STATISTICA, anno LXXIII, n. 3, 2013

IMPROVED CLASS OF ESTIMATORS OF FINITE POPULATION MEAN USING SAMPLING FRACTION AND INFORMATION ON TWO AUXILIARY VARIABLES IN SAMPLE SURVEYS

H. P. Singh,

A. Rathour,

R. S. Solanki*

School of Studies in Statistics, Vikram University, Ujjain-456010, M.P., India.

1. INTRODUCTION

Auxiliary information has been widely used in sample surveys to increase the efficiency of the estimators. For this reason, many authors used the auxiliary information at the estimation stage. When the correlation between study variable and auxiliary variable is positive (high) ratio method of estimation due to (Cochran, 1940) is used. On the other hand if the correlation is negative (high) product method of estimation due to (Robson, 1957) and revisited by (Murthy, 1964) is preferred.

In addition several authors such as (Srivastava, 1980); (Srivenkataramana and Tracy, 1980); (Prasad, 1989); (Bahl and Tuteja, 1991); (Mohanty and Sahoo, 1995); (Singh and Gangele, 1998); (Upadhyaya and Singh, 1999); (Singh and Ruiz-Espejo, 2003); (Singh *et al.* , 2004, 2009, 2010); (Singh and Tailor, 2005); (Singh *et al.* , 2007); (Singh and Karpe, 2009); (Tailor and Sharma, 2009); (Upadhyaya *et al.* , 2011) (Singh and Solanki, 2012)and (Solanki *et al.* , 2012) suggested some modified estimators of population mean using information on single auxiliary. (Singh, 1967); suggested a ratio-cum-product estimator of population mean using information on two auxiliary variables. In this paper we have made the use of information on two auxiliary variables and the sampling fraction in defining the estimators of population mean and study their properties under large sampling approximation. An empirical is carried out to judge the merits of the suggested estimators over usual unbiased estimator and (Singh, 1967); ratio-cum-product estimator.

Consider a finite population $U=(U_1,U_2,...,U_N)$ containing N units. Let Y, X_1 and X_2 are denote the study variable and the auxiliary variables taking values Y_j, X_{1j} and X_{2j} respectively on the unit U_j (j = 1, 2, ..., N) of population U, where X_1 is positively and X_2 is negatively correlated with the study variable Y. Let a sample of size n is drawn by simple random sampling without replacement (SRSWOR) from population U to estimate the population mean $\bar{Y}$ of the study variable Y. It is very well known that sample mean $\bar{y}\ [= n^{-1}\sum_{i=1}^{n} Y_i]$ of the study variable Y is an unbiased estimator of population mean $\bar{Y}$ and under SRSWOR the variance of $\bar{y}$ is given by

* Corresponding author: ramkssolanki@gmail.com

$$Var\left(\bar{y}\right) = MSE\left(\bar{y}\right) = \theta S_y^2 = \theta \bar{Y}^2 C_y^2 \tag{1}$$

where $f = (n / N)$, $\theta = \{(1-f)/n\}$, $C_y^2 = (S_y^2 / \bar{Y}^2)$, $S_y^2 = (N-1)^{-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$ and $MSE(\bullet)$ stands for the mean square error of $(\bullet)$.

Assuming the knowledge of the population mean $\bar{X}_1$ and $\bar{X}_2$ of the auxiliary variables X_1 and X_2 respectively, (Singh, 1967); proposed a ratio-cum-product estimator for population mean $\bar{Y}$ as

$$t_{RP} = \bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2}) \tag{2}$$

where $\bar{x}_1 = n^{-1}\sum_{i=1}^{n} X_{1i}$ and $\bar{x}_2 = n^{-1}\sum_{i=1}^{n} X_{2i}$ are sample mean of auxiliary variables $\bar{X}_1$ and $\bar{X}_2$ respectively.

