STATISTICS IN TRANSITION new series, Spring 2015
Vol. 16, No. 1, pp. 37–52

SOME CLASSES OF MODIFIED RATIO TYPE ESTIMATORS IN SAMPLE SURVEYS

A. K. P. C. Swain[1], **Manjula Das**[2]

ABSTRACT

In this paper some classes of modified ratio type estimators with additive and multiplicative adjustments made to the simple mean per unit estimator and classical ratio estimator are suggested to obtain more efficient ratio type estimators compared to the classical one. Their biases and mean square errors are obtained and compared with first order approximations.

Key words: ratio type estimator, simple random sampling, bias, mean square error, efficiency.

1. Introduction

In sample surveys it is usual practice to look for information on auxiliary variables which are either available from official records or can be collected inexpensively in the course of investigation. In the case of single auxiliary variable the ratio estimator and the regression estimator are two classical estimators making use of the auxiliary information to improve the efficiency of the finite population parameters such as population mean, total, variance, etc. Although simple to compute, the ratio estimator is always less efficient than the linear regression estimator in large samples.

But the theory of linear regression is not very much appropriate for the sample survey situations(Cochran,1953) and requires that the assumptions such as: (a) existence of linearity of regression of y on x in the population;

(b) constancy of residual variance;

(c) infinite nature of population;

should be approximately satisfied, but are rarely satisfied in finite population sampling.

[1] Former Professor of Statistics, Utkal University, Bhubaneswar-751004, India.
E-mail:akpcs@rediffmail.com.

[2] Associate Professor, Department of Mathematics, ITER, SOA University, Bhubaneswar-751030, India.

This has motivated some research workers to look for different techniques to form ratio type estimators whose mean square errors approximate to that of the approximate mean square error of the linear regression estimate in large samples.Srivastava(1967) modified the ratio estimator with power transformation of the ratio of the population mean to the sample mean whose minimum mean square error to first approximation equals to that of the linear regression estimator.Srivastava(1971) proposed a class of estimators having minimum mean square error equal to that of the linear regression estimator,provided certain regularity conditions are satisfied. In this paper we make a variety of additive and multiplicative adjustments to the simple mean per unit estimator and classical ratio estimator so that their large sample mean square errors attain the minimum mean square bound of Srivastava's class of estimators, which is in fact the large sample mean square error of the linear regression estimator. The proposed classes of estimators are compared as regards their large sample biases.

Let $U = (U_1, U_2, ..., U_N)$ be a finite population of N distinct and identifiable units. Let y and x denote the study variable and auxiliary variable respectively taking paired values (Y_i, X_i) on the unit $U_i (i = 1, 2, ..., N)$ .Assume x to be positively correlated with y. Further, assume that y and x are positively measured.

Define $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$, $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$, $R = \frac{\bar{Y}}{\bar{X}}$

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 \quad , S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2 ,$$

$$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X}) , C_y^2 = \frac{S_y^2}{\bar{Y}^2} \quad , C_x^2 = \frac{S_x^2}{\bar{X}^2} \text{ and}$$

$C_{yx} = \frac{S_{yx}}{\overline{YX}} = \rho C_y C_x , \rho$ being the correlation coefficient between y and x .The population regression coefficient of y on x is defined as $\beta = \frac{S_{yx}}{S_x^2}$.

Let $u_1, u_2, ..., u_n$ be a simple random sample s of size n units drawn without replacement from U .We observe paired values $(y_i, x_i), i = 1, 2, ..., n$ on the sampled units.

Define $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i , \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$,

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2, s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

$s_{yx} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})$ and $r = \frac{\bar{y}}{\bar{x}}$. The sample regression coefficient of y on x is defined as $b = \frac{s_{yx}}{s_x^2}$.

To estimate the population mean $\bar{Y}$ of the study variable y, the classical ratio estimator $\hat{\bar{Y}}_R$ is defined by

$$\hat{\bar{Y}}_R = \bar{y}(\frac{\bar{X}}{\bar{x}}), \quad (1.1)$$

where $\bar{X}$ is assumed to be known in advance.

It is well known (Cochran,1953)that $\hat{\bar{Y}}_R$ is a biased estimate of the population mean $\bar{Y}$ with bias to $O(1/n)$ given by

$$Bias(\hat{\bar{Y}}_R) = \theta\bar{Y}(C_x^2 - \rho C_y C_x) = -\theta\bar{Y}(\frac{\beta}{R} - 1)C_x^2$$

$$= -\theta\bar{Y}(K-1)C_x^2 \text{ ,where } K = \frac{\beta}{R} \quad (1.2)$$

$\theta = (\frac{1}{n} - \frac{1}{N}), C_y$ and C_x being the coefficients of variation of y and x respectively,.

