Original Research Paper

Volume - 11 | Issue - 03 | March - 2021 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Statistics

A MODIFIED RATIO TYPE ESTIMATOR OF FINITE POPULATION MEAN UNDER STRATIFIED RANDOM SAMPLING SCHEME

Naziru Fadisanku Haruna

Prof Ran Vijay Kumar Singh* *Corresponding Author

Samsudeen Dahiru

ABSTRACT In This paper a modified ratio-type estimator for finite population mean under stratified random sampling using single auxiliary variable has been proposed. The expression for mean square error and bias of the proposed estimator are derived up to the first order of approximation. The expression for minimum mean square error of proposed estimator is also obtained. The mean square error the proposed estimator is compared with other existing estimators theoretically and condition are obtained under which proposed estimator performed better. A real life population data set has been considered to compare the efficiency of the proposed estimator numerically.

KEYWORDS : Auxiliary variable, Bias, Mean square error, Ratio estimator, Stratified Random Sampling

1. INTRODUCTION

The problems of estimation of population parameters and search of the efficient estimators in present day have brought about rapid development in the theory and practice of sample surveys. An important objective in any estimation problem is to obtain an estimator of a population parameter which can take care of the salient features of the population. If the population is homogeneous with respect to the characteristic under study, then the method of simple random sampling will yield a homogeneous sample and in turn, the sample mean will serve as a good estimator of population mean. Moreover, the variance of sample mean not only depends on the sample size and sampling fraction but also on the population variance. In order to increase the precision of an estimator, there is need to use a sampling scheme which can reduces the heterogeneity in the population. If the population is heterogeneous with respect to the characteristic under study, then best sampling procedure is stratified random sampling.

In the theory of survey sampling, the auxiliary information is frequently used to increase the precision of estimators. In many such approaches and practices, the classical ratio, product and regression type estimators are widely used for estimating the unknown population parameters, provided there exists a sufficient correlation between the study variable and the auxiliary variable. Auxiliary variable have been used by various authors in various estimation situation. In sample surveys the scientific technique for selecting a sample is that of selecting a probability sample that is usually based upon a stratification of the population. Stratification is one of the design tools that give increased precision. In the progression for improving the performance of the ratio estimators, authors have proposed various improved ratio-type estimators in stratified sampling. Notably among them include Sisodia and Dwivedi (1981), Singh and Kakran (1993), Upadhyaya and Singh (1999), Singh and Singh (2007), Vishwakarma and Singh (2011), Singh and Ahmed (2013) , Singh (2015). The work incorporated in this research is to propose a modified ratio type estimator of finite population mean under stratified random sampling scheme to estimate the finite population mean of characteristics under study.

2. Notation

Let U = $\{u_{1,\dots}u_n\}$ Be A Finite Population Of Size *N* And The Units Are Partitioned Into K Distinct Strata With Hth Stratum Containing *Hn* Units (h = 1 . . . K) Such That $\sum_{h=1}^{k} N_h = N$.

Let A Sample Of Size *Hn* Units (h = 1. . . K) Be Drawn From The Population Using A Simple Random Sampling Without Replacement Such That $\sum_{h=1}^{k} n_h = n$. Let (y_{hi}, x_{hi}) be the observed values of (y, x) on the *thi*unit of the *thh*stratum (i = 1, 2, . . . ,Nh).

The population means of the variables *Y* and *X* in the *thh* stratum are

$$\bar{Y}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} Y_{hi}\,,\ \bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} X_{hi}.$$

And the corresponding sample means of the variables Y and X in the h^{th} stratum are;

$$\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi}, \bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$$

The sample means of the variables Y and X in stratified random sampling, is given by

$\bar{y}_{st} = \sum_{h=1}^{k} W_h \bar{y}_h$, and $\bar{y}_{st} = \sum_{h=1}^{k} W_h \bar{x}_h$, where $W_h = \frac{N_h}{N}$ denotes the stratum weight

Also, $\bar{y}_{st}$ and $\bar{x}_{st}$ are the unbiased estimators of the population means $\bar{Y} = \sum_{h=1}^{k} W_h \bar{Y}_h$ and $\bar{X} = \sum_{h=1}^{k} W_h \bar{X}_h$ respectively.

