A FAMILY OF PRODUCT- CUM- DUAL TO RATIO ESTIMATORS OF FINITE POPULATION MEAN IN SIMPLE RANDOM SAMPLING

DIGANTA KALITA

Department of Statistics,
North Lakhimpur College (Autonomous), North Lakhimpur, Assam, India-787031

ABSTRACT

This paper presents a family of product- cum- dual to ratio estimator for estimating finite population mean of the variable under study. The bias and mean square error (MSE) of the proposed estimator is obtained to the first degree approximation under simple random sampling without replacement (SRSWOR) scheme. The asymptotically optimum estimators (AOEs) are recognised with its bias and mean square error. A comparison has been made with some existing estimators viz.Sample mean per unit estimator, usual ratio estimator Cochran, product estimator Robson & Murthy, dual to ratio estimator Srivenkataraman and dual to product estimator Bandyopadhya. The proposed estimators are found to be more efficient theoretically and numerically.

Keywords: Auxiliary variable, product-cum-dual to ratio estimator, Bias, MSE and Efficiency.

1. INTRODUCTION:

In sample survey theory, it is seen that the use of auxiliary information in sample survey increases the precision of the estimate of population mean of study variate. Consider a simple random sample of size n, which is drawn by without replacement from a finite population of size N. Let Y_i and X_i denote the values of the study and auxiliary variables respectively for the ith unit $(i = 1, 2, 3, \dots N)$ of the population. Let $\bar{y} = \sum_{i=1}^{n} y_i / n$ and $\bar{x} = \sum_{i=1}^{n} x_i / n$ be the sample means of the study and auxiliary variable y and x respectively.

When the correlation between the study variable y and auxiliary variable x is highly positive, Cochran (1940) used auxiliary information and proposed the usual ratio estimator for estimating population mean $\bar{Y} = \sum_{i=1}^{N} Y_i / N$ of y as $\bar{y}_R = \frac{\bar{y}}{\bar{x}} \bar{X}$,where $\bar{X} = \sum_{i=1}^{N} X_i / N$ is known. When the correlation is highly negative between y and x, with known population mean $\bar{X}$, Robson (1957) and Murthy (1964) worked independently and proposed product estimator as $\bar{y}_P = \frac{\bar{y}}{\bar{X}}\bar{x}$.

Using the transformation

$x_i^* = (N\bar{X} - nx_i)/(N-n)$,$(i = 1,2,3...,N)$ or $x_i^* = (1+g)\bar{X} - gx_i, (i = 1,2,3...,N)$,

Srivenkataramana (1980) and Bandopadhya (1980), suggested dual to ratio and dual to product estimator as:

$\bar{y}_R^* = \bar{y}\left(\frac{\bar{x}^*}{\bar{X}}\right)$ and $\bar{y}_P^* = \left(\frac{\bar{y}}{\bar{x}^*}\right)\bar{X}$ respectively, where $\bar{x}^* = (N\bar{X} - n\bar{x})/(N-n)$ and $g = n/(N-n)$

In this paper, we have proposed an estimator of combination of product estimator and dual to ratio estimator for estimating population mean $\bar{Y}$ for its efficiency over other estimators.

2. PROPOSED ESTIMATOR:

Based on the estimators $\bar{y}_p$ and $\bar{y}_R^*$, we proposed the following estimator as,

$$T = \bar{y}\left\{\alpha\left(\frac{a\bar{x}+b}{a\bar{X}+b}\right) + \beta\left(\frac{a\bar{x}^*+b}{a\bar{X}+b}\right)\right\} \qquad (1)$$

where $a(\neq 0), b$ and α & β are suitably chosen as constant such that $\alpha + \beta = 1$

Remark:

(i) if $(\alpha,\beta) = (1,0)$ & $(a,b) = (1,0)$ then the estimator T reduces to the usual product estimator $\bar{y}_p$ and its properties.

(ii) if $(\alpha,\beta) = (0,1)$ & $(a,b) = (1,0)$ then the estimator T reduces to the dual to ratio estimator $\bar{y}_R^*$ and its properties.