Further, up to terms of $O(1/n)$, the mean square error of $\hat{\bar{Y}}_R$ is given by

$$MSE(\hat{\bar{Y}}_R) = \theta\bar{Y}^2(C_y^2 + C_x^2 - 2\rho C_y C_x), \quad (1.3)$$

$\hat{\bar{Y}}_R$ is more efficient than $\bar{y}$

$$\text{if} \quad \rho > \frac{1}{2}(\frac{C_x}{C_y}) \quad (1.4)$$

Besides ratio method of estimation, linear regression method of estimation is another early method initiated by Watson (1937), making use of auxiliary information in sample surveys. The simple regression estimate of the population mean $\bar{Y}$ is given by

$$\bar{y}_{lr} = \bar{y} + b(\bar{X} - \bar{x}), \quad (1.5)$$

where b is the linear regression coefficient of y on x, calculated from the sample.

The mean square error of $\bar{y}_{lr}$ to $O(1/n)$ is given by

$$MSE(\bar{y}_{lr}) = \theta\bar{Y}^2 C_y^2 (1-\rho^2) \tag{1.6}$$

Both ratio and regression estimators are biased estimators and biases decrease with increase in the sample size. Comparing (1.3) and (1.6) it may be seen that in large samples $\hat{\bar{Y}}_R$ is always less efficient than $\bar{y}_{lr}$ unless the regression line of y on x passes through the origin, in which case they have equal efficiency.

Srivastava (1967) suggested a class of power transformation estimator

$$\hat{\bar{Y}}_{SR} = \bar{y}(\frac{\bar{X}}{\bar{x}})^{\alpha}, \tag{1.7}$$

where α is a real number to be suitably chosen. The optimum value of $MSE(\hat{\bar{Y}}_{SR})$,when optimized with respect to α, gives the expression given in (1.6).

Walsh (1970) suggested an alternative class of ratio-type estimator where x_i is transformed to z_i such that

$z_i = \alpha x_i + (1-\alpha)\bar{X}$

Hence, $\bar{z} = \alpha\bar{x} + (1-\alpha)\bar{X}$ and $\bar{Z} = \bar{X}$.

As such, modified ratio type estimator is formed as

$$\hat{\bar{Y}}_{WR} = \frac{\bar{y}}{\bar{z}}\bar{Z} = \frac{\bar{y}}{\alpha\bar{x} + (1-\alpha)\bar{X}}\bar{X} \tag{1.8}$$

To first order approximations of the optimum mean square error of $\hat{\bar{Y}}_{WR}$ is given by

$MSE(\hat{\bar{Y}}_{WR}) = \theta\bar{Y}^2 C_y^2 (1-\rho^2)$, as given in (1.6).

Srivastava (1971) proposed a generalized class of estimators given by

$$t_g = \bar{y}H(u) \tag{1.9}$$

where $u = \bar{x}/\bar{X}$ and $H(.)$ is a parametric function satisfying certain regularity conditions as given in Srivastava (1971), such as

(i) $H(1) = 1$

(ii) The first and second order derivatives of H with respect to u exist and are known constants at a given point $u = 1$.

He also showed that the asymptotic mean square error of t_g cannot be reduced further than $\min.MSE(t_g) = \theta\bar{Y}^2 C_y^2 (1-\rho^2)$, which is the approximate mean square error of the linear regression estimator, which is the lower bound to mean square error of class of estimators t_g .Prabhu-Ajgaonkar(1993) has noted that an optimum estimator does not exist uniformly in the class t_g .

Srivastava (1980) defined another wider class of estimators as

$$t_w = H(\bar{y}, u) \tag{1.10}$$

where $H(\bar{y}, u)$ is a function of $\bar{y}$ and u, satisfying certain regularity conditions specified by him. He showed that asymptotic minimum mean square error of t_w cannot be reduced further than that given in (1.6).

Thus, ratio estimator $\hat{\bar{Y}}_R$, Srivastava's power transformation estimator $\hat{\bar{Y}}_{SR}$ and Walsh's estimator $\hat{\bar{Y}}_{WR}$ are the special cases of t_g .The wider class t_w includes regression estimator besides ratio estimator, power transformation estimator and many others.

