

Investigation of some estimators via taylor series approach and an application

Tolga Zaman¹, Vedat Sağlam¹, Murat Sagır¹, Erdinc Yucesoy¹, Mujgan Zobu²

¹Department of Statistics, Faculty of Science and Arts, Ondokuz Mayıs University, Kurupelit, Turkey

²Department of Statistics, Faculty of Science and Arts, Amasya University, Turkey

Email address:

tolga.zaman@omu.edu.tr (T. Zaman)

To cite this article:

Tolga Zaman, Vedat Sağlam, Murat Sağır, Erdinç Yücesoy, Müjgan Zobu. Investigation of Some Estimators Via Taylor Series Approach and an Application. *American Journal of Theoretical and Applied Statistics*. Vol. 3, No. 5, 2014, pp. 141-147.

doi: 10.11648/j.ajtas.20140305.14

Abstract: In this study, the use of taylor series method in the calculation of some means with single auxiliary variable developed in a simple random sampling and mean square error unit ratio estimators having certain properties was investigated. An application was performed in respect thereof. The study population (mass) included 111 secondary schools from 18 districts of Trabzon province. Auxiliary variable (x) was taken as the number of students whereas the main variable (y) was taken as the number of teachers. Sample size was calculated as 45 for unit ratio that has certain features. Afterwards, theoretically proposed mean and units ratio estimators having certain properties were compared numerically. Random sampling was performed using the SPSS 20 program thus giving an equal chance to the units sampled and variability in the population was protected.

Keywords: Simple Random Sampling, Taylor Series Approach, Mean, Proportional Estimators, Mean Square Error, Units Ratio having Certain Features, Ancillary Information

1. Introduction

Proportional estimator method is not a sampling method but a method for finding estimator. In the case where the correlation between two variables is positive, information from auxiliary variables is used. Newly proposed ratio estimators are frequently encountered in the literature. Data related to auxiliary variable were examined in detail by Cochran (1977). Singh-Tailor (2003) estimation is based on mean ratio estimation proposed in the case where mass correlation coefficient (ρ) of main and auxiliary variables is known. Kadilar and Çingi (2004) proposed new ratio estimators by implementing linear regression estimator rather than simple estimates and using coefficient of variation information (C_x) from auxiliary variable. Kadilar and Çingi (2006) developed a new mean ratio estimator by typing regression estimation itself in Singh and Taylor estimator and using coefficient of mass variable from coefficient of mass correlation and auxiliary variable rather than simple estimator. Naik-Gupta estimator (1996) is the unit ratio estimator having a certain property. Singh et al. (2008) proposed an estimation method based on the

implementation of linear regression estimation having certain features rather than simple conventional estimation in Naik-Gupta estimation and the use of data such as variation coefficient and correlation coefficient.

In our application, the estimators proposed in the literature were compared numerically.

2. Taylor Series Approach

Observation of non-linear estimates is frequently encountered in sampling. It may not be possible to calculate the variance of these non-linear estimators exactly. Estimations such as ratio estimate or estimation of regression coefficient may be used. To calculate the variance of a non-linear estimator, observations of non-linear estimator are approximated to a linear function and an appropriate variance equation is written. This linearization is usually performed by Taylor Series Method. In other words, Taylor series method can be used in the calculation of variance of an estimator.

In general, it can be given for the variable k as follows:

$$h(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k) = h(Y_1, Y_2, \dots, Y_k) + \sum_{j=1}^k d_j (\hat{Y}_j - Y_j) + R_k(Y_k, \alpha) \quad (1)$$

where $h(Y_1, Y_2, \dots, Y_k)$ is a non-linear function and $h(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k)$ is the estimation of this function.

$$d_j = \left. \frac{\partial h(\alpha_1, \alpha_2, \dots, \alpha_k)}{\partial \alpha_j} \right|_{Y_1, Y_2, \dots, Y_k} \quad (2)$$

where ∂ represents the notation of partial derivative and

$$R_k(Y_k, \alpha) = \sum_{j=1}^k \sum_{i=1}^k \frac{1}{2!} \frac{\partial^2 h(Y_1, Y_2, \dots, Y_k)}{\partial Y_j \partial Y_i} (\hat{Y}_j - Y_j)(\hat{Y}_i - Y_i) + O_k \quad (3)$$

Here, O_k represents the terms after second order in Taylor series method.

