Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666,
Vol. 5, Issue 1 (2012): 65-71

DUAL TO RATIO-CUM-PRODUCT ESTIMATOR USING KNOWN PARAMETERS OF AUXILIARY VARIABLES

Rajesh Tailor, Ritesh Tailor[1], Rajesh Parmar and Manish Kumar[2]
School of Studies in Statistics, Vikram University, Ujjain, M.P., India
[1] Wood Properties and Uses Division, Institute of Wood Science and Technology (IWST), Bangalore, India
[2] Rain Forest Research Institute Jorhat, Assam, India
(Received February 28, 2012)

Abstract

This paper deals with the dual to ratio-cum-product estimator for population mean using known parameters of auxiliary variables. In this paper, dual to ratio-cum-product estimator of Singh and Tailor (2005) has been suggested. The Bias and mean squared error expressions have also been obtained up to the first degree of approximation. Suggested estimator has been compared theoretically as well as empirically.

Keywords: Dual, Ratio-cum-product, Bias, Mean Squared Error.

1. Introduction

Use of auxiliary information has shown its significance in improvement of efficiency of estimators of unknown population parameters. Cochran (1940) used auxiliary information in the form of population mean of auxiliary variate at estimation stage for the estimation of population parameters when study and auxiliary variates are positively correlated. In case of negative correlation between study variate and auxiliary variate, Robson (1957) defined product estimator for the estimation of population mean which was revisited by Murthy (1964). Ratio estimator performs better than simple mean estimator in case of positive correlation between study variate and auxiliary variate. Sisodia and Dwivedi (1981) used known value of coefficient of variation of auxiliary variate whereas Upadhya and Singh (1999). Singh and Tailor (2003) used both coefficient of kurtosis as well as coefficient of variation for estimating the population mean of study variate. Used correlation coefficient between study variate and auxiliary variate. Work done by above Cochran (1940), Robson (1957), Murthy (1964), Sisodia and Dwivedi (1981) and Singh and Tailor (2003) were based on use of single auxiliary variate.

Singh (1967) used information on two auxiliary variates and suggested a ratio-cum-product estimator for population mean. Singh and Tailor (2005) utilized correlation coefficient between study variate and auxiliary variate beside population mean of auxiliary variate and suggested improved ratio-cum-product estimator for population mean.

Srivenkataramana (1980) proposed dual to ratio and product estimators to estimate population mean. Singh et. al (2005) suggested dual to Singh (1967) ratio-cum-product estimator.

Singh et. al (2005) and Singh and Tailor (2005) motivates authors to propose dual to Singh and Tailor (2005) ratio-cum-product estimator for population mean and study its properties.

Consider a finite population U of size N consisting of units $U_1, U_2, ..., U_N$. Let y and (x, z) be the study variate and auxiliary variates respectively. A random sample of size n is drawn from U using simple random sampling without replacement. Let y_i and (x_i, z_i) are observations taken on study variate y and auxiliary variates (x, z) respectively.

The classical ratio and product estimators for population mean $\bar{Y}$ are

$$\bar{y}_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \tag{1.1}$$

and

$$\bar{y}_P = \bar{y}\left(\frac{\bar{z}}{\bar{Z}}\right) \tag{1.2}$$

where $\bar{y} = \sum_{i=1}^{n} \frac{y_i}{n}$, $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$ and $\bar{z} = \sum_{i=1}^{n} \frac{z_i}{n}$ are sample mean of y, x and z and unbiased estimator of population mean $\bar{Y} = \sum_{i=1}^{N} \frac{y_i}{N}$, $\bar{X} = \sum_{i=1}^{N} \frac{x_i}{N}$ and $\bar{Z} = \sum_{i=1}^{N} \frac{z_i}{N}$.

Population means of auxiliary variate x and z i.e. $\bar{X}$ and $\bar{Z}$ respectively are assumed to be known.

Singh (1967) suggested a ratio-cum-product estimator for population mean $\bar{Y}$ as

$$\bar{y}_{RP} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)\left(\frac{\bar{z}}{\bar{Z}}\right) \tag{1.3}$$

