

# RATIO ESTIMATOR IN STRATIFIED RANDOM SAMPLING USING HODGES LEHMANN AS AUXILIARY INFORMATION

<sup>1</sup>S. Maqbool, <sup>2</sup>Mir Subzar, <sup>3</sup>M A Bhat

Division of Agricultural Statistics, SKUAST-Kashmir (190025), India.

**Abstract:** The present study was taken into consideration to propose new modified Ratio estimator in stratified random sampling using the auxiliary information of Hodges Lehmann. The properties associated with the proposed estimators are assessed by mean square error and bias. We also provide the Numerical example to support the theoretical results.

**Keywords:** Hodges Lehmann; Ratio-type estimators; stratified sampling; Mean square error; Efficiency.

## 1. INTRODUCTION

In practice, it is very common that the use of the auxiliary variable ( $x$  or  $z$ ) improves the efficiency of the estimators for the population parameters of the study variable ( $y$ ). The auxiliary information supplied by the auxiliary variables is used both at the design and estimation stages of the survey. We have used it in the estimation stage in this article. The auxiliary variable is highly correlated (positively or negatively) with the study variable. The ratio method of estimation providing ratio type estimators is used for the estimation of population parameters when the study variable and the auxiliary variables are highly positively correlated to each other; whereas, the product method of estimation giving product type estimators is used when  $y$  and  $x$  are highly negatively correlated to each other.

Following common notations of the stratified random sampling can be given by

### Notations

$N$	Population size	$n_h$	Sample size in stratum $h$
$N_h$	No. of units in stratum $h$	$W_h = \frac{N_h}{N}$	Stratum weight
$C_{xh}$	Coefficient of variation of stratum $h$ ,	$HL_h$	Hodges lemma stratum $h$
$\beta_{x^2h}$	Coefficient of kurtosis of stratum $h$ ,	$l$	No. of stratum
$\beta_{x^3h}$	Coefficient of skewness of stratum $h$		
$\bar{y}_h$	Sample mean of the study variable in stratum $h$ ,		
$\gamma_h = (1 - f_h)/n_h$ ,	$f_h = \frac{n_h}{N_h}$		sampling fraction in stratum $h$
$S_{yh}^2$	Variance of the study variable in stratum $h$ ,		
$\bar{x}_{st} = \sum_{h=1}^l W_h \bar{x}_h$	Is an unbiased estimator of $\bar{X}$		
$\bar{x}_h$	Sample mean of auxiliary variable in stratum $h$		
$S_{xh}^2$	Variance of the auxiliary variable in stratum $h$		
$S_{xyh}$	Covariance between auxiliary and study variables in stratum		

## 2. EXISTING ESTIMATOR IN STRATIFIED SAMPLING

### Combined ratio estimator in stratified sampling

$$\bar{y}_{RC-KC} = \bar{y}_{st} \frac{\sum_{h=1}^l W_h (\bar{X}_h + C_{xh})}{\sum_{h=1}^l W_h (\bar{x}_h + C_{xh})}. \quad (\text{Kadilar and Cingi, 2005})$$

$$\text{Bias}(\bar{y}_{RC\_KC}) = \sum_{h=1}^l W_h^2 \gamma_h \left( \frac{R_{KC}}{\bar{X}_{KC}} S_{xh}^2 - \frac{1}{\bar{X}_{KC}} S_{xyh} \right)$$

$$\text{MSE}(\bar{y}_{RC\_KC}) = \sum_{h=1}^l W_h^2 \gamma_h (S_{yh}^2 + R_{KC}^2 S_{xh}^2 - 2R_{KC} S_{xyh})$$

$$\text{Where } R_{KC} = \frac{\bar{Y}_{st}}{\bar{X}_{KC}} = \frac{\bar{Y}_{st}}{\sum_{h=1}^l W_h (\bar{X}_h + C_{xh})}$$

