Journal of Reliability and Statistical Studies [ISSN: 0974-8024 (Print), 2229-5666 (Online)]
Vol. 3, Issue 2 (2010): 93-101

ON CLASSES OF UNBIASED SAMPLING STRATEGIES

Shashi Bhushan[1] and Shubhra Katara[2]

1. Department of Statistics, PUC Campus, Mizoram University, Aizawl – Mizoram, India.
E Mail: shashi.py@gmail.com
2. Department of Statistics, Bareilly College, Bareilly, U.P., India.
E Mail: shubhraonthenet@gmail.com.

Abstract

In this paper, we have proposed to use two classes of sampling strategies based on the modified ratio estimator using the standard deviation and the coefficient of skewness of the auxiliary variable by Singh (2003) for estimating the population mean (total) of the study variable in a finite population. The properties of the proposed sampling strategies are studied and some concluding remarks are given. Also, an empirical study is included as an illustration.

Keywords: Ratio estimator, Variance, Unbiasedness, Mean square error.

1. Introduction

The use of information on an auxiliary variable for increasing the efficiency of sampling strategy is quite well established in sampling from finite populations. Many transformed estimation procedures are also available in the literature for increasing the efficiency but, unfortunately, this is always done at the cost of unbiasedness. There are numerous instances where the bias of these estimation procedures is very large with respect to their mean square error and therefore may not be advisable. The aim of this paper is to introduce some efficient as well as unbiased sampling strategies for finite population. Further, it may be of interest to note that there are very few works in sampling literature wherein the focus is on devising better sampling strategies both in terms of efficiency and unbiasedness. This may be done by some innovations in both the aspects of sampling strategy viz namely sampling scheme and estimation procedure. In this paper, we have used the prior information about coefficient of coefficient of skewness and standard deviation at both the stages – estimation and sampling scheme so that the accuracy of the sampling strategy is improved. Recently, some attempts have been made by various authors, including Singh (2003), Senapati (2005), Singh (2005) among others, to improve the existing sampling strategies by using information about auxiliary variable.

The use of prior value of coefficient of skewness and standard deviation in estimating the population mean of characteristic under study y was made by Singh (2003). It was mentioned by the above author that the use of such prior information about the coefficient of skewness and standard deviation leads to more efficient estimation population mean (total) of the study variable.

The ratio estimator for estimating the population mean of the study variable y is given by

$$\bar{y}_R = \frac{\bar{y}}{\bar{x}}\bar{X} = \hat{R}\bar{X} \qquad (1.1)$$

where $\hat{R} = \frac{\bar{y}}{\bar{x}}$, $\bar{y}$ is the sample mean of the study variable y, $\bar{x}$ is the sample mean of the some auxiliary variable x and $\bar{X}$ is the population mean of the x which is assumed to be known.

If the population standard deviation of x denoted by σ_x and population coefficient of skewness denoted by β_{1x} is known, then Singh (2003) proposed a modified product estimator for estimating the population mean $\bar{Y}$ of the study variable given by

$$\bar{y}_{SP} = \bar{y}\frac{\left(\bar{x}\beta_{1x} + \sigma_x\right)}{\left(\bar{X}\beta_{1x} + \sigma_x\right)} \tag{1.2}$$

The bias and mean square error of $\bar{y}_{SP}$ under simple random sampling are given by

$$Bias\left(\bar{y}_{SP}\right) = (\frac{1}{n} - \frac{1}{N})\bar{Y}\rho\nu C_x C_y \tag{1.3}$$

and

$$MSE\left(\bar{y}_{SP}\right) = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2\left\{C_y^2 + \nu C_x^2\left(\nu + 2K\right)\right\} \tag{1.4}$$

where C_y is the population coefficient of variation of y, $\nu = \bar{X}\beta_{1x}/\left(\bar{X}\beta_{1x} + \sigma_x\right)$, $K = \rho\, C_y/C_y$ and ρ is the population correlation coefficient between x and y. Similar to (1.2), a modified ratio estimator can be given by

$$\bar{y}_{SR} = \bar{y}\frac{\left(\bar{X}\beta_{1x} + \sigma_x\right)}{\left(\bar{x}\beta_{1x} + \sigma_x\right)}$$

having bias and mean square error given by

$Bias\left(\bar{y}_{SR}\right) = (\frac{1}{n} - \frac{1}{N})\bar{Y}\left(\nu^2 C_x^2 - \rho\nu C_x C_y\right)$ and

$MSE\left(\bar{y}_{SR}\right) = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2\left\{C_y^2 + \nu C_x^2\left(\nu - 2K\right)\right\}$ respectively.