Here, the term $R_k(Y_k, \alpha)$ does not change the value of mean square error dramatically; but $R_k(Y_k, \alpha)$ is important to calculate the bias. In simple random sampling, second order terms, in other words, the term $R_k(Y_k, \alpha)$ is frequently neglected when the sample size is large enough. Therefore, the mean square error can be obtained by Taylor Series Method as follows;

$$h(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k) \cong h(Y_1, Y_2, \dots, Y_k) + \sum_{j=1}^k d_j (\hat{Y}_j - Y_j) \quad (4)$$

By using Equation (4), mean square error is obtained as follows;

$$E[h(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k) - h(Y_1, Y_2, \dots, Y_k)]^2 \cong E[\sum_{j=1}^k d_j (\hat{Y}_j - Y_j)]^2 \quad (5)$$

$$MSE[h(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k)] \cong E[\sum_{j=1}^k d_j (\hat{Y}_j - Y_j)]^2 \quad (6)$$

3. Mean Square Error of the Proposed Estimators

3.1. Simple Conventional Estimation and Mean Square Error

When n -width sample is taken from N -width mass via simple random sampling method, variance of \bar{y} can be obtained as follows; Variance of mean estimator,

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \frac{(N-n)}{N} \frac{S_y^2}{n} \quad (7)$$

Where $f = \frac{n}{N}$ ve $\lambda = \frac{1-f}{n}$ is taken, Equation (8) becomes

$$V(\bar{y}) = \frac{(N-n)}{N} \frac{S_y^2}{n} = \frac{1-f}{n} S_y^2 = \lambda S_y^2 \quad (8)$$

$$S_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1} [1]$$

3.2. Conventional Propotional Estimation and Mean Square Error

Conventional ratio estimation for (\bar{Y}) mean mass of variable of interest y is

$$\bar{y}_o = \frac{\bar{y}}{\bar{x}} \bar{X} = \hat{R} \bar{X} \quad (9)$$

Here it is equal to $\hat{R} = \frac{\bar{y}}{\bar{x}}$ and \bar{y} is the sample mean of the variable of interest and \bar{x} is the sample mean of the auxiliary variable. We assume that mass mean of auxiliary variable x is known,

$$\bar{x} = \sum_{i=1}^n x_i \text{ ve } \bar{y} = \sum_{i=1}^n y_i \quad (10)$$

where n is the number of sample units.

Mean square error of this proposed estimator was obtained as

$$MSE(\bar{y}_o) \cong \lambda (S_y^2 - 2R\rho_{yx}S_yS_x + R^2S_x^2) = \lambda (S_y^2 + R^2S_x^2(1 - 2\theta)) \quad (11)$$

Where $\rho = \frac{C_y}{C_x}$, C_y and C_x are the variable of interest and coefficient of variation of the auxiliary variable respectively. $\lambda = \frac{1-f}{n}$ is the sampling fraction f and it is obtained by division of sample width to the mass width. In other words, $\lambda = \frac{n}{N}$; is the number of N mass units. $R = \frac{\bar{y}}{\bar{x}}$ is the mass ratio; S_y^2 and S_x^2 are the mass variance of the variable of interest and auxiliary variable respectively [2].

3.3. Investigation of Singh and Tailor Ratio Estimation and Mean Square Error

If the mass correlation coefficient of variables of interest and auxiliary variable (ρ) is known, ratio estimate proposed by Singh-Tailor can be defined as follows;

$$\bar{y}_{ST} = \frac{\bar{y}}{\bar{x} + \rho} (\bar{X} + \rho) \quad (12)$$

Mean square error of this estimator is expressed by

$$MSE(\bar{y}_{ST}) \cong \lambda (S_y^2 + R^2S_x^2w(w - 2\theta)) \quad (13)$$

where $w = \frac{\bar{x}}{\bar{x} + \rho}$ ve $\theta = \frac{\rho C_y}{C_x}$ [3].

