A NEW CLASS OF EXPONENTIAL REGRESSION CUM RATIO ESTIMATOR IN TWO PHASE SAMPLING

Nilgun Ozgul * † and Hulya Cingi * ‡

Abstract

In this paper, we propose a new class of exponential regression cum ratio estimator using the auxiliary variable for the estimation of the finite population mean under two phase sampling scheme. The Bias and Mean Square Error (MSE) equations of the proposed estimator are obtained and compared with the MSE equations of some existing estimators in two phase sampling. We find theoretically the proposed estimator is always more efficient than classical ratio and regression estimators, Singh and Vishwakarma [17] ratio type exponential estimator in two phase sampling. In addition, theoric results are supported by a numerical example using original data sets.

1. Introduction

In the sampling theory, the use of auxiliary information results in considerable improvement in the precision of estimators of population mean. The ratio and regression methods have been widely used when auxiliary information is available. In literature, number of authors introduced many ratio and regression type estimators by using general linear transformation of the auxiliary variable. For recent development, exponential estimators have been widely studied by several authors such as Bahl and Tuteja [2], Singh et al. [19] and Grover and Kaur [6].

Under various sampling schemes, many exponential estimators, using the population information of the auxiliary variable, have been proposed. However, the knowledge on the population mean of the auxiliary variable is not always available. In this situation, two phase sampling method is the most popular sampling scheme in literature. Two

---
*Hacettepe University, Department of Statistics, Beytepe, 06800, Ankara, Turkey.

†Email: (N. Ozgul) nozgul@hacettepe.edu.tr

‡Email: (H. Cingi) hcingi@hacettepe.edu.tr

phase sampling, first introduced by Neyman [13], is a cost effective technique in survey sampling. It is typically used when it is very expensive to collect data on the variables of interest, but it is relatively inexpensive to collect data on variables that are correlated with the variables of interest. By these aspects two phase sampling is a powerful and cost economical procedure for finding the reliable estimate in first phase sample for the unknown parameters of the auxiliary variable x. Simply, a field survey is to be undertaken to determine the average value of some characters of a population. For example, the amount of money families spend on food. As the collection of data requires long interviews by specially trained enumerators, the cost per family is quite high. The cost of survey is constrained within a specified amount but the sample does not appear to yield an estimate of desired precision because of the great variability of the character. Nevertheless, the character is correlated with a second character that can be determined at a lower cost per family so that a precise estimate of the distribution of this second character is readily obtained. Hence, a more precise estimate of the original character can be found by first estimating the distribution of the second character alone from a large random sample [10]. In literature, many authors improved ratio and regression estimators using at least one auxiliary variable under two phase sampling scheme. Singh and Espejo [16] suggested a class of ratio-product estimators in two phase sampling with its properties and identified asymptotically optimum estimators from proposed class of estimators. Samiuddin and Hanif [14] proposed ratio and regression estimation procedures to estimate the population mean in two-phase sampling using idea of partial and no information cases. Ahmad [1] has proposed various estimators for two phase and multiphase sampling using information on several auxiliary variables. Hanif et al. [7] proposed regression estimator using several auxiliary variables. In recent years, exponential estimators have not been studied sufficiently in two phase sampling. Singh and Vishwakarma [17] adapted Bahl and Tuteja [2] exponential ratio type estimator into two phase sampling. We, here, give the notations about two phase sampling and various estimators of the population mean in two phase sampling method in Section 2. We propose a class of exponential regression cum ratio estimator in Section 3. In Section 4, the proposed estimator is compared with other existing estimators in two phase sampling and we obtain certain conditions that proposed estimator is found to be more efficient than other estimators. In Section 5, the theoretical results are supported by a numerical example. In Section 6, we give conclusion.

2. Notations and Various Existing Estimators

Consider a finite population $U = U_1, U_2, \ldots, U_N$, of size N units. Let y denote the study variable taking the values y_i on the unit U_i, $(i = 1, 2, \ldots, N)$ and $\overline{Y}$ is its unknown population mean. Let x denotes the auxiliary variable taking the values x_i on the unit U_i, $(i = 1, 2, \ldots, N)$ positively correlated with y and $\overline{X}$ is its unknown population mean.

