

Improved Exponential Estimator in Stratified Random Sampling

Rajesh Singh
Department of Statistics
B.H.U., Varanasi (U.P.)
India
rsinghstat@yahoo.com

Mukesh Kumar
Department of Statistics
B.H.U., Varanasi (U.P.)
India

Manoj K. Chaudhary
Department of Statistics
B.H.U., Varanasi (U.P.)
India

Cem Kadilar
Department of Statistics
Hacettepe University
Beytepe 06800, Ankara
Turkey
kadilar@hacettepe.edu.tr

Abstract

In this article we have considered the problem of estimating the population mean (\bar{Y}) in the stratified random sampling using the information of an auxiliary variable x which is correlated with y and suggested improved exponential ratio estimators in the stratified random sampling. The mean square error (MSE) equations for the proposed estimators have been derived and it is shown that the proposed estimators under optimum condition performs better than estimators suggested by Singh et al. (2008). Theoretical and empirical findings are encouraging and support the soundness of the proposed estimators for mean estimation.

Keywords: Auxiliary variable, exponential ratio-type estimates, mean square error, stratified random sampling.

2000 AMS Classification: 62D05

1. Introduction

In survey sampling, it is well established that the use of auxiliary information results in substantial gain in efficiency over the estimators which do not use such information. However, in planning surveys, the stratified sampling has often proved needful in improving the precision of estimates over simple random sampling. In some cases, in addition to mean of auxiliary variable, various parameters related to auxiliary variable such as coefficient of variation, kurtosis, correlation coefficient, etc. may also be known. For these cases, many authors,

such as Upadhyaya and Singh (1999), Sisodia and Dwivedi (1981), Singh and Tailor (2003), Singh et al. (2007), developed various estimators to improve the ratio estimators in the simple random sampling. Kadilar and Cingi (2003) adapted the estimators in Upadhyaya and Singh (1999) to the stratified random sampling.

In this study, under stratified random sampling without replacement scheme, we suggest two improved exponential estimators which are more efficient than estimators proposed by Singh et al. (2008). There are some recent studies proposing the family of estimators without using the exponential function in literature, such as Koyuncu and Kadilar (2009(a,b)), Khoshnevisan et al. (2007), Singh and Vishwakarma (2006, 2008) etc. However, in this article, the proposed family of estimators depends on the exponential function.

We assume that the population comprises N units, which can be uniquely partitioned into L strata of size N_1, N_2, \dots, N_L such that $\sum_{h=1}^L N_h = N$. The strata weights $W_h = N_h / N$ ($h=1, 2, \dots, L$) are assumed known. Let (y_{hi}, x_{hi}) ($i=1, 2, \dots, N_h$) denote the values of variates (y, x) respectively for the i -th unit in the h -th stratum and \bar{Y}_h and \bar{X}_h denote stratum means.

When the population mean of the auxiliary variable, \bar{X} , is known, Hansen et al. (1946) suggested a combined ratio estimator for estimating the population mean of the study variable (\bar{Y}) :

$$t_1 = \frac{\bar{y}_{st}}{\bar{x}_{st}} \bar{X}, \quad (1.1)$$

where $\bar{y}_{st} = \sum_{h=1}^L w_h \bar{y}_h$, $\bar{x}_{st} = \sum_{h=1}^L w_h \bar{x}_h$, and $\bar{X} = \sum_{h=1}^L w_h \bar{X}_h$.

Here $\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$, and $\bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$.

The MSE of t_1 to a first degree of approximation is given by

$$MSE(t_1) \cong \sum_{h=1}^L w_h^2 \gamma_h \left[S_{yh}^2 + R^2 S_{xh}^2 - 2RS_{yxh} \right], \quad (1.2)$$

where $\gamma_h = \left(\frac{1}{n_h} - \frac{1}{N_h} \right)$, $R = \frac{\bar{Y}}{\bar{X}} = \frac{\bar{Y}_{st}}{\bar{X}_{st}}$ is the population ratio, S_{yh}^2 is the population variance of a study variable, S_{xh}^2 is the population variance of the auxiliary variable and S_{yxh} is the population covariance between study and auxiliary variables in the stratum h .

2. Kadilar and Cingi estimator

Kadilar and Cingi (2003) introduced an estimator for population mean using known value of some population parameters in stratified random sampling given by

$$t_2 = \frac{\bar{y}_{st}}{\bar{X}_{st,a,b}} \bar{X}_{st,a,b} \quad (2.1)$$

where $\bar{X}_{st,a,b} = \sum_{h=1}^L w_h (a_h \bar{X}_h + b_h)$, $\bar{X}_{st,a,b} = \sum_{h=1}^L w_h (a_h \bar{X}_h + b_h)$ and a_h, b_h are the functions of the known parameters of the auxiliary variable such as coefficient of variation C_{xh} , coefficient of kurtosis $\beta_{2h}(x)$ etc., or some constants in the h^{th} stratum.

The MSE of the estimator t_2 is given by

$$MSE(t_2) = \sum_{h=1}^L W_h^2 \gamma_h \left[S_{yh}^2 + R_{a,b}^2 a_h^2 S_{xh}^2 - 2 R_{a,b} a_h S_{yxh} \right], \quad (2.2)$$

where $R_{a,b} = \frac{\bar{Y}_{st}}{\bar{X}_{st,a,b}}$.

