Pakistan Journal of Statistics and Operation Research

Improved Estimators using Exponential Function for the Population Mean in Simple and Stratified Random Samplings

Ceren Unal^{1*}, Cem Kadilar²

* Corresponding Author



- 1. Department of Statistics, Hacettepe University, Turkey, cerenunal@hacettepe.edu.tr
- 2. Department of Statistics, Hacettepe University, Turkey, kadilar@hacettepe.edu.tr

Abstract

In this article, we investigated estimators with the exponential function for the estimation of the population mean in the simple and stratified random samplings. Family of estimators based on the exponential function is proposed for both sampling methods. The proposed estimators are compared with estimators in literature. Moreover, we provide an application on different data sets to demonstrate the efficiency of the proposed estimators. As a result, the proposed estimators are more efficient than other estimators in literature under the obtained conditions in theory.

Key Words: Simple Random Sampling; Stratified Random Sampling; Exponential Type Estimators; First Order Approximation; Auxiliary Information.

Mathematical Subject Classification: 62D05

1. Introduction

Working on a population may cause facing with some difficulties about money, time, and manpower. Mostly, it is impossible to reach all of the population. At this point, sample surveys play an important role. After the sample, which has the quality of symbolizing population, is determined, the process of estimating the population parameters follows. By benefiting from the chosen sample, mathematical equality, which is identified for estimating population features, is called estimator. One of the requested features of an estimator is efficiency. It is well known that the information of auxiliary variable is commonly used in order to increase efficiency and precision in sample surveys. It has also a role in the related methods of estimation, such as ratio, product, and regression. At the same time, these different type of estimators are important in field sampling, especially in natural sciences (TÜBİTAK, 2020). If the correlation between the study variable (y) and the auxiliary variable (x) is highly positive, the ratio method of estimation is used. If not, the product method of estimation is employed effectively provided this correlation is highly negative.

For the simple random sampling, there are some assumptions such as the population, $S = (S_1, S_2, ..., S_N)$, consists of N units and a sample of size n was drawn with the simple random sampling without replacement (SRSWOR) and (y_i, x_i) are the values of the study and auxiliary variables for the i^{th} unit (i = 1, 2, ..., 10) of the population, respectively.

The classical ratio type estimator was proposed by Cochran (1940) for estimation of the population mean as

$$\overline{y}_R = \frac{\overline{y}}{\overline{x}} \overline{X} , \qquad (1)$$

where \overline{y} and \overline{x} are the sample means of the study and auxiliary variables, respectively, and \overline{X} is the population mean of the auxiliary variable.

The bias and Mean Square Error (MSE), up to the first order of approximation, are given by

$$Bias(\bar{y}_{R}) = \gamma \bar{Y} (C_{x}^{2} - C_{xy}), \qquad (2)$$

$$MSE(\bar{y}_R) = \gamma \bar{Y}^2 \left(C_y^2 + C_x^2 - 2C_{xy} \right), \tag{3}$$

where \overline{Y} is the population mean of the study variable and $\gamma = \frac{1-f}{n}$, $C_x^2 = \frac{S_x^2}{\overline{X}^2}$, $C_y^2 = \frac{S_y^2}{\overline{Y}^2}$, $C_{xy} = \rho C_x C_y$. Here,

$$f = \frac{n}{N}, S_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_i - \bar{X} \right)^2, S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(y_i - \bar{Y} \right)^2.$$

Robson (1957) proposed the classical product estimator as

$$\bar{y}_{p} = \frac{\bar{y}}{\bar{X}} \bar{x} , \qquad (4)$$

whose bias and MSE are given by

$$Bias(\overline{y}_{P}) = \gamma \overline{Y} C_{xy}, \qquad (5)$$

$$MSE\left(\overline{y}_{P}\right) = \gamma \overline{Y}^{2}\left(C_{y}^{2} + C_{x}^{2} + 2C_{xy}\right).$$
(6)

When the relation between the study variable and the auxiliary variable is a straight line passing through the origin, the ratio and product estimators have same efficiencies equal to the classical regression estimator (Solanki et al., 2012). However, this situation is not valid most of the time. At this point, many authors propose estimators by using the exponential function. Bahl and Tuteja (1991) are the first to propose an estimation using the exponential function for the estimation of the population mean which is given by

$$\overline{y}_{BT} = \overline{y}exp\left(\frac{\overline{X} - \overline{x}}{\overline{X} + \overline{x}}\right).$$
(7)

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

$$Bias(\overline{y}_{BT}) = \gamma \frac{Y}{8} (3C_x^2 - 4C_{xy}), \qquad (8)$$

$$MSE\left(\overline{y}_{BT}\right) = \overline{Y}^{2} \gamma \left(C_{y}^{2} + \frac{C_{x}^{2}}{4} - C_{xy}\right).$$

$$\tag{9}$$

Singh and Pal (2015) proposed a general class of ratio type exponential estimator as

$$\overline{y}_{SP3,i} = \overline{y} \left(\frac{a\overline{X} + b}{a\overline{x} + b} \right) \exp\left\{ \frac{a\left(\overline{X} - \overline{x}\right)}{a\left(\overline{X} + \overline{x}\right) + 2b} \right\}, i = 1, 2, ..., 10,$$
(10)

whose bias and MSE are given by

$$Bias(\bar{y}_{SP3,i}) = \gamma \bar{Y} \frac{3}{8} (5\theta_i^2 C_x^2 - 4\theta_i C_{yx}), i = 1, 2, ..., 10,$$
(11)

$$MSE(\bar{y}_{SP3,i}) = \gamma \bar{Y}^{2} \left(C_{y}^{2} + \frac{9\theta_{i}^{2}C_{x}^{2}}{4} - 3\theta_{i}C_{xy} \right), i = 1, 2, ..., 10.$$
(12)