3.4. Some Estimators Proposed by Kadilar-Çingı and Mean Square Error

Kadilar and Çingı (2004) proposed the following ratio estimators.

$$\bar{y}_{K\zeta 1} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X} \quad (14)$$

$$\bar{y}_{K\zeta 3} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + C_x} [\bar{X} + C_x] \quad (15)$$

$\hat{R}_{K\zeta 1} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}}$ and $b = \frac{s_{yx}}{s_x^2}$ represents the sample variance of auxiliary variable and s_{yx} is the sample covariance between variables of interest y and auxiliary variable x . General form of this expression was proposed by Ray and Singh (1981) [4]. Proportional estimator proposed by Ray and Singh is as follows;

$$\bar{y}_{RS} = \frac{\bar{y} + b(\bar{X}^\alpha - \bar{x}^\alpha)}{\bar{x}^\xi} \bar{X}^\xi \quad (16)$$

In this estimator, if $b = 0$ and $\xi = 1$, it becomes conventional proportional estimation. In addition, in the case of $\alpha = 1$ and $\xi = 1$, proportional estimator proposed by Ray and Singh corresponds to that proposed by Kadilar and Çingü. Mean square error of Kadilar and Çingü estimator is expressed as follows;

$$MSE(\bar{y}_{K\zeta i}) = \lambda[(B + R_{K\zeta i})^2 S_x^2 + S_y^2 - 2(B + R_{K\zeta i})S_{xy}] \quad (17)$$

Here, if $B = \frac{S_{yx}}{S_x^2}$ and is squared

It is expressed as

$$MSE(\bar{y}_{K\zeta i}) \cong \lambda[(R_{K\zeta i}^2 S_x^2 + S_y^2(1 - \rho^2)] ; i = 1, 3 \quad (18)$$

where $R_{K\zeta 1} = R = \frac{\bar{y}}{\bar{x}}$, $R_{K\zeta 3} = \frac{\bar{y}}{\bar{x} + C_x}$ [5]

Inference: In Equation (17), if $B = -R_{K\zeta i}$, mean square error of $\bar{y}_{K\zeta i}$ estimator is equal to the \bar{y} variance of simple estimator.

If $B = -R_{K\zeta i}$,

$$MSE(\bar{y}_{K\zeta i}) = \lambda S_y^2 = V(\bar{y})$$

Kadilar and Çingü (2006) developed a new ratio estimator by utilizing regression estimation rather than simple estimation in Singh and Taylor Estimator and using coefficient of mass correlation and variation coefficient of auxiliary variable. This estimator is as follows;

$$\bar{y}_{pr1} = \frac{\bar{y} + b(\bar{x} - \bar{x})}{\bar{x} + \rho} (\bar{X} + \rho) \quad (19)$$

$$\bar{y}_{pr2} = \frac{\bar{y} + b(\bar{x} - \bar{x})}{\bar{x}C_x + \rho} (\bar{X}C_x + \rho) \quad (20)$$

$$\bar{y}_{pr3} = \frac{\bar{y} + b(\bar{x} - \bar{x})}{\bar{x}\rho + C_x} (\bar{X}\rho + C_x) \quad (21)$$

Mean square error of this estimator can be expressed as follows

$$MSE(\bar{y}_{pri}) \cong \lambda[R_{pri}^2 S_x^2 + S_y^2(1 - \rho^2)] ; i = 1, 2, 3 \quad (22)$$

where $R_{pr1} = \frac{\bar{y}}{\bar{x} + \rho}$, $R_{pr2} = \frac{\bar{y}}{\bar{x}C_x + \rho}$, $R_{pr3} = \frac{\bar{y}\rho}{\bar{x}\rho + C_x}$ [6].

3.5. Some Ratio Estimators that have Certain Feature which were Developed in Simple Random Sampling Using Auxiliary Information and Mean Square Error

Here, R in ratio estimation is replaced by B in regression estimation. The use of auxiliary variables can improve the efficiency of estimation in the case of where auxiliary variables and the variable of interest are highly correlated. Suppose that n-width sample is taken from n-width mass (population) via simple random sampling without replacement. y_i and ϕ_i represent the observation values in the variables for $i = 1, 2, \dots, N$ respectively. Suppose that

$$\phi_i = \begin{cases} 1, & \text{if the desire feauture is achieved} \\ 0, & \text{other cases} \end{cases} \quad (23)$$

By using this Equation, $P = \frac{A}{N}$ and $p = \frac{a}{n}$ mass and sample ratio estimations and ratio equation of the desired events can be obtained. Here, $A = \sum_{i=1}^N \phi_i$ and $a = \sum_{i=1}^n \phi_i$ show the total number of mass and sampling units respectively [1].