Bahl and Tuteja (1991) suggested an exponential ratio type estimator for population mean in simple random sampling as

$$t_3 = \bar{y} \exp \left[\frac{\bar{X} - \bar{X}}{\bar{X} + \bar{X}} \right]. \quad (2.3)$$

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

$$t_{4i} = \bar{y}_{st} \exp \left[\frac{\bar{X}_{4i} - \bar{X}_{4i}}{\bar{X}_{4i} + \bar{X}_{4i}} \right], \quad i = 0, 1, 2, 3, 4. \quad (2.4)$$

where

$$\begin{aligned} \bar{X}_{40} &= \sum_{h=1}^L w_h \bar{X}_h, & \bar{X}_{40} &= \sum_{h=1}^L w_h \bar{X}_h, \\ \bar{X}_{41} &= \sum_{h=1}^L w_h (\bar{X}_h + C_{xh}), & \bar{X}_{41} &= \sum_{h=1}^L w_h (\bar{X}_h + C_{xh}), \\ \bar{X}_{42} &= \sum_{h=1}^L w_h [\bar{X}_h + \beta_{2h}(x)], & \bar{X}_{42} &= \sum_{h=1}^L w_h [\bar{X}_h + \beta_{2h}(x)], \\ \bar{X}_{43} &= \sum_{h=1}^L w_h [\bar{X}_h \beta_{2h}(x) + C_{xh}], & \bar{X}_{43} &= \sum_{h=1}^L w_h [\bar{X}_h \beta_{2h}(x) + C_{xh}], \\ \bar{X}_{44} &= \sum_{h=1}^L w_h [\bar{X}_h C_{xh} + \beta_{2h}(x)], & \bar{X}_{44} &= \sum_{h=1}^L w_h [\bar{X}_h C_{xh} + \beta_{2h}(x)]. \end{aligned}$$

The bias and MSE of the estimators t_{4i} ($i=0,1,2,3,4$) can be obtained respectively by following expressions:

$$\text{Bias}(t_{4i}) = \frac{1}{\bar{X}_{4i}} \sum_{h=1}^L w_h^2 \gamma_h \left(\frac{3}{8} R_{4i} a_{hi}^2 S_{xh}^2 - \frac{1}{2} a_{hi} S_{yhx} \right), (i=0,1,2,3,4) \quad (2.5)$$

$$\text{MSE}(t_{4i}) = \sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{4i} a_{hi} S_{yhx} + \frac{R_{4i}^2}{4} a_{hi}^2 S_{xh}^2 \right), (i=0,1,2,3,4) \quad (2.6)$$

where,

$$\begin{aligned} R_{40} &= \frac{\bar{Y}_{st}}{\sum_{h=1}^L w_h \bar{X}_h}, & R_{41} &= \frac{\bar{Y}_{st}}{\sum_{h=1}^L w_h (\bar{X}_h + C_{xh})}, \\ R_{42} &= \frac{\bar{Y}_{st}}{\sum_{h=1}^L w_h (\bar{X}_h + \beta_{2h}(x))}, & R_{43} &= \frac{\bar{Y}_{st}}{\sum_{h=1}^L w_h (\bar{X}_h \beta_{2h}(x) + C_{xh})}, \\ R_{44} &= \frac{\bar{Y}_{st}}{\sum_{h=1}^L w_h (\bar{X}_h C_{xh} + \beta_{2h}(x))}, & \bar{Y}_{st} &= \sum_{h=1}^L w_h \bar{Y}_h = \bar{Y}. \end{aligned}$$

We would like to remark that for various values of parameters in (2.4), we get five estimators (for $l = 0, 1, 2, 3, 4$) as shown in Table 1. The MSE expressions for these estimators are also given in the Table 1.

Table 1: Some members of family of estimators t_{4i}

The values of a_h, b_h	Estimator	MSE of the Estimator
$a_h=1, b_h=0$	$t_{40} = \bar{y}_{st} \exp \left[\frac{\bar{X} - \bar{x}_{st}}{\bar{X} + \bar{x}_{st}} \right]$	$\sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{40} S_{yhx} + \frac{R_{40}^2}{4} S_{xh}^2 \right)$
$a_h=1, b_h=C_{xh}$	$t_{41} = \bar{y}_{st} \exp \left[\frac{\bar{X}_{41} - \bar{x}_{41}}{\bar{X}_{41} + \bar{x}_{41}} \right]$	$\sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{41} S_{yhx} + \frac{R_{41}^2}{4} S_{xh}^2 \right)$
$a_h=1, b_h=\beta_{2h}(x)$	$t_{42} = \bar{y}_{st} \exp \left[\frac{\bar{X}_{42} - \bar{x}_{42}}{\bar{X}_{42} + \bar{x}_{42}} \right]$	$\sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{42} S_{yhx} + \frac{R_{42}^2}{4} S_{xh}^2 \right)$
$a_h=\beta_{2h}(x), b_h=C_{xh}$	$t_{43} = \bar{y}_{st} \exp \left[\frac{\bar{X}_{43} - \bar{x}_{43}}{\bar{X}_{43} + \bar{x}_{43}} \right]$	$\sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{43} \beta_{2h}(x) S_{yhx} + \frac{R_{43}^2}{4} \beta_{2h}^2(x) S_{xh}^2 \right)$
$a_h=C_{xh}, b_h=\beta_{2h}(x)$	$t_{44} = \bar{y}_{st} \exp \left[\frac{\bar{X}_{44} - \bar{x}_{44}}{\bar{X}_{44} + \bar{x}_{44}} \right]$	$\sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{44} C_{xh} S_{yhx} + \frac{R_{44}^2}{4} C_{xh}^2 S_{xh}^2 \right)$

Remark 1 - Here we would like to mention that the choice of the estimator depends on the availability and values of the various parameter(s) used (for choice of the parameters a_h and b_h refer to Singh et al. (2008) and Singh and Kumar(2009)).

