

5-1-2014

Separate Ratio-type Estimators of Population Mean in Stratified Random Sampling


Rajesh Tailor

Vikram University, Ujjain, M.P., India, tailorraj@gmail.com

Hilal A. Lone

Vikram University, Ujjain, M.P., India, hilalstat@gmail.com

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Tailor, Rajesh and Lone, Hilal A. (2014) "Separate Ratio-type Estimators of Population Mean in Stratified Random Sampling," *Journal of Modern Applied Statistical Methods*: Vol. 13 : Iss. 1 , Article 14.

DOI: 10.22237/jmasm/1398917580

Available at: <http://digitalcommons.wayne.edu/jmasm/vol13/iss1/14>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Separate Ratio-type Estimators of Population Mean in Stratified Random Sampling

Rajesh Tailor
Vikram University
Ujjain, M.P., India

Hilal A. Lone
Vikram, University
Ujjain, M.P., India

Separate ratio-type estimators for population mean with their properties are considered. Some separate ratio-type estimators for population mean using known parameters of auxiliary variate are proposed. The bias and mean squared error of the proposed estimators are obtained up to the first degree of approximation. It is shown that the proposed estimators are more efficient than unbiased estimators in stratified random sampling and usual separate ratio estimators under certain obtained conditions. To judge the merits of the proposed estimators, an empirical study was conducted.

Keywords: Finite population mean, separate ratio estimator, auxiliary variable, bias, mean squared error, stratified random sampling

Introduction

The use of auxiliary information improves the efficiency of estimators. Cochran (1940) used auxiliary information at the estimation stage and envisaged the ratio estimation method. This method provides a ratio estimator which assumes that the population mean of the auxiliary variate is known. The ratio estimator performs well when a study and auxiliary variate are positively correlated. When these variates are negatively correlated, Robson's (1957) product method, which was independently given by Murthy (1964), is used. Searls (1964) utilized the coefficient of variation of an auxiliary variate to estimate the population mean of a study variate. Based on the work of Searls (1964), Sisodia and Dwivedi (1981) used a coefficient of variation of an auxiliary variate. Singh, et al. (2004) proposed ratio and product type estimators using the coefficient of kurtosis of an auxiliary variate, whereas Upadhyaya and Singh (1999) utilized both the

Dr. Tailor is a Reader in the School of Studies in Statistics. Email him at tailorraj@gmail.com. H. Lone is a Research Scholar in the School of Studies in Statistics. Email at hilalstat@gmail.com.

SEPARATE RATIO TYPE ESTIMATORS OF POPULATION MEAN

coefficients of variation and kurtosis of an auxiliary variate. Kadilar and Cingi (2003), Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999) and Singh, et al. (2004) defined estimators in stratified random sampling. This article develops separate ratio-type estimators along the lines of Kadilar and Cingi (2003).

Consider a population U of size N consisting of units $U_1, U_2, U_3, \dots, U_N$. Let x and y be the auxiliary variate and study variate, respectively. If population U is divided into L homogenous strata of sizes n_h ($h=1, 2, 3 \dots, l$), and a sample of size n_h is drawn from the h^{th} stratum, then the usual separate ratio estimator for population mean \bar{Y} is defined as

$$\hat{Y}_{RS} = \sum_{h=1}^L W_h \bar{y}_h \left(\frac{\bar{X}_h}{\bar{x}_h} \right),$$

where \bar{x}_h is the sample mean of the auxiliary variate in stratum h , and \bar{y}_h is the sample mean of a study variate of interest in stratum h

To the first degree of approximation, the bias and mean squared of the usual separate ratio estimator are

$$B\left(\hat{Y}_{RS}\right) = \sum_{h=1}^L W_h \gamma_h \bar{Y}_h \left(C_{xh}^2 - \rho_{yxh} C_{xh} C_{yh} \right) \quad (1)$$

and

$$MSE\left(\hat{Y}_{RS}\right) = \sum_{h=1}^L W_h^2 \gamma_h \left(S_{yh}^2 + R_h^2 S_{xh}^2 - 2R_h S_{yxh} \right). \quad (2)$$

