

A Ratio-cum-Dual to Ratio Estimator of Population Variance Using Qualitative Auxiliary Information Under Simple Random Sampling

Subhash Kumar Yadav

Dept of Maths & Stats, Dr RML Avadh University,
Faizabad-224001, U.P., India
Email: drskystats@gmail.com

Himanshu Pandey

Dept of Maths & Stats, DDU University,
Gorakhpur- 273001, U.P., India

Abstract

In this paper we have proposed a class of ratio-cum-dual to ratio estimators for estimating population variance of the variable under study, using known values of some population parameters of auxiliary variable, which is available in the form of an attribute. The expressions for the bias and mean squared error of the proposed estimators have been derived upto the first order of approximation. A comparison has been made with some well known estimators of population variance available in the literature when auxiliary information is in qualitative form. It has been shown that the proposed estimator is better than the existing estimators under the optimum condition. For illustration an empirical study has been carried out.

Keywords: Class of estimators, dual to ratio estimator, bias, mean squared error, efficiency.

Mathematics Subject Classification 2010: 62D05

1 INTRODUCTION

The use of auxiliary information increases the precision of an estimator when study variable Y is highly correlated with auxiliary variable X . When the variable under study y is highly positively correlated with the auxiliary variable x , then the ratio type estimators are used to estimate the population parameter and product estimators are used when the variable under study y is highly negatively correlated with the auxiliary variable x for improved estimation of parameters of variable under study. But there are situations when information on auxiliary variable is not available in quantitative form but in practice, the information regarding the population proportion possessing certain attribute ψ is easily available (see Jhaji et.al.), which is highly correlated with the study variable Y . For example (i) Y may be the use of drugs and ψ may be the gender (ii) Y may be the production of a crop and ψ may be the particular variety. (iii) Y may be the Amount of milk produced and ψ a particular breed of

Mathematical Journal of
Interdisciplinary Sciences
Vol. 1, No. 2,
March 2013
pp. 91–96



©2013 by Chitkara
University. All Rights
Reserved.

Yadav, S. K.
 Pandey, H.

cow. (iv) Y may be the yield of wheat crop and ψ a particular variety of wheat etc. (see Shabbir and Gupta).

Let there be N units in the population. Let (y_i, ψ_i) , $i = 1, 2, \dots, N$ be the corresponding observation values of the i^{th} unit of the population of the study variable Y and the auxiliary variable ψ respectively. Further we assume that $\psi_i = 1$ and $\psi_i = 0$, $i = 1, 2, \dots, N$ if it possesses a particular characteristic or does not possess it. Let $A = \sum_{i=1}^N \psi_i$ and $a = \sum_{i=1}^n \psi_i$ denote the total number of units in the population and sample respectively possessing the attribute ψ . Let $P = \frac{A}{N}$ and $p = \frac{a}{n}$ denote the proportion of units in the population and sample respectively possessing the attribute ψ . Let a simple random sample of size n from this population is taken without replacement having sample values (y_i, ψ_i) , $i = 1, 2, \dots, n$.

Ratio type estimator for population variance of the study variable based on ISAKI (1983) estimator is as follows

$$t_R = s_y^2 \left(\frac{S_\psi^2}{s_\psi^2} \right) \quad (1.1)$$

where

$$s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2, \quad s_\psi^2 = \frac{1}{n-1} \sum_{i=1}^n (\psi_i - p)^2, \quad S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (\psi_i - P)^2$$

$$P = \frac{1}{N} \sum_{i=1}^N \psi_i, \quad \bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \quad p = \frac{1}{n} \sum_{i=1}^n \psi_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

The mean squared error (MSE) of the estimator t_R , up to the first order of approximation is given as

$$MSE(t_R) = f S_y^4 [(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)] \quad (1.2)$$

where $\lambda_{rs} = \frac{\mu_{rs}}{\mu_{r/2} \mu_{s/2}}$, $\mu_{rs} = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^r (\psi_i - P)^s$ and $f = \frac{1}{N} - \frac{1}{n}$

A dual to ratio type estimator for the population variance of the study variable y is defined as

$$t^{(d)} = s_y^2 \left(\frac{S_\psi^{*2}}{S_\psi^2} \right) \quad (1.3)$$

where $s_{\psi}^{*2} = \frac{NS_{\psi}^2 - ns_{\psi}^2}{N - n} = (1 + g)S_{\psi}^2 - gs_{\psi}^2$ with $g = \frac{n}{N - n}$

with the mean squared error (MSE) of the estimator $t^{(d)}$, up to the first order of approximation as

$$MSE(t^{(d)}) = f S_y^4 [(\lambda_{40} - 1) + g^2(\lambda_{04} - 1) - 2g(\lambda_{22} - 1)] \quad (1.4)$$

2 PROPOSED ESTIMATOR

Motivated by Sharma & Tailor, we propose the following ratio-cum-dual to ratio type estimator for the population variance of the study variable as

$$t_R^d = s_y^2 \left[\alpha \left(\frac{S_{\psi}^2}{s_{\psi}^2} \right) + (1 - \alpha) \left(\frac{s_{\psi}^{*2}}{S_{\psi}^2} \right) \right] \quad (2.1)$$

Where α is a suitably chosen constant to be determined such that MSE of the estimator t_R^d is minimum. For $\alpha = 1$, the estimator t_R^d reduces to the estimator t_R and for $\alpha = 0$ it is the estimator $t^{(d)}$.