Singh et al. (2008) showed that all of these modified estimators (t_{4i}) have a smaller MSE than the MSE of the stratified version of exponential ratio type estimator t_{3st} under certain conditions:

$$t_{3st} = \bar{y}_{st} \exp \left[\frac{\bar{X} - \bar{x}_{st}}{\bar{X} + \bar{x}_{st}} \right] \quad (2.7)$$

Note that t_{3st} is the same estimator with t_{40} .

Singh et al. (2008) reported the minimum value of mean square error of the

$$\text{estimator } t_{4i} \text{ as } \text{MSE}(t_{4i})_{\min} = \sum_{h=1}^L w_h^2 \gamma_h S_{yh}^2 (1 - \rho_c^2) \quad (2.8)$$

where, ρ_c is combined correlation coefficient in stratified sampling across all

$$\text{strata. It is calculated as } \rho_c^2 = \frac{\left(\sum_{h=1}^L w_h^2 \gamma_h \rho_h S_{yh} S_{xh} \right)^2}{\sum_{h=1}^L w_h^2 \gamma_h S_{yh}^2 \sum_{h=1}^L w_h^2 \gamma_h S_{xh}^2}.$$

3. Suggested modified exponential ratio estimator

In stratified random sampling, Kadilar and Cingi (2005) introduced a ratio estimator as

$$t_5 = k t_1, \quad (3.1)$$

where the constant k is obtained by minimizing the MSE of t_5 .

Motivated by Kadilar and Cingi (2005), we propose the following modified estimator

$$\begin{aligned} t_{6i} &= k t_{4i} \\ &= k \bar{y}_{st} \exp \left[\frac{\bar{X}_{4i} - \bar{x}_{4i}}{\bar{X}_{4i} + \bar{x}_{4i}} \right], \quad i = 0, 1, 2, 3, 4. \end{aligned} \quad (3.2)$$

To obtain the MSE of t_{6i} to the first degree of approximation, let us define

$$e_0 = \frac{\bar{y}_{st} - \bar{Y}}{\bar{Y}} \text{ and } e_{li}^* = \frac{a_i^* - A_i^*}{A_i^*}. \text{ Using these notations we have}$$

$$\bar{y}_{st} = \bar{Y}(1 + e_0) \text{ and } a_i^* = A_i^*(1 + e_{li}^*),$$

where

$$A_i^* = \sum_{h=1}^L w_h a_{hi} \bar{X}_h, \quad a_i^* = \sum_{h=1}^L w_h a_{hi} \bar{x}_h,$$

$$E(e_0) = E(e_{1i}^*) = 0,$$

$$E(e_0^2) = \frac{1}{\bar{Y}^2} \sum_{h=1}^L w_h^2 \theta_h S_{yh}^2,$$

$$E(e_{li}^2) = \frac{1}{A_i^{*2}} \sum_{h=1}^L w_h^2 \gamma_h a_{hi} S_{xh}^2,$$

$$E(e_0 e_{li}^*) = \frac{1}{\bar{Y} A_i^*} \sum_{h=1}^L w_h^2 \gamma_h a_{hi} S_{yxh}.$$

Expressing (3.2) in terms of e's, we have

$$\begin{aligned} t_{6i} &= k_i \bar{Y} (1 + e_0) \exp \left[-\frac{\theta_i^* e_{li}^*}{2} \left(1 + \frac{\theta_i^* e_{li}^*}{2} \right)^{-1} \right] \\ &\cong k_i \bar{Y} \left[1 + e_0 - \frac{\theta_i^* e_{li}^*}{2} + \frac{3}{8} \theta_i^{*2} e_{li}^{*2} - \frac{\theta_i^* e_0 e_{li}^*}{2} \right]. \end{aligned} \quad (3.3)$$

where $\theta_i^* = \frac{A_i^*}{\bar{X}_{4i}}$.

$$\begin{aligned} \text{Bias}(t_{6i}) &= E(t_{6i} - \bar{Y}) \\ &= \bar{Y}_{st} (k_i - 1) + k_i \bar{Y}_{st} \left(\frac{3\theta_i^* E(e_{li}^{*2})}{8} - \frac{\theta_i^* E(e_0 e_{li}^*)}{2} \right) \\ &= \bar{Y}_{st} (k_i - 1) + \frac{k_i}{\bar{X}_{4i}} \sum_{h=1}^L w_h^2 \gamma_h \left[\frac{3}{8} R_{4i}^2 a_{hi}^2 S_{xh}^2 - \frac{1}{2} a_{hi} S_{yxh} \right] \\ &= \bar{Y}_{st} (k_i - 1) + k_i \text{Bias}(t_{4i}) \end{aligned} \quad (3.4)$$