Proposed separate ratio-type estimator using coefficient of variation

Sisodia and Dwivedi (1981) defined a ratio-type estimator using coefficient of variation C_x of auxiliary variate x as

$$\hat{Y}_{SD} = \bar{y} \left(\frac{\bar{X} + C_x}{\bar{x} + C_x} \right) \quad (3)$$

Kadilar and Cingi (2003) further defined Sisodia and Dwivedi's (1981) estimator in stratified random sampling as

$$\hat{Y}_{SD}^{ST} = \bar{y}_{st} \left(\frac{\sum_{h=1}^L W_h (\bar{X}_h + C_{xh})}{\sum_{h=1}^L W_h (\bar{x}_h + C_{xh})} \right) \quad (4)$$

Motivated by Sisodia and Dwivedi (1981), Kadilar and Cingi (2003), suggested a separate ratio-type estimator using coefficient of variation C_{xh} of auxiliary variate in h^{th} stratum as

$$\hat{Y}_{RS}^{SD} = \sum_{h=1}^L W_h \bar{y}_h \left(\frac{\bar{X}_h + C_{xh}}{\bar{x}_h + C_{xh}} \right) \quad (5)$$

To obtain the bias and mean squared error of the proposed separate ratio-type estimator \hat{Y}_{RS}^{SD} :

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

such that $E(e_{oh}) = E(e_{1h}) = 0$ and

$$E(e_{oh}^2) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) C_{yh}^2 = \gamma_h C_{yh}^2 ,$$

$$E(e_{1h}^2) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) C_{xh}^2 = \gamma_h C_{xh}^2 ,$$

$$E(e_{oh} e_{1h}) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \rho_{yxh} C_{yh} C_{xh} = \gamma_h \rho_{yxh} C_{yh} C_{xh} .$$

Expressing \hat{Y}_{RS}^{SD} in terms of e_i 's results in

$$\hat{Y}_{RS}^{SD} = \sum_{h=1}^L W_h \bar{Y}_h + \sum_{h=1}^L W_h \bar{Y}_h (e_{oh} - \lambda_{1h} e_{1h} + \lambda_{1h}^2 e_{1h}^2 - \lambda_{1h} e_{oh} e_{1h}) .$$

SEPARATE RATIO TYPE ESTIMATORS OF POPULATION MEAN

Thus, the bias and mean squared error of the proposed ratio estimator \hat{Y}_{RS}^{SD} up to the first degree of approximation is obtained as

$$B\left(\hat{Y}_{RS}^{SD}\right) = \sum_{h=1}^L W_h \bar{Y}_h \gamma_h \left(\lambda_{1h}^2 C_{xh}^2 - \lambda_{1h} \rho_{yxh} C_{yh} C_{xh} \right) \quad (6)$$

$$MSE\left(\hat{Y}_{RS}^{SD}\right) = \sum_{h=1}^L W_h^2 \gamma_h \left(S_{yh}^2 + \lambda_{1h}^2 R_h^2 S_{xh}^2 - 2R_h \lambda_{1h} S_{yxh} \right) \quad (7)$$

where $\lambda_{1h} = \frac{\bar{X}_h}{\bar{X}_h + C_{xh}}$, $R_h = \frac{\bar{Y}_h}{\bar{X}_h}$ and $\rho_{yxh} = \frac{S_{yxh}}{S_{xh} S_{yh}}$.

