STATISTICS IN TRANSITION new series, March 2020
Vol. 21, No. 1 pp. 159–168, DOI 10.21307/stattrans-2020-009
Submitted – 17.10.2019; accepted for publishing – 30.01.2020

Generalized exponential estimators for the finite population mean

Tolga Zaman[1]

ABSTRACT

This study proposes a new class of exponential-type estimators in simple random sampling for the estimation of the population mean of the study variable using information of the population proportion possessing certain attributes. Theoretically, mean squared error (MSE) equations of the suggested ratio exponential estimators are obtained and compared with the Naik and Gupta (1996) ratio and product estimators, the ratio and product exponential estimator presented in Singh et al. (2007) and the ratio exponential estimators presented in Zaman and Kadilar (2019a). As a result of these comparisons, it is observed that the proposed estimators always produce more efficient results than the others. In addition, these theoretical results are supported by the application of original datasets.

Key words: ratio-exponential estimators, auxiliary attribute, mean square error, efficiency.

Mathematical classification: 62D05

1. Introduction

When there is a positive correlation between the study variable and the auxiliary variable in the simple random sampling method, ratio-type estimators are used to estimate the population mean. But when the correlation coefficient between the study and auxiliary variables is negative, the product estimator is used. Many authors have proposed estimators based on auxiliary attribute. Bahl and Tuteja (1991) estimator and a family of estimators considered by Jhajj et al. (2006), Singh et al. (2008), Koyuncu (2012), Malik and Singh (2013), Shabbir and Gupta (2010), Solanki and Singh (2013), Zaman (2018), Zaman and Kadilar (2019b) suggested a class of estimators by using auxiliary attributes. In this paper, a new class of exponential type estimators is proposed to estimate the population mean for the study variable using information on auxiliary attributes.

[1] Çankırı Karatekin University, Faculty of Science, Department of Statistics, 18100 Çankırı, Turkey.
E-mail: tolgazaman@karatekin.edu.tr. ORCID: https://orcid.org/0000-0001-8780-3655.

Let y_i be i-th characteristic of the population and ϕ_i is the case of possessing certain attributes. If ith unit has the desired characteristic, it takes the value 1; if not then the value 0. That is:

$$\phi_i = \begin{cases} 1 \quad , & if\ ith\ unit\ of\ the\ population\ possesses\ attribute \\ 0 \quad , & otherwise. \end{cases}$$

Let $A = \sum_{i=1}^{N} \phi_i$ and $a = \sum_{i=1}^{n} \phi_i$ be the total count of the units that possess certain attribute in the population and the sample, respectively. And $P = \frac{A}{N}$ and $p = \frac{a}{n}$ are the ratio of these units, respectively.

The Naik and Gupta (1996) estimator for the population mean $\bar{Y}$ of the variate of study, which makes use of information concerning the population proportion possessing certain attribute, is defined by

When the relationship between the study variable and the auxiliary attribute is positive, the ratio estimator of the population proportion of the study variable,

$$\bar{y}_{NG1} = \frac{\bar{y}}{p} P \tag{1.1}$$

When the relationship between the study variable and the auxiliary attribute is negative, the product estimator of the population proportion of the study variable,

$$\bar{y}_{NG2} = \frac{\bar{y}}{P} p \tag{1.2}$$

where it is assumed that the population proportion P of the form of attribute ϕ is known.

The expressions of MSE of the Naik and Gupta estimators are

$$MSE(\bar{y}_{NG1}) \cong \frac{1-f}{n} \bar{Y}^2 \left(C_y^2 - 2\rho_{pb} C_y C_p + C_p^2\right) \tag{1.3}$$

$$MSE(\bar{y}_{NG2}) \cong \frac{1-f}{n} \bar{Y}^2 \left(C_y^2 + 2\rho_{pb} C_y C_p + C_p^2\right) \tag{1.4}$$

Considering Bahl and Tuteja (1991), Singh et al. (2007) suggested the following ratio estimator when the study variable and the auxiliary attribute are positively correlated.

$$t_1 = \bar{y} exp\left(\frac{P-p}{P+p}\right) \tag{1.5}$$

Singh et al. (2007) suggested the following ratio estimator when the study variable and the auxiliary attribute are negatively correlated.