Thus the above estimators are the particular case of the proposed estimator t_R^d .

In order to study the large sample properties of the proposed family of estimators, we define

$$s_y^2 = S_y^2(1 + \varepsilon_0) \text{ and } s_{\psi}^2 = S_{\psi}^2(1 + \varepsilon_1) \text{ such that } E(\varepsilon_i) = 0, i = 0, 1,$$

and $E(\varepsilon_0^2) = f(\lambda_{40} - 1)$, $E(\varepsilon_1^2) = f(\lambda_{04} - 1)$, $E(\varepsilon_0\varepsilon_1) = f(\lambda_{22} - 1)$

Expressing (2.1) in terms of ε_i 's, ($i = 0, 1$), we have

$$t_R^d = S_y^2(1 + \varepsilon_0) [\alpha(1 + \varepsilon_1)^{-1} + (1 - \alpha)(1 - g\varepsilon_1)] \quad (2.2)$$

Expanding terms on right hand side of (2.2), we get

$$\begin{aligned} t_R^d &= S_y^2(1 + \varepsilon_0) [\alpha(1 - \varepsilon_1 + \varepsilon_1^2) + (1 - \alpha)(1 - g\varepsilon_1)] \\ &= S_y^2 [\alpha(1 + \varepsilon_0 - \varepsilon_1 + \varepsilon_1^2 - \varepsilon_0\varepsilon_1) + (1 - \alpha)(1 + \varepsilon_0 - g\varepsilon_1 - g\varepsilon_0\varepsilon_1)] \end{aligned} \quad (2.3)$$

Taking expectation, after subtracting S_y^2 on both sides, upto the first order of approximation, we get

$$\begin{aligned} B(t_R^d) &= f S_y^2 [\alpha(\lambda_{04} - 1) - \{g + (1 - g)\alpha\}(\lambda_{22} - 1)] \\ &= f S_y^2 [\alpha(\lambda_{04} - 1) - \alpha_1(\lambda_{22} - 1)] \end{aligned} \quad (2.4)$$

where $\alpha_1 = \{g + (1 - g)\alpha\}$

From equation (2.3), we have

$$\begin{aligned} t_R^d - S_y^2 &\cong S_y^2(\varepsilon_0 - \alpha_1\varepsilon_1) \\ \Rightarrow (t_R^d - S_y^2)^2 &\cong S_y^4(\varepsilon_0 - \alpha_1\varepsilon_1)^2 \\ &= S_y^4(\varepsilon_0^2 + \alpha_1^2\varepsilon_1^2 - 2\varepsilon_0\varepsilon_1) \end{aligned} \quad (2.5)$$

Taking expectation on both sides of (2.5), we get the MSE of t_R^d , upto the first order of approximation as

$$MSE(t_R^d) = f S_y^4[(\lambda_{04} - 1) - \alpha_1^2(\lambda_{04} - 1) - 2\alpha_1(\lambda_{22} - 1)] \quad (2.6)$$

The MSE of the estimator t_R^d is minimized for the optimum value of α as

$$\alpha = \frac{K - g}{1 - g} \text{ or } \alpha_1 = K = \alpha_1^{opt} \text{ (say)} \quad (2.7)$$

where $K = \frac{(\lambda_{22} - 1)}{(\lambda_{04} - 1)}$

Substituting this value of α_1^{opt} in (2.6), we get the minimum $MSE(t_R^d)$ as

$$MSE_{\min}(t_R^d) = f S_y^4 \left[(\lambda_{40} - 1) - \frac{(\lambda_{22} - 1)^2}{(\lambda_{04} - 1)} \right] \quad (2.8)$$

