DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 11 | Issue 1 Article 26

5-1-2012

Ratio Type Estimator of Ratio of Two Population Means in Stratified Random Sampling

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

Sunil Chouhan
ShriVaishnav Institute of Management, Indore, M.P., India

Journal of Modern Applied Statistical Methods
May 2012, Vol. 11, No. 1, 279-283

1538 – 9472/12/$95.00

Ratio Type Estimator of Ratio of Two Population Means in Stratified Random Sampling

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

Sunil Chouhan
ShriVaishnav Institute of Management,
Indore, M.P., India

A ratio estimator is proposed for the ratio of two population means using auxiliary information in stratified random sampling. Bias and mean squared error expressions are obtained under large sample approximation, and the proposed estimator is compared both theoretically and empirically with the conventional estimator of ratio for two population means in stratified random sampling.

Key words: Ratio of two population means, auxiliary information, stratified random sampling, bias, mean squared error.

Introduction

The problem of estimating population means is considered in a wide range of applications in many areas of human activities. A ratio of two population means is also applicable in many situations, such as per hectare production of crops, per capita income and per kilometer population density. Many researchers have studied the estimation of the ratio of two population means in simple random sampling (Singh, 1965; Rao&Pareira, 1968; Shah & Shah, 1978: Ray & Singh, 1985; Upadhyaya& Singh, 1985; Upadhyaya, et al., 1985; Singh & Rani, 2005, 2006; Sindhu, et al., 2009). Other sampling designs have not attracted much attention; in many situations, it has been observed that stratified random sampling provides efficient estimators compared to those of simple random sampling. Thus, this article estimates the ratio of two population means under stratified random sampling.

Rajesh Tailor is a Reader in the School of Studies in Statistics. His research interests are in the field of sampling techniques. Email him at: tailorraj@gmail.com. Sunil Chouhan is an Assistant Professor at the Vaishnav Institute of Management, Indore, M. P. India. Email him at: chouhansd@gmail.com.

Consider a finite population $U = U_1, U_2, ..., U_N$ of size N. This population U is divided into L strata each of size N_h and sample of size n_h is drawn from each stratum such that $n = \sum_{h=1}^{L} n_h \ (h = 1,2,.....L)$. If y_0 and y_1 are the study variates, x is an auxiliary variate, and y_{0hi}, y_{1hi} and x_{hi} $(h = 1,2,....L;\ i = 1,2,....,N_h)$ are the observations taken from the i^{th} unit of the h^{th} stratum on study variates y_0 , y_1 and auxiliary variate x respectively, then the following are defined:

h^{th} stratum mean for study variate y_0 :

$$\bar{Y}_{0h} = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{0hi}$$

h^{th} stratum mean for study variate y_1 :

$$\bar{Y}_{1h} = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{1hi}$$

h^{th} stratum mean for auxiliary variate *x*:

$$\bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{1hi}$$

Population mean of study variate y_0 :

$$\bar{Y}_0 = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} y_{0hi} = \frac{1}{N}\sum_{h=1}^{L} N_h \bar{Y}_{0h} = \sum_{h=1}^{L} W_h \bar{Y}_{0h}$$

Population mean of study variate y_1 :

$$\bar{Y}_1 = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} y_{1hi} = \frac{1}{N}\sum_{h=1}^{L} N_h \bar{Y}_{1h} = \sum_{h=1}^{L} W_h \bar{Y}_{1h}$$

Population mean of auxiliary variate *x*:

$$\bar{X} = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} x_{hi} = \frac{1}{N}\sum_{h=1}^{L} W_h \bar{X}_h$$

Sample mean of study variate y_0 for h^{th} stratum:

$$\bar{y}_{0h} = \frac{1}{n_h}\sum_{i=1}^{n_h} \bar{y}_{0hi}$$

Sample mean of study variate y_1 for h^{th} stratum:

$$\bar{y}_{1h} = \frac{1}{n_h}\sum_{i=1}^{n_h} \bar{y}_{1hi}$$

Sample mean of auxiliary variate *x* for h^{th} stratum:

$$\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} \bar{x}_{hi}$$

Stratum weight of h^{th} stratum:

$$W_h = \frac{N_h}{N}$$

Ratio of two population means:

$$R = \frac{\bar{Y}_0}{\bar{Y}_1}$$

The usual unbiased estimators of population means $\bar{Y}_0$, $\bar{Y}_1$ and $\bar{X}$ are

$$\bar{y}_{0st} = \sum_{h=1}^{L} W_h \bar{y}_{0h} \qquad (1.1)$$

$$\bar{y}_{1st} = \sum_{h=1}^{L} W_h \bar{y}_{1h} \qquad (1.2)$$

$$\bar{x}_{st} = \sum_{h=1}^{L} W_h \bar{x}_h \qquad (1.3)$$

and the estimator of the ratio of two population means *R* in stratified random sampling is:

$$\hat{R}_{st} = \left(\frac{\bar{y}_{0st}}{\bar{y}_{1st}}\right).$$

Suggested Ratio Estimator

When the population mean $\bar{X}$ of auxiliary variate *x* is known, Singh (1965) suggested an estimator for *R* in simple random sampling as

$$T_1 = \left(\frac{\bar{y}_0}{\bar{y}_1}\right)\left(\frac{\bar{X}}{\bar{x}}\right) = R\left(\frac{\bar{X}}{\bar{x}}\right).$$

In stratified random sampling T_1 is defined as

$$T_{1st} = \left(\frac{\bar{y}_{0st}}{\bar{y}_{1st}}\right)\left(\frac{\bar{X}}{\bar{x}_{st}}\right)$$

In order to derive the bias and mean squared error expressions of estimators $\hat{R}_{st}$ and T_{1st} it is assumed that $\bar{y}_{0h} = \bar{Y}_{0h}(1+e_{0h})$, $\bar{y}_{1h} = \bar{Y}_{1h}(1+e_{1h})$, and $\bar{x}_h = \bar{X}_h(1+e_{2h})$, such that, $E(e_{0h}) = E(e_{1h}) = E(e_{2h}) = 0$.

To the first degree of approximation the bias and mean squared error are

$$B(\hat{R}_{st}) = R\sum_{h=1}^{L} W_h^2 \gamma_h \left(\frac{S_{1h}^2}{\overline{Y}_1^2} - \frac{S_{01h}}{\overline{Y}_0 \overline{Y}_1} \right) \tag{2.1}$$

$$B(T_{1st}) = R\sum_{h=1}^{L} W_h^2 \gamma_h \left(\frac{S_{1h}^2}{\overline{Y}_1^2} + \frac{S_{xh}^2}{\overline{X}^2} - \frac{S_{01h}}{\overline{Y}_0 \overline{Y}_1} - \frac{S_{0xh}}{\overline{Y}_0 \overline{X}} + \frac{S_{1xh}}{\overline{Y}_1 \overline{X}} \right) \tag{2.2}$$

$$MSE(\hat{R}_{st}) = R^2 \sum_{h=1}^{L} W_h^2 \gamma_h \left(\frac{S_{0h}^2}{\overline{Y}_0^2} + \frac{S_{1h}^2}{\overline{Y}_1^2} - 2\frac{S_{01h}}{\overline{Y}_0 \overline{Y}_1} \right) \tag{2.3}$$

$$MSE(T_{1st}) = R^2 \sum_{h=1}^{L} W_h^2 \gamma_h \left\{ \begin{array}{l} \frac{S_{0h}^2}{\overline{Y}_0^2} + \frac{S_{1h}^2}{\overline{Y}_1^2} + \frac{S_{xh}^2}{\overline{X}^2} \\ -2\left(\frac{S_{01h}}{\overline{Y}_0 \overline{Y}_1} + \frac{S_{0xh}}{\overline{Y}_0 \overline{X}} - \frac{S_{1xh}}{\overline{Y}_1 \overline{X}} \right) \end{array} \right\} \tag{2.4}$$

where

$$S_{0h}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{ohi} - \overline{y}_{oh} \right)^2 ,$$

$$S_{1h}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{1hi} - \overline{y}_{1h} \right)^2 ,$$

$$S_{xh}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(x_{hi} - \overline{x}_{h} \right)^2 ,$$

$$S_{01h} = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{0hi} - \overline{y}_{0h} \right)\left(y_{1hi} - \overline{y}_{1h} \right),$$

$$S_{0xh} = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{0hi} - \overline{y}_{0h} \right)\left(x_{hi} - \overline{x}_{h} \right)$$

and

$$S_{1xh} = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{1hi} - \overline{y}_{1h} \right)\left(x_{hi} - \overline{x}_{h} \right).$$

Efficiency Comparison of Estimators

A comparison of equations (2.3) and (2.4) shows that the suggested estimator T_{1st} will be more efficient than conventional estimator $\hat{R}_{st}$ if

$$MSE(T_{1st}) - MSE(\hat{R}_{st}) < 0$$

that is, if

$$\sum_{h=1}^{L} W_h^2 \gamma_h \left\{ \frac{S_{xh}^2}{\overline{X}^2} - 2\left(\frac{S_{0xh}}{\overline{Y}_0 \overline{X}} - \frac{S_{1xh}}{\overline{Y}_1 \overline{X}} \right) \right\} < 0 \tag{3.1}$$

If condition (3.1) is satisfied, the suggested estimator would be more efficient than the conventional estimator.

