Continental J. Applied Sciences 10 (1): 43 - 49, 2015
© Wilolud Journals, 2015
Printed in Nigeria

ISSN: 1597 – 9928
http://www.wiloludjournal.com
doi:10.5707/cjapplsci.2015.10.1.43.49

RESEARCH PAPER

IMPROVED CLASS OF RATIO-TYPE ESTIMATOR UNDER DOUBLE SAMPLING

R. V. K. Singh[1] A. Ahmed [2]and A. Y. Shero[1]
[1]Department of Mathematics, Kebbi State University of Science and Technology, Aliero.
[2] Department of Mathematics Usmanu Danfodiyo University, Sokoto.

ABSTRACT

In this paper we have proposed a class of ratio -type estimator under double sampling to estimate the population mean of the characteristic under study. The expression for bias and the mean square error of the proposed estimator have been derived up to the first order of approximation. The proposed estimator includes a number of ratio -type estimator under double sampling, thus we have discussed its particular cases and their mean square error have also been compared theoretically with the mean square error of the proposed estimators and conditions are found under which the proposed estimator is more efficient. Numerical illustration has been carried out to observe the efficiency of proposed estimator.

KEYWORDS: Bias, Double Sampling, Efficiency, Estimator, Mean square error

Received for Publication: 22/05/15 Accepted for Publication: 25/07/15
Corresponding Author: singhrvk13@gmail.com

INTRODUCTION

In sampling survey, use of the information on auxiliary characteristic to increase the precision of the estimator of finite population mean or total of characteristic under study is well recognized in literature as well as in practice. Cochran (1942) first introduced use of auxiliary information by suggesting classical ratio estimator which perform better when characteristic under study is highly correlated with auxiliary characteristic. Murthy (1964) proposed product estimator, which is utilized to estimate the population mean of characteristic under study contrary to the situation of ratio estimator if characteristic under study is negatively correlated with auxiliary characteristic. Later on, several researchers diverted their attention in the direction of modifying estimation procedure so that estimators that are more efficient and less biased than classical ratio and product estimator could be obtained. Thus, a number of modified ratio type estimator came into existence and these modified estimators were constructed by using one or more unknown constants or by taking a convex linear combination of sample and population means of auxiliary characteristic with unknown weights. Walsh (1970), Ray *et al* (1979), Srivevkataramana and Tracy (1979), Ray and Sahai (1980), Srivevkataramana (1980) and Vos (1980) have done the work in the same direction. They have introduced, generally, a class of estimator with gain in precision than the classical ratio and product estimator.

However, due to the stronger intuitive appeal, statisticians are more inclined towards the use of ratio and product estimators. Singh and Singh (2007) studied the factor type estimators under polynomial regression model which include several estimators and is based on one parameter. Isaki (1983) introduced some estimators for estimation of population variance. All these estimators require the knowledge of population mean of auxiliary characteristic X. But in practical situation, generally it is not known. Therefore to estimate the population mean $\bar{X}$ by using technique of double sampling, first the sample of size n_1 is drawn from the population, on which only the auxiliary variable X is observed. Then a sub-sample of size n ($n < n_1$) is drawn to observe both study variable Y and auxiliary variable X. Let $\bar{x}_1$ be the sample mean based on n_1 units and $\bar{y}$ and $\bar{x}$ be the sample means of based on n units. Then ratio and product estimators of population mean $\bar{Y}$ under double sampling are given by

$$\hat{\bar{Y}}_R^d = \bar{y}\,\frac{\bar{x}_1}{\bar{x}} \text{ and } \hat{\bar{Y}}_P^d = \bar{y}\,\frac{\bar{x}}{\bar{x}_1}$$

In the present paper a class of ratio type estimator under double sampling scheme has been proposed and the expression for bias and mean square error has been derived. Properties of the proposed estimator have been studied theoretically and empirically.

Notation

Let Y_i denotes the value of characteristic under study for the i[th] unit in population of size N (i=1,2…N). and X_i, The value of auxiliary characteristic for the i[th] unit in population. Then

$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$; The population mean of characteristic under study.