For the stratified random sampling there are also studies in Literature as follows:

Hansen et al. (1946) proposed classical ratio estimator for the population mean of the study variable as

$$\overline{y}_{R(st)} = \frac{\overline{y}_{st}}{\overline{x}_{st}} \overline{X} , \qquad (13)$$

where $\overline{y}_{st} = \sum_{h=1}^{L} W_h \overline{y}_h$ and $\overline{x}_{st} = \sum_{h=1}^{L} W_h \overline{x}_h$. Here, *L* is the number of strata, \overline{y}_h and \overline{x}_h are the sample means of the

study and auxiliary variables, respectively, in the h^{th} stratum. The bias and MSE of the estimator in Eq. (13) are given by

$$Bias\left(\overline{y}_{R(st)}\right) = \sum_{h=1}^{L} W_h^2 \gamma_h \frac{1}{X} \left(RS_{sh}^2 - S_{ysh}\right),\tag{14}$$

$$MSE\left(\overline{y}_{R(st)}\right) = \sum_{h=1}^{L} W_{h}^{2} \gamma_{h} \left(S_{yh}^{2} + R^{2} S_{xh}^{2} - 2RS_{yxh}\right),$$
(15)

where
$$W_h = \frac{N_h}{N}$$
, $\gamma_h = \frac{1 - f_h}{n_h}$, $R = \frac{Y_{st}}{\bar{X}_{st}}$, $\bar{Y}_{st} = \bar{Y} = \sum_{h=1}^{L} W_h \bar{Y}_h$, $\bar{X}_{st} = \bar{X} = \sum_{h=1}^{L} W_h \bar{X}_h$, $S_{xh}^2 = \frac{1}{N - 1} \sum_{i=1}^{L} \sum_{i=1}^{N_h} (X_{hi} - \bar{X}_h)^2$,
 $S_{yxh} = \frac{1}{N - 1} \sum_{i=1}^{L} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h) (X_{hi} - \bar{X}_h)$, $S_{yh}^2 = \frac{1}{N - 1} \sum_{i=1}^{L} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$. Here, $f_h = \frac{n_h}{N_h}$, \bar{Y}_h and \bar{X}_h are the population

means of the study and auxiliary variables, respectively, in the h^{th} stratum.

Motivated by Bahl and Tuteja (1991), Singh et al. (2008) adapted the estimator to the stratified random sampling as

$$\overline{y}_{BT(st)} = \overline{y}_{st} \exp\left[\frac{\overline{X}_{st} - \overline{x}_{st}}{\overline{X}_{st} + \overline{x}_{st}}\right],\tag{16}$$

whose bias and MSE are given by

$$Bias\left(\bar{y}_{BT(st)}\right) = \sum_{h=1}^{L} W_h^2 \gamma_h \frac{1}{X} \left(\frac{3R}{4} S_{sth}^2 - \frac{S_{ysh}}{2}\right),$$
(17)

$$MSE\left(\overline{y}_{BT(st)}\right) = \sum_{h=1}^{L} W_{h}^{2} \gamma_{h} \left(S_{yh}^{2} + \frac{R^{2}}{4}S_{xh}^{2} - RS_{yxh}\right).$$
(18)

Hansen et al. (1946) proposed the classical regression estimator as

$$\overline{y}_{reg(st)} = \overline{y}_{st} + b_c \left(\overline{X} - \overline{x}_{st} \right), \tag{19}$$

where b_c is the regression coefficient in the stratified random sampling method. The MSE of $\overline{y}_{reg(st)}$, up to the first order of approximation, is given by

$$MSE\left(\overline{y}_{reg(st)}\right) = \sum_{h=1}^{L} W_h^2 \gamma_h S_{yh}^2 \left(1 - \rho_c^2\right), \tag{20}$$

where $\rho_c^2 = \frac{\left(\sum_{h=1}^{L} W_h^2 \gamma_h \rho_h S_{yh} S_{xh}\right)^2}{\sum_{k=1}^{L} W_h^2 \gamma_h S_{yh}^2 \sum_{k=1}^{L} W_h^2 \gamma_h S_{xh}^2}$.

In this paper, we have proposed a new exponential family of estimators for the population mean in the simple and stratified random samplings.

2. The Suggested Exponential Family of Estimators

2.1. Simple Random Sampling

Motivated by Yadav and Kadilar (2013), Singh and Pal (2015) and Ünal and Kadilar (2019), we propose a new exponential family of estimators for the population mean in the Simple Random Sampling as

$$\overline{y}_{CCi} = k\overline{y} \left(\frac{a_i \overline{X} + b_i}{a_i \overline{x} + b_i} \right)^{\alpha} \exp\left(\frac{a_i \left(\overline{X} - \overline{x} \right)}{a_i \left(\overline{X} + \overline{x} \right) + 2b_i} \right), i = 1, 2, ..., 10,$$
(21)

where k is a suitably chosen constant to make the Mean Squared Error (MSE) of \overline{y}_{CCi} , i = 1, 2, ..., 10 minimum, α is a constant which takes values (-1, 0, 1) for designing the different types of estimators and a_i, b_i are either real numbers or the functions of the known parameters of the auxiliary variable, such as the coefficient of variation (C_{r}) , the coefficient of kurtosis $(\beta_2(x))$, and the correlation coefficient (ρ) .

