Gazi University Journal of Science
GU J Sci
26(2):181-193 (2013)

Families of Estimators for Population Mean Using Information on Auxiliary Attribute in Stratified Random Sampling

Nursel KOYUNCU[1,♠]

[1]*Hacettepe University. Department of Statistics, Ankara, TURKEY*

Received: 13.02.2012 Revised: 19.03.2012 Accepted:10.08.2012

ABSTRACT

Grover and Kaur [4], Haq et al. [5], Abd-Elfattah et al. [1], Jhajj et al. [6], Shabbir and Gupta [8] and Koyuncu [7] have suggested some families of estimators by using the known population proportion of elements possessing attribute in the simple random sampling and in the two phase sampling. In this paper, after adapting these families of estimators to the stratified random sampling, we have proposed a family of exponential ratio type estimators which uses the information regarding the population proportion possessing certain attribute. For the proposed family of estimators, the expressions of bias and mean square error (MSE) up to the first order approximations are derived and the optimum case of the proposed family is discussed in theory. Also an empirical study is carried out to show its properties.

Key words: Attribute, auxiliary information, efficiency, exponential ratio estimator, stratified random sampling.

1. INTRODUCTION

Grover and Kaur [4], Haq et al. [5], Abd-Elfattah et al. [1], Jhajj et al. [6], Shabbir and Gupta [8] and Koyuncu [7] suggested some estimators of population mean taking the advantage of point bi-serial correlation between the study variable and the auxiliary attribute. In simple random sampling, Bahl and Tuteja [2] introduced exponential estimators for finite population mean. Following Bahl and Tuteja [2], Singh and Vishwakarma [10] suggested exponential ratio and product estimators of population mean in double sampling. Shabbir and Gupta [9] proposed a new ratio-type exponential estimator for population variance in simple random sampling. Singh et al. [11] proposed some exponential ratio-type estimators for estimating the population mean using known values of certain population parameter(s) under simple random sampling scheme. In this paper, we proposed a family of exponential ratio type estimators for estimating the population mean using known values of two auxiliary attributes.

Assume that the population of size, *N*, is stratified into *L* strata with the *h*-th stratum containing N_h units,

♠Corresponding author, e-mail: nkoyuncu@hacettepe.edu.tr

where $h=1,2,...,L$ such that $\sum_{h=1}^{L} N_h = N$. A simple random sample of size n_h is drawn without replacement from the *h*-th stratum such that $\sum_{h=1}^{L} n_h = n$. Let $(y_{hi}, \psi_{1hi}, \psi_{2hi})$ denote observed values of *Y* and auxiliary attributes ψ_1, ψ_2 on the *i*-th unit of the *h*-th stratum, respectively, where $i=1,2,...,N_h$ and $h=1,2,...,L$. Moreover,

$$\psi_{jhi} = \begin{cases} 1 & , h\text{-th stratum } i\text{-th unit possesses the attribute } \Psi_j \\ 0 & , \text{otherwise} \end{cases}.$$

Supposing that $A_{jh} = \sum_{i=1}^{N_h} \psi_{jhi}$ and $a_{jh} = \sum_{i=1}^{n_h} \psi_{jhi}$ denote the *h*-th total number of units,

$$P_{jh} = \frac{\sum_{i=1}^{N_h} \psi_{jhi}}{N_h} = \frac{A_{jh}}{N_h}, \quad p_{jh} = \frac{\sum_{i=1}^{n_h} \psi_{jhi}}{n_h} = \frac{a_{jh}}{n_h}$$

$(j=1,2)$ are the *h*-th proportion of units in population and sample, respectively. Let

$$s_{yh}^2 = \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2}{n_h - 1},$$

$$s_{\psi_j h}^2 = \frac{\sum_{i=1}^{n_h} (\psi_{jhi} - p_{jh})^2}{n_h - 1}$$ be the *h*-th sample variances and let $S_{yh}^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1}$,

$$S_{\psi_j h}^2 = \frac{\sum_{i=1}^{N_h} (\psi_{jhi} - P_{jh})^2}{N_h - 1}$$ be the population variances of the study variable and auxiliary attributes, respectively. Also

$$s_{y\psi_j h} = \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)(\psi_{jhi} - p_{jh})}{n_h - 1}$$ and

$\hat{\rho}_{y\psi_j h} = \frac{s_{y\psi_j h}}{s_{yh} s_{\psi_j h}}$ $(j=1,2)$ are, respectively, the *h*-th sample covariance and point bi-serial correlation, assume that $\bar{y}_h = \sum_{i=1}^{n_h} \frac{y_{hi}}{n_h}$, $\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h$, and $\bar{Y}_h = \sum_{i=1}^{N_h} \frac{Y_{hi}}{N_h}$, $\bar{Y} = \sum_{h=1}^{L} W_h \bar{Y}_h$ are the sample and population means of *Y*, respectively, where $W_h = \frac{N_h}{N}$ is the stratum weight. Further let

$$S_{y\psi_j h} = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)(\psi_{jhi} - P_{jh})}{N_h - 1}$$ and

$\hat{\rho}_{y\psi_j h} = \frac{S_{y\psi_j h}}{S_{yh} S_{\psi_j h}}$ are, respectively, population covariance and point bi-serial correlation between the study variable and the auxiliary attribute. Let

$$s_{\psi_1\psi_2 h} = \frac{\sum_{i=1}^{n_h} (\psi_{1hi} - p_{1h})(\psi_{2hi} - p_{2h})}{n_h - 1}$$ and

$\hat{\rho}_{\phi h} = \frac{s_{\psi_1\psi_2 h}}{s_{\psi_1 h} s_{\psi_2 h}}$ be the *h*-th sample phi-covariance and phi correlation, respectively, and let

$$S_{\psi_1\psi_2 h} = \frac{\sum_{i=1}^{N_h} (\psi_{1hi} - P_{1h})(\psi_{2hi} - P_{2h})}{N_h - 1}$$ and

$\rho_{\phi h} = \frac{S_{\psi_1\psi_2 h}}{S_{\psi_1 h} S_{\psi_2 h}}$ be the *h*-th population phi covariance and phi correlation, respectively, between the first and the second auxiliary attributes.

