Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/070105

Dual to Ratio Estimators for Mean Estimation in Successive Sampling using Auxiliary Information on Two Occasion

Nazeema Beevi

Department of Statistics, University of Calicut, Kerala - 673 635, India.

Received: 14 Jan. 2016, Revised: 30 Dec. 2017, Accepted: 22 Jan. 2018
Published online: 1 Mar. 2018

Abstract: In this paper, consider dual to ratio estimator for estimating mean using auxiliary information on both occasions in successive sampling scheme. Dual to ratio estimators have been developed by Srivenkataramana (1980) under simple random sampling. Using this estimator under successive sampling scheme, the bias and mean squared error are obtained upto the first order of approximation and show theoretically that the proposed estimator is more efficient than the Cochran's estimator using no auxiliary variable and simple mean per unit estimator. Optimum replacement strategy is also discussed. Results have been justified through empirical interpretation.

Keywords: Auxiliary Information, Dual to Ratio Estimator, Optimum Replacement, Successive Sampling.

1 Introduction

Now a days, it is often seen that sample surveys are not limited to one time inquiries. A survey carried out on a finite population is subject to change overtime if the value of study character of a finite population is subject to change (dynamic) overtime. A survey carried out on a single occasion will provide information about the characteristics of the surveyed population for the given occasion only and can not give any information on the nature or the rate of change of the characteristics over different occasions and the average value of the characteristics over all occasions or most recent occasion. A part of the sample is retained being replaced for the next occasion (or sampling on successive occasions, which is also called successive sampling or rotation sampling).

The successive method of sampling consists of selecting sample units on different occasions such that some units are common with samples selected on previous occasions. If sampling on successive occasions is done according to a specific rule, with replacement of sampling units, it is known as successive sampling. Replacement policy was examined by Jessen (1942) who examined the problem of sampling on two occasions, without or with replacement of part of the sample in which what fraction of the sample on the first occasion should be replaced in order that the estimator of $\bar{Y}$ may have maximum precision. Yates (1949) extended Jessen's scheme to the situation where the population mean of a character is estimated on each of $(h > 2)$ occasions from a rotation sample design. These results were generalized by Patterson (1950) and Narain (1953), among others. Rao and Mudhdkar (1983) and Das (1982), used the information collected on the previous occasions for improving the current estimate. Data regarding changing properties of the population of cities or counties and unemployment statistics are collected regularly on a sample basis to estimate the changes from one occasion to the next or to estimate the average over a certain period. An important aspect of continuous surveys is the structure of the sample on each occasion. To meet these requirements, successive sampling provide a strong tool for generating the reliable estimates at different occasions.

Sen (1971) developed estimators for the population mean on the current occasion using information on two auxiliary variables available on previous occasion. Sen (1972, 1973) extended his work for several auxiliary variables. Singh *et.al* (1991) and Singh and Singh (2001) used the auxiliary information on current occasion for estimating the current population mean in two occasions successive sampling and Singh (2003) extended his work for $h-$occasions successive sampling. Feng and Zou (1997) and Biradar and Singh (2001) used the auxiliary information on both the occasions for

* Corresponding author e-mail: nazeemathaj@gmail.com

estimating the current mean in successive sampling. In many situations, information on an auxiliary variate may be readily available on the first as well as the second occasions; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation and number of beds in hospital surveys.

Most of the studies related to dual to ratio estimators have been developed by Srivenkataramana (1980). He considered, the relationship between the response y and the subsidiary variate x, is linear through the origin and variance of y is proportional to x. It is assumed that X is known. Motivated with the above argument and utilizing the information on an addition auxiliary variable is available on the both occasions, the dual to ratio estimator for estimating the population mean on current occasion in successive sampling has been proposed. It has been assumed that the additional auxiliary variable over two occasions.

The paper is spread over ten sections. Sample structure and notations have been discussed in section 2 and section 3 respectively. In section 4, the proposed estimators have been formulated. Properties of proposed including mean square error are derived under section 5. In section 6, optimum replacement policy is discussed. Section 7 contains comparison of the proposed estimator with Cochran (1977) and simple mean per unit when there is no matching from the previous occasion and the estimator when no additional auxiliary information has been used. In Section 8 and 9, the theoretical results are supported by a numerical interpretation and give conclusion in Section 10.