### 3. SUGGESTED ESTIMATOR

Motivated by the above mentioned estimators we propose following estimator using Hodges Lehmann as auxiliary information for estimating the population mean in survey sampling.

$$\bar{y}_{stSM} = \frac{\bar{y}_{st}}{(\bar{x}_{st} + HL_h)} (\bar{X} + HL_h).$$

$$\text{Bias}(\bar{y}_{stSM}) = \sum_{h=1}^l W_h^2 \gamma_h \left( \frac{R_{SM}}{\bar{X} + HL_h} S_{xh}^2 - \frac{S_{xyh}}{\bar{X} + HL_h} \right)$$

$$\text{MSE}(\bar{y}_{stSM}) = \sum_{h=1}^l W_h^2 \gamma_h (S_{yh}^2 + R_{SM}^2 S_{xh}^2 - 2R_{SM} S_{xyh})$$

$$\text{Where } R_{SM} = \frac{\bar{Y}_h}{(\bar{X}_h + HL_h)}.$$

### 4. EFFICIENCY COMPARISON

We compare here the suggested estimator with the existing estimators mentioned above. We will have the conditions as follows:

**Comparison of suggested estimator with Combined ratio estimator in stratified sampling given by Kadilar & Cingi, 2005.**

$$\text{MSE}(\bar{y}_{stSM}) < \text{MSE}(\bar{y}_{stRS})$$

$$\sum_{h=1}^l W_h^2 \gamma_h (S_{yh}^2 + R_{SM}^2 S_{xh}^2 - 2R_{SM} S_{xyh}) < \sum_{h=1}^l W_h^2 \gamma_h (S_{yh}^2 + R_{KC}^2 S_{xh}^2 - 2R_{KC} S_{xyh}) \quad (4.2.1)$$

$$\sum_{h=1}^l W_h^2 \gamma_h (R_{SM}^2 S_{xh}^2 - 2R_{SM} S_{xyh}) < \sum_{h=1}^l W_h^2 \gamma_h (R_{KC}^2 S_{xh}^2 - 2R_{KC} S_{xyh})$$

$$\text{Let } A = \sum_{h=1}^l W_h^2 \gamma_h S_{xyh} \quad \text{and} \quad B = \sum_{h=1}^l W_h^2 \gamma_h S_{xh}^2$$

Then (4.2.1) becomes

$$-2R_{SM}A + R_{SM}^2B < -2R_{KC}A + R_{KC}^2B$$

$$-2R_{SM}A + 2R_{KC}A + R_{SM}^2B - R_{KC}^2B < 0$$

$$-2A(R_{SM} - 2R_{KC}) + B(R_{SM}^2 - R_{KC}^2) < 0$$

$$-2A(R_{SM} - 2R_{KC}) + B(R_{SM} - R_{KC})(R_{SM} + R_{KC}) < 0$$

Where there are two conditions as follows

$$(i) \quad \text{When } (R_{SM} - R_{KC})(R_{SM} + R_{KC}) < 0$$

$$\frac{-2A}{R_{SM} + R_{KC}} + B < 0$$

$$B < \frac{2A}{R_{SM} + R_{KC}}$$

$$(ii) \quad \text{When } (R_{SM} - R_{KC})(R_{SM} + R_{KC}) > 0$$

$$\frac{-2A}{R_{SM} + R_{KC}} + B > 0$$

$$B > \frac{2A}{R_{SM} + R_{KC}}$$

## 5. EMPIRICAL STUDY

For the population, we use the data of cultivation and production of apple in district Baramulla of Kashmir in which the apple production (in tons) is denoted by Y (study variable) and number of apple trees are denoted by X (auxiliary variable, 1 unit = 100 trees) in 499 villages of the Baramulla region of Jammu and Kashmir in 2010-2011 (Source: RCM project, pilot survey for estimation of cultivation and production of apple in district Baramulla, RCM approved project). We apply the proposed and existing estimators to this data set and the data statistics of this population is given in Table 1. Statistical analysis of these estimators and percent relative efficiency is given in Table 2.

