AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 6(5): 4265–4279.
DOI:10.3934/math.2021252
Received: 31 December 2020
Accepted: 01 February 2021
Published: 07 February 2021

Research article

Exponential ratio and product type estimators of the mean in stratified two-phase sampling

Tolga Zaman[1,*] **and Cem Kadilar**[2]

[1] Department of Statistics, Çankırı Karatekin University, Çankırı, Turkey
[2] Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey

* **Correspondence:** Email: tolgazaman@karatekin.edu.tr; Tel: +9003762189540.

Abstract: In this study, we propose exponential ratio estimators in the stratified two-phase sampling utilizing an auxiliary attribute. The expressions for the mean squared error of these exponential-type estimators under two different cases are derived and theoretical comparisons are made with competing estimators. We show that the proposed estimators have a lower mean square error than the simple mean estimator, usual stratified two-phase sampling ratio, and product estimators, usual exponential ratio and product estimators for the stratified two-phase sampling under the obtained conditions in theory. In addition, these theoretical results are supported with the aid of a numerical example.

Keywords: stratified two-phase sampling; exponential ratio-product estimators; mean square error; efficiency.
Mathematics Subject Classification: 62D05

1. Introduction

Stratification improves the efficiency when the variance between strata is much larger than the variances within strata. It is used to estimate the population mean of the study variable using ratio or product estimator for positive or negative correlation, respectively, between the study variable and the auxiliary attribute. The auxiliary information is available in the form of an attribute. Some of those can be given as follows: The height of a person may depend on the fact that whether the person is male or female, the efficiency of a dog may depend on the particular breed of that dog, or the yield of wheat crop produced may depend on a particular variety of wheat, etc. [1]. There are some recent studies on the estimators using the information of the auxiliary variable under the stratified two-phase sampling in literature, such as Sahoo et al. [2], Samiuddin and Hanif [3], Singh and Vishwakarma [4], Singh and Choudhury [5], Choudhury and Singh [6], Hamad et al. [7], Sanaullah et al. [8], Malik and Singh [9], and Shabbir and Gupta [10]. The aim of this paper is to suggest the efficient exponential type estimators in the stratified two-phase sampling design. In addition, numerous of literature is available on exponential ratio-type estimators and stratified sampling for mean estimation, for examples, Vishwakarma and Kumar [11] provided families of separate and combined ratio-product estimators for estimating the mean in stratified random sampling. Pal and Singh [12] proposed a class of ratio-cum-ratio type exponential estimators for population mean along with its properties. Singh et al. [13] introduced some imputation methods to compensate the missing data in two-phase sampling. Saleem et al. [14] used the quantitative scrambled randomized-response model by measuring the efficiency of an estimator of the population mean or of the protection of privacy under two-stage sampling. Ünal and Kadilar [15] provided classes of estimators utilizing the exponential function for the population mean in the case of non-response in two different cases. Chaudhary et al. [16] provided a class of combined-type estimators of the population mean in stratified random sampling using the information on an auxiliary variable under the situation in which non-response is observed on both study and auxiliary variables. Irfan et al. [17] evaluated the performance of difference-type-exponential estimators based on dual auxiliary information for population mean. Ünal and Kadilar [18] adapted the estimator based on the exponential function for the estimation of the population mean in the presence of non-response on both the study and the auxiliary variables. Muneer et al. [19] proposed a parent-generalized class of chain exponential ratio type estimators in stratified random sampling to estimate the finite population mean utilizing known information on two supplementary variables. Sanaullah et al. [20] introduced a generalized randomized response technique model and use it to develop some exponential estimators in two-phase sampling. Sinha et al. [21] introduced the problem of estimating the product of two population means using the information of auxiliary character under two-phase sampling the non-respondents. Zaman [22] provided an efficient exponential ratio estimator that allows estimating the population mean in stratified random sampling using an auxiliary variable. Hassan et al. [23] provided the combination of exponential and ln ratio type estimator to estimate the mean of study variable by incorporating two auxiliary variables in two phase sampling design. Consider a finite population $U = \{1, 2, \ldots, N\}$ of N identifiable units divided into L strata with the hth stratum ($h = 1, 2, \ldots, L$) having N_h units such that $\sum_{h=1}^{L} N_h = N$. Let y_{hi} and p_{hi} be the values of the study variable (y) and the auxiliary attribute (p), respectively, for the ith ($i = 1, 2, \ldots, N_h$) population element of the hth stratum. Let $\bar{Y} = \sum_{h=1}^{L} \omega_h \bar{Y}_h$ and $P = \sum_{h=1}^{L} \omega_h P_h$ be populations means of the study and the auxiliary attribute, respectively, and let $\bar{Y}_h = \sum_{i=1}^{N_h} \frac{y_{hi}}{N_h}$, $P_h = \sum_{i=1}^{N_h} \frac{p_{hi}}{N_h}$be the hth population stratum means. When the information on P_h is not known, a first phase large sample of size $n_h^{'}(< N_h)$ is selected from each hth stratum to estimate P_h and the replace it with its unbiased estimator. To obtain the MSE equation under the stratified two-phase sampling, let us define $\Delta_{0st} = \frac{\bar{y}_{st}-\bar{Y}}{\bar{Y}}$, $\Delta_{1st} = \frac{p_{st}-P}{P}$, $\Delta_{1st}^{'} = \frac{p_{st}^{'}-P}{P}$, such that $E(\Delta_{ist}) = 0; (i = 1, 2)$ and $E(\Delta_{1st}^{'}) = 0$, $V_{rs} = \sum_{h=1}^{L} \omega_h^{r+t} \frac{E[(\bar{y}_h-\bar{Y}_h)^r (p_h-P_h)^s]}{\bar{Y}^r P^s}$ and $V_{rs}^{'} = \sum_{h=1}^{L} \omega_h^{r+t} \frac{E[(\bar{y}_h-\bar{Y}_h)^r (p_h^{'}-P_h)^s]}{\bar{Y}^r P^s}$, $E(\Delta_{0st}^2) = \sum_{h=1}^{L} \omega_h^2 \theta_h C_{yh}^2 = V_{20}$, $E(\Delta_{1st}^2) = \sum_{h=1}^{L} \omega_h^2 \theta_h C_{ph}^2 = V_{02}$, $E(\Delta_{1st}^{'2}) = E(\Delta_{1st}\Delta_{1st}^{'}) = \sum_{h=1}^{L} \omega_h^2 \theta_h^{'} C_{ph}^2 = V_{02}^{'}$, $E(\Delta_{0st}\Delta_{1st}) = \sum_{h=1}^{L} \omega_h^2 \theta_h C_{yph} = V_{11}$, $E(\Delta_{0st}\Delta_{1st}^{'}) = \sum_{h=1}^{L} \omega_h^2 \theta_h^{'} C_{yph} = V_{11}^{'}$, where $\theta_h = (\frac{1}{n_h} - \frac{1}{N_h})$ and $\theta_h^{'} = (\frac{1}{n_h^{'}} - \frac{1}{N_h})$. The procedure of the stratified two-phase sampling is given as: i) Select a sample size $n_h^{'}$ from the hth stratum using the simple random sampling without replacement (SRSWOR) sampling scheme such that $\sum_{h=1}^{L} n_h^{'} = n^{'}$ and $p_h^{'} = \frac{1}{n_h^{'}} \sum_{i=1}^{n_h^{'}} p_{hi}$. This is called a stratified first-phase sample. ii) Select another stratified random sample of size n_h from each $n_h^{'}(n_h < n_h^{'})$ using the SRSWOR such that $\sum_{h=1}^{L} n_h = n$ and collect the information on the study variable. This is called a second-phase sample. In the stratified random sampling, the sample mean estimator is as follows:

$$\bar{y}_{st} = \sum_{h=1}^{L} \omega_h \bar{y}_h \tag{1.1}$$

The variance is given by

$$Var(\bar{y}_{st}) = \bar{Y}^2 \sum_{h=1}^{L} \omega_h^2 \theta_h C_{yh}^2 = \bar{Y}^2 V_{20} \tag{1.2}$$

Two cases will be investigated for the selection of the needed sample as follows: Case I. when the second-phase sample of size n is a sub-sample of the first-phase sample of size $n^{'}$, and Case II. when the second-phase sample of size n is drawn independently of the first-phase sample of size $n^{'}$, see Bose [24]. The classical ratio estimator in the stratified two-phase sampling is

$$\bar{y}_{Rd} = \bar{y}_{st} \frac{p_{st}^{'}}{p_{st}} = \sum_{h=1}^{L} \omega_h \bar{y}_h \left(\frac{\sum_{h=1}^{L} \omega_h p_{st}^{'}}{\sum_{h=1}^{L} \omega_h p_{st}}\right) \tag{1.3}$$

The MSE equations of the estimator in (1.3) are

$$MSE(\bar{y}_{Rd})_I \cong \bar{Y}^2[V_{20} + (V_{02} - V_{02}^{'}) - 2(V_{11} - V_{11}^{'})], \tag{1.4}$$

$$MSE(\bar{y}_{Rd})_{II} \cong \bar{Y}^2[V_{20} + (V_{02} + V_{02}^{'}) - 2V_{11}], \tag{1.5}$$

for the, Case I and Case II, respectively.
Similarly, the classical product estimator in the stratified two-phase sampling is

$$\bar{y}_{Pd} = \bar{y}_{st} \frac{p_{st}}{p_{st}^{'}} = \sum_{h=1}^{L} \omega_h \bar{y}_h \left(\frac{\sum_{h=1}^{L} \omega_h p_{st}}{\sum_{h=1}^{L} \omega_h p_{st}^{'}}\right). \tag{1.6}$$

The MSE equations of the estimator in (1.6) are

$$MSE(\bar{y}_{Pd})_I \cong \bar{Y}^2[V_{20} + (V_{02} - V_{02}^{'}) + 2(V_{11} - V_{11}^{'})], \tag{1.7}$$

$$MSE(\bar{y}_{Pd})_{II} \cong \bar{Y}^2[V_{20} + (V_{02} + V_{02}^{'}) + 2V_{11}], \tag{1.8}$$

for the Case-I and the Case-II, respectively.
The classical exponential ratio estimator in the stratified two-phase sampling is as follows:

$$\bar{y}_{Exd} = \bar{y}_{st}exp(\frac{p'_{st} - p_{st}}{p'_{st} + p_{st}}). \tag{1.9}$$

The MSE equations of the estimator in (1.9) are

$$MSE(\bar{y}_{Exd})_I \cong \bar{Y}^2[V_{20} + \frac{1}{4}(V_{02} - V'_{02}) - (V_{11} - V'_{11})], \tag{1.10}$$

$$MSE(\bar{y}_{Exd})_{II} \cong \bar{Y}^2[V_{20} + \frac{1}{4}(V_{02} + V'_{02}) - V_{11}], \tag{1.11}$$

for the Case-I and the Case-II, respectively.
The classical exponential product estimator in the stratified two-phase sampling is as follows:

$$\bar{y}_{Exp} = \bar{y}_{st}exp(\frac{p_{st} - p'_{st}}{p_{st} + p'_{st}}). \tag{1.12}$$

The MSE equations of the estimator in (1.12) are

$$MSE(\bar{y}_{Exp})_I \cong \bar{Y}^2[V_{20} + \frac{1}{4}(V_{02} - V'_{02}) + (V_{11} - V'_{11})], \tag{1.13}$$

$$MSE(\bar{y}_{Exp})_{II} \cong \bar{Y}^2[V_{20} + \frac{1}{4}(V_{02} + V'_{02}) + V_{11}], \tag{1.14}$$

for the Case-I and the Case-II, respectively.