$$t_2 = \bar{y} exp\left(\frac{p-P}{p+P}\right) \tag{1.6}$$

The MSEs of these estimators are

$$MSE(t_1) \cong \frac{1-f}{n}\bar{Y}^2\left(C_y^2 - \rho_{pb}C_yC_p + \frac{C_p^2}{4}\right) \quad (1.7)$$

$$MSE(t_2) \cong \frac{1-f}{n}\bar{Y}^2\left(C_y^2 + \rho_{pb}C_yC_p + \frac{C_p^2}{4}\right) \quad (1.8)$$

Zaman and Kadilar (2019a) suggested modified exponential ratio estimators using auxiliary attribute information for estimating $\bar{Y}$ as

$$t_{ZKi} = \bar{y}exp\left[\frac{(kP+l)-(kp+l)}{(kP+l)+(kp+l)}\right] \quad i = 1,2,\ldots,9 \quad (1.9)$$

where $k(\neq 0)$ and l are either the real number or the functions of the known parameters of the attribute, C_p, $\beta_2(\phi)$ and the known parameter of the attribute with the study variable, ρ_{pb}. Note that the sum of k and l is not necessarily equal to one.

The MSE of these estimators

$$MSE(t_{ZKi}) \cong \frac{1-f}{n}\bar{Y}^2\left[\theta_i^2C_p^2 - 2\theta_i\rho_{pb}C_yC_p + C_y^2\right] \; ,i = 1,\ldots,9 \quad (1.10)$$

$$\theta_1 = \frac{P}{2(P+\beta_2(\phi))};\; \theta_2 = \frac{P}{2(P+C_p)};\; \theta_3 = \frac{P}{2(P+\rho_{pb})};\; \theta_4 = \frac{\beta_2(\phi)P}{2(\beta_2(\phi)P+C_p)};\; \theta_5 = \frac{C_pP}{2\left(C_pP+\beta_2(\phi)\right)}$$

$$\theta_6 = \frac{C_pP}{2(C_pP+\rho_{pb})};\; \theta_7 = \frac{\rho_{pb}P}{2(\rho_{pb}P+C_p)};\; \theta_8 = \frac{\beta_2(\phi)P}{2(\beta_2(\phi)P+\rho_{pb})};\; \theta_9 = \frac{\rho_{pb}P}{2\left(\rho_{pb}P+\beta_2(\phi)\right)}.$$

where $f = \frac{n}{N}$; N is the number of units in the population; C_p is the coefficient of population variation of the form of attribute and C_y is the coefficient of population variation of the study variable. ρ_{pb} is the point biserial correlation coefficient. $\beta_2(\phi)$ is the coefficient of population kurtosis of the auxiliary attribute.

2. Suggested estimators

Following Ozel (2016), the improved class of estimator $\bar{y}_{pr}$ for the population is proposed as follows

$$\bar{y}_{pri} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{(kP+l)-(kp+l)}{(kP+l)+(kp+l)}\right]; i = 1,2,\ldots,10 \quad (2.1)$$

A class of new estimators generated from Equation (2.1) is listed in Table 1.

Table 1. Proposed Estimators

Estimators	Values of	
	k	l
$\bar{y}_{pr1} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{P-p}{P+p}\right]$	1	0
$\bar{y}_{pr2} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{P-p}{P+p+2\beta_2(\phi)}\right]$	1	$\beta_2(\phi)$
$\bar{y}_{pr3} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{P-p}{P+p+2C_p}\right]$	1	C_p
$\bar{y}_{pr4} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{P-p}{P+p+2\rho_{pb}}\right]$	1	ρ_{pb}
$\bar{y}_{pr5} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{\beta_2(\phi)(P-p)}{\beta_2(\phi)(P+p)+2C_p}\right]$	$\beta_2(\phi)$	C_p
$\bar{y}_{pr6} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{C_p(P-p)}{C_p(P+p)+2\beta_2(\phi)}\right]$	C_p	$\beta_2(\phi)$
$\bar{y}_{pr7} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{C_p(P-p)}{C_p(P+p)+2\rho_{pb}}\right]$	C_p	ρ_{pb}
$\bar{y}_{pr8} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{\rho_{pb}(P-p)}{\rho_{pb}(P+p)+2C_p}\right]$	ρ_{pb}	C_p
$\bar{y}_{pr9} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{\beta_2(\phi)(P-p)}{\beta_2(\phi)(P+p)+2\rho_{pb}}\right]$	$\beta_2(\phi)$	ρ_{pb}
$\bar{y}_{pr10} = \bar{y}\left(\frac{p}{P}\right)^{\alpha} exp\left[\frac{\rho_{pb}(P-p)}{\rho_{pb}(P+p)+2\beta_2(\phi)}\right]$	ρ_{pb}	$\beta_2(\phi)$

In Table 1, C_p, $\beta_2(\phi)$ and ρ_{pb} are coefficient of variation, coefficient of population kurtosis of the form of the auxiliary attribute and the point biserial population correlation coefficient between the auxiliary attribute and the study variable, respectively. $\bar{y}$ and p are the sample mean belonging to the study variable and the sample proportion possessing certain attributes, respectively.