Mean squared errors of ratio estimator $\bar{y}_R$, product estimator $\bar{y}_P$ and ratio-cum-product estimator $\bar{y}_{RP}$ respectively are

$$MSE(\bar{y}_R) = \theta \ \bar{Y}^2\left[C_y^2 + C_x^2(1-2K_{yx})\right], \tag{1.4}$$

$$MSE(\bar{y}_P) = \theta \ \bar{Y}^2\left[C_y^2 + C_z^2(1+2K_{yz})\right], \tag{1.5}$$

$$MSE(\bar{y}_{RP}) = \theta \ \bar{Y}^2\left[C_y^2 + C_x^2(1-2K_{yx}) + C_z^2\left\{1+2(K_{yz}-K_{xz})\right\}\right], \tag{1.6}$$

where

$$\theta = \frac{1}{n} - \frac{1}{N},\ C_y = \frac{S_y}{\bar{Y}},\ C_x = \frac{S_y}{\bar{X}},\ C_z = \frac{S_z}{\bar{Z}},\ K_{yx} = \rho_{yx}\frac{C_y}{C_x},$$

$$K_{yz}=\rho_{yz}\frac{C_y}{C_z},\ K_{xz}=\rho_{xz}\frac{C_x}{C_z},\ S_y^2=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\overline{Y})^2,$$

$$S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\overline{X})^2,\ S_z^2=\frac{1}{N-1}\sum_{i=1}^{N}(z_i-\overline{Z})^2,\ \rho_{yx}=\frac{S_{yx}}{S_yS_x},$$

$$\rho_{yz}=\frac{S_{yz}}{S_yS_z},\quad \rho_{xz}=\frac{S_{xz}}{S_xS_z},\ S_{yx}=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\overline{Y})(x_i-\overline{X})$$

$$S_{yz}=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\overline{Y})(z_i-\overline{Z})\ \text{and}\ S_{yx}=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\overline{X})(z_i-\overline{Z}).$$

Srivenkataramana (1980) applied a transformation $x_i^*=\dfrac{N\overline{X}-n\overline{x}}{N-n}$ on auxiliary variate x and z suggested dual to ratio and product estimators $\overline{y}_R$ and $\overline{y}_P$ as

$$\overline{y}_R^*=\overline{y}\left(\frac{\overline{x}^*}{\overline{X}}\right) \tag{1.7}$$

and

$$\overline{y}_P^*=\overline{y}\left(\frac{\overline{Z}}{\overline{z}^*}\right) \tag{1.8}$$

where $\overline{x}^*=(1+g)\overline{X}-g\overline{x}$ and $\overline{z}^*=(1+g)\overline{Z}-g\overline{z}$ and $g=\dfrac{n}{N-n}$.

Singh et. al (2005) defined dual to ratio-cum-product estimator $\overline{y}_{RP}^*$ as

$$\overline{y}_{RP}^*=\overline{y}\left(\frac{\overline{x}^*}{\overline{X}}\right)\left(\frac{\overline{Z}}{\overline{z}^*}\right) \tag{1.9}$$

Mean squared error of $\overline{y}_R^*$, $\overline{y}_P^*$ and $\overline{y}_{RP}^*$ up to the first degree of approximation are

$$MSE(\overline{y}_R^*)=\theta\ \overline{Y}^2\left[C_y^2+gC_x^2(g-2K_{yx})\right], \tag{1.10}$$

$$MSE(\overline{y}_P^*)=\theta\ \overline{Y}^2\left[C_y^2+gC_z^2(g+2K_{yz})\right], \tag{1.11}$$

$$MSE(\overline{y}_{RP}^*)=\theta\ \overline{Y}^2\left[C_y^2+gC_x^2(g-2K_{yx})+gC_z^2\left\{g+2(K_{yz}-gK_{xz})\right\}\right]. \tag{1.12}$$

2. Proposed Estimator

Singh and Tailor (2005) defined a ratio-cum product estimator using known correlation coefficient ρ_{xz} between auxiliary variates x and z as

$$\bar{y}_{ST} = \bar{y}\left(\frac{\bar{X}+\rho_{xz}}{\bar{x}+\rho_{xz}}\right)\left(\frac{\bar{z}+\rho_{xz}}{\bar{Z}+\rho_{xz}}\right) \tag{2.1}$$

Mean squared error of Singh and Tailor (2005) estimator $\bar{y}_{ST}$ is

$$MSE(\bar{y}_{ST}) = \theta\ \bar{Y}^2\left[C_y^2 + \lambda_1 C_x^2(\lambda_1 - 2K_{yx}) + \lambda_2 C_z^2\left\{\lambda_2 + 2(K_{yz} - \lambda_1 K_{xz})\right\}\right] \tag{2.2}$$

where $\lambda_1 = \dfrac{\bar{X}}{\bar{X}+\rho_{xz}}$ and $\lambda_2 = \dfrac{\bar{Z}}{\bar{Z}+\rho_{xz}}$.

Let $x_i^* = N\bar{X} - nx_i /(N-n)$ and $z_i^* = N\bar{Z} - nz_i /(N-n)$ then $\bar{x}^* = (1-g)\bar{X} - g\bar{x}$ and $\bar{z}^* = (1+g)\bar{Z} - g\bar{z}$

where $g = \dfrac{n}{N-n}$.

