STATISTICS IN TRANSITION new series, Autumn 2013
Vol. 14, No. 3, pp. 391–398

A RATIO-CUM-PRODUCT ESTIMATOR OF FINITE POPULATION MEAN IN SYSTEMATIC SAMPLING

Rajesh Tailor, Narendra K. Jatwa, Housila P. Singh[1]

ABSTRACT

In this paper we consider the problem of estimation of population mean using information on two auxiliary variables in systematic sampling. We have extended Singh (1967) estimator for estimation of population mean in systematic sampling. We have derived the expressions for the bias and mean squared error of the suggested estimator up to the first degree of approximation. We have compared the suggested estimator with existing estimators and obtained the conditions under which the suggested estimator is more efficient. An empirical study has been carried out to demonstrate the performance of the suggested estimator.

Key words: systematic sampling, ratio-cum-product estimator, bias, mean squared error.

1. Introduction

The use of auxiliary information for the estimation of population parameters in simple random sampling and stratified random sampling has been widely used. But in systematic sampling which is frequently used when up-to-date sampling frame is available the use of auxiliary information in estimation of population parameters has not been discussed much.

Cochran (1940) used auxiliary information at estimation stage and suggested a ratio estimator. The ratio estimator is more efficient when study and auxiliary variates are positively correlated and regression line passes through the origin. In case of negative correlation, Robson (1957) developed product method of estimation that provides a product estimator which is more efficient than the simple mean estimator. Hansen et al. (1946) developed a combined ratio estimator in stratified sampling. In systematic sampling, Swain (1964) defined a ratio estimator whereas Shukla (1971) suggested a product estimator. Many authors including Kushwaha and Singh (1989), Singh and Singh (1998) and Singh and Solanki (2012) discussed various estimators of population mean. Singh et al. (2011) suggested a general family of estimators for estimating population mean in

[1] S. S. in Statistics, Vikram University, Ujjain-456010, M. P., India.

systematic sampling using auxiliary information in the presence of missing observations. Singh and Jatwa (2012) suggested a class of exponential-type estimators in systematic sampling. Singh et al. (2011) studied some modified ratio and product estimators for population mean in systematic sampling.

Singh (1967) used information on population means of two auxiliary variates and developed a ratio-cum-product estimator in simple random sampling. Motivated by Singh (1967), we have suggested a ratio-cum-product estimator in systematic sampling.

Suppose N units in the population are numbered from 1 to N in some order. Select a sample of n units, if a unit at random is taken from the first k units and every k^{th} subsequent unit, then $N = nk$. This sampling method is similar to that of selecting a cluster at random out of k clusters (each cluster containing n units), made such that i^{th} Cluster contains serially numbered units $i, i+k, i+2k, ..., i+(n-1)k$. After sampling of n units, observe both the study variate y and the auxiliary variate x. Let y_{ij} and x_{ij} denote the observations regarding the variate y and the variate x respectively on the unit bearing the serial number $i+(j-1)k$ in the population $(i = 1, 2, ..., k; j = 1, 2, ..., n)$. If the i^{th} sampling unit is taken at random from the first k units, then $\bar{y}_{sy}$ and $\bar{x}_{sy}$ are defined as

$$\bar{y}_{sy} = \bar{y}_{i.} = \frac{1}{n}\sum_{j=1}^{n} y_{ij}, \quad \bar{x}_{sy} = \bar{x}_{i.} = \frac{1}{n}\sum_{j=1}^{n} x_{ij} .$$

The usual ratio estimator for estimating the population mean $\bar{Y}$ in systematic sampling given by Swain (1964) is defined as

$$\hat{\bar{Y}}_{R}^{sys} = \bar{y}_{sys}\left(\frac{\bar{X}}{\bar{x}_{sys}}\right), \tag{1.1}$$

where $\bar{x}_{sys} = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$ is an unbiased estimator of population mean $\bar{X} = \frac{1}{N}\sum_{j=1}^{N} x_{ij}$, the population mean of the auxiliary variate x. Here, $\bar{X}$ is assumed to be known.

