Estimators Proposed by Geometric Mean, Harmonic Mean and Quadratic Mean

Vedat Sağlam, Tolga Zaman, Erdinç Yücesoy, Murat Sağır

Department of Statistics, Faculty of Science and Arts, Ondokuz Mayıs University, Kurupelit, Turkey

Email address:
vsaglam@omu.edu.tr (V. Sağlam)

Abstract: In the studies in literature up to date arithmetic population mean of auxiliary variable is used to obtain the proportional estimators. In this paper geometric mean, harmonic mean and quadratic mean is used as well as arithmetic population mean. Using geometric mean, harmonic mean and quadratic mean do not affect the variance of ratio estimator ($\hat{R}$). However new approaches are obtained for the estimation and variance of the dependent variable when these means are used as well as arithmetic population mean. In the application, the mean number of teaching staff of the departments in Ondokuz Mayıs University is estimated by auxiliary variable which is the number of students in the departments. In addition the variances of proportional estimation method are obtained and interpreted by using population arithmetic mean, geometric mean, harmonic and quadratic mean.

Keywords: Proportional Estimation, Geometric Mean, Harmonic Mean, Quadratic Mean, Inequalities

1. Introduction

One of the basic estimation method, using auxiliary variable to estimate the population mean, in simple random sampling is the proportional estimator method. In this study the proportional estimators, used to estimate population mean, and the variances of these estimators are given. In all studies regarding the proportional estimators, arithmetic mean is chosen as the population mean of the auxiliary variable. In this paper geometric mean, harmonic mean and quadratic mean are used as the population mean as well as arithmetic mean. Using these means as well as arithmetic will not change the variance of proportional estimator $\hat{R}$. However by using these means and under assumption $x_i > 0, (i = 1, 2, ..., N)$, new approaches are obtained for the estimation of the dependent variable and its variance. Geometric mean, harmonic mean and quadratic mean are investigated because the population mean is fixed. The aim is to minimize the variance for the estimators. In the application part the data is obtained from the departments of Ondokuz Mayıs University. The number of teaching staff at the departments are estimated via the number of student which is considered as the auxiliary variable. Some other means of the population also calculated and variances are calculated for

the proportional estimation method approaches. There are several new estimators in the literature proposed for population mean and population sum derived from known proportional estimation. "Cochran (1977) investigated the use of information obtained from the auxiliary variable [1]."

"Dayveh et al. (2003) proposed two estimators for the estimation of population mean using two auxiliary variable and generalized these estimators for more than two auxiliary variables [2]". "Sign and Tailor (2003) proposed new estimators by using correlation coefficient information [3]". "Ray and Singh (1981) developed a new proportional estimator using regression coefficient [4]". On the other hand "Kadılar and Çıngı (2004) investigated a special case of this estimator [5]". "Kadılar and Çıngı (2005) developed a new estimator, for population mean in two variable regression estimator, using the two auxiliary variable multiplicative estimator proposed by Dayveh et al. [6]". "İsaki (1983) proposed a proportional estimator for the population variance [10]" and "by developing this estimator Prasad and Singh (1992) proposed new estimators for the population variance [11]". "Garcia and Cebrian (1996) proposed a new estimator for the population variance by using auxiliary variable and generalized this estimator for more than one auxiliary variable by a iteration method [12]". Proceeding from here

the following estimators are proposed. "Beale (1962) and Tin (1965) used variation coefficients of dependent variable and independent variable as auxiliary variable to propose new proportional estimators for population mean [13, 14]". Additionally "Zaman et al. (2014), investigated mean square errors of some ratio estimators and means via Taylor Series approach and made comparisons [15]". Proceeding from here the following estimators are proposed.

2. Proportional Estimators

Suppose that a sample with a size n is chosen out of a population N by random simple sampling method. It is assumed that for the values of y and x the change of the proportion x/y is very small and the correlation between the variables is denoted by a line passing from the origin point. Then the sample means of the variables are as follows:

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} \tag{1}$$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2}$$

The population means of the variables are:

$$\overline{Y} = \frac{\sum_{i=1}^{N} y_i}{N} \tag{3}$$

$$\overline{Y} = \frac{\sum_{i=1}^{N} y_i}{N} \quad \overline{X} = \frac{\sum_{i=1}^{N} x_i}{N} \tag{4}$$