2 Selection of the sample

Consider a finite population $U = (U_1, U_2 \ldots U_N)$ which has been sampled over two occasions. Let x and y be the study variables on the first and second occasions respectively, further assumed that the information on the auxiliary variable z, whose population mean is known which is closely related (positively related) to x and y on the first and second occasions respectively available on the first as well as on the second occasion. For convenience, it is assumed that the population under consideration is large enough. Allowing SRSWOR (Simple Random Sampling without Replacement) design in each occasions, the successive sampling scheme as follows is carried out:

– n units which constitutes the sample on the first occasion. A random sub sample of $n_m = n\lambda$ $(0 < \lambda < 1)$ units is retained (matched) for use on the second occasion.
– In the second occasion $n_u = n\mu$ $(= n - n_m)$ $(0 < \mu < 1)$ units are drawn from the remaining $(N - n)$ units of the population. Where μ is the fraction of fresh sample on the current occasion.

So that the sample size on the second occasion is also n $(= n\lambda + n\mu)$.

3 Description of Notations

The following notations in this paper.

$\bar{X}$: The population mean of the study variable on the first occasion.
$\bar{Y}$: The population mean of the study variable on the second occasion.
$\bar{Z}$: The population mean of the auxiliary variable on both occasions.
S_y^2: Population mean square of y.
$\bar{z}_n$: The sample mean based on n units drawn on the first occasion.
$\bar{z}_{n_u}$: The sample mean based on n_u units drawn on the second occasion.
$\bar{x}_n$: The sample mean based on n units drawn on the first occasion.
$\bar{x}_{n_m}$: The sample mean based on n_m units observed on the second occasion and common with the first occasion.
$\bar{y}_{n_u}$: The sample mean based on n_u units drawn afresh on the second occasion.
$\bar{y}_{n_m}$: The sample mean based on n_m units common to both occasions and observed on the first occasion.
ρ_{yx}: The correlation coefficient between the variables y on x.
ρ_{xz}: The correlation coefficient between the variables x on z.
ρ_{yz}: The correlation coefficient between the variables y on z.
n_m: The sample units observed on the second occasion and common with the first occasion.
n_u: The sample size of the sample drawn afresh on the second occasion.
n: Total sample size.

4 Proposed Product Ratio Estimators in Successive sampling

In this section some dual to ratio estimators using one auxiliary variable have been proposed. To estimate the population mean $\bar{Y}$ on the second occasion, two different estimators are suggested. The first estimator is dual to ratio estimator based on sample of size $n_u(=n\mu)$ drawn afresh on the second occasion and is given by:

$$t_{n_u} = \bar{y}_{n_u}\frac{\bar{z}_{n_u}{}^{*}}{\bar{Z}}, \qquad (4.1)$$

where $\bar{z}_{n_u}{}^{*} = (1+g)\bar{Z} - g\bar{z}$ and $g = \frac{n}{N-n}$. The second estimator is a chain dual to ratio estimator based on the sample of size $n_m(=n\lambda)$ common with both the occasions and is defined as,

$$t_{n_m} = \bar{y}_{n_m}\frac{\bar{x}^{*}_{n_m}}{\bar{x}_n}\frac{\bar{z}_n{}^{*}}{\bar{Z}}, \qquad (4.2)$$

where $\bar{x}_{n_m}{}^{*} = (1+g)\bar{X} - g\bar{x}$, $\bar{x}_n{}^{*} = (1+g)\bar{X} - g\bar{x}$ and $g = \frac{n}{N-n}$. Combining the estimators t_{n_u} and t_{n_m}, the final estimator t_{dr} as follows

$$t_{dr} = \psi t_{n_u} + (1-\psi)t_{n_m}, \qquad (4.3)$$

where ψ is an unknown constant to be determined such that $MSE(t_{dr})$ is minimum and prove theoretically that the estimator is more efficient than the proposed estimator by (i) Cochran (1977) when no auxiliary variables are used at any occasion.This classical difference estimator is a widely used estimator to estimate the population mean $\bar{Y}$, in successive sampling. It is given by

$$\bar{y_2}' = \phi_2\bar{y}'_{2u} + (1-\phi_2)\bar{y}'_{2m},$$

where ϕ_2 is an unknown constant to be determined such that $V(\hat{\bar{Y}})_{opt}$ is minimum and $\bar{y_{2u}}' = \bar{y_{2u}}$ is the sample mean of the unmatched portion on the second occasion and $\bar{y}'_{2m} = \bar{y}_{2m} + b(\bar{y_1} - \bar{y}_{1m})$ is based on matched portion. The variance of this estimator is

$$V(\hat{\bar{Y}})_{opt} = [1+\sqrt{(1-\rho^2_{yx})}]\frac{S_y^2}{2n}.$$

Similarly, the variance of the mean per unit estimator is given by

$$V(\bar{y}) = \frac{S_y^2}{n}.$$

4.1 Properties of $\mathbf{t_{dr}}$

Since t_{n_u} and t_{n_m} both are biased estimators of t_{dr}, therefore, resulting estimator t_{dr} is also a biased estimator. The bias and MSE up to the first order of approximation are derived as using large sample approximation given below:

$\bar{y_{n_u}} = \bar{Y}(1+e_{\bar{y}_{n_u}})$, $\bar{y}_{n_m} = \bar{Y}(1+e_{\bar{y}_{n_m}})$,
$\bar{x}_{n_m} = \bar{X}(1+e_{\bar{x}_{n_m}})$, $\bar{x}_n = \bar{X}(1+e_{\bar{x}_n})$,
$\bar{z} = \bar{Z}(1+e_{\bar{z}_n})$, $\bar{z}_{n_u} = \bar{Z}(1+e_{\bar{z}_{n_u}})$

where $e_{\bar{y}_{n_u}}, e_{\bar{y}_{n_m}}, e_{\bar{x}_{n_m}}, e_{\bar{x}_n}, e_{\bar{z}_n}, \bar{z}_{n_u}$ are sampling errors and are of very small quantities. We assume that $E(e_{\bar{y}_{n_u}}) = E(e_{\bar{y}_{n_m}}) = E(e_{\bar{x}_{n_m}}) = E(e_{\bar{x}_n}) = E(e_{\bar{z}_n}) = E(e_{\bar{z}_{n_u}}) = 0.$ Then for simple random sampling without replacement for both first and second occasions, we write by using occasion wise operation of expectation as:

$E(e^2_{\bar{y}_{n_u}}) = \left(\frac{1}{n_u} - \frac{1}{N}\right)S_y^2, E(e^2_{\bar{y}_{n_m}}) = \left(\frac{1}{n_m} - \frac{1}{N}\right)S_y^2,$
$E(e^2_{\bar{x}_{n_m}}) = \left(\frac{1}{n_m} - \frac{1}{n}\right)S_x^2, E(e^2_{\bar{x}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_x^2,$
$E(e^2_{\bar{z}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_z^2,$

$$E(e_{\bar{y}_{n_u}}e_{\bar{z}_{n_u}}) = \left(\frac{1}{n_u} - \frac{1}{N}\right)S_{yz}, \qquad E(e_{\bar{y}_{n_m}}e_{\bar{x}_n}) = \left(\frac{1}{n_m} - \frac{1}{n}\right)S_{yx},$$
$$E(e_{\bar{y}_{n_m}}e_{\bar{x}_{n_m}}) = \left(\frac{1}{n_m} - \frac{1}{n}\right)S_{yx}, \qquad E(e_{\bar{y}_{n_m}}e_{\bar{x}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{yx},$$
$$E(e_{\bar{y}_{n_m}}e_{\bar{z}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{yz}, \qquad E(e_{\bar{x}_{n_m}}e_{\bar{x}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_x^2,$$
$$E(e_{\bar{x}_{n_m}}e_{\bar{z}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{xz}, \qquad E(e_{\bar{x}_n}e_{\bar{z}_n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{xz}.$$

Derive the bias of t_{n_u} in lemma 4.1.

Lemma 4.1*The bias of* t_{n_u} *denoted by* $B(t_{n_u})$ *is given by*

$$B(t_{n_u}) = -\bar{Y}\left(\frac{1}{n_u} - \frac{1}{N}\right)g\rho_{yz}S_yS_z$$

*Proof.*Expressing (4.1) in terms of e's and get

$$t_{n_u} = \overline{Y}(1+e_{\bar{y}_{n_u}})[(1+g)-g(1+e_{\bar{z}_{n_u}})]$$
$$= (1+e_{\bar{y}_{n_u}})(1-ge_{\bar{z}_{n_u}}).$$

Taking expectation on both sides and ignoring higher orders,

$$E(t_{n_u}-\overline{Y}) = -\overline{Y}g\left(\frac{1}{n_u}-\frac{1}{N}\right)E(e_{\bar{y}_{n_u}}e_{\bar{z}_{n_u}})$$

$$B(t_{n_u}) = -\overline{Y}\left(\frac{1}{u}-\frac{1}{N}\right)g\rho_{yz}S_yS_z. \tag{4.4}$$

The bias of t_{n_m} is derived in lemma 4.2.

