International Journal of Advanced Statistics and Probability, 1 (3) (2013) 82-87
©Science Publishing Corporation
www.sciencepubco.com/index.php/IJASP

Improved chain type estimators for population mean using two auxiliary variables and double sampling scheme

B.B. Khare and Habib Ur Rehman*

Department of Statistics, BHU, Varanasi-221005
**Corresponding author E-mail:hrmbd007@gmail.com*

Abstract

The chain type estimators for population mean using auxiliary and additional auxiliary variables using double sampling scheme have been proposed and their properties are studied. The proposed estimators are also found to more efficient than the relevant chain type estimators for population mean proposed by Chand [1] and Kiregyera [2,3]. An empirical study has been given in the support of the proposed estimators.

Keywords: *Chain Type and regression estimators, population mean, study variable, coefficient of variation.*

1 Introduction

The use of coefficient of variation of the study character helps for increasing the efficiency of the estimators. Using coefficient of variation, Searls [4,5] first proposed the estimator for population mean of the study character. Further, Das and Tripathi [6], Sisodia and Dwivedi [8] and Panday and Dubey [7] proposed estimators for population mean using coefficient of variation of auxiliary character.
In case when population mean of the auxiliary character is not known but population mean of another additional auxiliary character, Chand [1]proposed the chain ratio type estimator and Kiregyera [2,3] proposed chain ratio to regression, regression in regression and ratio in regression estimators for population mean of study character.
In this paper, we have proposed improved chain type estimators for population mean of study character using known coefficient of variation of study character. The expressions of bias and mean squarer error (MSE) are obtained and their properties are study. A comparative study of the proposed estimators with the relevant estimators is given. An empirical study is also given to show the performance of the proposed estimators with the relevant estimators.

2 The estimators

Let $\bar{Y}$, $\bar{X}$ and $\bar{Z}$ denote the population mean of study character y, auxiliary character x and additional auxiliary character z having jth value Y_j, X_j and Z_j: $j = 1,2,3.....N$.

In case when population mean of auxiliary character is not known, we draw a large preliminary sample of size $n'(< N)$ from population of size N by using simple random sampling without replacement (SRSWOR) scheme of sampling and estimate the population mean $\bar{X}$ by first phase sample mean $\bar{x}'$ based on n' units. Further, we draw a sub-sample of size $n(< n')$ from large preliminary sample of size n' by using SRSWOR and compute $\bar{y}$ and $\bar{x}$ based on n units. In such case the double sampling ratio and regression estimators are defined as

$$T_R = \bar{y}\frac{\bar{x}'}{\bar{x}} \tag{1}$$

and

$$T_{RR} = \bar{y} + b_{yx}(\bar{x}' - \bar{x}), \tag{2}$$

where $\bar{x}' = \frac{1}{n'}\sum_{j=1}^{n'} x_j$, $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$, $b_{yx} = \frac{\hat{S}_{yx}}{\hat{S}_x^2}$, $\hat{S}_x^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2, \hat{S}_{yx} = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})(y_j - \bar{y})$

The mean square errors of the estimators T_R and T_{RR} are given as follows:

$$MSE(T_R) = V(\bar{y}) + \bar{Y}^2\left(\frac{1}{n} - \frac{1}{n'}\right)\left(C_x^2 - 2\rho_{yx}C_yC_x\right), \tag{3}$$

and

$$MSE(T_{RR}) = V(\bar{y}) - \bar{Y}^2\left(\frac{1}{n} - \frac{1}{n'}\right)\rho_{yx}^2 C_y^2, \tag{4}$$

where $V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2 C_y^2$

In case when population mean $(\bar{X})$ of an auxiliary character x is not known but population mean $(\bar{Z})$ of another additional auxiliary character z is known. It is suggested to take a large preliminary sample of size $n'(< N)$ from population of size N by using simple random sampling without replacement scheme of sampling and estimate the population mean $\bar{X}$ by $\hat{\bar{X}} = \bar{x}'\bar{Z}/\bar{z}'$ which is more efficient in comparison to $\bar{x}'$ if $\rho_{xz} > C_z/2C_x$, where $\bar{x}'$ and $\bar{z}'$ is the preliminary sample mean based on n' units, ρ_{xz} is the correlation coefficient between x and z and (C_x, C_z) are coefficient of variations of auxiliary character and additional auxiliary character.