2. Materails and methods

Considering the development of a new family of estimators as in Zaman and Kadilar [25] for the classical exponential ratio estimator, given in (1.9), we propose the following families of ratio and product exponential estimators under the stratified two-phase sampling, respectively, to estimate the mean of the variable of study utilizing the information about the proportion of the population possessing the certain attributes as follows:

$$\bar{y}_{stZKi} = \bar{y}_{st}exp[\frac{(\psi p'_{st} + \varsigma) - (\psi p_{st} + \varsigma)}{(\psi p'_{st} + l) + (\psi p_{st} + l)}], \tag{2.1}$$

$$\bar{t}_{stZKi} = \bar{y}_{st}exp[\frac{(\psi p_{st} + \varsigma) - (\psi p'_{st} + \varsigma)}{(\psi p_{st} + \varsigma) + (\psi p'_{st} + \varsigma)}], \tag{2.2}$$

Table 1. Suggested ratio estimators.

Estimators	Values of ψ	Values of ς
$\bar{y}_{stZK1} = \bar{y}_{st}exp[\frac{p'_{st}-p_{st}}{p'_{st}+p_{st}+2\beta_2(\varphi)}]$	1	$\beta_2(\varphi)$
$\bar{y}_{stZK2} = \bar{y}_{st}exp[\frac{p'_{st}-p_{st}}{p'_{st}+p_{st}+2C_p}]$	1	C_p
$\bar{y}_{stZK3} = \bar{y}_{st}exp[\frac{p'_{st}-p_{st}}{p'_{st}+p+2\rho_{pb}}]$	1	ρ_{pb}
$\bar{y}_{stZK4} = \bar{y}_{st}exp[\frac{\beta_2(\varphi)(p'_{st}-p_{st})}{\beta_2(\varphi)(p'_{st}+p)+2C_p}]$	$\beta_2(\varphi)$	C_p
$\bar{y}_{stZK5} = \bar{y}_{st}exp[\frac{C_p(p'_{st}-p_{st})}{C_p(p'_{st}+p)+2\beta_2(\varphi)}]$	C_p	$\beta_2(\varphi)$
$\bar{y}_{stZK6} = \bar{y}_{st}exp[\frac{C_p(p'_{st}-p_{st})}{C_p(p'_{st}+p_{st})+2\rho_{pb}}]$	C_p	ρ_{pb}
$\bar{y}_{stZK7} = \bar{y}_{st}exp[\frac{\rho_{pb}(p'_{st}-p_{st})}{\rho_{pb}(p'_{st}+p_{st})+2C_p}]$	ρ_{pb}	C_p
$\bar{y}_{stZK8} = \bar{y}_{st}exp[\frac{\beta_2(\varphi)(p'_{st}-p_{st})}{\beta_2(\varphi)(p'_{st}+p_{st})+2\rho_{pb}}]$	$\beta_2(\varphi)$	ρ_{pb}
$\bar{y}_{stZK9} = \bar{y}_{st}exp[\frac{\rho_{pb}(p'_{st}-p_{st})}{\rho_{pb}(p'_{st}+p_{st})+2\beta_2(\varphi)}]$	ρ_{pb}	$\beta_2(\varphi)$

where $\psi(\neq 0)$ and ς are either real numbers or the functions of the known parameters of the auxiliary attribute, such as C_p, $\beta_2(\varphi)$, and the known parameter of the attribute with the study variable ρ_{pb}. Examples of nine proposed estimators for the population mean, which can be computed by taking the suitable choices of constants ψ and ς, are given in Tables 1 and 2.

Case-I. To obtain the MSE of the estimators $\bar{y}_{stZK}$, let $\bar{y}_{st} = \bar{Y}(1 + \Delta_{0st})$, $p_{st} = P(1 + \Delta_{1st})$, and $p'_{st} = P(1 + \Delta'_{1st})$ such that

$E(\Delta_{0st}) = E(\Delta_{1st}) = E(\Delta'_{1st}) = 0$,

$E(\Delta^2_{0st}) = V_{20}$,$E(\Delta^2_{1st}) = V_{02}$, $E(\Delta'^2_{1st}) = E(\Delta_{1st}\Delta'_{1st}) = V'_{02}$, $E(\Delta_{0st}\Delta_{1st}) = V_{11}$, $E(\Delta_{0st}\Delta'_{1st}) = V'_{11}$.

Expressing the estimators, $\bar{y}_{stZK}$ and $\bar{Y}_{stZKi}$, in terms of Δ_{ist} $(i = 0, 1)$, we can write (2.1) and (2.2) as

$$\bar{y}_{stZKi} = \bar{Y}(1 + \Delta_{0st})exp[\lambda_i(\Delta'_{1st} - \Delta_{1st})\{1 + \lambda_i(\Delta'_{1st} + \Delta_{1st})\}^{-1}], \quad (2.3)$$

$$\bar{t}_{stZKi} = \bar{Y}(1 + \Delta_{0st})exp[-\lambda_i(\Delta'_{1st} - \Delta_{1st})\{1 + \lambda_i(\Delta'_{1st} + \Delta_{1st})\}^{-1}], \quad (2.4)$$

respectively. Expanding the right hand sides of (2.3) and (2.4), to the first order of approximation, multiplying out and neglecting the terms of $\Delta' s$ greater than or equal to two, we get

$$\bar{y}_{stZKi} \cong \bar{Y}[1 + \Delta_{0st} + \lambda_i(\Delta'_{1st} - \Delta_{1st})] \Rightarrow \bar{y}_{stZKi} - \bar{Y} = \bar{Y}[\Delta_{0st} + \lambda_i(\Delta'_{1st} - \Delta_{1st})], \quad (2.5)$$

$$\bar{t}_{stZKi} \cong \bar{Y}[1 + \Delta_{0st} - \lambda_i(\Delta'_{1st} - \Delta_{1st})] \Rightarrow \bar{Y}_{stZKi} - \bar{Y} = \bar{Y}[\Delta_{0st} - \lambda_i(\Delta'_{1st} - \Delta_{1st})]. \quad (2.6)$$

Table 2. Suggested product estimators.

