Journal of Modern Applied Statistical Methods

Volume 15 | Issue 2 Article 15

11-1-2016

A New Exponential Type Estimator for the Population Mean in Simple Random Sampling

Gamze Özel Kadilar
Hacettepe University, Ankara, Turkey, gamzeozl@hacettepe.edu.tr

Journal of Modern Applied Statistical Methods
November 2016, Vol. 15, No. 2, 207-214.
doi: 10.22237/jmasm/1478002380

ISSN 1538 − 9472

A New Exponential Type Estimator for the Population Mean in Simple Random Sampling

Gamze Özel Kadilar
Hacettepe University
Ankara, Turkey

This paper provides a new exponential type estimator in simple random sampling for population mean. It is shown that proposed exponential type estimator is always more efficient than estimators considered by Bahl and Tuteja (1991) and Singh, Chauhan, Sawan, and Smarandache (2009). From numerical examples it is also observed that proposed modified ratio estimator performs better than existing estimators.

Keywords: Simple random sampling, Ratio and regression-type estimator, Auxiliary information, Mean squared error, Efficiency

Introduction

The auxiliary information in sampling theory is used for improved estimation of parameters enhancing the efficiencies of the estimators. The problem of estimating the population mean in the presence of an auxiliary variable has been widely discussed in finite population sampling literature. The use of auxiliary information is well-known to improve the precision of the estimate of the population mean for the study variable. In survey sampling, ratio, product and difference methods of estimation are good examples in this context. Ratio method of estimation is quite effective when there is high positive correlation between study and auxiliary variables. However, if correlation is negative (high), the product method of estimation can be employed efficiently. In recent years, a number of research papers on ratio-type, exponential ratio-type and regression-type estimators have appeared, based on different types of transformations. The main aspect of the paper is to obtain an estimator to predict the population mean

Dr. Kadilar is an Associate Professor in the Department of Statistics. Email her at: gamzeozl@hacettepe.edu.tr.

which is more efficient than the ratio, product, exponential estimators of Bahl and Tuteja (1991) and Singh et al. (2009).

Consider a finite population $U = U_1, U_2,\ldots, U_N$ of N units. Let Y and X stand for the variable under study and auxiliary variables, respectively. Let (y_i, x_i), $i = 1, 2,\ldots, n$, denote the n pair of sample observations for the study and auxiliary variables, respectively, drawn from the population size N using simple random sampling without replacement (SRSWOR). Let $\bar{X}$ and $\bar{Y}$ be the population means of auxiliary and study variables, respectively, and let $\bar{x}$ and $\bar{y}$ be the respective sample means. It is well known that the sample mean $\bar{y}$ is an unbiased estimator of $\bar{Y}$ and under SRSWOR its variance is given by

$$V(\bar{y}) = \gamma \bar{Y}^2 C_y^2 \tag{1}$$

where $\gamma = n^{-1}(1-f)$, $f = n/N$, $C_y^2 = S_y^2/\bar{Y}^2$, and S_y^2 is the variance of the study variable.

Ratio and product type estimators in the simple random sampling (SRS) were considered by Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999), Singh and Tailor (2003), Singh, Tailor, Tailor, and Kakran (2004), Singh and Tailor (2005), Yadav and Kadilar (2013a), Singh et al. (2009), Singh, Chauhan, Sawan, and Smarandache (2011), Yadav and Kadilar (2013b), etc. When information is available on X that is positively correlated with Y, the ratio estimator is suitable for estimating the population mean and it is given by

$$\bar{y}_R = \bar{y}\,\bar{X}/\bar{x}.$$

The mean squared error (MSE) of this estimator is

$$\text{MSE}(\bar{y}_R) = \gamma \bar{Y}^2 \left[C_y^2 - 2\rho C_x C_y + C_x^2\right], \tag{2}$$

where $C_x^2 = S_x^2/\bar{X}^2$, and S_x^2 is the variance of the auxiliary variable.