To obtain the bias and *MSE*, let us define

$\xi_0 = (\bar{y}_{st} - \bar{Y})/\bar{Y}$, $\xi_{\psi_1} = (p_{1st} - P_1)/P_1$, $\xi_{\psi_2} = (p_{2st} - P_2)/P_2$ where $\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h$, $p_{1st} = \sum_{h=1}^{L} W_h p_{1h}$ and $p_{2st} = \sum_{h=1}^{L} W_h p_{2h}$.

Using these notations,

$$E(\xi_0) = E(\xi_{\psi_1}) = E(\xi_{\psi_2}) = 0,$$

$$N_{rst}=\sum_{h=1}^{L}W_h^{r+s+t}\frac{E\left[(\bar{y}_h-\bar{Y}_h)^r(p_{1h}-P_{1h})^s(p_{2h}-P_{2h})^t\right]}{\bar{Y}^rP_1^sP_2^t} \quad (1.1)$$

Using (1.1), we can write

$$E(\xi_0^2)=\frac{\sum_{h=1}^{L}W_h^2\gamma_hS_{yh}^2}{\bar{Y}^2}=N_{200},$$

$$E(\xi_{\psi_1}^2)=\frac{\sum_{h=1}^{L}W_h^2\gamma_hs_{\psi_1h}^2}{P_1^2}=N_{020},$$

$$E(\xi_{\psi_2}^2)=\frac{\sum_{h=1}^{L}W_h^2\gamma_hs_{\psi_2h}^2}{P_2^2}=N_{002},$$

$$E(\xi_0\xi_{\psi_1})=\frac{\sum_{h=1}^{L}W_h^2\gamma_h\rho_{y\psi_1h}S_{yh}S_{\psi_1}}{\bar{Y}P_1}=N_{110},$$

$$E(\xi_0\xi_{\psi_2})=\frac{\sum_{h=1}^{L}W_h^2\gamma_h\rho_{y\psi_2h}S_{yh}S_{\psi_2h}}{\bar{Y}P_2}=N_{101},$$

$$E(\xi_{\psi_1}\xi_{\psi_2})=\frac{\sum_{h=1}^{L}W_h^2\gamma_h\rho_{\psi_1\psi_2h}S_{\psi_1h}S_{\psi_2h}}{P_1P_2}=N_{011}$$

where $\gamma_h=\frac{N_h-n_h}{N_hn_h}$.

For the one auxiliary attribute case, we can define similar expressions as

$$N_{rs}=\sum_{h=1}^{L}W_h^{r+s}\frac{E\left[(\bar{y}_h-\bar{Y}_h)^r(p_h-P_h)^s\right]}{\bar{Y}^rP^s} \quad (1.2)$$

Using (1.2), we can write

$$E(\xi_0^2)=\frac{\sum_{h=1}^{L}W_h^2\gamma_hS_{yh}^2}{\bar{Y}^2}=N_{20},$$

$$E(\xi_{\psi_1}^2)=\frac{\sum_{h=1}^{L}W_h^2\gamma_hs_{\psi_1h}^2}{P^2}=N_{02},$$

$$E(\xi_0\xi_{\psi_1})=\frac{\sum_{h=1}^{L}W_h^2\gamma_h\rho_{y\psi_1h}S_{yh}S_{\psi_1}}{\bar{Y}P}=N_{11}.$$

2. ADAPTIVE ESTIMATORS

The classical ratio estimator using a single auxiliary attribute in the stratified random sampling is given by

$$t_1=\bar{y}_{st}\frac{P}{p_{st}}. \quad (2.1)$$

The bias and *MSE* of t_1, to the first order of approximation, are given by

$$B(t_1)=\bar{Y}(N_{02}-N_{11}), \quad (2.2)$$

$$MSE(t_1)\cong\bar{Y}^2(N_{02}+N_{20}-2N_{11}). \quad (2.3)$$

(Çıngı and Kadılar [3]). Following Jhajj et al.[6], a general combined class in the stratified random sampling is defined by

$$\tilde{t}_{g(st)}=g(\bar{y}_{st},\upsilon_{st}), \quad (2.4)$$

where $\upsilon_{st}=\frac{p_{st}}{P}$ and $g(\bar{y}_{st},\upsilon_{st})$ is a function of $\bar{y}_{st}$ and υ_{st}. To study the properties of $\tilde{t}_{g(st)}$ we assume following regularity conditions:

1. The point $(\bar{y}_{st},\upsilon_{st})$ assumes the value in a closed convex subset R_2 of two dimensional real space containing the point $(\bar{Y},1)$,
2. The function $g(\bar{y}_{st},\upsilon_{st})$ is continuous and bounded in R_2,
3. $g(\bar{Y},1)=\bar{Y}$ and $g_0(\bar{Y},1)=1$, where $g_0(\bar{Y},1)$ denotes the first order partial derivative of g with respect to $\bar{y}_{st}$,

4. The first and the second order partial derivatives of $g(\bar{y}_{st},\upsilon_{st})$ exist and are continuous and bounded in R_2.