Singh et al. (2011) suggested a ratio type exponential estimator for estimating the population mean $\bar{Y}$ in systematic sampling as

$$\bar{y}_{Resy} = \bar{y}_{sy} \exp\left(\frac{\bar{X} - \bar{x}_{sy}}{\bar{X} + \bar{x}_{sy}}\right). \tag{1.2}$$

When the study variate and auxiliary variate are negatively correlated, Shukla (1971) suggested a product estimator for population mean $\overline{Y}$ as

$$\hat{\overline{Y}}_P^{sys} = \overline{y}_{sys}\left(\frac{\overline{z}_{sys}}{\overline{Z}}\right). \tag{1.3}$$

Singh et al. (2011) suggested a product type exponential estimator for estimating the population mean $\overline{Y}$ in systematic sampling as

$$\overline{y}_{Pesy} = \overline{y}_{sy}\exp\left(\frac{\overline{x}_{sy} - \overline{X}}{\overline{x}_{sy} + \overline{X}}\right). \tag{1.4}$$

Variances of the ratio estimators $\hat{\overline{Y}}_R^{sys}$ and $\overline{y}_{Resy}$ up to the first degree of approximation are respectively given by

$$V(\hat{\overline{Y}}_R^{sys}) = \left(\frac{N-1}{nN}\right)\overline{Y}^2\left[\rho_y^* c_y^2 + \rho_x^* c_x^2\left(1 - 2k\sqrt{\rho^{**}}\right)\right] \tag{1.5}$$

and

$$V\left(\overline{y}_{Resy}\right) = \left(\frac{N-1}{nN}\right)\overline{Y}^2\left[\rho_y^* c_y^2 + \rho_x^*\left(c_x^2/4\right)\left(1 - 4k\sqrt{\rho^{**}}\right)\right], \tag{1.6}$$

where

$$k = \rho_{yx} c_y / c_x,\ \rho_y^* = \{1 + \rho_y(n-1)\},\ \ \rho_x^* = \{1 + \rho_x(n-1)\},\ \ \rho^{**} = \{\rho_y^* / \rho_x^*\},$$

$$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \overline{X})(y_{ij} - \overline{Y}),\ \ \rho_{yx} = \frac{S_{yx}}{S_y S_x}\ ,$$

$$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \overline{X})^2, \qquad S_y^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \overline{Y})^2,$$

$$S_z^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(z_{ij} - \overline{Z})^2,$$

and (ρ_y, ρ_x, ρ_z) being the intra-class correlation between the units of a cluster corresponding the (y,x,z) variates.

Further, the variances of the product estimator $\hat{\overline{Y}}_P^{sys}$ and $\overline{y}_{Resy}$ up to the first degree of approximation are respectively given by

$$V(\hat{\overline{Y}}_P^{sys}) = \left(\frac{N-1}{nN}\right)\overline{Y}^2\left[\rho_y^* c_y^2 + \rho_z^* c_z^2\left(1 + 2k^*\sqrt{\rho_2^{**}}\right)\right] \tag{1.7}$$

and

$$Var\left(\overline{y}_{Pesy}\right) = \left(\frac{N-1}{nN}\right)\overline{Y}^2\left[\rho_y^* c_y^2 + \rho_z^*\left(c_z^2/4\right)\left(1 + 4k^*\sqrt{\rho_2^{**}}\right)\right], \tag{1.8}$$

where

$$k^* = \rho_{yz}\{c_y / c_z\} \text{ and } \quad \rho_2^{**} = \{\rho_y^* / \rho_z^*\}$$

In the situation when the study variate y is positively correlated with the auxiliary variable x and negatively correlated with another auxiliary variable z we have suggested a ratio-cum-product estimator of the population mean $\bar{Y}$ in line with Singh (1967) in systematic sampling. We have derived the bias and mean squared error of the suggested estimator up to the first degree of approximation. Conditions are obtained under which the suggested estimator is better than the usual unbiased estimator, Swain (1964) ratio estimator, Shukla's (1971) product estimator and Singh et al. (2011) estimators. The result of this paper has been supported through an empirical study.

