American Journal of Applied Mathematics
2014; 2(3): 92-95
Published online June 30, 2014 (http://www.sciencepublishinggroup.com/j/ajam)
doi: 10.11648/j.ajam.20140203.14

A difference type estimator for estimating population variance with possible applications to random stock and dividend growth

Rajesh Singh[1], Viplav Kumar Singh[1], Mohd Khoshnevisan[2]

[1]Department of Statistics, BHU, Varanasi, India
[2]Associate Professor of Finance, Ajman University of Science and Technology, UAE

Email address:
rsinghstat@gmail.com (R. Singh), viplavkumarsingh0802@gmail.com (V. K. Singh), m.khoshnevisan@ajman.ac.ae (Mhd Khoshnevisan)

To cite this article:
Rajesh Singh, Viplav Kumar Singh, Mohd Khoshnevisan. A Difference Type Estimator for Estimating Population Variance with Possible Applications to Random Stock and Dividend Growth. *American Journal of Applied Mathematics.* Vol. 2, No. 3, 2014, pp. 92-95.
doi: 10.11648/j.ajam.20140203.14

Abstract: In this paper we have suggested a difference type estimator for estimating the unknown population variance of the study variable y using auxiliary information. The optimum estimator in the suggested method has been identified along with its mean square error formula and it is seen that the suggested estimator performs better than other existing estimators. An empirical study is carried out to judge the merits of proposed estimator over other traditional estimators.

Keywords: Study Variable, Auxiliary Variable, Mean Square Error, Bias, Simple Random Sampling

1. Introduction

It ha been well recognized that the use of auxiliary information in sample survey design results in efficient estimators of population parameters. Out of many ratio, regression and product methods of estimation are good examples in this context. Estimating the finite population variance has great significance in various fields such as industry, agriculture, and medical and biological sciences, where we come across populations which are likely to be skewed. Variations are present everywhere in our day to day life. A manufacturer needs constant knowledge of the level of variations in people's reaction to his product to be able to know whether to reduce or increase his price, or improve the quality of the product.

Consider a finite population $U=(U_1,U_2,U_2.........U_N)$ having N units and let y and x are the study and auxiliary variable with means $\overline{X}$ and $\overline{Y}$ respectively. Let a sample of size n is drawn from the population using simple random sampling without replacement (SRSWOR) method. To measure the variations within the values of study variable y, the problem of estimating the population variance of S_y^2 of study variable y received a considerable attention of the statistician in survey sampling including Isaki (1983), Singh and Singh (2001, 2003), Jhajj et al. (2005), Kadliar and Cingi (2006), Singh et al. (2008), Grover (2010), Singh et al. (2011), Singh and Solanki (2012), Singh and Malik (2014) have suggested improved estimator for estimation of S_y^2.

Let $s_y^2=\sum_{i=1}^{n}(y_i-\bar{y})^2/(n-1)$ and $s_x^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1)$ be the sample variances for variables y and x, which are unbiased estimators for $S_Y^2(=\sum_{i=1}^{N}(y_i-\overline{Y})^2/(N-1))$ and $S_x^2(=\sum_{i=1}^{N}(x_i-\overline{X})^2/(N-1))$, respectively. Let $C_y=\frac{S_y}{\overline{Y}}$ and $C_x=\frac{S_x}{\overline{X}}$.

Also let $e_0=\left\{\frac{s_y^2}{S_y^2}-1\right\}$ and $e_1=\left\{\frac{s_x^2}{S_x^2}-1\right\}$, be such that $E(e_0)=E(e_0)=0$, also to the first order of approximation, we have $E(e_0^2)=\theta\beta_{2y}^*$, $E(e_1^2)=\theta\beta_{2x}^*$ and $E(e_0e_1)=\theta\lambda_{22}^*$. Where, $\theta=\frac{1}{n}$, $\lambda_{pq}=\frac{\mu_{pq}}{\mu_{20}^{p/2}\mu_{02}^{q/2}}$, $\mu_{pq}=\frac{1}{N}\sum_{i=1}^{N}(y_i-\overline{Y})^P(x_i-\overline{X})^q$ and $\beta_{2y}=\frac{\lambda_{40}}{\mu_{20}^2},\beta_{2x}=\frac{\lambda_{04}}{\mu_{02}^2}$ are the coefficient of kurtosis of y and x respectively.

