Revista Colombiana de Estadística
Junio 2013, volumen 36, no. 1, pp. 145 a 152

Improved Exponential Type Ratio Estimator of Population Variance

Estimador tipo razón exponencial mejorado para la varianza poblacional

Subhash Kumar Yadav[1,a], Cem Kadilar[2,b]

[1]Department of Mathematics & Statistics (A Centre of Excellence), Dr. Rml Avadh University, Faizabad, India

[2]Department of Statistics, Hacettepe University, Ankara, Turkey

Abstract

This article considers the problem of estimating the population variance using auxiliary information. An improved version of Singh's exponential type ratio estimator has been proposed and its properties have been studied under large sample approximation. It is shown that the proposed exponential type ratio estimator is more efficient than that considered by the Singh estimator, conventional ratio estimator and the usual unbiased estimator under some realistic conditions. An empirical study has been carried out to judge the merits of the suggested estimator over others.

***Key words*:** Auxiliary variable, Bias, Efficiency, Mean squared error.

Resumen

Este artículo considera el problema de estimar la varianza poblacional usando información auxiliar. Una versión mejorada de un estimador exponencial tipo razón de Singh ha sido propuesta y sus propiedades han sido estudiadas bajo aproximaciones de grandes muestras. Se muestra que el estimador exponencial tipo razón propuesto es más eficiente que el estimador de Singh, el estimador de razón convencional y el estimador insesgado usual bajo algunas condiciones realísticas. Un estudio empírico se ha llevado a cabo con el fin de juzgar los méritos del estimador sugerido sobre otros disponibles.

***Palabras clave*:** eficiencia, error cuadrático medio, sesgo, variable auxiliar.

[a]Assistant professor. E-mail: drskystats@gmail.com
[b]Professor. E-mail: kadilar@hacettepe.edu.tr

1. Introduction

The auxiliary information in sampling theory is used for improved estimation of parameters thereby enhancing the efficiencies of the estimators.

Let $(x_i,\, y_i),\; i = 1,\, 2,\, \ldots,\, n$ be the n pair of sample observations for the auxiliary and study variables, respectively, drawn from the population of size N using Simple Random Sampling without Replacement. Let $\overline{X}$ and $\overline{Y}$ be the population means of auxiliary and study variables, respectively, and let $\overline{x}$ and $\overline{y}$ be the respective sample means. Ratio estimators are used when the line of regression of y on x passes through the origin and the variables x and y are positively correlated to each other, while product estimators are used when x and y are negatively correlated to each other; otherwise, regression estimators are used.

The sample variance estimator of the population variance is defined as

$$t_0 = s_y^2 \tag{1}$$

which is an unbiased estimator of population variance $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$ and its variance is

$$V\left(t_0\right) = \gamma\, S_y^4\left(\lambda_{40} - 1\right) \tag{2}$$

where $\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}}$, $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^r(X_i - \overline{X})^s$, and $\gamma = \frac{1}{n}$.

Isaki (1983) proposed the ratio type estimator for estimating the population variance of the study variable as

$$t_R = s_y^2\left(\frac{S_x^2}{s_x^2}\right) \tag{3}$$

where

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2, s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2$$

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N}X_i, \overline{Y} = \frac{1}{N}\sum_{i=1}^{N}Y_i, \overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i, \overline{y} = \frac{1}{n}\sum_{i=1}^{n}y_i$$

The Bias (B) and Mean Squared Error (MSE) of the estimator in (3), up to the first order of approximation, are given, respectively, as

$$B(t_R) = \gamma S_y^2\left[(\lambda_{40} - 1) - \,(\lambda_{22} - 1)\right] \tag{4}$$

$$MSE(t_R) = \gamma S_y^4\left[(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2\left(\lambda_{22} - 1\right)\right] \tag{5}$$