2. Suggested estimator

Motivated by Singh (1967) we have suggested a ratio-cum-product estimator in systematic sampling for population mean $\bar{Y}$ as

$$\hat{\bar{Y}}_{RP}^{sys} = \bar{y}_{sys}\left(\frac{\bar{X}}{\bar{x}_{sys}}\right)\left(\frac{\bar{z}_{sys}}{\bar{Z}}\right) \quad . \tag{2.1}$$

To obtain the bias and the mean squared error of the suggested estimator $\hat{\bar{Y}}_{RP}^{sys}$, we write

$\bar{y}_{sys} = \bar{Y}(1+e_0),\ \bar{x}_{sys} = \bar{X}(1+e_1),\ \bar{z}_{sys} = \bar{Z}(1+e_2)$ such that

$$E(e_0) = E(e_1) = E(e_2) = 0$$

and $$E({e_0}^2) = \theta c_y^2 \rho_y^*,\ E({e_1}^2) = \theta c_x^2 \rho_x^*,\ E({e_2}^2) = \theta c_z^2 \rho_z^*,$$

$$E(e_0 e_1) = \theta k c_x^2 \sqrt{\rho_y^* \rho_x^*}\ ,\ E(e_0 e_2) = \theta k^* c_z^2 \sqrt{\rho_y^* \rho_z^*}\ ,$$

$$E(e_1 e_2) = \theta k^{**} c_z^2 \sqrt{\rho_x^* \rho_z^*}.$$

where

$$c_z^2 = S_z^2 / \bar{Z}^2,\ \rho_z^* = \{1+\rho_z(n-1)\},\ k^{**} = \rho_{xz}(c_x / c_z),$$

$$S_z^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(z_{ij}-\bar{Z})^2,\ \rho_{yz} = S_{yz}/(S_y S_z),\ \rho_{xz} = S_{xz}/(S_x S_z),$$

$$S_{xz} = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{X})(z_{ij}-\bar{Z}),\ S_{yz} = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{Y})(z_{ij}-\bar{Z}) \text{ and}$$

$$\theta = \left(\frac{N-1}{nN}\right).$$

Expressing the suggested estimator $\hat{\overline{Y}}_{RP}^{sys}$ in terms of e's we have

$$\hat{\overline{Y}}_{RP}^{sys} = \overline{Y}(1+e_0)(1+e_1)^{-1}(1+e_2)\,. \tag{2.2}$$

We assume that $|e_1| < 1$ so that $(1+ e_1)^{-1}$ is expandable. Expanding that right-hand side, multiplying out and neglecting terms of e's having power greater than two we have

$$\hat{\overline{Y}}_{RP}^{sys} = \overline{Y}\left[1 + e_0 - e_1 + e_2 - e_o e_1 + e_o e_2 - e_1 e_2 + e_1^2\right]$$

or

$$\left(\hat{\overline{Y}}_{RP}^{sys} - \overline{Y}\right) = \overline{Y}\left[e_0 + e_1 + e_2 - e_o e_1 + e_o e_2 - e_1 e_2 + e_1^2\right]. \tag{2.3}$$

Taking expectation of both sides of (2.3) we get the bias of the estimator $\hat{\overline{Y}}_{RP}^{sys}$ to the first degree of approximation as

$$B(\hat{\overline{Y}}_{RP}^{sys}) = \left(\frac{N-1}{nN}\right)\overline{Y}\left[\rho_x^* c_x^2\left(1 - k\sqrt{\rho^{**}}\right) + \sqrt{\rho_z^*}\, c_z^2\left(k^*\sqrt{\rho_y^*} - k^{**}\sqrt{\rho_x^*}\right)\right]. \tag{2.4}$$

Squaring both sides of (2.3) and neglecting terms of e's having power greater than two we have

$$\left(\hat{\overline{Y}}_{RP}^{sys} - \overline{Y}\right)^2 = \left[e_0^2 - e_1^2 + e_2^2 - 2e_o e_1 + 2e_o e_2 - 2e_1 e_2\right]$$