It is well known that when the population mean of the auxiliary variable is not known, two phase sampling is used. Two phase sampling consists of two phase. In first phase, a sample of fixed size is drawn by Simple Random Sampling Without Replacement (SRSWOR) from the finite population to estimate the mean of the auxiliary variable. The sample is drawn in first phase is named as primary sample and expressed by s'. The usual practice is to estimate the mean of the auxiliary variable by sample mean. In second phase, a sample s $(s \subset s')$ of fixed size n is drawn SRSWOR from the primary sample (s') to estimate the mean of the study variable. The sample is drawn in second phase is named as sub sample and expressed by s [14].

When information is not available on the auxiliary variable, x, that is positively correlated with the study variable, y, the classical ratio estimator is a widely used estimator to estimate the population mean, $\overline{Y}$, in two phase sampling as follows:

$$(2.1) \quad \overline{y}_R = \frac{\overline{y}}{\overline{x}}\overline{x}'$$

where $\overline{x}'$ is the primary sample mean of the auxiliary variable, $\overline{y}$ and $\overline{x}$ are the sub sample means of the study and auxiliary variables, respectively. It is well known that the MSE equation of the classical ratio estimator is given by

$$(2.2) \quad MSE\left(\overline{y}_R\right) \cong \overline{Y}^2\left[\lambda C_y^2 + \lambda^* C_x^2\left(1 - 2K_{yx}\right)\right]$$

where $K_{yx} = \rho_{yx}\dfrac{C_y}{C_x}$; $\lambda = \dfrac{1}{n} - \dfrac{1}{N}$; $\lambda^* = \dfrac{1}{n} - \dfrac{1}{n'}$; n' is the primary sample size; n is the sub sample size; N is the number of units in the population; ρ_{yx} is the population correlation coefficient between the auxiliary and the study variables, C_x and C_y are the population coefficients of variation of the auxiliary and study variables, respectively.

When auxiliary variable is correlated with the study variable, the classical unbiased regression estimator is used to estimate the population mean, in two phase sampling as follows:

$$(2.3) \quad \overline{y}_{lr} = \overline{y} + \beta_{yx}\left(\overline{x}' - \overline{x}\right)$$

where β_{yx} is the regression coefficient between the auxiliary and the study variables. It is well known that the variance of the classical regression estimator is given by

$$(2.4) \quad Var\left(\overline{y}_{lr}\right) = \overline{Y}^2 C_y^2\left(\lambda - \lambda^* \rho_{yx}^2\right)$$

Singh and Vishwakarma [17] suggested the following modified exponential ratio estimator in two phase sampling

$$(2.5) \quad \overline{y}_{svr} = \overline{y}\exp\left(\frac{\overline{x}' - \overline{x}}{\overline{x}' + \overline{x}}\right)$$

The MSE equation of the estimator can be given by

$$(2.6) \quad MSE\left(\overline{y}_{svr}\right) \cong \overline{Y}^2\left[\lambda C_y^2 + \lambda^*\left(\frac{C_x^2}{4} - \rho_{yx}C_yC_x\right)\right]$$

In sampling literature, the authors rarely consider the exponential estimators in two phase sampling scheme. For this reason, we improved a class of exponential regression cum ratio estimator in two phase sampling using the ratio and regression methods and their linear transformation in this study.

3. Suggested Exponential Estimator in Two Phase Sampling

Replacing regression estimator instead of sample mean and using linear transformation in exponential term in Singh and Vishwakarma [17] exponential ratio estimator given in (2.5), we improve a class of exponential regression cum ratio estimator as follows:

$$(3.1) \quad \overline{y}_{NH} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left(\frac{\overline{z}' - \overline{z}}{\overline{z} + \overline{z}'}\right)$$

where k_1 and k_2 are some known constants, $\overline{z}'$ is a transformation of the auxiliary variable at first phase as $\overline{z}' = a\overline{x}' + b$, and $\overline{z}$ is a transformation of the auxiliary variable at second phase as $\overline{z} = a\overline{x} + b$.