Singh, Chauhan, Sawan & Smarandache (2011) proposed the exponential ratio estimator for the population variance as

$$t_{Re} = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right] \tag{6}$$

The bias and MSE, up to the first order of approximation, respectively, are

$$B(t_{Re}) = \gamma\, S_y^2 \left[\frac{3}{8}(\lambda_{04}-1) - \frac{1}{2}(\lambda_{22}-1)\right] \tag{7}$$

$$MSE(t_{Re}) = \gamma\, S_y^4 \left[(\lambda_{40}-1) + \frac{(\lambda_{04}-1)}{4} - (\lambda_{22}-1)\right] \tag{8}$$

The usual linear regression estimator for population variance is

$$\widehat{S}_{lr}^2 = s_y^2 + b(S_x^2 - s_x^2) \tag{9}$$

where $b = \frac{s_y^2(\widehat{\lambda}_{22}-1)}{s_x^2(\widehat{\lambda}_{04}-1)}$ is the sample regression coefficient.

The MSE of $\widehat{S}_{lr}^2$, to the first order of approximation, is

$$MSE(\widehat{S}_{lr}^2) = \gamma\, S_y^4 \left[(\lambda_{40}-1) - \frac{(\lambda_{22}-1)^2}{\lambda_{04}-1}\right] \tag{10}$$

Many more authors, including Singh & Singh (2001, 2003), Nayak & Sahoo (2012), among others, have contributed to variance estimation.

2. Improved Exponential Type Ratio Estimator

Motivated by Upadhyaya, Singh, Chatterjee & Yadav (2011) and following them, we propose the improved ratio exponential type estimator of the population variance as follows:

The ratio exponential type estimator due to Singh et al. (2011) is given by

$$t_{Re} = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right] = s_y^2 \exp\left[1 - \frac{2s_x^2}{S_x^2 + s_x^2}\right]$$

which can be generalized by introducing a positive real constant 'α' (i.e. $\alpha \geq 0$) as

$$t_{Re}^{(\alpha)} = s_y^2 \exp\left[1 - \frac{\alpha s_x^2}{S_x^2 + (\alpha-1)s_x^2}\right] = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + (\alpha-1)s_x^2}\right] \tag{11}$$

Here, we note that: (i) For $\alpha = 0$, $t_{Re}^{(\alpha)}$ in (11) reduces to

$$t_{Re}^{(0)} = s_y^2 \exp[1] \tag{12}$$

which is a biased estimator with the MSE larger than s_y^2 utilizing no auxiliary information as the value of 'e' is always greater than unity.

(ii) For $\alpha = 1$, $t_{Re}^{(\alpha)}$ in (11) reduces to

$$t_{Re}^{(1)} = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2}\right] \tag{13}$$

(iii) For $\alpha = 2$, $t_{Re}^{(\alpha)}$ in (11) reduces to Singh et al. (2011) ratio exponential type estimator

$$t_{Re} = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right] \tag{14}$$

Then, we have investigated for which value of α, the proposed estimator, $t_{Re}^{(\alpha)}$, is more efficient than the estimators, t_0, t_R, and t_{Re}.

3. The First Degree Approximation to the Bias and Mean Squared Error of the Suggested Estimator

In order to study the large sample properties of the proposed class of estimator, $t_{Re}^{(\alpha)}$, we define $s_y^2 = S_y^2\left(1+\varepsilon_0\right)$ and $s_x^2 = S_x^2\left(1+\varepsilon_1\right)$ such that $E\left(\varepsilon_i\right) = 0$ for $(i = 0, 1)$ and $E\left(\varepsilon_0^2\right) = \gamma\left(\lambda_{40} - 1\right)$, $E\left(\varepsilon_1^2\right) = \gamma\left(\lambda_{04} - 1\right)$, $E\left(\varepsilon_0\varepsilon_1\right) = \gamma\left(\lambda_{22} - 1\right)$.