Swain (2013) proposed a class of estimators

$$\hat{\bar{Y}}_{SWR} = \bar{y}\left[\alpha\left(\frac{\bar{X}}{\bar{x}}\right)^g + (1-\alpha)\left(\frac{\bar{x}}{\bar{X}}\right)^h\right]^\delta, \tag{1.11}$$

which is also a subclass of t_g and where α, g, h and δ are free real constants to be suitably chosen and also the asymptotic mean square error $\hat{\bar{Y}}_{SWR}$ is equal to asymptotic mean square error of the linear regression estimator given in (1.6). Both $\hat{\bar{Y}}_{SR}$ and $\hat{\bar{Y}}_{WR}$ are the special cases of $\hat{\bar{Y}}_{SWR}$, which can be further generalized as

$$\hat{\bar{Y}}^*_{SWR} = \bar{y}\left[\alpha\left(\frac{A\bar{X}+B}{A\bar{x}+B}\right)^g + (1-\alpha)\left(\frac{A\bar{x}+B}{A\bar{X}+B}\right)^h\right]^\delta,$$

$$= \bar{y}\left[\alpha\left(\frac{\bar{X}+d}{\bar{x}+d}\right)^g + (1-\alpha)\left(\frac{\bar{x}+d}{\bar{X}+d}\right)^h\right]^\delta, \tag{1.12}$$

where $d = B/A$ and α, g, h, d and δ are free parameters to be suitably chosen.

We may arbitrarily specify any four of the aforesaid parameters and minimize the approximate mean square error with respect to the remaining one and the resulting mean square error equals the approximate mean square error of the linear regression estimator which is the lower bound to the mean square error of the class of estimators defined by t_g . To choose best estimator in this class the survey practitioner should select those set values for the unspecified parameters for which the first order bias is zero or approximately so.

In the following some adjustments are made to $\bar{y}$ and $\hat{\bar{Y}}_R$ to construct some classes of modified ratio type estimators to provide more efficient estimators of the population mean $\bar{Y}$, and the proposed classes of estimators, which are sub-classes of Srivastava's (1971,1980) classes of estimators,are compared as regards their biases and mean square errors.

2. Proposed classes of estimators

Consider the following classes of estimators, where $H(u)$ is as defined by Srivastava (1971).

$$T_1 = \bar{y}\left[H(u)\right]^{\mu}$$

$$T_2 = \bar{y}\left[\frac{1}{\mu H(u) + (1-\mu)}\right]$$

$$T_3 = \hat{\bar{Y}}_R\left[H(u)\right]^{\mu}$$

$$T_4 = \frac{\hat{\bar{Y}}_R}{\mu H(u) + (1-\mu)}$$

$$T_5 = \bar{y} + \mu(1 - H(u))$$

$$T_6 = \hat{\bar{Y}}_R + \mu(1 - H(u))$$

Expanding $H(u)$ by the value 1 in the second order Taylor's series we have

$$H(u) = H\left[1 + (u-1)\right] = H(1) + (u-1)(\frac{\partial H}{\partial u})_{u=1} + \frac{1}{2}(u-1)^2(\frac{\partial^2 H}{\partial u^2})_{u=1} + ... \quad (2.1)$$

Assuming $|u-1| < 1$, the higher order terms can be neglected and we write

$$T_1 = \bar{y}\left[1 + (u-1)H_1 + (u-1)^2 H_2 + ...\right]^{\mu}, \quad (2.2)$$

where $H_1 = (\frac{\partial H}{\partial u})_{u=1}$ and $H_2 = \frac{1}{2}(\frac{\partial^2 H}{\partial u^2})_{u=1}$ denote the first and second order partial derivatives of H with respect to u and are the known constants.

Thus, we write

$$T_1 = \bar{Y}(1+e_0)(1+e_1 H_1 + e_1^2 H_2 + ...)^{\mu}, \tag{2.3}$$

where $e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $e_1 = u - 1$.