In this paper, under SRSWOR, we have suggested a general family of estimators for estimating the population variance S_y^2. The expression of MSE, up to the first order of approximation have been obtained.

2. Existing Estimators

The variance of the usual unbiased estimator $\hat{S}_y^2$ is given by

$$\mathrm{var}(\hat{S}_y^2) = \theta\ S_y^4\beta_{2y}^* \qquad (1.1)$$

Isaki (1983), suggested the following ratio estimator for estimating population variance S_y^2

$$t_0 = s_y^2\left(\frac{S_x^2}{s_x^2}\right) \qquad (1.2)$$

The MSE of the ratio estimator t_0 is given by

$$MSE(t_0) = \theta S_y^4\left(\beta_{2y}^* + \beta_{2x}^* - 2\lambda_{22}^*\right) \qquad (1.3)$$

Garcia and Cebrian (1996) suggested following estimator for estimating S_y^2

$$t_1 = s_y^2\left[\frac{S_x^2}{s_x^2}\right]^\alpha \qquad (1.4)$$

MSE of the estimator t_1 is given by,

$$MSE(t_1) = \theta S_y^4\left(\beta_{2y}^* + \alpha^2\beta_{2x}^* - 2\alpha\lambda_{22}^*\right) \qquad (1.5)$$

where

$$\alpha_{OPT} = \frac{\lambda_{22}^*}{\beta_{2x}^*}.$$

Motivated by Khoshnevisan et al. (2007), Singh et al.(2007) proposed the following ratio-type estimator for the population variance

$$t_2 = \frac{aS_x^2 - b}{\left[\alpha(as_x^2 - b) + (1-\alpha)(aS_x^2 - b)\right]} \qquad (1.6)$$

where ($a \neq 0$), b are either real numbers or the function of the known parameters of the auxiliary variable x such as coefficient of variation C(x) and coefficient of kurtosis β_{2x}^* . The MSE of the estimator t_2 is given by

$$MSE(t_2) = \theta S_Y^4\left\{\beta_{2y}^* - 2\alpha\theta(h-1) + \alpha^2\theta^2\beta_{2x}^*\right\} \qquad (1.7)$$

where, $h = \dfrac{\mu_{22}}{\mu_{20}\mu_{02}}$.

Partially differentiating MSE of the estimator t_2 with respect to α, we get

$$\alpha_{OPT} = \frac{C}{\theta_1}$$

Where, $\theta_1 = \dfrac{aS_x^2}{aS_x^2 - b}$ and $C = \dfrac{h-1}{\beta_{2x}^*}$.

After substituting the optimum value of α in (1.7) minMSE of t_2 is obtain.

*Isaki (1983) suggested the regression estimator for estimating population variance S_y^2 as

$$t_3 = s_y^2 + b^*(S_x^2 - s_x^2) \qquad (1.8)$$

where b^* is the sample regression coefficient between s_x^2 and s_y^2.

The MSE of the regression estimator is given by

$$MSE(t_3) = \theta S_y^4\beta_{2y}^*\left(1 - \frac{\lambda_{22}^{*2}}{\beta_{2y}^*\beta_{2x}^*}\right) \qquad (1.9)$$

3. Suggested Estimator

Motivated by Dubey and Singh (2011), we propose an improved estimator for estimating population variance S_y^2 of study variable y, as

$$t = [\alpha_1 s_y^2 + \alpha_2 s_x^{*2} + (1-\alpha_1-\alpha_2)S_x^{*2}]\left(\frac{S_x^{*2}}{s_x^{*2}}\right)^\alpha \qquad (2.1)$$

where (α_1, α_2) are suitably chosen scalars such that MSE of proposed class of estimators t is minimum, $s_x^{*2}(=\eta s_x^2 + \lambda)$, $S_x^{*2}(=\eta S_x^2 + \lambda)$ with (η, λ) are either constants or function of some known population parameters such as mean $(\bar{X})$, mean square (S_x^2) , coefficient of variation (C_x) and coefficient of kurtosis $(\beta_2(x))$ of the auxiliary variable *x*.

