Heuristical Approach for Optimizing Population Mean Using Ratio Estimator in Stratified Random Sampling

G. R. V. Triveni and Faizan Danish∗

Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati AP-522237, India
E-mail: danishstat@gmail.com
∗Corresponding Author

Received 05 April 2023; Accepted 20 October 2023;
Publication 20 November 2023

Abstract
In this study, we have developed a Ratio type estimator in Stratified sampling to estimate the population average of study variable by using the information of a concomitant variable. By utilizing Taylor’s series, we have derived the expressions for Bias and MSE upto first degree of approximation. In numerical illustration, employing a real data set, we have demonstrated that the proposed estimator has highest Percentage relative efficiency when compared to the considered existing estimators. Furthermore, we have demonstrated that Separate ratio type estimators have the highest relative efficiency when contrasted with the Combined ratio type estimators.

Keywords: Relative efficiency, Bias, mean square error, stratified sampling, Taylors series, ratio estimators.

doi: 10.13052/jrss0974-8024.1617
© 2023 River Publishers
1 Introduction

In contemporary survey methodologies, the practice of stratification serves as a strategic tool aimed at strengthening the accuracy of estimations. In this approach, a diverse population is partitioned into distinct strata with the objective of achieving homogeneity within each stratum. Subsequently, a sample is drawn from each stratum utilizing various sampling designs. This method proves particularly effective when the selected units are not substantial in number, and the population demonstrates significant variability. The act of segmenting the population into strata offers the potential to yield more precise estimates of population parameters while concurrently reducing variance within each stratum.

Optimal stratification emerges as a technique that yields the minimal possible variance in this process. Ratio estimation constitutes a central endeavour, aiming to estimate the same ratio within the sample by utilizing the population total of a variable of interest and the population size. Ratio estimation stands out as a robust method for estimating population parameters, and its utility becomes especially apparent in situations where the population exhibits high levels of variability or when the selected units are limited in number. One of the key advantages of employing ratio estimation within the framework of stratified random sampling lies in its capacity to reduce estimator variance, thereby enhancing the precision of estimates.

The study of estimators has been supported by numerous researchers. Cochran (1971) gave a brief summary of the stratified sampling process for estimating the estimators using various sampling designs, whereas Sarjinder Singh (2003) spoke extensively about the evolution of estimators in general. A class of product cum ratio type estimators was presented by Sisodia and Dwivedi (1981). Ratio and Product type exponential estimators were shown to have higher precision by Bahl and Tuteja (1991) after Prasad (1989) worked on a class of ratio estimators. Upadhyaya and Singh (1999) provided a new approach by transforming the auxiliary variable by assuming auxiliary variable is known and they concluded that proposed estimator has lower MSE and it performs better than estimators in literature. In 2003, Kadilar and Cingi assessed a variety of ratio-type estimators before going on to create a new estimator. They expanded their study and contributed significantly to the field of survey sampling in 2005 by introducing a novel estimation method. In order to enhance several different forms of ratio estimators, Shabbir and Gupta (2005) worked on them. Later, Koyuncu and
Kadilar (2009) provided effective estimators for the population mean and demonstrated the effectiveness of their proposal by numerical illustration. Yadav et al. (2011) promoted an enhanced separate ratio exponential estimator and demonstrated the superiority of the estimator by utilizing theoretical and numerical comparisons.

Additionally, Tailor and Chouhan and Tailor and Lone (2014) researched the characteristics of distinct ratio-type estimators and worked on ratio cum product type exponential estimators. Yadav et al. (2014) developed ratio and product exponential estimators and evaluated its efficiency by using an empirical study. Later, using information from two auxiliary variables and a difference cum exponential ratio type estimator, Shabbir and Gupta (2017) were able to use stratified random sampling as well as basic random sampling. The efficiency of the novel estimators was supported by theoretical properties and numerical findings provided by Verma et al. (2017) in their proposal of a distinct class of estimators. A number of estimators in various sample designs were created by Koyuncu et al. (2018), Zaman (2019), Luengo et al. (2019), Kumar and Vishwakarma (2020), Yadav and Tailor (2020), Dahiru et al. (2021), and Khare et al. (2022). Rather and Kadilar (2022) demonstrated an increase in the proposed estimator’s efficiency by contrasting it with an unbiased estimator. Additionally, Kumar et al. (2023) created a group of estimators using stratified random sampling, compared those estimators to others already in use, and came to the conclusion that the new class of estimators outperformed the older ones. In this paper, we have developed an estimator in Section 4 and proved the efficiency of proposed and considered estimators in Section 6.