In this case Chand [1] proposed chain ratio type and Kiregyera [3] chain ratio in regression and regression in regression type estimators, which are given as follows:

$$T_C = \bar{y}\frac{\bar{x}'}{\bar{x}}\frac{\bar{Z}}{\bar{z}'} \tag{5}$$

$$T_{K1} = \frac{\bar{y}}{\bar{x}}\left[\bar{x}' + b_{xz}\left(\bar{Z} - \bar{z}'\right)\right] \tag{6}$$

and

$$T_{K2} = \bar{y} + b_{yx}\left[\left(\bar{x}' - \bar{x}\right) - b_{xz}\left(\bar{Z} - \bar{z}'\right)\right] \tag{7}$$

Mean square error of the estimators T_C, T_{K1} and T_{K2} are given as follows:

$$MSE(T_C) = \bar{Y}^2\left[\frac{f}{n}C_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left\{C_x^2 - 2C_{yx}\right\} + \left(\frac{1}{n'} - \frac{1}{N}\right)C_z^2 - 2\left(\frac{1}{n'} - \frac{1}{N}\right)C_{yz}\right]$$

$$MSE(T_{K1}) = \bar{Y}^2\left[\frac{f}{n}C_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left\{C_x^2 - 2C_{yx}\right\} + \beta^2 R^2\left(\frac{1}{n'} - \frac{1}{N}\right)C_z^2 - 2\beta R\left(\frac{1}{n'} - \frac{1}{N}\right)C_{yz}\right]$$

$$MSE(T_{K2}) = \bar{Y}^2\frac{f}{n}C_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left\{\beta_1^2\bar{X}^2C_x^2 - 2\bar{X}\bar{Y}\beta_1 C_{yx}\right\} + \beta_1^2\beta_2^2\bar{Z}^2\left(\frac{1}{n'} - \frac{1}{N}\right)C_z^2 - 2\bar{Y}\bar{Z}\beta_1\beta_2\left(\frac{1}{n'} - \frac{1}{N}\right)C_{yz}$$

Where $\beta = \beta_1 = \frac{\bar{Y}\rho_{yx}C_y}{\bar{X}Cx}$ and $\beta_2 = \frac{\bar{X}\rho_{xz}C_x}{\bar{Z}Cz}$

Serals [4] proposed an estimator for population mean of study character using known coefficient of variation are study character which is given as follows

$\bar{y}^* = a\,\bar{y}$ which has $MSE(\bar{y}^*) \le MSE(\bar{y})$ for the value of $a = \left[1 + \frac{f}{n}\bar{Y}^2C_y^2\right]^{-1}$,

where $f = \frac{N-n}{N}$

The $V(\bar{y}^*)$ is given by

$V(\bar{y}^*) = a\frac{f}{n}\bar{Y}^2C_y^2 = (1-B)\frac{f}{n}\bar{Y}^2C_y^2$ (After expanding a up to the term of order n^{-1})

Where $B = \frac{f}{n}\bar{Y}^2 C_y^2$

Now, using known coefficient of variation of the study character, we propose generalized ratio in regression and regression in regression type estimators for population mean which are given as follows:

$$T_1' = \frac{\bar{y}^*}{\bar{x}}[\bar{x}' + b(\bar{Z} - \bar{z}')] \tag{8}$$

and

$$T_2' = \bar{y}^* + b_1[(\bar{x}' - \bar{x}) - b_2(\bar{Z} - \bar{z}')], \tag{9}$$

where b, b_1 and b_2 are constant.