Expanding (2.3) in power series we have

$$T_1 = \bar{Y}(1+e_0)\left[1+\mu(H_1 e_1 + H_2 e_1^2) + \frac{\mu(\mu-1)}{2}(H_1 e_1 + H_2 e_1^2)^2 + ...\right].$$

To first order of approximations

$$MSE(T_1) = \theta\bar{Y}^2(C_y^2 + \mu^2 H_1^2 C_x^2 + 2\mu H_1 C_{yx}) \tag{2.4}$$

$$Bias(T_1) = B(T_1) = \theta\bar{Y}(\mu H_1 C_{yx} + \mu H_2 C_x^2 + \frac{\mu(\mu-1)}{2} H_1^2 C_x^2) \tag{2.5}$$

where $E(e_1^2) = \theta C_x^2$ and $E(e_0 e_1) = \theta C_{yx}$.

Differentiating $MSE(T_1)$ with respect to μ and putting it equal to zero, we have optimum μ given by

$$\mu_{opt} = -\frac{K}{H_1}.$$

Thus, the optimum mean square error of T_1 obtained by substituting the optimum value of H_1 in (2.4) is given by

$$MSE(T_1) = \theta\bar{Y}^2 C_y^2(1-\rho^2), \tag{2.6}$$

which is equal to that of the large sample mean square error of the linear regression estimator given by (1.6).

Also, the bias of T_1 with optimum μ is given by

$$\text{Bias }(T_1) = \theta\bar{Y}K\left[-K - (\frac{H_2}{H_1}) + \frac{(K+H_1)}{2}\right]C_x^2 \tag{2.7}$$

Proceeding as before we find to $O(1/n)$

$$MSE(T_2) = MSE(T_3) = MSE(T_4) = MSE(T_5) = MSE(T_6) = \theta\bar{Y}^2 C_y^2 (1-\rho^2)$$

$$Bias(T_2) = -\theta\bar{Y}K \frac{H_2}{H_1} C_x^2$$

$$Bias(T_3) = \theta\bar{Y}(1-K)\left[\left(\frac{H_2}{H_1} - \frac{H_1}{2}\right) + \frac{(1+K)}{2} \right] C_x^2$$

$$Bias(T_4) = \theta(1-K)\left(\frac{H_2}{H_1} + 1\right) C_x^2$$

$$Bias(T_5) = -\theta\bar{Y}K\left(\frac{H_2}{H_1}\right) C_x^2$$

$$Bias(T_6) = \theta\bar{Y}K\left(H_1 - \frac{H_2}{H_1}\right) C_x^2$$

The biases and mean square errors of different classes of estimators are summarized in Table 1.

Table 1. Biases and mean square errors

Classes of Estimators	μ_{opt}	$Bias$	MSE_{opt}
$T_1 = \bar{y}\left[H(u)\right]^{\mu}$	$-(\frac{K}{H_1})$	$\theta\bar{Y}K\left[-K - \frac{H_2}{H_1} + \frac{K+H_1}{2}\right]C_x^2$	$\theta\bar{Y}^2 C_y^2 (1-\rho^2)$
$T_2 = \bar{y}\left[\frac{1}{\mu H(u) + (1-\mu)}\right]$	$\frac{K}{H_1}$	$-\theta\bar{Y}K\left[\frac{H_2}{H_1}\right]C_x^2$	$\theta\bar{Y}^2 C_y^2 (1-\rho^2)$
$T_3 = \hat{\bar{Y}}_R\left[H(u)\right]^{\mu}$	$\frac{1-K}{H_1}$	$\theta\bar{Y}(1-K)(\frac{H_2}{H_1} - \frac{H_1}{2} + \frac{1+K}{2})C_x^2$	$\theta\bar{Y}^2 C_y^2 (1-\rho^2)$
$T_4 = \left[\frac{\hat{\bar{Y}}_R}{\mu H(u) + (1-\mu)}\right]$	$\frac{K-1}{H_1}$	$\theta\bar{Y}(1-K)\left[\frac{H_2}{H_1} + 1\right]C_x^2$	$\theta\bar{Y}^2 C_y^2 (1-\rho^2)$
$T_5 = \bar{y} + \mu\left[1 - H(u)\right]$	$\frac{K\bar{Y}}{H_1}$	$-\theta\bar{Y}K\left[\frac{H_2}{H_1}\right]C_x^2$	$\theta\bar{Y}^2 C_y^2 (1-\rho^2)$
$T_6 = \hat{\bar{Y}}_R + \mu\left[1 - H(u)\right]$	$\frac{(K-1)}{H_1}$	$\theta\bar{Y}K\left[H_1 - \frac{H_2}{H_1}\right]C_x^2$	$\theta\bar{Y}^2 C_y^2 (1-\rho^2)$

3. Some special cases of proposed classes of estimators

By defining $H(u)$ differently we may generate different classes of estimators and some of them related to ratio and product estimators are given in Table 2.