3 BIAS AND MSE OFPROPOSED ESTIMATOR 'T':

To obtain the bias and MSE of T to a first degree of approximation, we write

$\bar{y} = \bar{Y}(1+e_0)$, $\bar{x} = \bar{X}(1+e_1)$ such that

$$\left.\begin{aligned} &E(e_i) = 0, i = 0,1 \\ &E(e_0^2) = \frac{1-f}{n}C_y^2,\ E(e_1^2) = \frac{1-f}{n}C_x^2, E(e_0e_1) = \frac{1-f}{n}\rho_{yx}C_yC_x \end{aligned}\right\} \qquad (2)$$

where $f=\dfrac{n}{N}$, $C_y^2=S_y^2/\bar{Y}^2$, $C_x^2=S_x^2/\bar{X}^2$, $K_{yx}=\rho_{yx}\dfrac{C_y}{C_x}$, $S_y^2=\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^2/(N-1)$,

$S_x^2=\sum_{i=1}^{N}\left(X_i-\bar{X}\right)^2/\left(N-1\right)$, $S_{xy}=\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)\left(X_i-\bar{X}\right)/\left(N-1\right)$ and $\rho=S_{xy}/S_xS_y$.

Expanding the right hand side of (1) in terms of e_i 's $(i=0,1)$, we have

$$T\cong\bar{Y}\left(1+e_0\right)\left\{\alpha\left(1+\varphi e_1\right)+\left(1-\alpha\right)\left(1-\varphi e_1\right)\right\} \quad (3)$$

To the first degree approximation, the bias and MSE of T are respectively obtained as follows

$$B(T)=\frac{1-f}{n}\bar{Y}\lambda K_{yx}C_x^2. \quad (4)$$

$$M(T)=\bar{Y}^2\frac{1-f}{n}\left\{C_y^2+\lambda C_x^2\left(\lambda+2K_{yx}\right)\right\} \quad (5)$$

Where $\varphi=a\bar{X}/(a\bar{X}+b)$ and $\lambda=\varphi\alpha(1+g)-g$

Differentiation of equation (5) with respect to α and equating it to zero, we get optimum value of α as

$$\alpha=\alpha_{opt.}=(g-K_{yx})/\varphi(1+g) \quad (6)$$

Substituting (6) in (1) we get the asymptotically optimum estimator (AOE) for $\bar{Y}$ as

$$T^{opt.}=\bar{y}\left[\left\{\frac{g-K_{yx}}{\varphi(1+g)}\right\}\left(\frac{a\bar{x}+b}{a\bar{X}+b}\right)+\left\{1-\frac{g-K_{yx}}{\varphi(1+g)}\right\}\left(\frac{a\bar{x}^*+b}{a\bar{X}+b}\right)\right]$$

Therefore the resulting bias and the MSE of $T^{opt.}$ respectively as

$$B(T^{opt.})=-\frac{1-f}{n}\bar{Y}K_{yx}^2C_x^2$$

$$M(T^{opt.})=\bar{Y}^2\frac{1-f}{n}C_y^2\left(1-\rho_{yx}^2\right) \quad (7)$$

The MSE equation in (7) is same as the usual linear regression estimator $\bar{Y}_{lreg.}=\bar{y}+b_{yx}\left(\bar{X}-\bar{x}\right)$, where b_{yx} is the sample regression coefficient y on x.

To the first degree approximation, the MSE of $\bar{y}_R,\bar{y}_P,\bar{y}_R^*,\bar{y}_P^*$ are

$$M(\bar{y}_R)=\bar{Y}^2\frac{1-f}{n}\left\{C_y^2+C_x^2(1-2K_{yx}\right\} \quad (8)$$

$$M(\bar{y}_P)=\bar{Y}^2\frac{1-f}{n}\left\{C_y^2+C_x^2(1+2K_{yx}\right\} \tag{9}$$

$$M(\bar{y}_R^*)=\bar{Y}^2\frac{1-f}{n}\left\{C_y^2+gC_x^2(g-2K_{yx}\right\} \tag{10}$$

$$M(\bar{y}_P^*)=\bar{Y}^2\frac{1-f}{n}\left\{C_y^2+gC_x^2(g+2K_{yx}\right\} \tag{11}$$

respectively.

