

# Dual to Ratio-Cum-Product Estimator in Simple and Stratified Random Sampling

Yunusa Olufadi  
Department of Statistics and Mathematical Sciences  
Kwara State University, P.M.B 1530, Malete, Nigeria  
olufadi14yunus@yahoo.com

## Abstract

New estimators for estimating the finite population mean using two auxiliary variables under simple and stratified sampling design is proposed. Their properties (e.g., mean square error) are studied to the first order of approximation. More so, some estimators are shown to be a particular member of this estimator. Furthermore, comparison of the proposed estimator with the usual unbiased estimator and other estimators considered in this paper reveals interesting results. These results are further supported with an empirical study using four natural data from literature.

**Keywords:** Auxiliary variable; Mean square error; Ratio-cum-product estimator; Simple random sampling; Stratified sampling.

## 1. Introduction

In real life, the problem of the estimation of population parameters like mean, proportion, variance, and ratio of two population means are common in virtually all discipline and facet of life. And sometimes, information on several variables is used to estimate or predict a characteristic of interest. For instance, an agriculturist might be interested in the total yield of maize taking into consideration the fertilizer levels, soil type, number of workers in a specific plot, regions etc. The use of this type of variables (known as auxiliary information in sample survey design) results in efficient estimate of population parameters (e.g. mean) under some realistic conditions. Ratio, product and regression methods of estimation are good examples in this context. Ratio and product type estimators take advantage of the correlation between the auxiliary variable,  $x$  and the study variable,  $y$  to improve the estimate of the characteristic of interest. For example, when information is available on the auxiliary variable that is positively (high) correlated with the study variable, the ratio method of estimation proposed by Cochran (1940) is a suitable estimator to estimate the population mean and when the correlation is negative the product method of estimation as envisaged by Robson (1957) and Murthy (1964) is appropriate.

Quite often information on many auxiliary variables is available in the survey which can be utilized to increase the precision of the estimate. In this situation, Olkin (1958) was the first author to deal with the problem of estimating the mean of a survey variable when auxiliary variables are made available. He suggested the use of information on more than one supplementary characteristic, positively correlated with the study variable, considering a linear combination of ratio estimators based on each auxiliary variable separately. The coefficients of the linear combination were determined so as to minimize

the variance of the estimator. Analogously to Olkin, Singh (1967) gave a multivariate expression of Murthy's (1964) product estimator, while Raj (1965) suggested a method for using multi-auxiliary variables through a linear combination of single difference estimators. More recently, Abu-Dayyeh *et al.* (2003), Kadilar and Cingi (2004, 2005), Perri (2005), Dianna and Perri (2007) among others have suggested estimators for  $\bar{Y}$  using information on several auxiliary variables.

Motivated by Srivenkataramana (1980), Bandyopadhyay (1980) and Singh *et al.* (2005) and with the aim of providing a more efficient estimator; we propose, in this paper, a new estimator for  $\bar{Y}$  when two auxiliary variables are available under simple and stratified sampling design.

## 2. Background to the Suggested Estimator

Consider a finite population  $P = (P_1, P_2, \dots, P_N)$  of  $N$  units. Let a sample  $s$  of size  $n$  be drawn from this population by simple random sampling without replacements (SRSWOR). Let  $y_i$  and  $(x_i, z_i)$  represents the value of a response variable  $y$  and two auxiliary variables  $(x, z)$  are available. The units of this finite population are identifiable in the sense that they are uniquely labeled from 1 to  $N$  and the label on each unit is known. Further, suppose in a survey problem, we are interested in estimating the population mean  $\bar{Y}$  of  $y$ , assuming that the population means  $(\bar{X}, \bar{Z})$  of  $(x, z)$  are known. The traditional ratio and product estimators for  $\bar{Y}$  are given as

$$\bar{y}_R = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right) \text{ and}$$

$$\bar{y}_P = \bar{y} \left( \frac{\bar{z}}{\bar{Z}} \right) \text{ respectively,}$$

where,  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ ,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  and  $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$  are the sample means of  $y$ ,  $x$  and  $z$  respectively.

Singh (1969) improved the ratio and product method of estimation given above and suggested the "ratio-cum-product" estimator for  $\bar{Y}$  as

$$\bar{y}_S = \bar{y} \left( \frac{\bar{X} \bar{z}}{\bar{x} \bar{Z}} \right)$$

In literature, it has been shown by various authors; see for example, Reddy (1974) and Srivenkataramana (1978) that the bias and the mean square error of the ratio estimator  $\bar{y}_R$ , can be reduced with the application of transformation on the auxiliary variable  $x$ . Thus, authors like, Srivenkataramana (1980), Bandyopadhyay (1980) and Singh *et al.* (2005) have improved on the ratio, product and ratio-cum-product method of estimation

using the transformation on the auxiliary information. We give below the transformations employed by these authors:

$$x_i^* = (1 + g)\bar{X} - gx_i \text{ and } z_i^* = (1 + g)\bar{Z} - gz_i, \text{ for } i = 1, 2, \dots, N, \quad (1)$$

where,  $g = \frac{n}{N - n}$ .