To obtain the bias and MSE of proposed class of estimators *t*, we define e_0 and e_1 as,

$s_y^2 = S_y^2(1+e_0)$ $s_x^2 = S_y^2(1+e_1)$

Expressing the estimator t in terms of *e*'s, we have

$$t = \left[\alpha_1 S_y^2(1+e_0) + \alpha_2 S_x^{*2}\tau e_1 + (1-\alpha_1)S_x^{*2}\right]\{1+\tau e_1\}^{-\alpha} \qquad (2.2)$$

Where $\tau = \eta S_x^2 (S_x^{*2})^{-1}$.

We assume $|\tau e_1| < 1$, so that the term $(1+\tau e_1)^{-\alpha}$ is expandable. Thus by expanding the right hand side of (2.2) and neglecting the terms of *e*'s having power greater than two, we have

$$t = \left[\alpha_1 S_Y^2 + \alpha_1 S_Y^2 e_0 + \alpha_2 S_x^{*2}\tau e_1 + (1-\alpha_1)S_x^{*2} - \alpha\alpha_1\tau S_Y^2 e_1 - \alpha\alpha_1\tau S_Y^2 e_1 e_0\right.$$

$$-\alpha\alpha_2\tau^2 S_x^{*2} e_1^2 - \alpha(1-\alpha_1)\tau S_x^{*2} e_1 + \frac{\alpha_1\alpha(1+\alpha)}{2}\tau^2 S_x^{*2} e_1^2 + \frac{\alpha(1+\alpha)}{2}(1-\alpha_1)\tau^2 S_x^{*2} e_1^2 \Bigg] \quad (2.3)$$

Expanding the right hand side of expression (2.3), to the first order of approximation and subtracting S_Y^2 from both sides, we have

$$t - S_Y^2 = S_Y^2\left[\alpha_1\left\{\gamma + e_0 - \alpha\tau e_0 e_1 - \alpha\tau\gamma e_1 + \frac{\alpha(\alpha+1)\tau^2 e_1^2\gamma}{2}\right\}\right.$$

$$\left. + \alpha_2\frac{\tau}{R^*}\left\{e_1 - \alpha\tau e_1^2\right\} - \left\{\gamma + \frac{\alpha\tau e_1}{R^*} - \frac{\alpha(\alpha+1)\tau^2 e_1^2}{2R^*}\right\}\right] \quad (2.4)$$

where, $\gamma = (1-(R^*)^{-1})$ and $R^* = (S_y^2/S_x^{*2})$

Taking expectation of both sides of (2.4), we get the bias of the estimator t to the first degree of approximation as

$$B(t) = S_Y^2\left[\alpha_1\left\{\gamma - \alpha\tau\theta\lambda_{22}^* + \frac{\alpha(\alpha+1)\tau^2\theta\gamma}{2}\beta_{2x}^*\right\} - \frac{\alpha\alpha_2\tau^2}{R^*}\theta\beta_{2x}^*\right.$$

$$\left. - \left\{\gamma - \alpha\tau\theta\lambda_{22}^* + \frac{\alpha(\alpha+1)\tau^2\theta}{2R^*}\beta_{2x}^*\right\}\right] \quad (2.5)$$

Squaring both sides of equation (2.4) and retaining terms up to the first order of approximation and then taking expectations, we get the MSE of proposed class of estimator *t*, to the first degree of approximation as

$$MSE(T) = S_y^4[A + \alpha_1^2 B + \alpha_2^2 C + 2\alpha_1\alpha_2 D - 2\alpha_1 E - 2\alpha_2 F] \quad (2.6)$$

Where,

$$A = \left[\gamma^2 + \theta\beta_{2x}^*\left\{\alpha^2\tau^2(1-\gamma)^2 - \alpha(\alpha+1)\tau^2\gamma(1-\gamma)\right\}\right]$$

$$B = \left[\gamma^2 + \theta\beta_{2y}^* + \theta\beta_{2x}^*\left\{\alpha^2\tau^2\gamma^2 + \alpha(\alpha+1)\tau^2\gamma^2\right\} - 4\alpha\tau\gamma\theta\lambda_{22}^*\right]$$

$$C = \tau^2(1-\gamma)^2\theta\beta_{2x}^*$$

$$D = \tau(1-\gamma)\left[\theta\lambda_{22}^* - 2\alpha\tau\gamma\theta\beta_{2x}^*\right]$$