If $\rho(x,y) > 0$ then the proportional estimation of the population mean of variable y is obtained as,

$$\widehat{\overline{Y}}_r = \frac{\overline{y}}{\overline{x}} \overline{X} \tag{5}$$

and proportion of the variables is

$$R = \frac{\overline{Y}}{\overline{X}}. \tag{6}$$

The estimator of R is given as follows

$$\hat{R} = \frac{\overline{y}}{\overline{x}} \tag{7}$$

Substituting equation (7) in equation (5), we have

$$\widehat{\overline{Y}}_r = \hat{R}\overline{X} \tag{8}$$

"Cochran, 1977, [1]". The variance of the estimator $\widehat{\overline{Y}}_r$ is

$$Var(\hat{R}) = \frac{N-n}{nN\overline{X}^2} \frac{\sum_{i=1}^{N}(y_i - Rx_i)^2}{N-1} \tag{9}$$

"Yamane, 2001, [7]"

3. Proportional Estimator and New Estimators Proposed by Other Means

New estimators are given by using geometric mean, harmonic mean and quadratic mean instead of arithmetic population mean in equation (8) under condition $x_i >$ $0 (i = 1, 2, \dots, N)$.

3.1. Estimator Proposed by Geometric Mean

Definition 1. The geometric mean of a population is given as,

$$\text{GM} = \sqrt[N]{\prod_{i=1}^{N} x_i} \tag{10}$$

where $x_1, x_2, \cdots, x_N$ are values of auxiliary variable of the population.

Replacing equation (10) instead of $\overline{X}$ in equation (8) we have following estimator

$$\widehat{\overline{Y}}_{r\text{GM}} = \hat{R}\text{GM} \tag{11}$$

Expected value of this estimator is

$$E\left(\widehat{\overline{Y}}_{r\text{GM}}\right) = E(\hat{R})\text{GM} \tag{12}$$

or

$$E\left(\widehat{\overline{Y}}_{r\text{GM}}\right) = E\left(\frac{\overline{y}}{\overline{x}}\right) \sqrt[N]{\prod_{i=1}^{N} x_i} \tag{13}$$

and its variance is,

$$Var\left(\widehat{\overline{Y}}_{r\text{GM}}\right) = \text{GM}^2 Var(\hat{R}) \tag{14}$$

or

$$Var\left(\widehat{\overline{Y}}_{r\text{GM}}\right) = \left[\sqrt[N]{\left(\prod_{i=1}^{N} x_i\right)}\right]^2 \left[\frac{N-n}{nN\overline{X}^2} \frac{\sum_{i=1}^{N}(y_i - Rx_i)^2}{N-1}\right] \tag{15}$$

3.2. Estimator Proposed by Harmonic Mean

Definition 2. The harmonic mean of a population is given as,

$$\text{HM} = N \Big/ \left(1/\sum_{i=1}^{N} \frac{1}{x_i}\right) \tag{16}$$

Where $x_1, x_2, \cdots, x_N$ are values of auxiliary variable of the population.

Replacing equation (16) instead of $\overline{X}$ in equation (8) the estimator,

$$\widehat{\overline{Y}}_{r\text{HM}} = \hat{R}\text{HM} \tag{17}$$

is obtained.

Expected value of this estimator is,

$$E\left(\widehat{\overline{Y}}_{r\text{HM}}\right) = E(\hat{R})\text{HM} \tag{18}$$

or

$$E\left(\widehat{\overline{Y}}_{r\text{HM}}\right) = E\left(\frac{\overline{y}}{\overline{x}}\right) \left[N \Big/ \left(1/\sum_{i=1}^{N} \frac{1}{x_i}\right)\right] \tag{19}$$

and its variance is obtained as,

$$Var\left(\widehat{\overline{Y}}_{r\text{HM}}\right) = \text{HM}^2 Var(\hat{R}) \tag{20}$$

or

$$Var\left(\widehat{\overline{Y}}_{r\mathbf{HM}}\right) = \left[N/\left(1/\sum_{i=1}^{N}\frac{1}{x_i}\right)\right]^2 \left[\frac{N-n}{n\overline{X}^2}\frac{\sum_{i=1}^{N}(y_i-Rx_i)^2}{N-1}\right] \quad (21)$$

3.3. Estimator Proposed by Quadratic Mean

Definition 3. The quadratic mean of a population is given as,

$$\mathbf{QM} = \sqrt{\frac{\sum_{i=1}^{N}x_i^2}{N}} \quad (22)$$

where $x_1, x_2, \cdots, x_N$ are values of auxiliary variable of the population.

Replacing equation (22) instead of $\overline{X}$ in equation (8) the estimator,

$$\widehat{\overline{Y}}_{r\mathbf{QM}} = \hat{R}\mathbf{QM} \quad (23)$$

is obtained.