Suggested separate ratio-type estimator using coefficient of kurtosis

Singh, et al. (2004) defined a modified ratio estimator using the coefficient of kurtosis $\beta_2(x)$ of an auxiliary variate x as

$$\hat{Y}_{SE} = \bar{y} \left(\frac{\bar{X} + \beta_2(x)}{\bar{x} + \beta_2(x)} \right) \quad (8)$$

Kadilar and Cingi (2003) defined Singh, et al's (2004) estimator in stratified random sampling as

$$\hat{Y}_{SE}^{ST} = \bar{y}_{st} \left(\frac{\sum_{h=1}^L W_h (\bar{X}_h + \beta_{2h}(x))}{\sum_{h=1}^L W_h (\bar{x}_h + \beta_{2h}(x))} \right) \quad (9)$$

Motivated by Kadilar and Cingi (2003) and Singh, et al. (2004), the proposed estimator using the coefficient of kurtosis $\beta_2(x)$ of auxiliary variate x in h^{th} stratum is

$$\hat{Y}_{RS}^{SE} = \sum_{h=1}^L W_h \bar{y}_h \left\{ \frac{\bar{X}_h + \beta_{2h}(x)}{\bar{x}_h + \beta_{2h}(x)} \right\} \quad (10)$$

The bias and mean squared error of the proposed estimator \hat{Y}_{RS}^{SE} are obtained as

$$B\left(\hat{Y}_{RS}^{SE}\right)=\sum_{h=1}^L W_h \bar{Y}_h \gamma_h \left(\lambda_{2h}^2 C_{xh}^2 - \lambda_{2h} \rho_{yxh} C_{yh} C_{xh}\right) \quad (11)$$

$$MSE\left(\hat{Y}_{RS}^{SE}\right)=\sum_{h=1}^L W_h^2 \gamma_h \left(S_{yh}^2 + \lambda_{2h}^2 R_h^2 S_{xh}^2 - 2R_h \lambda_{2h} S_{yxh}\right) \quad (12)$$

where $\lambda_{2h} = \frac{\bar{X}_h}{\bar{X}_h + \beta_{2h}(x)}$.

Proposed separate ratio-type estimator using coefficient of variation and coefficient of kurtosis

Upadhyaya and Singh (1999) suggested two different ratio-type estimators using the parameters coefficient of variation and coefficient of kurtosis as

$$\hat{Y}_{US1} = \bar{y} \left(\frac{\bar{X} \beta_2(x) + C_x}{\bar{x} \beta_2(x) + C_x} \right) \quad (13)$$

and

$$\hat{Y}_{US2} = \bar{y} \left(\frac{\bar{X} C_x + \beta_2(x)}{\bar{x} C_x + \beta_2(x)} \right) \quad (14)$$

Kadilar and Cingi (2003) defined Upadhyaya and Singh's (1999) estimators in stratified random sampling as

$$\hat{Y}_{US1}^{ST} = \bar{y}_{st} \left(\frac{\sum_{h=1}^L W_h \left(\bar{X}_h \beta_{2h}(x) + C_{xh} \right)}{\sum_{h=1}^L W_h \left(\bar{x}_h \beta_{2h}(x) + C_{xh} \right)} \right) \quad (15)$$

and

SEPARATE RATIO TYPE ESTIMATORS OF POPULATION MEAN

$$\hat{Y}_{US2}^{ST} = \bar{y}_{st} \frac{\sum_{h=1}^L W_h (\bar{X}_h C_{xh} + \beta_{2h}(x))}{\sum_{h=1}^L W_h (\bar{x}_h C_{xh} + \beta_{2h}(x))} \quad (16)$$

Based on Upadhyaya and Singh (1999) and Kadilar and Cingi (2003), the proposed separate ratio-type estimator using coefficients of kurtosis and variation in h^{th} stratum are

$$\hat{Y}_{RS}^{US1} = \sum_{h=1}^L W_h \bar{y}_h \left\{ \frac{\bar{X}_h \beta_{2h}(x) + C_{xh}}{\bar{x}_h \beta_{2h}(x) + C_{xh}} \right\} \quad (17)$$

and

$$\hat{Y}_{RS}^{US2} = \sum_{h=1}^L W_h \bar{y}_h \left\{ \frac{\bar{X}_h C_{xh} + \beta_{2h}(x)}{\bar{x}_h C_{xh} + \beta_{2h}(x)} \right\} \quad (18)$$