To obtain the MSE expression of these estimators $\bar{y}_{pri}$, let $\bar{y} = \bar{Y}(1+e_0)$ and $p = P(1+e_1)$ such that $E(e_0) = E(e_1) = 0$ and $E(e_0^2) = \frac{1-f}{n}C_y^2$, $E(e_1^2) = \frac{1-f}{n}C_p^2$ and $E(e_0e_1) = \frac{1-f}{n}C_{yp} = \frac{1-f}{n}\rho_{pb}C_yC_p$.

Expressing the estimator $\bar{y}_{pri}$ in terms of e_j, $(j = 0, 1)$ and shortened terms up to the first two and the first three components of e's as follows

$$\bar{y}_{pri} = \bar{Y}(1+e_0)(1+e_1)^{\alpha} exp\{-\theta e_1(1+\theta e_1)^{-1}\}$$
$$\bar{y}_{pri} = \bar{Y}(1+e_0)\left[1+\alpha e_1+\frac{\alpha(\alpha-1)}{2}e_1^2+\cdots\right][1+(-\theta e_1)(1+\theta e_1)^{-1}]$$

$$= \bar{Y}(1+e_0)\left[1+\alpha e_1+\frac{\alpha(\alpha-1)}{2}e_1^2+\cdots\right][1-\theta e_1+\theta^2 e_1^2] \tag{2.2}$$

Expanding the right-hand side of (2.2) to the first order approximation and subtracting $\bar{Y}$ from both sides, the following expression is obtained

$$\bar{y}_{pri}-\bar{Y} \cong \bar{Y}\left[-\theta_i e_1+\theta_i^2 e_1^2+\alpha e_1-\alpha\theta_i e_1^2+\frac{\alpha(\alpha-1)}{2}e_1^2+e_0-\theta_i e_0 e_1+\alpha e_0 e_1\right] \tag{2.3}$$

Squaring both sides of equation (2.3) and then taking the expectation of both sides, the following expression of MSE of the estimator $\bar{y}_{pri}$ is obtained

$$MSE(\bar{y}_{pri}) \cong \frac{1-f}{n}\bar{Y}^2\left[\alpha^2 C_p^2+\theta_i^2 C_p^2-2\alpha\theta_i C_p^2+C_y^2+2\alpha\rho_{pb}C_yC_p-2\theta_i C_yC_p\right]; i = 1,2,\ldots,10 \tag{2.4}$$

Using the following equations, the optimal value of α can be obtained:

$$\frac{\partial MSE(\bar{y}_{pri})}{\partial\alpha} = \frac{1-f}{n}\bar{Y}^2\left(2\alpha C_p^2-2\theta_i C_p^2+2\rho_{pb}C_yC_p\right)=0$$
$$\alpha_{opti} = \frac{\theta_i C_p-\rho_{pb}C_y}{C_p}; i = 1,2,..,10 \tag{2.5}$$

The minimum MSE of the proposed estimator can be calculated using α equations in (2.5). The suggested estimators have the same minimum MSE as follows:

$$MSE_{min}(\bar{y}_{pr}) \cong \frac{1-f}{n}\bar{Y}^2 C_y^2(1-\rho_{pb}^2) \tag{2.6}$$

3. Efficiency comparisons

The expressions of MSE of classical estimators $\bar{y}$, $\bar{y}_{NG1}$, $\bar{y}_{NG2}$, t_1, t_2 and zt_i are compared with the MSE of the suggested estimator $\bar{y}_{pri}$.