**Table 1:** Descriptive statistics of the population

Stratum 1	Stratum 2	Stratum 3	Stratum 4
$N_1 = 156$	$N_2 = 116$	$N_3 = 117$	$N_4 = 110$
$n_1 = 15$	$n_2 = 15$	$n_3 = 15$	$n_4 = 15$
$\bar{X}_1 = 10.317$	$\bar{X}_2 = 12.117$	$\bar{X}_3 = 6.053$	$\bar{X}_4 = 7.187$
$\bar{Y}_1 = 22.15$	$\bar{Y}_2 = 37.97$	$\bar{Y}_3 = 21.79$	$\bar{Y}_4 = 25.94$
$\beta_{x11} = 0.12$	$\beta_{x12} = 0.10$	$\beta_{x13} = 0.29$	$\beta_{x14} = 0.31$
$\beta_{x21} = 1.15$	$\beta_{x22} = 1.01$	$\beta_{x23} = 1.24$	$\beta_{x24} = 1.32$
$C_{x1} = 28.69$	$C_{x2} = 19.72$	$C_{x3} = 39.54$	$C_{x4} = 48.78$
$C_{y1} = 86.02$	$C_{y2} = 57.99$	$C_{y3} = 39.54$	$C_{y4} = 48.11$
$S_{x1} = 2.960$	$S_{x2} = 2.389$	$S_{x3} = 2.393$	$S_{x4} = 3.506$
$S_{y1} = 19.06$	$S_{y2} = 22.02$	$S_{y3} = 8.62$	$S_{y4} = 12.48$
$\rho_1 = 0.840$	$\rho_2 = 0.860$	$\rho_3 = 0.890$	$\rho_4 = 0.901$
$TM_1 = 10.338$	$TM_2 = 12.188$	$TM_3 = 6.073$	$TM_4 = 7.146$
$MR_1 = 10.175$	$MR_2 = 11.65$	$MR_3 = 5.925$	$MR_4 = 7.45$
$HL_1 = 10.2$	$HL_2 = 10.7$	$HL_3 = 6.1$	$HL_4 = 8.05$

**Table 2:** Statistical analysis and %RE of the estimators for the population

Estimators	Bias	Mean square error	%rRE
$\bar{y}_{RC\_KC}$	-0.147656756	438.1938213	100.00
$\bar{y}_{stSM}$	-0.333748038	162.5870631	269.51

## 6. CONCLUSION

From the above empirical study we assess that our proposed estimator is more efficient than the existing estimator by using the auxiliary information of Hodges Lehmann as their mean square error is much lower than the existing estimator and according to the percent relative efficiency criterion, hence preferred over the existing estimator for use in practical application.

## REFERENCES

- [1] Abid, M., Abbas, N., Nazir, H.Z. and Lin, Z. (2016a). Enhancing the mean ratio estimators for estimating population mean using non-conventional location parameters. *Revista Colombiana de Estadística*, 39(1), 63-79.
- [2] Cochran, W. G. (1977). *Sampling Techniques*, Third Edition, Wiley Eastern Limited.
- [3] Kadilar, C. and Cingi, H. (2003). Ratio estimators in stratified random sampling. *Biometrical journal*, 45(2): 218-225.
- [4] Kadilar, C. and Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, 151: 893-902.
- [5] Kadilar, C. and Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, 151: 893-902.
- [6] Kadilar, C. and Cingi, H. (2005). A new Ratio estimators in stratified random sampling. *Communication in statistics – theory and methods*, 34: 597-620.
- [7] Subzar, M., Raja, A. T., Maqbool, S. and Nazir, Nageena (2016). New Alternative to Ratio Estimator of population mean. *International Journal of Agricultural and Statistical Sciences*, 1: 221-225.
- [8] Upadhyaya, L.N. and Singh, H.P. (1999). Use of transformed auxiliary variable in estimating the finite population means. *Biometrical Journal*, 41(5): 627-636.