DOI: 10.15415/mjis.2017.52011

Variance Estimation Using Quality Characteristic

[1]PRABHAKAR MISHRA AND *[2]RAJESH SINGH

[1,2]Department of Statistics, Banaras Hindu University, Varanasi, 221005

***Email: rsinghstat@gmail.com**

Received: December 02, 2016| Revised: December 19, 2016| Accepted: January 11, 2017

Published online: March 05, 2017

Abstract This paper deals with estimation of unknown population variance of study variable (y) using auxiliary qualitative characteristic. The properties of the proposed class of estimators is studied and is supported through two real data sets.

Keywords: Mean square error; proportion; population variance; percent relative efficiency; simple random sampling without replacement.

1. INTRODUCTION

It is common in sampling to use auxiliary information to construct improved estimators. Ratio, product and difference estimators are common examples. One way to increase efficiency is the use of auxiliary quality characteristic. The common instance is-

(a) Gender and length of the persons,

(b) Yield of milk and variety of cow (see [3]).

Using point bi-serial correlation between the variables, many authors have suggested improved estimators for estimating unknown population mean (see [8] and [10]). Many a times in place of estimating population mean we are interested in estimation of unknown population variance (see [9]). Using known values of population parameter(s) Das and Tripathi [1], Isaki [2], Prasad and Singh [7], Kadilar and Cingi [4] and Singh et al. [10] have suggested estimators for estimating S_y^2.

Consider a finite population of size N with units U= $\{U_1, U_2, \ldots, U_N\}$.

Let y_i and ψ_i, be the values for y and ψ respectively (i=1,2,….,N). Let, $\psi_i = 1$, if ith unit of population has characteristic ψ and $\psi_i = 0$, otherwise. Let M = $\sum_{i=1}^{N}\psi_i$, m = $\sum_{i=1}^{n}\psi_i$, $P = \frac{M}{N}$ and $p = \frac{m}{n}$.

Mathematical Journal of Interdisciplinary Sciences
Vol-5, No-2,
March 2017
pp. 143–149

We use following values to obtain value of mean squared error (MSE) $g_0 = \frac{s_y^2}{S_y^2} - 1$ and $g_1 = \frac{p}{P} - 1$, such that $E(g_i) = 0$; i = (0,1) and $E(g_0^2) = \xi(\lambda_{40} - 1)$, $E(g_i) = 0$, $E(g_0 g_1) = \xi C_p \lambda_{21}$.

Where $\xi = \frac{1}{n} - \frac{1}{N}$, $C_p^2 = \frac{S_y^2}{P^2}$, $\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2} \mu_{02}^{s/2}}$, $\mu_{rs} = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^r (\psi_i - P)^s$.

Consider estimator $H_0 (= s_y^2)$ for estimation of S_y^2 -

$$H_0 = s_y^2 \tag{1}$$

The variance of H_o is given by

$$Var(H_0) = \xi S_y^4 (\lambda_{40} - 1) \tag{2}$$

2. PROPOSED ESTIMATOR

A ratio estimator H_1 for estimating S_y^2 is given by

$$H_1 = s_y^2 \frac{P}{p} \tag{3}$$

Expression of mean squared error for H_1 is given by

$$MSE(H_1) = \xi S_y^4 \left\{ (\lambda_{40} - 1) + C_p^2 - 2 C_p \lambda_{21} \right\} \tag{4}$$

Product Estimator: The conventional product estimator for estimating S_y^2 is given by

$$H_2 = s_y^2 \frac{p}{P} \tag{5}$$

Expression of mean squared error for H_2 is given by

$$MSE(H_2) = \xi S_y^4 \left\{ (\lambda_{40} - 1) + C_p^2 + 2 C_p \lambda_{21} \right\} \tag{6}$$

Difference-Type Estimator: The usual unbiased difference-type estimator is given by

$$H_3 = s_y^2 + k\left(P - p\right). \quad (7)$$

where k is an unknown constant.