Table 2. Estimators and their Biases excluding the common multiplier

Class of estimators	$H(u)=\bar{x}/\bar{X}$	*Bias*	$H(u)=\bar{X}/\bar{x}$	*Bias*
$T_1=\bar{y}[H(u)]^{\mu}$	$T_{11}=\bar{y}\left[\bar{x}/\bar{X}\right]^{\mu}$	$\frac{K(K-1)}{2}$	$T_{12}=\bar{y}\left[\bar{X}/\bar{x}\right]^{\mu}$	$\frac{K(K-1)}{2}$
$T_2=\bar{y}\left[\frac{1}{\mu H(u)+(1-\mu)}\right]$	$T_{21}=\bar{y}\left[\frac{1}{\mu(\bar{x}/\bar{X})+(1-\mu)}\right]$	0	$T_{22}=\bar{y}\left[\frac{1}{\mu(\bar{X}/\bar{x})+(1-\mu)}\right]$	K
$T_3=\hat{\bar{Y}}_R\left[H(u)\right]^{\mu}$	$T_{31}=\hat{\bar{Y}}_R\left[\bar{x}/\bar{X}\right]^{\mu}$	$\frac{K(K-1)}{2}$	$T_{32}=\hat{\bar{Y}}_R\left[\bar{X}/\bar{x}\right]^{\mu}$	$\frac{K(K-1)}{2}$
$T_4=\left[\frac{\hat{\bar{Y}}_R}{\mu H(u)+(1-\mu)}\right]$	$T_{41}=\left[\frac{\hat{\bar{Y}}_R}{\mu(\bar{x}/\bar{X})+(1-\mu)}\right]$	0	$T_{42}=\left[\frac{\hat{\bar{Y}}_R}{\mu(\bar{X}/\bar{x})+(1-\mu)}\right]$	$1-K$
$T_5=\bar{y}+\mu\left[1-H(u)\right]$	$T_{51}=\bar{y}+\mu\left[1-(\bar{x}/\bar{X})\right]$	0	$T_{52}=\bar{y}+\mu\left[1-(\bar{X}/\bar{x})\right]$	K
$T_6=\hat{\bar{Y}}_R+\mu\left[1-H(u)\right]$	$T_{61}=\hat{\bar{Y}}_R+\mu\left[1-(\bar{x}/\bar{X})\right]$	K	$T_{62}=\hat{\bar{Y}}_R+\mu\left[1-(\bar{X}/\bar{x})\right]$	0

We find (Table 2) the first order biases of T_{21}, T_{41}, T_{51} and T_{62} vanish, having the same approximate mean square error as that of the linear regression estimator. Since the optimum value of μ is a usually unknown parametric function $K=\frac{\beta}{R}$, we estimate it by its consistent estimator $\hat{K}=\frac{b}{r}$ from the sample.

Thus, the estimators T_{21}, T_{41}, T_{51} and T_{62} with estimated values of K are given by

$$\hat{T}_{21}=\frac{\bar{y}}{(\frac{b}{r})(\bar{x}/\bar{X})+(1-\frac{b}{r})}$$

$$\hat{T}_{41} = \frac{\hat{\bar{Y}}_R}{\left[(\frac{b-r}{r})(\bar{x}/\bar{X}) + (2-\frac{b}{r}) \right]}$$

$$\hat{T}_{51} = \bar{y} + b(\bar{X} - \bar{x})$$

$$\hat{T}_{62} = \hat{\bar{Y}}_R + \bar{y}(1-\frac{b}{r})(1-\frac{\bar{X}}{\bar{x}}) = \bar{X}\left[r + (r-b)(1-\frac{\bar{X}}{\bar{x}}) \right]$$

3.1. Bias and mean square error of $\hat{T}_{21}, \hat{T}_{41}, \hat{T}_{51}$ and $\hat{T}_{62}$

(i)
$$\hat{T}_{21} = \frac{\bar{y}}{\frac{b}{r}(\bar{x}/\bar{X}) + (1-\frac{b}{r})}$$

Or alternatively, $$\hat{T}_{21} = \frac{\bar{y}}{\bar{x}}\bar{X}\left[\frac{r}{b + (r-b)(\bar{X}/\bar{x})} \right] \quad (3.1)$$