Then clearly,  $\bar{x}^* = (1 + g)\bar{X} - g\bar{x}$  and  $\bar{z}^* = (1 + g)\bar{Z} - g\bar{z}$  are also unbiased estimate of  $\bar{X}$  and  $\bar{Z}$  respectively; and  $Corr(\bar{y}, \bar{x}^*) = -\rho_{yx}$  and  $Corr(\bar{y}, \bar{z}^*) = -\rho_{yz}$ . It is to be noted that by using the transformation above, the construction of the estimators for  $\bar{Y}$  requires the knowledge of unknown parameters, which restrict the applicability of these estimators. To overcome this restriction, in practice, information on these parameters can be obtained approximately from either past experience or pilot sample survey, inexpensively.

The following estimators  $\bar{y}_R^*$ ,  $\bar{y}_P^*$  and  $\bar{y}_{SE}$  are referred to as dual to ratio, dual to product and ratio-cum-product estimators and are due to Srivenkataramana (1980), Bandyopadhyay (1980) and Singh *et al.* (2005) respectively. They are as given below:

$$\bar{y}_R^* = \bar{y} \left( \frac{\bar{x}^*}{\bar{X}} \right),$$

$$\bar{y}_P^* = \bar{y} \left( \frac{\bar{Z}}{\bar{z}^*} \right) \text{ and}$$

$$\bar{y}_{SE} = \bar{y} \left( \frac{\bar{x}^*}{\bar{X}} \frac{\bar{Z}}{\bar{z}^*} \right)$$

It is well known that the variance of the simple mean estimator  $\bar{y}$ , under SRSWOR design is  $V(\bar{y}) = \lambda S_y^2$  and to the first order of approximation, the Mean Square Errors (MSE) of  $\bar{y}_R$ ,  $\bar{y}_P$ ,  $\bar{y}_S$ ,  $\bar{y}_R^*$ ,  $\bar{y}_P^*$  and  $\bar{y}_{SE}$  are, respectively, given by

$$MSE(\bar{y}_R) = \lambda(S_y^2 + R_1^2 S_x^2 - 2R_1 S_{yx})$$

$$MSE(\bar{y}_P) = \lambda(S_y^2 + R_2^2 S_z^2 + 2R_2 S_{yz})$$

$$MSE(\bar{y}_S) = \lambda[S_y^2 - 2D + C]$$

$$MSE(\bar{y}_R^*) = \lambda(S_y^2 + g^2 R_1^2 S_x^2 - 2gR_1 S_{yx})$$

$$MSE(\bar{y}_P^*) = \lambda(S_y^2 + g^2 R_2^2 S_z^2 + 2gR_2 S_{yz})$$

$$MSE(\bar{y}_{SE}) = \lambda(S_y^2 + g^2 C - 2gD)$$

where,

$$\lambda = \frac{1-f}{n}, \quad f = \frac{n}{N}, \quad S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2, \quad S_{yx} = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})(x_i - \bar{X}),$$

$$\rho_{yx} = \frac{S_{yx}}{S_y S_x}, \quad R_1 = \frac{\bar{Y}}{\bar{X}}, \quad R_2 = \frac{\bar{Y}}{\bar{Z}}, \quad C = R_1^2 S_x^2 - 2R_1 R_2 S_{zx} + R_2^2 S_z^2, \quad D = R_1 S_{yx} - R_2 S_{yz} \text{ and}$$

$S_j^2$  for  $(j = x, y, z)$  represents the variances of  $x$ ,  $y$  and  $z$  respectively; while  $S_{yx}$ ,  $S_{yz}$  and  $S_{zx}$  denote the covariance between  $y$  and  $x$ ,  $y$  and  $z$  and  $z$  and  $x$  respectively. Note that  $\rho_{yz}$ ,  $\rho_{zx}$ ,  $S_x^2$ ,  $S_z^2$ ,  $S_{yz}$  and  $S_{zx}$  are defined analogously and respective to the subscripts used.

More recently, Sharma and Tailor (2010) proposed a new ratio-cum-dual to ratio estimator of finite population mean in Simple Random Sampling (SRS), their estimator with its MSE are respectively given as,

$$\bar{y}_{ST} = \bar{y} \left[ \delta \left( \frac{\bar{X}}{\bar{x}} \right) + (1-\delta) \left( \frac{\bar{x}^*}{\bar{X}} \right) \right]$$

$$MSE(\bar{y}_{ST}) = \lambda S_y^2 (1 - \rho_{yx}^2).$$

where, MSE was computed for optimal choice of  $\delta$ .

### 3. Proposed Dual to Ratio-Cum-Product Estimator in SRS

Using the transformation given in (1), we suggest a new estimator for  $\bar{Y}$  as follows:

$$\bar{y}_{PR} = \bar{y} \left[ \theta \left( \frac{\bar{x}^*}{\bar{X}} \frac{\bar{Z}}{\bar{z}^*} \right) + (1-\theta) \left( \frac{\bar{X}}{\bar{x}^*} \frac{\bar{z}^*}{\bar{Z}} \right) \right]$$

where,  $\theta$  is a constant, chosen to minimize the variance of  $\bar{y}_{PR}$ .

We would like to remark here that some estimators could be shown to be a particular member of  $\bar{y}_{PR}$ . For instance, when information on the auxiliary variable  $z$  is not used (or variable  $z$  takes the value 'unity') and  $\theta = 1$ , the suggested estimator  $\bar{y}_{PR}$  reduces to the 'dual to ratio' estimator  $\bar{y}_R^*$  proposed by Srivenkataramana (1980). More so,  $\bar{y}_{PR}$  reduces to the 'dual to product' estimator  $\bar{y}_p^*$  proposed by Bandyopadhyay (1980) if the information on the auxiliary variate  $x$  is not used and  $\theta = 0$ . Furthermore, the suggested estimator reduces to the dual to ratio-cum-product estimator suggested by Singh *et al.* (2005) when  $\theta = 1$  and information on the two auxiliary variables  $x$  and  $z$  are been utilized.