In this paper, Singh (2003) estimator is modified using Walsh (1970) and Reddy (1973). We propose two classes of unbiased sampling strategies such that the estimator of population mean $\bar{Y}$ is

$$\bar{y}_{SA} = \bar{y}\frac{\left(\bar{X}\beta_{1x} + \sigma_x\right)}{\left\{A\left(\bar{x}\beta_{1x} + \sigma_x\right) + (1 - A)\left(\bar{X}\beta_{1x} + \sigma_x\right)\right\}} \tag{1.5}$$

Note that for $A = 1$ the proposed estimator reduces to $\bar{y}_{SR}$. We now consider this estimator under the following sampling schemes:

(i). Simple random sampling without replacement along with the jack-knife technique and denote the resulting estimator as $\hat{\bar{Y}}_{SS}^{*}$.

(ii). Midzuno (1952) - Lahiri (1951) - Sen (1952) type sampling scheme and denote the resulting estimator by $\bar{y}_{SM}$.

Both the sampling strategies aim at getting some classes of better sampling strategies than the existing ones in the sense of unbiasedness and lesser mean square error.

2. Bias and MSE under SRSWOR

Consider estimator $\bar{y}_{SA}$ under Simple random sampling without replacement and denote it by $\bar{y}_{SS}$.

Let $\bar{y} = \bar{Y} + e_0$ and $\bar{x} = \bar{X} + e_1$ such that $E(e_0) = E(e_1) = 0$ (2.1)

Putting these values in (1.5) and using Taylor’s series expansion, we have

$$\bar{y}_{SS} - \bar{Y} = e_0 - \frac{A\beta_{1x}\bar{Y}e_1}{\bar{X}\beta_{1x} + \sigma_x} + \frac{A^2\beta_{1x}^2\bar{Y}e_1^2}{(\bar{X}\beta_{1x} + \sigma_x)^2} - \frac{A\beta_{1x}e_0e_1}{\bar{X}\beta_{1x} + \sigma_x} + \mathbf{h} \quad (2.2)$$

Taking expectation on both sides and using (2.1) we have

$$Bias\left(\bar{y}_{SS}\right) = E\left(\bar{y}_{SS}\right) - \bar{Y}$$

$$= \bar{Y}\left\{\frac{A^2\beta_{1x}^2}{(\bar{X}\beta_{1x} + \sigma_x)^2}E(e_1^2) - \frac{A\beta_{1x}}{(\bar{X}\beta_{1x} + \sigma_x)}E(e_0e_1)\right\} \quad \text{(upto first order of approximation)}$$

Since $E(e_0^2) = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2C_y^2$

$$E(e_1^2) = (\frac{1}{n} - \frac{1}{N})\bar{X}^2C_x^2$$

$$E(e_0e_1) = (\frac{1}{n} - \frac{1}{N})\bar{X}\bar{Y}\rho C_xC_y \quad (2.3)$$

Therefore,

$$Bias\left(\bar{y}_{SS}\right) = \bar{Y}(\frac{1}{n} - \frac{1}{N})\left\{A^2\nu^2C_x^2 - A\nu\rho C_xC_y\right\} \quad (2.4)$$