Lemma 4.2*The bias of t_{n_m} is denoted by $B(t_{n_m})$given by*
$B(t_{n_m}) = \overline{Y}\left(\frac{1}{n_m}-\frac{1}{n}\right)(S_x^2-\rho_{yx}S_yS_x)+\left(\frac{1}{n}-\frac{1}{N}\right)[-g\rho_{yz}S_yS_z]$

*Proof.*Expressing (4.2) in terms of e's, get

$$t_{n_m} = \overline{Y}(1+e_{\bar{y}_{n_m}})(1+e_{\bar{x}_n})^{-1}(1-ge_{\bar{x}_{n_m}})(1-ge_{\bar{z}_n}).$$

Expanding the right hand side and neglecting the terms with power two or greater and get

$$t_{n_m} = \overline{Y}[(1+e_{\bar{y}_{n_m}}-e_{\bar{x}_n}+e_{\bar{x}_n}^2-e_{\bar{y}_{n_m}}e_{\bar{x}_n})$$

$$(1-ge_{\bar{y}_{n_m}}-ge_{\bar{z}_n}+g^2e_{\bar{y}_{n_m}}e_{\bar{z}_n})]. \tag{4.5}$$

Taking expectation (4.5) on both sides,

$$B(t_{n_m}) = \overline{Y}\left(\frac{1}{n_m}-\frac{1}{n}\right)(S_x^2-\rho_{yx}S_yS_x)+\left(\frac{1}{n}-\frac{1}{N}\right)[-g\rho_{yz}S_yS_z].$$

Using 4.1 and 4.2, derive the bias of t_{dr}.

Theorem 4.1*Bias of the estimator t_{dr} to the first order approximation is,*

$$B(t_{dr}) = \psi B(t_{n_u})+(1-\psi)B(t_{n_m}), \tag{4.6}$$

where

$$B(t_{n_u}) = -\overline{Y}\left(\frac{1}{u}-\frac{1}{N}\right)g\rho_{yz}S_yS_z.$$

and

$$B(t_{n_m}) = \overline{Y}\left(\frac{1}{n_m}-\frac{1}{n}\right)(S_x^2-\rho_{yx}S_yS_x)+\left(\frac{1}{n}-\frac{1}{N}\right)(-g\rho_{yz}S_yS_z).$$

*Proof.*The bias of the estimator t_{dr} is given by

$$B(t_{dr}) = E(t_{dr}-\overline{Y})$$

$$B(t_{dr}) = \psi E(t_{n_u}-\overline{Y})+(1-\psi)E(t_{n_m}-\overline{Y}), \tag{4.7}$$

Using lemma (4.1) and (4.2) into equation (4.7), the expression for the bias of the estimator t_{dr} as shown in (4.6)

We derive the MSE of t_{n_u} in lemma 4.3.

Lemma 4.3*The mean square error of t_{n_u} denoted by $M(t_{n_u})$ is given by*
$M(t_{n_u}) = \overline{Y}^2\left(\frac{1}{n_u}-\frac{1}{N}\right)\left[S_y^2+g^2S_z^2-2g\rho_{yz}S_yS_z\right]$

*Proof.*Expressing t_{n_u} in terms of e's, get

$$t_{n_u} = \bar{Y}(1+e_{\bar{y}_{n_u}})(1-ge_{\bar{z}_{n_u}}). \tag{4.8}$$

Expanding and squaring(4.8), the right hand side and neglecting the terms with power two or greater,

$$t_{n_u} = \bar{Y}(1+e_{\bar{y}_{n_u}} - ge_{\bar{z}_{n_u}})^2$$

$$t_{n_u} = \bar{Y}(1+e^2_{\bar{y}_{n_u}} + g^2 e^2_{\bar{z}_{n_u}} - 2ge_{\bar{y}_{n_u}} e_{\bar{z}_{n_u}}). \tag{4.9}$$

Taking expectation (4.9) on both sides and get $M(t_{n_u})$

$$M(t_{n_u}) = \bar{Y}^2\left(\frac{1}{n_u} - \frac{1}{N}\right)\left[S_y^2 + g^2 S_z^2 - 2g\rho_{yz} S_y S_z\right]. \tag{4.10}$$

The MSE of t_{n_m} is derived in lemma 4.4

Lemma 4.4*The mean square error of t_{n_m} denoted by $M(t_{n_m})$ is given by*

$$M(t_{n_m}) = \bar{Y}^2[\left(\frac{1}{n_m} - \frac{1}{N}\right)S_y^2 + \left(\frac{1}{n_m} - \frac{1}{n}\right)(g^2 {S_x}^2 - 2g\rho_{yx} S_y S_x) +$$

$$\left(\frac{1}{n} - \frac{1}{N}\right)(g^2 S_z^2 - 2g\rho_{yz} S_y S_z)].$$

*Proof.*Expressing t_{n_m} in terms of e's,

$$t_{n_m} = \bar{Y}(1+e_{\bar{y}_{n_m}})\left\{\frac{[(1+g)\bar{X} - g(1+e_{\bar{y}_{n_m}})\bar{X}]}{(1+e_{\bar{x}_n})\bar{X}}\frac{[(1+g)\bar{Z} - g(1+e_{\bar{z}_n})\bar{Z}]}{Z}\right\},$$

$$= \bar{Y}(1+e_{\bar{y}_{n_m}})(1-ge_{\bar{x}_{n_m}})(1-e_{\bar{x}_n})(1-e_{\bar{z}_n}). \tag{4.11}$$