#### Remark

To ensure the applicability of the estimator  $\bar{y}_{PR}$ , we assume the population value of the study variate is known. This is a reasonable assumption as survey samplers usually obtain such information inexpensively through pilot survey or past experience.

In order to study the properties of the suggested estimator  $\bar{y}_{PR}$  (e.g. MSE), we write

$$\bar{y} = \bar{Y}(1 + k_0); \quad \bar{x} = \bar{X}(1 + k_1); \quad \bar{z} = \bar{Z}(1 + k_2);$$

with  $E(k_0) = E(k_1) = E(k_2) = 0$  and

$$E(k_0^2) = \frac{\lambda S_y^2}{\bar{Y}^2}; \quad E(k_1^2) = \frac{\lambda S_x^2}{\bar{X}^2}; \quad E(k_2^2) = \frac{\lambda S_z^2}{\bar{Z}^2}; \quad E(k_0 k_1) = \frac{\lambda S_{yx}}{\bar{Y}\bar{X}}; \quad E(k_0 k_2) = \frac{\lambda S_{yz}}{\bar{Y}\bar{Z}};$$

$$E(k_1 k_2) = \frac{\lambda S_{zx}}{\bar{X}\bar{Z}},$$

Now expressing  $\bar{y}_{PR}$  in terms of  $k$ 's, we have

$$\bar{y}_{PR} = \bar{Y}(1 + k_0) \left[ \theta(1 - gk_1)(1 - gk_2)^{-1} + (1 - \theta)(1 - gk_1)^{-1}(1 - gk_2) \right] \quad (2)$$

We assume that  $|gk_1| < 1$  and  $|gk_2| < 1$  so that the right hand side of (2) is expandable.

Now expanding the right hand side of (2) to the first degree of approximation, we have

$$\bar{y}_{PR} - \bar{Y} = \bar{Y} \left[ k_0 + (1 - 2\theta)g(k_1 - k_2 + k_0 k_1 - k_0 k_2) + g^2(k_1^2 - k_1 k_2 - \theta(k_1^2 - k_2^2)) \right] \quad (3)$$

Taking expectations on both sides of (3), we get the bias of  $\bar{y}_{PR}$  to the first degree of approximation, as

$$B(\bar{y}_{PR}) = \lambda \bar{Y} \left[ gDA + g^2(R_1^2 S_x^2 - R_1 R_2 S_{zx} - \theta(R_1^2 S_x^2 - R_2^2 S_z^2)) \right]$$

where  $A = 1 - 2\theta$

Squaring both sides of (3) and neglecting terms of  $k$ 's involving power greater than two, we have

$$\begin{aligned} (\bar{y}_{PR} - \bar{Y})^2 &= \bar{Y}^2 [k_0 + Agk_1 - Agk_2]^2 \\ &= \bar{Y}^2 [k_0^2 + 2Agk_0 k_1 - 2Agk_0 k_2 - 2A^2 g^2 k_1 k_2 + A^2 g^2 k_1^2 + A^2 g^2 k_2^2] \end{aligned} \quad (4)$$

Taking expectations on both sides of (4), we get the MSE of  $\bar{y}_{PR}$ , to the first order of approximation, as

$$MSE(\bar{y}_{PR}) = \lambda [S_y^2 + 2AgD + A^2 g^2 C] \quad (5)$$

The MSE equation given in (5) is minimized for

$$\theta = \frac{D + Cg}{2Cg} = \theta_0 \text{ (say)} \quad (6)$$

Substituting (6) in  $\bar{y}_{PR}$ , we get the asymptotically optimum estimator as

$$\bar{y}_{PR}^* = \bar{y} \left[ \left( \frac{1}{2} + \frac{D}{2gC} \right) \left( \frac{\bar{x}^*}{\bar{X}} \frac{\bar{Z}}{\bar{z}^*} \right) + \left( \frac{1}{2} - \frac{D}{2gC} \right) \left( \frac{\bar{X}}{\bar{x}^*} \frac{\bar{z}^*}{\bar{Z}} \right) \right].$$

We can obtain the minimum MSE of  $\bar{y}_{PR}$ , by substituting (6) in (5), which after little algebraic simplification, yields:

$$MSE(\bar{y}_{PRS}^*) = \lambda \left( S_y^2 - \frac{D^2}{C} \right), \text{ for } C \neq 0 \tag{7}$$

#### 4. Extension of the Roposed Estimator to Stratified Random Sampling

The disadvantages of using SRS technique have been comprehensively documented in literature (see for instance, Cochran (1977)). More so, studies by several authors reveal that the ratio-cum-product estimator performs better than ratio and product type estimators in SRS under stratification and other certain conditions. This therefore motivates us to extend the proposed estimator in section 3 to stratified random sampling (STRS) design and study its properties.

Consider a finite population  $P = (P_1, P_2, \dots, P_N)$  of size  $N$  divided into  $L$  homogenous strata of size  $N_h$  ( $h=1, 2, \dots, L$ ). A sample of size  $n_h$  is drawn from each stratum using SRSWOR.