3.5.1. Investigation of Naik and Gupta ratio Estimators and Mean Square Error

$$\bar{y}_{NG} = \frac{\bar{y}}{p} P \quad (24)$$

where \bar{y} is the sample mean of the variable of interest. p is the unit ratio having certain features in the sampling. Mean square error of this estimation is

$$MSE(\bar{y}_{NG}) \cong \lambda(R_{NG}^2 S_\phi^2 - 2R_{NG}S_{y\phi} + S_y^2)$$

where $S_\phi^2 = \frac{1}{N-1} \sum_{i=1}^N (\phi_i - P)^2$ and $S_{y\phi} = \frac{1}{N-1} \sum_{i=1}^N (\phi_i - P)(y_i - \bar{y})$ [7].

3.5.2. The Proposed Singh, Chaucer, Sawan and Smarandache Estimates and Mean Square Error Error

Estimations following Ray and Singh Ratio Estimations are

$$t_1 = \frac{\bar{y} + b_\phi(P-p)}{p} P \quad (25)$$

$$t_2 = \frac{\bar{y} + b_\phi(P-p)}{p + C_p} [P + C_p] \quad (26)$$

$$t_3 = \frac{\bar{y} + b_\phi(P-p)}{p + \rho_o} [P + \rho_o] \quad (27)$$

In Equation (25), if $b_\phi = 0$, the proposed estimator becomes Naik and Gupta estimation in Equation (24). Mean square error of this estimator is

$$MSE(t_i) \cong \lambda[R_i^2 S_\phi^2 + S_y^2(1 - \rho_o^2)] ; i = 1, 2, 3$$

$$R_1 = \frac{\bar{y}}{p}, \quad R_2 = \frac{\bar{y}}{p + C_p}, \quad R_3 = \frac{\bar{y}}{p + \rho_o}$$

where $S_\phi^2 = \frac{1}{N-1} \sum_{i=1}^N (\phi_i - P)^2$ and $S_{y\phi} = \frac{1}{N-1} \sum_{i=1}^N (\phi_i - P)(y_i - \bar{y})$ [8].

4. Theoretical Comparisons

4.1. Comparison of Singh and Tailor Estimator and Conventional Proportional Estimation

Conventional proportional estimator;

$$MSE(\bar{y}_o) \cong \lambda(R^2 S_x^2 - 2R\rho_{yx}S_yS_x + S_y^2) = \lambda(S_y^2 + R^2 S_x^2(1 - \theta)) \quad (28)$$

We have seen conventional proportional estimator in Equation (28). Conventional proportional estimation is obtained by inequality of mean square error of Singh and Tailor Estimator

$$\lambda(S_y^2 + R^2 S_x^2 w(w - 2\theta)) < \lambda(S_y^2 + R^2 S_x^2 (1 - 2\theta))$$

$$w(w - 2\theta) < 1 - 2\theta$$

$$w^2 - 1 - 2w\theta + 2\theta < 0$$

$$(w - 1)(w + 1) - 2\theta(w - 1) < 0$$

$$(w - 1)(w + 1 - 2\theta) < 0$$

This inequality is satisfied when

$$w - 1 < 0 \text{ ve } w + 1 - 2\theta > 0 \quad (29)$$

or

$$w - 1 > 0 \text{ ve } w + 1 - 2\theta < 0 \quad (30)$$

Under these conditions, mean square error of Singh and Taylor Estimator is less than that of conventional proportional estimator. Therefore, it is more effective.

4.2. Comparison of Singh and Tailor Estimator and Simple Conventional Estimation

$$\lambda(S_y^2 + R^2 S_x^2 w(w - 2\theta)) < \lambda S_y^2$$

$$R^2 S_x^2 w^2 < R^2 S_x^2 w 2\theta$$

$$\frac{w}{\theta} < 2 \quad (31)$$

Under condition (31) Singh-Tailor estimation is more effective than simple conventional estimation.