$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$; The population mean of auxiliary characteristic.

$S_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$; The population mean square of characteristic under study.

$S_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$; The population mean square of auxiliary characteristic.

Let a preliminary sample of size n_1 is drawn by method of simple random sampling without replacement to estimate the auxiliary characteristic X Then a subsample sample of size n is selected from n_1 to observe both characteristic under study and auxiliary characteristic. Let us further denote

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$; The sample mean of study characteristic based on a sample of size n

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; The sample mean of auxiliary characteristic based on a sample of size n

$\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i$; The sample mean of auxiliary characteristic based on a preliminary sample of size n_1.

$C_Y = \frac{S_Y}{\bar{Y}}$; The coefficient of variation of characteristic under study.

$C_X = \frac{S_X}{\bar{X}}$; The coefficient of variation of auxiliary characteristic.

$\rho = \frac{S_{XY}}{S_X S_Y}$; The Correlation coefficient between the value of auxiliary variable and value of characteristic under study.

Proposed Estimators:

The proposed class of ratio type estimator is as follows:

$$\bar{y}_{Pd} = \bar{y}\left[\frac{\bar{x}_1}{\alpha\bar{x}+(1-\alpha)\bar{x}_1}\right]^K \quad (1)$$

where α and k are constants.

Bias and MSE of Proposed Estimator up to the first order of approximation.

To derive the expression for bias and mean square error of proposed estimator up to the first order of approximation let us denote:

$\bar{y} = \bar{Y}(1+e_0)$, $\bar{x} = \bar{X}(1+e_o')$ and $\bar{x}_1 = \bar{X}(1+e_1)$

Then clearly, we have

$E(e_0) = E(e_o') = E(e_1) = 0$

$E(e_0)^2 = \left(\frac{1}{n} - \frac{1}{N}\right) C_Y^2$

$E(e_o')^2 = \left(\frac{1}{n} - \frac{1}{N}\right) C_X^2$
$E(e_1)^2 = \left(\frac{1}{n_1} - \frac{1}{N}\right) C_X^2$
$E(e_0 e_o') \quad \left(\frac{1}{n} - \frac{1}{N}\right) \frac{S_{XY}}{\bar{X}\bar{Y}}$
$E(e_0 e_1) = \quad \left(\frac{1}{n_1} - \frac{1}{N}\right) \frac{S_{XY}}{\bar{X}\bar{Y}}$
$E(e_o' e_1) = \quad \left(\frac{1}{n_1} - \frac{1}{N}\right) C_X^2$

Now, up to the first order of approximation the expected valve of $\bar{y}_{pd}$ is obtained as

Thus, $\text{Bias}\,(\bar{y}_{Pd}) = \bar{Y}\left(\frac{1}{n} - \frac{1}{n_1}\right)\left(\frac{k\,(k+1)\alpha^2}{2} C_X^2 - k\alpha\rho C_X C_Y\right)$ (2)

$\text{MSE}\,(\bar{y}_{Pd}) = \bar{Y}^2\left(\frac{1}{n} - \frac{1}{n_1}\right)(C_Y^2 + k^2\alpha^2 C_X^2 - 2k\alpha\rho C_X C_Y) + \left(\frac{1}{n_1} - \frac{1}{N}\right) C_Y^2$ (3)

Particular Cases of Proposed Estimator and their Bias and Mean Square Error

For k=0 ,the proposed estimator becomes sample mean estimator with variance

$V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right) S_Y^2$ (4)

For α =1, k=1 and α =1, k=-1 it reduces to classical ratio and product estimator under double sampling respectively. From equation (2) and (3) the bias and MSE respectively obtained as

$\text{Bias}\,(\bar{y}_{rd}) = \bar{Y}\left(\frac{1}{n} - \frac{1}{n_1}\right)(C_X^2 - \rho C_X C_Y)$ (5)

$\text{MSE}\,(\bar{y}_{rd}) = \bar{Y}^2\left(\frac{1}{n} - \frac{1}{n_1}\right)(C_Y^2 + C_X^2 - 2\rho C_X C_Y) + \left(\frac{1}{n_1} - \frac{1}{N}\right) C_Y^2$ (6)