Then, we have

(3.2) $\qquad \left. \begin{array}{l} \overline{z}' = a\overline{x}' + b \\ \overline{z} = a\overline{x} + b \end{array} \right\}$

where $a\,(\neq 0)$ and b are either any known constants or functions of any known population parameters of the auxiliary variable, such as standard deviation (σ_x), coefficient of variation (C_x), coefficient of skewness $\{\beta_1(x)\}$, coefficient of kurtosis $\{\beta_2(x)\}$, coefficient of correlation (ρ_{yx}) [9]. The list of new exponential estimator generated from (3.1) is given in Table 1.

To obtain the Bias and MSE equations for the proposed estimator, we define following notations:

(3.3) $\qquad e_0 = \dfrac{(\overline{y} - \overline{Y})}{\overline{Y}}, e_1 = \dfrac{(\overline{x} - \overline{X})}{\overline{X}}, e_1' = \dfrac{\left(\overline{x}' - \overline{X}\right)}{\overline{X}}$

such that

(3.4) $\qquad \begin{array}{l} E(e_0) = E(e_1) = E\left(e_1'\right) = 0; E(e_0^2) = \lambda C_y^2; E(e_1^2) = \lambda C_x^2; \\[2mm] E\left(e_1'^{\,2}\right) = \lambda' C_x^2; E(e_0 e_1) = \lambda \rho_{yx} C_y C_x; E\left(e_0 e_1'\right) = \lambda' \rho_{yx} C_y C_x; \\[2mm] E\left(e_1 e_1'\right) = \lambda' C_x^2 \end{array}$

where

$$\lambda = \frac{1}{n} - \frac{1}{N}, \quad \lambda' = \frac{1}{n'} - \frac{1}{N}, \quad C_y^2 = \frac{S_y^2}{\overline{Y}^2}, \quad C_x^2 = \frac{S_x^2}{\overline{X}^2}, \quad S_y^2 = \frac{\sum\limits_{i=1}^{N} (y_i - \overline{Y})^2}{N-1},$$

$$S_x^2 = \frac{\sum\limits_{i=1}^{N} (x_i - \overline{X})^2}{N-1}, \quad \rho_{yx} = \frac{\sum\limits_{i=1}^{N} (y_i - \overline{Y})(x_i - \overline{X})}{\sqrt{\sum\limits_{i=1}^{N} (x_i - \overline{X})^2 \sum\limits_{i=1}^{N} (y_i - \overline{Y})^2}}$$

and we use Taylor series method [4] for two variables to solve the exponential term as

(3.5)
$$\begin{aligned} f\left(e_1, e_1'\right) = {}& f\left(e_1, e_1'\right)\big|_{e_1 = e_1' = 0} + \frac{1}{1!} \frac{\partial f\left(e_1, e_1'\right)}{\partial e_1}\bigg|_{e_1 = e_1' = 0} \\[2mm] & + \frac{1}{1!} \frac{\partial f\left(e_1, e_1'\right)}{\partial e_1'}\bigg|_{e_1 = e_1' = 0} + \frac{1}{2!} \frac{\partial f\left(e_1, e_1'\right)}{\partial e_1^2}\bigg|_{e_1 = e_1' = 0} \\[2mm] & + \frac{1}{2!} \frac{\partial f\left(e_1, e_1'\right)}{\partial e_1'^{\,2}}\bigg|_{e_1 = e_1' = 0} + \frac{1}{2!} \frac{\partial f\left(e_1, e_1'\right)}{\partial e_1 e_1'}\bigg|_{e_1 = e_1' = 0} \\[2mm] & + \frac{1}{2!} \frac{\partial f\left(e_1, e_1'\right)}{\partial e_1' e_1}\bigg|_{e_1 = e_1' = 0} + \cdots \end{aligned}$$