Now, for mean square error, considering (2.2) to the first order of approximation, we get

$$MSE(\bar{y}_{SS}) = E(\bar{y}_{SS} - \bar{Y})^2$$

$$= E\left\{e_0 - \frac{A\beta_{1x}\bar{Y}e_1}{\bar{X}\beta_{1x} + \sigma_x}\right\}^2$$

$$= E(e_0^2) + \frac{A^2\beta_{1x}^2\bar{Y}^2}{(\bar{X}\beta_{1x} + \sigma_x)^2}E(e_1^2) - \frac{2A\beta_{1x}\bar{Y}}{(\bar{X}\beta_{1x} + \sigma_x)}E(e_0e_1)$$

$$= \bar{Y}^2(\frac{1}{n} - \frac{1}{N})\left\{C_y^2 + \nu C_x^2\left(A^2\nu - 2AK\right)\right\} \quad (2.5)$$

Note that on putting $A = 1$ in (2.4) and (2.5) we get expressions of bias and mean square error of $\bar{y}_{SR}$ and the optimizing value of the characterizing scalar A is given by

$$A = \frac{K}{\nu} = A_{opt} \text{ (say)} \quad (2.6)$$

The minimum mean square error under optimizing value of $A = A_{opt}$ is

$$MSE\left(\bar{y}_{SS}\right) = \bar{Y}^2(\frac{1}{n} - \frac{1}{N})(1 - \rho^2)C_y^2 \quad (2.7)$$

which is same as the mean square error of the linear regression estimator. Also note that $Bias\left(\bar{y}_{SS}\right)=0$ under the optimizing value of A.

3. Unbiasedness and MSE of Jack-Knife Sampling Strategy

Let us now apply Quenouille's (1956) method of Jack-Knife such that the sample of size $n=2m$ from a population of size N is split up at random into two sub samples of size m each. For further details one may refer to Gray and Schucany (1972). Let us define

$$\bar{y}_{ss}^{(i)}=\bar{y}_i\,\frac{\bar{X}\beta_{1x}+\sigma_x}{A\left(\bar{x}_i\beta_{1x}+\sigma_x\right)+(1-A)\left(\bar{X}\beta_{1x}+\sigma_x\right)};\ i=1,2$$

$$\bar{y}_{ss}^{(3)}=\bar{y}_s\,\frac{\bar{X}\beta_{1x}+\sigma_x}{A\left(\bar{x}_s\beta_{1x}+\sigma_x\right)+(1-A)\left(\bar{X}\beta_{1x}+\sigma_x\right)} \tag{3.1}$$

where is the characterizing scalar to be chosen suitably such that $s=s_1+s_2$, s_1 and s_2 be the two sub samples of size m each and $+$ denotes the disjoint union. $\bar{y}_1$, $\bar{y}_2$ and $\bar{y}_s$ denote the sample means based on two sub samples of size m and the entire sample of size $n=2m$ for characteristic y. $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_s$ denote the sample means based on two sub samples of size m and the entire sample of size $n=2m$ for characteristic x.
It can be easily seen that

$$Bias\left(\bar{y}_{ss}^{(i)}\right)=\bar{Y}\left(\frac{1}{m}-\frac{1}{N}\right)\left\{A^2\nu^2C_x^2-A\nu\rho C_xC_y\right\};\ i=1,2$$

$$Bias\left(\bar{y}_{ss}^{(3)}\right)=\bar{Y}\left(\frac{1}{2m}-\frac{1}{N}\right)\left\{A^2\nu^2C_x^2-A\nu\rho C_xC_y\right\}=B_1(say) \tag{3.2}$$

Let us define $\hat{\bar{Y}}_{SS}{}'=\dfrac{\bar{y}_{SS}^{(1)}+\bar{y}_{SS}^{(2)}}{2}$ as an alternative estimator of the population mean $\bar{Y}$.