To the first degree of approximation, the bias and *MSE* of t_{RP} are respectively given by

$$B(t_{RP}) = \theta\bar{Y}[C_{x1}^2(1-K_{yx1}) + C_{x2}^2(K_{yx2} - K_{x1x2})] \tag{3}$$

$$MSE(t_{RP}) = \theta\bar{Y}^2[C_y^2 + C_{x1}^2(1-2K_{yx1}) + C_{x2}^2\{1+2(K_{yx2} - K_{x1x2})\}] \tag{4}$$

where

$$C_{xh}^2 = (S_{xh}^2 / \bar{X}_h^2), S_{xh}^2 = (N-1)^{-1}\sum_{j=1}^{N}(X_{hj} - \bar{X}_h)^2,\ K_{yxh} = \rho_{yxh}(C_y / C_{xh}),$$

$$\rho_{yxh} = (S_{yxh} / S_{xh}S_y),\ S_{yxh} = (N-1)^{-1}\sum_{j=i}^{N}(X_{hj} - \bar{X}_h)(Y_j - \bar{Y})\ (h = 1, 2);$$

$$K_{x1x2} = \rho_{x1x2}(C_{x1} / C_{x2}),\ \rho_{x1x2} = (S_{x1x2} / S_{x1}S_{x2})$$

and $S_{x1x2} = (N-1)^{-1}\sum_{j=1}^{N}(X_{1j} - \bar{X}_1)(X_{2j} - \bar{X}_2)$.

The remaining part of the paper is as follows. In Section 2 we have suggested a generalized class of estimators of population mean and studied their properties. Some members of suggested class have been generated. Section 3 made some theoretical comparisons of suggested class with other existing estimators of population mean. An empirical study is carried out in Section 4. We have end the paper with final conclusion.

2. SUGGESTED ESTIMATORS

We have proposed the following generalized class of estimators for estimating the population mean $\bar{Y}$ as a linear combination between $\bar{y}$ and t_{RP}

$$t = [w\bar{y} + (1-w)\bar{y}t_{RP}] = [w\bar{y} + (1-w)\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})] \qquad (5)$$

where $w = (a+bf)(c+df)^{-1}$ is a function of *f* and (*a*, *b*, *c*, *d*) are suitably chosen scalars for designing the different estimators which are shown in Table 1.
To obtain the bias and *MSE* of proposed class of estimators *t* we write

$$\bar{y} = \bar{Y}(1+e_0),\ \bar{x}_1 = \bar{X}_1(1+e_1) \text{ and } \bar{x}_2 = \bar{X}_2(1+e_2),$$

such that $E(e_0) = E(e_1) = E(e_2) = 0$
and to the first degree of approximation

$$E(e_0^2) = \theta\, C_y^2,\ E(e_1^2) = \theta\, C_{x1}^2,\ E(e_2^2) = \theta\, C_{x2}^2,\ E(e_0e_1) = \theta K_{yx1}C_{x1}^2,$$

$$E(e_0e_2) = \theta K_{yx2}C_{x2}^2 \text{ and } E(e_1e_2) = \theta K_{x1x2}C_{x2}^2$$

where

$$K_{yx1} = \rho_{yx1}(C_y / C_{x1}),\ K_{yx2} = \rho_{yx2}(C_y / C_{x2}),\ \rho_{yx1} = (S_{yx1} / S_y S_{x1}),$$

$$\rho_{yx2} = (S_{yx2} / S_y S_{x2}),\ S_{yx1} = (N-1)^{-1}\sum_{j=1}^{N}(Y_j - \bar{Y})(X_{1j} - \bar{X}_1) \text{ and}$$

