A New Ratio-Cum-Dual to Ratio Estimator of Finite Population Mean in Simple Random Sampling

GJMBR Classification (FOR) 010402,010202,010206 & 010207

Balkishan Sharma[1] Rajesh Tailor[2]

Abstract: **This paper proposes a ratio-cum-dual to ratio estimator of finite population mean. The bias and mean squared error of the proposed estimator are obtained. It has been shown that the proposed estimator is more efficient than the simple mean estimator, usual ratio estimator and dual to ratio estimator under certain given conditions. An Asymptotic optimum estimator in the class of estimator is identified with its mean squared error formula. To judge the merits of the proposed estimator over other estimators an empirical study is carried out.**

Keywords: Finite Population Mean, Dual to Ratio estimator, Bias and Mean squared error.

I. Introduction

In sample surveys, auxiliary information is used at both selections as well as estimation stages to improve the efficiency of the estimators. When the correlation between study variate and auxiliary variate is positive (high), the ratio method of estimation is used for estimating the population mean. On the other hand, if the correlation is negative, the product method of estimation is used. Cochran (1940) used auxiliary information at estimation stage and proposed ratio estimator. Robson (1957) and Murthy (1964) envisaged product estimator, Searls (1964) used coefficient of variation of study variate, motivated by Searls (1964), Sisodia and Dwivedi (1981) utilized coefficient of variation of auxiliary variate. Srivenkataramana (1980), first proposed dual to ratio estimator. Singh and Tailor (2005) and Tailor and Sharma (2009) worked on ratio-cum-product estimators. These motivates author to propose a new ratio-cum–dual to ratio estimator utilizing dual to ratio estimator of finite population mean. Let $U = (U_1, U_2 ..., U_N)$ be a finite population of size N and y and x be the study and auxiliary variates respectively. A sample of size n is drawn using simple random sampling without replacement to estimate the population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ of study variate y.

The classical ratio estimator for $\bar{Y}$ is given by

$$\bar{y}_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \quad (1.1)$$

where $\bar{x} = \sum_{i=1}^{n} x_i / n$ and $\bar{y} = \sum_{i=1}^{n} y_i / n$. Here it is assumed that $\bar{X} = \sum_{i=1}^{N} x_i / N$ and population mean of auxiliary variate is known.

Using the transformation $\bar{x}_i^* = (N\bar{X} - nx_i)/(N-n)$, (i=1,2,3,….,N), Srivenkataramana (1980) obtained dual to ratio estimator as

$$\bar{y}_R^{(d)} = \bar{y}\left(\frac{\bar{x}^*}{\bar{X}}\right) \quad (1.2)$$

where $\bar{x}^* = (N\bar{X} - n\bar{x})/(N-n)$

Bias and mean squared error of ratio estimator $\bar{y}_R$ and dual to ratio estimator $\bar{y}_R^{(d)}$ are respectively given as

$$B(\bar{y}_R) = \bar{Y}\frac{(1-f)}{n}\left[C_x^2 - \rho C_y C_x\right] \quad (1.3)$$

$$B(\bar{y}_R^{(d)}) = \frac{(1-f)}{n}\bar{Y} g C_x^2 (1-K) \quad (1.4)$$

$$MSE(\bar{y}_R) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2(1-2K)\right] \quad (1.5)$$

$$MSE(\bar{y}_R^{(d)}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + gC_x^2(g-2K)\right] \quad (1.6)$$

Where C_y, C_x and ρ are coefficient of variation of y, coefficient of variation of x and correlation coefficient between y and x respectively and defined as

About[1]-Sri Aurobindo Institute Of Medical Sciences, Indore, (Email- E-Mail-Bksnew@Rediffmail.Com)
About[2]-School Of Studies In Statistics, Vikram University, Ujjain- (E-Mail-Tailorraj@Gmail.Co)

$$C_y = \frac{S_y}{\bar{Y}}\; C_x = \frac{S_x}{\bar{X}}, K = \frac{\rho C_y}{C_x}\;\; \rho = \frac{Cov(y,x)}{\sqrt{V(x)V(y)}} = \frac{S_{yx}}{S_x S_y},$$

$$S_x^2 = \sum_{i=1}^{N}(x_i - \bar{X})^2/(N-1),\; S_y^2 = \sum_{i=1}^{N}(y_i - \bar{Y})^2/(N-1) \text{ and } S_{yx} = \sum_{i=1}^{N}(y_i - \bar{Y})(x_i - \bar{X})/(N-1).$$

In this paper we have suggested ratio-cum-dual to ratio type estimator.