2 Terminology

- \( N \): no. of units in population
- \( n \): no. of units in the sample
- \( P_h = \frac{N_h}{N} \): is stratum weight
- \( k \): no. of strata
- \( N_h \): population units in \( h^{th} \) stratum
- \( n_h \): sample units in \( h^{th} \) stratum
- \( Z_1 = \sum_{h=1}^{k} P_h \bar{z}_{1h} \): is the average of study variable
- \( Z_2 = \sum_{h=1}^{k} P_h \bar{z}_{2h} \): is the average of auxiliary variable
• $z_{1h} = \frac{1}{n_h} \sum_{j=1}^{n_h} z_{1hj}$ is the study variable’s average for the sample at $h^{th}$ stratum

• $z_{2h} = \frac{1}{n_h} \sum_{j=1}^{n_h} z_{2hj}$ is the auxiliary variable’s average for the sample at $h^{th}$ stratum

• $S_{hz1}^2 = \frac{1}{N_h - 1} \sum_{j=1}^{n_h} (z_{1hj} - \bar{Z}_{1h})^2$ is the study variable’s variance at $h^{th}$ stratum

• $S_{hz2}^2 = \frac{1}{N_h - 1} \sum_{j=1}^{n_h} (z_{2hj} - \bar{Z}_{2h})^2$ is the auxiliary variable’s variance at $h^{th}$ stratum

• $S_{hz1z2} = \frac{1}{N_h - 1} \sum_{j=1}^{n_h} (z_{1hj} - \bar{Z}_{1h})(z_{2hj} - \bar{Z}_{2h})$ is the covariance between variables at $h^{th}$ stratum

• $\rho_{hz1z2} = \frac{S_{hz1z2}}{S_{hz1}S_{hz2}}$ is correlation coefficient between variables at $h^{th}$ stratum

• $C_{hz1} = \frac{S_{hz1}}{\bar{Z}_{1h}}$ is Coefficient of variation for study variable at $h^{th}$ stratum

• $C_{hz2} = \frac{S_{hz2}}{\bar{Z}_{2h}}$ is Coefficient of variation for auxiliary variable at $h^{th}$ stratum

A population ‘U’ made up of units $U_1, U_2, U_3 \ldots U_N$ of size $N$. Let $\bar{Z}_1$ and $\bar{Z}_2$ represents the study and concomitant variable. If population $U$ is separated into $K$ homogeneous strata of sizes $n_h (h = 1, 2, 3, \ldots k)$ and $n_h$ represents the sample size drawn from $h^{th}$ stratum. The population mean $\bar{Z}_1$ for Separate Ratio estimator is expressed as

$$\bar{Z}_{1SR} = \sum_{h=1}^{k} P_h \bar{Z}_{1h} \left[ \frac{\bar{Z}_{2h}}{\bar{Z}_{2h}} \right]$$

3 Estimators in Literature

(1) The Traditional Separate ratio estimator is

$$\bar{Z}_{1SR} = \sum_{h=1}^{k} P_h \bar{Z}_{1hR}$$

$$= \sum_{h=1}^{k} P_h \left[ \bar{Z}_{1h} \bar{Z}_{2h} \right]$$
Heuristical Approach for Optimizing Population Mean Using Ratio Estimator