Empirical Study

Two natural datasets were used to compare the proposed estimator numerically (See Tables 1 and 2).

Conclusion

The theoretical comparison provides the condition under which the suggested estimator is more efficient than the conventional estimator $\hat{R}_{st}$. Table 3 shows that there is a considerable gain in efficiency by using the proposed ratio estimator T_{1ST} in comparison to the conventional estimator $\hat{R}_{st}$. Thus, if information regarding population mean $\overline{X}$ is available, the suggested estimator T_{1ST} is recommended for use in practice.

Table 1: Population 1 (y_0 : Number of Workers, y_1 : Fixed Capital, x : Output)

	$n_1 = 2$	$n_2 = 2$	$N_1 = 5$	$N_2 = 5$
N = 10 n = 4	$\overline{Y}_{01} = 51.80$	$\overline{Y}_{02} = 60.60$	$\overline{Y}_{11} = 214.4$	$\overline{Y}_{12} = 333.8$
	$\overline{X}_1 = 1925.8$	$\overline{X}_2 = 315.6$	$S_{y_{01}} = 0.75$	$S_{y_{02}} = 4.84$
	$S_{y_{11}} = 74.87$	$S_{y_{12}} = 66.35$	$S_{x1} = 615.92$	$S_{x_2} = 340.38$
	$S_{01_1} = 38.08$	$S_{01_2} = 287.92$	$S_{0x1} = 411.16$	$S_{0x2} = 1536.24$
	$S_{1x1} = 39360.68$	$S_{1x_2} = 22356.52$		

Source: Murthy, 1967

Table 2: Population 2 (y_0 : Area (in 000 Hectare), y_1 : Production (in 000 MT), x : Productivity (MT/Hectare))

	$n_1 = 4$	$n_2 = 4$	$N_1 = 10$	$N_2 = 10$
N = 20 n = 8	$\overline{Y}_{01} = 6.2$	$\overline{Y}_{02} = 80.67$	$\overline{Y}_{11} = 3.53$	$\overline{Y}_{12} = 111.61$
	$\overline{X}_1 = 0.5$	$\overline{X}_2 = 1.41$	$S_{y_{01}} = 1.2$	$S_{y_{02}} = 10.82$
	$S_{y_{11}} = 74.87$	$S_{y_{12}} = 66.35$	$S_{x1} = 615.92$	$S_{x_2} = 340.38$
	$S_{01_1} = 1.75$	$S_{01_2} = -92.02$	$S_{0x1} = -0.02$	$S_{0x2} = -7.04$
	$S_{1x1} = 1.60$	$S_{1x_2} = 144.87$		

Source: National Horticulture Board of India, Official Web Site, http://nhb.gov.in/statistics/area-production-statistics.html

Table 3: Percent Relative Efficiencies of $\widehat{R}_{st}$ and T_{1st} with respect to $\widehat{R}_{st}$

Estimator	$\widehat{R}_{st}$	T_{1st}
Population 1	100	9274.573
Population 2	100	400747.3

References

Murthy, & Singh, M. P. (1965).On the estimation of ratio and product of the population parameters. *Sankhya*, *B*, *27*, 321-328.

National Horticulture Board India. Official web site: http://nhb.gov.in/statistics/area-production-statistics.html.

Rao, J. N. K., & Pereira, N. P. (1968).On double ratio estimators. *Sankhya*, *A*, *30*, 83-90.

Ray, S. K., & Singh, R. K. (1985).Some estimators for the ratio and product of population parameters. *Journal of the Indian Society of Agricultural Statistics*, *37*(*1*), 1-10.

Shah, S. M., & Shah, D. N. (1978). Ratio-cum-product estimator for estimating ratio (product) of two population parameters. *Sankhya*, *C*, *40*(*2*), 156-166.

Sindhu, S. S., Tailor, R., & Singh, S. (2009). On the estimation of population proportion. *Applied Mathematical Science*, *3*(*35*), 1739-1744.

Singh, G. N., & Rani, R. (2005, 2006). Some linear transformations on auxiliary variable for estimating the ratio of two population means in sample surveys. *Model Assisted. Statistics and Applications*, *1*(*1*), *IOS Press*, 1-5.

Upadhyaya, L. N., & Singh, H. P. (1985). A class of estimators using auxiliary information for estimating ratio of two finite population means. *Gujarat Statistical Review*, *12*(2), 7-16.

Upadhyaya, L. N., Singh, H. P., &Vos, J. W. E. (1985). On the estimation of population means and ratios using supplementary information. *Statistica. Neerlandica*, *39*(*3*), 309-318.