Define
$$e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}, e_1 = \frac{\bar{x}-\bar{X}}{\bar{X}}, e_2 = \frac{s_{yx}-S_{yx}}{S_{yx}}, e_3 = \frac{s_x^2 - S_x^2}{S_x^2}$$

Expanding $\hat{T}_{21}$ using binomial series expansion with assumptions $|e_1| < 1$ and $|e_3| < 1$ for all possible samples and keeping terms up to second degree we have

$$\hat{T}_{21} = \bar{Y}\left[1 + (e_0 - \frac{\beta}{R}e_1) + \frac{\beta}{R}(e_1e_3 - e_1e_2) + \frac{\beta}{R}(\frac{\beta}{R}-1)e_1^2 \right] + \ldots$$

To first order approximations, that is to $O(1/n)$,

$$Bias(\hat{T}_{21}) = \bar{Y}\left[(\frac{\beta}{R}-1)(\frac{\beta}{R})\frac{V(\bar{x})}{\bar{X}^2} + \frac{\beta}{R}(\frac{Cov(s_x^2,\bar{x})}{S_x^2\bar{X}} - \frac{Cov(s_{yx},\bar{x})}{S_{yx}\bar{X}}) \right] \quad (3.2)$$

where $Cov(s_x^2,\bar{x}) = \frac{N(N-n)}{(N-1)(N-2)}\frac{\mu_{03}}{n}$ and

$$Cov(s_{yx},\bar{x}) = \frac{N(N-n)}{(N-1)(N-2)}\frac{\mu_{12}}{n}$$

with $\mu_{rs} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{Y})^r (x_i - \bar{X})^s, i = 1,2,3,4,....$ (see Sukhatme, Sukhatme and Asok,1984)

$$MSE(\hat{T}_{21}) = \theta \bar{Y}^2 C_y^2 (1-\rho^2) \tag{3.3}$$

Under bivariate normality of (y, x) or for symmetrical populations,

$$Bias(\hat{T}_{21}) = \frac{1}{n}\bar{Y}\left[(\frac{\beta}{R})(\frac{\beta}{R}-1)\right]C_x^2 \tag{3.4}$$

(ii) $$\hat{T}_{41} = \frac{\hat{\bar{Y}}_R}{(\frac{b}{r}-1)(\bar{x}/\bar{X}) + (2-\frac{b}{r})} = \frac{\bar{y}}{(\frac{b}{r}-1)(\bar{x}/\bar{X})^2 + (2-\frac{b}{r})(\bar{x}/\bar{X})} \tag{3.5}$$

On expansion $\hat{T}_{41} = \bar{Y} + \bar{Y}\left[(e_0 - \frac{\beta}{R}e_1) + \frac{\beta}{R}(e_3 e_1 - e_2 e_1) + (\frac{\beta}{R}-1)^2 e_1^2\right] + ...$

$$Bias(\hat{T}_{41}) = \theta \bar{Y}(\frac{\beta}{R}-1)^2 C_x^2 + \bar{Y}\frac{\beta}{R}\left[\frac{Cov(s_x^2, \bar{x})}{S_x^2 \bar{X}} - \frac{Cov(s_{yx}, \bar{x})}{S_{yx}\bar{X}}\right] \tag{3.6}$$

$$MSE(\hat{T}_{41}) = \theta \bar{Y}^2 C_y^2 (1-\rho^2) \tag{3.7}$$

Under bivariate normal populations or for symmetrical populations

$$Bias(\hat{T}_{41}) = \frac{1}{n}\bar{Y}(\frac{\beta}{R}-1)^2 C_x^2 \tag{3.8}$$

(iii) $$\hat{T}_{51}) = \bar{y} + b(\bar{X} - \bar{x}) \tag{3.9}$$

Expanding $\hat{T}_{51}$ using Binomial series with assumptions $|e_1| < 1$ and $|e_3| < 1$ we have

$$\hat{T}_{51} = \bar{Y} + \bar{Y}\left[(e_0 - \frac{\beta}{R}e_1) + \frac{\beta}{R}(e_3 e_1 - e_2 e_1)\right] + ...$$

Thus, $Bias(\hat{T}_{51}) = \bar{Y}\left[\frac{\beta}{R}(\frac{Cov(s_x^2, \bar{x})}{S_x^2 \bar{X}} - \frac{Cov(s_{yx}, \bar{x})}{S_{yx}\bar{X}})\right]$ (Sukhatme et al., 1984)

(3.10)

$$\text{and } MSE(\hat{T}_{51}) = \theta \bar{Y}^2 C_y^2 (1-\rho^2) \quad (3.11)$$

Under Bivariate normality $Bias(\hat{T}_{51})$ vanishes.