Where, $\bar{z}_{1hR}$ is the ratio estimator of $h^{th}$ stratum population mean $\bar{z}_{1h}$ and $\bar{z}_{2h}$ is the proportion of sample means of the study variable $z_1$ and auxiliary variable $z_2$ of $h^{th}$ stratum. The Bias and MSE are as follows:

\[
\text{Bias}(\bar{z}_{1SR}) = \sum_{h=1}^{k} P_h \bar{z}_{1h} \left( \frac{1 - f_h}{n_h} \right) \left[ C_{hz2}^2 - \rho_{hz1z2} C_{hz1} C_{hz2} \right]
\]

\[
\text{MSE}(\bar{z}_{1SR}) = \sum_{h=1}^{k} P_h^2 \bar{z}_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\times \left[ C_{hz1}^2 + C_{hz2}^2 - 2\rho_{hz1z2} C_{hz1} C_{hz2} \right]
\]

Rajesh and Lone (2014) modified the following estimators into Separate ratio estimators as:

(2) Separate ratio type estimator utilizing $C_{z2}$ of auxiliary variate in $h^{th}$ stratum as

\[
\bar{z}_{1RS}^{SD} = \sum_{h=1}^{k} P_h \bar{z}_{1h} \left[ \frac{z_{2h} + C_{z2h}}{z_{2h} + C_{z2h}} \right]
\]

where $C_{z2h}$ is the auxiliary variate’s coefficient of variation in $h^{th}$ stratum

\[
\text{Bias}(\bar{z}_{1RS}^{SD}) = \sum_{h=1}^{k} P_h \bar{z}_{1h} \left( \frac{1 - f_h}{n_h} \right) \left[ \lambda_{1h}^2 C_{hz2}^2 - \rho_{hz1z2} \lambda_{1h} C_{hz1} C_{hz2} \right]
\]

\[
\text{MSE}(\bar{z}_{1RS}^{SD}) = \sum_{h=1}^{k} P_h^2 \bar{z}_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\times \left[ C_{hz1}^2 + \lambda_{1h}^2 C_{hz2}^2 - 2\rho_{hz1z2} \lambda_{1h} C_{hz1} C_{hz2} \right]
\]

Where $\lambda_{1h} = \frac{z_{2h}}{z_{2h} + C_{z2h}}$

(3) Separate Ratio type estimator by utilizing kurtosis $\beta_2(z_2)$ of $z_2$ in $h^{th}$ stratum as

\[
\bar{z}_{1RS}^{SE} = \sum_{h=1}^{k} P_h \bar{z}_{1h} \left[ \frac{z_{2h} + \beta_{2h}(z_2)}{z_{2h} + \beta_{2h}(z_2)} \right]
\]
Where $\beta_2(z_2)$ is the auxiliary variate’s coefficient of kurtosis in $h^{th}$ stratum.

\[
\text{Bias}(Z_{1RS}^{SE}) = \sum_{h=1}^{k} P_h Z_{1h} \left( \frac{1 - f_h}{n_h} \right) \left[ \lambda_{2h}^2 C_{hz_2}^2 - \rho_{hz_1 z_2} \lambda_{2h} C_{hz_1} C_{hz_2} \right]
\]

\[
\text{MSE}(Z_{1RS}^{SE}) = \sum_{h=1}^{k} P_h^2 Z_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\times \left[ C_{hz_1}^2 + \lambda_{2h}^2 C_{hz_2}^2 - 2 \rho_{hz_1 z_2} \lambda_{2h} C_{hz_1} C_{hz_2} \right]
\]

(3)

where $\lambda_{2h} = \frac{z_{2h}}{\overline{z}_{2h} + \beta_{2h}(z_2)}$

(4) By utilizing coefficients of kurtosis and variation in $h^{th}$ stratum, preferred a separate ratio-type estimator as

\[
Z_{1RS}^{US1} = \sum_{h=1}^{k} P_h Z_{1h} \left[ \frac{\beta_{2h}(z_2) + C_{z_2 h}}{\overline{z}_{2h}+\beta_{2h}(z_2)} \right]
\]

Where $\beta_2(z_2)$ is the coefficient of kurtosis and $C_{z_2 h}$ is the auxiliary variate’s coefficient of variation in $h^{th}$ stratum.