$\text{Bias}\,(\bar{y}_{pd}) = \bar{Y}\left(\frac{1}{n} - \frac{1}{n_1}\right)\rho C_X C_Y$ (7)

$\text{MSE}\,(\bar{y}_{pd}) = \bar{Y}^2\left(\frac{1}{n} - \frac{1}{n_1}\right)(C_Y^2 + C_X^2 + 2\rho C_X C_Y) + \left(\frac{1}{n_1} - \frac{1}{N}\right) C_Y^2$ (8)

For α =w and k=1 the proposed estimator becomes Walsh (1979) estimator under double sampling

$$\bar{y}_{wd} = \bar{y}\left[\frac{\bar{x}_1}{w\bar{x} + (1-w)\bar{x}_1}\right] \quad (9)$$

and the expression for Bias and MSE are

$$\text{Bias}(\bar{y}_{wd}) = \bar{Y}\left(\frac{1}{n} - \frac{1}{N}\right)(w^2 C_X^2 - w\rho C_X C_Y) \quad (10)$$

$$\text{MSE}\,(\bar{y}_{wd}) = \bar{Y}^2\left[\left(\frac{1}{n} - \frac{1}{n_1}\right)(C_Y^2 + w^2 C_X^2 - 2w\rho C_X C_Y) + \left(\frac{1}{n_1} - \frac{1}{N}\right) C_Y^2\right] \quad (11)$$

When k = -1 and $\alpha = -\left(\frac{n}{N-n}\right)$ then proposed estimator $\bar{y}_{Pd}$ reduces to an estimator proposed by Srivenkataramana (1980) under double sampling

$$\bar{y}_{sd} = \bar{y}\left[\frac{N\bar{x}_1 + n\,\bar{x}}{(N-n)\bar{x}_1}\right] \quad (12)$$

The expression for Bias and MSE are

$$\text{Bias}(\bar{y}_{Sd}) = -\bar{Y}\left(\frac{1}{n} - \frac{1}{n_1}\right)\left(\frac{n}{N-n}\right)\rho C_X C_Y \quad (13)$$

$$\text{MSE}\,(\bar{y}_{Sd}) = \bar{Y}^2\left[\left(\frac{1}{n} - \frac{1}{n_1}\right)\left(C_Y^2 + \left(\frac{n}{N-n}\right)^2 C_X^2 - 2\left(\frac{n}{N-n}\right)\rho C_X C_Y\right) + \left(\frac{1}{n_1} - \frac{1}{N}\right) C_Y^2\right] \quad (14)$$

When k = -1 and $\alpha = 1 - w$ then proposed estimator $\bar{y}_p$ reduces to an estimator proposed by Vos (1980)

$$.\bar{y}_{Vd} = \bar{y}\left[\frac{w\bar{x}_1+(1-w)\bar{x}}{\bar{x}_1}\right] \tag{15}$$

The expression for Bias and MSE are

$$\text{Bias}(\bar{y}_{Vd}) = \bar{Y}\left(\frac{1}{n}-\frac{1}{n_1}\right)(1-\text{w})\rho C_X C_Y \tag{16}$$

$$\text{MSE}(\bar{y}_{Vd}) = \bar{Y}^2\left[\left(\frac{1}{n}-\frac{1}{n_1}\right)\left(C_Y^2+(1-\text{w})^2C_X^2+2(1-\text{w})\rho C_X C_Y\right)+\left(\frac{1}{n_1}-\frac{1}{N}\right)C_Y^2\right] \tag{17}$$

When k = -1 and $\alpha = -\gamma$ then proposed estimator $\bar{y}_p$ reduces to an estimator proposed by Ray et al (1979) .

$$\bar{y}_{RSd} = \bar{y}\left[\frac{(1+\gamma)\bar{x}_1-\gamma\bar{x}}{\bar{x}_1}\right] \tag{18}$$