(iv) $$\hat{T}_{62} = \hat{\bar{Y}}_R + \hat{\mu}_{opt}(1-\frac{\vec{X}}{\bar{x}}) = \bar{X}\left[r + (r-b)(1-\frac{\bar{X}}{\bar{x}})\right] \quad (3.12)$$

Expanding $\hat{T}_{62}$ using binomial series expansion with assumptions $|e_1|<1$ and $|e_3|<1$ for all possible samples and keeping terms up to second degree we have

$$\hat{T}_{62} = \bar{Y} + \bar{Y}\left[(e_0 - \frac{\beta}{R}e_1) + \frac{\beta}{R}(e_1e_3 - e_1e_2) + (\frac{\beta}{R}-1)e_1^2\right]$$

$$E(\hat{T}_{62}) = \bar{Y} + \bar{Y}\left[(\frac{\beta}{R}-1)\frac{V(\bar{x})}{\bar{X}^2} + \frac{\beta}{R}(\frac{Cov(s_x^2,\bar{x})}{S_x^2\bar{X}} - \frac{Cov(s_{yx},\bar{x})}{S_{yx}\bar{X}})\right]$$

$$Bias(\hat{T}_{62}) = \bar{Y}\left[(\frac{\beta}{R}-1)\frac{V(\bar{x})}{\bar{X}^2} + \frac{\beta}{R}(\frac{Cov(s_x^2,\bar{x})}{S_x^2\bar{X}} - \frac{Cov(s_{yx},\bar{x})}{S_{yx}\bar{X}})\right] \quad (3.13)$$

$$MSE(\hat{T}_{62}) = \theta \bar{Y}^2 C_y^2 (1-\rho^2) \quad (3.14)$$

Under bivariate normality of (y, x) or for symmetrical populations

$Cov(s_x^2,\bar{x})$ and $Cov(s_{yx},\bar{x})$ vanish and thus to $O(1/n)$

$$Bias(\hat{T}_{62}) = \frac{1}{n}\bar{Y}(\frac{\beta}{R}-1)C_x^2 \quad (3.15)$$

3.2. Comparison of biases and mean square errors of $\hat{T}_{21}, \hat{T}_{41}, \hat{T}_{51}$ and $\hat{T}_{62}$

To first order of approximations, that is to $O(1/n)$
the mean square errors of $\hat{T}_{21}, \hat{T}_{41}, \hat{T}_{51}$ and $\hat{T}_{62}$ are equal to that of the linear regression estimator. Further,

(i) $\hat{T}_{21}$ is less biased than the regression estimator $\hat{T}_{51}$ if $|A+B|<|A|$

(ii) $\hat{T}_{41}$ is less biased than the regression estimator $\hat{T}_{51}$ if $|A+C|<|A|$

(iii) $\hat{T}_{62}$ is less biased than the regression estimator $\hat{T}_{51}$ if $|A+D|<|A|$

where $$A=\bar{Y}\left[\frac{\beta}{R}\left(\frac{Cov(s_x^2,\bar{x})}{S_x^2\bar{X}}-\frac{Cov(s_{yx},\bar{x})}{S_{yx}\bar{X}}\right)\right]$$

$$B=\theta\bar{Y}\left[\frac{\beta}{R}\left(\frac{\beta}{R}-1\right)\right]C_x^2$$

$$C=\theta\bar{Y}\left(\frac{\beta}{R}-1\right)^2C_x^2$$

$$D=\theta\bar{Y}\left(\frac{\beta}{R}-1\right)C_x^2$$

Under bivariate normality or for symmetrical populations

$$Bias(\hat{T}_{21})=\frac{1}{n}\bar{Y}\left(\frac{\beta}{R}\right)\left(\frac{\beta}{R}-1\right)C_x^2$$

$$Bias(\hat{T}_{41})=\frac{1}{n}\bar{Y}\left(\frac{\beta}{R}-1\right)^2C_x^2$$

$$Bias(\hat{T}_{51})=0$$

$$Bias(\hat{T}_{62})=\frac{1}{n}\bar{Y}\left(\frac{\beta}{R}-1\right)C_x^2$$

$$Bias(\hat{Y}_R)=-\frac{1}{n}\bar{Y}\left(\frac{\beta}{R}-1\right)C_x^2$$

$\hat{T}_{21}$ is less biased than $\hat{T}_{41}$, if $\frac{\beta}{R}<1/2$, and less biased than $\hat{T}_{62}$, if $\frac{\beta}{R}<1$.