\[
\text{Bias}(Z_{1RS}^{US1}) = \sum_{h=1}^{k} P_h Z_{1h} \left( \frac{1 - f_h}{n_h} \right) \left[ \lambda_{3h}^2 C_{hz_2}^2 - \rho_{hz_1 z_2} \lambda_{3h} C_{hz_1} C_{hz_2} \right]
\]

\[
\text{MSE}(Z_{1RS}^{US1}) = \sum_{h=1}^{k} P_h^2 Z_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\times \left[ C_{hz_1}^2 + \lambda_{3h}^2 C_{hz_2}^2 - 2 \rho_{hz_1 z_2} \lambda_{3h} C_{hz_1} C_{hz_2} \right]
\]

(4)

where $\lambda_{3h} = \frac{\overline{z}_{2h} \beta_{2h}(z_2)}{\overline{z}_{2h}+\beta_{2h}(z_2)+C_{z_2 h}}$

(5) By using coefficients of variation and kurtosis in $h^{th}$ stratum, preferred a separate ratio-type estimator as

\[
Z_{1RS}^{US2} = \sum_{h=1}^{k} P_h Z_{1h} \left[ \frac{\beta_{2h}(z_2) + \beta_{2h}(z_2)}{\overline{z}_{2h} C_{z_2 h} + \beta_{2h}(z_2)} \right]
\]
Heuristical Approach for Optimizing Population Mean Using Ratio Estimator

Bias(\( \bar{Z}_{1RS}^{US2} \)) = \sum_{h=1}^{k} P_h \bar{Z}_{1h} \left( 1 - \frac{f_h}{n_h} \right) \left[ \lambda_{4h}^2 C_{hz2}^2 - \rho_{hz1z2} \lambda_{4h} C_{hz1} C_{hz2} \right]

MSE(\( \bar{Z}_{1RS}^{US2} \)) = \sum_{h=1}^{k} P_h^2 \bar{Z}_{1h}^2 \left( 1 - \frac{f_h}{n_h} \right) \times \left[ \sigma_{hz1}^2 + \lambda_{4h}^2 C_{hz2}^2 - 2 \rho_{hz1z2} \lambda_{4h} C_{hz1} C_{hz2} \right]

where \( \lambda_{4h} = \frac{\sigma_{z2h} C_{z2h}}{\sigma_{z2h}^2 + \beta_{2h}(z_2)} \)

(6) The traditional Combined Ratio estimate is

\( Z_{1RC} = \frac{Z_{1st} Z_2}{Z_{2st}} \)

Its MSE is given by

MSE(\( Z_{1RC} \)) = \sum_{h=1}^{k} P_h^2 \left( 1 - \frac{f_h}{n_h} \right) \left[ S_{hz1}^2 + R^2 S_{hz2}^2 - 2 R S_{hz1z2} \right]

(6)

(7) Kadilar and Cingi (2005) proposed a new Ratio estimator as

\( Z_{1KC} = m^* Z_{1RC} \)

MSE(\( Z_{1KC} \)) = \sum_{h=1}^{k} P_h^2 \left( 1 - \frac{f_h}{n_h} \right) \left[ S_{hz1}^2 + R^2 S_{hz2}^2 - 2 R S_{hz1z2} \right]

+ (m^* - 1)^2 \bar{Z}_{1}^2

(7)

Where \( m^* = \frac{\bar{Z}_2^2}{\bar{Z}_1^2 + \sum_{h=1}^{k} P_h^2 \left( 1 - \frac{f_h}{n_h} \right) \left[ S_{hz1}^2 + R^2 S_{hz2}^2 - 2 R S_{hz1z2} \right]} \)

4 Proposed Estimator

In Stratified random Sampling, we propose that

\( \bar{Z}_{1st} = \frac{t^2}{2} \sum_{h=1}^{k} P_h \left[ \frac{Z_{1h} Z_{2h}}{Z_{2h}} \right] \)

(8)
Proposed Estimator’s bias can be obtained as

\[
E(Z_{1st} - Z_1) = E \left( \frac{t^2}{2} \sum_{h=1}^{k} P_h \left[ \frac{Z_{1h} Z_{2h}}{Z_{2h}} \right] - Z_1 \right)
\]

\[
= Z_2 E \left[ \frac{t^2}{2} \sum_{h=1}^{k} P_h Z_{1h} - \left( \frac{Z_{1z}}{Z_2} \right) \right]
\]