Using the standard procedure for finding the bias and mean squared errors shown previously, the bias and mean squared error of the proposed separate ratio-type estimators up to the first degree of approximation are obtained as:

$$B(\hat{Y}_{RS}^{US1}) = \sum_{h=1}^L W_h \bar{Y}_h \gamma_h (\lambda_{3h}^2 C_{xh}^2 - \lambda_{3h} \rho_{yxh} C_{yh} C_{xh}) \quad (19)$$

$$MSE(\hat{Y}_{RS}^{US1}) = \sum_{h=1}^L W_h^2 \gamma_h (S_{yh}^2 + \lambda_{3h}^2 R_h^2 S_{xh}^2 - 2R_h \lambda_{3h} S_{yxh}) \quad (20)$$

$$B(\hat{Y}_{RS}^{US2}) = \sum_{h=1}^L W_h \bar{Y}_h \gamma_h (\lambda_{4h}^2 C_{xh}^2 - \lambda_{4h} \rho_{yxh} C_{yh} C_{xh}) \quad (21)$$

$$MSE(\hat{Y}_{RS}^{US2}) = \sum_{h=1}^L W_h^2 \gamma_h (S_{yh}^2 + \lambda_{4h}^2 R_h^2 S_{xh}^2 - 2R_h \lambda_{4h} S_{yxh}) \quad (22)$$

where $\lambda_{3h} = \frac{\bar{X}_h \beta_{2h}(x)}{\bar{X}_h \beta_{2h}(x) + C_{xh}}$ and $\lambda_{4h} = \frac{\bar{X}_h C_{xh}}{\bar{X}_h C_{xh} + \beta_{2h}(x)}$.

Efficiency comparisons

The variance of the usual unbiased estimators in stratified random sampling is

$$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \gamma_h S_{yh}^2 \quad (23)$$

From (7) and (23), it is observed that the proposed estimator \hat{Y}_{RS}^{SD} would be more efficient than the usual unbiased estimator \bar{y}_{st} if

$$A < 2B. \quad (24)$$

A comparison of (7) and (2) shows that the proposed estimator \hat{Y}_{RS}^{SD} would be more efficient than the usual ratio estimator \hat{Y}_{RS} if

$$C < 2D. \quad (25)$$

Comparing (12) and (23), it is observed that the proposed estimator \hat{Y}_{RS}^{SE} would be more efficient than the usual unbiased estimator \bar{y}_{st} if

$$E < 2F. \quad (26)$$

From (12) and (2) it is observed that the proposed estimator \hat{Y}_{RS}^{SE} would be more efficient than the usual separate ratio estimator \hat{Y}_{RS} if

$$G < 2H. \quad (27)$$

Comparison of (20) and (23) shows that the proposed estimator \hat{Y}_{RS}^{US1} would be more efficient than the usual unbiased estimator \bar{y}_{st} if

$$L < 2M. \quad (28)$$

SEPARATE RATIO TYPE ESTIMATORS OF POPULATION MEAN

From (20) and (2), it is observed that the proposed estimator \hat{Y}_{RS}^{US1} would be more efficient than the usual separate ratio estimator \hat{Y}_{RS} if

$$N < 2P. \quad (29)$$

Comparison of (22) and (23) shows that the proposed estimator \hat{Y}_{RS}^{US2} would be more efficient than the usual unbiased estimator \bar{y}_{st} if

$$Q < 2R. \quad (30)$$

From (22) and (2), it is observed that the proposed estimator \hat{Y}_{RS}^{US2} would be more efficient than the usual separate ratio estimator \hat{Y}_{RS} if

$$S < 2T, \quad (31)$$

where $A = \sum_{h=1}^L W_h^2 \gamma_h \lambda_{1h}^2 S_{xh}^2 R_h^2$, $B = \sum_{h=1}^L W_h^2 \gamma_h R_h \lambda_{1h} S_{yhx}$,

$$C = \sum_{h=1}^L W_h^2 \gamma_h S_{xh}^2 R_h^2 (\lambda_{1h}^2 - 1), \quad D = \sum_{h=1}^L W_h^2 \gamma_h R_h S_{yhx} (\lambda_{1h} - 1), \quad E = \sum_{h=1}^L W_h^2 \gamma_h \lambda_{2h}^2 S_{xh}^2 R_h^2,$$