The variance of the estimator H_3 is given by

$$Var.\left(H_3\right) = \xi S_y^4\left(\lambda_{40} - 1\right) + k^2P^2\xi C_p^2 - 2kPS_y^2\xi C_p\lambda_{21}. \quad (8)$$

At k ($= \frac{S_y^2\lambda_{21}}{PC_p}$), the minimum value of the equation (8) is given by

$$Var._{\min}\left(H_3\right) = \xi S_y^4\left(\lambda_{40} - 1 - \lambda_{21}^2\right). \quad (9)$$

3. ADAPTED ESTIMATOR

Motivated by Sahai and Ray [12], Koyuncu [5] and using the conventional product estimator of s_y^2, we propose the variance estimator using auxiliary attribute p, as-

$$H_4 = \left\{k_1 s_y^2\left(2 - \frac{p}{P}\right)^{\gamma}\left(\frac{p}{P}\right)^{1-\gamma} + k_2\left(\frac{p}{P}\right)^{\gamma}\right\}\exp\left\{\frac{\eta\left(P - p\right)}{\eta\left(P + p\right) + 2\lambda}\right\}. \quad (10)$$

k_1 and k_2 are constants. Here γ, η and λ are known population parameters of auxiliary attribute, like coefficient of variation C_p, population proportion P etc. (see [14]).

Expanding Eq. (10), we have

$$H_4 = k_1 S_y^2\left\{\begin{matrix} 1 + g_0 + \left(1 - \alpha - 2\gamma\right)g_1 + \left(2\gamma^2 + \frac{3}{2}\alpha^2 - \alpha - 2\gamma + 2\alpha\gamma\right) \\ \times g_1^2 + \left(1 - \alpha - 2\gamma\right)g_0 g_1 \end{matrix}\right\}$$

$$+ k_2\left\{1 + \left(\gamma - \alpha\right)g_1 + \left(\frac{3}{2}\alpha^2 - \gamma\alpha + \frac{\gamma^2}{2} - \frac{\gamma}{2}\right)g_1^2\right\} \quad (11)$$

here $\alpha = \frac{\eta P}{2\left(\eta P + \lambda\right)}$.

Solving (11), we get

$$MSE\left(H_4\right) = S_y^4 + k_1^2 A_1 + k_2^2 A_2 - 2k_1 A_3 - 2k_2 A_4 + 2k_1 k_2 A_5 \quad (12)$$

where,

$$A_1 = S_y^4\left[1+\xi\left\{\begin{matrix}(\lambda_{40}-1)+C_p^2\left(1+4\alpha^2+8\gamma^2-4\alpha-8\gamma+8\alpha\gamma\right)\\ +4C_p\lambda_{21}\left(1-\alpha-2\gamma\right)\end{matrix}\right\}\right],$$

$$A_2 = \left[1+\xi C_p^2\left(4\alpha^2+2\gamma^2-\gamma-4\alpha\gamma\right)\right],$$

$$A_3 = S_y^4\left[1+\xi\left\{C_p^2\left(\frac{3}{2}\alpha^2+2\gamma^2-\alpha-2\gamma+2\alpha\gamma\right)+C_p\lambda_{21}\left(1-\alpha-2\gamma\right)\right\}\right],$$

$$A_4 = S_y^2\left[1+\xi C_p^2\left(\frac{3}{2}\alpha^2+\frac{1}{2}\gamma^2-\frac{1}{2}\gamma-\alpha\gamma\right)\right],$$

$$A_5 = S_y^2\left[1+\xi\left\{C_p^2\left(4\alpha^2+\frac{1}{2}\gamma^2-2\alpha-\frac{3}{2}\gamma+2\alpha\gamma\right)+C_p\lambda_{21}\left(1-2\alpha-\gamma\right)\right\}\right]$$

On partially differentiating Eq. (12) with respect to the constants, we get

$$k_1^{opt.} = \frac{A_2A_3-A_4A_5}{A_1A_2-A_5^2} \text{ and } k_2^{opt.} = \frac{A_1A_4-A_3A_5}{A_1A_2-A_5^2}$$

Using optimum values of k_1 and k_2 in Eq. (12), we obtain

$$MSE_{\min.}\left(H_4\right) = S_y^4 + \frac{2A_3A_4A_5-A_2A_3^2-A_1A_4^2}{A_1A_2-A_5^2} \tag{13}$$

4. EFFICIENCY COMPARISONS

Comparison of estimators considered in this article-

From (2) and (9), $Var\left(H_0\right)-Var_{\min.}\left(H_3\right)\geq 0$, if $\xi S_y^4\lambda_{21}^2 \geq 0$.