Let  $y$  be the study variate taking values  $y_{hi}$  ( $i^{th}$  observation from  $h^{th}$  stratum) and  $(x, z)$  be the auxiliary variates taking values  $(x_{hi}, z_{hi})$ .

Moreover, let  $\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$ ,  $\bar{x}_{st} = \sum_{h=1}^L W_h \bar{x}_h$ , and  $\bar{z}_{st} = \sum_{h=1}^L W_h \bar{z}_h$ , be the unbiased estimators of the population mean  $\bar{Y}$  (the study variate) and  $(\bar{x}, \bar{z})$  the auxiliary variates respectively. where,

$$W_h = \frac{N_h}{N} : \text{is the weight of } h^{th} \text{ stratum}$$

$$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi} : \text{is the sample mean of the study variate } y \text{ in } h^{th} \text{ stratum.}$$

where,  $\bar{x}_h$  and  $\bar{z}_h$  are define analogously with respect to  $x$  and  $z$  respectively.

The direct generalization of dual transformation of the two auxiliary variates pioneered by Srivenkataramana (1980) and Bandyopadhyay (1980) is defined as follow:

$$x_{hi}^* = (1 + g_h) \bar{X}_h - g_h x_{hi} \text{ and}$$

$$z_{hi}^* = (1 + g_h) \bar{Z}_h - g_h z_{hi}$$

where,  $g_h = \frac{n_h}{N_h - n_h}$ , ( $h=1, 2, \dots, L$ )

We note here that the relation  $\text{corr}(\bar{y}, \bar{x}) = -\rho_{xy}$  and  $\text{corr}(\bar{y}, \bar{z}) = -\rho_{yz}$  is not valid in the case stratified sampling design.

Using the various definitions above, we propose the dual to ratio-cum-product estimator in stratified random sampling as follows:

$$\bar{y}_{PRS} = \bar{y}_{st} \left[ \alpha \left( \frac{\bar{x}_{st}^*}{\bar{X}} \frac{\bar{Z}}{\bar{z}_{st}^*} \right) + (1 - \alpha) \left( \frac{\bar{X}}{\bar{x}_{st}^*} \frac{\bar{z}_{st}^*}{\bar{Z}} \right) \right], \tag{8}$$

where,  $\alpha$  is a constant chosen to reduce the variance of  $\bar{y}_{PRS}$ .

**Remarks**

1. To ensure the applicability of the estimator  $\bar{y}_{PRS}$ , we assume the population values of the study variate are known in the entire stratum. This is a reasonable assumption as survey samplers usually obtain such information inexpensively through pilot survey or past experience.
2. We would also like to remark here that the stratified sampling case of the estimators  $\bar{y}_R^*$ ,  $\bar{y}_P^*$  and  $\bar{y}_{SE}$  due to Srivenkataramana (1980), Bandyopadhyay (1980) and Plikusas (2008) respectively, are particular member of  $\bar{y}_{PRS}$ .

To analyse the property of  $\bar{y}_{PRS}$ , for example, MSE. We define:

$$\bar{y}_{st} = \bar{Y}(1 + e_{0h}); \bar{x}_{st} = \bar{X}(1 + e_{1h}); \bar{z}_{st} = \bar{Z}(1 + e_{2h});$$

using these notations, it is easy to verify that  $E(e_{0h}) = E(e_{1h}) = E(e_{2h}) = 0$ .

$$V_{a,b,c} = \sum_{h=1}^L W_h^{a+b+c} \frac{E \left[ (\bar{y}_h - \bar{Y})^a (\bar{x}_h - \bar{X})^b (\bar{z}_h - \bar{Z})^c \right]}{\bar{Y}^a \bar{X}^b \bar{Z}^c} \tag{9}$$

From (9), we can write

$$\begin{aligned} E(e_{0h}^2) = V_{2,0,0} &= \frac{\sum_{h=1}^L W_h^2 \lambda_h S_{yh}^2}{\bar{Y}^2}, & E(e_{1h}^2) = V_{0,2,0} &= \frac{\sum_{h=1}^L W_h^2 \lambda_h S_{xh}^2}{\bar{X}^2}, \\ E(e_{2h}^2) = V_{0,0,2} &= \frac{\sum_{h=1}^L W_h^2 \lambda_h S_{zh}^2}{\bar{Z}^2}, & E(e_{0h}e_{1h}) = V_{1,1,0} &= \frac{\sum_{h=1}^L W_h^2 \lambda_h S_{yxh}}{\bar{Y}\bar{X}}, \\ E(e_{0h}e_{2h}) = V_{1,0,1} &= \frac{\sum_{h=1}^L W_h^2 \lambda_h S_{yzh}}{\bar{Y}\bar{Z}}, & E(e_{1h}e_{2h}) = V_{0,1,1} &= \frac{\sum_{h=1}^L W_h^2 \lambda_h S_{xzh}}{\bar{X}\bar{Z}}. \end{aligned}$$

where,

$$\lambda_h = \frac{1 - f_h}{n_h}, \quad f_h = \frac{n_h}{N_h}, \quad S_{yh}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2, \quad S_{yxh} = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)(x_{hi} - \bar{X}_h),$$

Note that  $S_{xh}^2$ ,  $S_{zh}^2$ ,  $S_{yzh}$  and  $S_{xzh}$  are defined analogously and respective to the subscripts used.