$$S_{yx2} = (N-1)^{-1}\sum_{j=1}^{N}(Y_j - \bar{Y})(X_{2j} - \bar{X}_2).$$

TABLE 1
Some members of generalized class of estimators t

Estimator	Values of constants				w
	a	b	c	d	
$t_0 = \bar{y}$	0	1	0	1	1
$t_1 = t_{RP}$	0	0	-	-	0
$t_2 = f\,\bar{y} + (1-f)\,\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	0	1	1	0	f
$t_3 = 2f\,\bar{y} + (1-2f)\,\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	2	1	0	2f
$t_4 = (1-f)\bar{y} + f\,\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	-1	1	0	(1-*f*)
$t_5 = (1-2f)\bar{y} + 2f\,\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	-2	1	0	(1-2*f*)
$t_6 = \frac{1}{(1+f)}\bar{y} + \frac{f}{(1+f)}\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	0	1	1	$\frac{1}{(1+f)}$
$t_7 = \frac{(1-f)}{(1+f)}\bar{y} + \frac{2f}{(1+f)}\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	-1	1	1	$\frac{(1-f)}{(1+f)}$
$t_8 = \frac{2f}{(1+f)}\bar{y} + \frac{(1-f)}{(1+f)}\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	0	2	1	1	$\frac{2f}{(1+f)}$
$t_9 = \frac{(1-2f)}{(1+f)}\bar{y} + \frac{3f}{(1+f)}\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	-2	1	1	$\frac{(1-2f)}{(1+f)}$
$t_{10} = \frac{(1-2f)}{(1+2f)}\bar{y} + \frac{4f}{(1+2f)}\bar{y}(\frac{\bar{X}_1}{\bar{x}_1})(\frac{\bar{x}_2}{\bar{X}_2})$	1	-2	1	2	$\frac{(1-2f)}{(1+2f)}$

Expressing (5) in terms of *e*'s we have

$$t = \bar{Y}[w(1+e_0) + (1-w)(1+e_0)(1+e_1)^{-1}(1+e_2)] \tag{6}$$

We assume $|e_1| < 1$ so that $(1+e_1)^{-1}$ is expendable, thus expanding, multiplying out and neglecting terms of *e*'s having power greater than two we have

$$t \cong \bar{Y}[1 + e_0 - (1-w)(e_1 - e_2 + e_0e_1 - e_0e_2 - e_1^2 + e_1e_2)]$$

or

$$(t - \bar{Y}) \cong \bar{Y}[e_0 - (1-w)(e_1 - e_2 + e_0e_1 - e_0e_2 - e_1^2 + e_1e_2)] \tag{7}$$

Taking expectation on both side of (7), we get the bias of t to the first degree of approximation as

$$B(t) = (1-w)\overline{Y}\theta[C_{x1}^2(1-K_{yx1}) + C_{x2}^2(K_{yx2} - K_{x1x2})]. \tag{8}$$

Squaring both side of (7) and neglecting terms of *e*'s having power greater than two we have

$$(t-\overline{Y})^2 \cong \overline{Y}^2[e_0^2 + (1-w)^2(e_1^2 + e_2^2 - 2e_1e_2) - 2(1-w)(e_0e_1 - e_0e_2)]. \tag{9}$$

Taking expectation on both side of (9), we get the *MSE* of the proposed class of estimators t to the first degree of approximation as

$$MSE(t) = \theta\overline{Y}^2[C_y^2 + A + w^2B - 2wC], \tag{10}$$

where

$$A = [C_{x1}^2(1-2K_{yx1}) + C_{x2}^2\{1+2(K_{yx2} - K_{x1x2})\}] = [B - 2(C_{x1}^2K_{yx1} - C_{x2}^2K_{yx2})],$$

$$B = [C_{x1}^2 + C_{x2}^2(1-2K_{x1x2})],$$

$$C = [C_{x1}^2(1-K_{yx1}) + C_{x2}^2(1+K_{yx2} - 2K_{x1x2})] = [B - (C_{x1}^2K_{yx1} - C_{x2}^2K_{yx2})],$$

being $K_{yx2} < 0 < K_{yx1}$, then $(C_{x1}^2K_{yx1} - C_{x2}^2K_{yx2}) > 0$ and $A < C < B$.

The *MSE*s of estimators which are members of suggested class of estimators *t* would be easily obtained through (2.6) putting relevant value of *w* (i.e. different values of scalars *a*, *b*, *c* and *d*) as shown in Table 1. In the next section have made some theoretical comparisons of suggested class *t* with usual unbiased estimator $\overline{y}$ and singh (1967) estimator t_{RP}.