Thus, $\hat{T}_{21}$ is less biased than both $\hat{T}_{41}$ and $\hat{T}_{62}$, if $\frac{\beta}{R}<1/2$.

$\hat{T}_{41}$ is less biased than $\hat{T}_{62}$ if $\left(\frac{\beta}{R}-1\right)^2<1$

Further , $\left|Bias\hat{\bar{Y}}_R\right|=\left|Bias\hat{T}_{62}\right|$

4. Numerical illustration

To estimate the total number of milch animals in 117 villages of zone 4 of Haryana state of India in 1977-78 a simple random sample of size 17 was selected. The number of milch animals in the survey (y) and the number of milch animals in the previous census (x) were observed for each village in the sample

(Singh and Chaudhary, 1986). The estimated values of approximate bias except the common multiplier are given in Table 3.

Table 3. Biases of estimators

Estimator	Absolute Bias excepting common multiplier
$\hat{T}_{21}$	0.01143
$\hat{T}_{41}$	0.01155
$\hat{T}_{51}$	0.01150
$\hat{T}_{62}$	0.01137

Comment: $\hat{T}_{62}$ is least biased among the competitors.

5. Conclusions

(i) Without assuming restrictive assumptions associated with the linear regression estimator ,the proposed modified ratio-type estimators $\hat{T}_{21}, \hat{T}_{41}$ and $\hat{T}_{62}$ are asymptotically as efficient as the linear regression estimator $\hat{T}_{51}(\bar{y}_{lr})$.

(ii) Under bivariate normality the first order bias of $\hat{T}_{51}$ is zero. Further, $\hat{T}_{21}$ is less biased than both $\hat{T}_{41}$ and $\hat{T}_{62}$ if $\frac{\beta}{R} < 1/2$, and $\hat{T}_{41}$ is less biased than $\hat{T}_{62}$ if $(\frac{\beta}{R} - 1)^2 < 1$

Further, $\hat{\bar{Y}}_R$ and $\hat{T}_{62}$ have same absolute bias

(iii) Numerical illustration shows that up to first order of approximations $\hat{T}_{62}$ is less biased than $\hat{T}_{21}, \hat{T}_{41}$ and $\hat{T}_{51}$ although the differences are marginal.

(iv) $H(u)$ may also be defined as exponential functions of u such as

$H(u) = Exp\left[\alpha(1-u)\right]$,where α is a real constant,

$H(u) = Exp\left[\frac{1-u}{1+u}\right]$

$H(u) = a^{1-u}$,where a is a non-zero positive real constant,etc.

Acknowledgement

The authors thank the esteemed referees for their helpful comments which helped improving the contents and presentation of results in the paper.

REFERENCES

COCHRAN, W. G., (1953). Sampling Techniques, Wiley, New York.

PRABHU-AJGAONKAR, S. G., (1993). Non-existence of an optimum estimator in a class of ratio estimators,Statistical Papers, 34, 1, 161−165.

SINGH, D., CHAUDHARY, F. S., (1986). Theory and Analysis of Sample Survey Designs, New Age International (P)Ltd. Publishers, New Delhi.

SRIVASTAVA, S. K., (1967). An estimator using auxiliary information in sample surveys, Bull. Cal. Stat Assoc., 16, 121−132.

SRIVASTAVA, S. K., (1971). Generalized estimator for mean of a finite population using multi auxiliary information. Journ. Amer. Stat. Assoc., 66, 404−407.

SRIVASTAVA, S. K., (1980). A class of estimator using auxiliary information in sample surveys. Canadian Journal of Statistics, 8, 253−254.

SUKHATME, P. V., SUKHATME, B. V., ASOK, C., (1984). Sampling theory of surveys with Applications, Asia Publishing House, New Delhi.

SWAIN, A. K. P. C., (2013). On some modified ratio and product type estimators-Revisited, Investigacion Operacional, Vol. 34, 1, 35−57.

WALSH, J. E., (1970). Generalization of ratio estimate of population total, Sankhya, A., 32, 141−46.

WATSON, D. J., (1937). The estimation of leaf areas, Journ. Ag. Sc., 27, 474.