Expressing  $\bar{y}_{PRS}$  in terms of  $e_{hi}$  ( $i = 0,1,2$ ), we can write (8) as

$$\bar{y}_{PRS} = \bar{Y}(1 + e_{0h}) \left[ \alpha (1 - g_h e_{1h})(1 - g_h e_{2h})^{-1} + (1 - \alpha)(1 - g_h e_{1h})^{-1}(1 - g_h e_{2h}) \right] \quad (10)$$

Assuming that  $|g_h e_{1h}| < 1$  and  $|g_h e_{2h}| < 1$ , so that  $(1 - g_h e_{1h})^{-1}$  and  $(1 - g_h e_{2h})^{-1}$  are expendable.

Expanding the right hand side of (10) and following the procedure in section 3 for SRS design, we obtain:

$$MSE(\bar{y}_{PRS}) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + 2MM_2 g_h + M^2 M_1 g_h^2) \quad (11)$$

where,  $M = 1 - 2\alpha$ ;  $M_1 = R_1^2 S_{xh}^2 - 2R_1 R_2 S_{zxh} + R_2^2 S_{zh}^2$ ;  $M_2 = R_1 S_{yxh} - R_2 S_{yzh}$

The MSE equation given in (11) is minimized for

$$\alpha_{(optimal)} = \frac{1}{2} + \frac{M_2}{2M_1 g_h}, \text{ for } M_1 > 0 \quad (12)$$

Substituting (12) in  $\bar{y}_{PRS}$ , we get the asymptotically optimum estimator as

$$\bar{y}_{PRS}^* = \bar{y}_{st} \left[ \left( \frac{1}{2} + \frac{M_2}{2M_1 g_h} \right) \left( \frac{\bar{x}_{st}^* \bar{Z}}{\bar{X} \bar{z}_{st}^*} \right) + \left( \frac{1}{2} - \frac{M_2}{2M_1 g_h} \right) \left( \frac{\bar{X} \bar{z}_{st}^*}{\bar{x}_{st}^* \bar{Z}} \right) \right].$$

We can obtain the minimum MSE of  $\bar{y}_{PRS}$ , by substituting (5) in (4), which after little algebraic simplification, yields:

$$MSE(\bar{y}_{PRS}^*) = \sum_{h=1}^L W_h^2 \lambda_h \left( S_{yh}^2 - \frac{M_2^2}{M_1} \right) \text{ for } M_1 > 0 \quad (13)$$

Assuming that the study variate  $y$  and the auxiliary variate  $x$  are positively correlated, Hansen *et al.* (1946) utilized known value of population mean  $\bar{X}$  of auxiliary variate  $x$  and defined combined ratio estimator for population mean  $\bar{Y}$  as:

$$\bar{y}_{R(S)} = \bar{y}_{st} \left( \frac{\bar{X}}{\bar{x}_{st}} \right).$$

When the study variate  $y$  and the auxiliary variate  $z$  are negatively correlated, assuming that the population mean  $\bar{Z}$  of auxiliary variate  $z$  is known, combined product estimator is defined as:

$$\bar{y}_{P(S)} = \bar{y}_{st} \left( \frac{\bar{z}_{st}}{\bar{Z}} \right).$$



Furthermore, Tailor *et al.* (2012) extended the ratio-cum-product estimator proposed by Singh (1967) to STRS as follows:

$$\bar{y}_{S(S)} = \bar{y}_{st} \left( \frac{\bar{X}}{\bar{x}_{st}} \frac{\bar{z}_{st}}{\bar{Z}} \right).$$

Here, the assumption is that the population means of the auxiliary variates  $\bar{X}$  and  $\bar{Z}$  are known in each stratum.

More so, Plikusas (2008) defined dual to ratio-cum-product estimator expressed in Singh *et al.* (2005) to STRS as

$$\bar{y}_{SE(S)} = \bar{y}_{st} \left( \frac{\bar{x}_{st}^*}{\bar{X}} \right) \left( \frac{\bar{Z}}{\bar{z}_{st}^*} \right)$$

The MSEs of these four estimators, up to the first degree of approximation, are as presented:

$$MSE(\bar{y}_{R(S)}) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + R_1^2 S_{xh}^2 - 2R_1 S_{yxh})$$

$$MSE(\bar{y}_{P(S)}) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + R_2^2 S_{zh}^2 + 2R_2 S_{yzh})$$

$$MSE(\bar{y}_{S(S)}) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + M_1 - 2M_2)$$

$$MSE(\bar{y}_{SE(S)}) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + g_h^2 M_1 - 2g_h M_2)$$

Moreover, we have extended the estimators  $\bar{y}_R^*$ ,  $\bar{y}_P^*$ , and  $\bar{y}_{ST}$  to stratified sampling design and denote as  $\bar{y}_{R(S)}^*$ ,  $\bar{y}_{P(S)}^*$  and  $\bar{y}_{ST(S)}$  respectively. They are presented as follows:

$$\bar{y}_{R(S)}^* = \bar{y}_{st} \left( \frac{\bar{x}_{st}^*}{\bar{X}} \right)$$

$$\bar{y}_{P(S)}^* = \bar{y}_{st} \left( \frac{\bar{Z}}{\bar{z}_{st}^*} \right)$$

$$\bar{y}_{ST(S)} = \bar{y}_{st} \left[ \delta \left( \frac{\bar{X}}{\bar{x}_{st}} \right) + (1 - \delta) \left( \frac{\bar{x}_{st}^*}{\bar{X}} \right) \right]$$