When there is a negative high correlation between Y and X, the product estimator for $\bar{Y}$ was defined by Robson (1957) as

$$\bar{y}_p = \bar{y}\,\bar{x}/\bar{X}$$

and the MSE of the product estimator is given by

$$\mathrm{MSE}\left(\bar{y}_p\right)=\gamma\bar{Y}^2\left[C_y^2+2\rho C_x C_y+C_x^2\right]. \tag{3}$$

Bahl and Tuteja (1991) suggested an exponential ratio type estimator for the population mean as

$$\bar{y}_{BT}=\bar{y}\exp\left[\frac{\left(\bar{X}-\bar{x}\right)}{\left(\bar{X}+\bar{x}\right)}\right]$$

and the MSE of the estimator is given by

$$\mathrm{MSE}\left(\bar{y}_{BT}\right)=\gamma\bar{Y}^2\left[C_y^2-2\rho C_x C_y+C_x^2/4\right]. \tag{4}$$

The auxiliary information associated with X such as mean, median, coefficient of variation, skewness, kurtosis or correlation coefficient can be used to improve the efficiency of the estimators. Singh et al. (2009) defined a modified exponential ratio estimator using auxiliary variable information for estimating $\bar{Y}$ as

$$\bar{y}\ \ =\bar{y}\exp\left[\frac{\left(a\bar{X}+b\right)-\left(a\bar{x}+b\right)}{\left(a\bar{X}+b\right)+\left(a\bar{x}+b\right)}\right],$$

where ($a\neq 0$), b are either real numbers or the functions of the known parameters of the auxiliary variable such as coefficient of variation (C_x), coefficient of kurtosis ($\beta_2(x)$), and correlation coefficient.

The MSE of the modified exponential estimator is given by

$$\mathrm{MSE}\left(\bar{y}_S\right)=\gamma\bar{Y}^2\left(C_y^2+\theta^2C_x^2-2\theta\rho C_x C_y\right), \tag{5}$$

where $\theta=a\bar{X}/2\left(a\bar{X}+b\right)$.

Suggested Estimator

Following Bahl and Tuteja (1991) and Singh et al. (2009), a modified exponential type estimator is defined for estimating $\bar{Y}$ as

$$\bar{y}_{PR} = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha} \exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right). \quad (6)$$

To obtain the MSE of $\bar{y}_{PR1}$, write $\bar{y} = \bar{Y}(1+e_0)$ and $\bar{x} = \bar{X}(1+e_1)$ such that $E(e_0) = E(e_1) = 0$ and $E(e_0^2) = \gamma C_Y^2$, $E(e_1^2) = \gamma C_X^2$, $E(e_0 e_1) = \gamma\rho C_y C_x$. Expressing (6), in terms of e's,

$$\begin{aligned}\bar{y}_{PR} &= \bar{Y}(1+e_0)(1+e_1)^{\alpha} \exp\left[\frac{\bar{X}-\bar{X}(1+e_1)}{\bar{X}+\bar{X}(1+e_1)}\right] \\ &= \bar{Y}(1+e_0)(1+e_1)^{\alpha} \exp\left[\frac{e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}\right].\end{aligned} \quad (7)$$

Expanding the right hand side of (7) and retaining terms up to the second power of e's,

$$\begin{aligned}\bar{y}_{PR} &= \bar{Y}(1+e_0)(1+e_1)^{\alpha} \exp\left[\frac{e_1}{2}\left(1+\frac{e_1}{2}+\frac{e_1^2}{4}+\ldots\right)\right] \\ &= \bar{Y}(1+e_0)\left[1+\alpha e_1+\frac{\alpha(\alpha-1)}{2}e_1^2+\ldots\right]\left(1+\frac{e_1}{2}+\frac{2e_1^2}{8}\right)\end{aligned} \quad (8)$$