Expected value of this estimator is,

$$E\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right) = E(\hat{R})\mathbf{QM} \quad (24)$$

or

$$E\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right) = E\left(\frac{\overline{y}}{\overline{x}}\right)\sqrt{\frac{\sum_{i=1}^{N}x_i^2}{N}} \quad (25)$$

and its variance is obtained as,

$$Var\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right) = \mathbf{QM}^2 Var(\hat{R}) \quad (26)$$

or

$$Var\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right) = \frac{\sum_{i=1}^{N}x_i^2}{N}\left[\frac{N-n}{nN\overline{X}^2}\frac{\sum_{i=1}^{N}(y_i-Rx_i)^2}{N-1}\right] \quad (27)$$

4. Comparisons

4.1. The Comparison of the Estimators

The following inequalities are given for the comparison of arithmetic mean against other means "Shahbazov, 2005, [8]"

$$\mathbf{HM} \leq \overline{X},$$

$$N/\left(1/\sum_{i=1}^{N}\frac{1}{x_i}\right) \leq \frac{1}{N}\sum_{i=1}^{N}x_i.$$

$$GM \leq \overline{X},$$

$$\sqrt[N]{\prod_{i=1}^{N}x_i} \leq \frac{1}{N}\sum_{i=1}^{N}x_i. \quad (28)$$

$$QM \geq \overline{X},$$

$$\sqrt{\frac{\sum_{i=1}^{N}x_i^2}{N}} \geq \frac{1}{N}\sum_{i=1}^{N}x_i.$$

where $x_1, x_2, \cdots, x_N$ are values of auxiliary variable of the population.

4.2. Comparison of the Means

Using inequalities (28), the following inequalities are obtained:

$$E\left(\widehat{\overline{Y}}_r\right) \geq E\left(\widehat{\overline{Y}}_{r\mathbf{HM}}\right)$$

$$E\left(\widehat{\overline{Y}}_r\right) \geq E\left(\widehat{\overline{Y}}_{r\mathbf{GM}}\right) \quad (29)$$

$$E\left(\widehat{\overline{Y}}_r\right) \leq E\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right)$$

4.3. The Comparison of the Variances

Similarly using inequalities (28), the inequalities for the variances of estimators are given below:

$$Var\left(\widehat{\overline{Y}}_r\right) \geq Var\left(\widehat{\overline{Y}}_{r\mathbf{HM}}\right)$$

$$Var\left(\widehat{\overline{Y}}_r\right) \geq Var\left(\widehat{\overline{Y}}_{r\mathbf{GM}}\right) \quad (30)$$

$$Var\left(\widehat{\overline{Y}}_r\right) \leq Var\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right)$$

In equation (30) the inequality $Var\left(\widehat{\overline{Y}}_r\right) \leq Var\left(\widehat{\overline{Y}}_{r\mathbf{QM}}\right)$ is not a desirable situation therefore the estimator $\widehat{\overline{Y}}_{rQM}$ is ignored. In this manner, using the following inequality

$$N/\left(1/\sum_{i=1}^{N}\frac{1}{x_i}\right) \leq \sqrt[N]{\prod_{i=1}^{N}x_i} \leq \frac{1}{N}\sum_{i=1}^{N}x_i \quad (31)$$

we have,

$$Var\left(\widehat{\overline{Y}}_{r\mathbf{HM}}\right) \leq Var\left(\widehat{\overline{Y}}_{r\mathbf{GM}}\right) \leq) Var\left(\widehat{\overline{Y}}_r\right) \quad (32)$$

According to these results, if the population values of the auxiliary variable are known and positive then the estimators $\widehat{\overline{Y}}_{rHM}$ or $\widehat{\overline{Y}}_{rGM}$ can be proposed. An actual data application, presented in section 4, supports these results.

5. Application

The actual data needed for application is obtained from the actively teaching departments of Ondokuz Mayıs University. The number of teaching staff in the university is denoted by the variable y and the number of students in the university is denoted by the variable x. When the correlation between these two variables is positive the variable, considered by auxiliary variable, can be estimated. Each of the department in the university is population unit and the sample size estimation is calculated by equation (33) "Çıngı, 2009, [9]"

$$n_0 = \frac{t^2 S^2}{d^2}, n = \frac{n_0}{1+\frac{n_0}{N}} \quad (33)$$

In equation (33) when $\alpha = 0.05$, $d = 0.30$ the sample equals to 14. Arithmetic mean, geometric mean and harmonic mean of the number of students are given in Table 1 respectively.

Table 1. *The population means of the variable x.*

$\bar{X} = 475.7083$
$GM = 302.7891$
$HM = 152.064$

Data of the chosen sample is obtained as in Table 2 by using random numbers table.