Taking expectation of both sides of the above equation, we get the mean square error of $\hat{\overline{Y}}_{RP}^{sys}$ to the first degree of approximation as

$$MSE(\hat{\overline{Y}}_{RP}^{sys}) = \left(\frac{N-1}{nN}\right)\overline{Y}^2\left[\rho_y^* c_y^2 + \rho_x^* c_x^2\left(1 - 2k\sqrt{\rho^{**}}\right) + \sqrt{\rho_z^*}\, c_z^2\left(1 - 2k^{**}\sqrt{\rho_1^{**}}\right) + 2k^* c_z^2\sqrt{\rho_y^* \rho_z^*}\right], \tag{2.5}$$

where $\rho_1^{**} = \left(\rho_x^* / \rho_z^*\right)$.

3. Efficiency comparisons

In systematic sampling, the variance of the usual unbiased estimator $\overline{y}_{sys}$ is given by

$$V(\overline{y}_{sys}) = \left(\frac{N-1}{nN}\right)\overline{Y}^2 \rho_y^* c_y^2 \tag{3.1}$$

From (2.5), (1.5), (1.6), (1.7) and (1.8) it is observed that the suggested estimator $\hat{\overline{Y}}_{RP}^{sys}$ would be more efficient than

(i) $V(\bar{y}_{sys})$ if

$$\left[\rho_x^* c_x^2 + \left(1 - 2k\sqrt{\rho^{**}}\right) + \sqrt{\rho_z^*} c_z^2 \left(1 - 2k^{**}\sqrt{\rho_1^{**}}\right) + 2k^*\sqrt{\rho^{**}}\right] < 0 \tag{3.2}$$

(ii) $\hat{\bar{Y}}_R^{sys}$ if

$$d > 1/2 \tag{3.3}$$

(iii) $\hat{\bar{Y}}_P^{sys}$ if

$$d_1 < 1/2 \tag{3.4}$$

(iv) $\bar{y}_{Resy}$ if

$$\left[\rho_x^* c_x^2 \left(\frac{3}{4} - k\sqrt{\rho^{**}}\right) + \rho_z^* c_z^2 \left(1 - 2k^{**}\sqrt{\rho_1^{**}}\right) + 2k^*\sqrt{\rho_2^{**}}\right] < 0 \tag{3.5}$$

(v) $\bar{y}_{Pesy}$ if

$$\left[\rho_z^* c_z^2 \left(\frac{3}{4} + k^*\sqrt{\rho_2^{**}}\right) + \rho_x^* c_x^2 \left(1 - 2k\sqrt{\rho^{**}}\right) + 2k_1^*\sqrt{\rho_3^*}\right] < 0 \tag{3.6}$$

where $d = \left(k^{**}\sqrt{\rho_1^{**}} - k^*\sqrt{\rho_2^{**}}\right), d_1 = \left(k\sqrt{\rho^{**}} + k_1^*\sqrt{\rho_3^*}\right), \rho_1^{**} = (\rho_x^* / \rho_z^*)$, $\rho_2^{**} = (\rho_y^* / \rho_z^*), \rho_3^{**} = (\rho_z^* / \rho_x^*)$ and $k_1^* = \rho_{xz}(c_z / c_x)$.

Expressions (3.2), (3.3), (3.4), (3.5) and (3.6) provide the conditions under which the suggested estimator $\hat{\bar{Y}}_{RP}^{sys}$ would be more efficient than usual unbiased estimator $\bar{y}_{sys}$, ratio estimator $\hat{\bar{Y}}_R^{sys}$, product estimator $\hat{\bar{Y}}_P^{sys}$, ratio type exponential estimator $\bar{y}_{Resy}$ and product type exponential estimator $\bar{y}_{Pesy}$.

4. Empirical study

To illustrate the performance of the proposed ratio-cum-product estimator $\hat{\bar{Y}}_{RP}^{sys}$ relative to the usual unbiased estimator, we assumed (artificially) the following values of parameters:

$\bar{X} = 44.47$, $S_x^2 = 149.55$, $c_x = 0.28$, $S_{xy} = 538.57$, $\bar{Y} = 80$, $S_y^2 = 2000$, $c_y = 0.56$, $S_{yz} = -902.86$, $\bar{Z} = 48.40$, $S_z^2 = 427.83$, $c_z = 0.43$, $S_{xz} = -241.06$, $\rho_{xy} = 0.9848$ $\rho_{yz} = -0.9760$, $\rho_{zx} = -0.9530$, $\rho_x = 0.707$, $\rho_y = 0.6652$, $\rho_z = 0.5487$, N=15, n=3.

