
Chain Ratio Type Estimators Using Known Parameters of Auxiliary Variates in Double Sampling

Priya Mehta* and Rajesh Tailor

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

E-mail: priya90128@gmail.com; tailorraj@gmail.com

**Corresponding Author*

Received 06 July 2020; Accepted 14 October 2020;

Publication 26 December 2020

Abstract

This paper discusses chain ratio type estimator for estimation of population mean in double sampling. The developed estimator uses two auxiliary variates associated with study variate in order to increase its efficiency. The developed estimator has been compared with usual unbiased estimator and other existing estimators. The expression for the bias and mean squared error of the developed estimator is obtained under large sample approximation. We have considered the natural population data set to examine the merits of the developed estimator and carried out the empirical study in support of theoretical findings. Numerical illustration shows that the proposed estimator is more efficient.

Keywords: Auxiliary variate, study variate, double sampling, coefficient of kurtosis, bias and MSE.

1 Introduction

It is well known fact that the auxiliary information can increase the efficiency of the estimators of parameters of importance in survey sampling. Designs

Journal of Reliability and Statistical Studies, Vol. 13, Issues 2–4 (2020), 243–252.

doi: 10.13052/jrss0974-8024.13242

© 2020 River Publishers

in which initially a sample of units is selected for obtaining auxiliary information only, and then a second sample is selected in which the variable of interest is observed in addition to the auxiliary information. It is useful in obtaining auxiliary variables for ratio estimation. In many practical situations, information on population mean of auxiliary variate is not available. In such situations double sampling procedure is quite efficient. For example, in a survey to estimate the production of lime crop based on orchards under the crop while the yield rate is determined from only a sub-sample of the orchards selected for determining acreage.

Bose (1943) established sampling error in the method of double sampling. Sukhatme (1962) discussed ratio type estimators in two-phase sampling. Srivastava (1970) A two phase sampling estimator in sample surveys.

Chand (1975) envisaged chain ratio estimator for population mean using ratio estimate for population mean of auxiliary variate using known population mean of another auxiliary variate. Rao (1975) discusses two phase ratio estimator in finite population. Singh and Upadhyaya (1995) used coefficient of variance of second auxiliary variate. In this paper a chain ratio in ratio estimator is being developed using coefficient of kurtosis of second auxiliary variate instead of coefficient of variate used by Singh and Upadhyaya (1995).

Kiregyera (1980) suggested chain ratio type estimator in finite population double sampling using two auxiliary variables. Hidioglou et al. (1998) suggested use of auxiliary information for two-phase sampling. Hidioglou (2001) discussed double sampling. Singh and Vishwakarma (2007) proposed modified exponential ratio and product estimators for finite population mean in double sampling. Tailor and Sharma (2009) suggested ratio-cum-product estimator of finite population mean using known coefficient of variation & coefficient of kurtosis.. Tailor and Sharma (2013) proposed ratio-cum-product estimator of finite population mean in double sampling when coefficient of variation & coefficient of kurtosis of auxiliary variable are known. Malik and Tailor (2013) developed ratio type estimator of population mean in double sampling. In double sampling procedure, first a large sample is peaked to estimate population mean of auxiliary variate and then a sub-sample is peaked either from 1st sample (Case-I) or from the population (Case-II).

Case-I: when the second phase sample of size n is a sub-sample of the first sample of size n' , and

Case-II: when the second phase sample of size n is drawn independently of the first-phase sample of size n' , see Bose (1943).

Let us consider a finite population $P = \{P_1, P_2, \dots, P_N\}$ of size N . Let y be the study variate and x and z are the auxiliary variates.

When population mean of auxiliary variate \bar{X} is not known, in double sampling usual ratio estimator of population mean \bar{Y} is defined as

$$\hat{Y}_R^d = \bar{y} \left(\frac{\bar{x}'}{\bar{x}} \right) \quad (1)$$

where $\bar{x}' = \frac{1}{n'} \sum_{i=1}^{n'} x_i$ is an unbiased estimator of population mean \bar{X} based on the sample of size n' and $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ and $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ are unbiased estimators of population means $\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i$ and $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$ respectively.

Chand (1975) defined a chain ratio-type estimator for \bar{Y} in double sampling as

$$\hat{Y}_{CR} = \bar{y} \left(\frac{\bar{x}'}{\bar{x}} \right) \left(\frac{\bar{Z}}{\bar{z}'} \right) \quad (2)$$

where $\bar{x}' = \frac{1}{n'} \sum_{i=1}^{n'} x_i$ and $\bar{z}' = \frac{1}{n'} \sum_{i=1}^{n'} z_i$ are unbiased estimators of population means \bar{X} and \bar{Z} based on sample size n' .