To obtain the Bias and MSE of \overline{y}_{CCi} , i = 1, 2, ..., 10, we consider

$$\overline{y} = \overline{Y}(1+e_y), \ \overline{x} = \overline{X}(1+e_x)$$

then, we have
$$E(e_x) = E(e_y) = 0$$
, $E(e_x^2) = \gamma C_x^2$, $E(e_y^2) = \gamma C_y^2$, $E(e_x e_y) = \gamma \rho_{xy} C_x C_y$.

Now, expressing the family of estimators \overline{y}_{cci} , i = 1, 2, ..., 10 in terms of e_i (i = x, y), we have

$$\overline{y}_{CCi} = k\overline{Y}\left(1+e_y\right) \left(\frac{a_i \overline{X}+b_i}{a_i \overline{X}+a_i \overline{X}e_x+b_i}\right)^{\alpha} \exp\left(\frac{a_i \left(\overline{X}-\overline{X}-\overline{X}e_x\right)}{a_i \left(\overline{X}+\overline{X}+\overline{X}e_x\right)+2b_i}\right), i = 1, 2, ..., 10,$$
(22)

$$= k\overline{Y}(1+e_{y})(1-\theta_{i}e_{x})^{\alpha}\exp\left(-\frac{\theta_{i}e_{x}}{2}\left(1+\frac{\theta_{i}e_{x}}{2}\right)^{-1}\right), i = 1, 2, ..., 10,$$
(23)

where $\theta_i = \frac{a_i \overline{X}}{a_i \overline{X} + b_i}$, i = 1, 2, ..., 10.

Expanding Eq. (23) on right hand side to the first degree of approximation and retaining the terms of $e_i(i = x, y)$ up to the second power, we have

$$\left(\overline{y}_{CCi} - \overline{Y}\right) = \overline{Y}\left(k - \frac{k\theta_i e_x}{2} + \frac{3k\theta_i^2 e_x^2}{8} - k\alpha\theta_i e_x + \frac{k\alpha^2\theta_i^2 e_x^2}{2} + ke_y - \frac{k\theta_i e_y e_x}{2} - k\alpha\theta_i e_y e_x - 1\right)$$
(24)

Taking the expectation of both sides, we get

$$Bias(\overline{y}_{CCi}) = \overline{Y}(k-1) + \overline{Y}\gamma\left(\left(\frac{k\theta_i^2}{2}C_x^2\left(\frac{3}{4}+\alpha^2\right)-k\theta_i C_{yx}\left(\frac{1}{2}+\alpha\right)\right)\right), i = 1, 2, ..., 10$$
(25)

Squaring both sides of Eq. (24) and then taking expectation, we get the MSE of the proposed family of estimators as

$$\left(\bar{y}_{CCi} - \bar{Y}\right)^{2} = \bar{Y}^{2} \left(\left(k-1\right)^{2} + k^{2} e_{y}^{2} + k^{2} \theta_{i}^{2} e_{x}^{2} + 2\alpha^{2} k^{2} \theta_{i}^{2} e_{x}^{2} - k\alpha^{2} \theta_{i}^{2} e_{x}^{2} + \alpha k^{2} \theta_{i}^{2} e_{x}^{2} - \frac{3k \theta_{i}^{2} e_{x}^{2}}{4} + k \theta_{i} e_{y} e_{x} + 2k \alpha \theta_{i} e_{y} e_{x} - 2k^{2} \theta_{i} e_{y} e_{x} - 4\alpha k^{2} \theta_{i} e_{y} e_{x} \right)$$

$$MSE\left(\bar{y}_{CCi}\right) = \bar{Y}^{2} \left(k-1\right)^{2} + \bar{Y}^{2} \gamma \left(k^{2} C_{y}^{2} + \theta_{i}^{2} C_{x}^{2} \left(k^{2} + 2\alpha^{2} k^{2} + \alpha k^{2} - \frac{3k}{4} - k\alpha^{2}\right)$$

$$(26)$$

$$MSE\left(\bar{y}_{CCi}\right) = \bar{Y}^{2} \left(k-1\right)^{2} + \bar{Y}^{2} \gamma \left(k^{2} C_{y}^{2} + \theta_{i}^{2} C_{x}^{2} \left(k^{2} + 2\alpha^{2} k^{2} + \alpha k^{2} - \frac{3k}{4} - k\alpha^{2}\right)$$

$$(27)$$

$$+ \theta_i \rho_{xy} C_y C_x (k + 2k\alpha - 2k^2 - 4\alpha k^2)), \quad i = 1, 2, ..., 10$$

The optimal value of k is obtained by

 $\frac{\partial MSE\left(\overline{y}_{CCi}\right)}{\partial k} = 0$

and we obtain the optimal value of k as

$$k^* = \frac{A}{B},$$
(28)
where $A = \left(\gamma \left(\frac{3}{4}\theta_i^2 C_x^2 + \alpha^2 \theta_i^2 C_x^2 - \theta_i C_{xy} - 2\alpha \theta_i C_{xy}\right) + 2\right)$ and $B = 2\left(\gamma \left(C_y^2 + \theta_i^2 C_x^2 + 2\alpha^2 \theta_i^2 C_x^2 - 2\theta_i C_{xy} - 4\alpha \theta_i C_{xy}\right) + 1\right).$