Expanding $g(\bar{y}_{st},\upsilon_{st})$ about the point $(\bar{Y},1)$ in a second order Taylor series and using the above regularity conditions, we have

$$\tilde{t}_{g(st)} = \bar{y}_{st} + (\upsilon_{st}-1)g_1 + (\upsilon_{st}-1)^2 g_2 + (\bar{y}_{st}-\bar{Y})(\upsilon_{st}-1)g_3 + (\bar{y}_{st}-\bar{Y})^2 g_4 \tag{2.5}$$

where

$$g_1 = \left.\frac{\partial g}{\partial \upsilon_{st}}\right|_{\bar{y}_{st}=\bar{Y},\upsilon_{st}=1}, g_2 = \left.\frac{1}{2}\frac{\partial^2 g}{\partial \upsilon_{st}^2}\right|_{\bar{y}_{st}=\bar{Y},\upsilon_{st}=1},$$

$$g_3 = \left.\frac{1}{2}\frac{\partial^2 g}{\partial \bar{y}_{st}\partial \upsilon_{st}}\right|_{\bar{y}_{st}=\bar{Y},\upsilon_{st}=1}, g_4 = \left.\frac{1}{2}\frac{\partial^2 g}{\partial \bar{y}_{st}^2}\right|_{\bar{y}_{st}=\bar{Y},\upsilon_{st}=1}.$$

The class of estimators $\tilde{t}_{g(st)}$, in terms of ξ_0 and ξ_{ψ_1}, can be written as,

$$\tilde{t}_{g(st)} - \bar{Y} = \bar{Y}\xi_0 + \xi_{\psi_1}g_1 + \xi_{\psi_1}^2 g_2 + \bar{Y}\xi_0\xi_{\psi_1}g_3 + \bar{Y}^2\xi_0^2 g_4 \tag{2.6}$$

and the bias and the *MSE* of $\tilde{t}_{g(st)}$ are, respectively, given by

$$B(\tilde{t}_{g(st)}) \cong g_2 N_{02} + g_3\bar{Y}N_{11} + g_4\bar{Y}^2 N_{20}, \tag{2.7}$$

$$MSE(\tilde{t}_{g(st)}) \cong \bar{Y}^2 N_{20} + N_{02}g_1^2 + 2\bar{Y}N_{11}g_1. \tag{2.8}$$

By using the optimal value of $g_1^* = -\frac{\bar{Y}N_{11}}{N_{02}}$, the minimum *MSE* of the estimators in the class $\tilde{t}_{g(st)}$ is found as

$$MSE_{min}(\tilde{t}_{g(st)}) \cong \bar{Y}^2 N_{20}\left[1 - \frac{N_{11}^2}{N_{02}N_{20}}\right] \tag{2.9}$$

Motivated by Shabbir and Gupta [8], we present a combined estimator using one auxiliary attribute as

$$\bar{y}_{M(st)} = \bar{y}_{st}\left[d_{1st} + d_{2st}(P - p_{st})\right]\frac{P}{p_{st}} \tag{2.10}$$

where d_1 and d_2 are constants to be determined later whose sum is not necessarily equal to one. Expressing the estimator, $\bar{y}_{M(st)}$, in terms of ξ_i (i = 0, ψ_1), we can write (2.10) as

$$\bar{y}_{M(st)} \cong \bar{Y}(1+\xi_0)(d_{1st} - Pd_{2st}\xi_{\psi_1})(1 - \xi_{\psi_1} + \xi_{\psi_1}^2) \tag{2.11}$$

The bias and *MSE* of the estimator to the first order of approximations are, respectively, given by

$$B(\bar{y}_{M(st)}) \cong \bar{Y}\left[(d_{1st}-1) - (Pd_{2st}+d_{1st})N_{11} + (Pd_{2st}+d_{1st})N_{02}\right] \tag{2.12}$$

$$\begin{aligned} MSE(\bar{y}_{M(st)}) \cong \bar{Y}^2 \big[&1 + d_{1st}^2(1 + N_{20} + 3N_{02} - 4N_{11}) + 4Pd_{1st}d_{2st}(N_{02} - N_{11}) \\ &- 2d_{1st}(1 + N_{02} - N_{11}) + P^2 d_{2st}^2 N_{02} + 2Pd_{2st}(N_{11} - N_{02})\big] \end{aligned} \tag{2.13}$$

The optimum values of d_{1st}, d_{2st} and minimum *MSE* of $\bar{y}_{M(st)}$ are given by

$$d_{1st}^* = \frac{N_{02} + N_{02}^2 - N_{11}N_{02} - 2N_{02}^2 - 2N_{11}^2 + 4N_{02}N_{11}}{N_{02} + N_{20}N_{02} - N_{02}^2 - 4N_{11}^2 + 4N_{02}N_{11}}, \tag{2.14}$$

$$d^*_{2st} = \frac{N_{20}N_{02} - N_{11}N_{20} - 3N_{11}N_{02} - N_{02} + N_{02}^2 + N_{11} + 2N_{11}^2}{P\left(N_{02} + N_{02}N_{20} - N_{02}^2 - 4N_{11}^2 + 4N_{02}N_{11}\right)}, \quad (2.15)$$

$$MSE_{\min}\left(\bar{y}_{M(st)}\right) \cong \bar{Y}^2\left[1 - \frac{N_{02} - N_{02}^2 + 4N_{11}N_{02} - 3N_{11}^2 + N_{20}N_{02}^2 + N_{20}N_{11}^2 - 2N_{02}N_{11}N_{20}}{N_{02} + N_{20}N_{02} - N_{02}^2 - 4N_{11}^2 + 4N_{02}N_{11}}\right] \quad (2.16)$$