Thus the mean squared error of the \hat{Y}_R^d and \hat{Y}_{CR} up to the first degree of approximation under case I and case II respectively which are given as

$$MSE(\hat{Y}_R^d)_I = \bar{Y}^2 (f_1 C_y^2 + f_3 C_x^2 - 2f_3 \rho_{yx} C_y C_x), \quad (3)$$

$$MSE(\hat{Y}_R^d)_{II} = \bar{Y}^2 (f_1 C_y^2 + f_1 C_x^2 + f_2 C_x^2 - 2f_1 \rho_{yx} C_y C_x), \quad (4)$$

$$MSE(\hat{Y}_{CR})_I = \bar{Y}^2 (f_1 C_y^2 + f_3 C_x^2 - 2f_3 \rho_{yx} C_y C_x + f_2 C_z^2 - 2f_2 \rho_{yz} C_y C_z) \quad (5)$$

$$MSE(\hat{Y}_{CR})_{II} = \bar{Y}^2 (f_1 C_y^2 + f_2 C_x^2 + f_1 C_x^2 + f_2 C_z^2 - 2f_1 \rho_{yx} C_y C_x - 2f_2 \rho_{xz} C_x C_z). \quad (6)$$

It is well known under simple random sampling without replacement variance of unbiased estimator is defined as

$$V(\bar{y}) = \bar{Y}^2 f_1 C_y^2. \quad (7)$$

2 Developed Estimator

Assuming that coefficient of kurtosis of second auxiliary variate i.e. $\beta_2(Z)$ is known, motivated by Chand (1975) and Singh and Upadhyaya (1995) our developed estimator is defined as

$$\hat{Y}_{BC} = \bar{y} \left(\frac{\bar{x}'}{\bar{x}} \right) \left(\frac{\bar{Z} + \beta_2(Z)}{\bar{z}' + \beta_2(Z)} \right). \quad (8)$$

To obtain the Bias and mean squared error of the developed estimator \hat{Y}_{BC} in case-I and case-II we write

$$\begin{aligned} \bar{y} &= \bar{Y}(1 + e_0), & \bar{x} &= \bar{X}(1 + e_1), & \bar{x}' &= \bar{X}(1 + e'_1), \\ \bar{z} &= \bar{Z}(1 + e_2) & \text{and} & & \bar{z}' &= \bar{Z}(1 + e'_2) \end{aligned}$$

such that

$$\begin{aligned} E(e_0) &= E(e_1) = E(e'_1) = E(e_2) = E(e'_2) = 0, \\ E(e_0^2) &= \left(\frac{1}{n} - \frac{1}{N} \right) C_y^2, & E(e_1^2) &= \left(\frac{1}{n} - \frac{1}{N} \right) C_x^2, & E(e'_1{}^2) &= \left(\frac{1}{n'} - \frac{1}{N} \right) C_x^2, \\ E(e_2^2) &= \left(\frac{1}{n} - \frac{1}{N} \right) C_z^2, & E(e_0 e_1) &= \left(\frac{1}{n} - \frac{1}{N} \right) \rho_{yx} C_y C_x, \\ E(e_0 e'_1) &= \left(\frac{1}{n'} - \frac{1}{N} \right) \rho_{yx} C_y C_x, & E(e_0 e_2) &= \left(\frac{1}{n} - \frac{1}{N} \right) \rho_{yz} C_y C_z, \\ E(e_0 e'_2) &= \left(\frac{1}{n'} - \frac{1}{N} \right) \rho_{yz} C_y C_z, & E(e_1 e'_1) &= \left(\frac{1}{n'} - \frac{1}{N} \right) C_x^2, \\ E(e_1 e_2) &= \left(\frac{1}{n} - \frac{1}{N} \right) \rho_{xz} C_x C_z, & E(e_1 e'_2) &= \left(\frac{1}{n'} - \frac{1}{N} \right) \rho_{xz} C_x C_z & \text{and} \\ E(e_2 e'_2) &= \left(\frac{1}{n'} - \frac{1}{N} \right) C_z^2 \end{aligned}$$

Expressing \hat{Y}_{BC} in terms of e_i 's we have

$$\hat{Y}_{BC} = \bar{Y}(1 + e_0)(1 + e'_1)(1 + e_1)^{-1}(1 + \lambda e'_2)^{-1}. \quad (9)$$