The variance of the usual unbiased estimator in stratified sampling,  $\bar{y}_{st}$ , is given as

$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \lambda_h S_{yh}^2$ . And to the first order of approximation, the MSEs of  $\bar{y}_{R(S)}^*$ ,  $\bar{y}_{P(S)}^*$ ,  $\bar{y}_{SE(S)}$  and  $\bar{y}_{ST(S)}$  are, respectively, given by:

$$MSE(\bar{y}_{R(S)}^*) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + g_h^2 R_1^2 S_{xh}^2 - 2g_h R_1 S_{yxh})$$

$$MSE(\bar{y}_{P(S)}^*) = \sum_{h=1}^L W_h^2 \lambda_h (S_{yh}^2 + g_h^2 R_2^2 S_{zh}^2 + 2g_h R_2 S_{yzh})$$

$$MSE(\bar{y}_{ST(S)}) = \sum_{h=1}^L W_h^2 \lambda_h S_{yh}^2 (1 - \rho_{yxh}^2).$$

### 5. Efficiency Comparison

In this section, the efficiency of the suggested estimator  $\bar{y}_{PR}$  over the following estimator,  $\bar{y}$ ,  $\bar{y}_R$ ,  $\bar{y}_P$ ,  $\bar{y}_S$ ,  $\bar{y}_R^*$ ,  $\bar{y}_P^*$ ,  $\bar{y}_{SE}$  and  $\bar{y}_{ST}$  are investigated. We will have the conditions as follows:

- (a)  $MSE(\bar{y}_{PR}) - V(\bar{y}) < 0$  if  $\frac{-D^2}{C} < 0$ . This is always true because  $C > 0$ .
- (b)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_R) < 0$  if  $\frac{-D^2}{C} < R_1^2 S_x^2 - 2R_1 S_{yx}$ , is always true since  $C > 0$ .
- (c)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_P) < 0$  if  $\frac{-D^2}{C} < R_2^2 S_z^2 + 2R_2 S_{yz}$ , is always true since  $C > 0$ .
- (d)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_S) < 0$  if  $(C - D)^2 > 0$
- (e)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_R^*) < 0$  if  $\frac{-D^2}{C} < g^2 R_1^2 S_x^2 - 2gR_1 S_{yx}$ , is always true since  $C > 0$ .
- (f)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_P^*) < 0$  if  $\frac{-D^2}{C} < g^2 R_2^2 S_z^2 + 2gR_2 S_{yz}$ , is always true since  $C > 0$ .
- (g)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_{SE}) < 0$  if  $(D - gC)^2 > 0$
- (h)  $MSE(\bar{y}_{PR}) - MSE(\bar{y}_{ST}) < 0$  if  $-\frac{D^2}{C} < \rho_{yx}^2 S_y^2$ , where  $C > 0$ .

The efficiency of the suggested estimator in stratified random sampling over the estimator  $\bar{y}_{st}$ ,  $\bar{y}_{R(S)}$ ,  $\bar{y}_{P(S)}$ ,  $\bar{y}_{S(S)}$ ,  $\bar{y}_{R(S)}^*$ ,  $\bar{y}_{P(S)}^*$ ,  $\bar{y}_{SE(S)}$  and  $\bar{y}_{ST(S)}$  are as presented.

- (i)  $MSE(\bar{y}_{PRS}) - V(\bar{y}_{st}) < 0$  if  $-\sum_{h=1}^L W_h^2 \lambda_h \frac{M_2^2}{M_1} < 0$  for  $M_1 > 0$ .

- (ii)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{R(S)}) < 0$  if  

$$-\sum_{h=1}^L W_h^2 \lambda_h \frac{M_2^2}{M_1} < \sum_{h=1}^L W_h^2 \lambda_h (R_1^2 S_{xh}^2 - 2R_1 S_{yxh}), \text{ for } M_1 > 0.$$
- (iii)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{P(S)}) < 0$  if  

$$-\sum_{h=1}^L W_h^2 \lambda_h \frac{M_2^2}{M_1} < \sum_{h=1}^L W_h^2 \lambda_h (R_2^2 S_{zh}^2 - 2R_2 S_{yzh}), \text{ for } M_1 > 0.$$
- (iv)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{S(S)}) < 0$  if  $\sum_{h=1}^L W_h^2 \lambda_h (M_1 - M_2)^2 > 0$
- (v)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{R(S)}^*) < 0$  if  

$$-\sum_{h=1}^L W_h^2 \lambda_h \frac{M_2^2}{M_1} < \sum_{h=1}^L W_h^2 \lambda_h (g_h^2 R_1^2 S_{xh}^2 - 2g_h R_1 S_{yxh}), \text{ for } M_1 > 0.$$
- (vi)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{P(S)}^*) < 0$  if  

$$-\sum_{h=1}^L W_h^2 \lambda_h \frac{M_2^2}{M_1} < \sum_{h=1}^L W_h^2 \lambda_h (g_h^2 R_2^2 S_{zh}^2 - 2g_h R_2 S_{yzh}), \text{ for } M_1 > 0.$$
- (vii)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{SE(S)}) < 0$  if  $\sum_{h=1}^L W_h^2 \lambda_h (g_h M_1 - M_2)^2 > 0$
- (viii)  $MSE(\bar{y}_{PRS}) - MSE(\bar{y}_{ST(S)}) < 0$  if  $-\sum_{h=1}^L W_h^2 \lambda_h \left( \frac{M_2^2}{M_1} - \rho_{yxh}^2 \right) < 0$ , where  $M_1 > 0$ .