Table 4.1. Percent Relative Efficiency of $\bar{y}_{sys}$, $\hat{\bar{Y}}_{R}^{sys}$, $\hat{\bar{Y}}_{P}^{sys}$ and $\hat{\bar{Y}}_{RP}^{sys}$ with respect to $\bar{y}_{sys}$.

Estimator	PRE $(., \bar{y}_{sys})$
$\bar{y}_{sys}$	100.00
$\hat{\bar{Y}}_{R}^{sys}$	389.620
$\hat{\bar{Y}}_{p}^{sys}$	189.452
$\bar{y}_{Re\,sy}$	177.434
$\bar{y}_{Pesy}$	139.318
$\hat{\bar{Y}}_{RP}^{sys}$	777.790

5. Conclusions

Section 3 provides the conditions under which the suggested estimator has less mean squared errors in comparison to the simple mean estimator, the ratio estimator and the product estimator in systematic sampling. Table 4.1 exhibits that the suggested estimator $\hat{\bar{Y}}_{RP}^{sys}$ has the largest percent relative efficiency as compared to the simple mean estimator $\bar{y}_{sys}$, the ratio estimator $\hat{\bar{Y}}_{R}^{sys}$ and the product estimator $\hat{\bar{Y}}_{p}^{sys}$, the ratio type exponential estimator $\bar{y}_{Re\,sy}$ and the product type exponential estimator $\bar{y}_{Pesy}$. Thus, the suggested estimator $\hat{\bar{Y}}_{RP}^{sys}$ is recommended for its use in practice.

Acknowledgements

The authors are thankful to the Editor and to the learned referee for his valuable suggestions regarding improvement of the paper.

REFERENCES

COCHRAN, W. G., (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. J. Agri. Sci., 30, 262–275.

HANSEN, M. H., HURWITZ, W. N., GURNEY, M., (1946). Problems and methods of the sample survey of business. J. Amer. Statist. Assoc., 41, 173–189.

KUSHWAHA, K. S., SINGH H. P., (1989). Class of almost unbiased ratio and product estimators in systematic sampling. J. Ind. Soc. Agri. Statist, 41, 2, 193–205.

ROBSON, D. S., (1957). Application of multivariable polykays to the theory of unbiased ratio type estimation. J. Amer. Statist. Assoc., 50, 1225–1226.

SHUKLA, N. D., (1971). Systematic sampling and product method of estimation. In *Proceedings of all India Seminar on Demography and Statistics*. B.H.U. Varanasi: India.

SINGH, H. P., SINGH R., (1998). Almost unbiased ratio and product type estimators in systematic sampling. Questiio, 22, 3, 403–416.

SINGH, H. P., SOLANKI, R. S., (2012). An efficient class of estimators for the population mean using auxiliary information in systematic sampling. Journal of Statistical Theory and Practice, 6, (2), 274–285.

SINGH, R., MALIK, S., CHAUDHARY, M. K., VERMA, H., ADEWARA, A. A., (2012). A general family of ratio type estimators in systematic sampling. J. Reli. & Stat. Studs., 5, 1, 73–82.

SINGH, H. P., JATWA, N. K., (2012). A class of exponential type estimators in systematic sampling. Eco. Qulty. Control., 27, 195–208.

SINGH, H. P., TAILOR, R., JATWA, N. K., (2011). Modified ratio and product estimators for population mean in systematic sampling. J. Modern. Appl. Statist. Meth. 10 , 2, 424–435.

SWAIN, A. K. P. C., (1964). The use of systematic sampling ratio estimate. J. Ind. Statist. Assoc., 2, 160–164.

SINGH, M. P., (1967). Ratio-cum-product method of estimation. Metrika, 12, 34–42.