3. EFFICIENCY COMPARISON

From (1) and (10) we have

$$MSE(t) < MSE(\overline{y}) \text{ if}$$
$$A - 2wC + w^2B < 0 \tag{11}$$

or

$$\left(\frac{2C}{B} - 1\right) < w < 1. \tag{12}$$

From (4) and (10) we have

$$MSE(t) < MSE(t_{RP}) \text{ if}$$
$$\min.\left(0, \frac{2C}{B}\right) < w < \max.\left(0, \frac{2C}{B}\right). \tag{13}$$

in order to judge the biasedness of proposed class of estimators t we have defined the quantity as

$$Q=\left|\frac{B(t)}{\theta\bar{Y}}\right| =\left|(1-w)[C_{x1}^2(1-K_{yx1})+C_{x2}^2(K_{yx2}-K_{x1x2})]\right|. \quad (14)$$

the quantity q for different estimators which are members of suggested class of estimators t would be easily obtained through (4) by putting relevant value of w as shown in Table 1.

4. EMPIRICAL STUDY

To observe the relative performance of different estimators of $\bar{Y}$, we have considered the natural population data set given in Steel and Torrie (1960, p. 282). The population description is given below:

Y = log of leaf burn in sec, X_1 = Potassium percentage, X_2 = Chlorine percentage.

$N=30$, $\bar{Y}=0.6860$, $C_y=0.4803$, $\rho_{yx1}=0.1797$, $n=6$, $\bar{X}_1=4.6537$, $C_{x1}=0.2295$, $\rho_{yx2}=-0.4996$, $\bar{X}_2=0.8077$, $C_{x2}=0.7493$, $\rho_{x1x2}=0.4074$.

We have computed the percent relative efficiency (*PRE*) with respect to $\bar{y}$ and quantity Q of different estimators of population mean $\bar{Y}$ and summarized in Table 2.

TABLE 2
PRE and quantity Q of different estimators of $\bar{Y}$

Estimator	PRE	QuantityQ
$t_0=\bar{y}$	100.00	0.0000
$t_1=t_{RP}$	75.52	0.2170
t_2	107.46	0.1736
t_3	142.58	0.1302
t_4	135.85	0.0434
t_5	157.10	0.0868
t_6	130.10	0.0362
t_7	153.50	0.0723
t_8	131.66	0.1447
t_9	154.22	0.1085
t_{10}	146.62	0.1240

It is observed from Table 2 that the estimators t_i, (i = 2, 3, ..., 10) which are members of suggested class of estimators t are biased but marginal in magnitude. The estimator t_6 is least biased estimator followed by t_4 and t_7 among estimators t_i, (i = 1, 2, ..., 10). Further it is observed from Table 2 that the estimators t_i, (i = 2, 3, ..., 10) are more efficient than the usual unbiased estimator $\bar{y}$ and Singh (1967) ratio-cum-product estimator t_{RP} in respect of *PRE*. The estimator t_5 is best in the sense of having largest *PRE* [i.e. $PRE(t_5) = 157.10$] among all the estimators t_i, ($i =$ 0, 1, ..., 10) followed by t_9 and t_7.

5. CONCLUSION

The present article suggests a generalized class of estimators of population mean using information on two auxiliary variables and sampling fraction. The members of suggested class have been generated by choosing suitable values of scalars involved in the class. The usual unbiased estimator and (Singh, 1967) ratio-cum-product estimator are identified as members of suggested class. The bias and *MSE* expressions of the suggested class have been obtained to the first degree of approximation. The properties of the members of the suggested class have been obtained through the properties of the suggested class. Thus the suggested class unifies several results at one place.

The conditions are obtained under which the proposed class of estimators is more efficient than the usual unbiased estimator and the (Singh, 1967) ratio-cum-product estimator. The theoretical findings have been examined with an empirical study. It has been observed that the members of suggested class are more efficient than the usual unbiased estimator and (Singh, 1967) ratio-cum-product estimator. Thus the proposal of the suggested class of estimators are justified.

ACKNOWLEDGEMENTS

Authors are thankful to the learned referee for his valuable suggestions regarding improvement of the paper.

REFERENCES

B. PRASAD (1989). *Some improved ratio type estimators of population mean and ratio in finite population sample surveys*. Commun. Statist. Theo. Meth., 18, (1), pp. 379–392.