(9)

We have,

\[
\frac{1}{Z_{2h}} = \frac{1}{Z_2} \left( 1 + \frac{Z_{2h} - Z_2}{Z_2} \right)^{-1}
\]

By using Taylor Series expansion, we get

\[
\frac{1}{Z_{2h}} \approx \frac{1}{Z_2} \left( 1 - \frac{Z_{2h} - Z_2}{Z_2} \right)
\]

(10)

Substituting Equation (10) in Equation (9), we get

\[
E(Z_{1st} - Z_1) = Z_2 E \left\{ \left( \frac{t^2}{2} \sum_{h=1}^{k} P_h Z_{1h} - \frac{Z_{1z} Z_2}{Z_2} \right) \frac{1}{Z_2} \left( 1 - \frac{Z_{2h} - Z_2}{Z_2} \right) \right\}
\]

\[
= \left[ \frac{t^2}{2} - 1 \right] Z_1 + \frac{1}{Z_2} \sum_{h=1}^{k} P_h Z_{1h} \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\times [C_{hz_2}^2 - t \rho_{hz_1 z_2} C_{hz_1} C_{hz_2}]
\]

Therefore,

\[
\text{Bias}(Z_{1st}) = (t - 1) Z_1 + \frac{1}{Z_2} \sum_{h=1}^{k} P_h Z_{1h} \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\times \left[ C_{hz_2}^2 - \frac{t^2}{2} \rho_{hz_1 z_2} C_{hz_1} C_{hz_2} \right]
\]

\[
\text{MSE}(Z_{1st}) = \frac{t^4}{4} \text{var}(Z_{1SR}) + Z_1^2 \left[ \frac{t^2}{2} - 1 \right]^2
\]

(11)
Thus, the MSE of the proposed estimator will take the following form as:

\[
\text{MSE}(Z_{1st}) = \frac{t^4}{4} \sum_{h=1}^{k} P_h Z_{1h}^2 \left( \frac{1 - f_h}{n_h} \right) \\
\times [C_{hz1}^2 + C_{hz2}^2 - 2\rho_{hz1z2} C_{hz1} C_{hz2}] + Z_1^2 \left[ \frac{t^2}{2} - 1 \right]^2 \quad (12)
\]

For minimizing MSE differentiate above equation with respect to ‘t’ and equating to zero, we get

\[
\frac{d\text{MSE}(Z_{1st})}{dt} = 0
\]

\[
t^2 \left\{ \sum_{h=1}^{k} P_h^2 Z_{1h} \left( \frac{1 - f_h}{n_h} \right) \left[ C_{hz1}^2 + C_{hz2}^2 - 2\rho_{hz1z2} C_{hz1} C_{hz2} \right] + Z_1^2 \right\} = 2Z_1^2
\]

We obtained “\(t^2\)” from the above equation as follows:

\[
t^2 = \frac{2Z_1^2}{\sum_{h=1}^{k} P_h^2 Z_{1h} \left( \frac{1 - f_h}{n_h} \right) \left[ C_{hz1}^2 + C_{hz2}^2 - 2\rho_{hz1z2} C_{hz1} C_{hz2} \right]}
\]

where, \(t\) lies between 0 and 1.

5 Empirical Studies

In the empirical study, we compare our proposed estimator with traditional Separate ratio estimator, Combined Ratio estimator and the other estimators listed below. Additionally, we examine various conditions as outlined below:

From Equations (1) and (12), we get

\[
\text{MSE}(Z_{1st}) < \text{MSE}(Z_{1SR})
\]

\[
Z_1^2 \left[ \frac{t^2}{2} - 1 \right]^2 < \sum_{h=1}^{k} P_h^2 Z_{1h} \left( \frac{1 - f_h}{n_h} \right) \\
\times [C_{hz1}^2 + C_{hz2}^2 - 2\rho_{hz1z2} C_{hz1} C_{hz2}] \left( 1 - \frac{t^4}{4} \right) \quad (13)
\]
From (2) and (12), we have the condition as