The MSE of usual unbiased estimator $\bar{y}$ under SRSWOR scheme is

$$M\left(\bar{y}\right)=\bar{Y}^2\frac{1-f}{n}C_y^2 \tag{12}$$

4 EFFICIENCY COMPARISONS

(A) Comparison of T

In this sub-section, we have presented the comparisons of proposed estimator with other estimators to investigate the ranges of α for which the proposed estimator is better than the others.

Now from the equations (5) and (12), we observe that the proposed estimator T is more efficient than usual unbiased estimator $\bar{y}$, under the condition

either $\frac{g-2K_{yx}}{\varphi(1+g)}<\alpha<\frac{g}{\varphi(1+g)}$ or $\frac{g}{\varphi(1+g)}<\alpha<\frac{g-2K_{yx}}{\varphi(1+g)}$.

Therefore, the ranges of α under which the proposed estimator T is more efficient than the sample mean $\bar{y}$ is $\left\{\min\left(\frac{g}{\varphi(1+g)},\frac{g-2K_{yx}}{\varphi(1+g)}\right),\max\left(\frac{g}{\varphi(1+g)},\frac{g-2K_{yx}}{\varphi(1+g)}\right)\right\}$

From the equations (5) and (8), we observe that the proposed estimator T is more efficient than usual ratio estimator $\bar{y}_R$, if

either $\frac{g-1}{\varphi(1+g)}<\alpha<\frac{(1+g)-2K_{yx}}{\varphi(1+g)}$ or $\frac{(1+g)-2K_{yx}}{\varphi(1+g)}<\alpha<\frac{g-1}{\varphi(1+g)}$.

Therefore, the ranges of α under which the proposed estimator T is more efficient than the usual ratio estimator $\bar{y}_R$ is $\left\{\min\left(\frac{(1+g)-2K_{yx}}{\varphi(1+g)},\frac{g-1}{\varphi(1+g)}\right),\max\left(\frac{(1+g)-2K_{yx}}{\varphi(1+g)},\frac{g-1}{\varphi(1+g)}\right)\right\}$

From the equations (5) and (9), we observe that the proposed estimator T is more efficient than usual product estimator $\bar{y}_P$, if

Either $\frac{g-(1+2K_{yx})}{\varphi(1+g)} < \alpha < \frac{1+g}{\varphi(1+g)}$ or $\frac{1+g}{\varphi(1+g)} < \alpha < \frac{g-(1+2K_{yx})}{\varphi(1+g)}$.

Therefore, the ranges of α under which the proposed estimator T is more efficient than the product estimator $\bar{y}_P$, is $\left[\min\left\{\frac{g-(1+2K_{yx})}{\varphi(1+g)}, \frac{1+g}{\varphi(1+g)}\right\}, \max\left\{\frac{g-(1+2K_{yx})}{\varphi(1+g)}, \frac{1+g}{\varphi(1+g)}\right\}\right]$.

From the equations (5) and (10), we observe that the proposed estimator T is more efficient than dual to ratio estimator $\bar{y}_R^*$, if

either $0 < \alpha < \frac{2(g-K_{yx})}{\varphi(1+g)}$ or $\frac{2(g-K_{yx})}{\varphi(1+g)} < \alpha < 0$.

Therefore, the ranges of α under which the proposed estimator T is more efficient than the dual to ratio estimator $\bar{y}_R^*$, is

$$\left[\min\left\{0, \frac{2(g-K_{yx})}{\varphi(1+g)}\right\}, \max\left\{0, \frac{2(g-K_{yx})}{\varphi(1+g)}\right\}\right].$$

From the equations (5) and (11), we observe that the proposed estimator T is more efficient than dual to product estimator $\bar{y}_P^*$, if

either $-\frac{2K_{yx}}{\varphi(1+g)} < \alpha < \frac{2g}{\varphi(1+g)}$ or $\frac{2g}{\varphi(1+g)} < \alpha < -\frac{2K_{yx}}{\varphi(1+g)}$.