Estimators	Values of	
	ψ	ς
$\bar{t}_{stZK1} = \bar{y}_{st}exp[\frac{p'_{st}-p_{st}}{p'_{st}+p_{st}+2\beta_2(\varphi)}]$	1	$\beta_2(\varphi)$
$\bar{t}_{stZK2} = \bar{y}_{st}exp[\frac{p'_{st}-p_{st}}{p'_{st}+p_{st}+2C_p}]$	1	C_p
$\bar{t}_{stZK3} = \bar{y}_{st}exp[\frac{p'_{st}-p_{st}}{p'_{st}+p+2\rho_{pb}}]$	1	ρ_{pb}
$\bar{t}_{stZK4} = \bar{y}_{st}exp[\frac{\beta_2(\varphi)(p'_{st}-p_{st})}{\beta_2(\varphi)(p'_{st}+p)+2C_p}]$	$\beta_2(\varphi)$	C_p
$\bar{t}_{stZK5} = \bar{y}_{st}exp[\frac{C_p(p'_{st}-p_{st})}{C_p(p'_{st}+p)+2\beta_2(\varphi)}]$	C_p	$\beta_2(\varphi)$
$\bar{t}_{stZK6} = \bar{y}_{st}exp[\frac{C_p(p'_{st}-p_{st})}{C_p(p'_{st}+p_{st})+2\rho_{pb}}]$	C_p	ρ_{pb}
$\bar{t}_{stZK7} = \bar{y}_{st}exp[\frac{\rho_{pb}(p'_{st}-p_{st})}{\rho_{pb}(p'_{st}+p_{st})+2C_p}]$	ρ_{pb}	C_p
$\bar{t}_{stZK8} = \bar{y}_{st}exp[\frac{\beta_2(\varphi)(p'_{st}-p_{st})}{\beta_2(\varphi)(p'_{st}+p_{st})+2\rho_{pb}}]$	$\beta_2(\varphi)$	ρ_{pb}
$\bar{t}_{stZK9} = \bar{y}_{st}exp[\frac{\rho_{pb}(p'_{st}-p_{st})}{\rho_{pb}(p'_{st}+p_{st})+2\beta_2(\varphi)}]$	ρ_{pb}	$\beta_2(\varphi)$

Squaring both sides of (2.5) and (2.6) and taking the expectation, we get the MSE of the proposed families of estimators in (2.1) and (2.2) as follows:

$$MSE(\bar{y}_{stZKi})_I \cong \bar{Y}^2[V_{20} + \lambda_i^2(V_{02} - V'_{02}) - 2\lambda_i(V_{11} - V'_{11})], \tag{2.7}$$

$$MSE(\bar{t}_{stZKi})_I \cong \bar{Y}^2[V_{20} + \lambda_i^2(V_{02} - V'_{02}) + 2\lambda_i(V_{11} - V'_{11})], \tag{2.8}$$

where $\lambda_1 = \frac{P}{2(P+\beta_2(\varphi))}$; $\lambda_2 = \frac{P}{2(P+C_p)}$; $\lambda_3 = \frac{P}{2(P+\rho_{pb})}$; $\lambda_4 = \frac{\beta_2(\varphi)P}{2(\beta_2(\varphi)P+C_p)}$; $\lambda_5 = \frac{C_pP}{2(C_pP+\beta_2(\varphi))}$; $\lambda_6 = \frac{C_pP}{2(C_pP+\rho_{pb})}$; $\lambda_7 = \frac{\rho_{pb}P}{2(\rho_{pb}P+C_p)}$; $\lambda_8 = \frac{\beta_2(\varphi)P}{2(\beta_2(\varphi)P+\rho_{pb})}$; $\lambda_9 = \frac{\rho_{pb}P}{2(\rho_{pb}P+\beta_2(\varphi))}$.

Case-II. To obtain the MSE of the estimators $\bar{y}_{stZK}$, let $\bar{y}_{st} = \bar{Y}(1 + \Delta_{0st})$, $p_{st} = P(1 + \Delta_{1st})$, and $p'_{st} = P(1 + \Delta'_{1st})$ such that

$E(\Delta_{0st}) = E(\Delta_{1st}) = E(\Delta'_{1st}) = 0$,

$E(\Delta^2_{0st}) = V_{20}$, $E(\Delta^2_{1st}) = V_{02}$, $E(\Delta'^2_{1st}) = V'_{02}$, $E(\Delta_{0st}\Delta_{1st}) = V_{11}$, $E(\Delta_{0st}\Delta'_{1st}) = 0$, $E(\Delta_{1st}\Delta'_{1st})$.

Squaring both sides of (2.5) and (2.6) and taking the expectation of both sides, we get the MSE of the estimators $\bar{y}_{stZKi}$ and $\bar{Y}_{stZKi}$, respectively, as

$$MSE(\bar{y}_{stZKi})_{II} \cong \bar{Y}^2[V_{20} + \lambda_i^2(V_{02} + V'_{02}) - 2\lambda_i V_{11}], \tag{2.9}$$

$$MSE(\bar{t}_{stZKi})_{II} \cong \bar{Y}^2[V_{20} + \lambda_i^2(V_{02} + V'_{02}) + 2\lambda_i V_{11}], \tag{2.10}$$

where $\lambda_1 = \frac{P}{2(P+\beta_2(\varphi))}$; $\lambda_2 = \frac{P}{2(P+C_p)}$; $\lambda_3 = \frac{P}{2(P+\rho_{pb})}$; $\lambda_4 = \frac{\beta_2(\varphi)P}{2(\beta_2(\varphi)P+C_p)}$; $\lambda_5 = \frac{C_pP}{2(C_pP+\beta_2(\varphi))}$; $\lambda_6 = \frac{C_pP}{2(C_pP+\rho_{pb})}$; $\lambda_7 = \frac{\rho_{pb}P}{2(\rho_{pb}P+C_p)}$; $\lambda_8 = \frac{\beta_2(\varphi)P}{2(\beta_2(\varphi)P+\rho_{pb})}$; $\lambda_9 = \frac{\rho_{pb}P}{2(\rho_{pb}P+\beta_2(\varphi))}$.

3. Results

We compare the proposed estimators with other competing estimators in two different cases in the stratified two-phase sampling as the following subsections.