II. Proposed Ratio-Cum-Dual To Ratio Estimator

The proposed ratio-cum-dual to ratio estimator of population mean $\bar{Y}$ is

$$\hat{\bar{Y}}_{bk1} = \bar{y}\left[\alpha\left(\frac{\bar{X}}{\bar{x}}\right) + (1-\alpha)\left(\frac{\bar{x}^*}{\bar{X}}\right)\right] \qquad (2.1)$$

where α is a suitably chosen scalar. It is to be noted that for α =1 and α = 0, $\hat{\bar{Y}}_{bk1}$ reduces to the estimators $\bar{y}_R$ and $\bar{y}_R^{(d)}$ respectively. Thus $\bar{y}_R$ and $\bar{y}_R^{(d)}$ are particular case of proposed estimator $\hat{\bar{Y}}_{bk1}$.

To obtain the bias and MSE of $\hat{\bar{Y}}_{bk1}$, we write

$\bar{y} = \bar{Y}(1+e_0)$ and $\bar{x} = \bar{X}(1+e_1)$ suchthat

$E(e_0) = E(e_1) = 0$ and

$E(e_0^2) = \left(\frac{1}{n} - \frac{1}{N}\right)C_y^2, E(e_1^2) = \left(\frac{1}{n} - \frac{1}{N}\right)C_x^2$ and

$$E(e_0 e_1) = \left(\frac{1}{n} - \frac{1}{N}\right)\rho\, C_y\, C_x$$

Expressing (2.1) in terms of $e_i^{'}s$ we have (i=1,2)

$$\hat{\bar{Y}}_{bk1} = \bar{Y}(1+e_0)\left[\alpha(1+e_1)^{-1} + (1-\alpha)(1-g\,e_1)\right] ..(2.2)$$

To the first degree of approximation, the bias and mean squared error of $\hat{\bar{Y}}_{bk1}$ are respectively obtained as

$$B(\hat{\bar{Y}}_{bk1}) = \frac{(1-f)}{n}\bar{Y}C_x^2\left[\alpha - \{(1-g)\alpha + g\}K\right] \quad (2.3)$$

and

$$MSE(\hat{\bar{Y}}_{bk1}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + \{(1-g)\alpha) + g\}^2 C_x^2\{(1-g)\alpha + g - 2K\}\right]$$

………………………………………………………. (2.4)

Thus the estimator $\hat{\bar{Y}}_{bk1}$ with $\alpha = gK/(gK + K - 1)$ is almost unbiased. It is also observed from (2.3) that the bias of $\hat{\bar{Y}}_{bk1}$ is negligible for large sample.

Mean squared error of $\hat{\bar{Y}}_{bk1}$ in (2.4) is minimized for

$$\alpha = \frac{K-g}{1-g} = \alpha_0 \text{ (say)} \qquad (2.5).$$

Substitution of (2.5) in (2.1) yields the asymptotically optimum estimator (AOE) for $\bar{Y}$ as

$$\hat{\bar{Y}}_{bk1}^{(opt)} = \frac{\bar{y}}{1-g}\left[(K-g)\left(\frac{\bar{X}}{\bar{x}}\right) + (1-K)\left(\frac{\bar{x}^*}{\bar{X}}\right)\right]$$

Putting (2.5) in (2.3) and (2.4), we get the bias and mean squared error of $\hat{\bar{Y}}_{bk1}^{(opt)}$ respectively as

$$B\left(\hat{\bar{Y}}_{bk1}^{(opt)}\right) = \frac{(1-f)}{n}\bar{Y}\,C_x^2\left[\frac{K-g}{1-g} - K^2\right], \quad (2.7)$$

and

$$MSE\left(\hat{\bar{Y}}_{bk1}^{(opt)}\right) = \frac{(1-f)}{n}S_y^2(1-\rho^2) \qquad (2.8)$$

Equation (2.8) shows that mean squared error of $\hat{\bar{Y}}_{bk1}^{(opt)}$ is same as that of the approximate variance of the usual linear regression estimator $\bar{y}_{lr} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x})$, where $\hat{\beta}$ is the sample regression coefficient of y on x.