Expressing (3.1) in terms of e's and using (3.5) for the exponential term, we have

(3.6) $\qquad \overline{y}_{NH} = \left[k_1 \overline{Y}(1 + e_0) + k_2 \overline{X}\left(e_1' - e_1\right)\right] \exp\left\{ \dfrac{a\overline{X}\left(e_1' - e_1\right)}{a\overline{X}\left(e_1 + e_1' + 2\right) + 2b} \right\}$

where $f(e_1, e_1') = \exp\left\{\dfrac{a\overline{X}\left(e_1' - e_1\right)}{a\overline{X}\left(e_1 + e_1' + 2\right) + 2b}\right\}$ and we solve the exponential term from

(3.5) as

$$\overline{y}_{NH} = \left[k_1\overline{Y}\left(1 + e_0\right) + k_2\overline{X}\left(e_1' - e_1\right)\right]$$
$$\left\{1 - \theta\left(e_1 - e_1'\right) + \frac{3\theta^2}{2}e_1^2 - \frac{\theta^2}{2}e_1'^{\,2} - \theta^2 e_1 e_1' + ...\right\}$$

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

Assuming $|e_1| < 1$, expanding the right hand side of (3.6), and retaining terms up to the second degree of e's, we have

$$(3.7) \quad \overline{y}_{NH} - \overline{Y} \cong \overline{Y}\left[(k_1 - 1) - k_1\theta\left(e_1 - e_1'\right) - \frac{3\theta^2}{2}\left(e_1^2 - e_1'^{\,2}\right) + k_1 e_0 - \right.$$
$$\left. k_1\theta\left(e_0 e_1 - e_0 e_1'\right)\right] + k_2\overline{X}\left[e_1' - e_1 + \theta\left(e_1^2 - e_1'^{\,2}\right)\right]$$

Squaring both sides of (3.7), retaining terms of e's up to the second degree and taking expectation, we get the Bias and MSE Equations of $\overline{y}_{NH}$ to the second degree of approximation as

$$(3.8) \quad Bias\left(\overline{y}_{NH}\right) \cong E\left(\overline{y}_{NH} - \overline{Y}\right) \cong \overline{Y}\left[(k_1 - 1) + k_1\lambda^*\theta C_x^2\right]\left(\frac{3\theta}{2} - K_{yx}\right)$$

$$(3.9) \quad MSE\left(\overline{y}_{NH}\right) \cong E\left(\overline{y}_{NH} - \overline{Y}\right)^2$$
$$\cong \bar{Y}^2\left[(k_1 - 1)^2 + k_1^2\left\{\lambda C_y^2 + 4\lambda^*\theta C_x^2\left(\theta - K_{yx}\right)\right\} + \right.$$
$$k_1\lambda^*\theta C_x^2\left(2K_{yx} - 3\theta\right) + k_2^2\lambda^*\overline{X}^2 C_x^2 + 2k_2\overline{X}\,\overline{Y}\lambda^* C_x^2\left\{k_1\left(2\theta - K_{yx}\right)\right\}$$

To obtain the minimum $MSE\left(\overline{y}_{NH}\right)$, we get

$$(3.10) \quad \frac{\partial}{\partial k_i}\left\{MSE\left(\overline{y}_{NH}\right)\right\} = 0;\, i = 1,\, 2.$$

Solving two equations simultaneously, the optimum values of k_1 and k_2 are respectively,

$$(3.11) \quad k_1 = 1 - \frac{2 - \lambda^*\theta^2 C_x^2}{1 + \left(\lambda - \lambda^*\rho_{yx}^2\right)}$$

$$(3.12) \quad k_2 = \frac{\overline{Y}}{\overline{X}}\left\{(\theta - 1) + \frac{2 - \lambda^*\theta^2 C_x^2}{1 + \left(\lambda - \lambda^*\rho_{yx}^2\right)}\left(2\theta - K_{yx}\right)\right\}$$

k_1 and k_2 quantities can be guessed quite accurately through a pilot sample survey or sample data or experience gathered in due course of time, see Das and Tripathi [5], Singh and Ruiz-Espejo [16], Singh, H.P. et al. [18] and Koyuncu and Kadilar [11].

