2006. A new estimator of population mean in stratified sampling. Commun. Stat. Theo. Meth., Vol: 35 (7).
10. Singh, R., P. Cauhan, N. Sawan and F. Smarandache, 2007. Auxiliary information and a priori values in construction of improved estimators. Renaissance High Press.
11. Singh, R., M. Kumar and F. Smarandache, 2008. Almost Unbiased Estimator for Estimating Population Mean Using Known Value of Some Population Parameter (s). Pak. J. Stat. Oper. Res., 4 (2): 63-76.
12. Singh, H.P. and R. Tailor, 2003. Use of known correlation coefficient in estimating the finite population mean. Stat. in Trans., 6 (4): 555-560.
13. Singh, H.P., R. Tailor, S. Singh and J.M. Kim, 2008. A modified estimator of population mean using power transformation. Statis., 49: 37-58.
14. Singh, H.P. and N. Agnihotri, 2008. A general procedure of estimating population mean using auxiliary information in sample surveys. Stat. in Trans.-new series, 9 (1): 71-87.
15. Tracy, D.S., H.P. Singh and R. Singh, 1998. A class of almost unbiased estimators for finite population mean using two auxiliary variables. Biom. Jour., 40: 753-766.