From (3.3), we have

$$\begin{aligned} \text{MSE}(t_{6i}) &= (k_i - 1)^2 \bar{Y}^2 + k_i^2 \sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{4i} a_{hi} S_{yxh} + \frac{R_{4i}^2}{4} a_{hi}^2 S_{xh}^2 \right) \\ &\quad + 2k_i (k_i - 1) \sum_{h=1}^L w_h^2 \gamma_h \left(\frac{3}{8} R_{4i}^2 a_{hi}^2 S_{xh}^2 - \frac{1}{2} R_{4i} a_{hi} S_{yxh} \right). \end{aligned} \quad (3.5)$$

MSE of the estimator t_{6i} can be rewritten as

$$\text{MSE}(t_{6i}) = (k_i - 1)^2 \bar{Y}^2 + k_i^2 \text{MSE}(t_{4i}) + 2k_i(k_i - 1)\bar{Y}_{st} \text{Bias}(t_{4i}) \quad (3.6)$$

To obtain the optimum value of k , we partially differentiate the expression (3.6) with respect to k and put it equal to zero as follows:

$$\begin{aligned} \frac{\partial}{\partial k} \text{MSE}(t_{6i}) &= 0, \\ \frac{\partial}{\partial k} [k^2 \text{MSE}(t_{4i}) + (k_i - 1)^2 \bar{Y}^2 + 2k_i(k_i - 1)\bar{Y}_{st} \text{Bias}(t_{4i})] &= 0, \\ k_i^* &= \frac{\bar{Y}^2 + A_i}{\bar{Y}^2 + \text{MSE}(t_{4i}) + 2A_i}, \quad i = 0, 1, 2, 3, 4. \end{aligned} \quad (3.7)$$

where

$$A_i = \bar{Y}_{st} \text{Bias}(t_{4i}) = \sum_{h=1}^L w_h^2 \gamma_h \left(\frac{3}{8} R_{4i}^2 a_{hi}^2 S_{xh}^2 - \frac{1}{2} R_{4i} a_{hi} S_{yxh} \right).$$

Note that $0 < k_i^* < 1$.

Remark 2

Shahbaz and Hanif (2009) proposed a general class of shrinkage estimator in survey sampling as

$$\hat{t}_s = \frac{\hat{t}}{1 + T^{-2} \text{MSE}(\hat{t})} \quad (3.8)$$

where \hat{t} is any available estimator of parameter T . The minimum MSE of \hat{t}_s reported by Shahbaz and Hanif (2009) as

$$\text{MSE}(\hat{t}_s) = \frac{\text{MSE}(\hat{t})}{1 + T^{-2} \text{MSE}(\hat{t})} \quad (3.9)$$

Following Shahbaz and Hanif (2009), the estimator t_{6i} is written as

$$\hat{t}_{6iS} = \frac{\hat{t}_{4i}}{1 + \bar{Y}^{-2} \text{MSE}(\hat{t}_{4i})}, \quad i=0,1,2,3,4. \quad (3.10)$$

and the minimum MSE of \hat{t}_{6iS} is given by

$$\text{MSE}(t_{6iS}) = \frac{\text{MSE}(\hat{t}_{4i})}{1 + \bar{Y}^{-2} \text{MSE}(\hat{t}_{4i})}, \quad i=0,1,2,3,4. \quad (3.11)$$

4. Improved estimator

We propose classes of estimators for estimating the population mean (\bar{Y}), in the stratified random sampling as

$$t_{7i} = \omega_{1i} \bar{y}_{st} + \omega_{2i} t_{4i}, \quad i = 0, 1, 2, 3, 4.$$

$$= \omega_{1i} \bar{y}_{st} + \omega_{2i} \bar{y}_{st} \exp \left[\frac{\bar{X}_{4i} - \bar{x}_{4i}}{\bar{X}_{4i} + \bar{x}_{4i}} \right]. \quad (4.1)$$

where ω_{1i} and ω_{2i} are real constants. There are two choices in terms of how to select ω_{1i} and ω_{2i} . Some authors, such as Kadilar and Cingi (2006), use the constraint $\omega_{1i} + \omega_{2i} = 1$, while others use an unconstrained selection of ω_{1i} and ω_{2i} . This latter group includes Upadhyaya et al. (1985), Singh et al. (1988) and Shabbir and Gupta (2207). These authors choose ω_{1i} and ω_{2i} which minimize the MSE for the proposed estimator and do not insist on having $\omega_{1i} + \omega_{2i} = 1$. In this paper we have made unconstrained selection of ω_{1i} and ω_{2i} .

Expressing (4.1) in terms of e's by using the same notations as defined earlier we have

$$t_{7i} = \omega_{1i} \bar{Y}(1 + e_0) + \omega_{2i} \bar{Y} \exp \left[-\frac{\theta_i^* e_0}{2} \left(1 + \frac{\theta_i^* e_{li}}{2} \right)^{-1} \right] \quad (4.2)$$

Expanding the right hand side of (4.2), and retaining the terms to the first order of approximation, we have

$$t_{7i} \cong \omega_{1i} \bar{Y}(1 + e_0) + \omega_{2i} \bar{Y} \left(1 + e_0 - \frac{\theta_i^* e_{li}}{2} + \frac{3\theta_i^{*2} e_{li}^2}{8} - \frac{\theta_i^* e_0 e_{li}}{2} \right) \quad (4.3)$$