The bias of $\hat{\bar{Y}}_{SS}{}'$ is

$$Bias(\hat{\bar{Y}}_{SS}{}')=\bar{Y}\left(\frac{1}{m}-\frac{1}{N}\right)\left\{A^2\nu^2C_x^2-A\nu\rho C_xC_y\right\}=B_2(say) \tag{3.3}$$

We propose the jackknife estimator $\hat{\bar{Y}}_{SS}{}^{*}$ for estimating population mean $\bar{Y}$ given by

$$\hat{\bar{Y}}_{SS}{}^{*}=\frac{\bar{y}_{SS}^{(3)}-R\hat{\bar{Y}}_{SS}{}'}{1-R}=\frac{\bar{y}_{SS}^{(3)}-\left\{\dfrac{N-2n}{2(N-n)}\right\}\hat{\bar{Y}}_{SS}{}'}{1-\left\{\dfrac{N-2n}{2(N-n)}\right\}}\quad\text{where } R=\frac{B_1}{B_2} \tag{3.4}$$

Taking expectation of (3.4) and using (3.2) and (3.3) we obtain
$E(\hat{\bar{Y}}_{SS}{}^{*})=\bar{Y}$ showing that $\hat{\bar{Y}}_{SS}{}^{*}$ is an unbiased estimator of population mean $\bar{Y}$ to the first order of approximation.
Consider

$$MSE(\hat{\bar{Y}}_{SS}{}^{*})=E\left(\frac{\bar{y}_{SS}^{(3)}-R\hat{\bar{Y}}_{SS}{}'}{1-R}-\bar{Y}\right)^2$$

$$= \frac{1}{(1-R)^2} E\left\{\bar{y}_{SS}^{(3)} - R\hat{\bar{Y}}_{SS}' - \bar{Y} + R\bar{Y}\right\}^2$$

$$= \frac{1}{(1-R)^2}\left\{E(\bar{y}_{SS}^{(3)} - \bar{Y})^2 + R^2 E(\hat{\bar{Y}}_{SS}' - \bar{Y})^2 - 2RE(\bar{y}_{SS}^{(3)} - \bar{Y})(\hat{\bar{Y}}_{SS}' - \bar{Y})\right\} \quad (3.5)$$

Also,

$$E(\bar{y}_{SS}^{(3)} - \bar{Y})^2 = MSE(\bar{y}_{SS}^{(3)}) = \bar{Y}^2\left(\frac{1}{2m} - \frac{1}{N}\right)\left\{C_y^2 + \nu C_x^2\left(A^2\nu - 2AK\right)\right\} \quad (3.6)$$

Further,

$$E(\hat{\bar{Y}}_{SS}' - \bar{Y})^2 = E\left(\frac{\bar{y}_{SS}^{(1)} + \bar{y}_{SS}^{(2)}}{2} - \bar{Y}\right)^2$$

$$= \frac{1}{4}E\left\{(\bar{y}_{SS}^{(1)} - \bar{Y}) + (\bar{y}_{SS}^{(2)} - \bar{Y})\right\}^2$$

$$= \frac{1}{4}\left\{E(\bar{y}_{SS}^{(1)} - \bar{Y})^2 + E(\bar{y}_{SS}^{(2)} - \bar{Y})^2 + 2E(\bar{y}_{SS}^{(1)} - \bar{Y})(\bar{y}_{SS}^{(2)} - \bar{Y})\right\} \quad (3.7)$$

Since

$$E(\bar{y}_{SS}^{(i)} - \bar{Y})^2 = MSE(\bar{y}_{SS}^{(i)}) = \bar{Y}^2\left(\frac{1}{m} - \frac{1}{N}\right)\left\{C_y^2 + \nu C_x^2\left(A^2\nu - 2AK\right)\right\}; i = 1, 2 \quad (3.8)$$

Let $\bar{y}_i = \bar{Y} + e_0^{(i)}$ and $\bar{x}_i = \bar{X} + e_1^{(i)}$ such that $E\left(e_0^{(i)}\right) = E\left(e_1^{(i)}\right) = 0 \quad i = 1, 2$