\[
\text{MSE}(\bar{Z}_{1_{\text{st}}}) < \text{MSE}(\bar{Z}_{1_{\text{RS}}})
\]

\[
\bar{Z}_1^2 \left[ \frac{t^2}{2} - 1 \right]^2 < \sum_{h=1}^{k} P_h^2 \bar{Z}_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\left\{ [C_{hz_1}^2 + \lambda_{1h}^2 C_{hz_2}^2 - 2 \rho_{hz_1z_2} \lambda_{1h} C_{hz_1} C_{hz_2}] - \frac{t^4}{4} [C_{hz_1}^2 + C_{hz_2}^2 - 2 \rho_{hz_1z_2} C_{hz_1} C_{hz_2}] \right\}
\]

(14)

From (3) and (12),

\[
\text{MSE}(\bar{Z}_{1_{\text{st}}}) < \text{MSE}(\bar{Z}_{1_{\text{RS}}})
\]

\[
\bar{Z}_1^2 \left[ \frac{t^2}{2} - 1 \right]^2 < \sum_{h=1}^{k} P_h^2 \bar{Z}_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\times \left\{ [C_{hz_1}^2 + \lambda_{2h}^2 C_{hz_2}^2 - 2 \rho_{hz_1z_2} \lambda_{2h} C_{hz_1} C_{hz_2}] - \frac{t^4}{4} [C_{hz_1}^2 + C_{hz_2}^2 - 2 \rho_{hz_1z_2} C_{hz_1} C_{hz_2}] \right\}
\]

(15)

From Equations (4) and (12),

\[
\text{MSE}(\bar{Z}_{1_{\text{st}}}) < \text{MSE}(\bar{Z}_{1_{\text{US1}}})
\]

\[
\bar{Z}_1^2 \left[ \frac{t^2}{2} - 1 \right]^2 < \sum_{h=1}^{k} P_h^2 \bar{Z}_{1h}^2 \left( \frac{1 - f_h}{n_h} \right)
\]

\[
\times \left\{ [C_{hz_1}^2 + \lambda_{3h}^2 C_{hz_2}^2 - 2 \rho_{hz_1z_2} \lambda_{3h} C_{hz_1} C_{hz_2}] - \frac{t^4}{4} [C_{hz_1}^2 + C_{hz_2}^2 - 2 \rho_{hz_1z_2} C_{hz_1} C_{hz_2}] \right\}
\]

(16)
From (5) and (12), we have condition as

\[
\text{MSE}(Z_{1st}) < \text{MSE}(Z_{1RS})
\]

\[
\text{MSE}(Z_{1st}) = \sum_{h=1}^{k} p_h^2 Z_{1th}^2 \left( \frac{1 - f_h}{n_h} \right) \times \left\{ C_{hz_1}^2 + \lambda_{hz_2}^2 C_{hz_2}^2 - 2\rho_{hz_12} \lambda_{hz_1} C_{hz_1} C_{hz_2} \right\}
\]

\[
- \frac{t^4}{4} \left[ C_{hz_1}^2 + C_{hz_2}^2 - 2\rho_{hz_12} C_{hz_1} C_{hz_2} \right] \right\} \tag{17}
\]

From (6) and (12), we have condition as

\[
\text{MSE}(Z_{1st}) < \text{MSE}(Z_{1RC})
\]

\[
\frac{t^4}{4} \sum_{h=1}^{k} p_h^2 Z_{1th}^2 \left( \frac{1 - f_h}{n_h} \right) \left[ C_{hz_1}^2 + C_{hz_2}^2 - 2\rho_{hz_12} C_{hz_1} C_{hz_2} \right]
\]

\[
+ Z_{1}^2 \left[ \frac{t^2}{2} - 1 \right]^2 \]

\[
< \sum_{h=1}^{k} p_h^2 \left( \frac{1 - f_h}{n_h} \right) \left[ S_{hz_1}^2 + R^2 S_{hz_2}^2 - 2R S_{hz_12} \right] \tag{18}
\]