Using the transformation $\bar{x}^*$ and $\bar{z}^*$, suggested dual to Singh and Tailor (2005) estimator $\bar{y}_{ST}$ is proposed as

$$\bar{y}^*_{ST} = \bar{y}\left(\frac{\bar{x}^*+\rho_{xz}}{\bar{X}+\rho_{xz}}\right)\left(\frac{\bar{Z}+\rho_{xz}}{\bar{z}^*+\rho_{xz}}\right) \tag{2.3}$$

To obtain the bias and mean squared error of $\bar{y}^*_{ST}$, we write

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

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

$E(e_0^2) = \theta C_y^2$, $\quad E(e_0e_1) = \theta\rho_{yx} C_y C_x$,

$E(e_1^2) = \theta C_x^2$, $\quad E(e_0e_2) = \theta\rho_{yz} C_y C_z$,

$E(e_2^2) = \theta C_z^2$, $\quad E(e_1e_2) = \theta\rho_{xz} C_x C_z$.

Expressing (2.1) in terms of $e_i s$ we have

$$\bar{y}^*_{ST} = \bar{Y}(1+e_0)\left\{(1-\lambda_1 e_1)(1-\lambda_2 g e_2)^{-1}\right\}$$

Finally, the bias and mean squared error of proposed estimator $\bar{y}^*_{ST}$ is

$$B(\bar{y}^*_{ST}) = g\bar{Y}\theta\left[\lambda_2 g C_z^2(\lambda_2 + K_{yz}) - \lambda_1 C_x^2(K_{yx} + \lambda_2 g K_{xz})\right] \tag{2.3}$$

$$MSE(\bar{y}^*_{ST}) = \bar{Y}^2\left[C_y^2 + g\lambda_1 C_x^2(g\lambda_1 - 2K_{yx}) + \lambda_2 g C_z^2\left\{\lambda_2 g + 2(K_{yz} - \lambda_1 g K_{xz})\right\}\right] \tag{2.4}$$

3. Efficiency Comparison

Variance of simple mean estimator $\bar{y}$ in simple random sampling without replacement is defined as

$$V(\bar{y}) = \theta S_y^2 \tag{3.1}$$

From (1.4), (1.5), (1.6), (1.10), (1.11), (1.12), (2.4) and (3.1) it follows that the mean squared error of the suggested estimator $\bar{y}_{ST}^*$ is less than

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

$$g < \frac{2(\lambda_1 C_x^2 K_{yx} - \lambda_2^2 C_z^2 K_{yz})}{\lambda_1^2 C_x^2 + \lambda_2^2 C_z^2 - 2\lambda_1\lambda_2 K_{xz} C_x^2} \tag{3.2}$$

(ii) $V(\bar{y}_R)$ if

$$g < \frac{C_x^2\{1 + K_{yx}(\lambda_1 - 2)\} - 2\lambda_2 C_z^2 K_{yz}}{\lambda_1^2 C_x^2 + \lambda_2^2 C_z^2 - 2\lambda_1\lambda_2 C_z^2 K_{xz}} \tag{3.3}$$

(iii) $V(\bar{y}_P)$ if

$$g < \frac{C_z^2\{1 + K_{yz}(1 - \lambda_2)\} + 2\lambda_1 C_x^2 K_{yx}}{\lambda_1^2 C_x^2 + \lambda_2^2 C_z^2 - 2\lambda_1\lambda_2 C_z^2 K_{xz}} \tag{3.4}$$

(iv) $V(\bar{y}_{RP})$ if

$$g < \frac{C_x^2\left[1 + 2K_{yx}(\lambda_1 - 1) - C_z^2\{1 - 2K_{xz} + 2K_{yz}(\lambda_2 + 1)\}\right]}{\lambda_1^2 C_x^2 + \lambda_2^2 C_z^2 - 2\lambda_1\lambda_2 C_z^2 K_{xz}} \tag{3.5}$$

(v) $V(\bar{y}_R^*)$ if

$$g < \frac{2C_x^2\{K_{yx}(\lambda_1 - 1)\} - 2\lambda_2 C_z^2 K_{yz}}{C_x^2(\lambda_1^2 - 1) + \lambda_2^2 C_z^2 - 2\lambda_1\lambda_2 C_z^2 K_{xz}} \tag{3.6}$$

(vi) $V(\bar{y}_P^*)$ if

$$g < \frac{C_x^2(1 + \lambda_1 K_{yx}) + 2K_{yz} C_z^2(1 - \lambda_2)}{C_x^2\lambda_1^2 + C_z^2(\lambda_2^2 - 1) - 2\lambda_1\lambda_2 C_z^2 K_{xz}} \tag{3.7}$$