4.3. Comparison of Kadılar-Çingı (2004) Estimation and Conventional Proportional Estimation

From Equation (11) and (18); Under the condition

$$\lambda[(R_{K\zeta i}^2 S_x^2 + S_y^2 (1 - \rho^2)) < \lambda(S_y^2 - 2R\rho S_y S_x + R^2 S_x^2)$$

$$(R_{K\zeta i}^2 - R^2) S_x^2 < \rho^2 S_y^2 - 2R\rho S_y S_x)$$

$$(R_{K\zeta i}^2 - R^2) S_x^2 < \rho S_y (\rho S_y - 2R S_x) \quad (32)$$

Kadılar-Çingı estimation is more effective.

For example, consider the case $i = 1$

$$\text{If } R_{K\zeta 1} = R = \frac{\bar{y}}{\bar{x}}$$

Under the condition

$$0 < \rho S_y (\rho S_y - 2R S_x) \quad 2R S_x < \rho S_y$$

$$\rho > 2 \frac{c_x}{c_y} \quad (33)$$

Kadılar-Çingı estimation is more effective.

4.4. Comparison of the Proposed Singh, Chaucer, Sawan, Smarandache Estimations and Simple Estimate

Comparison of the proposed estimators and simple estimates and variance of the sample mean was examined in simple random sampling as follows;

$$V(\bar{y}) = \lambda S_y^2 \quad (34)$$

$$V(\bar{y}) - MSE(t_i) \geq 0 \quad ; i = 1, 2, 3 \quad (35)$$

$$\lambda S_y^2 > \lambda [R_i^2 S_\phi^2 + S_y^2 (1 - \rho_o^2)] \quad (36)$$

$$\rho_o^2 > \frac{S_\phi^2}{S_y^2} R_i^2 \quad (37)$$

condition is obtained. Under this condition, the proposed estimator is more efficient than simple estimates [8].

4.5. Comparison of the Proposed Estimators with Each other

In the case of

$$\lambda[R_1^2 S_\phi^2 + S_y^2 (1 - \rho_o^2)] < \lambda[R_j^2 S_\phi^2 + S_y^2 (1 - \rho_o^2)]; \quad j = 2, 3 \quad (38)$$

$$R_1^2 S_\phi^2 < R_j^2 S_\phi^2$$

$$R_1^2 S_\phi^2 - R_j^2 S_\phi^2 < 0$$

$$S_\phi^2 (R_1^2 - R_j^2) < 0$$

$$|R_1| < |R_j| \quad (39)$$

t_1 estimate is more effective than t_2 and t_3 .

4.6. Comparison of the Proposed Singh, Chaucer, Sawan, Smarandache Estimates and Naik and Gupta Estimations

With the proposed estimation, comparison of Naik and Gupta estimate is achieved under the condition of

$$MSE(\bar{y}_{NG}) - MSE(t_i) > 0 \quad ; i = 1, 2, 3 \quad (40)$$

$$\rho_o^2 > \frac{S_\phi^2}{S_y^2} \left[R_i^2 - R_\phi^2 + 2R_\phi \rho_o \frac{c_y}{c_p} \right] \quad (41)$$

Under this condition, mean square error of the proposed estimator is less and therefore more effective [8].

4.7. Comparison of Naik and Gupta Estimate and Simple Estimation

$$\lambda S_y^2 > \lambda (R_{NG}^2 S_\phi^2 - 2R_{NG} S_y S_\phi + S_y^2) \quad (42)$$

$$2\rho_o S_y S_\phi > R_{NG} S_\phi^2 \quad (43)$$

$$\rho_o > \frac{R_{NG} S_\phi}{2S_y} \quad (44)$$

condition is obtained.

Under this condition, we can conclude that mean square error of Naik and Gupta estimation is smaller and therefore more effective.

5. Application

School ranking of 18 districts in Trabzon province in the 2011-2012 academic years was used (Trabzon MEM, 2011-2012) [9]. The data from 18 districts of Trabzon province

included the number of students and teachers from secondary schools in the 2011-2012 academic years. Information regarding the number of teachers and students was obtained from Trabzon provincial Directorate of National Education.

Main variable (y) is taken as the number of teachers and auxiliary variable (x) is taken as the number of students.

The relationship between the number of teachers and the number of students can be demonstrated by a linear equation ($y = bx$) passing through the point of origin and the related variables can be estimated using auxiliary variables. Mass unit is the 111 secondary schools in Trabzon province. Data regarding the mass is given in Table 5.1.