The expression for Bias and MSE are

$$\text{Bias}(\bar{y}_{RSd}) = -\bar{Y}\left(\frac{1}{n}-\frac{1}{n_1}\right)\gamma\rho C_X C_Y \tag{19}$$

$$\text{MSE}(\bar{y}_{RSd}) = \bar{Y}^2\left[\left(\frac{1}{n}-\frac{1}{n_1}\right)\left(C_Y^2+\gamma^2C_X^2-2\gamma\rho C_X C_Y\right)+\left(\frac{1}{n_1}-\frac{1}{N}\right)C_Y^2\right] \tag{20}$$

When k = -1 and $\alpha = 1-f$ then proposed estimator $\bar{y}_p$ reduces to an estimator proposed by Srivenkataramana and Tracy (1979)

$$\bar{y}_{STd} = \bar{y}\left[\frac{f\bar{x}_1+(1-f)\bar{x}}{\bar{x}_1}\right] \tag{21}$$

The expression for Bias and MSE are

$$\text{Bias}(\bar{y}_{STd}) = \bar{Y}\left(\frac{1}{n}-\frac{1}{n_1}\right)(1-f)\rho C_X C_Y \tag{22}$$

$$\text{MSE}(\bar{y}_{STd}) = \bar{Y}^2\left[\left(\frac{1}{n}-\frac{1}{n_1}\right)\left({C_Y^{\bar{y}_{STd}}}^2+(1-f)^2C_X^2+2(1-f)\rho C_X C_Y\right)+\left(\frac{1}{n_1}-\frac{1}{N}\right)C_Y^2\right] \tag{23}$$

Optimum value of k

From mean square error, the optimum value of k is obtained as

$$\text{k} = \frac{\rho C_Y}{\alpha C_X} \tag{24}$$

For optimum value of k the mean square error of proposed estimator reduces to

$$\text{MSE}(\bar{y}_{Pd})_{min} = \bar{Y}^2\left[\left(\frac{1}{n}-\frac{1}{n_1}\right)C_Y^2(1-\rho^2)+\left(\frac{1}{n_1}-\frac{1}{N}\right)C_Y^2\right] \tag{25}$$

This is the mean square error of linear regression estimator under double sampling.

Comparison of proposed Estimator with Existing Estimators

In this section, condition has been derived under which proposed class of estimator is more efficient than the exiting ratio-type estimators.

First proposed class of estimator is compared with Walsh (1970) estimator

MSE $(\bar{y}_{wd})$ - MSE $(\bar{y}_{Pd}) \geq 0$ if

$$\min\left(1, \frac{2\rho C_Y}{WC_X}-1\right) \leq K \leq \max\left(1, \frac{2\rho C_Y}{WC_X}-1\right) \tag{26}$$

Thus Proposed class of estimator is more efficient than Walsh (1970) estimator if condition (26) is satisfied.

Proposed class of estimator is compared with Ray et al (1979) estimator
MSE ($\bar{y}_{RSd}$) - MSE ($\bar{y}_{Pd}$) ≥ 0 if

$$\min\left(1, \frac{2\rho C_Y}{\gamma C_X} - 1\right) \leq K \leq \max\left(1, \frac{2\rho C_Y}{\gamma C_X} - 1\right) \tag{27}$$

Thus Proposed class of estimator is more efficient than Ray et al (1979) estimator if condition (27) is satisfied.

Proposed class of estimator is compared with Srivenkataramana and Tracy (1979) estimator
MSE ($\bar{y}_{STd}$) − MSE ($\bar{y}_{Pd}$) ≥ 0 if

$$\min\left(1, -\left(\frac{2\rho C_Y}{\left(1-\frac{n}{N}\right)C_X} + 1\right)\right) \leq K \leq \max\left(1, -\left(\frac{2\rho C_Y}{\left(1-\frac{n}{N}\right)C_X} + 1\right)\right) \tag{28}$$

Thus Proposed class of estimator is more efficient than Srivenkataramana and Tracy (1979) estimator if condition (28) is satisfied.