3.1. For Case-I and the proposed ratio estimators

(i) With the sample mean estimator,

By (1.2) and (2.9)

$$MSE(\bar{y}_{stZKi})_I < Var(\bar{y}_{st}); i = 1, 2, \ldots, 9$$

$$\begin{aligned} \bar{Y}^2[V_{20} + \lambda_i^2(V_{02} - V_{02}^{'}) - 2\lambda_i(V_{11} - V_{11}^{'})] &< \bar{Y}^2 V_{20}, \\ \lambda_i^2(V_{02} - V_{02}^{'}) - 2\lambda_i(V_{11} - V_{11}^{'}) &< 0, \\ \lambda_i[\lambda_i(V_{02} - V_{02}^{'}) - 2(V_{11} - V_{11}^{'})] &< 0. \end{aligned}$$

For $\lambda_i > 0$,

$$\lambda_i(V_{02} - V_{02}^{'}) - 2(V_{11} - V_{11}^{'}) < 0,$$

$$\lambda_i < \frac{2(V_{11} - V_{11}^{'})}{(V_{02} - V_{02}^{'})}. \tag{3.1}$$

Similarly, for $\lambda_i < 0$,

$$\lambda_i > \frac{2(V_{11} - V_{11}^{'})}{(V_{02} - V_{02}^{'})}. \tag{3.2}$$

When the condition (3.1) or (3.2) is satisfied, the proposed exponential ratio estimators, given in Table 1, perform better than the sample mean.

(ii) With the classical ratio estimator

By (1.4) and (2.9)

$$MSE(\bar{y}_{stZKi})_I < MSE(\bar{y}_{Rd})_I; i = 1, 2, \ldots, 9$$

$$\bar{Y}^2[V_{20} + \lambda_i^2(V_{02} - V_{02}^{'}) - 2\lambda_i(V_{11} - V_{11}^{'})] < \bar{Y}^2[V_{20} + (V_{02} - V_{02}^{'}) - 2(V_{11} - V_{11}^{'})],$$

$$\begin{aligned}
\lambda_i^2(V_{02}-V_{02}^{'})-2\lambda_i(V_{11}-V_{11}^{'})-(V_{02}-V_{02}^{'})+2(V_{11}-V_{11}^{'}) &< 0,\\
(V_{02}-V_{02}^{'})(\lambda_i-1)(\lambda_i+1)-2(V_{11}-V_{11}^{'})(\lambda_i-1) &< 0,\\
(\lambda_i-1)[(\lambda_i+1)(V_{02}-V_{02}^{'})-2(V_{11}-V_{11}^{'})] &< 0.
\end{aligned}$$

For $\lambda_i > 1$,

$$\begin{aligned}
(\lambda_i+1)(V_{02}-V_{02}^{'})-2(V_{11}-V_{11}^{'}) &< 0,\\
(\lambda_i+1) &< \frac{2(V_{11}-V_{11}^{'})}{(V_{02}-V_{02}^{'})},
\end{aligned}$$

$$\lambda_i < \frac{2(V_{11}-V_{11}^{'})-(V_{02}-V_{02}^{'})}{(V_{02}-V_{02}^{'})}. \tag{3.3}$$

Similarly, for $\lambda_i < 1$,

$$\lambda_i > \frac{2(V_{11}-V_{11}^{'})-(V_{02}-V_{02}^{'})}{(V_{02}-V_{02}^{'})}. \tag{3.4}$$

When the condition (3.3) or (3.4) is satisfied, the proposed exponential ratio estimators, given in Table 1, perform better than the classical ratio estimator, given in (1.3).

(iii) With the classical exponential ratio estimator

By (1.10) and (2.9)

$$MSE(\bar{y}_{stZKi})_I < MSE(\bar{y}_{Exd})_I; i = 1,2,\ldots,9$$

$$\begin{aligned}
\bar{Y}^2[V_{20}+\lambda_i^2(V_{02}-V_{02}^{'})-2\lambda_i(V_{11}-V_{11}^{'})] &< \bar{Y}^2[V_{20}+\frac{1}{4}(V_{02}-V_{02}^{'})-(V_{11}-V_{11}^{'})],\\
\lambda_i^2(V_{02}-V_{02}^{'})-2\lambda_i(V_{11}-V_{11}^{'})-\frac{1}{4}(V_{02}-V_{02}^{'})+(V_{11}-V_{11}^{'}) &< 0,\\
(V_{02}-V_{02}^{'})(\lambda_i-\frac{1}{2})(\lambda_i+\frac{1}{2})-2(V_{11}-V_{11}^{'})(\lambda_i-\frac{1}{2}) &< 0,\\
(\lambda_i-\frac{1}{2})[(\lambda_i+\frac{1}{2})(V_{02}-V_{02}^{'})-2(V_{11}-V_{11}^{'})] &< 0.
\end{aligned}$$

For $\lambda_i > \frac{1}{2}$,

$$(\lambda_i+\frac{1}{2})(V_{02}-V_{02}^{'})-2(V_{11}-V_{11}^{'}) < 0,$$

$$(\lambda_i + \frac{1}{2}) \quad < \quad \frac{2(V_{11} - V_{11}^{'})}{(V_{02} - V_{02}^{'})},$$

$$\lambda_i < \frac{4(V_{11} - V_{11}^{'}) - (V_{02} - V_{02}^{'})}{2(V_{02} - V_{02}^{'})}. \tag{3.5}$$

Similarly, for $\lambda_i < \frac{1}{2}$,

$$\lambda_i > \frac{4(V_{11} - V_{11}^{'}) - (V_{02} - V_{02}^{'})}{2(V_{02} - V_{02}^{'})}. \tag{3.6}$$

When the condition (3.5) or (3.6) is satisfied, the proposed exponential ratio estimators, given in Table 1, perform better than the classical exponential ratio estimator, given in (1.9).