Expanding (4.10), the right hand side and neglecting the terms with power two or greater, get

$$t_{n_m} = \bar{Y}(1+e_{\bar{y}_{n_m}} - ge_{\bar{x}_n} - ge_{\bar{x}_{n_m}} - ge_{\bar{z}_n}). \tag{4.12}$$

Squaring on both sides (4.11) and taking expectation, MSE of the estimator t_{n_m} upto first order of approximation as,

$$M(t_{n_m}) = \bar{Y}^2[\left(\frac{1}{n_m} - \frac{1}{N}\right)S_y^2 + \left(\frac{1}{n_m} - \frac{1}{n}\right)(g^2 {S_x}^2 - 2g\rho_{yx} S_y S_x) +$$

$$\left(\frac{1}{n} - \frac{1}{N}\right)(g^2 S_z^2 - 2g\rho_{yz} S_y S_z)] \tag{4.13}$$

.

Using lemma 4.3 and 4.4, we derive the MSE of t_{dr}

Theorem 4.2*The mean square error of the estimator t_{dr} to the first order approximation is,*

$$M(t_{dr}) = \psi^2 M(t_{n_u}) + (1-\psi)^2 M(t_{n_m}) + 2\psi(1-\psi)Cov(t_{n_u}, t_{n_m}) \tag{4.14}$$

where

$$M(t_{n_u}) = \bar{Y}^2\left(\frac{1}{n_u} - \frac{1}{N}\right)\left[S_y^2 + g^2 S_z^2 + 2g\rho_{yz} S_y S_z\right]$$

$$M(t_{n_m}) = \bar{Y}^2[\left(\frac{1}{n_m} - \frac{1}{N}\right)S_y^2 + \left(\frac{1}{n_m} - \frac{1}{n}\right)(g^2 {S_x}^2 - 2g\rho_{yx} S_y S_x) +$$

$$\left(\frac{1}{n} - \frac{1}{N}\right)(g^2 S_z^2 - 2g\rho_{yz} S_y S_z)].$$

and

$$Cov(t_{n_u}, t_{n_m}) = 0.$$

*Proof.*The mean square error of the estimator $\mathbf{t_{dr}}$ is given by

$M(t_{dr}) = E(t_{dr} - \bar{Y})^2$

$$M(t_{dr}) = E[\psi(t_{n_u} - \bar{Y}) + (1-\psi)(t_{n_m} - \bar{Y})]^2, \tag{4.15}$$

$M(t_{dr}) = \psi^2 M(t_{n_u}) + (1-\psi)^2 M(t_{n_m}) + 2\psi(1-\psi)Cov(t_{n_u}, t_{n_m}),$

using lemma (4.3) and (4.4) into the equation (4.15), the expression for the *MSE* of the estimator t_{dr} as shown in (4.14)

Remark 4.3*The estimators, t_{n_u} and t_{n_m} are based on two independent samples of sizes n_u and n_m respectively, therefore the covariance term has been vanished.*

5 Minimum Mean Square Error of $\mathbf{t_{dr}}$

To obtain the optimum value of ψ, partially differentiate the expression (4.14) with respect to ψ, and put it equal to zero, we get

$$\psi_{opt} = \frac{M(t_{n_m})}{M(t_{n_u}) + M(t_{n_m})} \tag{5.1}$$

substituting the values of $M(t_{n_u})$ and $M(t_{n_m})$ from (4.10) and (4.13) in (5.1), get

$$\psi_{opt} = \frac{(k_1 + \mu k_2)}{(k_1 + \mu^2 k_2)}$$

$$= \frac{\mu[(k_1 + \mu k_2)]}{(k_1 + \mu^2 k_2)}$$

. Substitution of ψ_{opt} from (5.1) into (4.14) gives optimum value of *MSE* of t_{dr} as:

$$M(t_{dr})_{opt} = \frac{M(t_{n_m})M(t_{n_u})}{M(t_{n_u}) + M(t_{n_m})}. \tag{5.2}$$

Substituting the values of $M(t_{n_m})$ and $M(t_{n_u})$ from (4.9) and (4.10) in (5.2), get

$$M(t_{dr})_{opt} = \frac{1}{n}\left[\frac{k_1^2 + \mu k_1 k_2}{k_1 + \mu^2 k_2}\right], \tag{5.3}$$

, where $k_1 = 1 + g^2 - 2g\rho_{yz}$, $k_2 = 2g(\rho_{yx} - \rho_{yz})$, here $\mu (= \frac{u}{n})$ is the fraction of fresh sample drawn on the second occasion. Again $M(t_{dr})_{opt}$ derived in equation (5.3) is the function of μ. To estimate the population mean on each occasion the better choice of μ is 1 (no matching). However, to estimate the change in mean from one occasion to the other, μ should be 0 (complete matching).