From (8),

$$\bar{y}_{PR} - \bar{Y} \cong \bar{Y}\left[\alpha e_1 + \frac{\alpha(\alpha-1)}{2}e_1^2 + \frac{e_1}{2} + \frac{\alpha e_1^2}{2} + \frac{3e_1^2}{8} + e_0 + \alpha e_1 e_0 + \frac{e_0 e_1}{2}\right]. \quad (9)$$

Squaring (9) and then taking expectation of both sides, the MSE of the estimator $\bar{y}_{PR}$ is

$$\text{MSE}(\bar{y}_{PR}) = \gamma \bar{Y}^2\left(C_y^2 + C_x^2/4 + 2\alpha\rho C_x C_y + \rho C_x C_y + \alpha^2 C_x^2 + \alpha C_x^2\right). \quad (10)$$

Obtain the optimum α to minimize $\text{MSE}(\bar{y}_{PR})$. Differentiating $\text{MSE}(\bar{y}_{PR})$ with respect to α and equating the derivative to zero, optimum value of α is given by

$$\alpha_{opt} = \left(C_x - 2\rho C_y\right)\big/2C_x .$$

Substituting the value of α_{opt} in (10), we get the minimum value of $\text{MSE}(\bar{y}_{PR})$ as

$$\text{MSE}_{\min}\left(\bar{y}_{PR}\right) = \gamma \bar{Y}^2 C_y^2 \left(1-\rho^2\right). \tag{11}$$

It follows from (11) that the proposed estimator $\bar{y}_{PR}$ at its optimum condition is equal efficient as that of the usual linear regression estimator.

Efficiency Comparisons

In this section, the MSE of traditional estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}_P$, $\bar{y}_S$, and $\bar{y}_{BT}$ are compared with the MSE of the proposed estimator $\bar{y}_{PR}$. From (1)-(5) and (11),

$$\left[\text{Var}\left(\bar{y}\right) - \text{MSE}_{\min}\left(\bar{y}_{prl}\right)\right] = \lambda \bar{Y}^2 \rho^2 > 0, \tag{12}$$

$$\left[\text{MSE}\left(\bar{y}_R\right) - \text{MSE}_{\min}\left(\bar{y}_{prl}\right)\right] = \lambda \bar{Y}^2 \left(C_x - \rho C_y\right)^2 > 0, \tag{13}$$

$$\left[\text{MSE}\left(\bar{y}_P\right) - \text{MSE}_{\min}\left(\bar{y}_{prl}\right)\right] = \lambda \bar{Y}^2 \left(C_x + \rho C_y\right)^2 > 0, \tag{14}$$

$$\left[\text{MSE}\left(\bar{y}_{BT}\right) - \text{MSE}_{\min}\left(\bar{y}_{prl}\right)\right] = \lambda \bar{Y}^2 \left(C_x/2 - \rho C_y\right)^2 > 0, \tag{15}$$

$$\left[\text{MSE}\left(\bar{y}_S\right) - \text{MSE}_{\min}\left(\bar{y}_{prl}\right)\right] = \lambda \bar{Y}^2 \left(C_x\theta - \rho C_y\right)^2 > 0, \tag{16}$$

It is observed that $\bar{y}_{PR}$ is always more efficient than the traditional estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}_P$, $\bar{y}_S$, and $\bar{y}_{BT}$, because the conditions from (12) to (16) are always satisfied.

Numerical Illustrations

The appropriateness of the proposed estimator has been verified with the help of the following data sets given in Table 1.

Table 1. Statistics of populations

Parameters	Population 1	Population 2	Population 3	Population 4
N	80	104	80	10
n	20	20	20	4
$\bar{Y}$	11.264	6.254	51.826	5.920
$\bar{X}$	51.826	13931.680	2.851	3.590
ρ	0.941	0.860	0.915	1.680
C_y	0.750	1.860	0.354	0.144
C_x	0.354	1.650	0.948	0.128
$\beta_2(x)$	0.063	17.516	1.300	0.381

The explanation of the data sets in Table 1 from various sources is given as follows:

Population 1. Source Murthy (1967): Y is the fixed capital and X is the output of the 80 factories.