Therefore, the ranges of α under which the proposed estimator T is more efficient than the dual to product estimator $\bar{y}_P^*$, is

$$\left[\min\left\{-\frac{2K_{yx}}{\varphi(1+g)}, \frac{2g}{\varphi(1+g)}\right\}, \max\left\{-\frac{2K_{yx}}{\varphi(1+g)}, \frac{2g}{\varphi(1+g)}\right\}\right].$$

(B) Comparison of $T^{opt.}$

From the equations (7), (8), (9),(10), (11),and (12), it is found that the AOE $T^{opt.}$ is more efficient than the estimators $\bar{y}_R, \bar{y}_P, \bar{y}_R^*, \bar{y}_P^*$ and $\bar{y}$,since

$$V(\bar{y}_R) - M(T^{opt.}) = \bar{Y}^2 \frac{1-f}{n} C_x^2 \left(1-K_{yx}\right)^2 > 0$$

$$V(\bar{y}_P)-M(T^{opt.})=\bar{Y}^2\frac{1-f}{n}C_x^2\left(1+K_{yx}\right)^2>0$$

$$V(\bar{y}_R^*)-M(T^{opt.})=\bar{Y}^2\frac{1-f}{n}C_x^2\left(g-K_{yx}\right)^2>0$$

$$V(\bar{y}_P^*)-M(T^{opt.})=\bar{Y}^2\frac{1-f}{n}C_x^2\left(g+K_{yx}\right)^2>0 \text{ and}$$

$$V(\bar{y})-M(T^{opt.})=\bar{Y}^2\frac{1-f}{n}C_x^2K_{yx}^2>0$$

Hence, the proposed estimator 'T'is better than the other estimators in case of its optimality.

EMPERICAL STUDY

To analyze the performance of the proposed class of estimators T (or $T_{(opt.)}$) over other estimators, eight natural population data sets have been taken into consideration. The sources of the population, the nature of variates *y* and *x*; and the values of the various parameters are shown as follows.

Population I: [Source: Steel and Torrie, 1960, p. 282]

x: Chlorine percentage

y: Log of leaf burn in secs

$N=30,\ n=6,\ \bar{Y}=0.6860,\ \bar{X}=0.8077,\ \rho_{yx}=-0.4996,\ C_y=0.700123,\ C_x=0.7493$

Population II: [Source: Pandey and Dubey, 1988]

$N=20,\ n=8,\ \bar{Y}=19.55,\ \bar{X}=18.8,\ \rho_{yx}=-0.9199,\ C_y=0.3552,\ C_x=0.3943$

Population III: [Source: Kadilar and Cingi, 2006a, p. 78]

$N=106,\ n=20,\ \bar{Y}=2212.59,\ \bar{X}=27421.70,\ \rho_{yx}=0.86,\ C_y=5.22,\ C_x=2.10$

Population IV: [Source: Kadilar and Cingi, 2006b, p. 1054]

$N=106,\ n=20,\ \bar{Y}=15.37,\ \bar{X}=243.76,\ \rho_{yx}=0.82,\ C_y=4.18,\ C_x=2.02$

Population V: [Source: Sukhatme and Sukhatme, 1970, p. 256]

x: A circle consisting more than five villages

y: Number of villages in the circles.

$N=89,\ n=12,\ \bar{Y}=3.36,\ \bar{X}=0.1236,\ \rho_{yx}=0.766,\ C_y=0.60400,\ C_x=2.19012$

Population VI: [Source: Maddala, 1977]

x: Deflated prices of veal

y: Consumption per capita.