3.2. For Case-II and the proposed ratio estimators

(i) With the sample mean estimator,
By (1.2) and (2.11)

$$MSE(\bar{y}_{stZKi})_{II} < Var(\bar{y}_{st}); i = 1, 2, \ldots, 9$$

$$\begin{aligned}
\bar{Y}^2[V_{20} + \lambda_i^2(V_{02} + V_{02}^{'}) - 2\lambda_i V_{11}] &< \bar{Y}^2 V_{20}, \\
\lambda_i^2(V_{02} + V_{02}^{'}) - 2\lambda_i V_{11} &< 0, \\
\lambda_i[\lambda_i(V_{02} + V_{02}^{'}) - 2V_{11}] &< 0.
\end{aligned}$$

For $\lambda_i > 0$,

$$\begin{aligned}
\lambda_i(V_{02} + V_{02}^{'}) - 2V_{11} &< 0, \\
\lambda_i &< \frac{2V_{11}}{(V_{02} + V_{02}^{'})}.
\end{aligned}$$

Similarly, for $\lambda_i < 0$,

$$\lambda_i < \frac{2V_{11}}{(V_{02} + V_{02}^{'})}. \tag{3.7}$$

When the condition (3.7) or (3.8) is satisfied, the proposed exponential ratio estimators, given in Table 1, perform better than the sample mean.

(ii) With the classical ratio estimator

By (1.5) and (2.11)

$$MSE(\bar{y}_{stZKi})_{II} < MSE(\bar{y}_{Rd})_{II}; i = 1, 2, \ldots, 9$$

$$\begin{aligned}
\bar{Y}^2[V_{20} + \lambda_i^2(V_{02} + V_{02}^{'}) - 2\lambda_i V_{11}] &< \bar{Y}^2[V_{20} + (V_{02} + V_{02}^{'}) - 2V_{11}], \\
\lambda_i^2(V_{02} + V_{02}^{'}) - 2\lambda_i V_{11} - (V_{02} + V_{02}^{'}) + 2V_{11} &< 0, \\
(V_{02} + V_{02}^{'})(\lambda_i - 1)(\lambda_i + 1) - 2V_{11}(\lambda_i - 1) &< 0, \\
(\lambda_i - 1)[(\lambda_i + 1)(V_{02} + V_{02}^{'}) - 2V_{11}] &< 0.
\end{aligned}$$

For $\lambda_i > 1$

$$\begin{aligned}
(\lambda_i + 1)(V_{02} + V_{02}^{'}) - 2V_{11} &< 0, \\
(\lambda_i + 1) &< \frac{2V_{11}}{(V_{02} + V_{02}^{'})},
\end{aligned}$$

$$\lambda_i < \frac{2V_{11} - (V_{02} + V_{02}^{'})}{(V_{02} + V_{02}^{'})}. \tag{3.8}$$

Similarly, for $\lambda_i < 1$

$$\lambda_i > \frac{2V_{11} - (V_{02} + V_{02}^{'})}{(V_{02} + V_{02}^{'})}. \tag{3.9}$$

When the condition (3.9) or (3.10) is satisfied, the proposed exponential ratio estimators, given in Table 1, perform better than the classical ratio estimator, given in (1.3).

(iii) With the classical exponential ratio estimator

By (1.11) and (2.11)

$$MSE(\bar{y}_{stZKi})_{II} < MSE(\bar{y}_{Exd})_{II}; i = 1, 2, \ldots, 9$$

$$\bar{Y}^2[V_{20} + \lambda_i^2(V_{02} + V_{02}^{'}) - 2\lambda_i V_{11}] < \bar{Y}^2[V_{20} + \frac{1}{4}(V_{02} + V_{02}^{'}) - V_{11}],$$

$$\begin{aligned} \lambda_i^2(V_{02}+V_{02}^{'})-2\lambda_i V_{11}-\frac{1}{4}(V_{02}+V_{02}^{'})+V_{11} &< 0, \\ (V_{02}+V_{02}^{'})(\lambda_i-\frac{1}{2})(\lambda_i+\frac{1}{2})-2V_{11}(\lambda_i-\frac{1}{2}) &< 0, \\ (\lambda_i-\frac{1}{2})[(\lambda_i+\frac{1}{2})(V_{02}+V_{02}^{'})-2V_{11}] &< 0. \end{aligned}$$

For $\lambda_i > \frac{1}{2}$,

$$\begin{aligned} (\lambda_i+\frac{1}{2})(V_{02}+V_{02}^{'})-2V_{11} &< 0, \\ (\lambda_i+\frac{1}{2}) &< \frac{2V_{11}}{(V_{02}+V_{02}^{'})}, \end{aligned}$$

$$\lambda_i < \frac{4V_{11}-(V_{02}+V_{02}^{'})}{2(V_{02}+V_{02}^{'})}. \tag{3.10}$$

Similarly, for $\lambda_i < \frac{1}{2}$,

$$\lambda_i > \frac{4V_{11}-(V_{02}+V_{02}^{'})}{2(V_{02}+V_{02}^{'})}. \tag{3.11}$$

When the condition (3.11) or (3.12) is satisfied, the proposed exponential ratio estimators, given in Table 1, perform better than the classical exponential ratio estimator, given in (1.9).

We would like to remark that similar comparisons could be made for product estimators.

4. Discussion

We have used the data of Kadilar and Cingi [26], to examine the efficiencies of the proposed exponential estimators for the population mean in the stratified two-phase sampling. This data is defined as: Study variable is apple production amount in 1999; auxiliary attribute is the number of apple trees more than 15000 in 1999. (Source: Institute of Statistics, Republic of Turkey). We have stratified the data by regions of Turkey (as 1: Marmara 2: Agean 3: Mediterranean 4: Central Anatolia 5: Black Sea 6: East and Southeast Anatolia) and from each stratum; we have randomly selected the samples whose sizes are computed by using the Neyman allocation method. The summary statistics of the data are given in Table 3. As the correlation between the variable of study and the auxiliary attribute is positive for the population, we only use the ratio estimators. We use the Eq (4.1) to obtain the percent relative efficiency (PRE) of different estimators,

$$PRE = \frac{Var(\bar{y}_{st})}{MSE(y)} \times 100, \tag{4.1}$$

where $MSE(y)$ is the MSE values of the proposed estimators in Section 2 ($\bar{y}_{stZKi}, i = 1, 2, \ldots, 9$) and other estimators mentioned in Section 1 ($\bar{y}_{Rd}, \bar{y}_{Exd}$) for both cases.