3. SUGGESTED ESTIMATORS

Following Bahl and Tuteja [2] and Singh et al. [11], we propose the following estimator

$$\bar{y}_N = \bar{y}_{st}\left\{K_{1st}\left(\frac{P_1}{p_{1st}}\right)^{\gamma_{1st}} \exp\left[\frac{a_{st}(p_{1st} - P_1)}{P_1 + b_{st}(p_{1st} - P_1)}\right] + K_{2st}\left(\frac{P_2}{p_{2st}}\right)^{\gamma_{2st}} \exp\left[\frac{c_{st}(p_{2st} - P_2)}{P_2 + d_{st}(p_{2st} - P_2)}\right]\right\} \quad (3.1)$$

where K_{1st}, K_{2st}, γ_{1st}, γ_{2st} are suitable constants, a_{st} and b_{st} are either real numbers or the functions of the known parameters for the *h*-th stratum of the auxiliary attribute, ψ_1, such as $C_{p_1(st)} = \sum_{h=1}^{L} W_h C_{p_1 h}$, $\beta_{1(p_1)st} = \sum_{h=1}^{L} W_h \beta_{1h}(\phi_1)$, $\beta_{2(p_1)st} = \sum_{h=1}^{L} W_h \beta_{2h}(\phi_1)$, $\rho_{(yp_1)st} = \sum_{h=1}^{L} W_h \rho_{(yp_1)h}$. For c_{st} and d_{st}, we can define similar expressions for the auxiliary attribute ψ_2, as a_{st} and b_{st}. Some new estimators can be generated from (3.1) for different combinations of K_{1st}, K_{2st}, γ_{1st}, γ_{2st}, a_{st}, b_{st}, c_{st}, and d_{st}.

Case 1: For this family of estimators firstly we consider the case of $K_{1st} + K_{2st} = 1$ then we can rewrite (3.1) in terms of ξ_i s $i = 0, \psi_1, \psi_2$ as

$$\begin{aligned}\bar{y}_N - \bar{Y} \cong \bar{Y}\Bigg\{&K_{1st}\Bigg(a_{st}\xi_{\psi_1} + \frac{a_{st}(a_{st} - 2b_{st})}{2!}\xi_{\psi_1}^2 - \gamma_{1st}\xi_{\psi_1} - \gamma_{1st}a_{st}\xi_{\psi_1}^2 \\ &+ \frac{\gamma_{1st}(\gamma_{1st}+1)}{2}\xi_{\psi_1}^2 + \xi_0 + a_{st}\xi_0\xi_{\psi_1} - \gamma_{1st}\xi_0\xi_{\psi_1}\Bigg) \\ &+ K_{2st}\Bigg(c_{st}\xi_{\psi_2} + \frac{c_{st}(c_{st} - 2d_{st})}{2!}\xi_{\psi_2}^2 - \gamma_{2st}\xi_{\psi_2} - \gamma_{2st}c_{st}\xi_{\psi_2}^2 \\ &+ \frac{\gamma_{2st}(\gamma_{2st}+1)}{2}\xi_{\psi_2}^2 + \xi_0 + c_{st}\xi_0\xi_{\psi_2} - \gamma_{2st}\xi_0\xi_{\psi_2}\Bigg)\Bigg\}\end{aligned} \quad (3.2)$$

The bias and *MSE* of the estimator in (3.1), to the first-order of approximation, are, respectively, given by

$$\begin{aligned}B(\bar{y}_N) \cong \bar{Y}\Bigg\{&K_{1st}\left(\left(\frac{a_{st}(a_{st} - 2b_{st})}{2!} - \gamma_{1st}a_{st} + \frac{\gamma_{1st}(\gamma_{1st}+1)}{2}\right)N_{020} + (a_{st} - \gamma_{1st})N_{110}\right) \\ &+ K_{2st}\left(\left(\frac{c_{st}(c_{st} - 2d_{st})}{2!} - \gamma_{2st}c_{st} + \frac{\gamma_{2st}(\gamma_{2st}+1)}{2}\right)N_{002} + (c_{st} - \gamma_{2st})N_{101}\right)\Bigg\}\end{aligned} \quad (3.3)$$

$$MSE\left(\bar{y}_N\right) \cong \bar{Y}^2\left\{N_{200} + t_{1st}^2 N_{020} + t_{2st}^2 N_{002} + 2t_{1st}N_{110} + 2t_{2st}N_{101} + 2t_{1st}t_{2st}N_{011}\right\} \quad (3.4)$$

where $t_{1st} = K_{1st}\left(a_{st} - \gamma_{1st}\right)$ and $t_{2st} = K_{2st}\left(c_{st} - \gamma_{2st}\right)$.

From (3.4), we get the optimum values of t_{1st} and t_{2st}, given by

$$t_{1st} = \frac{N_{101}N_{011} - N_{110}N_{002}}{N_{020}N_{002} - N_{011}^2},$$

$$t_{2st} = \frac{N_{110}N_{011} - N_{101}N_{020}}{N_{002}N_{020} - N_{011}^2}. \quad (3.5)$$

Substituting the optimum values of t_{1st} and t_{2st}, we get minimum *MSE* of $\bar{y}_N$, as given by

$$MSE_{\min}\left(\bar{y}_N\right) \cong \bar{Y}^2\left\{N_{200} - \frac{\left(N_{020}N_{101}^2 + N_{110}^2 N_{002} - 2N_{110}N_{011}N_{101}\right)}{\left(N_{020}N_{002} - N_{011}^2\right)}\right\} \quad (3.6)$$

Case 2: Now we consider the case of $K_{1st} + K_{2st} \neq 1$. Some new estimators which are generated from (3.1) for different combinations of K_{1st}, K_{2st}, γ_{1st}, γ_{2st}, a_{st}, b_{st}, c_{st} and d_{st} are given in Table1. Then, we can rewrite (3.1) in terms of ξ_i, $i = 0, \psi_1, \psi_2$ as