(vii) $V(\bar{y}_{RP}^*)$ if

$$g < \frac{C_x^2(\lambda_1^2 - 1) + C_z^2(\lambda_2^2 - 1) + 2C_x^2 K_{xz}}{2\left[C_x^2 K_{yx}(\lambda_1 + 1) + \{\lambda_1\lambda_2 K_{xz} - K_{yz}(\lambda_2 + 1)\}\right]} \tag{3.8}$$

Expressions (3.2) to (3.8) provide the conditions under which proposed estimator $\bar{y}_{ST}^*$ would be more efficient than $\bar{y}$, $\bar{y}_R$, $\bar{y}_P$, $\bar{y}_{RP}$ $\bar{y}_R^*$, $\bar{y}_P^*$, $\bar{y}_{RP}^*$ and $\bar{y}_{ST}$.

4. Empirical Study

To see the performance of the proposed estimator we are considering two natural population data sets.

Population I [Source: Singh p.377]
y : Number of females employed
x : Number of females service
z: Number of educated females
$\overline{Y} = 7.46$, $\overline{X} = 5.31$, $\overline{Z} = 179.00$, $C_y^2 = 0.5046$, $C_x^2 = 0.5737$,
$C_z^2 = 0.0633$, $\rho_{yx} = 0.7737$, $\rho_{yz} = -0.2070$
$\rho_{xz} = -0.0033$, $N = 61$ and $n = 20$.

Population II [Source: Johnston p. 171]
y : Percentage of hives affected by disease
x : Mean January temperature
z: Date of flowering of a particular summer species (number of days from January 1)
$\overline{Y} = 52$, $\overline{X} = 42$, $\overline{Z} = 200$, $C_y^2 = 0.0244$, $C_x^2 = 0.0170$, $C_z^2 = 0.0021$,
$\rho_{yx} = 0.80$, $\rho_{yz} = -0.94$
$\rho_{xz} = -0.73$, $N = 10$ and $n = 4$.

Estimators	Percent Relative Efficiencies	
	Population I	Population II
$\overline{y}$	100.00	100.00
$\overline{y}_R$	205.34	276.85
$\overline{y}_P$	102.16	187.08
$\overline{y}_{RP}$	213.54	394.86
$\overline{y}_R^*$	214.74	238.49
$\overline{y}_P^*$	104.35	149.13
$\overline{y}_{RP}^*$	235.52	401.98
$\overline{y}_{ST}$	213.36	383.49
$\overline{y}_{ST}^*$	235.61	405.83

Table 4.1: Percent Relative Efficiencies of $\overline{y}, \overline{y}_R, \overline{y}_P, \overline{y}_{RP}$ $\overline{y}_R^*, \overline{y}_P^*, \overline{y}_{RP}^*$ $\overline{y}_{ST}$ and $\overline{y}_{ST}^*$ with respect to $\overline{y}$

Section 3 provides the conditions under which mean squared error of proposed estimator $\overline{y}_{ST}^*$ would be less than mean squared error of $\overline{y}, \overline{y}_R, \overline{y}_P, \overline{y}_R^*, \overline{y}_P^*$ and $\overline{y}_{RP}^*$. Table 4.1 reveals that suggested estimator $\overline{y}_{ST}^*$ has maximum percent relative efficiency in comparison to all other estimators considered in this paper. Thus if the correlation coefficient between auxiliary variates is known, the proposed estimator $\bar{y}^*_{ST}$ is recommended for use in practice for estimating the population mean.

Acknowledgement

Authors are thankful to the referees for their valuable suggestions regarding the improvement of the paper.

References

1. 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, p. 262-275.
2. Murthy, M.N. (1964). Product method of estimation, Sankhya, A, 26, p. 69-74.
3. Robson, D.S. (1957). Application of multivariate polykays to the theory of unbiased ratio-type estimation, J. Amer. Statist. Asso., 59, p. 1225-1226.
4. Singh, H.P. and Tailor, R. (2005). Estimation of finite population mean with known coefficient of variation of an auxiliary character, Statistica, LXV, 3, p. 301-313.
5. Singh, H.P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean, Statistics in Transition, 6, 4, p. 555-560.
6. Singh, M.P. (1964). Product method of estimation, Sankhya A, 26, p. 69-74.
7. Singh, H.P., Singh. R., Espejo, M. R. and Pineda, M.D. (2005). On the efficiency of a dual to ratio-cum-product estimator in sample surveys, Math. Proceed. Royal Irish Academy, 105 A (2), p. 51-56.
8. Sisodia, B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable, J. Ind. Soc. Agri. Statati., 33(1), p. 13-18.
9. Srivenkatramana, T. (1980). A dual to ratio estimator in sample surveys, Biometrika, 67, p. 194-204.
10. Upadhyaya, L. N. and Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean, Biometrical J., 41(5), p. 627-636.