To the first degree of approximation, the bias and the MSE of the estimator, $t_{Re}^{(\alpha)}$, are respectively given by

$$B(t_{Re}^{(\alpha)}) = \gamma S_y^2 \frac{(\lambda_{04} - 1)}{2\alpha^2}\left[2\alpha(1-\lambda) - 1\right] \tag{15}$$

$$MSE(t_{Re}^{(\alpha)}) = \gamma\, S_y^4\left[(\lambda_{40} - 1) + \frac{(\lambda_{04} - 1)}{\alpha^2}(1 - 2\alpha\lambda)\right] \tag{16}$$

where $\lambda = \frac{\lambda_{22}-1}{\lambda_{04}-1}$.

The $MSE(t_{Re}^{(\alpha)})$ is minimum for

$$\alpha = \frac{1}{\lambda} = \alpha_0 \text{ (say)} \tag{17}$$

Substituting $\alpha = \frac{1}{\lambda}$ into (11), we get the asymptotically optimum estimator (AOE) in the class of estimators $(t_{Re}^{(\alpha)})$ as

$$(t_{Re}^{(\alpha_0)}) = s_y^2 \exp\left[\frac{\lambda(S_x^2 - s_x^2)}{\lambda S_x^2 + (1-\lambda)s_x^2}\right] \tag{18}$$

The value of λ can be obtained from the previous surveys or the experience gathered in due course of time, for instance, see Murthy (1967), Reddy (1973, 1974) and Srivenkataramana & Tracy (1980), Singh & Vishwakarma (2008), Singh & Kumar (2008) and Singh & Karpe (2010).

The mean square error of AOE $(t_{Re}^{(\alpha_0)})$, to the first degree of approximation, is given by

$$MSE(t_{Re}^{(\alpha_0)}) = \gamma\, S_y^4\left[(\lambda_{40} - 1) - \frac{(\lambda_{22} - 1)^2}{\lambda_{04} - 1}\right] \tag{19}$$

which equals to the approximate MSE of the usual linear regression estimator of population variance.

In case the practitioner fails to guess the value of 'λ' by utilizing all his resources, it is worth advisable to replace λin (18) by its consistent estimate

$$\widehat{\lambda} = \frac{\widehat{\lambda}_{22} - 1}{\widehat{\lambda}_{04} - 1} \tag{20}$$

Thus, the substitution of $\widehat{\lambda}$ in (18) yields an estimator based on estimated 'λ' as

$$(t_{Re}^{(\widehat{\alpha}_0)}) = s_y^2 \exp\left[\frac{\widehat{\lambda}(S_x^2 - s_x^2)}{\widehat{\lambda}S_x^2 + (1 - \widehat{\lambda})s_x^2}\right] \tag{21}$$

It can be shown to the first degree of approximation that

$$MSE(t_{Re}^{(\alpha_0)}) = MSE(t_{Re}^{(\widehat{\alpha}_0)}) = \gamma\, S_y^4 \left[(\lambda_{40} - 1) - \frac{(\lambda_{22} - 1)^2}{\lambda_{04} - 1}\right] \tag{22}$$

Thus, the estimator $t_{Re}^{(\widehat{\alpha}_0)}$, given in (21), is to be used in practice as an alternative to the usual linear regression estimator.

4. Efficiency Comparisons of the Proposed Estimator with the Mentioned Existing Estimators

From (16) and (2), we have $MSE(t_0) - MSE(t_{Re}^{(\alpha)}) = \gamma\, S_y^4 \frac{(\lambda_{04}-1)}{\alpha^2}\,(1 - 2\alpha\lambda) > 0$, if

$$\alpha > \frac{1}{2\lambda} \tag{23}$$

From (16) and (5), we have $MSE(t_R) - MSE(t_{Re}^{(\alpha)}) = \gamma\, S_y^4(\lambda_{04} - 1)(1 - \frac{1}{\alpha})\,(1 + \frac{1}{\alpha} - 2\lambda) > 0$, if either