Let ρ_{YXh} be the correlation coefficient between the study variable and the auxiliary variable in hth stratum. Also, $C_{Xh} = \frac{S_{Xh}}{\bar{X}_h}$ and $C_{Yh} = \frac{S_{Yh}}{\bar{Y}_h}$ are the coefficient of variation the study variable Y and the auxiliary variable X in hth stratum.

3. Literature Review

3.1 Separate and combined ratio estimators under stratified sampling;

The stratified sample mean $\bar{y}_{st}$ under SRSWOR is given by;

$$\bar{y}_{st} = \sum_{h=1}^{k} W_h \bar{y}_h \tag{1}$$

Which is an unbiased estimator of population mean with variance

$$Var\left(\bar{y}_{st}\right) = \sum_{h=1}^{k} W_h^2 \theta_h S_{Yh}^2 \quad \text{where } \theta_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \tag{2}$$

By utilizing the concept of ratio method of estimation, the separate ratio estimator for the population mean $\bar{Y}$ is defined as;

$$\bar{y}_{SR} = \sum_{h=1}^{k} W_h \bar{y}_h \frac{\bar{X}_h}{\bar{x}_h} \tag{3}$$

The bias and MSE of separate ratio estimator are given by

$$B\left(\bar{y}_{SR}\right) = \sum_{h=1}^{k} W_h \theta_h \bar{Y}_h \left(C_{Xh}^2 - \rho_h C_{Yh} C_{Xh}\right) \tag{4}$$

$$MSE\left(\bar{y}_{SR}\right) = \sum_{h=1}^{k} W_h^2 \theta_h \left(S_{Yh}^2 + R_h^2 S_{Xh}^2 - 2R_h S_{YXh}\right), \quad \text{where } R_h = \frac{\bar{Y}_h}{\bar{X}_h} \tag{5}$$

Combined ratio estimator for the population mean Y is defined as;

$$y_{CR} \quad \left(\frac{y_{st}}{x_{st}}\right) X \tag{6}$$

and its Bias and MSE are given as;

$$B\left(\bar{y}_{CR}\right)=\frac{1}{\bar{X}}\sum_{h=1}^{K}W_h^2\theta_h\left(RS_{Xh}^2-S_{YXh}\right) \tag{7}$$

$$MSE\left(\bar{y}_{CR}\right)=\sum_{h=1}^{K}W_h^2\theta_h\left(S_{Yh}^2+R^2S_{Xh}^2-2R\,S_{YXh}\right); \tag{8}$$

3.2 Some existing modified ratio-type estimators in stratified sampling

Kadilar and Cingi (2003) developed ratio-types estimators in stratified random sampling which are as follows;

Sisodia and Dwivedi (1981) estimator was modified by Kadilar and Cingi (2003) under stratified random sampling as;

$$y_{stSD} \quad y_{st}\frac{\sum_{h=1}^{K}W_h\left(X_h+C_{Xh}\right)}{\sum_{h=1}^{K}W_h\left(x_h+C_{Xh}\right)} \tag{9}$$

and its MSE is given as;

$$MSE\left(y_{stSD}\right)=\sum_{h=1}^{K}W_h^2\theta_h\left(S_{Yh}^2+\eta_h^2R_h^2S_{Xh}^2-2\eta_hR_hS_{YXh}\right) \tag{10}$$

Where; $\eta_h=\frac{\bar{X}_h}{\bar{X}_h+C_{Xh}}$

Singh and Kakran (1993) estimator was modified by kadilar and Cingi (2003) under stratified random sampling given as;

$$\bar{y}_{stSK}=\bar{y}_{st}\frac{\sum_{h=1}^{K}W_h\left(\bar{X}_h+\beta_{2h}(x)\right)}{\sum_{h=1}^{K}W_h\left(\bar{x}_h+\beta_{2h}(x)\right)} \tag{11}$$

and its MSE is given as;

$$MSE\left(y_{stSK}\right)-\sum_{h=1}^{L}W_h^2\theta_h\left(S_{Yh}^2+\vartheta_h^2R_h^2S_{Xh}^2-2\vartheta_hR_hS_{YXh}\right) \tag{12}$$