D.S. ROBSON (1957). *Applications of multivariate polykays to the theory of unbiased ratio-type estimation*. Jour. Amer. Statist. Assoc., 52, (280), pp. 511-522.

H.P. SINGH L.N. UPADHYAYA,R. TAILOR (2009). *Ratio-cum-product type exponential estimator.* Statistica, 69, (4), pp. 299-310.

H.P. SINGH, M. RUIZ-ESPEJO (2003). *On linear regression and ratio-product estimation of a finite population mean*. The Statistician, 52, (1), pp. 59-67.

H.P. SINGH, N. KARPE (2009). *On the estimation of ratio and product of two population means using supplementary information in presence of measurement errors*. Statistica, 69, (1), pp. 27-47.

H.P. SINGH, R. TAILOR (2005). *Estimation of finite population mean using known correlation coefficient between auxiliary characters*. Statistica, 65, (4), pp. 407-418.

H.P. SINGH, R.K. GANGELE (1998). *Classes of almost unbiased ratio and product-type estimators in two-phase sampling*. Statistica, 59, (1), pp. 109-124.

H.P. SINGH, R.S. SOLANKI (2012). *Improved estimation of population mean in simple random sampling using information on auxiliary attribute*. Appl. Math. Comp., 218, (15), pp. 7798-7812.

H.P. SINGH, RAJESH TAILOR, RITESH TAILOR, M.S. KAKRAN (2004). *An improved estimator of population mean using power transformation.* Jour. Ind. Soc. Agri. Statist., 58, (2), pp. 223-230.

H.P. SINGH, RITESH TAILOR, RAJESH TAILOR (2010). *On ratio and product methods with certain known population parameters of auxiliary variable in sample surveys*. SORT, 34, (2), 157-180.

L.N. UPADHYAYA, H.P. SINGH (1999). *Use of transformed auxiliary variable in estimating the finite population mean*. Biometrical Jour., 41, (5), pp. 627-636.

L.N. UPADHYAYA, H.P. SINGH, S. CHATTERJEE, R. YADAV, R. (2011). *Improved ratio and product exponential type estimators.* Jour. Statist. Theo. Prac., 5, (2), pp. 285-302.

M.N. MURTHY (1964). *Product method of estimation*. Sankhya, A, 26, (1), pp. 69-74.

M.P. SINGH (1967). *Ratio cum product method of estimation.* Metrika, 12, (1), pp. 34-42.

R. SINGH, P. CHAUHAN, N. SAWAN (2007). *A family of estimators for estimating population mean using known correlation coefficient in two-phase sampling*. Statist. Trans. new series, 8, (1), pp. 89-96.

R. TAILOR, B. SHARMA (2009). *A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of kurtosis*. Statist. Trans. new series, 10, (1), pp. 15-24.

R.G.D. STEEL, J.H. TORRIE (1960). *Principles and Procedures of Statistics.* McGraw, New York, USA.

R.S. SOLANKI,H.P. SINGH, A. RATHOUR (2012). *An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys*. ISRN Prob. Statist., DOI: 10.5402/2012/657682.

S. MOHANTY, J. SAHOO (1995). *A note on improving the ratio method of estimation through linear transformation using certain known population parameters*. Sankhya, B, 57, (1), pp. 93-102.

S.K. SRIVASTAVA (1980). *A class of estimators using auxiliary information in sample surveys*. Canad. Jour. Statist., 8, (2), pp. 253-254.

T. SRIVENKATARAMANA, D.S. TRACY (1980). *An alternative to ratio method in sample surveys.* Ann. Inst. Statist. Math., 32, (1), pp. 111-120.

W.G. COCHRAN (1940). *The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce*. Jour. Agri. Sci., 30, pp. 262-275.

SUMMARY

Improved class of estimators of finite population mean using sampling fracttion and information on two auxiliary variables in sample surveys

This paper suggested a generalized class of estimators using information on two auxiliary variables and sampling fraction in simple random sampling. The bias and mean squared error formulae of suggested class have been derived under large sample approximation and compared with usual unbiased estimator and Singh's (1967) ratio-cum-product estimator. The theoretical findings have been satisfied with an empirical study.