$$F = \sum_{h=1}^L W_h^2 \gamma_h R_h \lambda_{2h} S_{yhx}, \quad G = \sum_{h=1}^L W_h^2 \gamma_h S_{xh}^2 R_h^2 (\lambda_{2h}^2 - 1), \quad H = \sum_{h=1}^L W_h^2 \gamma_h R_h S_{yhx} (\lambda_{2h} - 1),$$

$$L = \sum_{h=1}^L W_h^2 \gamma_h \lambda_{3h}^2 S_{xh}^2 R_h^2, \quad M = \sum_{h=1}^L W_h^2 \gamma_h R_h \lambda_{3h} S_{yhx}, \quad N = \sum_{h=1}^L W_h^2 \gamma_h S_{xh}^2 R_h^2 (\lambda_{3h}^2 - 1),$$

$$P = \sum_{h=1}^L W_h^2 \gamma_h R_h S_{yhx} (\lambda_{3h} - 1), \quad Q = \sum_{h=1}^L W_h^2 \gamma_h \lambda_{4h}^2 S_{xh}^2 R_h^2, \quad R = \sum_{h=1}^L W_h^2 \gamma_h R_h \lambda_{4h} S_{yhx}$$

$$S = \sum_{h=1}^L W_h^2 \gamma_h S_{xh}^2 R_h^2 (\lambda_{4h}^2 - 1) \text{ and } T = \sum_{h=1}^L W_h^2 \gamma_h R_h S_{yhx} (\lambda_{4h} - 1).$$

Empirical study

To examine the performance of the proposed estimator in comparison to other estimators considered in this study, three natural population data sets were

considered (see Populations 1-3). The estimators based on the population data are compared in Table 1.

Population 1. (Singh and Mangat, 1996, p. 208)

	$n_1=14$	$n_2=9$	$n_3=12$	$n_4=17$
	$N_1=400$	$N_2=216$	$N_3=364$	$N_4=364$
	$\bar{X}_1=76.21$	$\bar{X}_2=58.11$	$\bar{X}_3=69.08$	$\bar{X}_4=63.71$
	$\bar{Y}_1=79.35$	$\bar{Y}_2=59.44$	$\bar{Y}_3=76.66$	$\bar{Y}_4=64.57$
$N = 1344,$ $n = 52$	$\beta_{21}(x)=2.22$	$\beta_{22}(x)=2.29$	$\beta_{23}(x)=1.96$	$\beta_{24}(x)=2.47$
	$C_{x_1}=0.1906$	$C_{x_2}=0.2416$	$C_{x_3}=0.201$	$C_{x_4}=0.1908$
	$S_{x_1}^2=210.9938$	$S_{x_2}^2=197.1041$	$S_{x_3}^2=192.7954$	$S_{x_4}^2=147.7651$
	$S_{y_1}^2=166.70$	$S_{y_2}^2=174.28$	$S_{y_3}^2=226.60$	$S_{y_4}^2=170.61$
	$S_{yx_1}=148.76$	$S_{yx_2}=161.19$	$S_{yx_3}=192.21$	$S_{yx_4}=143.83$

Population 2. (Murthy, 1967, p. 228)

	$n_1=2$	$n_2=2$
	$N_1=5$	$N_2=5$
	$\bar{X}_1=214.4$	$\bar{X}_2=333.8$
	$\bar{Y}_1=1925.8$	$\bar{Y}_2=3115.6$
$N = 10,$ $n = 4$	$\beta_{21}(x)=1.88$	$\beta_{22}(x)=2.32$
	$\rho_{yx_1}=0.85$	$\rho_{yx_2}=0.98$
	$C_{x_1}=0.34$	$C_{x_2}=0.19$
	$S_{x_1}^2=5605.84$	$S_{x_2}^2=4401.76$
	$S_{y_1}^2=379360.16$	$S_{y_2}^2=115860.24$
	$S_{yx_1}=39360.69$	$S_{yx_2}=22356.52$