Where $\vartheta_h-\frac{X_h}{\bar{X}_h+\beta_{2h}(x)}$,

Upadhyaya and Singh (1999) estimators are modified by Kadilar and Cingi (2003) under stratified random sampling and given as;

$$\bar{y}_{stUS1}-\bar{y}_{st}\frac{\sum_{h=1}^{K}W_h\left(\bar{X}_h\beta_{2h}(x)+C_{Xh}\right)}{\sum_{h=1}^{L}W_h\left(x_h\beta_{2h}(x)+C_{Xh}\right)} \tag{13}$$

$$y_{stUS2}-y_{st}\frac{\sum_{h=1}^{K}W_h\left(X_hC_{Xh}+\beta_{2h}(x)\right)}{\sum_{h=1}^{K}W_h\left(\bar{x}_hC_{Xh}+\beta_{2h}(x)\right)} \tag{14}$$

MSE of both the estimators $\bar{y}_{stUS1}$ and $\bar{y}_{stUS2}$ are obtained as;

$$MSE\left(\bar{y}_{stUS1}\right)=\sum_{h=1}^{K}W_h^2\theta_h\left(S_{Yh}^2+\tau_h^2R_h^2S_{Xh}^2-2\tau_hR_hS_{YXh}\right) \tag{15}$$

$$MSE\left(\bar{y}_{stUS2}\right)=\sum_{h=1}^{K}W_h^2\theta_h\left(S_{Yh}^2+\phi_h^2R_h^2S_{Xh}^2-2\phi_hR_hS_{YXh}\right) \tag{16}$$

Where $\tau_h=\frac{\bar{X}_h\beta_{2h}(x)}{\bar{X}_h\beta_{2h}(x)+C_{Xh}}$ and $\phi_h=\frac{\bar{X}_hC_{Xh}}{\bar{X}_hC_{Xh}+\beta_{2h}(x)}$

Tailor *et al* (2011) suggested a ratio-cum-product estimator of finite population mean under stratified random sampling using information on coefficient of variation and coefficient of kurtosis of auxiliary variable in population as;

$$\bar{\bar{Y}}_T=\bar{y}_{st}\left[\alpha\left\{\frac{\sum_{h=1}^{k}W_h\left(\bar{X}_hC_{Xh}+\beta_{2h}(x)\right)}{\sum_{h=1}^{k}W_h\left(\bar{X}_hC_{Xh}+\beta_{2h}(x)\right)}\right\}+(1-\alpha)\left\{\frac{\sum_{h=1}^{k}W_h\left(\bar{X}_hC_{Xh}+\beta_{2h}(x)\right)}{\sum_{h=1}^{k}W_h\left(\bar{X}_hC_{Xh}+\beta_{2h}(x)\right)}\right\}\right] \tag{17}$$

The mean squared error of the proposed estimator is given as;

$$MSE\left(\bar{\bar{Y}}_T\right)=\sum_{h=1}^{k}W_h\gamma_h\left[S_{Yh}^2+2R_T(1-2\alpha)C_{Xh}S_{YXh}+R^2{}_T(1-2\alpha)^2C^2{}_{Xh}S_{Xh}^2\right]$$

$$R_T=\frac{\bar{Y}}{\sum_{h=1}^{k}W_h\left(\bar{X}_hC_{Xh}+\beta_{2h}(x)\right)}$$

and the optimum value of α is obtained as $\alpha_{opt}=\frac{1}{2}\left[1+\frac{\left\{\sum_{h=1}^{k}W_h^2\gamma_hC_{Xh}S_{YXh}\right\}}{\left\{R_T\sum_{h=1}^{k}W_h^2\gamma_hC_{Xh}^2S_{Xh}^2\right\}}\right]$

By substituting the optimum value of α, the minimum MSE of $\bar{\bar{Y}}_T$ is obtained as

$$\left(\bar{\bar{Y}}_T\right)=\sum_{h=1}^{k}W_h^2\gamma_hS_{Yh}^2-\frac{\left(\sum_{h=1}^{k}w_h^2\gamma_hC_{Xh}S_{YXh}\right)^2}{\sum_{h=1}^{k}W_h^2\gamma_hC_{Xh}^2S_{YXh}} \tag{18}$$