$$\min\left\{1, \frac{1}{2\lambda - 1}\right\} < \alpha < \max\left\{1, \frac{1}{2\lambda - 1}\right\}, \quad \lambda > \frac{1}{2} \tag{24}$$

or

$$\alpha > 1, \quad 0 \le \lambda \le \frac{1}{2}. \tag{25}$$

From (16) and (8), we have $MSE(t_{Re}) - MSE(t_{Re}^{(\alpha)}) = \gamma\, S_y^4(\lambda_{04} - 1)(\frac{1}{2} - \frac{1}{\alpha})\,(\frac{1}{2} + \frac{1}{\alpha} - 2\lambda) > 0$, if either

$$\min\left\{2, \frac{2}{4\lambda - 1}\right\} < \alpha < \max\left\{2, \frac{2}{4\lambda - 1}\right\}, \quad \lambda > \frac{1}{4} \tag{26}$$

or

$$\alpha > 2, \quad 0 \le \lambda \le \frac{1}{4} \tag{27}$$

From (16) and (10), we have

$$MSE(t_{Re}^{(\alpha)}) - MSE(\widehat{S}_{lr}^{2}) = \gamma\, S_y^4(\lambda_{04}-1)\left[\frac{(\lambda_{04}-1)}{\alpha} - (\lambda_{22}-1)\right]^2 < 0$$

if

$$\lambda_{04} < 1 \tag{28}$$

5. Numerical Illustrations

The appropriateness of the proposed estimator has been verified with the help of the four data sets, given in Table 1 (Subramani & Kumarapandiyan 2012). In Table 2, which gives the range of α and also the optimal value, α_0, for the efficiency condition of the proposed estimator, we see that $(t_{Re}^{(\alpha)})$, is quite wide as t_{Re}; whereas, from Table 3, which provides the Percent Relative Efficiencies (PREs) of different estimators of the population variance with respect to the sample variance, we observe that the proposed estimator is more efficient than t_{Re}.

TABLE 1: Parameters of populations.

Parameters	Population 1	Population 2	Population 3	Population 4
N	103	103	80	49
n	40	40	20	20
$\overline{Y}$	626.2123	62.6212	51.8264	116.1633
$\overline{X}$	557.1909	556.5541	11.2646	98.6765
ρ	0.9936	0.7298	0.9413	0.6904
S_y	913.5498	91.3549	18.3569	98.8286
C_y	1.4588	1.4588	0.3542	0.8508
S_x	818.1117	610.1643	8.4563	102.9709
C_x	1.4683	1.0963	0.7507	1.0435
λ_{04}	37.3216	17.8738	2.8664	5.9878
λ_{40}	37.1279	37.1279	2.2667	4.9245
λ_{22}	37.2055	17.2220	2.2209	4.6977
λ	0.9969	0.9635	0.7748	0.7846

TABLE 2: Range of 'α' for $(t_{Re}^{(\alpha)})$ to be more efficient than different estimators of the population variance.

	Populations			
Estimators	**1**	**2**	**3**	**4**
t_0	$\alpha > 0.50$	$\alpha > 0.52$	$\alpha > 0.65$	$\alpha > 0.64$
t_R	$\alpha \in (1.00, 1.01)$	$\alpha \in (1.00, 1.08)$	$\alpha \in (1.00, 1.82)$	$\alpha \in (1.00, 1.76)$
t_{Re}	$\alpha \in (0.67, 2.00)$	$\alpha \in (0.70, 2.00)$	$\alpha \in (0.95, 2.00)$	$\alpha \in (0.94, 2.00)$
Common Range of α for $\left(t_{Re}^{(\alpha)}\right)$ to be more efficient than t_0, t_R, t_{Re}	$\alpha \in (0.67, 2.00)$	$\alpha \in (0.70, 2.00)$	$\alpha \in (0.95, 2.00)$	$\alpha \in (0.94, 2.00)$
Optimum value of α	$\alpha_0 = 1.003$	$\alpha_0 = 1.038$	$\alpha_0 = 1.291$	$\alpha_0 = 1.275$

TABLE 3: Percent relative efficiencies (PREs) of different estimators of population variance with respect to sample variance $t_0 = s_y^2$.