III. Efficiency Comparisons

Under simple random sampling without replacement (SRSWOR), variance of sample mean $\bar{y}$ is

$$V(\bar{y}) = \frac{(1-f)}{n}\bar{Y}^2 C_y^2, \qquad (3.1)$$

From (2.4) and (3.1), it is observed that $\hat{\bar{Y}}_{bk1}$ is more efficient than usual unbiased estimator $\bar{y}$ if

$$\left.\begin{array}{l} \text{either } \dfrac{g}{g-1} < \alpha < \left(\dfrac{2K-g}{1-g}\right) \\ \text{or } \left(\dfrac{2K-g}{1-g}\right) < \alpha < \dfrac{g}{g-1} \end{array}\right\} \qquad (3.2)$$

Comparison of (2.4) and (1.5) shows that $\hat{\bar{Y}}_{bk1}$ is more efficient than usual ratio estimator

$\bar{y}_R$ if

$$\left.\begin{aligned} &either \quad \frac{2K-g-1}{1-g} < \alpha < 1 \\ &or \quad 1 < \alpha < \frac{2K-g-1}{1-g} \end{aligned}\right\} \quad (3.3)$$

Comparing (2.4) and (1.6), it is observed that $\hat{\hat{Y}}_{bk1}$ is more efficient than usual dual to ratio estimator $\bar{y}_R^{(d)}$ if

$$\left.\begin{aligned} &either \quad \frac{2(K-g)}{1-g} < \alpha < \infty \\ &or \quad \infty < \alpha < \frac{2(K-g)}{1-g} \end{aligned}\right\} \quad (3.4)$$

IV. Empirical Study

To analyze the performance of the proposed estimator in comparison to other estimators, eleven natural population data sets are being considered. The descriptions of the populations are given below.

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

N=30 y : log of leaf burn in secs
n=6 x : chlorine perecentage

$\bar{Y}$ =0.6860, $\bar{X}$ =0.8077, C_y =0.700123 , C_x =0.7493, ρ =-0.4996, and f=0.20

Population II. [Source: Sukhatme and Sukhatme (1970), p. 256]

N = 89 y = Number of villages in the circles and
n=12 x = A circle consisting more than five villages.

$\bar{Y}$ =3.360, $\bar{X}$ = 0.1236, Cy = 0.60400, C_x = 2.19012 and ρ_{yx} = 0.766,

Population-III [Source: Maddala (1977)]

N=30 y : Consumption per capita
n=6 x : Deflated prices of veal

$\bar{Y}$ =7.6375, $\bar{X}$ =75.4313, C_y =0.2278 C_x =0.0986 and ρ =-0.6823

Population-IV [Source: Das (1988)]

x : The number of agricultural labourers for 1961
y : The number of agricultural labourers for 1971

$\bar{Y}$ =39.0680, $\bar{X}$ =25.1110, C_y =1.4451, C_x =1.6198 and ρ =0.7213

Population-V [Source: Das (1988)]

It consists of 142 cities of India with population (number of persons) 100,000 and above; the character x and y being
x : Census population in the year 1961
y : Census population in the year 1971

$\bar{Y}$ =4015.2183, $\bar{X}$ =2900.3872, C_y =2.1118, C_x =2.1971 and ρ =.9948

Population-VI [Source: Cochran (1977)]

y : The number of persons per block
x : The number of rooms per block

$\bar{Y}$ =101.1, $\bar{X}$ =58.80, C_y =0.14450, C_x =0.1281 and ρ =0.6500

Population-VII [Source: Pandey and Dubey (1988)]

N=20, n=8, $\bar{Y}$ = 19.55, $\bar{X}$ =18.8, C_y^2 =0.1262, C_x^2 =0.1555 and ρ_{yx} =-0.9199,

Population- VIII [Source: Kadilar and Singi (2006)]

N=104, y: the level of apple production
n=20, x: the number of apple trees

$\bar{Y}$ =625.37, $\bar{X}$ =13.93 C_y =1.866, C_x =1.653 and ρ =0.865

Population –IX : Murthy (1967, p. 228)

N= 80, y: Output
n = 20, x: Fixed Capital

$\bar{Y} = 51.8264, \quad \bar{X} = 11.2646, \quad C_y = 0.3542, \quad C_x = 0.7507$ and $\rho = 0.9413,$

Population –X : Murthy (1967, p. 228)

N= 80, y: Output
n = 20, x: Number of Workers $\bar{Y} = 51.8264, \quad \bar{X} = 2.8513, \quad C_y = 0.3542, \quad C_x = 0.9484$

and $\rho = 0.9150$

Population- XI [Source: Kadilar and Cingi (2004)]

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

Table – 4.1

Ranges of α in which $\hat{\hat{Y}}_{bk1}$ is better than $\bar{y}$, $\bar{y}_r$ and $\bar{y}_R^{(d)}$