$$E = \left[\gamma^2 + \theta\beta_{2x}^*\left\{\frac{\alpha(\alpha+1)}{2}\tau^2\gamma^2 - \frac{\alpha(\alpha+1)}{2}\tau^2\gamma(1-\gamma) + \alpha^2\tau^2\gamma(1-\gamma)\right\}\right.$$

$$\left. + \theta\lambda_{22}^*\left[\alpha\tau(1-\gamma) - \alpha\tau\gamma\right]\right]$$

$$F = \tau(1-\gamma)\theta\beta_{2x}^*\left[\alpha\tau(1-\gamma) - \alpha\tau\gamma\right]$$

Minimizing expression (2.6) with respect to α_1 and α_2, we get the optimum values of α_1 and α_2 as

$$\alpha_1 = \frac{CE - DF}{BC - D^2} \quad (2.7)$$

$$\alpha_2 = \frac{BF - DE}{BC - D^2} \quad (2.8)$$

Using these optimum values of α_1 and α_2 in expression (2.5) and (2.6), we get the minimum Bias and minimum MSE of the proposed class of estimator t.

In addition some new members of suggested difference-type class of estimator *t* [for different values of (α, η, λ)] have been summarized in Table 1.

Table 1. *Some new members of suggested class of estimators t*

Estimators	Values of constants		
	α	η	λ
$T_1 = [\alpha_1 s_y^2 + \alpha_2(\beta_{2x}^* - c_x s_x^2) + (1-\alpha_1-\alpha_2)(\beta_{2x}^* - c_x S_x^2)]\left[\frac{(\beta_{2x}^* - c_x S_x^2)}{(\beta_{2x}^* - c_x s_x^2)}\right]$	1	$-c_x$	β_{2x}^*
$T_2 = [\alpha_1 s_y^2 + \alpha_2(c_x^2 s_x^2 - 1) + (1-\alpha_1-\alpha_2)(c_x^2 S_x^2 - 1)]\left[\frac{(c_x^2 S_x^2 - 1)}{(c_x^2 s_x^2 - 1)}\right]$	1	c_x^2	-1
$T_3 = [\alpha_1 s_y^2 - \alpha_2(\beta_{2x}^* + c_x s_x^2) - (1-\alpha_1-\alpha_2)(\beta_{2x}^* + c_x S_x^2)]\left[\frac{(\beta_{2x}^* + c_x S_x^2)}{(\beta_{2x}^* + c_x s_x^2)}\right]$	1	$-c_x$	$-\beta_{2x}^*$
$T_4 = [\alpha_1 s_y^2 + \alpha_2(\beta_{2x}^* s_x^2 + \rho^2) + (1-\alpha_1-\alpha_2)(\beta_{2x}^* S_x^2 + \rho^2)]\left[\frac{(\beta_{2x}^* S_x^2 + \rho^2)}{(\beta_{2x}^* s_x^2 + \rho^2)}\right]$	1	β_{2x}^*	ρ^2
$T_5 = [\alpha_1 s_y^2 - \alpha_2 c_x s_x^2 - (1-\alpha_1-\alpha_2)c_x S_x^2]\left[\frac{S_x^2}{s_x^2}\right]$	1	$-c_x$	0
$T_6 = [\alpha_1 s_y^2 - \alpha_2(\beta_{2x}^* + c_x^2 s_x^2) - (1-\alpha_1-\alpha_2)(\beta_{2x}^* + c_x^2 S_x^2)]\left[\frac{(\beta_{2x}^* + c_x^2 S_x^2)}{(\beta_{2x}^* + c_x^2 s_x^2)}\right]$	1	$-C_x^2$	$-\beta_{2x}^*$
$T_7 = [\alpha_1 s_y^2 + \alpha_2(\beta_{2x}^* s_x^2 - \rho) + (1-\alpha_1-\alpha_2)(\beta_{2x}^* S_x^2 - \rho)]\left[\frac{(\beta_{2x}^* S_x^2 - \rho)}{(\beta_{2x}^* s_x^2 - \rho)}\right]$	1	β_{2x}^*	$-\rho$
$T_8 = [\alpha_1 s_y^2 + \alpha_2 c_x^2 s_x^2 + (1-\alpha_1-\alpha_2)c_x^2 S_x^2]\left[\frac{S_x^2}{s_x^2}\right]$	1	c_x^2	0

4. Empirical Study

To illustrate the performance of estimators t_i, (i = 1, 2, …, 8) which are members of the suggested class of estimators t, over other existing estimators, we have considered the data set of Singh and Solanki (2012).