3 Bias and mean square error of the estimators T_1' and T_2'

$$B(T_1') = \bar{Y}\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\{C_x^2 - (1-B)C_{yx}\} - \left(\frac{1}{n'} - \frac{1}{N}\right)bR\{C_{xz} - (1-B)C_{yz}\}\right]$$

and

$$MSE(T_1') = \bar{Y}^2\left[(1-B)\frac{f}{n}C_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\{C_x^2 - 2(1-B)C_{yx}\} + b^2R^2\left(\frac{1}{n'} - \frac{1}{N}\right)C_z^2 - 2bR(1-B)\left(\frac{1}{n'} - \frac{1}{N}\right)C_{yz}\right] \tag{10}$$

$$MSE(T_2') = \bar{Y}^2(1-B)\frac{f}{n}C_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\{b_1^2\bar{X}^2C_x^2 - 2(1-B)\bar{X}\bar{Y}b_1C_{yx}\} + b_1^2b_2^2\bar{Z}^2\left(\frac{1}{n'} - \frac{1}{N}\right)C_z^2$$

$$- 2\bar{Y}\bar{Z}b_1b_2(1-B)\left(\frac{1}{n'} - \frac{1}{N}\right)C_{yz}, \tag{11}$$

Where $R = \frac{\bar{Z}}{\bar{X}}$ and $B = \frac{f}{n}\bar{Y}^2C_y^2$

The optimum values of b, b_1 and b_2 for which the $MSE\ (T_1')$ and $MSE\ (T_2')$ are minimum, and are given as follows:

$$b_{(opt)} = \frac{(1-B)C_{yz}}{RC_z^2},$$

$$b_{1(opt)} = \frac{-k_2 \pm \sqrt{k_2^2 - 4k_1k_3}}{2k_1}$$

and

$$b_{2(opt)} = \frac{M}{b_{1(opt)}}$$

where $M = \frac{\bar{Y}(1-B)C_{yz}}{\bar{Z}C_z^2}$, $k_1 = \left(\frac{1}{n} - \frac{1}{n'}\right)b_1^2\bar{X}^2C_x^2$, $k_2 = -\left(\frac{1}{n} - \frac{1}{n'}\right)(1-B)\bar{X}\bar{Y}b_1C_{yx}$

and $k_3 = M\bar{Z}\left(\frac{1}{n'} - \frac{1}{N}\right)\{M\bar{Z}C_z^2 - \rho\bar{Y}(1-B)C_{yz}\}$

4 Comparisons of the proposed estimators with competitive estimators

$MSE(T_1') < MSE(T_{K1})$

if $\frac{-M_2 - \sqrt{M_2^2 - 4M_1M_3}}{2M_1} < b < \frac{-M_2 + \sqrt{M_2^2 - 4M_1M_3}}{2M_1}$

$MSE(T_1') < MSE(T_C)$

if $\dfrac{-M_2-\sqrt{M_2^2-4M_1N_1}}{2M_1}<b<\dfrac{-M_2+\sqrt{M_2^2-4M_1N_1}}{2M_1}$

$MSE(T_1')<MSE(T_{RR})$

if $\dfrac{-M_2-\sqrt{M_2^2-4M_1P_1}}{2M_1}<b<\dfrac{-M_2+\sqrt{M_2^2-4M_1P_1}}{2M_1}$

$MSE(T_2')<MSE(T_{K2})$

if $\dfrac{-R_2-\sqrt{R_2^2-4R_1R_3}}{2R_1}<b_1<\dfrac{-R_2+\sqrt{R_2^2-4R_1R_3}}{2R_1}$

and

$\dfrac{-A_2-\sqrt{A_2^2-4A_1A_3}}{2A_1}<b_2<\dfrac{-A_2+\sqrt{A_2^2-4A_1A_3}}{2A_1}$

$MSE(T_{21}')<MSE(T_C)$

if $\dfrac{-R_2-\sqrt{R_2^2-4R_1S_1}}{2R_1}<b_1<\dfrac{-R_2+\sqrt{R_2^2-4R_1S_1}}{2R_1}$

and

$\dfrac{-A_2-\sqrt{A_2^2-4B_1A_3}}{2B_1}<b_2<\dfrac{-A_2+\sqrt{A_2^2-4B_1A_3}}{2B_1}$

$MSE(T_2')<MSE(T_{RR})$

if $\dfrac{-R_2-\sqrt{R_2^2-4R_1T_1}}{2R_1}<b_1<\dfrac{-R_2+\sqrt{R_2^2-4R_1T_1}}{2R_1}$

and

$\dfrac{-A_2-\sqrt{A_2^2-4C_1A_3}}{2C_1}<b_2<\dfrac{-A_2+\sqrt{A_2^2-4C_1A_3}}{2C_1}$