From Equations (7) and (12), we have

\[
\text{MSE}(Z_{1st}) < \text{MSE}(Z_{1KC})
\]

\[
\frac{t^4}{4} \sum_{h=1}^{k} p_h^2 Z_{1th}^2 \left( \frac{1 - f_h}{n_h} \right) \left[ C_{hz_1}^2 + C_{hz_2}^2 - 2\rho_{hz_12} C_{hz_1} C_{hz_2} \right]
\]

\[
+ Z_{1}^2 \left[ \frac{t^2}{2} - 1 \right]^2 \]

\[
< m^* \sum_{h=1}^{k} p_h^2 \left( \frac{1 - f_h}{n_h} \right) \left[ S_{hz_1}^2 + R^2 S_{hz_2}^2 - 2R S_{hz_12} \right]
\]

\[
+ (m^* - 1)^2 Z_{1}^2 \tag{19}
\]
6 Application of Real-Life Data Set

We used ratio estimators to analyze data from the number of apple trees and the amount of apple output in 854 villages throughout Turkey in 1999 (Source: Institute of Statistics, Republic of Turkey). First, we divided the data into six strata based on the various areas of Turkey, and then we randomly chose samples (villages) from each stratum (region). Neyman allocation (Cochran, 1977) is used

\[ n_h = n \frac{N_h S_h}{\sum_{h=1}^{k} N_h S_h} \]

and a total of 140 units selected as sample from a population of 854, the data represented in Table 1.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Estimator</th>
<th>MSE</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>( Z_{ISR} )</td>
<td>165108.3</td>
<td>100</td>
</tr>
<tr>
<td>2.</td>
<td>( Z_{ISR} )</td>
<td>165112.2</td>
<td>99.99</td>
</tr>
<tr>
<td>3.</td>
<td>( Z_{ISR} )</td>
<td>165335.7</td>
<td>99.86</td>
</tr>
<tr>
<td>4.</td>
<td>( Z_{ISR} )</td>
<td>165100.4</td>
<td>100.00</td>
</tr>
</tbody>
</table>

The suggested estimator outperforms the separate ratio, combined ratio, and taken into account existing estimators in terms of percentage relative efficiency (PRE). Table 2 shows that the proposed estimator’s MSE value is significantly lower than that of the other estimators. Further, it can be concluded that, in Table 2, the estimators 6 & 7 are combined ratio type estimators. It is observed that, they have highest PRE values than others.
7 Conclusion

In this research, we have introduced a novel separate-type ratio estimator designed for application within stratified random sampling. Our investigation involved a comprehensive comparative analysis, contrasting this newly developed estimator with traditional separate ratio estimators, combined ratio estimators, and certain existing estimators that were previously modified by Rajesh and Lone in 2014. Utilizing a real-world dataset, we empirically validated the efficiency and effectiveness of our proposed estimator.

The findings, as presented in Table 2, prominently demonstrate the superiority of the proposed estimator. Specifically, our estimator consistently yielded lower Mean Squared Error (MSE) values, affirming its ability in achieving higher precision in estimation. Moreover, these results unequivocally underscore the dominance of separate ratio estimators over their combined ratio counterparts, reinforcing the significance of this method in optimizing accuracy within the realm of stratified random sampling.

References


Biographies

G. R. V. Triveni currently a Ph.D Research Scholar in the Department of Mathematics, School of Advanced Sciences, VIT-AP University, Amaravati, Andhra Pradesh, India. Her Research area is Sampling theory. She had done M.Sc from Andhra University, Andhra Pradesh. She possessed a limited knowledge of R software.

Faizan Danish is currently working as an Assistant Professor (Statistics) in VIT-AP University, Vijayawada Andhra Pradesh, India and has done Ph.D. from Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, J&K, India. He has worked as Postdoctoral Fellow in New York University, New York, USA. Dr. Danish has around 3 years of Teaching experience and published around 25 research papers in reputed journals. Further, he has worked as Biostatistician under Research Consultation Services, Doha Qatar. He has research expertise in Sampling Theory, Mathematical, Applied Statistics and Biostatistics. Dr. Danish is well versed with Statistical Software’s: R, STATA, SPSS, Fortran 77, O.P STAT, WINDOSTAT, Mathematica.