Table 2. *Obtained values of the sample.*

Department	Number of teaching staff	Number of students
Faculty of Medicine	290	1700
Fine Arts	11	439
Turkish Language	4	301
Foreign Languages	17	1164
Archaeology	5	204
Biology	23	262
Art History	3	196
Sociology	7	240
Economy	9	404
International Trade and Logistics	3	118
Communication	11	304
Chemistry Engineering	7	209
Music	9	28
Civil Aviation	11	189

The estimated values and estimated variances of the proposed new estimators are calculated by Table 2 and given in Table 3.

Table 3. *The estimators and numerical values of the variances.*

$\hat{\bar{Y}}_r = 33.870$	$Var\left(\hat{\bar{Y}}_r\right) = 153.017$
$\hat{\bar{Y}}_{r\mathbf{GM}} = 21.558$	$Var\left(\hat{\bar{Y}}_{r\mathbf{GM}}\right) = 61.992$
$\hat{\bar{Y}}_{r\mathbf{HM}} = 10.826$	$Var\left(\hat{\bar{Y}}_{r\mathbf{HM}}\right) = 15.635$

In Table 3, we notice that the variances of estimators, obtained by using geometric and harmonic means, are smaller.

6. Conclusions

In this paper new mean estimators are proposed by the population geometric mean, population harmonic mean and population quadratic mean of the auxiliary variable. The expected values and variances of these proposed new estimators theoretically obtained. The estimators proposed by the quadratic mean are ignored since the variances of these estimators are greater than or equal to the variances of the estimators obtained by population arithmetic mean. However the estimators proposed by geometric mean and harmonic mean are analyzed since the variances of these estimators are smaller than or equal to the variances of estimators obtained by population arithmetic mean. When the obtained results are discussed it is seen that the variance of the estimator, proposed by harmonic mean, is the smallest. The aim of this paper, assuming that the population values of the auxiliary variable are positive, is to propose new estimators and to find

their variances by using inequalities concerning the means. Finally an application is done using actual data to show that the numerical results tend to theoretically obtained results.

References

[1] Cochran, W. G., "Sampling Techniques", John Wiley and Sons, New-York, 1977.

[2] Dayyeh, W. A. A., Ahmed M. S., Ahmed R. A., Mutlak, H. A, "Some estimators of a finite population mean using auxiliary information", Applied Mathematics and Computation, 139, 287-298, 2003.

[3] Singh, H. P. and Tailor, R., "Use of known correlation coefficient in estimating the finite population mean", Statistics in Transition, Vol. 6, No. 4, 555-560, 2003.

[4] Ray, S. K. and Singh R. K., "Difference-cum-ratio type estimators, Journal of Indian Statistical Association", 19, 147-151, 1981.

[5] Kadılar, C. and Çıngı, H., "Ratio estimators in simple random sampling", Applied Mathematics and Computation, 151, 893-902, 2004.

[6] Kadılar, C. and Çıngı, H, "A new estimator using two auxiliary variables", Applied Mathematics and Computation, 162, 901-908, 2005.

[7] Yamane, T., "Temel Örnekleme Yöntemleri, (Çevirinler: Esin, Aydın, C. Bakır, M. A., Gürbüzsel, E.)" Literatür Yayınları, No: 53, İstanbul, 2001.

[8] Shahbazov, A., "Introduction to Probability Theory", Birsen Yayınevi Ltd. Şti. Kod No: Y. 0029, ISBN: 975-511-415-7, İstanbul, 2005.

[9] Çıngı, H., "Sampling Theory", H. Ü. Fen Fakültesi Basımevi, Beytepe, 2009.

[10] Isaki, C. T., "Variance estimation using auxiliary information, Journal of American Statistical Assosiation", 78, 117-123, 1983.

[11] Prasad, B., Singh, H. P., "Unbiased estimators of finite population variance using auxiliary information in sample surveys", 21, 5, 1367-1376, 1992.

[12] Garcia, M. R., Cebrian, A. A., "Repeated subsitution method: The ratio estimator for the population variance", Metrika, 43, 101-105, 1996.

[13] Beale, E. M. L., "Some use of computers in operational research, Industrielle Organisation", 31, 27-28, 1962.

[14] Tin, M., "Comparison of some ratio estimators", Journal of American Statistical Assosiation, 60, 294-307, 1965.

[15] Zaman, T., Sağlam, V., Sağir, M., Yücesoy, E., and Zobu, M., "Investigation of some estimators Taylor Series approach and an application", Science Publishing Group, 3 (5): 141-147, 2014.