Where $M_1=\left(\dfrac{1}{n'}-\dfrac{1}{N}\right)R^2C_z^2, M_2=-\left(\dfrac{1}{n'}-\dfrac{1}{N}\right)2R(1-B)C_{yz},$

$M_3=-B\dfrac{f}{n}C_y^2+\left(\dfrac{1}{n}-\dfrac{1}{n'}\right)2BC_{yx}-\left(\dfrac{1}{n'}-\dfrac{1}{N}\right)R\beta\left\{R\beta C_z^2+2C_{yz}\right\}$

$N_1=-B\dfrac{f}{n}C_y^2+\left(\dfrac{1}{n}-\dfrac{1}{n'}\right)2BC_{yx}-\left(\dfrac{1}{n'}-\dfrac{1}{N}\right)\left\{C_z^2-2C_{yz}\right\}$

$P_1=-B\dfrac{f}{n}C_y^2+\left(\dfrac{1}{n}-\dfrac{1}{n'}\right)\left\{C_x^2-2(1-B)C_{yx}+2BC_{yx}\right\}$

$R_1=\left(\dfrac{1}{n}-\dfrac{1}{n'}\right)\bar{X}^2C_x^2$

$R_2=-\left(\dfrac{1}{n}-\dfrac{1}{n'}\right)2(1-B)\bar{X}\bar{Y}C_{yx}$

$$R_3 = -B\bar{Y}^2\frac{f}{n}C_y^2 - \left(\frac{1}{n}-\frac{1}{n'}\right)\left\{\beta_1^2\bar{X}^2C_x^2 - 2\bar{X}\bar{Y}\beta_1 C_{yx}\right\} - \left(\frac{1}{n'}-\frac{1}{N}\right)\bar{Z}\left\{M^2-\beta_1^2\beta_2^2\bar{Z}\right\}C_z^2 - \left(\frac{1}{n'}-\frac{1}{N}\right)2\bar{Y}\bar{Z}\left\{M(1-B)-\beta_1\beta_2\right\}C_{yz}$$

$$S_1 = -B\bar{Y}^2\frac{f}{n}C_y^2 - \left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}^2\left\{C_x^2 - 2C_{yx}\right\} + \left(\frac{1}{n'}-\frac{1}{N}\right)\left\{\bar{Z}^2M^2C_z^2 - C_z^2 + 2C_{yz}\right\} - \left(\frac{1}{n'}-\frac{1}{N}\right)2\bar{Y}\bar{Z}M(1-B)C_{yz}$$

$$T_1 = -B\bar{Y}^2\frac{f}{n}C_y^2 + \left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}^2\rho_{yx}^2C_y^2 + \left(\frac{1}{n'}-\frac{1}{N}\right)\left\{\bar{Z}^2M^2C_z^2 - 2\bar{Y}\bar{Z}M(1-B)C_{yz}\right\}$$

$$A_1 = -B\bar{Y}^2\frac{f}{n}C_y^2 - \left(\frac{1}{n}-\frac{1}{n'}\right)\left\{\beta_1^2\bar{X}^2C_x^2 - 2\bar{X}\bar{Y}\beta_1 C_{yx}\right\} + \left(\frac{1}{n'}-\frac{1}{N}\right)\bar{Z}^2\left\{M^2-\beta_1^2\beta_2^2\right\}C_z^2 - \left(\frac{1}{n'}-\frac{1}{N}\right)2\bar{Y}\bar{Z}\left\{M(1-B)+\beta_1\beta_2\right\}C_{yz}$$

$$A_2 = -2\left(\frac{1}{n}-\frac{1}{n'}\right)(1-B)\bar{X}\bar{Y}MC_{yx}$$

$$A_3 = \left(\frac{1}{n}-\frac{1}{n'}\right)M^2\bar{X}^2C_x^2$$

$$B_1 = -B\bar{Y}^2\frac{f}{n}C_y^2 - \left(\frac{1}{n}-\frac{1}{n'}\right)\left\{C_x^2 - C_{yx}\right\} + \left(\frac{1}{n'}-\frac{1}{N}\right)\left\{M^2\bar{Z}^2-1\right\}C_z^2 + \left(\frac{1}{n'}-\frac{1}{N}\right)2\left\{1-\bar{Y}\bar{Z}M(1-B)\right\}C_{yz}$$