3.3 Adapted estimators

Prasad (1989) proposed a ratio type estimator under simple random sampling as

$$\bar{y}_{pr}=k\bar{y}_R=k\frac{\bar{y}}{\bar{x}}\bar{X}\text{ , where } k \text{ is constant} \tag{19}$$

The Bias and MSE of the estimator $\bar{y}_{pr}$ are given by;

$$Bias\left(\bar{y}_{pr}\right)=\bar{Y}\left[(k-1)+\left(\frac{1}{n}-\frac{1}{N}\right)\left\{C_X^2-\rho C_YC_X\right\}\right] \tag{20}$$

$$MSE\left(\bar{y}_{pr}\right)=\bar{Y}^2\left[(k-1)^2+\left(\frac{1}{n}-\frac{1}{N}\right)\left\{C_Y^2+C_X^2-\rho C_YC_X\right\}\right] \tag{21}$$

And optimum value of k is obtained as $k=\frac{1+\theta\rho C_YC_X}{1+\theta C_Y^2}$

Prasad (1989) estimator, was modified by Kadilar and Cingi (2005) under stratified random sampling as;

$$y_{KC} \quad K^{*}y_{RC} \tag{22}$$

And the Bias and MSE of the estimator are given as;

$$Bias\left(y_{KC}\right) \quad \left(k^{*}-1\right)Y+\frac{1}{X}\sum_{k=1}^{k}W_k\gamma_k\left(R_kS_{Xh}^2-k^{*}S_{YXh}\right) \tag{23}$$

$$MSE\left(y_{KC}\right) \quad k^{*2}\sum_{k=1}^{k}W_k^2\theta_h\left(S_{Yh}^2+R_k^2S_{Xk}^2-2R_hS_{YXh}\right)+\left(k^{*}-1\right)^2Y^2 \tag{24}$$

Where $k^{*}_{opt} \quad \frac{Y^2}{Y^2+\sum_{h=1}^{K}W_h^2\theta_h\left(S_{Yh}^2+R_k^2S_{Xh}^2-2R_hS_{YXh}\right)}$, $0<k^{*}<1$

4. Proposed estimator

Following Prasad (1989) and Kadilar and Cingi (2005) in this paper a modified separate ratio -type estimator has been suggested as;

$$\bar{y}_{drs} = \sum_{h=1}^{k} d_h W_h \frac{\bar{y}_h}{\bar{x}_h} \bar{X}_h \qquad (25)$$

Where d_h is constant for h=1, 2,…k

5. Bias and mean square error of proposed estimator

To, obtain the bias and mean square error of $\bar{y}_{drs}$ up to first order of approximation let us assume that

$$\bar{y}_h = \bar{Y}_h(1+e_{oh}) \text{ and } \bar{x}_h = \bar{X}_h(1+e_{1h}) \qquad (26)$$

Then $E(e_{oh}) = E(e_{1h}) = 0$ and further we have

$$\left.\begin{aligned} E(e_{oh}^2) &= \left(\frac{1}{n_h} - \frac{1}{N_n}\right)\frac{S_{Yh}^2}{\bar{Y}_h^2} = \theta_h C_{Yh}^2 \\ E(e_{1h}^2) &= \left(\frac{1}{n_h} - \frac{1}{N_n}\right)\frac{S_{Xh}^2}{\bar{X}_h^2} = \theta_h C_{Xh}^2 \\ E(e_{oh}e_{1h}) &= \left(\frac{1}{n_h} - \frac{1}{N_n}\right)\frac{S_{YXh}}{\bar{X}\bar{Y}_h} = \theta_h \rho_h C_{Yh} C_{Xh} \end{aligned}\right\} \qquad (27)$$