The bias of the estimator t_{7i} , to the first order of approximation, is obtained as

$$\text{Bias}(t_{7i}) = E(t_{7i} - \bar{Y})$$

$$\begin{aligned} &\cong (\omega_{1i} + \omega_{2i} - 1) \bar{Y}_{st} + \omega_{2i} \sum_{h=1}^L w_h^2 \gamma_h \left(\frac{3}{8} R_{4i}^2 a_{hi}^2 S_{xh}^2 - \frac{1}{2} R_{4i} a_{hi} S_{yhx} \right) \\ &= (\omega_{1i} + \omega_{2i} - 1) \bar{Y}_{st} + \omega_{2i} \bar{Y}_{st} \text{Bias}(t_{4i}) \\ &= (\omega_{1i} + \omega_{2i} - 1) \bar{Y}_{st} + \omega_{2i} A_i \end{aligned} \quad (4.4)$$

From equation (4.4), we observe that to the first order of approximation the proposed estimator is biased.

From (4.3), squaring and then retaining the terms of e's upto power two, we have

$$\text{MSE}(t_{7i}) = E(t_{7i} - \bar{Y})^2$$

$$\begin{aligned}
&= E \left\{ \omega_{1i} \bar{Y}(1 + e_0) + \omega_{2i} \bar{Y} \left(1 + e_0 - \frac{\theta_i^* e_{1i}^*}{2} + \frac{3\theta_i^{*2} e_{1i}^{*2}}{8} - \frac{\theta_i^* e_0 e_{1i}^*}{2} \right) - \bar{Y}_{st} \right\}^2 \\
&= E \left\{ \bar{Y}(\omega_{1i} + \omega_{2i} - 1) + \omega_{1i} e_0 \bar{Y} + \omega_{2i} \bar{Y} \left(e_0 - \frac{\theta_i^* e_{1i}^*}{2} + \frac{3\theta_i^{*2} e_{1i}^{*2}}{8} - \frac{\theta_i^* e_0 e_{1i}^*}{2} \right) \right\}^2 \\
&\cong \bar{Y}^2 (\omega_{1i} + \omega_{2i} - 1)^2 + \omega_{1i}^2 D + \omega_{2i}^2 \text{MSE}(t_{4i}) \\
&\quad + 2\omega_{1i}\omega_{2i}C_i + 2(\omega_{1i} + \omega_{2i} - 1)\omega_{2i}A_i
\end{aligned} \tag{4.5}$$

where

$$C_i = \sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - \frac{R_{4i}}{2} a_{hi} S_{yhx} \right) \text{ and } D = \sum_{h=1}^L w_h^2 \gamma_h S_{yh}^2.$$

In order to find optimum value of estimator t_{7i} , we differentiate MSE (t_{7i}) with respect to ω_{1i} and ω_{2i} . The optimum values of ω_{1i} and ω_{2i} are

$$\omega_{1i}^* = \frac{(\bar{Y}^2 + A_i)(\bar{Y}^2 + C_i + A_i) - \bar{Y}^2 [\bar{Y}^2 + \text{MSE}(t_{4i}) + 2A_i]}{(\bar{Y}^2 + C_i + A_i)^2 - (\bar{Y}^2 + D)[\bar{Y}^2 + \text{MSE}(t_{4i}) + 2A_i]} \tag{4.6}$$

and

$$\omega_{2i}^* = \frac{(\bar{Y}^2 + A_i + C_i)\bar{Y}^2 - (\bar{Y}^2 + D)(\bar{Y}^2 + A_i)}{(\bar{Y}^2 + C_i + A_i)^2 - (\bar{Y}^2 + D)[\bar{Y}^2 + \text{MSE}(t_{4i}) + 2A_i]}. \tag{4.7}$$

Using these optimum values, ω_{1i}^* and ω_{2i}^* , in the place of ω_{1i} and ω_{2i} in (4.5), respectively, we get the minimum MSE of the estimator t_{7i} .

Remark 3 - Here we would like to mention that the optimum values of ω_{1i} and ω_{2i} require information about population parameters. However, in applications, to have this information is sometimes impossible. However in repeated surveys or studies based on multiphase sampling, where information is gathered on several occasions (or based on past experience) it may be possible to guess the value of these parameters quite accurately. Even though this approach may reduce the precision, it may be satisfactory provided the relative decrease in precision is marginal. (For further details refer to Sukhatme et. al. ((1984), pp. 260), Singh and Vishwakarma (2006) and Koyuncu and Kadilar (2009(b)).

5. Efficiency comparisons

For the estimators, efficiency comparisons are given below.

In this section, we first compare modified estimator t_{6i} given in (3.2), with the estimator t_{4i} given in (2.4).

$$\begin{aligned} \text{MSE}(t_{6i}) &< \text{MSE}(t_{4i}), \quad i = 0, 1, 2, 3, 4, \\ (k_i - 1)^2 \bar{Y}^2 + k_i^2 \sum_{h=1}^L w_h^2 \gamma_h &\left(S_{yh}^2 - R_{4i} a_h S_{yxh} + \frac{R_{4i}^2}{4} a_h^2 S_{yxh}^2 \right) \\ &+ 2k_i(k_i - 1) \sum_{h=1}^L w_h^2 \gamma_h \left(\frac{3R_{4i}^2}{8} a_{hi}^2 S_{xh}^2 - \frac{1}{2} R_{4i} a_{hi} S_{xyh} \right) \\ &< \sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{4i} a_{hi} S_{yxh} + \frac{R_{4i}^2}{4} a_{hi}^2 S_{xh}^2 \right), \\ k_i &> \frac{\bar{Y}^2 - \text{MSE}(t_{4i})}{\bar{Y}^2 + \text{MSE}(t_{4i}) + 2A_i}. \end{aligned} \quad (5.1)$$

As the condition (5.1) is always satisfied, we can say that the suggested estimator is more efficient than exponential ratio estimator in stratified random sampling in all conditions. Note that when $i = 0$, $\text{MSE}(t_{40}) = \text{MSE}(t_{3st})$.