Proposed class of estimator is compared with Srivenkataramana (1980) duel to ratio estimator
MSE ($\bar{y}_{Sd}$) − MSE ($\bar{y}_{Pd}$) ≥ 0 if

$$\min\left(1, \left(\frac{2\rho C_Y}{\left(\frac{n}{N-n}\right)C_X} - 1\right)\right) \leq K \leq \max\left(1, \left(\frac{2\rho C_Y}{\left(\frac{n}{N-n}\right)C_X} - 1\right)\right) \tag{29}$$

Thus Proposed class of estimator is more efficient than Srivenkataramana (1980) estimator if condition (29) is satisfied.

Proposed class of estimator is compared with Vos (1980) estimator
MSE ($\bar{y}_{Vd}$) - MSE ($\bar{y}_{Pd}$) ≥ 0 if

$$\min\left(1, -\left(\frac{2\rho C_Y}{(1-W)C_X} + 1\right)\right) \leq K \leq \max\left(1, -\left(\frac{2\rho C_Y}{(1-W)C_X} + 1\right)\right) \tag{30}$$

Thus Proposed class of estimator is more efficient than Vos (1980) estimator if condition (30) is satisfied.

Empirical Comparison

In order to investigate the performance of proposed class of estimator from Hina Khan *et al* (2012) two real population data sets have been considered and description of populations have been given in Table 1. Population I has high positive correlation whereas population II has low but positive while population III has high negative correlation between auxiliary variable and study variable. The percent relative efficiency of proposed class of estimator has been computed with respect to sample mean estimator for different value of α and K and given in Table 2.

Table 1 Description of the Population

Parameters	Population I Source: Nachtshemin,neter and Kumar Advance Applied Linear Model,2004	Population I Source: Nachtshemin,neter and Kumar Advance Applied Linear Model,2004
N	440	705
n_1	88	141
n	18	28
$\bar{Y}$	119.50	2.9773
$\bar{X}$	906.76	76.95
C_Y	1.9955	0.213123
C_X	1.7501	0.242157
ρ_{XY}	0.956	0.398

Table 2: Population I

k \ α	-0.1	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
-0.1	404.58	338.22	253.04	183.23	133.80	100	76,71	60.30	48.45	39.67	33.02
-0.8	338.22	269.39	208.59	161.21	125.98	100	80.72	66.21	55.11	46.48	39.67
-0.6	253.04	208.59	171.82	142.26	118.72	100	85.03	72.97	63.16	55.11	48.85
-0.4	183.23	161.21	142.26	125.98	112.01	100	89.66	80.72	72.97	66.21	60.30
-0.2	133.80	125.98	118.72	112.01	105.78	100	94.64	89.66	85.03	80.72	76.71
0.0	100	100	100	100	100	100	100	100	100	100	100
0.2	76,71	80.72	85.03	89.66	94.64	100	105.78	112.01	118.72	125.98	133.80
0.4	60.30	66.21	72.97	80.72	89.66	100	112.01	125.98	142.26	161.21	183.23
0.6	48.45	55.11	63.16	72.97	85.03	100	118.72	142.26	171.82	208.59	253.04
0.8	39.67	46.48	55.11	66.21	80.72	100	125.98	161.21	208.59	269.39	338.22
1.0	33.02	39.67	48.85	60.30	76.71	100	133.80	183.23	253.04	338.22	404.58
1.2	27.87	34..21	42.88	55.11	72.97	100	142.26	208.59	303.63	395.59	400.47
1.4	23.82	29.78	38.19	50.53	69.47	100	151.38	237.40	354.76	412.26	329.73
1.6	20.58	26.14	34.21	46.48	66.21	100	161.22	269.39	395.59	377.34	245.16
1.8	17.95	23.12	30.81	42.88	63.16	100	171.82	303.63	413.12	312.38	177.45
2.0	15.79	20.58	27.87	39.67	60.30	100	183.23	338.22	400.47	245.16	129.83