$$\begin{aligned}\bar{y}_N^* - \bar{Y} \cong \bar{Y}\Bigg\{ & K_{1st}\Bigg(1 + a_{st}\xi_{\psi_1} + \frac{a_{st}\left(a_{st} - 2b_{st}\right)}{2!}\xi_{\psi_1}^2 - \gamma_{1st}\xi_{\psi_1} - \gamma_{1st}a_{st}\xi_{\psi_1}^2 \\ & + \frac{\gamma_{1st}\left(\gamma_{1st}+1\right)}{2}\xi_{\psi_1}^2 + \xi_0 + a_{st}\xi_0\xi_{\psi_1} - \gamma_{1st}\xi_0\xi_{\psi_1}\Bigg) \\ & + K_{2st}\Bigg(1 + c_{st}\xi_{\psi_2} + \frac{c_{st}\left(c_{st} - 2d_{st}\right)}{2!}\xi_{\psi_2}^2 - \gamma_{2st}\xi_{\psi_2} - \gamma_{2st}c_{st}\xi_{\psi_2}^2 \\ & + \frac{\gamma_{2st}\left(\gamma_{2st}+1\right)}{2}\xi_{\psi_2}^2 + \xi_0 + c_{st}\xi_0\xi_{\psi_2} - \gamma_{2st}\xi_0\xi_{\psi_2}\Bigg) - 1\Bigg\}\end{aligned} \quad (3.7)$$

The bias and *MSE* of the estimator in (3.1), to the first-order of approximation, are, respectively, given by

$$\begin{aligned}B\left(\bar{y}_N^*\right) \cong \bar{Y}\Bigg\{ & K_{1st}\Bigg(1 + \left\{\frac{a_{st}\left(a_{st} - 2b_{st}\right)}{2!} - \gamma_{1st}a_{st} + \frac{\gamma_{1st}\left(\gamma_{1st}+1\right)}{2}\right\}N_{020} + \left(a_{st} - \gamma_{1st}\right)N_{110}\Bigg) \\ & + K_{2st}\Bigg(1 + \left\{\frac{c_{st}\left(c_{st} - 2d_{st}\right)}{2!} - \gamma_{2st}c_{st} + \frac{\gamma_{2st}\left(\gamma_{2st}+1\right)}{2}\right\}N_{002} + \left(c_{st} - \gamma_{2st}\right)N_{101}\Bigg) - 1\Bigg\}\end{aligned} \quad (3.8)$$

$$MSE\left(\bar{y}_N^*\right) \cong \bar{Y}^2\left\{1 + K_{1st}^2 A + K_{2st}^2 B + 2K_{1st}K_{2st}C - 2K_{1st}D - 2K_{2st}E\right\} \quad (3.9)$$

where

$$A = 1 + \left\{2a_{st}^2 + 2\gamma_{1st}^2 - 2a_{st}b_{st} - 4\gamma_{1st}a_{st} + \gamma_{1st}\right\}N_{020} + N_{200} + 4\left\{a_{st} - \gamma_{1st}\right\}N_{110}$$

$$B = 1 + \left\{2c_{st}^2 + 2\gamma_{2st}^2 - 2c_{st}d_{st} - 4\gamma_{2st}c_{st} + \gamma_{2st}\right\}N_{002} + N_{200} + 4\left\{c_{st} - \gamma_{2st}\right\}N_{101}$$

$$C = 1 + \left\{ \frac{a_{st}(a_{st} - 2b_{st})}{2} - \gamma_{1st}a_{st} + \frac{\gamma_{1st}(\gamma_{1st}+1)}{2} \right\} N_{020} + (a_{st} - \gamma_{1st}) N_{110}$$
$$+ (a_{st}c_{st} - \gamma_{1st}c_{st} - \gamma_{2st}a_{st} + \gamma_{2st}\gamma_{1st}) N_{011} + 2(c_{st} - \gamma_{2st}) N_{101}$$
$$+ \left\{ \frac{c_{st}(c_{st} - 2d_{st})}{2} - \gamma_{2st}c_{st} + \frac{\gamma_{2st}(\gamma_{2st}+1)}{2} \right\} N_{002} + N_{200}$$

$$D = 1 + \left\{ \frac{a_{st}(a_{st} - 2b_{st})}{2} - \gamma_{1st}a_{st} + \frac{\gamma_{1st}(\gamma_{1st}+1)}{2} \right\} N_{020} + (a_{st} - \gamma_{1st}) N_{110}$$

$$E = 1 + \left\{ \frac{c_{st}(c_{st} - 2d_{st})}{2} - \gamma_{2st}c_{st} + \frac{\gamma_{2st}(\gamma_{2st}+1)}{2} \right\} N_{002} + (c_{st} - \gamma_{2st}) N_{101}$$

From (3.9), we get the optimum values of K_{1st} and K_{2st}, given by

$$K_{1st}^* = \frac{(DB - EC)}{(AB - C^2)}, K_{2st}^* = \frac{(EA - DC)}{(AB - C^2)}. \quad (3.10)$$

Substituting the optimum values of K_{1st} and K_{2st}, we get minimum *MSE* of $\bar{y}_N^*$, as given by

$$MSE_{min}(\bar{y}_N^*) \cong \bar{Y}^2 \left(1 - \frac{BD^2 + AE^2 - 2CED}{AB - C^2} \right). \quad (3.11)$$

4. COMPARISON OF ESTIMATORS

The $\bar{y}_N$ -family of estimators is more efficient than t_1 if

$$MSE_{min}(\bar{y}_N) < MSE(t_1),$$

$$2N_{11} - N_{02} - \frac{N_{020}N_{101}^2 + N_{110}^2 N_{002} - 2N_{110}N_{011}N_{101}}{N_{020}N_{002} - N_{011}^2} < 0 \quad (4.1)$$

When the condition (4.1) is satisfied, we can infer that the $\bar{y}_N$ family is more efficient than t_1 estimator.

The $\bar{y}_N$ -family of estimators is more efficient than $\tilde{t}_{g(st)}$ if

$$MSE_{min}(\bar{y}_N) < MSE_{min}(\tilde{t}_{g(st)}),$$

$$\frac{N_{11}^2}{N_{02}} - \frac{N_{020}N_{101}^2 + N_{110}^2 N_{002} - 2N_{110}N_{011}N_{101}}{N_{020}N_{002} - N_{011}^2} < 0 . \quad (4.2)$$

When the condition (4.2) is satisfied, we can infer that the $\bar{y}_N$ family is more efficient than $\tilde{t}_{g(st)}$ estimator.