From (2) and (13), $Var\left(H_0\right)-MSE_{\min.}\left(H_4\right)\geq 0$, if

$$S_y^4\left\{\xi\left(\lambda_{40}-1\right)-1\right\}-\frac{2A_3A_4A_5-A_2A_3^2-A_1A_4^2}{A_1A_2-A_5^2}\geq 0 .$$

From (9) and (13), $Var_{\min.}\left(H_3\right)-MSE_{\min.}\left(H_4\right)\geq 0$, if

$$S_y^4\left\{\xi\left(\lambda_{40}-1-\lambda_{21}^2\right)-1\right\}-\frac{2A_3A_4A_5-A_2A_3^2-A_1A_4^2}{A_1A_2-A_5^2}\geq 0 .$$

5. EMPIRICAL STUDY

In this section theoretical results are verified with help of following two data sets:

Population Data Sets:

Population Data 1	Population Data 2
[Source: [6]]	[Source: [11]]
Y: Household size and	Y: Number of villages in the circles and
Ψ: Household took an agricultural loan from a bank.	Ψ: A circle consisting more than five villages.
N=25, n=8, P=0.4, C_p^2=1.5625,	N=89, n=23, P=0.1235, C_p^2=7.7172,
λ_{40}=2.3077, λ_{21}=-0.16801, S_y^2=17.4233.	λ_{40}=3.8542, λ_{21}=1.3228, S_y^2=4.0738.

6. CONCLUSION

We have proposed some new estimators for estimating population variance using auxiliary quality characteristics. Tables 1, 2 and 3 reveal that the estimators proposed performs better than other estimators considered in the present study.

Table 1: Percent relative efficiency

Estimators	PRE (population 1)	PRE (population 2)
H_0	100.00	100.00
H_1	39.7453	88.5871
H_2	53.3712	15.9268
H_3	102.2059	258.4389

Table 2: Percent relative efficiency of the proposed estimator H_4 for Population 1.

η	Λ	PRE ($\gamma = 0$)	PRE ($\gamma = 1$)
1	1	7896.457	767.5542
1	0	727.0434	**2388.9827**
C_p	1	5819.49	811.4687
1	C_p	10944.1369	730.8589

P	1	33706.1568	653.1480
1	P	2629.017	1001.8551
P	C_p	**49789.2678**	636.9053
C_p	P	2145.2016	1081.5429

Table 3: Percent relative efficiency of the proposed estimator H_4 for Population 2.

η	Λ	**PRE (γ = 0)**	**PRE (γ = 1)**
1	1	369192.9438	1005.0804
C_p	1	87858.40	1172.6728
1	C_p	2247258.882	932.6347
P	1	17438366.19	904.0739
1	P	49156.0955	**1482.9417**
P	C_p	**130508750.5**	894.3079

CONFLICT OF INTEREST

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

REFERENCE

[1] Das, A. K. and Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance. Sankhya, **40**(C), 139–148.

[2] Isaki, C. T. (1983). Variance estimation using auxiliary information. Journal of the American Statistical Association, **78**(381), 117–123.

[3] Jhajj, H. S., Sharma, M. K. and Grover, L. K., (2006). A family of estimators of population mean using information on auxiliary attribute. Pakistan Journal of Statistics, **22**(1), 43–50.

[4] Kadilar, C. and Cingi, H. (2007). Improvement in variance estimation in simple random sampling. Communications in Statistics—Theory and Methods, **36**(11), 2075–2081.

[5] Koyuncu, N. (2012). Efficient estimators of population mean using auxiliary attributes. Applied Mathematics and Computation, Vol. **218**, no.22, pp. 10900–10905.

[6] Mukhopadhyaya, P. (2009). Theory and methods of survey sampling. Prentice Hall of India, New Delhi, India.

[7] Prasad, B. and Singh, H. P. (1990). Some improved ratio-type estimators of finite population variance in sample surveys. Communications in Statistics-Theory and Methods, **19**(3), 1127–1139.

[8] Singh, R. and Malik, S. (2014). Improved estimation of population variance using information on auxiliary attributes in simple random sampling. Applied Mathematics and Computation, **235**, 43–49.

[9] Singh, H. P., Upadhyaya, L. N. and Namjoshi, U. D. (1988). Estimation of finite population variance. Current Science, **57**(24), 1331–1334.

[10] Singh, R., Chauhan, P., Smarandache, F., & Sawan, N. (2007). Auxiliary information and a priori values in construction of improved estimators. Infinite Study.

[11] Sukhatme, P. V. and Sukhatme, B. V. (1970). Sampling theory of surveys with applications. Iowa State University Press, Ames, U.S.A.

[12] Sahai, A. and Ray, S. K. (1980). An efficient estimator using auxiliary information. Metrika, Vol. **27**, no.1, pp. 271–275.

[13] Singh, S. (2003). Advanced Sampling Theory with Application: How Michael "Selected'' Amy Vol. **1 & 2**, pp. 1–1247, Kluwer Academic Publisher, The Netherlands.

[14] Singh, R. and Kumar, M. (2011): A note on transformations on auxiliary variable in survey sampling. Mod. Assis. Stat. Appl., **6**:1, 17–19.