Using the optimal value of k in the MSE of the proposed estimator, given in Eq. (27), we obtain the minimum MSE of the proposed family of estimators as

$$MSE_{min}(\bar{y}_{CCi}) = \bar{Y}^{2}\left(1 - \frac{A^{2}}{2B}\right), \ i = 1, 2, ..., 10$$
(29)

Some members of the proposed family of estimators \overline{y}_{CCi} , i = 1, 2, ..., 10 are given in Table 1. Table 1: Members of the proposed family of estimators

ole	e I: Members	of the proposed	family of estimation
	$ heta_i$	a_i	b_i
_	$ heta_{\scriptscriptstyle 1}$	1	1
	$ heta_2$	1	$\beta_2(x)$
	$\theta_{_3}$	1	C_{x}
	$ heta_{_4}$	1	ho
	θ_5	$\beta_2(x)$	C_x
	$ heta_6$	C_{x}	$\beta_2(x)$

θ_7	C_x	ρ
$ heta_{8}$	ρ	C_{x}
$ heta_9$	$\beta_2(x)$	ρ
$ heta_{10}$	ρ	$\beta_2(x)$

Using Table 1, we can write

$$\begin{aligned} \theta_{1} &= \frac{X}{\left(\overline{X}+1\right)}, \quad \theta_{2} &= \frac{X}{\left(\overline{X}+\beta_{2}\left(x\right)\right)}, \quad \theta_{3} &= \frac{X}{\left(\overline{X}+C_{x}\right)}, \quad \theta_{4} &= \frac{X}{\left(\overline{X}+\rho\right)} \\ \theta_{5} &= \frac{\beta_{2}\left(x\right)\overline{X}}{\left(\beta_{2}\left(x\right)\overline{X}+C_{x}\right)}, \quad \theta_{6} &= \frac{C_{x}\overline{X}}{\left(C_{x}\overline{X}+\beta_{2}\left(x\right)\right)}, \quad \theta_{7} &= \frac{C_{x}\overline{X}}{\left(C_{x}\overline{X}+\rho\right)}, \\ \theta_{8} &= \frac{\rho\overline{X}}{\left(\rho\overline{X}+C_{x}\right)}, \quad \theta_{9} &= \frac{\beta_{2}\left(x\right)\overline{X}}{\left(\beta_{2}\left(x\right)\overline{X}+\rho\right)}, \quad \theta_{10} &= \frac{\rho\overline{X}}{\left(\rho\overline{X}+\beta_{2}\left(x\right)\right)}. \end{aligned}$$

2.2. Stratified Random Sampling

We adapt the proposed exponential family of estimators for the population mean in the simple random sampling to the stratified random sampling as follows:

$$\overline{y}_{CCi}^{(st)} = k\overline{y}_{st} \left(\frac{a_i \overline{X}_{st} + b_i}{a_i \overline{x}_{st} + b_i} \right)^{\alpha} \exp\left(\frac{a_i \left(\overline{X}_{st} - \overline{x}_{st} \right)}{a_i \left(\overline{X}_{st} + \overline{x}_{st} \right) + 2b_i} \right), \ i = 1, 2, \dots, 10.$$
(30)

In order to obtain the Bias and MSE of $\overline{y}_{CCi}^{(st)}$, i = 1, 2, ..., 10, we define

$$\overline{y}_{st} = \overline{Y}(1+e_y), \quad \overline{x}_{st} = \overline{X}(1+e_x),$$
then, we have $E(e_x^2) = \frac{1}{\overline{X}^2} \sum_{h=1}^{L} W_h^2 \gamma_h S_{xh}^2, \quad E(e_y^2) = \frac{1}{\overline{Y}^2} \sum_{h=1}^{L} W_h^2 \gamma_h S_{yh}^2, \quad E(e_x e_y) = \frac{1}{\overline{X}\overline{Y}} \sum_{h=1}^{L} W_h^2 \gamma_h S_{yxh}, \quad E(e_x) = E(e_y) = 0.$
Now, rewriting the expression for the family of estimators $\overline{y}^{(st)}$ $i = 1, 2, \dots, 10$ in terms of $e_x(i = x, y)$, we can write estimators $\overline{y}^{(st)}$.

Now, rewriting the expression for the family of estimators $\overline{y}_{CCi}^{(st)}$, i = 1, 2, ..., 10 in terms of $e_i(i = x, y)$, we can write as

$$\overline{y}_{CCi}^{(st)} = k\overline{Y}\left(1 + e_y\right) \left(\frac{a_i \overline{X} + b_i}{a_i \overline{X} + b_i} + \frac{a_i \overline{X}e_x}{a_i \overline{X} + b_i}\right)^{-\alpha} \exp\left(\frac{-a_i \overline{X}e_x}{2\left(a_i \overline{X} + b_i\right) + a_i \overline{X}e_x}\right),\tag{31}$$

$$= k\overline{Y}\left(1+e_{y}\right)\left(1-\theta_{i}e_{x}\right)^{\alpha}\exp\left(-\frac{\theta_{i}e_{x}}{2}\left(1+\frac{\theta_{i}e_{x}}{2}\right)^{-1}\right),$$
(32)

where $\theta_i = \frac{a_i \overline{X}}{a_i \overline{X} + b_i}$, i = 1, 2, ..., 10.