Now using the standard technique, the mean squared error of the developed estimator \hat{Y}_{BC} up to the first degree of approximation under case-I and

case-II are respectively obtained as

$$MSE(\hat{Y}_{BC})_I = \bar{Y}^2 [C_y^2 (f_1 - 2f_3 K_{yx}) + C_x^2 (f_1 - f_2) + \lambda f_2 C_z^2 (\lambda - 2K_{yz})], \quad (10)$$

and

$$MSE(\hat{Y}_{BC})_{II} = \bar{Y}^2 [f_1 C_y^2 + f_1 C_x^2 + f_2 C_x^2 + \lambda^2 f_2 C_z^2 - 2f_1 K_{yx} - 2\lambda f_2 K_{xz}]. \quad (11)$$

where

$$\begin{aligned} C_y &= \frac{S_y}{\bar{Y}}, \quad C_x = \frac{S_x}{\bar{X}}, \quad C_z = \frac{S_z}{\bar{Z}}, \quad K_{yx} = \rho_{yx} \frac{C_y}{C_x}, \\ K_{yz} &= \rho_{yz} \frac{C_y}{C_z}, \quad K_{xz} = \rho_{xz} \frac{C_x}{C_z}, \quad \lambda = \frac{\bar{Z}}{\bar{Z} + \beta_2(Z)}, \\ f_1 &= \left(\frac{1}{n} - \frac{1}{N} \right), \quad f_2 = \left(\frac{1}{n'} - \frac{1}{N} \right) \quad \text{and} \quad f_3 = f_1 - f_2 \end{aligned}$$

3 Theoretical Comparisons of $(\hat{Y}_{BC})_I$

This section provides the situations under which developed chain ratio type estimator would have less mean squared error in comparison to alternative considered estimators.

Comparisons of (3), (5), (10) and (7) exhibits that the developed estimator \hat{Y}_{BC} under case-I $(\hat{Y}_{BC})_I$ would be higher efficient than other considered estimators.

- (i) It is observed from (7) and (10) that the developed estimator $(\hat{Y}_{BC})_I$ is more efficient than the usual unbiased estimator \bar{y} if

$$[f_1 C_x^2 - f_2 C_x^2 + \lambda^2 f_2 C_z^2 - 2f_3 \rho_{yx} C_y C_x - 2\lambda f_2 \rho_{yz} C_y C_z] < 0. \quad (12)$$

- (ii) It is observed from (3) and (10) that the developed estimator $(\hat{Y}_{BC})_I$ is more efficient than the usual ratio estimator \hat{Y}_R^d if

$$[f_1 C_x^2 - f_2 C_x^2 - f_3 C_x^2 + \lambda^2 f_2 C_z^2 - 2\lambda f_2 \rho_{yz} C_y C_z] < 0. \quad (13)$$

- (iii) It is observed from (5) and (10) that the developed estimator $(\hat{Y}_{BC})_I$ is more efficient than the chain ratio-type estimator \hat{Y}_{CR} if

$$[f_1 C_x^2 - f_2 C_x^2 - f_3 C_x^2 - f_2 C_z^2 + \lambda^2 f_2 C_z^2 - 2\lambda f_2 \rho_{yz} C_y C_z + 2f_2 \rho_{yz} C_y C_z] < 0. \quad (14)$$

4 Theoretical Comparisons of $(\hat{Y}_{BC})_{II}$

Comparisons of (4), (6), (11) and (7) exhibits that the developed estimator \hat{Y}_{BC} under case-II $(\hat{Y}_{BC})_{II}$ would be higher efficient than other considered estimators.

- (i) It is observed from (7) and (11) that the developed estimator $(\hat{Y}_{BC})_{II}$ is more efficient than the usual unbiased estimator \bar{y} if

$$[f_1 C_x^2 + f_1 C_x^2 + \lambda^2 f_2 C_z^2 - 2f_1 \rho_{yx} C_y C_x - 2\lambda f_2 \rho_{xz} C_x C_z] < 0. \quad (15)$$

- (ii) It is observed from (4) and (10) that the developed estimator $(\hat{Y}_{BC})_{II}$ is more efficient than the usual ratio estimator \hat{Y}_R^d if

$$[\lambda^2 f_2 C_z^2 - 2\lambda f_2 \rho_{xz} C_x C_z] < 0. \quad (16)$$

- (iii) It is observed from (6) and (10) that the developed estimator $(\hat{Y}_{BC})_{II}$ is more efficient than the chain ratio-type estimator \hat{Y}_{CR} if

$$[f_1 C_y^2 + \lambda^2 f_2 C_z^2 - f_2 C_z^2 - 2\lambda f_2 \rho_{xz} C_x C_z] < 0. \quad (17)$$

5 Numerical Comparison

To show the superiority of the developed estimator \hat{Y}_{BC} over simple mean estimator \bar{y} , double sampling ratio estimator \hat{Y}_R^d and chain ratio-type estimator \hat{Y}_{CR} , a natural population data set is given below:

Population [Source: Murthy 1967, p. 228]

z : Number of workers

y : Output and

x : Fixed capital

N	10
n	2
n'	4
\bar{Y}	1925.8
\bar{X}	214.4
\bar{Z}	51.8
C_y	0.319
C_x	0.349
C_z	0.014
ρ_{yx}	0.853
ρ_{yz}	0.89
ρ_{xz}	0.678
$\beta_2(x)$	1.88
$\beta_2(z)$	1.84