Table 5.1. Mass units regarding the number of teachers (y) and the number of students (x) variables.

N = 111	$\bar{Y} = 29.28$	$\bar{X} = 395.06$	$\rho = 0.939$
$R = \frac{\bar{Y}}{\bar{X}} = 0.07$	$S_y^2 = 651.312$	$S_y = 25.521$	$C_y = \frac{S_y}{\bar{Y}} = 0.872$
$S_x^2 = 159948.937$	$S_x = 399.936$	$C_x = \frac{S_x}{\bar{X}} = 1.012$	$C_{yx} = \frac{S_{xy}}{\bar{X}\bar{Y}} = 0.829$

Samples are obtained from a mass via simple random sampling without replacement using SPSS-20 program. First, sample size is calculated.

$$n = \frac{t^2 NPQ}{d^2(N-1) + t^2 PQ} \quad (45)$$

can be used in the estimation of sample size [10].

In this equation, P ratio of units having the same features is ($Q = 1 - P$). If we do not have information about units ratio which has the same feature, the largest sample that can be selected for $P = Q = 0.5$ could be reached. However, unit ratio having specific features is calculated as the ratio

of the schools in the city. where d is the amount of tolerance. Considering the above equation (45), sample size reduces when the the amount of tolerance increases. The amount of tolerance (d) depends on the researchers. In the present application, if $d = 0.09$ sample size is obtained as

$$n = \frac{1.96^2 * 111 * 0.81 * 0.19}{0.09^2 * 110 + 1.96^2 * 0.81 * 0.19} \cong 45 \quad (46)$$

The statistics and the data regarding the schools location in the city are given in Table 5.2 and Table 5.3 respectively.

Table 5.2. Sample units of the number of teachers (y) and the number of students (x) variable.

n = 45	$\bar{y} = 32.02$	$\bar{x} = 413.82$	$r = 0.955$
$\hat{R} = \frac{\bar{y}}{\bar{x}} = 0.077$	$S_y^2 = 875.522$	$S_y = 29.589$	$f = \frac{n}{N} = 0.405$
$S_x^2 = 183282.5586$	$S_x = 428.115$	$\lambda = \frac{1-f}{n} = 0.013$	$b = \frac{S_{yx}}{S_x^2} = 0.066$

Table 5.3. Mass and sample data required for unit ratio having certain features.

A = 90	a = 39	P = 0.810	$\rho_o = \frac{S_{y\phi}}{S_y S_\phi} = 0.216$
$b_\phi = \frac{S_{y\phi}}{S_\phi^2} = 14.006$	$S_y^2 = 651.312$	$S_y = 25.521$	$C_y = \frac{S_y}{\bar{Y}} = 0.872$
$S_\phi^2 = 0.155$	$S_\phi = 0.394$	$C_p = \frac{S_\phi}{P} = 0.486$	$C_{yp} = \frac{S_{y\phi}}{P\bar{Y}} = 0.092$

Table 5.4. Mean Ratio Estimators and Mean Square Error

Mean Estimators	Mean Values	Mean Square Error
Simple Estimation (\bar{y})	32.02	8.467
Conventional Proposional Estimation (\bar{y}_o)	30.568	2.168
Singh-Tailor Estimation (\bar{y}_{ST})	30.572	2.163
Kadilar-Çingir-1 Estimation (\bar{y}_{KC1})	29.386	11.190
Kadilar-Çingir-3 Estimation (\bar{y}_{KC3})	29.389	12.388
Kadilar-Çingir Estimation (\bar{y}_{pr1})	29.389	12.357
Kadilar-Çingir Estimation (\bar{y}_{pr2})	29.389	12.082
Kadilar-Çingir Estimation (\bar{y}_{pr3})	29.390	13.880

According to the mean ratio estimators, the average number of teachers per secondary schools in Trabzon

province was 30. In addition, Singh- Tailor estimation (\bar{y}_{ST}) showed the highest sensitivity.

Table 5.5. Ratio Estimators having certain features and Mean Square Error.