When k_1 and k_2 are replaced in (3.9), the minimum MSE of the proposed estimator can be written as

$$(3.13) \quad MSE_{\min}\left(\overline{y}_{NH}\right) \cong \overline{Y}^2 \frac{C_y^2\left(\lambda - \lambda^* \rho_{yx}^2\right)\left(1 - \lambda^* \theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{\left\{1 + C_y^2\left(\lambda - \lambda^* \rho_{yx}^2\right)\right\}}$$

$$\cong \overline{Y}^2 \frac{Var\left(\overline{y}_{lr}\right)\left(1 - \lambda^* \theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\overline{Y}^2 \theta^4 C_x^4}{4}}{\left\{\overline{Y}^2 + Var\left(\overline{y}_{lr}\right)\right\}}$$

Table 1. Some Members of the Suggested Estimator $\overline{y}_{NH}$

A subset of $\overline{y}_{NH}$	**a**	**b**
$\overline{y}_{NH1} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left(\dfrac{\overline{x}' - \overline{x}}{\overline{x} + \overline{x}'}\right)$	1	0
$\overline{y}_{NH2} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left(\dfrac{\overline{x}' - \overline{x}}{\overline{x} + \overline{x}' + 2}\right)$	1	1
$\overline{y}_{NH3} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left(\dfrac{\overline{x}' - \overline{x}}{\overline{x} + \overline{x}' + 2\beta_2(x)}\right)$	1	$\beta_2(x)$
$\overline{y}_{NH4} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\dfrac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x} + \overline{x}'\right) + 2}\right\}$	$\beta_2(x)$	1
$\overline{y}_{NH5} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\dfrac{C_x\left(\overline{x}' - \overline{x}\right)}{C_x\left(\overline{x} + \overline{x}'\right) + 2\beta_2(x)}\right\}$	C_x	$\beta_2(x)$
$\overline{y}_{NH6} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\dfrac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x} + \overline{x}'\right) + 2C_x}\right\}$	$\beta_2(x)$	C_x
$\overline{y}_{NH7} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\dfrac{\rho_{yx}\left(\overline{x}' - \overline{x}\right)}{\rho_{yx}\left(\overline{x} + \overline{x}'\right) + 2\beta_2(x)}\right\}$	ρ_{yx}	$\beta_2(x)$
$\overline{y}_{NH8} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\dfrac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x} + \overline{x}'\right) + 2\rho_{yx}}\right\}$	$\beta_2(x)$	ρ_{yx}
$\overline{y}_{NH9} = \left[k_1 \overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\dfrac{C_x\left(\overline{x}' - \overline{x}\right)}{C_x\left(\overline{x} + \overline{x}'\right) + 2\rho_{yx}}\right\}$	C_x	ρ_{yx}

4. Efficiency Comparisons in Two Phase Sampling

In this section, we obtain the efficiency conditions for the proposed estimator by comparing the MSE of the proposed estimators with the MSE of classical ratio and regression estimators and the exponential ratio estimator suggested by Singh and Vishwakarma [17].

We compare the MSE of the proposed estimator, $\overline{y}_{NH}$, given in (3.13), with the MSE of the existing estimators, $\overline{y}_R, \overline{y}_{lr}, \overline{y}_{svr}$.