Now, substituting the values of $\bar{y}_h$ and $\bar{x}_h$ from (26) in (25), we have

$$\bar{y}_{drs} = \sum_{h=1}^{k} d_h W_h \bar{Y}_h (1+e_{oh})\left(\frac{\bar{X}_h}{\bar{X}_h(1+e_{1h})}\right)$$

$$\bar{y}_{drs} = \sum_{h=1}^{k} d_h W_h \bar{Y}_h (1+e_{oh})\left(1 - e_{1h} + e_{1h}^2 - \ldots\right)$$

$$\bar{y}_{drs} = \sum_{h=1}^{k} d_h W_h \bar{Y}_h \left(1 + e_{0h} - e_{1h} - e_{0h}e_{1h} + e_{1h}^2\right)$$

$$\bar{y}_{drs} - \bar{Y} = \sum_{h=1}^{k} W_h \bar{Y}_h \left\{d_h\left(1 + e_{0h} - e_{1h} - e_{0h}e_{1h} + e_{1h}^2\right) - 1\right\}$$

$$\text{Thus, } Bias(\bar{y}_{drs}) = \sum_{h=1}^{k} W_h \bar{Y}_h \left\{(d_h - 1) + \theta_h d_h\left(C_{Xh}^2 - \rho_h C_{Xh} C_{Yh}\right)\right\} \qquad (28)$$

The mean square error of proposed estimator ($\bar{y}_{drs}$) can be obtained as

$$MSE(\bar{y}_{drs}) = E\left[\sum_{h=1}^{k} W_h \bar{y}_h \left\{(d_h - 1) + d_h\left(e_{oh} - e_{1h} - e_{oh}e_{1h} + e_{1h}^2\right)\right\}\right]^2$$

Squaring, taking expectation inside and substituting the values from (27) we have

$$MSE(\bar{y}_{drs}) = \sum_{h=1}^{k} W_h^2 \bar{Y}_h^2 \left[(d_h - 1)^2 + \theta_h\left\{d_h^2\left(C_{Yh}^2 + 3C_{Xh}^2 - 4\rho_h C_{Xh} C_{Yh}\right) - 2d_h\left(C_{Xh}^2 - \rho_h C_{Yh} C_{Xh}\right)\right\}\right] \qquad (29)$$

and the optimum value of d_h is obtained as

$$d_h = \frac{1 + \theta_h\left(C_{Xh}^2 - \rho_h C_{Yh} C_{Xh}\right)}{1 + \theta_h\left(C_{Yh}^2 + 3C_{Xh}^2 - 4\rho_h C_{Yh} C_{Xh}\right)}$$

By substituting the optimum value of d_h in (29), the minimum mean square error of the proposed estimator $\bar{y}_{drs}$ is obtained as

$$MSE(\bar{y}_{drs})_{min} = \sum_{h=1}^{k} W_h^2 \bar{Y}_h^2 \theta_h \left(C_{Yh}^2 + C_{Xh}^2 - 2\rho_h C_{Xh} C_{Yh}\right) - \sum_{h=1}^{k} W_h \bar{Y}^2 \left\{\frac{(A_h - B_h)^2}{1 + A_h}\right\} \qquad (30)$$

Where $A_h = \theta_h\left(C_{Yh}^2 + 3C_{Xh}^2 - 4\rho_h C_{Yh} C_{Xh}\right)$ and $B_h = \theta_h\left(C_{Xh}^2 - \rho_h C_{Yh} C_{Xh}\right)$

6. Numerical efficiency comparison

Kadilar and Cingi (2003) data is used to compare the efficiency of proposed estimator numerically with other estimators under study. Table 1 showed data of apple production amount as study variable Y and number of apple trees as auxiliary variable X in 854 villages of Turkey in 1999 (Source: Institute of Statistics, Republic of Turkey). The data is stratified by region of Turkey and from each stratum (region) samples (villages) are selected by using Neyman allocation.

TABLE 1: Statistics of population

$N = 854$ $N_1 = 106$ $N_2 = 106$ $N_3 = 94$ $N_4 = 171$ $N_5 = 204$ $N_6 = 173$

$n = 140$ $n_1 = 9$ $n_2 = 17$ $n_3 = 38$ $n_4 = 67$ $n_5 = 7$ $n_6 = 2$

$\bar{X} = 37600$ $\bar{X}_1 = 24375$ $\bar{X}_2 = 27421$ $\bar{X}_3 = 72409$ $\bar{X}_4 = 74365$ $\bar{X}_5 = 26441$ $\bar{X}_6 = 9844$