6 Replacement Policy

In order to estimate t_{dr} with maximum precision an optimum value of μ should be determined so as to know what fraction of the sample on the first occasion should be replaced and minimize, $M(t_{dr})_{opt}$ in (5.3) with respect to μ, the optimum value of μ is obtained as,

$$\hat{\mu} = \frac{-k_1 \pm \sqrt{k_1^2 + k_1 k_2}}{k_2}, \tag{6.1}$$

where $k_1 = 1 + g^2 - 2g\rho_{yz}$, $k_2 = 2g(\rho_{yx} - \rho_{yz})$. From (6.1) it is obvious that for $\rho_{yz} \neq \rho_{yx}$ two values of $\hat{\mu}$ are possible, therefore to choose a value of $\hat{\mu}$, it should be remembered that $0 \leq \hat{\mu} \leq 1$. All other values of $\hat{\mu}$ are inadmissible. If both the real values of $\hat{\mu}$ are admissible, the lowest one will be the best choice as it reduces the total cost of the survey. Substituting the value of $\hat{\mu}$ from (6.1) in (5.3),

$$M(t_{dr})_{opt} = \frac{1}{n}\left[\frac{k_1^2 + \hat{\mu} k_1 k_2}{k_1 + \hat{\mu^2} k_2}\right]. \tag{6.2}$$

7 Efficiency Comparisons

In this section, to compare t_{dr} with respect to $\bar{y}$, (i) sample mean of y, when a sample units are selected at second occasion without any matched portion. (ii) difference estimator (Cochran 1977) when no auxiliary information is used at any occasion, have been obtained for known values of ρ_{yx} and ρ_{yz}. Since $\bar{y}$ and $\hat{\bar{Y}}$ are unbiased estimators of $\bar{Y}$, their variances for large N are respectively given by

$$V(\bar{y}) = \frac{S_y^2}{n}, \tag{7.1}$$

$$V(\hat{\bar{Y}})_{opt} = [1+\sqrt{(1-\rho_{yx}{}^2)}]\frac{S_Y{}^2}{2n}. \tag{7.2}$$

For different values ρ_{yx} and ρ_{yz}, the below shows the optimum value of μ. That is $\hat{\mu}$. The percent relative efficiencies, R_1 and R_2 of t_{opt} with respect to $\bar{y}$ and $\hat{\bar{Y}}$ respectively, where

$$R_1 = \frac{V(\bar{y})}{M(t_{dr})_{opt}} \times 100$$

and

$$R_2 = \frac{V(\hat{\bar{Y}})_{opt}}{M(t_{dr})_{opt}} \times 100.$$

The estimator t_{pr} (at optimal conditions) is also compared with respect to the estimators $V(\bar{y})$ and $V(\hat{\bar{Y}})$, respectively. Where

$$M(t_{pr})_{opt} = \frac{1}{n}\left[\frac{k_1^2+\hat{\mu}k_1k_2}{k_1+\hat{\mu}^2k_2}\right] \tag{7.3}$$

and

$$\hat{\mu} = \frac{-k_1 \pm \sqrt{k_1^2+k_1k_2}}{k_2}, \tag{7.4}$$

where $k_1 = 1+g^2-2g\rho_{yz}$, $k_2 = 2g(\rho_{yx}-\rho_{yz})$.

7.1 Empirical Study

The expressions of the optimum value of μ (i.e. $\hat{\mu}$) and the percent relative efficiencies R_1 and R_2 are in terms of population correlation coefficients ρ_{yx} and ρ_{yz}. The Table 7.1. shows that the values of $\hat{\mu}$, R_1 and R_2 for different choices of ρ_{yx}, ρ_{yz} and μ.

Table 1: Table 7.1. optimum values of μ and percent relative efficiencies of t_{dr} with respect to $\bar{y}$ and $\hat{\bar{Y}}$

ρ_{yz}		ρ_{yx}			
		0.2	0.4	0.6	0.8
0.3	$\hat{\mu}$	0.4973	0.5026	0.5081	0.5139
	R_1	104.66	105.82	106.98	108.19
	R_2	91.24	96.23	101.41	86.55
0.5	$\hat{\mu}$	0.4920	0.4973	0.5028	0.5085
	R_1	108.14	109.17	110.50	109.89
	R_2	107.03	104.62	108.44	87.79
0.7	$\hat{\mu}$	0.4864	0.4916	0.4972	0.5029
	R_1	111.18	113.03	114.29	115.62
	R_2	110.67	108.31	102.86	92.49
0.9	$\hat{\mu}$	0.4805	0.4858	0.4913	0.4970
	R_1	115.79	117.06	118.38	119.76
	R_2	114.61	72.92	106.53	95.80