Which is equal to the variance of the linear regression estimator of population variance, $\hat{S}_{lr}^2 = s_y^2 + b(S_\psi^2 - s_\psi^2)$ with the sample regression coefficient

$$b = \frac{s_y^2(\lambda_{22} - 1)}{s_\psi^2(\lambda_{04} - 1)}$$

3 EFFICIENCY COMPARISON

It is well known that the variance of estimator $t_0 = s_y^2$ is given as

$$V(t_0) = f S_y^4[\lambda_{40} - 1] \quad (3.1)$$

The estimator t_R^d is better than the estimator t_0 if,

$$MSE(t_0) > MSE_{\min}(t_R^d)$$

i.e.,
$$\frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)} > 0 \quad \text{or} \quad K > 0$$

or
$$(\lambda_{04} - 1) > 0 \tag{3.2}$$

The estimators t_R^d is more efficient than the estimator t_R , under the condition

$$MSE(t_R) > MSE_{\min}(t_R^d)$$

i.e.,
$$[(\lambda_{04}-1) - (\lambda_{22}-1)]^2 > 0$$

or
$$[1 - K]^2 > 0 \tag{3.3}$$

Now suggested estimator t_R^d is better than the estimators $t^{(d)}$ if,

$$MSE(t^{(d)}) > MSE_{\min}(t_R^d)$$

i.e.,
$$[g(\lambda_{04}-1) - (\lambda_{22}-1)]^2 > 0$$

or
$$[g - K]^2 > 0 \tag{3.4}$$

4 EMPIRICAL STUDY:

To analyze the performance of various estimators of population variance S_y^2 of study variable y , we have considered the following two data:

Data [Source: Mukhopadhyay, page 44]

Y : Household size.

ψ : Whether households have agricultural loan.

$N = 515, n = 25, \lambda_{40} = 2.3077, \lambda_{04} = 1.12, \lambda_{22} = 0.892795$

The percentage relative efficiency (PRE) of estimators $t_0 = s_y^2, t_R, t^{(d)}$ and t_R^d with respect to s_y^2 have been computed and given in the table1 below.

Table1: Efficiencies of different estimators with respect to s_y^2

Estimator	PRE ($., S_y^2$)
$t_0 = s_y^2$	100.00
t_R	79.64
$t^{(d)}$	99.15
$(t_R^d)_{opt}$	107.90

5 CONCLUSION

From the strong theoretical discussions and empirical study, it is concluded that the proposed estimator t_R^d for estimating population variance under optimum condition perform better than the usual estimator s_y^2 , estimator based on ISAKI (1983) ratio type estimator t_R and the dual to ratio type estimator $t^{(d)}$ of population variance using qualitative auxiliary information in the present example.

REFERENCES

- Cochran, W.G., 1977. Sampling Techniques. New-York, John Wiley and Sons.
- Isaki, C.T. (1983): Variance estimation using auxiliary information. Jour. Amer. Stat. Assoc., 78, 117-123. <http://dx.doi.org/10.1080/01621459.1983.10477939>
- Jhajj, H. S. Sharma, M. K. and Grover, L. K. (2006) A family of estimators of population mean using information on auxiliary attribute. *Pak. J. Statist.*, 22 (1), 43-50.
- Kadilar, C. and Cingi, H. (2006): Ratio estimators for the population variance in simple and stratified random sampling. *Applied Mathematics and Computation*, 173, 1047-1059. <http://dx.doi.org/10.1016/j.amc.2005.04.032>
- Misra, S. and Yadav, S.K. (2010): Ratio type estimator of population variance using qualitative auxiliary information. *International transactions in mathematical sciences and computers*, Vol. 3, 2, 313-322.
- Mukhopadhyay, P., 1998. Theory and methods of survey sampling, Prentice Hall of India private limited, New Delhi, India.
- Singh, H.P., Upadhyaya, L.N. and Namjoshi, U.D. (1988). Estimation of finite population variance. *Curr. Sc.* 57:1331-1334.
- Singh, S. (2003). *Advanced sampling theory with applications*. Kluwer Academic. Press. <http://dx.doi.org/10.1007/978-94-007-0789-4>
- Sukhatme, P.V., and Sukhatme, B.V. 1970. *Sampling theory of surveys with applications*. Asia Publishing House, New Delhi.
- Shabbir, J. and Gupta, S. On estimating the finite population mean with known population proportion of an auxiliary variable, *Pak. J. Statist.*, 23(1), 1-9, 2007.
- Sharma, B. and Tailor, R. A New Ratio-Cum-Dual to Ratio Estimator of Finite Population Mean in Simple Random Sampling, *Global Journal of Science Frontier Research*, Vol. 10, 1, 27-31, 2010.
- Upadhyaya, L.N. and Singh, H.P. 2001. Estimation of the population standard deviation using information. *Am. Jour. Math. Manag. Sci.* 21:345-358 <http://dx.doi.org/10.1080/01966324.2001.10737565>

Subhash Kumar Yadav is a Asstt. Professor, In the Deptt. of Mathematics & Statistics, Dr R M L Avadh University, Faizabad (UP) INDIA. He has been working in field of Sampling, Non-Linear Regression Analysis. He has Publish several paper in this fields.

Himanshu Pandey is a Associate. Professor, In the Deptt. of Mathematics & Statistics, D.D.U.Gorakhpur University, Gorakhpur (UP) INDIA. He has been working in field of Sampling, Demography & O.R.. He has Publish several research paper in these fields.