Expanding Eq. (32) on right hand side to the first degree of approximation and subtracting \overline{Y} from both sides, we get

$$\left(\overline{y}_{CCi}^{(st)} - \overline{Y}\right) = \overline{Y}\left(\left(k-1\right) + ke_{y} - \frac{k\theta_{i}e_{x}}{2} - k\alpha\theta_{i}e_{x} + \frac{3k\theta_{i}^{2}e_{x}^{2}}{8} + \frac{k\alpha^{2}\theta_{i}^{2}e_{x}^{2}}{2} - \frac{k\theta_{i}e_{y}e_{x}}{2} - k\alpha\theta_{i}e_{y}e_{x}\right)$$
(33)

Taking the expectation of both sides, we have

$$Bias\left(\overline{y}_{CCi}^{(st)}\right) = (k-1)\sum_{h=1}^{L} W_{h}^{2} \gamma_{h} \frac{k\theta_{i}}{\overline{X}} \left(\frac{\theta_{i}RS_{xh}^{2}}{2} \left(\frac{3}{4} + \alpha^{2}\right) - S_{yxh}\left(\frac{1}{2} + \alpha\right)\right), \ i = 1, 2, ..., 10$$
(34)

Squaring both sides of Eq. (33), we have

$$\left(\bar{y}_{CCi}^{(st)} - \bar{Y}\right)^{2} = \bar{Y}^{2} \left(\left(k-1\right)^{2} + k^{2} e_{y}^{2} + k^{2} \theta_{i}^{2} e_{x}^{2} + 2\alpha^{2} k^{2} \theta_{i}^{2} e_{x}^{2} - k\alpha^{2} \theta_{i}^{2} e_{x}^{2} \right) + \alpha k^{2} \theta_{i}^{2} e_{x}^{2} - \frac{3k \theta_{i}^{2} e_{x}^{2}}{4} + k \theta_{i} e_{y} e_{x} + 2k \alpha \theta_{i} e_{y} e_{x} - 4\alpha k^{2} \theta_{i} e_{y} e_{x} - 2k^{2} \theta_{i} e_{y} e_{x} \right)$$

$$(35)$$

Taking the expectation of both sides of Eq. (35), we get the MSE of the estimator $\overline{y}_{CCi}^{(st)}$, i = 1, 2, ..., 10 to the first degree of approximation as

$$MSE\left(\bar{y}_{CCi}^{(st)}\right) = \left(\bar{Y}^{2}\left(k-1\right)^{2} + \sum_{h=1}^{L} W_{h}^{2} \gamma_{h}\left(k^{2} S_{yh}^{2} + \theta_{i}^{2} R^{2} S_{xh}^{2}\left(k^{2} + 2\alpha^{2} k^{2} + \alpha k^{2} - \frac{3k}{4} - k\alpha^{2}\right) + \theta_{i} RS_{yxh}\left(k + 2k\alpha - 2k^{2} - 4\alpha k^{2}\right)\right), \quad i = 1, 2, ..., 10$$
(36)

For minimizing the MSE of the proposed estimators, given in Eq. (36), we obtain the optimum value of k by

$$k_{st}^* = \frac{A_{st}}{B_{st}}, \qquad (37)$$

where $A_{st} = \left(\sum_{h=1}^{L} W_h^2 \gamma_h \left(\frac{3}{4} \theta_i^2 R^2 S_{xh}^2 + \alpha^2 \theta_i^2 R^2 S_{xh}^2 - \theta_i R S_{xyh} - 2\alpha \theta_i R S_{xyh}\right) + 2\overline{Y}^2\right)$ and $B_{st} = \left(\sum_{h=1}^{L} W_h^2 \gamma_h \left(2S_{yh}^2 + 2\theta_i^2 R^2 S_{xh}^2 + 4\alpha^2 \theta_i^2 R^2 S_{xh}^2 + 2\alpha \theta_i^2 R^2 S_{xh}^2 - 4\theta_i R S_{xyh} - 8\alpha \theta_i R S_{xyh}\right) + 2\overline{Y}^2\right).$

Replacing the value of k_{st}^* instead of k in Eq. (36), we obtain the minimum value of $MSE(\overline{y}_{CCi}^{(st)}), i = 1, 2, ..., 10$ as

$$MSE_{min}\left(\bar{y}_{CCi}^{(st)}\right) = \left(\bar{Y}^2 - \frac{A_{st}^2}{2B_{st}}\right), \ i = 1, 2, ..., 10$$
(38)

We can also write some members of $\overline{y}_{CCI}^{(st)}$, i = 1, 2, ..., 10 using Table 1 in the stratified random sampling method.

3. Efficiency Comparisons

3.1. Simple Random Sampling

We compare the MSE of proposed family of estimators with the MSE of classical ratio type estimator \overline{y}_R (Cochran, 1940), Bahl and Tuteja estimator \overline{y}_{BT} (1991), classical regression type estimator \overline{y}_{reg} (Cochran, 1977), and Singh and Pal $\overline{y}_{SP3,i}$, i = 1, 2, ..., 10 estimators (2015). The MSE of the mentioned estimators are given in Table 2.