Estimators of units ratio having certain features	Estimates	Mean square error
Naik-Gupta Estimate (\bar{y}_{NG})	29.949	9.058
Singh-Chaucan-Sawan-Smarandache 1 Estimate (t_1)	29.216	10.705
Singh-Chaucan-Sawan-Smarandache 2 Estimate (t_2)	29.940	9.10
Singh-Chaucan-Sawan-Smarandache 3 Estimate (t_3)	29.619	9.713

According to ratio estimators having certain units feature, the average number of teacher per secondary schools in Trabzon was 29. Naik-Gupta Estimate (\bar{y}_{NG}) estimator was the most efficient estimator with the minimum mean square error.

Comparison of Singh and Tailor Estimataion with Conventional Proportional Estimation

$$MSE(\bar{y}_{ST}) \cong \lambda(S_y^2 + R^2 S_x^2 w(w - 2\theta)) \cong 2.163$$

$$\text{where } w = \frac{\bar{x}}{\bar{x} + p} = 0.998 \text{ and } \theta = \frac{p C_y}{C_x} = 0.809$$

To ensure inequality $MSE(\bar{y}_{ST}) < MSE(\bar{y}_O)$ $w - 1 < 0$ and $w + 1 - 2\theta > 0$ or $w - 1 > 0$ and $w + 1 - 2\theta < 0$ condition should be fulfilled. Thus, since $w - 1 = 0.998 - 1 < 0$ and $w + 1 - 2\theta = 0.998 + 1 - 2 * 0.809 = 1.967 > 0$ condition is satisfied,

$$MSE(\bar{y}_{ST}) \cong 2.163 < MSE(\bar{y}_O) \cong 2.168$$

mean square error of Singh and Tailor Estimation is smaller and therefore more effective.

Comparison of Singh and Tailor Estimation and Simple Estimation

The variance of simple estimation is $V(\bar{y}) = \lambda S_y^2 = 8.467$

$MSE(\bar{y}_{ST}) < MSE(\bar{y})$ inequality is valid when the condition $\frac{w}{\theta} = \frac{0.998}{0.809} = 1.234 < 2$ is satisfied. Since this condition is fulfilled, the Tailor Singh estimation is more effective than simple classical estimation.

Comparison of Kadilar-Çingı-1 Estimation and Proportional Estimation

Under the condition $MSE(\bar{y}_{KÇ1}) \cong \lambda[(R^2 S_x^2 + S_y^2) - p^2] = 11.190$ $p > 2C_x C_y$, Kadilar

Çingı1 estimation is more effective than conventional proportional estimate. However, condition is not satisfied, so conventional proportional estimate is more efficient.

Comparison of Naik-Gupta Estimation and Simple Estimation

Mean square error is found as $MSE(\bar{y}_{NG}) \cong \lambda(R_{NG}^2 S_{\phi}^2 - 2R_{NG} S_{y\phi} + S_y^2) = 9.058$. When compared with simple estimates, if the condition $\rho_o > \frac{R_{NG} S_{\phi}}{2S_y}$ is satisfied, Naik and Gupta Estimation is smaller and therefore more effective than simple estimation. However, since $\rho_o = 0.216 < \frac{R_{NG} S_{\phi}}{2S_y}$ condition is not satisfied, error of simple estimate is smaller and therefore more effective.

Comparison of the Proposed t_1 Estimation and Simple Estimation

To ensure $V(\bar{y}) - MSE(t_i) \geq 0$ inequality, $\rho_o^2 > \frac{S_{\phi}^2}{S_y^2} R_i^2$ condition must be fulfilled.

For this condition $i = 1$;

$$\rho_o^2 = 0.216^2 = 0.046 < \frac{S_{\phi}^2}{S_y^2} R_1^2 = \frac{0.155 * (36.148)^2}{651.312} = 0.311$$

Since $\rho_o^2 > \frac{S_{\phi}^2}{S_y^2} R_1^2$ condition is not satisfied, error of simple estimation is smaller and therefore more effective.

For $i = 2$;

$$\rho_o^2 = 0.216^2 = 0.046 < \frac{S_{\phi}^2}{S_y^2} R_2^2 = \frac{0.155 * (22.593)^2}{651.312} =$$

0.121. Therefore, simple estimation is more effective. In the same way, simple estimate is more effective for $i = 3$ case.