Consider

$$E(\bar{y}_{SS}^{(1)} - \bar{Y})(\bar{y}_{SS}^{(2)} - \bar{Y}) = E\left(e_0^{(1)} - \frac{A\beta_{1x}\bar{Y}e_1^{(1)}}{\bar{X}\beta_{1x} + \sigma_x}\right)\left(e_0^{(2)} - \frac{A\beta_{1x}\bar{Y}e_1^{(2)}}{\bar{X}\beta_{1x} + \sigma_x}\right)$$

$$= E(e_0^{(1)}e_0^{(2)}) - \frac{A\beta_{1x}\bar{Y}}{\bar{X}\beta_{1x} + \sigma_x}\left\{E(e_0^{(2)}e_1^{(1)}) + E(e_0^{(1)}e_1^{(2)})\right\} + \left(\frac{A\beta_{1x}\bar{Y}}{\bar{X}\beta_{1x} + \sigma_x}\right)^2 E(e_1^{(1)}e_1^{(2)})$$

Substituting the results in Sukhatme and Sukhatme

$$E(e_0^{(1)}e_0^{(2)}) = -\frac{1}{N}\bar{Y}^2 C_y^2$$

$$E(e_1^{(1)}e_1^{(2)}) = -\frac{1}{N}\bar{X}^2 C_x^2$$

$E(e_0^{(1)}e_1^{(2)}) = E(e_0^{(2)}e_1^{(1)}) = -\frac{1}{N}\bar{X}\bar{Y}\rho C_x C_y$ we have

$$E(\bar{y}_{SS}^{(1)} - \bar{Y})(\bar{y}_{SS}^{(2)} - \bar{Y}) = -\frac{1}{N}\bar{Y}^2\left\{C_y^2 + A^2\nu^2 C_x^2 - 2A\nu\rho C_x C_y\right\}$$

$$= -\frac{1}{N}\bar{Y}^2\left\{C_y^2 + \nu C_x^2\left(A^2\nu - 2AK\right)\right\} \quad (3.9)$$

Putting the values from (3.8) and (3.9) in (3.7) we have

$$E(\hat{\bar{Y}}_{SS}' - \bar{Y})^2 = \bar{Y}^2\frac{1}{4}\left[2\left(\frac{1}{m} - \frac{1}{N}\right) - \frac{2}{N}\right]\left\{C_y^2 + \nu C_x^2\left(A\nu - 2AK\right)\right\}$$

$$= \bar{Y}^2\left(\frac{1}{2m} - \frac{1}{N}\right)\left\{C_y^2 + \nu C_x^2\left(A^2\nu - 2AK\right)\right\} \quad (3.10)$$

Now consider

$$E(\overline{y}_{SS}^{(3)}-\overline{Y})(\hat{Y}_{SS}^{'}-\overline{Y})=E\{(\overline{y}_{SS}^{(3)}-\overline{Y})(\frac{\overline{y}_{SS}^{(1)}+\overline{y}_{SS}^{(2)}}{2}-\overline{Y})\}$$

$$=\frac{1}{2}\{E(\overline{y}_{SS}^{(3)}-\overline{Y})(\overline{y}_{SS}^{(1)}-\overline{Y})+E(\overline{y}_{SS}^{(3)}-\overline{Y})(\overline{y}_{SS}^{(2)}-\overline{Y})\}$$

Since

$$E(\overline{y}_{SS}^{(3)}-\overline{Y})(\overline{y}_{SS}^{(i)}-\overline{Y})=E\left(e_0-\frac{A\beta_{1x}\overline{Y}e_1}{\overline{X}\beta_{1x}+\sigma_x}\right)\left(e_0^{(i)}-\frac{A\beta_{1x}\overline{Y}e_1^{(i)}}{\overline{X}\beta_{1x}+\sigma_x}\right)\text{where } i=1,2$$

$$=E(e_0e_0^{(i)})-\frac{A\beta_{1x}\overline{Y}}{\overline{X}\beta_{1x}+\sigma_x}\left\{E(e_0^{(i)}e_1)+E(e_0e_1^{(i)})\right\}+\left(\frac{A\beta_{1x}\overline{Y}}{\overline{X}\beta_{1x}+\sigma_x}\right)^2E(e_1e_1^{(i)})$$