From (2.2) and (3.13), we have the condition

$$MSE\left(\overline{y}_{NH}\right) < MSE\left(\overline{y}_R\right)$$

$$\overline{Y}^2 \frac{C_y^2\left(\lambda - \lambda^* \rho_{yx}^2\right)\left(1 - \lambda^* \theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{1 + C_y^2\left(\lambda - \lambda^* \rho_{yx}^2\right)} < \overline{Y}^2\left[\lambda C_y^2 + \lambda^* C_x^2\left(1 - 2K_{yx}\right)\right]$$

Table 1 Continued: Some Members of the Suggested Estimator $\overline{y}_{NH}$

A subset of $\overline{y}_{NH}$	a	b
$\overline{y}_{NH10} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\rho_{yx}\left(\overline{x}' - \overline{x}\right)}{\rho_{yx}\left(\overline{x}+\overline{x}'\right)+2C_x}\right\}$	ρ_{yx}	C_x
$\overline{y}_{NH11} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\sigma_x\left(\overline{x}' - \overline{x}\right)}{\sigma_x\left(\overline{x}+\overline{x}'\right)+2\rho_{yx}}\right\}$	σ_x	ρ_{yx}
$\overline{y}_{NH12} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\rho_{yx}\left(\overline{x}' - \overline{x}\right)}{\rho_{yx}\left(\overline{x}+\overline{x}'\right)+2\sigma_x}\right\}$	ρ_{yx}	σ_x
$\overline{y}_{NH13} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2\sigma_x}\right\}$	$\beta_2\left(x\right)$	σ_x
$\overline{y}_{NH14} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\sigma_x\left(\overline{x}' - \overline{x}\right)}{\sigma_x\left(\overline{x}+\overline{x}'\right)+2\beta_2(x)}\right\}$	σ_x	$\beta_2\left(x\right)$
$\overline{y}_{NH15} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\beta_1(x)\left(\overline{x}' - \overline{x}\right)}{\beta_1(x)\left(\overline{x}+\overline{x}'\right)+2\beta_2(x)}\right\}$	$\beta_1\left(x\right)$	$\beta_2\left(x\right)$
$\overline{y}_{NH16} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left\{\frac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2\beta_1(x)}\right\}$	$\beta_2\left(x\right)$	$\beta_1\left(x\right)$

Note: In addition to estimators listed in Table 1, a large number of estimators can also be generated from (3.1) by putting 1, C_x, $\beta_2\left(x\right)$, ρ_{yx}, σ_x, $\beta_1\left(x\right)$ values for a and b.

$$\frac{C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)\left(1 - \lambda^*\theta^2 C_x^2\right) - \frac{\lambda^{*2}\theta^4 C_x^4}{4}}{1 + C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)} < \lambda C_y^2 + \lambda^*\left(C_x - \rho_{yx}C_y\right)^2 - \lambda^*\rho_{yx}C_y$$

$$(4.1) \qquad \left\{\frac{\lambda^*\theta^2 C_x^2}{2} + \frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}\right\}^2 + \lambda^*\left(C_x - \rho_{yx}C_y\right)^2\left\{1 + \frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}\right\} > 0$$

The condition (4.1) is always satisfied, the proposed estimator, $\overline{y}_{NH}$, is always more efficient than the classical ratio estimator, $\overline{y}_R$.

From (2.4) and (3.13), we have the condition

$$MSE\left(\overline{y}_{NH}\right) < Var\left(\overline{y}_{lr}\right)$$

$$\overline{Y}^2\frac{Var\left(\overline{y}_{lr}\right)\left(1 - \lambda^*\theta^2 C_x^2\right) - \frac{\lambda^{*2}\overline{Y}^2\theta^4 C_x^4}{4}}{\overline{Y}^2 + Var\left(\overline{y}_{lr}\right)} < Var\left(\overline{y}_{lr}\right)$$

$$(4.2) \qquad \left\{\frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2} + \frac{\lambda^*\theta^2 C_x^2}{2}\right\}^2 > 0$$

The condition (4.2) is always satisfied, the proposed estimator, $\overline{y}_{NH}$, is always more efficient than the classical regression estimator, $\overline{y}_{lr}$.