**Remark:** Efficiency comparisons in case of proportional allocation

We would like to remark here that when the units from the  $h^{th}$  stratum are selected according to proportional allocation i.e.,  $n_h \propto N_h$  then  $\frac{n_h}{N_h} = \frac{n}{N}$ . In this case, the conditions (i) – (viii) holds but we now replace  $\lambda_h$  with  $\lambda$ .

## 6. Numerical Illustration

In this section, we analyze the performance of the suggested estimator with respect to other estimators considered in this paper. To achieve this, four natural population data sets from the literature are considered. The sources and brief descriptions of these populations are presented below. We note that the first two populations are used for estimators under SRS while the last two are used for STRS.

- (1) **Population I** [Singh (1969, p. 377)]; a detailed description can be found in Singh (1965)

$y$  : Number of females employed  
 $x$  : Number of females in service  
 $z$  : Number of educated females

$$N = 61, n = 20, \bar{Y} = 7.46, \bar{X} = 5.31, \bar{Z} = 179, S_y^2 = 28.0818, S_x^2 = 16.1761, S_z^2 = 2028.1953, \rho_{xy} = 0.7737, \rho_{yz} = -0.2070, \rho_{zx} = -0.0033,$$

(2) **Population II** [Source: Johnston 1972, p. 171]; A detailed description of these variables is shown in Table 1.

y : Percentage of hives affected by disease

x : Mean January temperature

z : Date of flowering of a particular summer species (number of days from January 1)

$$N = 10, n = 10, \bar{Y} = 52, \bar{X} = 42, \bar{Z} = 200, S_y^2 = 65.9776, S_x^2 = 29.9880, S_z^2 = 84, \rho_{xy} = 0.8, \rho_{yz} = -0.94, \rho_{zx} = -0.73,$$

For population II, the population size is small, we therefore used the whole sample method to determine the MSEs of the various estimators (this will lead to a more precise results). We note that when  $N = n$ , the MSEs of the proposed and all the estimators considered in this study would be zero. However, looking at the MSEs of these estimators, it is easy to find that they don't depend on sample size. Thus, to calculate the MSEs we can ignore the constant  $\lambda$  common to all these estimators.

**Table 1: Description of Population II**

	y	x	z
	49	35	200
	40	35	212
	41	38	211
	46	40	212
	52	40	203
	59	42	194
	53	44	194
	61	46	188
	55	50	196
	64	50	190

(3) **Population III** [Source: National Horticulture Board (2010), retrieved in Tailor *et al.* (2012)]

y : Productivity (MT/Hectare)

x : Production in thousand tons

z : Area in thousand hectares

Strata	$n_h$	$N_h$	$S_{xh}$	$S_{yh}$	$S_{zh}$	$S_{yxh}$	$S_{yzh}$	$S_{xzh}$	$\rho_{yxh}$	$\bar{X}_h$	$\bar{Y}_h$	$\bar{Z}_h$
1	3	10	3.53	0.54	1.19	1.6	-0.02	1.75	0.839366	10.41	1.7	6.2
2	4	10	80.54	1.41	10.81	83.47	-7.06	68.57	0.735021	309.1	3.67	80.67

(4) **Population IV** [Source: Murthy (1967)]

y : Output      x : Fixed capital      z : Number of workers

Strata	$n_h$	$N_h$	$S_{xh}$	$S_{yh}$	$S_{zh}$	$S_{yxh}$	$S_{yzh}$	$S_{xzh}$	$\rho_{yxh}$	$\bar{X}_h$	$\bar{Y}_h$	$\bar{Z}_h$
1	2	5	74.87	615.9	0.75	39360.7	411.16	38.08	0.854	214.4	1926	51.8
2	3	5	66.35	340.4	4.84	22356.5	1536.24	287.92	0.99	333.8	315.6	60.6

For these comparisons, the Percent Relative Efficiencies (PREs) of the different estimators are computed with respect to the usual unbiased estimator  $\bar{y}$ , using the formula

$$PRE(\bar{y},.) = \frac{V(\bar{y})}{MSE(.)} \times 100$$

and they are as presented in Table 2.

Table 2 shows clearly that the proposed dual to ratio-cum-product estimator  $\bar{y}_{PR}$  has the highest PRE than other estimators; therefore, we can conclude based on the study populations that the suggested estimator is more efficient than the usual unbiased estimators, the traditional ratio and product estimator, ratio-cum-product estimator by Singh (1969), Srivenkataramana (1980) estimator, Bandyopadhyay (1980) estimator, Singh *et al.* (2005) estimator and Sharma and Tailor (2010).

Under SRSWOR, the suggested estimator as demonstrated through the theory and empirical results (populations I and II) is always better than estimators considered in this study when one of the auxiliary variate is positively correlated with the study variate, the other is negatively correlated with the study variable and the two (auxiliary variates) are negatively correlated with each other. It is also observed from Table 2, that the product method of estimators  $\bar{y}_P$ ,  $\bar{y}_P^*$ ,  $\bar{y}_{P(S)}$ , and  $\bar{y}_{P(S)}^*$  perform poorly for the study populations especially under the stratified sampling design.