Population II

k \ α	-0.1	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
-0.1	62.53	76.51	90.83	101.93	105.55	100	87.82	73.33	69.71	48.26	39.13
-0.8	76.51	88.09	98.17	104.16	105.15	100	90.57	79.15	67.67	57.23	48.26
-0.6	90.83	98.17	103.36	105.57	104.37	100	93.21	84.97	76.23	67.67	59.71
-0.4	101.93	104.16	105.57	105.15	103.24	100	95.68	90.57	84.97	79.15	73.33
-0.2	105.55	105.15	104.37	103.24	101.77	100	97.96	95.68	93.21	90.57	87.82
0.0	100	100	100	100	100	100	100	100	100	100	100
0.2	87.82	90.57	93.21	95.68	97.96	100	101.77	103.24	104.37	105.15	105.55
0.4	73.33	79.15	84.97	90.57	95.68	100	103.24	105.15	105.57	104.16	101.93
0.6	69.71	67.67	76.23	84.97	93.21	100	104.37	105.57	103.36	98.17	90.83
0.8	48.26	57.23	67.67	79.15	90.57	100	105.15	104.16	98.17	88.09	76.51
1.0	39.13	48.26	59.71	73.33	87.82	100	105.55	101.93	90.83	76.51	62.53
1.2	32-00	40.78	52-55	67-67	84.97	100	105.57	98.17	82.36	65.20	50.57
1.4	26.45	34.64	46.26	62.28	82.08	100	105.20	93.45	73.61	55.07	40.95
1.6	22.11	29.61	40.78	57.23	79.15	100	104.46	88.09	65.20	46.44	33.41
1.8	18.69	25.49	36.06	52.55	76.23	100	103.36	82.36	57.46	39.29	27.55
2.0	15.96	22.11	32.00	48.26	73.33	100	101.93	76.51	50.57	33.41	22.97

CONCLUSION

The proposed class of ratio type estimator provides different ratio and product type estimator under double sampling scheme for different valve of k and α which can be utilized according to the nature of the population to estimate the population mean. Table 2 show that the efficiency of $\bar{y}_{wd}$ increases from 33.02% to 404.58% as w increases from -1.0 to1.0. Similar way the efficiency of $\bar{y}_{Vd}$ and $\bar{y}_{RSd}$ also increases from 33.02% to 404.58% as w increases from 0 to2.0 and γ decreases from 1.0 to -1.0 respectively. $\bar{y}_{Sd}$ has also gain in efficiency than Sample mean estimator. For k=1.8 and α =0.6 the proposed estimator $\bar{y}_{pd}$ has gain in efficiency of 413.12% .

REFERENCES

Cochran,W.J. (1942): The estimation of the yields of the cereals experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*. 30,262-275.

Hina Khan, Saleha, S.,Aamir,S.(2012): A Class of Improved Estimators for Estimating Population Mean Regarding Partial Information in Double Sampling. *Global Journal* Inc. , Vol. 12

Isaki, C. T. (1983): Variance estimation using auxiliary information, *Journal of the American Statistical Association*, 78, pp 117–123.

Murthy, M. N. (1964): Product method of estimation. *Sankhya* 26A, pp 69 – 74.

Ray, S. K., Sahai, A. (1980): Efficient families of ratio and product estimators. *Biomerica*, 67,215-217.

Ray,S.K.,Sahai, A. and Sahai, A (1979). A note on ratio and product estimators. *Annals of the Institute of Mathematical Statistics.* 31,141- 144

Singh, R. V. K., Singh, B. K. (2007): Study of a Class of Ratio Type Estimator under Polynomial Regression Model. *Proceedings of Mathematical Society*, B.H.U, Vol. 23.

Srivenkataramana, T. (1980): A dual to ratio estimators in sample surveys. *Biometrica* 67(1), pp 199 – 204.

Srivenkataramana, T. and Tracy, D.S. (1979): On ratio and product methods of estimation in sampling. *Statistica Neerlandica,* 33 pp 47-49.

Vos, J. W. E (1980): Mixing of direct ratio and product estimators. S*tatistica Neerlandica*, 34, 209-218.

Walsh, J. K. (1970): Generalization of ratio estimator for population total. *Sankhya*, 32A, 99-106.