$\bar{Y} = 2930$ $\bar{Y}_1 = 1536$ $\bar{Y}_2 = 2212$ $\bar{Y}_3 = 9384$ $\bar{Y}_4 = 5588$ $\bar{Y}_5 = 967$ $\bar{Y}_6 = 404$

$S_x = 144794$ $S_{x1} = 49189$ $S_{x2} = 57461$ $S_{x3} = 160757$ $S_{x4} = 285603$ $S_{x5} = 45403$ $S_{x6} = 18794$

$S_y = 17106$ $S_{y1} = 6425$ $S_{y2} = 11552$ $S_{y3} = 29907$ $S_{y4} = 28643$ $S_{y5} = 2390$ $S_{y6} = 946$

$R = 0.07793$ $\rho_1 = 0.82$ $\rho_2 = 0.86$ $\rho_3 = 0.90$ $\rho_4 = 0.99$ $\rho_5 = 0.71$ $\rho_6 = 0.89$

$K^* = 0.975$ $\gamma_1 = 0.102$ $\gamma_2 = 0.049$ $\gamma_3 = 0.016$ $\gamma_4 = 0.009$ $\gamma_5 = 0.138$ $\gamma_6 = 0.006$

$\gamma = 215710.432$ $\alpha_1^2 = 0.015$ $\alpha_2^2 = 0.015$ $\alpha_3^2 = 0.012$ $\alpha_4^2 = 0.04$ $\alpha_5^2 = 0.057$ $\alpha_6^2 = 0.041$

TABLE 2: Mean Square Error of the Existing Estimator

ESTIMATORS	MSE	PRE
$\bar{y}_{st}$	673477.7	100
$\bar{y}_{CR}$	212011.9	317.66
$\bar{y}_{CP}$	1824330	36.91
$\bar{y}_{SR}$	159137.4	423.20
$\bar{y}_{SP}$	1790757	37.61
$\bar{y}_{drs}$	135013.3	498.81

7. CONCLUSION

The result obtained is given in Table 2 The proposed estimator is more efficient than the other existing estimators by having smaller mean square errors and higher percent relative efficiency for the population data set and therefore it gives a better estimate than the other estimators.

REFERENCES

1. Kadilar, C and Cingi, H. (2003). Ratio Estimators in Stratified Random Sampling. *Biometrical Journal*, 45(2), 218-225.
2. Kadilar, C. and Cingi, H. (2005). A new Ratio Estimator in Stratified Random Sampling. *Communication in Statistics- Theory and Methods*, 34, 1-6.
3. Prasad, B. (1989). Some improved ratio-type estimators of population mean and ratio in finite population sample surveys, *Communications in Statistics*: *Theory and Methods* 18, 379–392.
4. Singh, H. P. and Kakran, M. S. (1993). A modified ratio estimator using coefficient of kurtosis of an auxiliary character. (unpublished).
5. Sisodia, B. V. S. and Dwivedi, V. K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable, *Journal of Indian Society of Agricultural Statistics* 33(1), 13-18.
6. Singh, R.V.K. and Singh B.K. (2007): Study of a class of Ratio-Type Estimator under Polynomial Regression Model. *Proceeding of Mathematical Society*, *BHU*, Vol 23.
7. Singh, R. V. K. and Ahmed, A. (2013). Efficiency of Ratio Estimators in Stratified Random Sampling Using Information on Auxiliary Attribute. *International Journal of Engineering Science and innovative Technology.* 2 (1), 166-172.
8. Singh, R. V. K. (2015). Improved Ratio Type Estimator for Population Mean, *International Journal of Scientific & Engineering Research*, 6 (7), 55-60.
9. Tailor, R., B. Sharma,B and Kim, G.M. (2011). A generalized ratio-cum-product estimator of finite population mean in stratified random sampling, *Commun. Korean Statist.Soc.* 18, 111–118.
10. Upadhyaya, L.N. and Singh, H. P. (1999).Use of transformed auxiliary variable estimating the finite population mean. *Biometrical Journal.* 41 (5), 627-636.
11. Vishwakarma. G.K. and Singh, H. P. (2011). Separate ratio-Product estimator for estimating Population mean using auxiliary information. *Journal of Statistical Theory and Applications*, 10(10), 653-664.