**Table 2: PRE of the different estimators with respect to  $\bar{y}$**

Estimators	SRS design case		Estimators	STRS design case	
	Population I	Population II		Population III	Population IV
$\bar{y}$	100	100	$\bar{y}_{st}$	100	100
$\bar{y}_R$	205	277	$\bar{y}_{R(S)}$	176	314
$\bar{y}_P$	102	187	$\bar{y}_{P(S)}$	121	85
$\bar{y}_S$	214	395	$\bar{y}_{S(S)}$	298	258
$\bar{y}_R^*$	215	239	$\bar{y}_{R(S)}^*$	181	232
$\bar{y}_P^*$	105	150	$\bar{y}_{P(S)}^*$	119	80
$\bar{y}_{SE}$	236	402	$\bar{y}_{SE(S)}$	270	216
$\bar{y}_{ST}$	250	278	$\bar{y}_{ST(S)}$	233	470
$\bar{y}_{PR}^*$	279	457	$\bar{y}_{PRS}^*$	390	391

Furthermore, it is observed from the empirical analysis that the proposed estimator under STRS is better than the usual unbiased estimators in STRS, Hansen *et al.* (1946), stratified product estimator, Singh (1969), Srivenkataramana (1980), Bandyopadhyay (1980), and Plikusas (2008) estimator except for Sharma and Tailor (2010) in the population IV.

## 7. Conclusion

The use of auxiliary information to increase the precision of the estimate has received numerous attentions from several authors. In this paper, we continue this research by developing a new estimator under SRSWOR, this estimator is further extended to STRS design. These estimators are found to be more efficient than the usual unbiased estimator, the traditional ratio [Cochran (1940)] and product estimators and the estimators proposed by Singh (1969), Srivenkataramana (1980), Bandyopadhyay (1980), Singh *et al.* (2005), Sharma and Tailor (2010) and Tailor *et al.* (2012) under simple and stratified random sampling. The theoretical inference (results) is supported by the result of an application with original data.

## References

1. Abu-Dayyeh, W. A., Ahmed, M. S., Ahmed, R. A., and Muttlak, H. A. (2003): Some estimators of finite population mean using auxiliary information, *Applied Mathematics and Computation*, 139, 287–298.
2. Bandyopadhyay, S. (1980): Improved ratio and product estimators, *Sankhya series C* 42, 45- 49.
3. Cochran, W. G. (1940): The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, *The Journal of Agricultural Science*, 30, 262–275.
4. Cochran, W. G. (1977): *Sampling techniques*. Wiley, Third edition
5. Diana, G. and Perri, P.F. (2007): Estimation of finite population mean using multi-auxiliary information. *Metron*, Vol. LXV, Number 1, 99-112
6. Hansen, M.H., Hurwitz, W.N., and M.Gurney, M. (1946): Problems and methods of the sample survey of business, *Journal of American Statistical Association*, 41, 173-189.
7. Johnston, J. (1972): *Econometric methods*, (2nd edn), McGraw-Hill, Tokyo.
8. Kadilar, C. and Cingi, H. (2004): Estimator of a population mean using two auxiliary variables in simple random sampling, *International Mathematical Journal*, 5, 357–367.
9. Kadilar, C. and Cingi, H. (2005): A new estimator using two auxiliary variables, *Applied Mathematics and Computation*, 162, 901–908.
10. Murthy, M. N. (1964): Product method of estimation. *Sankhya*, A, 26, 294-307.
11. Olkin, I. (1958): Multivariate ratio estimation for finite populations, *Biometrika*, 45, 154–165.
12. Perri, P. F. (2005): Combining two auxiliary variables in ratio-cum-product type estimators. *Proceedings of Italian Statistical Society, Intermediate Meeting on Statistics and Environment*, Messina, 21-23 September 2005, 193–196.

13. Plikusas, A. (2008): Some overview of the ratio type estimators In: Workshop on survey sampling theory and methodology, Statistics Estonia.
14. Raj, D. (1965): On a method of using multi-auxiliary information in sample surveys, *Journal of the American Statistical Association*, 60, 154–165.
15. Reddy, V. N. (1974): On a transformed ratio method of estimation, *Sankhya C*, 36, 59–70.
16. Sharma, B. and Tailor, R. (2010): A New Ratio-Cum-Dual to Ratio Estimator of Finite Population Mean in Simple Random Sampling. *Global Journal of Science Frontier Research*, Vol. 10, Issue 1, 27-31
17. Singh, M.P. (1965): On the estimation of ratio and product of the population parameters, *Sankhya series B* 27, 321-328.
18. Singh, M. P. (1967): Multivariate product method of estimation for finite populations, *Journal of the Indian Society of Agricultural Statistics*, 31, 375–378.
19. Singh, M.P. (1969): Comparison of some ratio-cum-product estimators, *Sankhya series B* 31, 375-378.
20. 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), 51-56
21. Srivenkataramana, T. (1978): Change of origin and scale in ratio and difference methods of estimation in sampling, *Canad. J. Stat.*, 6, 79–86.
22. Srivenkataramana, T. (1980): A dual to ratio estimator in sample surveys, *Biometrika*, 67(1), 199–204.
23. Tailor, R., Chouhan, S., Tailor, R., and Garg, N. (2012): A ratio-cum-product estimator of population mean in stratified random sampling using two auxiliary variables, *STATISTICA*, anno LXXII, n. 3, 287 – 297.

### **Acknowledgement**

The author is most grateful to the four anonymous referees for their valuable comments and suggestions which have greatly improved the earlier draft of this paper.