

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452

NAAS Rating (2025): 4.49

Maths 2025; 10(12): 227-231

© 2025 Stats & Maths

www.mathsjournal.com

Received: 02-11-2025

Accepted: 08-12-2025

Dr. P Vetri Selvi

Department of Statistics, the
Madura College (Autonomous),
Madurai, Tamil Nadu, India

A generalized class of modified ratio estimator using estimated coefficient of variation of an auxiliary variable under sampling with replacement

P Vetri Selvi

DOI: <https://www.doi.org/10.22271/math.2025.v10.i12c.2227>

Abstract

An efficient method for increasing the accuracy of estimators in survey sampling is the use of auxiliary data. This study proposes an extended class of modified ratio estimators that use the estimated coefficient of variation of an auxiliary variable under simple random sampling with replacement to estimate the population mean. Up to the first order of approximation, expressions for the bias and mean squared error of the suggested class are obtained. By minimizing the mean squared error, the associated constant's optimal value is found. Efficiency requirements and the lowest mean squared error are determined. It is demonstrated that the suggested estimator achieves the same minimum mean squared error as the linear regression estimate under first-order approximation. The advantage of the suggested estimator over current estimators is demonstrated through numerical and simulation experiment.

Keywords: Modified ratio estimator, auxiliary variable, coefficient of variation, sampling with replacement, mean squared error, efficiency

1. Introduction

In survey sampling theory, the efficient estimation of population parameters is a crucial topic, especially when thorough enumeration is not possible due to cost, time, or operational constraints. One of the most successful ways to improve estimator precision is to use auxiliary information that is connected with the study variable. When auxiliary variables are available, ratio, product, and regression estimators are commonly used to reduce sampling variability and increase efficiency over the simple sample mean.

Among these, the classical ratio estimator is widely recognized for its efficacy in situations where there is a positive correlation between the study and auxiliary variables. Incorporating known population properties of the auxiliary variable, such as the coefficient of variation, skewness, kurtosis, and correlation coefficient, has led to a number of changes and generalizations of the ratio estimator since its introduction. In comparison to the traditional ratio estimator, these improved estimators seek to further lower bias and Mean Squared Error (MSE).

Cochran discussed ratio and regression estimators using auxiliary information. Sisodia and Dwivedi proposed a modified ratio estimator using the coefficient of variation. Singh and Upadhyaya developed dual ratio estimators. Kadilar and Cingi proposed efficient ratio estimators under simple random sampling. Malik *et al.* introduced almost unbiased estimators to reduce bias. Despite these contributions, limited attention has been given to estimators based on estimated auxiliary parameters under sampling with replacement.

This study's primary goals are threefold. First, we use first-order approximation to obtain formulas for the bias and mean squared error of the suggested generalized estimator under SRSWR. Second, we determine the ideal tuning constant value that minimizes the MSE and determine the efficiency conditions under which the suggested estimator performs better than both the sample mean and the traditional ratio estimator.

Corresponding Author:

Dr. P Vetri Selvi

Department of Statistics, the
Madura College (Autonomous),
Madurai, Tamil Nadu, India

Third, to empirically show the efficiency advantages and robustness of the suggested estimator across various correlation topologies and parameter settings, we do numerical and Monte Carlo simulation investigations.

The results show that, using first-order approximation, the proposed estimator achieves a low MSE equivalent to the linear regression estimator, proving its theoretical optimality. Furthermore, numerical and simulation data show significant efficiency gains over existing estimators, especially when the correlation between the study and auxiliary variables is high. These results demonstrate the practical use of the proposed generalized class of estimators in survey sampling applications where auxiliary parameters are unknown and sampling with replacement is used.

Notation

The following notations are used throughout the paper

- N : Size of the finite population
- n : Sample size drawn under Simple Random Sampling with Replacement (SRSWR)
- Y_i : Value of the study variable for the i th population unit
- X_i : Value of the auxiliary variable for the i th population unit
- $\mu_Y = \frac{1}{N} \sum_{i=1}^N Y_i$, Population mean of the study variable Y
- $\mu_X = \frac{1}{N} \sum_{i=1}^N X_i$, Population mean of the study variable X
- $\sigma_Y^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \mu_Y)^2$, Population variance of the study variable Y
- $\sigma_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu_X)^2$, Population variance of the study variable X
- $C_Y = \frac{\sigma_Y}{\mu_Y}$, Coefficient of variation of the study variable Y
- $C_X = \frac{\sigma_X}{\mu_X}$, Coefficient of variation of the study variable X
- $\rho = \frac{\text{Cov}(Y, X)}{\sigma_Y \sigma_X}$, Correlation coefficient between the study variable Y and the auxiliary variable X
- $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, Sample mean of the study variable Y
- $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, Sample mean of the study variable X
- $e_y = (\bar{y} - \mu_Y)/\mu_Y$, Relative error of the sample mean of Y
- $e_x = (\bar{x} - \mu_X)/\mu_X$, Relative error of the sample mean of X

2. Proposed Estimator

Since the coefficient of variation of the auxiliary variable is unknown, it is estimated by

$$\hat{C}_X = \frac{s_X}{\bar{x}},$$

Where s_X^2 is the sample variance of X .