Second, we compare the improved estimator t_{7i} , given in (4.1), with estimator t_{4i} , given in (2.4) as follows:

$$\begin{aligned} \text{MSE}(t_{7i}) &< \text{MSE}(t_{4i}), \quad i = 0, 1, 2, 3, 4, \\ \bar{Y}^2 (\omega_{1i} + \omega_{2i} - 1)^2 + \omega_{1i}^2 D + \omega_{2i}^2 \text{MSE}(t_{4i}) &+ 2\omega_{1i}\omega_{2i}C_i + 2(\omega_{1i} + \omega_{2i} - 1)\omega_{2i}A_{ii} \\ &< \sum_{h=1}^L w_h^2 \gamma_h \left(S_{yh}^2 - R_{4i} a_{hi} S_{yxh} + \frac{R_{4i}^2}{4} a_{hi}^2 S_{xh}^2 \right). \end{aligned}$$

Let $(\omega_{1i} + \omega_{2i} - 1) = B_i$. Then the expressions can be obtained on solving as

$$\bar{Y}^2 < \frac{\Delta_{1i}^* + \Delta_{2i}^*}{\Delta_{3i}^*}, \quad (5.2)$$

where

$$\begin{aligned} \Delta_{1i}^* &= \text{MSE}(t_{4i})(1 - \omega_{2i}^2), \\ \Delta_{2i}^* &= 2A_i[-\omega_{2i}B_i] - \omega_{1i}^2 D - 2\omega_{1i}\omega_{2i}C_i, \\ \Delta_{3i}^* &= B_i^2. \end{aligned}$$

Next, we compare the improved estimator t_{7i} given in (4.1) with the estimator t_{6i} given in (3.2).

$$\begin{aligned} \text{MSE}(t_{7i}) &< \text{MSE}(t_{6i}), \quad i = 0, 1, 2, 3, 4, \\ \bar{Y}^2(\omega_{1i} + \omega_{2i} - 1)^2 + \omega_{1i}^2 D + \omega_{2i}^2 \text{MSE}(t_{4i}) + 2\omega_{1i}\omega_{2i}C_i + 2(\omega_{1i} + \omega_{2i} - 1)\omega_{2i}A_i \\ &< \bar{Y}^2(k_i - 1)^2 + k_i^2 \text{MSE}(t_{4i}) + 2k_i(k_i - 1) \sum_{h=1}^L w_h^2 \gamma_h \left(\frac{3R_{4i}^2}{8} a_{hi}^2 S_{xh}^2 - \frac{1}{2} R_{4i} a_{hi} S_{xyh} \right), \\ \bar{Y}^2 &< \frac{\Delta_{1i} + \Delta_{2i}}{\Delta_{3i}}, \end{aligned} \quad (5.3)$$

where

$$\begin{aligned} \Delta_{1i} &= \text{MSE}(t_{4i})(k_i^2 - \omega_{2i}^2), \\ \Delta_{2i} &= 2A_i[k_i(k_i - 1) - \omega_{2i}B] - \omega_{1i}^2 D - 2\omega_{1i}\omega_{2i}C_i, \\ \Delta_{3i} &= B_i^2 - (k_i - 1)^2. \end{aligned}$$

Finally, we compare the improved estimator t_{7i} , given in (4.1), with estimator $\text{MSE}(t_{4i})_{\min}$ given in (2.8).

$$\begin{aligned} \text{MSE}(t_{7i}) &< \text{MSE}(t_{4i})_{\min}, \quad i = 0, 1, 2, 3, 4, \\ \bar{Y}^2(\omega_{1i} + \omega_{2i} - 1)^2 + \omega_{1i}^2 D + \omega_{2i}^2 \text{MSE}(t_{4i}) + 2\omega_{1i}\omega_{2i}C_i + 2(\omega_{1i} + \omega_{2i} - 1)\omega_{2i}A_i \\ &< \sum_{h=1}^L w_h^2 \gamma_h S_{yh}^2 (1 - \rho_c^2). \end{aligned}$$

or

$$\bar{Y}^2 < \frac{\Delta^* + \Delta_{1i}^* + \Delta_{2i}^*}{\Delta_{3i}^*} \quad (5.4)$$

where,

$$\begin{aligned} \Delta^* &= \text{MSE}(t_{4i})_{\min} \\ \Delta_{1i}^* &= \text{MSE}(t_{4i})(-\omega_{2i}^2), \\ \Delta_{2i}^* &= 2A_i[-\omega_{2i}B_i] - \omega_{1i}^2 D - 2\omega_{1i}\omega_{2i}C_{ii}, \\ \Delta_{3i}^* &= B_i^2. \end{aligned}$$

In all above expressions we used optimum values of ω_{1i} , ω_{2i} , and k_i , i.e., $\omega_{1i} = \omega_{1i}^*$, $\omega_{2i} = \omega_{2i}^*$, and $k_i = k_i^*$, respectively.