$$C_1 = -B\bar{Y}^2\frac{f}{n}C_y^2 + \left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}^2\rho_{yx}^2C_y^2 + \left(\frac{1}{n'}-\frac{1}{N}\right)M\bar{Z}\left\{M\bar{Z}C_z^2 - 2\bar{Y}(1-B)C_{yz}\right\}$$

5 Empirical study

The data from the population of 100 records of resale of homes from Feb 15 to Apr 30, 1993 from the files maintained by the Albuquerque Board of Realtors on Selling price ($hundreds) as a study variable (y), Square feet of living space as an auxiliary variable (x) and annual taxes ($) as an additional variable (z) have been taken. The values of the parameters of the population are given as follows:

$\bar{X} = 1697.44, \bar{Y} = 1093.41,\ \bar{Z} = 801.58,\ S_x = 535.01, S_y = 391.90,\ R = \frac{\bar{Z}}{\bar{X}},$
$S_z = 316.62, \rho_{yx} = 0.84, \rho_{yz} = 0.64, \rho_{xz} = 0.86$

Table 1: Relative efficiency of the estimators $\bar{y}$,T_R ,T_{RR} ,T_C , T_{K1},T_{K2},T_1' *and* T_2' (in %) with respect to $\bar{y}$.

Estimators	$n'=60, n=30$	$n'=65, n=35$	$n'=70, n=40$
$\bar{y}$	100(3583.66)*	100(2852.30)*	100(2303.78)*
T_R	211.87(1691.43)	210.48(1355.15)	211.87(1087.35)
T_{RR}	211.99(1690.47)	210.60(1354.39)	211.99(1086.73)
T_C	240.41(1490.62)	239.09(1192.96)	240.41(958.26)
T_{K1}	279.54(1282.01)	278.42(1024.46)	279.54(824.15)
T_{K2}	280.41(1278.03)	279.29(1021.27)	280.41(821.59)
T_1'	281.04(1275.16)	280.09(1018.37)	281.28(819.03)
T_2'	281.26(1274.17)	280.30(1017.59)	281.50(818.40)

*Figures in parenthesis give the $MSE(.)$.

From table, we observed that the proposed estimator T_1' and T_2' are more efficient than the relevant estimators $\bar{y}$, T_R , T_{RR} , T_C, T_{K1} and T_{K2} proposed by Chand [1] and Kiregyera [3].

References

[1] Chand, L. (1975). Some Ratio-type Estimators based on two or more Auxiliary Variables. Ph.D. Thesis submitted to Iowa State University, Ames, IOWA 1975.

[2] Kiregyera, B. (1980). A Chain Ratio-type Estimator in Finite Population Double Sampling using two Auxiliary Variables. Metrika, 27, 217-223.

[3] Kiregyera, B. (1984). Regression-type Estimator using two Auxiliary Variables and Model of Double Sampling from Finite Population. Metrika, 31, 215-226.

[4] Searls, D. T. (1964). The utilization of a known coefficient of variation in the estimation procedure. J. Amer. Statist. Assoc., 59, 1225-1226.

[5] Searls, D. T. (1967). A note on the use of an approximately known coefficient of variation. J. Amer. Statist. Assoc. 21:20–21.

[6] Das, A.K. and Tripathi, T.P. (1980). Sampling strategies for population mean when the coefficient of variation of an auxiliary character is known. Sankhya, C, 42, pp. 76-86.

[7] Pandey, B.N. and Dubey, V. (1988). Modified product estimator using coefficient of variation of auxiliary vitiates. Assam Statistical Review, 2, part 2, pp. 64-66.

[8] Sisodia, B. V. and Dwivedi, V. K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable. Journal Indian Society of Agricultural Statistics, New Delhi, 33, pp.13-18.