8 Numerical Illustration

The results obtained in previous selections are now examined with the help of one natural population set of data.
Population Source: [Free access to the data by the Statistical Abstracts of the United States.] Let Y (study variable) be the level of corn production (in per acre) and Z (auxiliary variate) be the dosage of fertilizer using in corn filed in 50 counties in the United states in 2007 and X be the corn production in the year 2006 in the States of United states.
Based on the above description, the values of the different required parameters for population is,
$N = 50, \bar{X} = 139, \bar{Y} = 118, S_Y{}^2 = 3314.2, \rho_{yx} = 0.9827, \rho_{yz} = 0.2141, \hat{\mu} = 0.4796, \psi_{opt} = 0.4996.$

Table 2: Table 8.1. Percent Relative Efficiencies of t_{dr} with Respect to $\bar{y}$ and $\hat{\bar{Y}}$

f	g	Relative Efficiencies
0.5	1	R_1 = 4018.18
		R_2 =2381.82
0.7	2.33	R_1 = 4063.95
		R_2 = 2408.15
0.9	9	R_1 = 4068.33
		R_2 = 2410.87

where

$$R_1 = \frac{V(\bar{y})}{M(t_{dr})_{opt}}$$

and

$$R_2 = \frac{V(\hat{\bar{Y}})_{opt}}{M(t_{dr})_{opt}}.$$

ρ_{yz}	ρ_{yx}				
		-0.2	-0.4	- 0.6	-0.8
-0.3	$\hat{\mu}$	0.4973	0.5026	0.5081	0.5139
	R_1	104.66	105.82	106.98	108.19
	R_2	91.24	96.23	101.41	86.55
-0.5	$\hat{\mu}$	0.4920	0.4973	0.5028	0.5085
	R_1	108.14	109.17	110.50	109.89
	R_2	107.03	104.62	108.44	87.79
-0.7	$\hat{\mu}$	0.4864	0.4916	0.4972	0.5029
	R_1	126.18	135.03	142.29	159.62
	R_2	135.67	142.31	159.86	192.49
-0.9	$\hat{\mu}$	0.4805	0.4858	0.4913	0.4970
	R_1	115.79	117.06	118.38	119.76
	R_2	114.61	72.92	106.53	95.80

Table 3: Table 8.2. MSE and Bias of Different Estimators.

MSE	Bias	$\frac{\lvert Bias \rvert}{\bar{Y}}$
$M(t_{dr})_{opt} = 3.2585$ $V(\bar{y}) = 132.5680$ $V(\hat{\bar{Y}})_{opt} = 78.5600$	-734816.16	6227.255
$M(t_{dr})_{opt} =2.3275$ $V(\bar{y}) = 94.6914$ $V(\hat{\bar{Y}})_{opt} =56.1143$	-484314.82	4104.36
$M(t_{dr})_{opt} =$ **1.8103** $V(\bar{y}) = 73.6500$ $V(\hat{\bar{Y}})_{opt} = 43.6440$	-163808.63	1388.20

9 Interpretations of Empirical Results of t_{d_r}

From Table 7.1., the relative efficiency is observed that the suggested estimator is compared with mean per unit estimator and Cochran (1977) estimator. So, the use of auxiliary information at both occasions is justified.

1. For fixed values of ρ_{yx}, the value of R_1 and R_2 are increasing with increasing values of μ and the increasing ρ_{yz}.
2. The values of R_1, R_2 and μ are increasing with increasing values of ρ_{yz}. This is an agreement with the results Sukhatme *et.al* (1984), which justifies that higher the value of ρ_{yx}, higher the fraction of fresh sample required at the second (current) occasion.
3. For fixed values of ρ_{yx} and ρ_{yz}, there is appreciable gain in the performance of the proposed estimator t_{d_r} over $\bar{y}$ and $\hat{\bar{Y}}$ with the increasing value of μ.

10 Conclusion

From Table 7.1. clearly seen that the value of $\hat{\mu}$ (at optimum condition) also exist for both the considered populations. Hence, it justifies that the suggested family of estimators t_{d_r} is feasible under optimal conditions.
Tables 8.1. and 8.2. indicates that the suggested estimators t_{d_r} at optimum conditions is preferable over sample mean per unit estimator and also performs better than the Cochran's estimator.