Table 2: MSE equations of the estimators in the simple random sampling

Estimators	MSE
$\overline{y}_R = \overline{y}\left(\frac{\overline{X}}{\overline{x}}\right)$	$MSE(\overline{y}_{R}) = \gamma \overline{Y}^{2} \left(C_{y}^{2} + C_{x}^{2} - 2C_{xy} \right)$
$\overline{y}_{BT} = \overline{y} \exp\left(\frac{\overline{X} - \overline{x}}{\overline{X} + \overline{x}}\right)$	$MSE\left(\overline{y}_{BT}\right) = \gamma \overline{Y}^{2} \left(C_{y}^{2} + \frac{C_{x}^{2}}{4} - C_{xy}\right)$
$\overline{y}_{reg} = \overline{y} + b\left(\overline{X} - \overline{x}\right)$	$MSE\left(\overline{y}_{reg}\right) = \gamma \overline{Y}^2 C_y^2 \left(1 - \rho_{xy}^2\right)$
$\overline{y}_{SP3,i}, i = 1, 2,, 10$	$MSE\left(\overline{y}_{SP3,i}\right) = \gamma \overline{Y}^2 \left(C_y^2 + \frac{9\theta_i^2 C_x^2}{4} - 3\theta_i C_{xy}\right), i = 1, 2,, 10$
$\overline{y}_{CCi}, i = 1, 2,, 10$	$MSE\left(\overline{y}_{CCi}\right) = \overline{Y}^{2}\left(1 - \frac{A^{2}}{2B}\right), i = 1, 2,, 10$

We find the efficiency conditions of the proposed estimators as follows:

$$MSE_{min}\left(\overline{y}_{CCi}\right) < MSE\left(\overline{y}_{R}\right), i = 1, 2, ..., 10$$

$$\left(1 - \frac{A^{2}}{2B}\right) < \gamma\left(C_{y}^{2} + C_{x}^{2} - 2C_{xy}\right)$$
(39)

$$MSE_{min}\left(\overline{y}_{CCi}\right) < MSE\left(\overline{y}_{BT}\right), \ i = 1, 2, ..., 10$$

$$\left(1 - \frac{A^2}{2B}\right) < \gamma \left(C_y^2 + \frac{C_x^2}{4} - C_{xy}\right)$$

$$(40)$$

$$MSE_{min}\left(\overline{y}_{CCi}\right) < MSE\left(\overline{y}_{reg}\right), \ i = 1, 2, ..., 10$$

$$\left(1 - \frac{A^2}{2B}\right) < \gamma C_y^2 \left(1 - \rho_{xy}^2\right)$$
(41)

$$MSE_{min}(\bar{y}_{CCi}) < MSE(\bar{y}_{SP3,i}), i = 1, 2, ..., 10$$

$$\left(1 - \frac{A^2}{2B}\right) < \gamma \left(C_y^2 + \frac{9\theta_i^2 C_x^2}{4} - 3\theta_i C_{xy}\right), i = 1, 2, ..., 10$$
(42)

When the conditions (39) – (42) are satisfied, we infer that the proposed estimators \overline{y}_{CCi} , i = 1, 2, ..., 10 are more efficient than the classical ratio estimator (Cochran, 1940), the Bahl and Tuteja estimator (1991), the classical regression estimator (Cochran, 1977), and the Singh and Pal estimator (2015), respectively.

3.2. Stratified Random Sampling

In this section, we compare the MSE of proposed family of estimators with the MSE of Hansen et al. classical ratio estimator $\overline{y}_{R(st)}$ (1946), Singh et al. estimator $\overline{y}_{BT(st)}$ (2008), and the classical regression type estimator $\overline{y}_{reg(st)}$ (1946) in the stratified random sampling method.

Using (15), (18), (20), and (38), we have

$$MSE_{min}\left(\bar{y}_{CCi}^{(st)}\right) < MSE\left(\bar{y}_{R(st)}\right), \ i = 1, 2, ..., 10$$

$$\left(\bar{Y}^{2} - \frac{A^{2}}{2B}\right) < \sum_{h=1}^{L} W_{h}^{2} \gamma_{h} \left(S_{yh}^{2} + R^{2} S_{xh}^{2} - 2RS_{yxh}\right)$$

$$MSE_{min}\left(\bar{y}_{CCi}^{(st)}\right) < MSE\left(\bar{y}_{BT(st)}\right), \ i = 1, 2, ..., 10$$

$$\left(\bar{Y}^{2} - \frac{A^{2}}{2B}\right) < \sum_{h=1}^{L} W_{h}^{2} \gamma_{h} \left(S_{yh}^{2} + \frac{R^{2}}{4} S_{xh}^{2} - RS_{yxh}\right)$$

$$MSE_{min}\left(\bar{y}_{CCi}^{(st)}\right) < MSE\left(\bar{y}_{reg(st)}\right), \ i = 1, 2, ..., 10$$

$$\left(\bar{Y}^{2} - \frac{A^{2}}{2B}\right) < MSE\left(\bar{y}_{reg(st)}\right), \ i = 1, 2, ..., 10$$

$$\left(\bar{Y}^{2} - \frac{A^{2}}{2B}\right) < MSE\left(\bar{y}_{reg(st)}\right), \ i = 1, 2, ..., 10$$

$$\left(\bar{Y}^{2} - \frac{A^{2}}{2B}\right) < \sum_{h=1}^{L} W_{h}^{2} \gamma_{h} S_{yh}^{2} \left(1 - \rho_{c}^{2}\right)$$

$$(45)$$

If the conditions (43) - (45) are satisfied, the proposed family of estimators is more efficient than the mentioned estimators
$$\overline{y}_{R(st)}$$
, $\overline{y}_{BT(st)}$, $\overline{y}_{reg(st)}$.

4. Numerical Illustrations

If

4.1. Simple Random Sampling

We have used Kadilar and Cingi (2003) data to show the appropriateness of the proposed family of estimators in practice as well. Since the correlation between the study variable (y) and the auxiliary variable (x) is positive in the data set, value of α is taken as 1. The descriptive statistics of the data are given in Table 3. Note that for this data set, y= apple production in the Marmara Region, x= the number of apple trees in the Marmara Region.