Population	Range of α in which is $\hat{\hat{Y}}_{bk1}$ better than			Optimum value of α
	$\bar{y}$	$\bar{y}_R$	$\bar{y}_R^{(d)}$	α_0
I	(-0.33, 1.464)	(0.13, 1)	(∞ , 1.130)	0.5653
II	-0.184, 0.316)	(-0.869, 1))	(∞ , 0.131)	0.0656
III	(-0.5, -0.522)	(-6.729 , 1))	(∞ , -5.729)	2.8645
IV	(-0.138, 1.327)	(0.1889 , 1))	(∞ , 1.1889)	0.5944
V	(-0.138,2.038)	(0.900. 1)	(∞ , 1.900)	0.9501
VI	(-0.138, 1.053)	(.3930, 1)	(∞ , 1.393)	0.6965
VII	(-2.0, -6.972)	(-9.9720, 1)	(∞ , -8.972)	-4.4860
VIII	(-0.303, 2.241)	(0.938, 1)	(∞ , 1.939)	0.9693
IX	(-0.5, 0.832)	(-0.667, 1)	(∞ , 0.332)	0.1662
X	(-0.5, 0.525)	(-0.975, 1)	(∞ , 0.025)	0.0126
XI	(-0.303, 4.119)	(2.816, 1)	(∞ , 3.816)	1.9078

Table -4. 2

Percent relative efficiencies of $\bar{y}$, $\bar{y}_r$, $\bar{y}_R^{(d)}$ and $\hat{\hat{Y}}_{bk1}$ ($\hat{\hat{Y}}_{bk1}^{(opt)}$) with respect to $\bar{y}$

Population / Estimator	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
$\bar{y}$	100	100	100	100	100	100	100	100	100	100	100
$\bar{y}_R$	166.26	11.64	56.24	156.40	8031.10	157.87	23.40	396.50	66.59	30.59	226.76
$\bar{y}_R^{(d)}$	145.86	220.45	82.12	121.54	130.64	114.67	34.37	145.76	591.38	612.44	120.73
$\hat{\bar{Y}}_{bk1}$ ($\hat{\hat{Y}}_{bk1}^{(opt)}$)	208.45	241.99	187.10	208.45	9640.45	173.16	650.26	397.18	877.54	614.34	305.25

Table 4.2 shows that there is a significant gain in efficiency by using proposed estimator $\hat{\bar{Y}}_{bk1}$ or ($\hat{\bar{Y}}_{bk1}^{(opt)}$) over unbiased estimator $\bar{y}$, usual ratio estimator $\bar{y}_R$ and dual to ratio $\bar{y}_R^{(d)}$.

Table 4.1 provides the wide range of α in which suggested estimator $\hat{\bar{Y}}_{bk1}$ or ($\hat{\bar{Y}}_{bk1}^{(opt)}$) is more efficient then all estimators considered in this paper. This shows that even if the scalar α deviates from its optimum value (α_{opt}), the suggested estimator $\hat{\bar{Y}}_{bk1}$ will yield better estimates than $\bar{y}$, $\bar{y}_R$ and $\bar{y}_R^{(d)}$. Therefore, suggested estimator $\hat{\bar{Y}}_{bk1}$ or ($\hat{\bar{Y}}_{bk1}^{(opt)}$) is recommended for use in practice.

V. References

1) Cochran, W. G. (1977). Sampling Techniques. Third U. S. Edition. Wiley Eastern Limited.
   a. 325.
2) Das, A.K. (1988). Contribution to the theory of sampling strategies based on auxiliary information. Ph. D. Thesis, BCKV, West Bengal, India.
3) Maddala, G.S (1977). Econometrics, "Mcgraw Hills Pub.Co." New York.
4) Murty, M. N. (1964). Product method of estimation. Sankhya, A, 26, 294-307.
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) Pandey, B.N. and Dubey, V. (1988). Modified product estimator using coefficient of variation of auxiliary variable. Assam Stat. Rev., 2, 64-66.
8) Singh, H. P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. Statistics in Transition, 6 (4): 555-560.
9) Singh, H. P. and Tailor, R. (2005). Estimation of finite population mean using known correlation coefficient between auxiliary characters. Statistica, Anno LXV, 4, 407-418.
10) Sisodia, B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable. Jour. Ind. Soc. Agri. Stat., 33(1), 13-18.
11) Srivenkataramana , T. (1980). A dual of ratio estimator in sample surveys. Biometrika, 67, 1, 199-204.
12) Steel, R.G.D and Torrie, J. H.(1960). Principles and procedures of statistics, McGraw hill book.
13) Sukhatme, P. V. and Sukhatme, B. V. (1970): Sampling theory of surveys with
14) applications. Iowa State University Press, Ames, U. S. A.
15) 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, Jul-09, Vol. 10, No. 1, 15-24.