Table 3. Data statistics.

Total		Stratum →	1	2	3	4	5	6
$N = 854$	$\lambda_1 = -0.138$	N_h	106	106	94	171	204	173
$n' = 370$	$\lambda_2 = 0.121$	n'_h	35	50	75	145	45	20
$n = 200$	$\lambda_3 = 0.332$	n_h	13	24	55	95	10	3
$\bar{Y} = 2930.127$	$\lambda_4 = -0.687$	$\bar{Y}_h$	1536.78	2212.59	9384.30	5588.01	966.96	404.40
$S_y = 17105.3$	$\lambda_5 = -0.184$	S_{yh}	6425.09	11551.53	29907.48	28643.42	2389.77	945.75
$C_y = 5.839$	$\lambda_6 = 0.356$	C_{yh}	4.181	5.221	3.187	5.126	2.471	2.339
$C_p = 1.239$	$\lambda_7 = 0.029$	C_{ph}	1.401	1.372	0.943	0.929	1.128	2.105
$\beta_2(\varphi) = -1.82$	$\lambda_8 = 0.691$	C_{yph}	1.725	1.720	0.863	0.826	0.966	3.024
$V_{20} = 0.144$	$\lambda_9 = -0.022$	θ'_h	0.019	0.011	0.003	0.001	0.017	0.044
$V_{02} = 0.069$		θ_h	0.067	0.032	0.008	0.005	0.095	0.328
$V'_{02} = 0.010$		ω_h^2	0.015	0.015	0.012	0.040	0.057	0.041
$V_{11} = 0.049$								
$V'_{11} = 0.007$								

When we examine the efficiency conditions, determined in Section 3, for the data set, we obtain that they are satisfied for the suggested estimators, say $\bar{y}_{stZK8}$, as follows:

For $\lambda_8 = 0.69062 > 0$, $\lambda_8 = 0.69062 < \frac{2(V_{11}-V'_{11})}{(V_{02}-V'_{02})} = 1.39543 \Longrightarrow$ Condition (3.1) is satisfied.

For $\lambda_8 = 0.69062 < 1$, $\lambda_8 = 0.69062 > \frac{2(V_{11}-V'_{11})-(V_{02}-V'_{02})}{(V_{02}-V'_{02})} = 0.39543 \Longrightarrow$ Condition (3.3) is satisfied.

For $\lambda_8 = 0.69062 > 0.5$, $\lambda_8 = 0.69062 < \frac{4(V_{11}-V'_{11})-(V_{02}-V'_{02})}{2(V_{02}-V'_{02})} = 0.89543 \Longrightarrow$ Condition (3.4) is satisfied.

For $\lambda_8 = 0.69062 > 0$, $\lambda_8 = 0.69062 < \frac{2V_{11}}{(V_{02}+V'_{02})} = 1.22005 \Longrightarrow$ Condition (3.6) is satisfied.

For $\lambda_8 = 0.69062 < 1$, $\lambda_8 = 0.69062 > \frac{2V_{11}-(V_{02}+V'_{02})}{(V_{02}+V'_{02})} = 0.22005 \Longrightarrow$ Condition (3.9) is satisfied.

For $\lambda_8 = 0.69062 > 0.5$ $\lambda_8 = 0.69062 < \frac{4V_{11}-(V_{02}+V'_{02})}{2(V_{02}+V'_{02})} = 0.72005 \Longrightarrow$ Condition (3.10) is satisfied.

The comparisons of the proposed exponential ratio estimators have been made with respect to the stratified two-phase ratio estimators in literature. The MSE values of the estimators in Table 4 are obtained using (1.2), (1.4), (1.5), (1.10), (1.11), (2.7) and (2.9), respectively, and the relative efficiency values of the estimators for both of the cases are given in Table 5. From Table 5, it is shown that the proposed exponential ratio estimators $\bar{y}_{stZK2}, \bar{y}_{stZK3}, \bar{y}_{stZK6}, \bar{y}_{stZK7}$ and $\bar{y}_{stZK8}$ perform better than all other mentioned estimators for both cases. Note that the most efficient estimator is the proposed estimator, $\bar{y}_{stZK8}$,for the dataset.

Table 4. MSE values of the classical and suggested exponential estimators.

Estimators	Case I	Case II
$\bar{y}_{st}$	1238296.35	1238296.35
$\bar{y}_{Rd}$	1036469.19	1087246.96
$\bar{y}_{Exd}$	1009782.9	991166.596
$\bar{y}_{stZK1}$	1346860.5	1367609.66
$\bar{y}_{stZK2}$	1159733.69	1147176.64
$\bar{y}_{stZK3}$	1057738.07	1035551.26
$\bar{y}_{stZK4}$	1968437.26	2137542.13
$\bar{y}_{stZK5}$	1386671.02	1415684.56
$\bar{y}_{stZK6}$	1049457.46	1027179.13
$\bar{y}_{stZK7}$	1217627.94	1214069.3
$\bar{y}_{stZK8}$	989856.486	987315.085

Table 5. Percent relative efficiencies (PREs) of the estimators with respect to $\bar{y}_{st}$.

Estimators	Case I	Case II
$\bar{y}_{st}$	100	100
$\bar{y}_{Rd}$	119.47257	113.8928
$\bar{y}_{Exd}$	122.62996	124.9332
$\bar{y}_{stZK1}$	*	*
$\bar{y}_{stZK2}$	106.7741976	107.9429536
$\bar{y}_{stZK3}$	117.0702254	119.578469
$\bar{y}_{stZK4}$	*	*
$\bar{y}_{stZK5}$	*	87.46979249
$\bar{y}_{stZK6}$	117.9939529	120.5531064
$\bar{y}_{stZK7}$	101.6974321	101.9955244
$\bar{y}_{stZK8}$	125.0985737	125.420584
$\bar{y}_{stZK9}$	*	*

* Estimator is not applicable.