Under simple random sampling without replacement (SRSWOR) the variance of the sample mean is

$$V(\bar{y}) = \frac{1-f}{n}\bar{Y}^2 C_y^2 \tag{3.1}$$

From (2.6) and (3.1), we have

$$MSE_{min}(\bar{y}_{pri}) < V(\bar{y})$$
$$\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{pb}^2) < \frac{1-f}{n}\bar{Y}^2C_y^2$$
$$\rho_{pb}^2 > 0 \tag{3.2}$$

From (2.6) and (1.3), we have

$$MSE_{min}(\bar{y}_{pri}) < MSE(\bar{y}_{NG1})$$
$$\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{pb}^2) < \frac{1-f}{n}\bar{Y}^2(C_y^2 - 2\rho_{pb}C_yC_p + C_p^2)$$
$$(C_y\rho_{pb} - C_p)^2 > 0 \tag{3.3}$$

From (2.6) and (1.4), we have

$$MSE_{min}(\bar{y}_{pri}) < MSE(\bar{y}_{NG2})$$
$$\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{pb}^2) < \frac{1-f}{n}\bar{Y}^2(C_y^2 + 2\rho_{pb}C_yC_p + C_p^2)$$
$$(C_y\rho_{pb} + C_p)^2 > 0 \tag{3.4}$$

From (2.6) and (1.7), we have

$$MSE_{min}(\bar{y}_{pri}) < MSE(t_1)$$
$$\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{pb}^2) < \frac{1-f}{n}\bar{Y}^2\left(C_y^2 - \rho_{pb}C_yC_p + \frac{C_p^2}{4}\right)$$
$$\left(C_y\rho_{pb} - \frac{C_p}{2}\right)^2 > 0 \tag{3.5}$$

From (2.6) and (1.8), we have

$$MSE_{min}(\bar{y}_{pri}) < MSE(t_2)$$
$$\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{pb}^2) < \frac{1-f}{n}\bar{Y}^2\left(C_y^2 + \rho_{pb}C_yC_p + \frac{C_p^2}{4}\right)$$
$$\left(C_y\rho_{pb} + \frac{C_p}{2}\right)^2 > 0 \tag{3.6}$$

From (2.6) and (1.10), we have

$$MSE_{min}(\bar{y}_{pri}) < MSE(t_{ZKi})$$
$$\frac{1-f}{n}\bar{Y}^2C_y^2(1-\rho_{pb}^2) < \frac{1-f}{n}\bar{Y}^2[\theta_i^2C_p^2 - 2\theta_i\rho_{pb}C_yC_p + C_y^2]$$
$$(C_y\rho_{pb} - \theta_iC_p)^2 > 0 \tag{3.7}$$

We infer that the suggested exponential estimators $\bar{y}_{pr}$ perform better than the competing estimators in all conditions because the condition given in (3.2)-(3.7) is always satisfied.

4. Empirical study

For empirical study, we use the data given in Sukhatme and Sukhatme (1970) and Zaman et al. (2014).

The values of the required parameters for these data sets are given in Tables 2 and 3.

Population 1: The data is defined as follows:

$$y = \text{the number of villages in the circles}$$
$$\phi_i = \begin{cases} 1 \quad , & \text{if a circle consists of more than five villages} \\ 0 \quad , & \text{otherwise.} \end{cases}$$

Table 2. Data Statistics of Population 1

N=89	$\bar{Y}$ = 3.3596	θ_1 =0.0171	θ_5 =0.0433	θ_9 =0.0132
n= 20	P = 0.1236	θ_2 =0.0221	θ_6 =0.1508	
$\beta_2(\phi)$ = 3.492	C_y = 0.6008	θ_3 =0.0695	θ_7 =0.0171	
ρ_{pb} = 0.766	C_p = 2.6779	θ_4 =0.0694	θ_8 =0.1802	

Population 2: The data is defined as follows:

$$y = \text{the number of teachers}$$
$$\phi_i = \begin{cases} 1 \quad , & \text{if the number of teachers is more than 60} \\ 0 \quad , & \text{otherwise.} \end{cases}$$

Table 3. Data Statistics of Population 2

N=111	$\bar{Y}$ = 29.279	θ_1 =0.0146	θ_5 =0.0382	θ_9 =0.0117
n= 30	P = 0.117	θ_2 =0.0203	θ_6 =0.1441	
$\beta_2(\phi)$ = 3.898	C_y = 0.872	θ_3 =0.0640	θ_7 =0.0164	
ρ_{pb} = 0.797	C_p = 2.758	θ_4 =0.0709	θ_8 =0.1819	