SEPARATE RATIO TYPE ESTIMATORS OF POPULATION MEAN

Population 3. (Singh and Mangat, 1996, p. 219)

N = 10, n = 4	$n_1 = 2$	$n_2 = 2$
	$N_1 = 5$	$N_2 = 5$
	$\bar{X}_1 = 214.4$	$\bar{X}_2 = 333.8$
	$\bar{Y}_1 = 1925.8$	$\bar{Y}_2 = 3115.6$
	$\beta_{21}(x) = 1.88$	$\beta_{22}(x) = 2.32$
	$\rho_{yx_1} = 0.85$	$\rho_{yx_2} = 0.98$
	$C_{x_1} = 0.34$	$C_{x_2} = 0.19$
	$S_{x_1}^2 = 5605.84$	$S_{x_2}^2 = 4401.76$
	$S_{y_1}^2 = 379360.16$	$S_{y_2}^2 = 115860.24$
	$S_{yx_1} = 39360.69$	$S_{yx_2} = 22356.52$

Table 1. Percent Relative Efficiency of \bar{y}_{st} , \hat{Y}_{RS} , \hat{Y}_{RS}^{SD} , \hat{Y}_{RS}^{SE} , \hat{Y}_{RS}^{US1} and \hat{Y}_{RS}^{US2} with respect to \bar{y}_{st}

Estimators	Population I	Population II	Population III
\bar{y}_{st}	100.00	100.00	100.00
\hat{Y}_{RS}	350.08	239.76	254.99
\hat{Y}_{RS}^{SD}	351.53	240.35	255.22
\hat{Y}_{RS}^{SE}	364.51	244.55	258.23
\hat{Y}_{RS}^{US1}	350.76	240.05	255.10
\hat{Y}_{RS}^{US2}	397.29	260.33	275.69

Conclusion

The conditions under which the proposed estimators have less mean squared error in comparison to the usual unbiased estimator in stratified random sampling and usual separate ratio estimator were described. Table 1 shows that the proposed estimators have the highest percent relative efficiency compared to the usual

unbiased estimator and separate ratio estimator, in all three populations. Thus, the proposed estimators \hat{Y}_{RS}^{SD} , \hat{Y}_{RS}^{SE} , \hat{Y}_{RS}^{US1} and \hat{Y}_{RS}^{US2} are recommended for use in practice for estimating the population mean when the described proper conditions are satisfied.

References

- Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio gain to produce. *Journal of Agricultural Science*, 30: 262-275.
- Kadilar, C., & Cingi, H. (2003). Ratio estimators in stratified sampling. *Biometrical Journal*, 45: 218-225.
- Murthy, M. N. (1967). *Sampling Theory and Methods*. Calcutta, India: Statistical Publishing Society.
- Robson, D. S. (1957). Application of multivariate polykeys to the theory of unbiased ratio-type estimation. *Journal of the American Statistical Association*, 52: 511-522.
- Searls, D. T. (1964). The utilization of a known coefficient of variation in the estimating procedure. *Journal of the American Statistical Association*, 59: 1225-1226. doi: 10.1080/01621459.1964.10480765
- Singh, H. P., Tailor, R., Tailor, R., & Kakran, M. S. (2004). An improved estimator of population mean using power transformation. *Journal of the Indian Society of Agricultural Statistics*, 58(2): 223-230.
- Singh, R., & Mangat, N. S. (1996). *Elements of survey sampling*. Boston, MA: Kluwer Academic.
- Sisodia, B. V. S., & Dwivedi, V. K. (1981). A modified Ratio Estimator using coefficient of variation of auxiliary variable. *Journal of the Indian Society of Agricultural Statistics*, 33(1): 13-18.
- Upadhyaya, L. N., & Singh, H. P. (1999). Use of transformed Auxiliary Variable in estimating the finite Population Mean. *Biometrical Journal*, 41(5): 627-636.