$N=30,\ n=6,\ \bar{Y}=7.6375,\ \bar{X}=75.4313,\ \rho_{yx}=-0.6823,\ C_y=0.2278,\ C_x=0.0986$

Population VII: [Source: Murthy, 1967, p. 228]

x : Fixed capital

y : Output

$$N = 80, \quad n = 20, \quad \bar{Y} = 51.8264, \quad \bar{X} = 11.2646, \quad \rho_{yx} = 0.9413, \quad C_y = 0.3542, \quad C_x = 0.7507$$

Population VIII: [Source: Murthy, 1967, p. 228]

x : Number of workers

y : Output

$$N = 80, \quad n = 20, \quad \bar{Y} = 51.8264, \quad \bar{X} = 2.8513, \quad \rho_{yx} = 0.9150, \quad C_y = 0.3542, \quad C_x = 0.9484$$

To observe the relative performance of different estimators of $\bar{Y}$, we have computed the percentage relative efficiencies (PREs) of these estimators with respect to $\bar{y}$ by the formula

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

Table 1: Percentage relative efficiencies of different estimators with respect to $\bar{y}$

Population	$\bar{y}$	$\bar{y}_R$	$\bar{y}_P$	$\bar{y}_R^*$	$\bar{y}_P^*$	T or $T^{opt.}$
I	100	*	*	*	124.34	133.26
II	100	*	526.50	*	537.23	650.26
III	100	212.82	*	117.95	*	384.02
IV	100	*	*	220.46	*	241.99
V	100	*	167.59	*	115.73	187.10
VI	100	*	*	591.38	*	877.54
VII	100	*	*	612.44	*	614.34
VIII	100	226.76	*	120.73	*	305.25

*percentage relative efficiency less than 100

CONCLUSION:

Table 1 clearly indicates that the percentage relative efficiencies (PRE) of proposed estimator $(T \text{ or } T^{opt.})$ are higher than all other estimators considered in this paper. Therefore we may conclude that the proposed product cum dual to ratio estimator $(T \text{ or } T^{opt.})$ is more efficient than the estimators $\bar{y}$ $\bar{y}_R, \bar{y}_P, \bar{y}_R^*, \bar{y}_P^*$ .So the proposed estimator $(T \text{ or } T^{opt.})$ is preferable in practice.

REFERENCES

1. 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(2), 262–275
2. Kadilar, C. and Cingi, H. (2006a). Improvement in estimating the Population mean in Simple random sampling. Applied Mathematics Letters, 19, 75-79.
3. Maddala, G. S. (1977). Econometrics, Mcgraw Hills Pub.Co., New York.
4. Murthy, M. N. (1964). Product method of estimation. Sankhya A, 26, 69-74.
5. Murthy, M. N. (1967). Sampling Theory and Methods, Statistical Publishing Society, Calcutta.
6. Pandey, B.N. and Dubey, V. (1988). Modified product estimator using coefficient of variation of auxiliary variable. Assam Stat. Rev., 2, 64-66.
7. Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimation. Journal of the American Statistical Association, 52, 511–522.
8. S. Bandyopadhya,(1980) Improved Ratio and and product estimators, Sankhya Series C,42(2),15-24.
9. Singh, Rajeh and Kumar, Mukesh.(2011). A General family of Dual to Ratio- cum-Dual to Product estimator in Sample Survey, Statistics in Transition-new series, 12(3), 587-594
10. Srivenkataramana, T. (1980). A dual to ratio estimator in sample surveys. Biometrika, 67(1), 199-204.
11. Steel, R. G. D. and Torrie, J. H. (1960). Principles and Procedures of Statistics, McGraw Hill Book.
12. Sukhatme, P. V. and Sukhatme, B. V. (1970). Sampling Theory of Surveys with Applications, Iowa State University Press, Ames, U. S. A.
13. Tailor R. and Sharma, B. K. (2009). A modified ratio-cum-product estimator of finite population mean using known coefficient of Variation and coefficient of Kurtosis. Statistics in Transition-new series, 10(1), 15-24.
14. Yadav, S.K.,Misra, Sheela.(2011). Improved family of Ratio-cum-Dual to Ratio estimator of finite population Mean. International Journal of Computing, 1(1),17-23.
15. Sarmah,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, 10(1), 27-31.