using the following results given in Sukhatme and Sukhatme

$$E(e_0e_0^{(i)})=\left(\frac{1}{2m}-\frac{1}{N}\right)\overline{Y}^2C_y^{\,2}$$

$$E(e_1e_1^{(i)})=\left(\frac{1}{2m}-\frac{1}{N}\right)\overline{X}^2C_x^{\,2}$$

$$E(e_0e_1^{(i)})=E(e_0^{(i)}e_1)=\left(\frac{1}{2m}-\frac{1}{N}\right)\overline{X}\overline{Y}\rho C_xC_y \text{ for } i=1,2 \text{ we have}$$

$$E(\overline{y}_{SS}^{(3)}-\overline{Y})(\overline{y}_{SS}^{(i)}-\overline{Y})=\overline{Y}^2\left(\frac{1}{2m}-\frac{1}{N}\right)\left\{C_y^{\,2}+\nu C_x^{\,2}\left(A^2\nu-2AK\right)\right\} \tag{3.11}$$

Putting these values from (3.6), (3.10), (3.11) in (3.5) we have

$$MSE(\hat{Y}_{SS}^{*})=\frac{1}{(1-R)^2}\overline{Y}^2\left(\frac{1}{2m}-\frac{1}{N}\right)(1+R^2-2R)\left\{C_y^{\,2}+\nu C_x^{\,2}\left(A^2\nu-2AK\right)\right\}$$

$$=\overline{Y}^2(\frac{1}{n}-\frac{1}{N})\left\{C_y^{\,2}+\nu C_x^{\,2}\left(A^2\nu-2AK\right)\right\} \tag{3.12}$$

which is equal to the mean square error of $\overline{y}_{SS}$. Therefore, the optimizing value of the characterizing scalar A is given by (2.6) and the minimum mean square error under optimizing value of $A=A_{opt}$ is given by (2.7).

4. Unbiasedness and MSE of Midzuno-Lahiri-Sen Type Sampling Strategy

Let us consider $\overline{y}_{SA}$ under Midzuno (1952)-Lahiri (1951)-Sen (1952) type sampling scheme and denote it by $\overline{y}_{SM}$. The proposed Midzuno-Lahiri-Sen type sampling scheme for selecting a sample s of size n deals with selecting first unit with probability proportional to $\overline{X}\beta_{1x}+\sigma_x+A\beta_{1x}(x_i-\overline{X})$ where x_i is the size of the first selected unit such that

$$P(i)=\text{P(selecting first unit } i \text{ with size } x_i)=\frac{\overline{X}\beta_{1x}+\sigma_x+A\beta_{1x}(x_i-\overline{X})}{N(\overline{X}\beta_{1x}+\sigma_x)} \tag{4.1}$$

and selecting the remaining $n-1$ units in the sample from $N-1$ units in the population by simple random sampling without replacement. Thus the probability of selecting the sample s of size n is

$$P(s)=\frac{\overline{X}\beta_{1x}+\sigma_x+A\beta_{1x}(\overline{x}_s-\overline{X})}{(\overline{X}\beta_{1x}+\sigma_x).{}^NC_n} \tag{4.2}$$

where $\bar{x}_s$ and $\bar{y}_s$ are the sample mean of x and y respectively based on the sample s.
Consider

$$E(\bar{y}_{SM}) = E\{\frac{\bar{y}_s(\bar{X}\beta_{1x}+\sigma_x)}{\bar{X}\beta_{1x}+\sigma_x + A\beta_{1x}(\bar{x}_s-\bar{X})}\}$$

$$= \sum_{s=1}^{{}^N C_n} \frac{\bar{y}_s(\bar{X}\beta_{1x}+\sigma_x)}{\bar{X}\beta_{1x}+\sigma_x + A\beta_{1x}(\bar{x}_s-\bar{X})} P(s)$$

$$= \sum_{s=1}^{{}^N C_n} \frac{\bar{y}_s}{{}^N C_n}$$

$$= E(\bar{y}_s) \qquad \text{(under simple random sampling without replacement)}$$

$$= \bar{Y} \qquad (4.3)$$

showing that $\bar{y}_{SM}$ is an unbiased estimator of population mean $\bar{Y}$ for all values of A under the proposed Midzuno-Lahiri-Sen type sampling scheme.
Now, since