Table 4. MSE Values of the Classical and Suggested Estimators

Estimator	MSE	
	Population 1	Population 2
$\bar{y}$	0.1579	15.8557
$\bar{y}_{NG1}$	2.2168	94.532
$\bar{y}_{NG2}$	4.3742	254.4077
t_1	0.4030	15.5403
t_2	1.4817	95.4782
t_{ZK1}	0.1404	14.7247
t_{ZK2}	0.1357	14.2948
t_{ZK3}	0.0981	11.3891
t_{ZK4}	0.0982	10.9827
t_{ZK5}	0.1171	13.0318
t_{ZK6}	0.0667	7.6304
t_{ZK7}	0.1404	14.5909
t_{ZK8}	0.0654	6.5614
t_{ZK9}	0.1442	14.9436
$\boldsymbol{\bar{y}_{pr}}$	**0.0652**	**5.7840**

From the Table 4, it is observed that the suggested exponential estimators $\bar{y}_{pri}$, $(i = 1,2,\dots,10)$ perform better than the usual unbiased estimator $\bar{y}$, ratio and product estimators of Naik and Gupta (1996), ratio and product estimators suggested by Singh et al. (2007) and the ratio exponential estimators presented in Zaman and Kadilar (2019a). Finally, it is inferred that the suggested estimators perform better than the considered ratio estimators in all conditions, because the conditions given in Section 3 are always satisfied.

5. Conclusion

This paper proposes exponential families of estimators and presents a theoretical argument. All of the suggested estimators have the same minimum MSE equation. Considering MSE equations, the suggested exponential estimators are always more efficient than those of the sample mean, the ratio and product estimators of Naik and Gupta (1996), the ratio and product estimators of Singh et al. (2007) and the ratio exponential estimators of Zaman and Kadilar (2019a), under all the conditions. The results presented here support these conclusions by theoretical development and numerical analysis. In forthcoming studies, we hope to extend the suggested class of estimators presented in this article to the stratified random sampling.

REFERENCES

BAHL, S., TUTEJA, R. K., (1991). Ratio and product type estimator, Information and Optimization Science, Vol. XII, pp. 159–163.

JHAJJ, H. S., SHARMA, M. K., GROVE, L. K., (2006). A family of estimators of population mean using information on auxiliary attribute, Pakistan Journal of Statistics 22 (1), pp. 43–50.

KOYUNCU, N., (2012). Efficient estimators of population mean using auxiliary attributes, Applied Mathematics and Computation, 218(22), pp. 10900–10905.

MALIK, S., SINGH, R., (2013). An improved estimator using two auxiliary attributes, Applied Mathematics and Computation, 219(23), pp. 10983–10986

NAIK, 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), pp. 151–158.

OZEL, K. G., (2016). A New Exponential Type Estimator for the Population Mean in Simple Random Sampling, Journal of Modern Applied Statistical Methods, 15(2), pp. 207–214

SHABBIR, J., GUPTA, S., (2010). Estimation of the finite population mean in two phase sampling when auxiliary variables are attributes, Hacettepe Journal of Mathematics and Statistics, 39(1).

SINGH, R., CHAUHAN, P., SAWAN, N., SMARANDACHE, F., (2007). Ratio-product type exponential estimator for estimating finite population mean using information on auxiliary attribute, in "Auxiliary Information and a Priori Values in Construction of Improved Estimators" edited by Singh, R., Chauhan, P., Sawan, N. and Smarandache, F., Renaissance High Press, pp. 18–32.

SINGH, R., CHAUHAN, P., SAWAN, N., (2008). On linear combination of Ratio-product type exponential estimator for estimating finite population mean, Statistics in Transition, 9(1), pp. 105–115.

SOLANKI, R. S., SINGH, H. P., (2013). Improved estimation of population mean using population proportion of an auxiliary character, Chilean journal of Statistics, 4(1), pp. 3–17.

SUKHATME, P. V., SUKHATME, B. V., (1970). Sampling Theory of Surveys with Applications, Iowa State University Press, Ames, U.S.A.

ZAMAN, T., SAGLAM, V., SAGIR, M., YUCESOY, E., ZOBU, M., (2014). Investigation of some estimators via Taylor series approach and an application, American Journal of Theoretical and Applied Statistics, 3(5), pp. 141–147.

ZAMAN, T., (2018). New family of estimators using two auxiliary attributes, International Journal of Advanced Research I Engineering & Management (IJAREM), 4(11), pp. 11–16.

ZAMAN, T., KADILAR, C., (2019a). Novel family of exponential estimators using information of auxiliary attribute, Journal of Statistics and Management Systems, pp. 1–11.

ZAMAN, T., KADILAR, C., (2019b). New class of exponential estimators for finite population mean in two-phase sampling, Communications in Statistics – Theory and Methods, pp. 1–16.