Population 2. Source Shabbir, Haq, and Gupta (2014): The study variable Y is the level of apple production (in 1000 tons) and the auxiliary variable X is the number of apple trees in 104 villages in 1999.

Population 3. Source Murthy (1967): The auxiliary variable X is the number of workers and the study variable is the output for 80 factories in a region.

Population 4. Source Cochran (1977): The auxiliary variable X is the number of rooms and the study variable is the number of persons.

The Percent Relative Efficiencies (PREs) of different estimators of the population mean with respect to the sample mean based on Populations 1-4 are given in Table 2.

Table 2. PREs of different estimators of population mean with respect to sample mean $\bar{y}$.

Estimators	Population 1	2	3	4
$\bar{y}$	100.000	100.000	100.000	100.000
$\bar{y}_R$	298.972	382.945	30.586	158.823
$\bar{y}_P$	47.369	30.186	7.651	34.089
$\bar{y}_{BT}$	163.521	230.504	292.078	161.440
$\bar{y}_S$	104.054	323.244	319.840	8.197
$\bar{y}_{PR1}$	873.218	384.025	614.345	173.748

From the values of Table 2, it is observed that the MSE of the proposed estimator is less than the mean squared errors of all the existing estimators. Note that $\bar{y}_S$ requires the auxiliary variable information, on the other hand, one can reach the minimum MSE value using the proposed estimator without auxiliary variable information.

Conclusion

As an improved exponential ratio estimator for estimating the population mean was proposed. The proposed estimator is better than the mentioned existing estimators in literature, in the sense of having lesser mean squared error. Hence, the proposed estimator is recommended for its practical use for estimating the population mean when the auxiliary information is available.

References

Bahl, S., & Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Journal of Information and Optimization Sciences*, *12*(1), 159-163. doi:10.1080/02522667.1991.10699058

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). Wiley Eastern Limited.

Murthy, M. N. (1967). *Sampling theory and methods*. Calcutta, India: Statistical Publishing Society.

Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimation. *Journal of American Statistical Association, 52*(280), 511-522. doi:10.1080/01621459.1957.10501407

Shabbir, J., Haq, A., & Gupta, S. (2014). A new difference-cum-exponential type estimator of finite population mean in simple random sampling. *Revista Colombiana de Estadística*, *37*(1), 197-209. doi:10.15446/rce.v37n1.44366

Singh, H. P. & Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. *Statistics in Transition, 6*(4), 555-560.

Singh, H. P., & Tailor, R. (2005). Estimation of finite population mean with coefficient of variation of an auxiliary character. *Statistica, 65*(3), 407-413. doi:10.6092/issn.1973-2201/1299

Singh, H. P., Tailor, R., Tailor, R., & Kakran, M. S. (2004). An improved estimator of population mean using power transformation. *Journal of the Indian Society Agricultural Statistics, 58*(2), 223-230.

Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2009). Improvement in estimating the population mean using exponential estimator in simple random sampling. *International Journal of Statistics and Economics*, *3*(A09), 13-18.

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*, 101-109.

Sisodia, B. V. S., & Dwivedi, V. K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable. *Journal of the Indian Society Agricultural Statistics, 33*(2), 13-18.

Upadhyaya, L. N., & Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal, 41*(5), 627-636. doi:10.1002/(SICI)1521-4036(199909)41:5<627::AID-BIMJ627>3.0.CO;2-W

Yadav, S. K., & Kadilar, C. (2013a). Improved Class of Ratio and Product Estimators. Applied Mathematics and Computation, 219 (22), 10726-10731

Yadav, S. K., & Kadilar, C. (2013b). Improved exponential type ratio estimator of population variance. *Revista Colombiana de Estadística, 36*(1), 145-152.