Table 1 MSE and PREs of \bar{y} , \hat{Y}_R^d , \hat{Y}_{CR} and \hat{Y}_{BC} under case-I and case-II with respect to \bar{y}

Estimators	MSE (Case-I)	MSE (Case-II)	PRE (Case-I)	PRE (Case-II)
\bar{y}	150960.6		100.00	
\hat{Y}_R^d	87792.4	117650.1	171.951	128.313
\hat{Y}_{CR}	83479.09	114073.3	180.836	132.336
\hat{Y}_{BC}	83623.44	114192.4	182.524	133.198

6 Conclusion

Sections 3 and 4 are provide the conditions under which the developed estimator \hat{Y}_{BC} is more efficient than usual unbiased estimator \bar{y} , usual ratio estimator \hat{Y}_R^d and chain ratio-type estimator \hat{Y}_{CR} . Table 1 clearly shows that the developed estimators \hat{Y}_{BC} has maximum percent relative efficiency as compared to other considered estimators under Case-I and Case-II.. Thus larger gain in efficiency is noticed using developed estimator \hat{Y}_{BC} over other estimators. It is also observed that $(\hat{Y}_{BC})_I$ is giving better result as compare to $(\hat{Y}_{BC})_{II}$. Hence, developed estimator \hat{Y}_{BC} may be recommended to use in practice for estimating the population mean provided conditions given

Sections 3 and 4 where coefficient of kurtosis of second auxiliary variate is known.

References

- Bose, C. (1943). Note on the sampling error in the method of double sampling. *Sankhya*, 6, 330.
- Chand, L. (1975). Some ratio type estimators based on two or more auxiliary variables. Unpublished Ph.D. Thesis, Iowa, State University, Ames, Iowa, U. S. A.
- Hidiroglou, M. A. (2001). Double sampling. *Survey Methodology*, 27, 143–154.
- Hidiroglou, M. A. and Sarndal, C. E. (1998). Use of auxiliary information for two-phase sampling. *Survey Methodology*, 24, 11–20.
- Kiregyera, B. (1980). A chain ratio-type estimator in finite population in double using two auxiliary variable, *Metrika*, 27, 217–223.
- Malik, K. A. and Tailor, R. (2013). Ratio type estimator of population mean in double sampling. *Int. J. Adv. Math. Statis.*, 1, 1, 34–39.
- Murthy, M. N. (1967). *Sampling theory and methods*. Statistical Publishing Society, Calcutta, India.
- Rao, P. S. R. S. (1975). On the two-phase ratio estimator in finite population. *J. Amer. Statist. Assoc.*, 70, 839–845.
- Singh, G. N. and Upadhyaya, L. N. (1995). A class of modified chain type estimator using two auxiliary variables in two-phase sampling. *Metron*, L III, 117–125.
- Singh, H. P. and Vishwakarma G. K. (2007). Modified exponential ratio and product estimators for finite population mean in double sampling. *Austrian J. Statist*, 36, 217–225.
- Srivastava, S. K. (1970). A two phase sampling estimator in sample surveys. *Austral. J. Statist.*, 12, 23–27.
- Sukhatme, B.V. (1962). Some ratio-type estimators in two-phase sampling. *J. Amer. Statist. Assoc.* 57, 628–632.
- Tailor, R. and Sharma, B. K. (2009). A modified ratio-cum-product estimator of finite population mean using known coefficient of variation & coefficient of kurtosis. *Stat. in Transition new series*, 10, 1, 15–24.
- Tailor, R. and Sharma, B.K. (2013). On the efficiency of ratio-cum-product estimator of finite population mean in double sampling when coefficient of variation & coefficient of kurtosis of auxiliary variable are known. *J. Appl. Statist. Sci.*, 20, 1, 1–20.

Biographies



Priya Mehta is a Doctorate from Vikram University, Ujjain in 2019, my undergraduate studies completed at Vikram University, Ujjain in 2010, I received my Post Graduate in 2012 from Vikram University, Ujjain. My research interest lies in the area of Sampling Techniques. I have completed my Ph.D. on Contribution of Auxiliary Information in Estimation of Population Parameters at Vikram University, Ujjain.



Rajesh Tailor received his M.Sc. degree in 1998 from Vikram University, Ujjain. He did his M.Phil. in 1999 and Ph.D. in 2002 from Vikram University, Ujjain. He started his career as a lecturer at NCERT, New Delhi and became Reader in 2008 Vikram University, Ujjain. He is a life member of Sankhya, Indian Science Congress, Indian Society of Agriculture Statistics and he is reviewer of various National and International Journals of Statistics. He is Nodal officer of All India Survey on Higher Education, a programmer of MHRD, Government of India, Calcutta Statistical Association.