5. Conclusions

Tables 4 and 5 clearly show that the stratified two-phase sampling of Eq (2.1) for estimating the population mean using an auxiliary attribute are more efficient. The estimator in Eq (2.1) lower mean square error than the sample mean estimator $\bar{y}_{st}$ of Eq (1.1), the estimator $\bar{y}_{Rd}$ of Eq (1.3), the estimator $\bar{y}_{Exd}$ of Eq (1.9) in stratified two-phase sampling. This means that the proposed exponential estimators for both cases is more efficient than the ratio and the exponential estimators. Finally, it is concluded that the efficiencies of the suggested estimators are better in theory and for the data set used in the article. We hope that in the future we will expand the estimators presented here to ranked set sampling as in Mahdizadeh and Zamanzade [27].

Conflict of interest

The authors have no conflict of interest.

References

1. P. Sharma, R. Singh, Improved estimators in simple random sampling when study variable is an attribute, *J. Stat. Appl. Pro. Lett.*, **2** (2015), 51–58.
2. J. Sahoo, L. N. Sahoo, S. Mohanty, A regression approach to estimation in two-phase sampling using two auxiliary variables, *Curr. Sci.*, **65** (1993), 73–75.
3. M. Samiuddin, M. Hanif, Estimation of population mean in single and two phase sampling with or without additional information, *Pak. J. Stat.*, **23** (2007), 99.
4. H. P. Singh, G. K. Vishwakarma, Modified exponential ratio and product estimators for finite population mean in double sampling, *Austrian J. Stat.*, **36** (2007), 217–225.
5. B. K. Singh, S. Choudhury, Exponential chain ratio and product type estimators for finite population mean under double sampling scheme, *Glob. J. Sci. Front. Res. Math. Decis. Sci.*, **12** (2012), 13–24.
6. S. Choudhury, B. K. Singh, A class of chain ratio-product type estimators with two auxiliary variables under double sampling scheme, *J. Korean Stat. Soc.*, **41** (2012), 247–256.
7. N. Hamad, M. Hanif, N. Haider, A regression type estimator with two auxiliary variables for two-phase sampling, *Open J. Stat.*, **3** (2013), 74–78.
8. A. Sanaullah, H. A. Ali, M. N. ul Amin, M. Hanif, Generalized exponential chain ratio estimators under stratified two-phase random sampling, *Appl. Math. Comput.*, **226** (2014), 541–547.
9. S. Malik, R. Singh, Estimation of population mean using information on auxiliary attribute in two-phase sampling, *Appl. Math. Comput.*, **261** (2015), 114–118.
10. J. Shabbir, S. Gupta, On generalized exponential chain ratio estimators under stratified two-phase random sampling, *Commun. Stat.–Theor. M.*, **46** (2017), 2910–2920.
11. G. K. Vishwakarma, M. Kumar, Efficient classes of ratio-cum-product estimators of population mean in stratified random sampling, *Probab. Math. Stat.*, **36** (2016), 267–278.
12. S. K. Pal, H. P. Singh, A class of ratio-cum-ratio-type exponential estimators for population mean with sub sampling the non-respondents, *JJMS*, **10** (2017), 73–94.
13. G. N. Singh, S. Suman, C. Kadilar, On the use of imputation methods for missing data in estimation of population mean under two-phase sampling design, *Hacet. J. Math. Stat.*, **47** (2018), 1715–1729.
14. I. Saleem, A. Sanaullah, M. Hanif, Double-sampling regression-cum-exponential estimator of the mean of a sensitive variable, *Math. Popul. Stud.*, **26** (2019), 163–182.
15. C. Ünal, C. Kadilar, Improved family of estimators using exponential function for the population mean in the presence of non-response, *Commun. Stat.–Theor. M.*, **50** (2021), 237–248.
16. M. K. Chaudhary, A. Kumar, G. K. Vishwakarma, C. Kadilar, Family of combined-type estimators for population mean using stratified two-phase sampling scheme under non-response, *J. Stat. Manage. Syst.*, **23** (2020), 915–928.

17. M. Irfan, M. Javed, S. H. Bhatti, Difference-type-exponential estimators based on dual auxiliary information under simple random sampling, *Sci. Iran.*, 2020, DOI: 10.24200/sci.2020.53592.3318.
18. C. Ünal, C. Kadilar, Exponential type estimator for the population mean in the presence of non-response, *J. Stat. Manage. Syst.*, **23** (2020), 603–615.
19. S. Muneer, A. Khalil, J. Shabbir, A parent-generalized family of chain ratio exponential estimators in stratified random sampling using supplementary variables, *Commun. Stat.–Simul. Comput.*, 2020, DOI: 10.1080/03610918.2020.1748887.
20. A. Sanaullah, I. Saleem, S. Gupta, M. Hanif, Mean estimation with generalized scrambling using two-phase sampling, *Commun. Stat.–Simul. Comput.*, 2020, DOI: 10.1080/03610918.2020.1778032.
21. R. R. Sinha, S. Gangwar, S. Sharma, Regression-cum-exponential estimators for product of two population means under double sampling the non-respondents, *Proc. Natl. Acad. Sci. India Sect. A Phys. Sci.*, 2020, DOI: 10.1007/s40010-020-00693-x.
22. T. Zaman, An efficient exponential estimator of the mean under stratified random sampling, *Math. Popul. Stud.*, 2020, DOI: 10.1080/08898480.2020.1767420.
23. Y. Hassan, M. Ismail, W. Murray, M. Q. Shahbaz, Efficient estimation combining exponential and ln functions under two phase sampling, *AIMS Mathematics*, **5** (2020), 7605–7623.
24. C. Bose, Note on the sampling error in the method of double sampling, *Sankhya*, **6** (1943), 329–330.
25. T. Zaman, C. Kadilar, New class of exponential estimators for finite population mean in two-phase sampling, *Commun. Stat.–Theor. Methods*, **50** (2021), 874–889.
26. C. Kadilar, H. Cingi, Ratio estimators in stratified random sampling, *Biometrical J.*, **45** (2003), 218–225.
27. M. Mahdizadeh, E. Zamanzade, Stratified pair ranked set sampling, *Commun. Stat.–Theor. Methods*, **47** (2018), 5904–5915.