From (2.6) and (3.13), we have the condition

$$MSE\left(\overline{y}_{NH}\right) < MSE\left(\overline{y}_{svr}\right)$$

$$\overline{Y}^2 \frac{C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)\left(1 - \lambda^*\theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{1 + C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)} < \overline{Y}^2\left[\lambda C_y^2 + \lambda^*\left(\frac{C_x^2}{4} - \rho_{yx}C_yC_x\right)\right]$$

$$\frac{C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)\left(1 - \lambda^*\theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{1 + C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)} < \lambda C_y^2 + \lambda^*\left(\frac{C_x}{2} - \rho_{yx}C_y\right) - \lambda^*\rho_{yx}^2 C_y^2$$

$$(4.3) \qquad \left(\frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2} + \lambda^*\theta^2 C_x^2\right)^2 + \lambda^*\left(\frac{C_x}{2} - \rho_{yx}C_y\right)^2\left\{1 + \frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}\right\} > 0$$

The condition (4.3) is always satisfied, the proposed estimator, $\overline{y}_{NH}$, is always more efficient than Singh and Vishwakarma [17] exponential ratio estimator, $\overline{y}_{svr}$.

Thus, finally, we conclude from the efficiency comparisons that the class of exponential regression cum ratio estimator, $\overline{y}_{NH}$, is always more efficient than the estimators, $\overline{y}_R$, $\overline{y}_{lr}$ and $\overline{y}_{svr}$.

5. Numerical Example

To show the performance of the proposed estimator in comparison to other estimators in two phase sampling, four original data sets used by other authors in literature has been considered. The descriptions of the populations are given below.

Population I : Cingi et. al. [3],
   y : the number of teachers
   x : the number of student in both primary and secondary school for 923 districts
   $N = 923$, $n^{'} = 400$, $n = 200$, $\overline{Y} = 436,3$, $\overline{X} = 11440,50$, $C_y = 1,72$, $C_x = 1,86$, $\rho_{yx} = 0,955$.

Population II : Sukhatme and Sukhatme [20],
   y: No. of villages in the circle.
   x: A circle consisting more than five villages.
   $N = 89$, $n^{'} = 30$, $n = 20$, $\overline{Y} = 3,360$, $\overline{X} = 0,124$, $C_y = 0,604$, $C_x = 2,190$, $\rho_{yx} = 0,766$.

Population III : Kadilar and Cingi [9],
   y: Level of apple production.
   x: No. of apple trees.
   $N = 104$, $n^{'} = 40$, $n = 20$, $\overline{Y} = 625,37$, $\overline{X} = 13,930$, $C_y = 1,866$, $C_x = 1,653$, $\rho_{yx} = 0,865$.

Population IV : Murthy [12],
   y: Output
   x: fixed capital
   $N = 80$, $n^{'} = 40$, $n = 20$, $\overline{Y} = 51,826$, $\overline{X} = 11,265$, $C_y = 0,354$, $C_x = 0,751$, $\rho_{yx} = 0,9413$.

We compute the MSE values of classical ratio and regression estimators, Singh and Vishwakarma [17] estimator and proposed estimator using the equations, (2.2), (2.4), (2.6), and (3.13), respectively. We have taken $a = b = 1$, that is, $\theta = \dfrac{\overline{X}}{2\left(\overline{X}+1\right)}$, just for the sake of simplicity.

These MSE values are shown in Table 2. We observe that the most efficient estimator is the proposed exponential regression cum ratio estimator as compared to those existing ones.

Table 2. MSE Values of Estimators in Two Phase Sampling

Estimators	Population			
	I	**II**	**III**	**IV**
Classical Ratio($\bar{y}_R$)	807,59	0,30	54993,75	12,64
Classical Regression($\bar{y}_{lr}$)	780,89	1,86	29536,17	16,87
Singh and Vishwakarma ($\bar{y}_{svr}$)	1045,59	0,40	35586,14	5,29
Proposed Est.($\bar{y}_{NH}$)	**774,71**	**0,12**	**26960,89**	**5,12**

6. Conclusion

We propose a class of regression cum estimator using the exponential function for the population mean in two phase sampling improving the exponential ratio estimator suggested in Singh and Vishwakarma [17]. Theoretically, we demonstrate that the proposed estimator is always the most efficient estimator in two phase sampling and numerically, for various specific data sets, we show that the proposed estimator has small MSE value according to other estimators. In future work, we will improve the proposed estimator, presented here, with using several auxiliary variables and adding more parameters for other sampling schemes.