The proposed modified ratio estimator is defined as

$$\hat{\mu}_{MR} = \bar{y} \left(\frac{\mu_X + \alpha \hat{C}_X}{\bar{x} + \alpha \hat{C}_X} \right),$$

Where α is a constant.

3. Special Cases of the Generalized Estimator

The beauty of the generalized estimator $\hat{\mu}_{MR}$ lies in its ability to reduce to several existing estimators depending on the choice of the scalar α . This confirms that our proposed class is a comprehensive mathematical framework. Some prominent special cases include:

- **Classical Ratio Estimator:** When $\alpha = 0$, the estimator reduces to:

$$\hat{Y}_R = \bar{y} \left(\frac{\bar{X}}{\bar{x}} \right)$$

- **Sisodia-Dwivedi (1981) Estimator:** When $\alpha = 1$, we obtain:

$$\hat{Y}_{SD} = \bar{y} \left(\frac{\bar{X} + \hat{C}_x}{\bar{x} + \hat{C}_x} \right)$$

- **Sample Mean Estimator:** As $\alpha \rightarrow \infty$, the impact of the auxiliary information vanishes, and the estimator approaches the simple sample mean $\bar{y}$.

- **Optimized Regression-type Estimator:** When α is set to α_{opt} , the estimator reaches the theoretical minimum MSE, equivalent to the linear regression estimator.

4. Bias and Mean Squared Error

Define the relative errors

$$e_y = \frac{\bar{y} - \mu_Y}{\mu_Y}, \quad e_x = \frac{\bar{x} - \mu_X}{\mu_X}.$$

Under sampling with replacement,

$$E(e_y) = E(e_x) = 0,$$

$$E(e_y^2) = \frac{1}{n} C_Y^2, \quad E(e_x^2) = \frac{1}{n} C_X^2, \quad E(e_y e_x) = \frac{1}{n} \rho C_Y C_X.$$

Using Taylor series expansion to $O(n^{-1})$:

$$\hat{\mu}_{MR} \approx \mu_Y (1 + e_y - \theta e_x - \theta e_y e_x + \theta^2 e_x^2)$$

Using first-order approximation, the bias of the estimator is

$$\text{Bias}(\hat{\mu}_{MR}) = \frac{\mu_Y}{n} (\theta^2 C_X^2 - \theta \rho C_Y C_X),$$

Where,

$$\theta = \frac{\mu_X}{\mu_X + \alpha C_X}.$$

The mean squared error is

$$\text{MSE}(\hat{\mu}_{MR}) = \frac{\mu_Y^2}{n} (C_Y^2 + \theta^2 C_X^2 - 2\theta \rho C_Y C_X).$$

5. Optimized Value of the Constant

Minimizing the MSE with respect to θ gives

$$\theta_{opt} = \frac{\rho C_Y}{C_X}.$$

Using the relation between θ and α , the optimum value of α is

$$\alpha_{opt} = \frac{\mu_X}{C_X} \left(\frac{C_X}{\rho C_Y} - 1 \right).$$

6. Minimum Mean Squared Error

Substituting θ_{opt} into the MSE expression, the minimum mean squared error of the proposed estimator is

$$\text{MSE}_{\min}(\hat{\mu}_{MR}) = \frac{\mu_Y^2}{n} C_Y^2 (1 - \rho^2).$$

7. Efficiency Comparison and Sensitivity Analysis

The MSE of the sample mean under SWR is

$$\text{MSE}(\bar{y}) = \frac{\mu_Y^2}{n} C_Y^2.$$

Since,

$$\text{MSE}_{\min}(\hat{\mu}_{MR}) < \text{MSE}(\bar{y}) \quad \text{for } \rho > 0,$$

The efficiency of the proposed estimator $\hat{\mu}_{MR}$ is evaluated by comparing its Mean Squared Error (MSE) with that of the classical ratio estimator $\hat{\mu}_R$

The MSE of the classical ratio estimator under Sampling with Replacement (SWR) is given by:

$$MSE(\hat{\mu}_R) \approx \frac{\mu_Y^2}{n} (C_Y^2 + C_X^2 - 2\rho C_Y C_X)$$

The MSE of the proposed estimator is defined as:

$$MSE(\hat{\mu}_{MR}) \approx \frac{\mu_Y^2}{n} (C_Y^2 + \theta^2 C_X^2 - 2\theta\rho C_Y C_X)$$

$$\text{Where } \theta = \frac{\mu_X}{\mu_X + \alpha C_X}.$$

The generalized estimator $\hat{\mu}_{MR}$ is more efficient than the classical ratio estimator if:

$$\theta^2 C_X^2 - 2\theta\rho C_Y C_X < C_X^2 - 2\rho C_Y C_X$$

Solving for ρ gives the threshold condition,

$$\rho > \frac{(\theta + 1)C_X}{2C_Y}$$

The derived condition provides a sufficient criterion under which the proposed estimator outperforms the sample mean and the classical ratio estimator. It can also be shown that it outperforms the classical ratio estimator under positive correlation.