The $\bar{y}_N$ -family of estimators is more efficient than $\bar{y}_{M(st)}$ if

$$MSE_{min}(\bar{y}_N) < MSE_{min}(\bar{y}_{M(st)}),$$

$$\left\{ N_{200} - \frac{N_{020}N_{101}^2 + N_{110}^2 N_{002} - 2N_{110}N_{011}N_{101}}{N_{020}N_{002} - N_{011}^2} + \frac{N_{02} - N_{02}^2 + 4N_{11}N_{02} - 3N_{11}^2 + N_{20}N_{02}^2 + N_{20}N_{11}^2 - 2N_{02}N_{11}N_{20}}{N_{02} + N_{20}N_{02} - N_{02}^2 - 4N_{11}^2 + 4N_{02}N_{11}} \right\} < 1 \quad (4.3)$$

When the condition (4.3) is satisfied, we can infer that the $\bar{y}_N$ family is more efficient than $\bar{y}_{M(st)}$ estimator.

The $\bar{y}_N^*$ -family of estimators is more efficient than t_1 if

$$MSE_{\min}\left(\bar{y}_N^*\right) < MSE\left(t_1\right),$$

$$1 - N_{20} - N_{02} + 2N_{11} - \frac{BD^2 + AE^2 - 2CED}{AB - C^2} < 0 \cdot \quad (4.4)$$

When the condition (4.4) is satisfied, we can infer that the $\bar{y}_N^*$ family is more efficient than t_1 estimator.

The $\bar{y}_N^*$ -family of estimators is more efficient than $\tilde{t}_{g(st)}$ if

$$MSE_{\min}\left(\bar{y}_N^*\right) < MSE_{\min}\left(\tilde{t}_{g(st)}\right),$$

$$1 - \frac{BD^2 + AE^2 - 2CED}{AB - C^2} - N_{20}\left(1 - \frac{N_{11}^2}{N_{02}N_{20}}\right) < 0 \cdot \quad (4.5)$$

When the condition (4.5) is satisfied, we can infer that the $\bar{y}_N^*$ family is more efficient than $\tilde{t}_{g(st)}$ estimator.

The $\bar{y}_N^*$ -family of estimators is more efficient than $\bar{y}_{M(st)}$ if

$$MSE_{\min}\left(\bar{y}_N^*\right) < MSE_{\min}\left(\bar{y}_{M(st)}\right),$$

$$\frac{N_{02} - N_{02}^2 + 4N_{11}N_{02} - 3N_{11}^2 + N_{20}N_{02}^2 + N_{20}N_{11}^2 - 2N_{02}N_{11}N_{20}}{N_{02} + N_{20}N_{02} - N_{02}^2 - 4N_{11}^2 + 4N_{02}N_{11}} - \frac{BD^2 + AE^2 - 2CED}{AB - C^2} < 0 \quad (4.6)$$

When the condition (4.6) is satisfied, we can infer that the $\bar{y}_N^*$ family is more efficient than $\bar{y}_{M(st)}$ estimator.

5. NUMERICAL EXAMPLE

In this study we use the data of apple production amount in 1999 as study variable, number of apple trees less than 20.000 in 1999 and apple production amount less than 1000 tons in 1998 as auxiliary attributes in 854 villages of Turkey (Source: Institute of Statistics, Republic of Turkey).

Firstly, 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 Neyman allocation method. The summary of the data is given in Table 2.

The *MSE* values of the adapted estimators (t_1, $\tilde{t}_{g(st)}$, $\bar{y}_{M(st)}$) and suggested estimators ($\bar{y}_N$, $\bar{y}_N^*$) have been obtained using (2.3), (2.9), (2.16), (3.6), and (3.11), respectively. These values are given in Table 3. When we examine Table 3, we observe that the $\bar{y}_{N1}^*$, $\bar{y}_{N9}^*$, $\bar{y}_{N13}^*$ estimators have the smallest *MSE* values and are more efficient than adapted and other suggested estimators. From this result, we can conclude that using suitable known parameters of auxiliary attributes in estimators give more efficient results.

We also compute the percent relative efficiency (PRE) using the following expression

$$PRE = \frac{MSE\left(t_1\right)}{MSE\left(\alpha\right)} * 100$$

where $\alpha = \bar{y}_{N1}^*, \bar{y}_{N2}^*, \ldots, \bar{y}_{N16}^*, \bar{y}_{M(st)}, \bar{y}_N, \tilde{t}_{g(st)}$

.

Table 3.Values of MSE of Estimators

Estimator	MSE	PRE	Estimator	MSE	PRE
$\bar{y}^*_{N1}$	**248035.4***	**197.5690****	$\bar{y}^*_{N11}$	391892.3	125.0448
$\bar{y}^*_{N2}$	338505.9	144.7659	$\bar{y}^*_{N12}$	330934	148.0782
$\bar{y}^*_{N3}$	391892.3	125.0448	$\bar{y}^*_{N13}$	**248035.4***	**197.5690****
$\bar{y}^*_{N4}$	330934	148.0782	$\bar{y}^*_{N14}$	338505.9	144.7659
$\bar{y}^*_{N5}$	325497.8	150.5512	$\bar{y}^*_{N15}$	391892.3	125.0448
$\bar{y}^*_{N6}$	471536.3	103.9243	$\bar{y}^*_{N16}$	330934	148.0782
$\bar{y}^*_{N7}$	351286.2	139.4991	t1	490041	100
$\bar{y}^*_{N8}$	342801.7	142.9517	$\bar{y}_{M(st)}$	347726.3	140.9272
$\bar{y}^*_{N9}$	**248035.4***	**197.5690****	$\bar{y}_N$	356766.1586	137.3564
$\bar{y}^*_{N10}$	338505.9	144.7659	$\tilde{t}_{g(st)}$	367247.1604	133.4363

When we examine Table 3. all suggested estimators are more efficient than classical ratio estimator. From Table 3 we can see that $\bar{y}_N$ family is more efficient than t_1; $\bar{y}_N$ family is more efficient than $\tilde{t}_{g(st)}$ estimator. We can infer that conditions given in (4.1) and (4.2) are satisfied. But $\bar{y}_N$ family isn't more efficient than $\bar{y}_{M(st)}$ estimator. So we can say that the condition in (4.3) isn't satisfied for this data set.