References

[1] Ahmad, Z. *Generalized Multivariate Ratio and Regression Estimators for Multi-phase Sampling* (PhD thesis, National College of Business Administration and Economics, Lahore, Pakistan, 2008).

[2] Bahl, S. and Tuteja, R.K. *Ratio and product exponential estimators*, Journal of Information and Optimization Sciences, **12** (1), 159-164, 1991.

[3] Cingi, H., Kadilar, C. and Kocberber, G. *Determination of the Opportunities of the Primary and the Secondary Schools in Turkey and Suggestions to Solve the Determined Problems* (TUBITAK, SOBAG106K077.(in Turkish), 2007).

[4] Cingi, H. and Kadilar, C. *Advances in Sampling Theory-Ratio Method of Estimation* (Bentham Science Publishers, 2009).

[5] Das, A.K. and Tripathi, T.P. *Use of auxiliary information in estimating the finite population variance*, Sankhya C **40** (2): 139-148, 1978.

[6] Grover, L.K. and Kaur, P. *An improved estimator of the finite population mean in simple random sampling*, Model Assisted Statistics and Applications **6** (1), 47-55, 2011.

[7] Hanif, M., Shahbaz, M.Q. and Ahmad, Z. *Some Improved Estimators in Multiphase Sampling*, Pak. J. Statist. **26** (1), 195-202, 2010.

[8] Kadilar, C. and Cingi, H. *Ratio Estimators in Stratified Random Sampling*, Biometrical Journal **45** (2), 218-225, 2003.

[9] Kadilar, C. and Cingi, H. *An Improvement in Estimating The Population Mean By Using The Correlation Coefficient*, Hacettepe Journal of Mathematics and Statistics **35** (1), 103-109, 2006.

[10] Keen, K. J. *Two-Phase Sampling*, Encyclopedia of Biostatistics, 2005.

[11] Koyuncu, N. and Kadilar, C. *On the family of estimators of population mean in stratified random sampling*, Pak. Jour. Statist. **26** (2), 427-443, 2010.

[12] Murthy, M.N. *Sampling Theory and methods* (Statistical Publishing Society, Calcutta, India, 1967).

[13] Neyman, J. *Contribution to the theory of sampling human populations*, Journal of the American Statistical Association **33**, 101-116, 1938.

[14] Ozgul, N. *Mean Estimators in Two Phase Sampling* (MSc.Thesis (in Turkish), Department of Statistics, Hacettepe University, Ankara, Turkey, 2007).

[15] Samiuddin, M. and Hanif, M. *Estimation of population mean in single and two phase sampling with or without additional information*, Pak. J. Stat. **23** (2), 99-118, 2007.

[16] Singh, H.P. and Ruiz-Espejo, M. *On linear regression and ratio-product estimation of a finite population mean*, The Statistician **52** (1), 59-67, 2003.

[17] Singh, H.P. and Vishwakarma, G.K. *Modified exponential ratio and product estimators for finite population mean in two phase sampling*, Austrian Journal of Statistics **36** (3), 217-225, 2007.

[18] Singh, H.P., Tailor, R. and Tailor, R. O*n ratio and product methods with certain known population parameters of auxiliary variable in sample surveys*, SORT **34** (2), 157-180, 2010.

[19] Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. *Improvement in estimating the population mean using exponential estimator in simple random sampling*, Bulletin of Statistics and Economics **3** (9), 13-18, 2009.

[20] Sukhatme, P.V. and Sukhatme, B.V. *Sampling theory of surveys with applications (Second Edition)* (Asia Publishing House, Bombay, India, 1970).