Table 2: Data Statistics

N=854	n=140	$\beta_x = 312.07$
$C_x=3.85$	$C_y=5.84$	$S_x=144794$
$S_y=17106$	$\rho = 0.92$	
$N_1=106$	$N_2=106$	$N_3=94$
$N_4=171$	$N_5=204$	$N_6=173$
$n_1=9$	$n_2=17$	$n_3=38$
$n_4=67$	$n_5=7$	$n_6=2$
$\bar{X}_1 = 24375$	$\bar{X}_2 = 27421$	$\bar{X}_3 = 72409$
$\bar{X}_4 = 74365$	$\bar{X}_5 = 26441$	$\bar{X}_6 = 9844$
$\bar{Y}_1 = 5336$	$\bar{Y}_2 = 2212$	$\bar{Y}_3 = 9384$
$\bar{Y}_4 = 5588$	$\bar{Y}_5 = 967$	$\bar{Y}_6 = 404$
$\beta_{x1} = 25.71$	$\beta_{x2} = 34.57$	$\beta_{x3} = 26.14$
$\beta_{x4} = 97.60$	$\beta_{x5} = 27.47$	$\beta_{x6} = 28.10$
$C_{x1}=2.02$	$C_{x2}=2.10$	$C_{x3}=2.22$
$C_{x4}=3.84$	$C_{x5}=1.72$	$C_{x6}=1.91$
$C_{y1}=4.18$	$C_{y2}=5.22$	$C_{y3}=3.19$
$C_{y4}=5.13$	$C_{y5}=2.47$	$C_{y6}=2.34$
$S_{x1}=49189$	$S_{x2}=57461$	$S_{x3}=160757$
$S_{x4}=285603$	$S_{x5}=45403$	$S_{x6}=18794$
$S_{y1}=6425$	$S_{y2}=11552$	$S_{y3}=29907$
$S_{y4}=28643$	$S_{y5}=2390$	$S_{y6}=946$
$\rho_1 = 0.82$	$\rho_2 = 0.86$	$\rho_3 = 0.90$
$\rho_4 = 0.99$	$\rho_5 = 0.71$	$\rho_6 = 0.89$
$\gamma_1 = 0.102$	$\gamma_2 = 0.049$	$\gamma_3 = 0.016$
$\gamma_4 = 0.009$	$\gamma_5 = 0.138$	$\gamma_6 = 0.006$
$w_1^2 = 0.015$	$w_2^2 = 0.015$	$w_3^2 = 0.012$
$w_4^2 = 0.04$	$w_5^2 = 0.057$	$w_6^2 = 0.041$
$\bar{X} = 37600$	$\bar{Y} = 2930$	$\bar{X}_{41}=37642$
$\bar{X}_{42}=37602$	$\bar{X}_{43}=2086644$	$\bar{X}_{44}=102815$
$R_{40}=0.07790$	$R_{41}=0.07789$	$R_{42}=0.07780$
$R_{43}=0.00140$	$R_{44}=0.02850$	
$A_{40}= -68606.40$	$A_{41}= -68611.10$	$A_{42}= -68692.40$
$A_{43}= -23961.80$	$A_{44}= -63674.00$	
$C_{40}= 476871.32$	$C_{41}= 476884.00$	$C_{42}= 477111.29$
$C_{43}= 439626.58$	$C_{44}= 457109.01$	$D = 691035.70$

6. Numerical study

For empirical study, we use the data set earlier presented by Kadilar and Cingi (2003). In this data set, Y is the apple production amount and X is the number of apple trees in 854 villages of Turkey in 1999. The population information about this data set is given in Table 2.

By using this data, we have calculated MSE values of suggested estimators and compared them with MSE values of Singh et al. (2008) estimators. For the different values of a_h and b_h , we can find the MSE equations of all members of the improved estimator t_{7i} by simply changing R_{4i} , respectively.

Table 3: MSE Values of Estimators

Values of R	MSE(t_{4i}) (Bias)	MSE(t_{6i}) (Bias)	MSE(t_{7i}) (Bias)	Values of k_i^* , * _{1i} and * _{2i}
$R_{40}=0.07790$	359745.60 (-23.5173)	350039.68 (-119.9887)	217839.68 (-74.6724)	$k_0^* = 0.9667$ $\omega_{10}^* = -1.1917$ $\omega_{20}^* = 2.1837$
$R_{41}=0.07789$	359760.09 (-23.5189)	350053.50 (-119.9934)	217839.99 (-74.6725)	$k_1^* = 0.9667$ $\omega_{11}^* = -1.1918$ $\omega_{21}^* = 2.1839$
$R_{42}=0.07780$	360008.22 (-23.4940)	350290.60 (-120.0747)	217845.36 (-71.3671)	$k_2^* = 0.9665$ $\omega_{12}^* = -1.1942$ $\omega_{22}^* = 2.1863$
$R_{43}=0.00140$	339848.98 (-8.2138)	328512.90 (-112.6096)	267874.93 (-91.8238)	$k_3^* = 0.9641$ $\omega_{13}^* = -0.6360$ $\omega_{23}^* = 1.6091$
$R_{44}=0.0285$	336684.09 (-21.8266)	328136.37 (-112.4806)	208197.14 (-74.6743)	$k_4^* = 0.9687$ $\omega_{14}^* = -1.0460$ $\omega_{24}^* = 2.0368$