Hence, it may be concluded that the estimation of mean at current using auxiliary information on both occasions in successive sampling is highly in terms of precision and reducing the cost of survey.clearly indicates that the proposed estimators is more efficient than simple arithmetic mean estimator and Cochran (1977) estimator. The following conclusion can be formed from Tables 7.1. For fixed ρ_{yx}, ρ_{yxz} and μ, the values of R_1 and $R_2$are increasing. This phenomenon indicates that smaller fresh sample at current occasion is required, if a highly positively correlated auxiliary characters is available. That is the performance of precision of the estimates also reduces the cost of the survey.

11 Perspective

Table 7.1. clearly indicates that the suggested estimators is more efficient than simple arithmetic mean and Cochran (1977) estimators. The following conclusion can be made from Table 7.1. Fixed ρ_{yx}, the values of R_1 and R_2 are increasing while $\hat{\mu}$ is decreasing with the increasing values of ρ_{yz}. This phenomenon indicates that smaller fresh sample at current occasion is required, if a highly positively correlated auxiliary characters is available. For Fixed ρ_{yz}, the values of R_1 and $\hat{\mu}$ are increasing while R_2 is decreasing for initial values of the increasing values of ρ_{yx}. Thus behavior is in agreement with Cochran (1977) results, which explains that more the value of ρ_{yx}, more fraction of fresh sample is required at current occasion. That is the performance of precision of the estimates as well as reduces the cost of the survey.

Under the given framework (Tables 8.1.and 8.2.) tables it is possible to reduce the bias and mean square error of the estimator, the analytical and empirical results support the theoretical justification of the work. The estimation of population mean on successive occasions should be encouraged as there are numerous practical situations that require the estimate of mean at different points of time as the characters are time dependent. Hence, the proposed estimators should be recommended for their use in practice.

References

[1] Biradar, R.S. and Singh, H.P., Successive sampling using auxiliary information on both occasions. Cal. Stat. Assoc. Bull., 51, 243-251, (2001).
[2] Cochran, W.G.: Sampling Techniques. 3^{rd} edition, (1977).
[3] Chaturvedi, D.K. and Tripathi, T.P. (1983). Estimation of Population ratio on two occasions using multivariate auxiliary information. Jour. Ind. Stat. Asso., 21, 113-120.
[4] Das, K.: Estimation of population ratio on two occasions. Jour. Ind. Soc. Agri. Stat. 34(2), 1-9, (1982).
[5] Feng, G.S, and Zou, G.: Sample rotation method with auxiliary variable. Communication in Statistics-Theory and Methods, 26, 6, 1497-1509, (1997).
[6] Jessen, R.J : Statistical Investigation of a sample survey for obtaining farm facts. Iowa Agricultural Experiment Statistical Research Bulletin, 304, (1942).
[7] Narain, R.D.: On the recurrence formula in sampling on successive occasions. Jour. Ind. Soc. Agri. Stat., 5, 96-99, (1953).
[8] Okafor, F.C. (1992) The theory and application of sampling over two occasions for the estimation of current population ratio , Statistica1, 137-147.
[9] Patterson, H.D.: Sampling on successive occasions with partial replacement of units. Jour. Roy. Stat. Soc., B_{12}, 241-255, (1950).
[10] Rao, J.N.K. and Mudholkar, G.S.: Generalized multivariate estimators for the mean of finite population parameters . Jour. Amer. Stat. Asso., 62, 1008-1012, (1967).
[11] Sen, A.R.: Successive sampling with two auxiliary variables. Sankhya, Series B, 33, 371-378, (1971).
[12] Sen, A.R.: Theory and application of sampling on repeated occasions. Jour. Amer. Stat. Asso., 59, 492- 509, (1973).
[13] Sen, A.R., Sellers, S. and Smith, G.E.J.: The use of a ratio estimate in successive sampling.Biometrics, 31, 673-683, (1975).
[14] Srivenkataramana, T.: A dual to ratio estimator in sample surveys, Biometrika, 67(1), 199?204, (1980).
[15] Singh, G.N.: Estimation of population mean using auxiliary information on recent occasion in h occasions successive sampling. Statistics in Transition, 6(4), 523-532, (2003).
[16] Singh, G.N. and Singh, V.K.: On the use of auxiliary information in successive sampling. Jour. Ind. Soc. Agri. Stat. 54 (1), 1-12, (2001).
[17] Yates, F.: Sampling Methods for Censuses and Surveys. Charles Grifﬂn and Co.,London, (1949).

Nazeema Beevi is a statistical investigator at Directorate of Economics and Statistics and done the Ph.D. in the Department of Statistics, University of Calicut. My research interests are Sampling Theory (Two-Phase Sampling, Super Population Model, Predictive Model), Statistical Inference and Data Analysis in particular, Small Area Estimation and Item Response Theory. I am a lifelong member of Kerala Statistical Association (KSA) and also Reviewer of Journal of Applied Probability and Statistics, UK (2016).