$$MSE(\bar{y}_{SM}) = V(\bar{y}_{SM}) = E(\bar{y}_{SM}^2) - \bar{Y}^2$$

where

$$E(\bar{y}_{SM}^2) = \sum_{s=1}^{{}^N C_n} \bar{y}_{SM}^2 . P(s)$$

$$= \sum_{s=1}^{{}^N C_n} \left\{ \frac{\bar{y}_s(\bar{X}\beta_{1x}+\sigma_x)}{\bar{X}\beta_{1x}+\sigma_x + A\beta_{1x}(\bar{x}_s-\bar{X})} \right\}^2 P(s)$$

$$= \sum_{s=1}^{{}^N C_n} \bar{y}_s^2 \frac{(\bar{X}\beta_{1x}+\sigma_x)}{\bar{X}\beta_{1x}+\sigma_x + A\beta_{1x}(\bar{x}_s-\bar{X})} \frac{1}{{}^N C_n}$$

$$= E\left[\left(\bar{Y}+e_0\right)^2 \left\{ 1 + \frac{A\beta_{1x}e_1}{\bar{X}\beta_{1x}+\sigma_x} \right\}^{-1} \right]$$

$$= E\left[\left(\bar{Y}^2 + e_0^2 + 2\bar{Y}e_0\right) \left\{ 1 - \frac{A\beta_{1x}e_1}{\bar{X}\beta_{1x}+\sigma_x} + \frac{A^2\beta_{1x}^2 e_1^2}{(\bar{X}\beta_{1x}+\sigma_x)^2} - \ldots \right\} \right]$$

$$= \bar{Y}^2 + E(e_0^2) + \frac{A^2\beta_{1x}^2\bar{Y}^2}{(\bar{X}\beta_{1x}+\sigma_x)^2} E(e_1^2) - 2\frac{A\beta_{1x}\bar{Y}}{\bar{X}\beta_{1x}+\sigma_x} E(e_0 e_1) \quad \text{(upto } 1^{st} \text{ order of}$$

approximation)
Therefore, by using (2.3), we have

$$MSE(\bar{y}_{SM}) = \bar{Y}^2(\frac{1}{n}-\frac{1}{N})\left\{C_y^2 + \nu C_x^2\left(A^2\nu - 2AK\right)\right\} = MSE(\bar{y}_{SS}) = MSE(\hat{\bar{Y}}_{SS}^{*})$$

$$= MSE(\bar{y}_{SA}) \text{ (say)} \qquad (4.4)$$

Thus, the optimizing value of the characterizing scalar A is given by (2.6) and the minimum mean square error under optimizing value of $A = A_{opt}$ is given by (2.7). Further, let the minimum mean square error under optimizing value of $A = A_{opt}$ be denoted by $MSE(\bar{y}_{SA})_{\min}$.

5. Concluding Remarks

If the minimizing value $A = K/\delta = A_{opt}$ is known then, we have

$$MSE(\bar{y}) - MSE(\bar{y}_{SA})_{\min} = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2 \rho^2 C_y^{\ 2} \geq 0 \tag{5.1}$$

$$MSE(\bar{y}_R) - MSE(\bar{y}_{SA})_{\min} = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2 \left(C_x - \rho C_y\right)^2 \geq 0 \tag{5.2}$$

$$MSE(\bar{y}_P) - MSE(\bar{y}_{SA})_{\min} = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2 \left(C_x + \rho C_y\right)^2 \geq 0 \tag{5.3}$$

$$MSE(\bar{y}_{SR}) - MSE(\bar{y}_{SA})_{\min} = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2 \left(\delta C_x - \rho C_y\right)^2 \geq 0 \tag{5.4}$$

$$MSE(\bar{y}_{SP}) - MSE(\bar{y}_{SA})_{\min} = (\frac{1}{n} - \frac{1}{N})\bar{Y}^2 \left(\delta C_x + \rho C_y\right)^2 \geq 0 \tag{5.5}$$

Hence, under the optimizing value of the characterizing scalar $A = K/\nu = A_{opt}$, the proposed sampling strategies are always better than $\bar{y}$, $\bar{y}_{SR}$, $\bar{y}_{SP}$, $\bar{y}_P$, $\bar{y}_R$ and $\bar{y}_{lr}$ in sense of unbiasedness and gain in efficiency.