Comparison of the proposed t_1 estimation and t_2 estimation

$$\text{Since } R_1 = \frac{Y}{P} = 36.15 > R_2 = \frac{Y}{P + C_p} = 22.59,$$

$MSE(t_2) = 9.10 < MSE(t_1) = 10.71$ error of t_2 estimation is smaller and therefore more sensitive.

6. Conclusion

When $B = -R_{KÇi}$ in Kadilar-Çingı (2004) and $B = -R_{pri}$ in Kadilar-Çingı (2006) estimation, mean square error of Kadilar-Çingı (2004) and Kadilar-Çingı (2006) estimates are equal to that of simple estimation.

When comparing Singh and Tailor Estimator with the conventional proportional estimators, under the condition $w - 1 < 0$ and $w + 1 - 2\theta > 0$ or $w - 1 > 0$ and $w + 1 - 2\theta < 0$ Singh-Tailor estimation is more effective than conventional proportional estimation.

In our numerical example, since $w - 1 = 0.998 - 1 < 0$ and $w + 1 - 2\theta = 0.998 + 1 - 2 * 0.809 = 1.967 > 0$ condition is satisfied, $MSE(\bar{y}_{ST}) < MSE(\bar{y}_O)$ inequality is obtained.

In the comparison of Singh and Tailor Estimators with the simple estimate, $\frac{w}{\theta} < 2$ condition is satisfied and under this condition, Singh-Tailor Estimator is more effective than simple estimator. In our applications, $\frac{w}{\theta} = \frac{0.998}{0.809} = 1.234 < 2$ condition is fulfilled and Singh-Tailor Estimator was found to be more sensitive.

In the ratio of units having certain features, estimators proposed by Singh et al., t_1 , t_2 and t_3 estimators were compared and $R_1 < R_j$ ($j = 2, 3$) condition was obtained. Under this condition, the proposed t_1 estimation is more effective than t_2 and t_3 estimations. In our application,

$R_1 = \frac{Y}{P} = 36.15 > R_2 = \frac{Y}{P + C_p} = 22.59$ inequality is valid and expressed as $MSE(t_2) < MSE(t_1)$.

Considering the numerical examples, according to the

estimators proposed in the literature, the average number of teachers per secondary schools in Trabzon province in 2011-2012 academic years ranged from 29 to 32. Singh-Tailor estimation (\bar{y}_{ST}), is the mean ratio estimator having the smallest root mean square error (2,163) whereas the estimator proposed by Kadilar and Çingı (\bar{y}_{pr3}) has the greatest mean square error (13,880). Among the estimators having certain features, Naik-Gupta estimator (9.058) has the smallest mean square error.

The numerical results obtained are valid only for this application.

References

- [1] Pawliszyn, Janusz., 2012. Comprehensive Sampling and Sample Preparation, 1-30.
- [2] Cochran, W. G., 1977. *Sampling Techniques*, John Wiley and Sons, New-York.
- [3] Singh, H. P., Tailor, R., 2003. *Use of Known Correlation Coefficient in Estimating the Finite Population Mean*, *Statistics in Transition*, 6, 555-560.
- [4] Ray, S. K., Singh, R. K., 1981. *Difference-cum-ratio type estimators*, *J. Ind. Statist. Assoc.*, 19, 147-151.
- [5] Kadilar, C., Çingı H., 2004. *Ratio Estimators in Simple Random Sampling*, *Applied Mathematics and Computation*, 151, 893-902.
- [6] Kadilar, C., Çingı, H., 2006. *New Ratio Estimators Using Correlation Coefficient*, *İnterstat* 4, pp, 1-11.
- [7] Naik, G., V. D., Gupta, P. C., 1996. *A note on estimation of mean with known population proportion of an auxiliary character*. *Jour. Ind. Soc. Agr. Stat.*, 48(2), 151-158.
- [8] Singh, R., Chauhan, P., Sawan, N., Smarandache, F., 2008. *Ratio Estimators in Simple Random Sampling Using Information on Auxiliary Attribute*, *Pak. j. stat. oper. res.* Vol.IV No. 1 2008.
- [9] Number of teachers and students in Secondary Schools, 2012. Trabzon Provincial Directorate of National Education.
- [10] Çingı, H., 2009. *Örnekleme Kuramı*, H.Ü. Fen Fakültesi Basımevi, Beytepe.