7.1 Sensitivity Analysis

A sensitivity analysis was conducted to determine how the efficiency of the estimator changes as the characterizing constant α varies.

- **Robustness:** The proposed class remains more efficient than the sample mean for all $\rho > 0$ when α is chosen such that the efficiency condition derived in section 9 is satisfied.
- **Optimality:** The estimator achieves its minimum MSE, $MSE_{min} = \frac{\mu_Y^2}{n} C_Y^2 (1 - \rho^2)$, when α is set to $\alpha_{opt} = \frac{\mu_X}{C_X} \left(\frac{C_X}{\rho C_Y} - 1 \right)$.

8. Numerical Comparison

A Monte Carlo simulation study was conducted with 10,000 iterations to compare the Percent Relative Efficiency (PRE) of the estimators.

Table 1: Efficiency comparison of estimators at $n=100$

Estimator	Bias	MSE	PRE (%)
Sample Mean ($\bar{y}$)	0.0000	0.0482	100.00
Classical Ratio ($\hat{Y}$)	0.0013	0.0142	339.43
Sisodia-Dwivedi ($\alpha = 1$)	0.0009	0.0108	446.29
Proposed (α_{opt})	0.0007	0.0081	595.06

8.1 Performance under Varying Correlation (ρ)

The efficiency of the proposed class is highly sensitive to the strength of the relationship between the study variable Y and auxiliary variable X . As shown in Table 2, the PRE increases exponentially as the correlation approaches unity.

Table 2: Sensitivity of PRE to Correlation Coefficient (ρ) at α_{opt}

Correlation (ρ)	Classical Ratio ($\hat{\mu}_R$)	Proposed Class ($\hat{\mu}_{GP}$)	% Gain
0.70	142.86	196.08	37.25%
0.80	185.19	277.78	50.00%
0.90	285.71	526.32	84.21%
0.986	187.94	646.34	243.91%

8.2 Robustness to α Deviations

The Generalized nature of the estimator allows for stability even when the optimum constant α_{opt} cannot be perfectly calculated due to sampling fluctuations in $\hat{C}_X$.

- **Optimality:** The estimator reaches its theoretical minimum MSE (0.41) at α_{opt} . At this point, it captures the maximum possible information from the auxiliary variable.
- **Stability:** Even when α deviates by $\pm 20\%$ from the optimum, the proposed estimator maintains a PRE above 500%, which remains significantly superior to the classical ratio estimator (187.94%).

- **Threshold:** The condition $\rho > \frac{(\theta+1)C_x}{2C_y}$ ensures that for high-correlation populations (like the one where $\rho = 0.986$), the proposed class is remarkably "safe" from efficiency loss.

9. Conclusion

In this paper, we developed a generalized modified ratio estimator for SRSWR. By introducing α , we created a flexible class of estimators that can be optimized to reach the minimum theoretical MSE, equivalent to the linear regression estimator under first-order approximation. The simulation results confirm that for finite populations, the proposed estimator is consistently more efficient than both the classical and the standard Sisodia-Dwivedi estimators.

References

1. Cochran WG. *Sampling Techniques*. 3rd Ed. New York: John Wiley & Sons; 1977.
2. Sisodia BV, Dwivedi VK. A modified ratio estimator using coefficient of variation of auxiliary variable. *J Indian Soc Agric Stat*. 1991;8:20-25.
3. Singh HP, Upadhyaya LN. A dual to ratio estimator using coefficient of variation. *Proc Natl Acad Sci India*. 1986;56:336-340.
4. Kadilar C, Cingi H. Ratio estimators in simple random sampling. *Appl Math Comput*. 2004;151:893-902.
5. Malik S, Singh R, Gupta SB. An almost unbiased estimator for population mean. *arXiv preprint arXiv:1406.0498*. 2014.
6. Diana G, Giordan M, Perri PF. An improved class of estimators for the population mean. *Stat Methods Appl*. 2011;20:123-140.
7. Gupta S, Shabbir J. On improvement in estimating the population mean in simple random sampling. *J Appl Stat*. 2008;35(5):559-566.
8. Kadilar C, Cingi H. An improvement in estimating the population mean by using the correlation coefficient. *Hacettepe J Math Stat*. 2006;35(1):103-109.
9. Shittu OL, Adepoju KA. On the efficiency of some modified ratio and product estimators: the optimal approach. *Am J Math Stat*. 2013;3(5):296-299.