6. Empirical Study

Let us consider the following example considered by Singh and Chaudhary (1986) wherein the following values were obtained $\bar{Y}$ = 1467.545, $\bar{X}$ = 22.62, C_x = 1.460904, C_y = 1.745871, β_{1x} = 3.307547, σ_x =32.28717, ρ = 0.902147. The bias, mean square errors and percent relative efficiency (PRE) w.r.t. $\bar{y}$ of the sample mean $\bar{y}$, ratio estimator $\bar{y}_R$, product estimator $\bar{y}_P$, $\bar{y}_{SR}$, $\bar{y}_{SP}$ and $\bar{y}_{lr}$ are given by

Est. (t)	$\bar{y}$	$\bar{y}_R$	$\bar{y}_P$	$\bar{y}_{lr}$	$\bar{y}_{SR}$	$\bar{y}_{SP}$	$\bar{y}_{SM}$ / $\hat{\bar{Y}}^{*}_{SS}$
Bias (t)	0	-244.68	3376.77	3867.58	-830.47	2358.86	0
MSE (t)	6564590	1249925	21072230	1221872	1884104	15731012	1221872
PRE(t : $\bar{y}$)	100	525.20	31.15	537.26	348.42	41.73	537.26

Note that the above results are scaled by the factor $(\frac{1}{n} - \frac{1}{N})$. It can be easily observed from the table that only sample mean $\bar{y}$ and the proposed sampling strategies are unbiased. Further, it can be easily observed that $\bar{y}_{lr}$ and the proposed sampling strategies attains the minimum mean square error but $\bar{y}_{lr}$ is biased. It is evident from the above empirical study that the proposed sampling strategies are better than the remaining sampling strategies both in terms of unbiasedness and mean square error.

References

1. Gray H. L. and Schucany W. R. (1972). The Generalized Jack-knife Statistic, Marcel Dekker, New York.
2. Lahiri D. B. (1951). A method of sample selection providing unbiased ratio estimates, Bull. Int. Stat. Inst., 3, p. 133-140.

3. Midzuno H. (1952). On the sampling system with probability proportional to the sum of the sizes, Ann. Inst. Stat. Math., 3, p. 99-107.
4. Murthy M. N. (1967). Sampling theory and methods, Statistical Publishing Society, Calcutta, India.
5. Quenouille M. H. (1956). Notes on bias in estimation, Biometrika, 43, p. 353-360.
6. Reddy V. N. (1973). On ratio and product methods of estimation, Sankhya, B, 35(3), p. 307 – 316.
7. Sen A. R. (1952). Present status of probability sampling and its use in the estimation of a characteristic, Econometrika, 20, p. 103.
8. Senapati S. C., Sahoo L. N. and Misra G. (2006). On a Scheme of Sampling of Two Units With Inclusion Probability Proportional to Size, Austrian Journal of Statistics, 35(4), p. 445–454.
9. Singh G. N. (2003). On the improvement of product method of estimator in sample surveys, J. I. S. A. S., 56(3), p. 267-275.
10. Singh H. P., Singh R., Espejo M. R., Pineda M. D. and Nadarajah S. (2005). On The Efficiency Of A Dual To Ratio-Cum-Product Estimator In Sample Surveys, Mathematical Proceedings of the Royal Irish Academy, 105A (2), p. 51–56.
11. Singh D. and Chaudhary F. S. (1986). Theory and analysis of sampling survey design, Wiley Eastern Limited, New Delhi.
12. Sukhatme P. V. and Sukhatme B. V. (1997). Sampling Theory of Surveys with Applications, Piyush Publications, Delhi.
13. Walsh J. E. (1970). Generalization of ratio estimator for population total, Sankhya, A, 32, p. 99-106.