All member of $\bar{y}^*_N$ family are more efficient than t_1 estimator and the condition in (4.4.) is satisfied. Member of $\bar{y}^*_N$ such as $\bar{y}^*_{N1}$, $\bar{y}^*_{N2}$, $\bar{y}^*_{N4}$, $\bar{y}^*_{N5}$, $\bar{y}^*_{N7}$, $\bar{y}^*_{N8}$, $\bar{y}^*_{N9}$, $\bar{y}^*_{N10}$, $\bar{y}^*_{N12}$, $\bar{y}^*_{N13}$, $\bar{y}^*_{N14}$ and $\bar{y}^*_{N16}$ are more efficient than $\tilde{t}_{g(st)}$ estimator and the condition in (4.5) is satisfied for these estimators. Member of $\bar{y}^*_N$ such as $\bar{y}^*_{N3}$, $\bar{y}^*_{N6}$, $\bar{y}^*_{N11}$ and $\bar{y}^*_{N15}$ aren't more efficient than $\tilde{t}_{g(st)}$ and condition in (4.5) isn't satisfied for these estimators. Member of $\bar{y}^*_N$ such as $\bar{y}^*_{N1}$, $\bar{y}^*_{N2}$, $\bar{y}^*_{N4}$, $\bar{y}^*_{N5}$, $\bar{y}^*_{N8}$, $\bar{y}^*_{N9}$, $\bar{y}^*_{N10}$, $\bar{y}^*_{N12}$, $\bar{y}^*_{N13}$, $\bar{y}^*_{N14}$, $\bar{y}^*_{N16}$ are more efficient than $\bar{y}_{M(st)}$ and the condition in (4.6) is satisfied for these estimators. But member of $\bar{y}^*_N$ such as $\bar{y}^*_{N3}$, $\bar{y}^*_{N6}$, $\bar{y}^*_{N7}$, $\bar{y}^*_{N11}$ and $\bar{y}^*_{N15}$ aren't more efficient than $\bar{y}_{M(st)}$ and the condition in (4.6) isn't satisfied for these estimators.

6. CONCLUSION

In this study. we have adapted some estimators to the stratified random sampling and we have proposed a family of exponential ratio type estimators which uses the information regarding the population proportion possessing certain attribute. The proposed estimators perform better than classical ratio estimator and adapted estimators known in sampling literature. Using suitable known parameters of auxiliary attributes in estimators give more efficient results in stratified random sampling.

REFERENCES

[1] Abd-Elfattah. A.M.. El-Sherpieny. E.A.. Mohamed. S.M. and Abdou. O.F.. "Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute". ***Applied Mathematics and Computation***. 215: 4198-4202 (2010).

[2] Bahl. S.. Tuteja. R.K.. "Ratio and product exponential estimators". ***J. Inform. Opt. Sci.***. 12(1): 159–164 (1991).

[3] Çıngı. H.. Kadılar. C.. "Advances in sampling theory-ratio method of estimation". ***Bentham Science Publishe***rs. (2009).

[4] Grover. L.K.. Kaur. P.. "An improved exponential estimator of finite population mean in simple random sampling using an auxiliary attribute". ***Applied Mathematics and Computation***. 218: 3093-3099 (2011).

[5] Haq. I.. Hanif. M.. Ahmad. M.. "A family of estimators for single and two-phase sampling using two auxiliary attributes". ***Pakistan Journal of Statistics***. 27(1): 81-94 (2011).

[6] Jhajj. H.S.. Sharma. M.K.. Grover. L.K.. "A family of estimators of population mean using information on auxiliary attribute". ***Pakistan Journal of Statistics***. 22(1): 43-50 (2006).

[7] Koyuncu. N.. "Efficient estimators of population mean using auxiliary attributes". ***Appl. Math. Comput.***. http://dx.doi.org/10.1016/j.amc.2012.04.050. (2012).

[8] Shabbir. J. and Gupta. S.. "On estimating the finite population mean with known population proportion of an auxiliary variable". ***Pakistan Journal of Statistics***. 23 (1): 1-9 (2007).

[9] Shabbir. J.. Gupta. S. "On Improvement in variance estimation using auxiliary information". ***Communications in Statistics – Theory and Methods***. 36(12): 2177-2185 (2007).