	Populations			
Estimators	**1**	**2**	**3**	**4**
$t_0 = s_y^2$	100.00	100.00	100.00	100.00
t_R	93,838.70	175.74	183.23	258.72
t_{Re}	401.30	149.76	247.21	266.29
$t_{Re}^{(\widehat{\alpha}_0)}$	**94,749.28**	**175.96**	**270.63**	**331.68**

6. Conclusion

We have suggested an improved exponential ratio estimator for estimating the population variance. From theoretical discussions, given in Section 4 and results in Table 3, we infer that the proposed estimator is better than the mentioned existing estimators in literature, the usual sample variance , traditional ratio estimator due to Isaki (1983) and Singh et al. (2011) exponential ratio estimator in the sense of having lesser mean squared error. We have also given the range of α along with its optimum value for to be more efficient than different estimators. Hence, the proposed estimator is recommended for its practical use for estimating the population variance when the auxiliary information is available. In future articles, we hope to adapt the proposed estimator here to the combined and separate methods in the stratifed random sampling.

Acknowledgements

The authors are grateful to the referees and the Editor-in-Chief for providing some useful comments on an earlier draft of the paper.

[Recibido: agosto de 2012 — Aceptado: mayo de 2013]

References

Isaki, C. (1983), ‘Variance estimation using auxiliary information’, *Journal of the American Statistical Association* **78**, 117–123.

Murthy, M. (1967), *Sampling Theory and Methods*, Calcutta Statistical Publishing Society, Kolkatta, India.

Nayak, R. & Sahoo, L. (2012), ‘Some alternative predictive estimators of population variance’, *Revista Colombiana de Estadistica* **35**(3), 507–519.

Reddy, V. (1973), ‘On ratio and product methods of estimation’, *Sankhya Serie B* **35**(3), 307–316.

Reddy, V. (1974), ‘On a transformed ratio method of estimation’, *Sankhya Serie C* **36**, 59–70.

Singh, H. & Karpe, N. (2010), 'Estimation of mean, ratio and product using auxiliary information in the presence of measurement errors in sample surveys', *Journal of Statistical Theory and Practice* **4**(1), 111–136.

Singh, H. & Kumar, S. (2008), 'A general family of estimators of finite population ratio, product and mean using two phase sampling scheme in the presence of non-response', *Journal of Statistical Theory and Practice* **2**(4), 677–692.

Singh, H. & Singh, R. (2001), 'Improved ratio-type estimator for variance using auxiliary information', *Journal of Indian Society of Agricultural Statistics* **54**(3), 276–287.

Singh, H. & Singh, R. (2003), 'Estimation of variance through regression approach in two phase sampling', *Aligarh Journal of Statistics* **23**, 13–30.

Singh, H. & Vishwakarma, G. (2008), 'Some families of estimators of variance of stratified random sample mean using auxiliary information', *Journal of Statistical Theory and Practice* **2**(1), 21–43.

Singh, R., Chauhan, P., Sawan, N. & Smarandache, F. (2011), 'Improved exponential estimator for population variance using two auxiliary variables', *Italian Journal of Pure and Applied Mathematics* **28**, 101–108.

Srivenkataramana, T. & Tracy, D. (1980), 'An alternative to ratio method in sample surveys', *Annals of the Institute of Statistical Mathematics* **32**, 111–120.

Subramani, J. & Kumarapandiyan, G. (2012), 'Variance estimation using median of the auxiliary variable', *International Journal of Probability and Statistics* **1**(3), 36–40.

Upadhyaya, L., Singh, H., Chatterjee, S. & Yadav, R. (2011), 'Improved ratio and product exponential type estimators', *Journal of Statistical Theory and Practice* **5**(2), 285–302.