 $\theta_{10} = 0,998715349, \ \overline{y}_{CC10}$

Table 3: Descriptive statistics of the population I			
N = 106, n = 20	$S_{yx} = 257778692.30$		
$C_x = 2.02$	$C_{y} = 4.18$		
$\bar{X} = 24375.59$	$\overline{Y} = 1536.77$		
$S_x = 49189.08$	$S_y = 6425.09$		
$\rho = 0.82$	$\beta_2(x) = 25.71$		

The MSE values of the proposed estimators and $\overline{y}_{R}, \overline{y}_{BT}, \overline{y}_{reg}, \overline{y}_{SP3,i}, i = 1, 2, ..., 10$ estimators are given in Table 4 and Table 5, respectively. We would like to remark that all of the proposed estimators are more efficient than the other estimators, including the regression estimator.

Table 4: MSE of the proposed estimators in the simple random sampling		
Proposed Estimators	MSE	
$\theta_1 = 0,999958976, \ \overline{y}_{CC1}$	$MSE_{\min}(\bar{y}_{CC1}) = 1313627,498$	
$\theta_2 = 0,998946342, \ \overline{y}_{CC2}$	$MSE_{\min}(\bar{y}_{CC2}) = 1315907,209$	
$\theta_3 = 0,999917135, \ \overline{y}_{CC3}$	$MSE_{\min}(\bar{y}_{CC3}) = 1313721,904$	
$\theta_4 = 0,99996636, \ \overline{y}_{CC4}$	$MSE_{\min}(\bar{y}_{CC4}) = 1313610,835$	
$\theta_5 = 0,999996777, \ \overline{y}_{CC5}$	$MSE_{\min}(\bar{y}_{CC5}) = 1313542,192$	
$\theta_6 = 0,99947811, \ \overline{y}_{CC6}$	$MSE_{\min}(\bar{y}_{CC6}) = 1314711,387$	
$\theta_7 = 0,999983346, \ \overline{y}_{CC7}$	$MSE_{\min}(\bar{y}_{CC7}) = 1313572,502$	
$\theta_8 = 0,999898947, \ \overline{y}_{CC8}$	$MSE_{\min}(\overline{y}_{CC8}) = 1313762,936$	
$\theta_9 = 0,999998692, \ \overline{y}_{CC9}$	$MSE_{\min}(\bar{y}_{CC9}) = 1313537,87$	

		1	.1 .	1 1	1.
I oblo St MISL of the	actimatore in	1 Itoroturo in	the com	alo rondor	n compling
	eximitations in	і шеганне ш	THE SHIII	петанов	n vanninno
	countators in	i moratare m	une binn	Jie rundor	n sampning

 $MSE_{\min}(\bar{y}_{CC10}) = 1316425,746$

Estimators in Literature	MSE		
\overline{y}_R	$MSE(\overline{y}_R) = 1854265,551$		
\overline{y}_{BT}	$MSE\left(\overline{y}_{BT}\right) = 2783998, 29$		
$\overline{\mathcal{Y}}_{reg}$	$MSE\left(\overline{y}_{reg}\right) = 1377460, 202$		
$\theta_1 = 0,999958976, \ \overline{y}_{SP3,\theta_1}$	$MSE\left(\overline{y}_{SP3,\theta_1}\right) = 1415527,045$		
$\theta_2 = 0,998946342, \ \overline{y}_{SP3,\theta_2}$	$MSE\left(\overline{y}_{SP3,\theta_2}\right) = 1416116,65$		
$\theta_3 = 0,999917135, \ \overline{y}_{SP3,\theta_3}$	$MSE\left(\overline{y}_{SP3,\theta_3}\right) = 1415551,317$		
$\theta_4 = 0,99996636, \ \overline{y}_{SP3,\theta_4}$	$MSE\left(\overline{y}_{SP3,\theta_4}\right) = 1415522,762$		
$\theta_5 = 0,999996777, \ \overline{y}_{SP3,\theta_5}$	$MSE\left(\overline{y}_{SP3,\theta_5}\right) = 1415505, 123$		
$\theta_6 = 0,99947811, \ \overline{y}_{SP3,\theta_6}$	$MSE\left(\overline{y}_{SP3,\theta_6}\right) = 1415806,464$		
$\theta_7 = 0,999983346, \overline{y}_{SP3,\theta_7}$	$MSE\left(\overline{y}_{SP3,\theta_{7}}\right) = 1415512,911$		
$\theta_8 = 0,999898947, \ \overline{y}_{SP3,\theta_8}$	$MSE\left(\overline{y}_{SP3,\theta_8}\right) = 1415561,87$		

$\theta_9 = 0,999998692, \ \overline{y}_{SP3,\theta_9}$	$MSE\left(\overline{y}_{SP3,\theta_9}\right) = 1415504,013$
$\theta_{10} = 0,998715349, \ \overline{y}_{SP3,\theta_{10}}$	$MSE\left(\overline{y}_{SP3,\theta_{10}}\right) = 1416251,781$

4.2. Stratified Random Sampling

We have used the data set of Lone and Tailor (2016) to show the proposed estimator applicable in the stratified random sampling. Note that for this data set,

y= output

x= fixed Capital

The descriptive statistics of the population are given in Table 6.