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

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

Table 1 .Some Members of $\bar{y}_N^*$ Estimator

	γ_{1st}	γ_{2st}	a_{st}	b_{st}	c_{st}	d_{st}
$\bar{y}_{N1}^*$	1	1	$\beta_{1(p_1)st}$	$\beta_{2(p_1)st}$	$\beta_{1(p_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N2}^*$	1	1	$\beta_{2(p_1)st}$	$\beta_{1(p_1)st}$	$\beta_{2(p_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N3}^*$	1	1	$\rho_{(yp_1)st}$	$\beta_{1(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N4}^*$	1	1	$\rho_{(yp_1)st}$	$\beta_{2(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N5}^*$	1	-1	$\beta_{1(p_1)st}$	$\beta_{2(p_1)st}$	$\beta_{1(p_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N6}^*$	1	-1	$\beta_{2(p_1)st}$	$\beta_{1(p_1)st}$	$\beta_{2(p_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N7}^*$	1	-1	$\rho_{(yp_1)st}$	$\beta_{1(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N8}^*$	1	-1	$\rho_{(yp_1)st}$	$\beta_{2(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N9}^*$	1	0	$\beta_{1(p_1)st}$	$\beta_{2(p_1)st}$	$\beta_{1(p_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N10}^*$	1	0	$\beta_{2(p_1)st}$	$\beta_{1(p_1)st}$	$\beta_{2(p_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N11}^*$	1	0	$\rho_{(yp_1)st}$	$\beta_{1(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N12}^*$	1	0	$\rho_{(yp_1)st}$	$\beta_{2(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N13}^*$	0	1	$\beta_{1(p_1)st}$	$\beta_{2(p_1)st}$	$\beta_{1(p_2)st}$	$\beta_{2(p_2)st}$
$\bar{y}_{N14}^*$	0	1	$\beta_{2(p_1)st}$	$\beta_{1(p_1)st}$	$\beta_{2(p_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N15}^*$	0	1	$\rho_{(yp_1)st}$	$\beta_{1(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{1(p_2)st}$
$\bar{y}_{N16}^*$	0	1	$\rho_{(yp_1)st}$	$\beta_{2(p_1)st}$	$\rho_{(yp_2)st}$	$\beta_{2(p_2)st}$

Table 2. Data Set

N_1=106	N_2=106	N_3=94
N_4=171	N_5=204	N_6=173
n_1=13	n_2=24	n_3=55
n_4=95	n_5=10	n_6=3
S_{y1}=6425.087216	S_{y2}=11551.53	S_{y3}=29907.48
S_{y4}=28643.42	S_{y5}=2389.77	S_{y6}=945.7486
$\overline{Y}_1$=1536.773585	$\overline{Y}_2$=2212.594	$\overline{Y}_3$=9384.309
$\overline{Y}_4$=5588.012	$\overline{Y}_5$=966.9559	$\overline{Y}_6$=404.3988
$S_{\psi_1(1)}$=0.432298	$S_{\psi_1(2)}$=0.45705	$S_{\psi_1(3)}$=0.501656
$S_{\psi_1(4)}$=0.501254	$S_{\psi_1(5)}$=0.481975	$S_{\psi_1(6)}$=0.320681
$P_{1(1)}$=0.24528	$P_{1(2)}$=0.29245	$P_{1(3)}$=0.46808
$P_{1(4)}$=0.48538	$P_{1(5)}$=0.36274	$P_{1(6)}$=0.11561
$S_{\psi_2(1)}$=0.4205	$S_{\psi_2(2)}$=0.43777	$S_{\psi_2(3)}$=0.49338
$S_{\psi_2(4)}$=0.49085	$S_{\psi_2(5)}$=0.41231	$S_{\psi_2(6)}$=0.28221
$P_{2(1)}$=0.22641	$P_{2(2)}$=0.25472	$P_{2(3)}$=0.40426
$P_{2(4)}$=0.39766	$P_{2(5)}$=0.21569	$P_{2(6)}$=0.08671
$\rho_{y\psi_1(1)}$=0.3588	$\rho_{y\psi_1(2)}$=0.26587	$\rho_{y\psi_1(3)}$=0.32486
$\rho_{y\psi_1(4)}$=0.19113	$\rho_{y\psi_1(5)}$=0.38986	$\rho_{y\psi_1(6)}$=0.67316
$\rho_{y\psi_2(1)}$=0.38094	$\rho_{y\psi_2(2)}$=0.29294	$\rho_{y\psi_2(3)}$=0.36653
$\rho_{y\psi_2(4)}$=0.22694	$\rho_{y\psi_2(5)}$=0.51919	$\rho_{y\psi_2(6)}$=0.71532
$\rho_{\psi_1\psi_2(1)}$=0.8442	$\rho_{\psi_1\psi_2(2)}$=0.86172	$\rho_{\psi_1\psi_2(3)}$=0.83468

$\rho_{\psi_1\psi_2(4)}$=0.7649	$\rho_{\psi_1\psi_2(5)}$=0.62069	$\rho_{\psi_1\psi_2(6)}$=0.59525
$\beta_2\left(\psi_{1(1)}\right)$=0.56846	$\beta_2\left(\psi_{1(2)}\right)$=1.16549	$\beta_2\left(\psi_{1(3)}\right)$=-2.02669
$\beta_2\left(\psi_{1(4)}\right)$=2.02028	$\beta_2\left(\psi_{1(5)}\right)$=1.68578	$\beta_2\left(\psi_{1(6)}\right)$=3.927807
$\beta_1\left(\psi_{1(1)}\right)$=1.20109	$\beta_1\left(\psi_{1(2)}\right)$=0.92567	$\beta_1\left(\psi_{1(3)}\right)$=0.1300
$\beta_1\left(\psi_{1(4)}\right)$=0.05902	$\beta_1\left(\psi_{1(5)}\right)$=0.57519	$\beta_1\left(\psi_{1(6)}\right)$=2.42539
$\beta_2\left(\psi_{2(1)}\right)$=0.24603	$\beta_2\left(\psi_{2(2)}\right)$=0.70924	$\beta_2\left(\psi_{2(3)}\right)$=-1.88328
$\beta_2\left(\psi_{2(4)}\right)$=1.84369	$\beta_2\left(\psi_{2(5)}\right)$=0.06085	$\beta_2\left(\psi_{2(6)}\right)$=6.8594
$\beta_1\left(\psi_{2(1)}\right)$=1.32626	$\beta_1\left(\psi_{2(2)}\right)$=1.14215	$\beta_1\left(\psi_{2(3)}\right)$=0.39656
$\beta_1\left(\psi_{2(4)}\right)$=0.42192	$\beta_1\left(\psi_{2(5)}\right)$=1.39278	$\beta_1\left(\psi_{2(6)}\right)$=2.96314