In Table 3 we have reported the bias and MSE values of the different estimators. Here we would like to mention that the MSE of the different estimators depend on choice of a_h and b_h . The minimum MSE attained by Singh et al. (2008) estimator was 218374.8898. From Table 3 we observe that the estimator t_{7i} ($i=0,1,2,3,4$)

under optimum conditions performs better than all other estimators proposed by Singh et al. (2008). We also observe from the table that the estimator t_{73} is having larger MSE than the minimum MSE attained by Singh et al. (2008). So, for this data set choice of $a_h = \beta_{2h}(x)$ and $b_h = C_{xh}$ is not a good choice. Also, for the choice $a_h = C_{xh}$ and $b_h = \beta_{2h}(x)$ the MSEs of the estimators t_{44} , t_{64} and t_{74} is minimum.

7. Conclusion

As shown in theory, we also observe from Table 3 that suggested estimators t_{6i} (including optimal value) performs better than corresponding estimators without optimal values, t_{4i} . Besides, we see that improved estimator t_{7i} and its members are always more efficient than corresponding estimators (without optimal value of k_i and with k_i^*) for this data set.

Acknowledgements

The authors are thankful to the referee for his valuable comments and suggestions regarding the improvement of the paper. The second author (Mukesh Kumar) is grateful to UGC, New Delhi, India, for providing financial assistance.

References

1. Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimator. *Information and Optimization Science* XIII 159-163.
2. Hansen, M.H., Hurwitz, W.N. and Gurney, M. (1946). Problem and methods of the sample survey of business. *Journal of American Statistical Association* 41 174-189.
3. Kadilar, C. and Cingi, H. (2003). Ratio estimators in stratified random sampling, *Biometrical Journal* 45 (2), 218-225.
4. Kadilar, C. and Cingi, H. (2005). A new ratio estimator in stratified random sampling, *Communications in Statistics: Theory and Methods* 34 (3), 597-602.
5. Kadilar, C., Cingi, H. (2006). Improvement in variance estimation using auxiliary information. *Hacettepe Journal of Mathematical Statistics* 35(1):111–115.
6. Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N., and Smarandache, F. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s), *Far East Journal of Theoretical Statistics* 22 181–191.

7. Koyuncu, N. and Kadilar, C. (2009(a)). Ratio and product estimators in stratified random sampling, *Journal of Statistical Planning and Inference* 139 (8), 2552-2558.
8. Koyuncu, N. and Kadilar, C. (2009(b)). Family of estimators of population mean using two auxiliary variables in stratified random sampling. *Communications in Statistics: Theory and Methods* 38 (14), 2398-2417.
9. Shabiir, J. and Gupta, S. (2007): On improvement in variance estimation using auxiliary information. *Communications in Statistics: Theory and Methods* 36: 2177-2185.
10. Shahbaz, M. Q. and Hanif, M. (2009): A general shrinkage estimator in survey sampling. *World Applied Sciences Journal* 7(5): 593-596.
11. Singh, H. P. and Vishwakarma, G. K. (2006). Combined ratio-product estimator of finite population mean in stratified sampling, *Metodologia de Encuestas* 8 35-44.
12. Singh, H. P. and Vishwakarma, G. K. (2008). A family of estimators of population mean using auxiliary information in stratified sampling. *Communications in Statistics: Theory and Methods* 37 1038-1050.
13. Singh, H. P., Upadhyaya, L. N., Namjoshi, D. (1988). Estimation of finite population variance. *Current Science* 57(24):1331–1334.
14. Singh, H.P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. *Statistics in Transition* 6 555-560.
15. Singh, R. and Kumar, M. (2009): A note on transformations on auxiliary variable in survey sampling. *Model Assisted Statistics and Applications* (Accepted).
16. Singh, R., Cauhan, P., Sawan, N., and Smarandache, F. (2007). *Auxiliary Information and A Priori Values in Construction of Improved Estimators*. Renaissance High Press.
17. Singh, R., Kumar, M. and Smarandache, F. (2008): Almost Unbiased Estimator for Estimating Population Mean Using Known Value of Some Population Parameter(s). *Pakistan Journal of Statistics and Operations Research* 4(2) pp63-76.
18. Singh, R., Kumar, M., Singh, R.D., and Chaudhary, M.K. (2008): Exponential ratio type estimators in stratified random sampling, Presented in International Symposium on Optimisation and Statistics (I.S.O.S) at A.M.U., Aligarh, India, during 29-31 Dec. 2008.
19. Sisodia, B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable, *Journal of Indian Society of Agricultural Statistics* 33 13-18.
20. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Ashok, C. (1984). *Sampling Theory of Surveys with Applications*. Iowa State University Press, Ames, Iowa and ICAR, New Delhi.

21. Upadhyaya, L. N., Singh, H. P., Vos, J. W. E. (1985). On the estimation of population means and ratios using supplementary information. *Statistica Neerlandica* 39:309–318.
22. Upadhyaya, L.N. and Singh, H.P. (1999). An estimator for population variance that utilizes the kurtosis of an auxiliary variable in sample surveys. *Vikram Mathematical Journal* 19 14-17.