Table 6: Descriptive statistics of the population II			
Stratum	1	2	
N_h	5	5	_
n_h	2	2	
W_h	0.5	0.5	
\overline{X}_h	214.4	333.8	
$\overline{Y_h}$	1925.8	315.6	
S_{xh}	74.87	66.35	
S_{yh}	615.92	340.38	
ρ	0.8535	0.9899	
$\beta_{2h}(x)$	1.88	2.32	
S_{yxh}	39360.68	22356.50	

The MSE values of the proposed family of estimators and $\overline{y}_{R(st)}$, $\overline{y}_{BT(st)}$, $\overline{y}_{reg(st)}$ estimators are given in Table 7 and Table 8, respectively.

Proposed Estimators	MSE
$\theta_1 = 0,9963649582, \ \overline{y}_{CC1}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC1}^{(st)}\right) = 8535, 583484$
$\theta_2 = 0,9927562477, \ \overline{y}_{CC2}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC2}^{(s)}\right) = 8538,189888$
$\theta_3 = 0,9990014032, \ \overline{y}_{CC3}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC3}^{(s)}\right) = 8534,152727$
$\theta_4 = 0,9966486281, \ \overline{y}_{CC4}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC4}^{(st)}\right) = 8535, 410349$
$\theta_5 = 0,9995004522, \ \overline{y}_{CC5}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC5}^{(s)}\right) = 8533,926898$
$\theta_6 = 0,9727980631, \ \overline{y}_{CC6}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC6}^{(s)}\right) = 8566, 122956$
$\theta_7 = 0,9878758916, \ \overline{y}_{CC7}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC7}^{(st)}\right) = 8542,905769$
$\theta_8 = 0,9989166625, \ \overline{y}_{CC8}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC8}^{(st)}\right) = 8534,1922497$
$\theta_9 = 0,9983215014, \ \overline{y}_{CC9}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC9}^{(st)}\right) = 8534,483446$
$\theta_{10} = 0,9917562348, \ \overline{y}_{CC10}^{(st)}$	$MSE_{\min}\left(\overline{y}_{CC10}^{(st)}\right) = 8539,044653$

Table 7: MSE of the proposed estimators in the stratified random sampling

Estimators in Literature	MSE
$\overline{\mathcal{Y}}_{R(st)}$	$MSE\left(\overline{y}_{R(st)}\right) = 11837, 81832$
$\overline{\mathcal{Y}}_{BT(st)}$	$MSE\left(\overline{y}_{BT(st)}\right) = 21352,59727$
$\overline{\mathcal{Y}}_{reg(st)}$	$MSE\left(\overline{y}_{reg(st)}\right) = 8595,99973$

Table 8: MSE of the estimators in literature in the stratified random sampling

We would like to remark that all of the proposed estimators are more efficient than the other estimators in the stratified random sampling, as well.

5. Conclusion

We develop a new exponential family of estimators based on the estimators in Literature under the simple random sampling method and adapt this estimator to the stratified random sampling. The proposed families of estimators in both simple and stratified random samplings are found more efficient than the estimators in Literature under the obtained conditions in theory. Using the Kadilar and Cingi (2003) data in the simple random sampling and the Lone and Tailor (2016) data in the stratified random sampling, we show that the proposed family of estimators are quite efficient than other estimators in literature in both the simple and the stratified random samplings as seen in Tables 4-5 and Tables 7-8, respectively.

References

- 1. Bahl, S. & Tuteja, R. K. (1991). Ratio and product type exponential estimators. Journal of Information and Optimization Sciences, 12(1), 159-164.
- 2. Cochran, W. G. (1977). Sampling Techniques, John Wiley and Sons, New-York.
- 3. Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. The Journal of Agricultural Science, 30(02), 262-275.
- 4. Hansen, M. H., Hurwitz, W. N. & Gurney, M. (1946). Problem and methods of the sample survey of business. Journal of American Statistical Association, 41, 174-189.
- 5. Kadilar, C. & Cingi, H. (2003). Ratio estimators in stratified random sampling. Biometrical Journal, 45(2), 218-225.
- 6. Lone, H. A. & Tailor, R. (2016). A class of estimators of population mean in case of post stratification. Pakistan Journal of Statistics and Operation Research, 12(1), 111-124.
- 7. Robson, D. S. (1957). Application of multivariate polykays to the theory of unbiased ratio type estimation. Journal of American Statistical Association, 52(280), 511-522.
- 8. Singh, H. P. & Pal, S. K. (2015). A new chain ratio-ratio-type exponential estimator using auxiliary information in sample surveys. International Journal of Mathematics And its Applications, 3(4B), 37-46.
- 9. Singh, R., Kumar, M., Singh, R. D. & Chaudhary, M. K. (2008). Exponential ratio-type estimators in stratified random sampling. International Symposium on Optimisation and Statistics, Aligarh.
- Solanki, R. S., Singh, H. P. & Rathaur, A. (2012). An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. International Scholarly Research Network Probability and Statistics, 1-14.
- 11. TÜBİTAK 2237-A, Doğa Bilimlerinde İstatistiksel Modelleme Teknikleri ve Uygulamaları, Project Number: 1129B371900490, 2020.
- 12. Ünal, C. & Kadilar, C. (2019). Improved family of estimators using exponential function for the population mean in the presence of non-response. Communications in Statistics Theory and Methods, 50(1), 237-248.
- 13. Yadav, S. K. & Kadilar, C. (2013). Efficient family of exponential estimators for the population mean. Hacettepe Journal of Mathematics and Statistics, 42(6), 671-677.