In our first case, population data set the level of apple production amount(in 100 tones) is a study variable y and numbers of apple trees is an auxiliary variable x in 104 villages of the east Anatoliya Region in Turkey in 1999. We have also tested some random stock prices versus dividend growth and it seems that the MSE for S &P 500 follow the same pattern. i.e MSE was minimized for $t_1,\ldots,t_{354}$. Obviously, we shall employ more trials to confirm our findings with respect to the second case.

The required values of parameters are summarized below-

$$N = 104, \quad c_y = 1.866, \quad \overline{Y} = 6.254, \quad \beta_{2y} = 16.523, \quad \lambda_{22} = 14.398$$
$$n = 20\ , \quad c_x = 1.653, \quad \overline{X} = 13931.683, \quad \beta_{2x} = 17.516, \quad \rho = 0.837$$

Table 2. *Minimum MSE's of members of the class of estimator's t and other existing estimator's*

Estimators	Mean Square Error
s_y^2	14395.5773
t_0	4862.320715
t_1	4612.320715
t_2	4316.320715
t_3	4316.320715
T_1	3484.4405
T_2	3370.915447
T_3	3164.239127
T_4	3313.540299
T_5	2525.707811
T_6	2273.18508
T_7	1825.69442
T_8	356.102917

5. Conclusion

We observe from Table 2 that the proposed estimator 'T' under optimum condition performs better than usual estimator S_y^2, Isaki's (1983) estimator, Garcia and Cebrian (1996) estimators and other members included in this paper.

References

[1] AA.Ceberian and R.M Garcia(1997) .Variance estimation using auxiliary information ; an almost multivariate ratio estimator,metrika,vol45,pp.171-178.

[2] Grover, L. K . (2010): A correction note on improvement in variance estimation using auxiliary information. Commun. Stat. Theo. Meth. 39:753–764.

[3] Isaki, C.T.(1983): Variance estimation using auxiliary information. Journal of American Statistical Association 78, 117–123.

[4] Jhajj, H .S., Sharma, M. K. and Grover, L. K. (2005) : An efficient class of chain estimators of population variance under sub-sampling scheme. J. Japan Stat. Soc., 35(2), 273-286.

[5] Kadilar, C. and Cingi, H. (2006): Improvement in variance estimation using auxiliary information. Hacettepe Journal of Mathematics and Statistics 35 (1), 111–115.

[6] Singh, R.,Cauhan, P., Sawan, N. and Smarandache,F. (2007): Auxiliary information and a priori values in construction of improved estimators. Renaissance High press.

[7] Singh, H.P. and Singh, R. (2001): Improved ratio-type estimator for variance using auxiliary information. J Indian Soc Agric Stat 54(3):276–287.

[8] Singh,H.P. and Singh, R. (2003): Estimation of variance through regression approach in two phase sampling. Aligarh Journal of Statistics, 23, 13-30.

[9] Singh, H. P. and Solanki, R.S. (2012): A new procedure for variance estimation in simple random sampling using auxiliary information. Stat. Paps. DOI 10.1007/s00362-012-0445-2.

[10] Singh, R., Chauhan, P., Sawan, N. and Smarandache,F.(2008): Almost unbiased ratio and product type estimator of finite population variance using the knowledge of kurtosis of an auxiliary variable in sample surveys. Octogon Mathematical Journal, Vol. 16, No. 1, 123-130.

[11] Singh, R., Chauhan, P., Sawan, N. and Smarandache, F.(2011): Improved exponential estimator for population variance using two auxiliary variables. Italian Jour. Of Pure and Appld. Math., 28, 103-110.

[12] Singh, R. and Malik, S. (2014) : Improved estimator of population variance using information on auxiliary attribute in simple random sampling. Appl. Math. Comp., 235, 43-49.