Vol.4.Issue.4.2016 (October -December)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

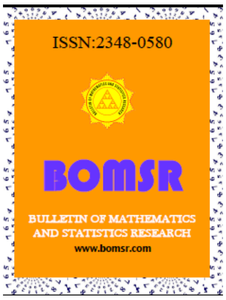
http://www.bomsr.com
Email:editorbomsr@gmail.com

RESEARCH ARTICLE

ISSN:2348-0580

THE CHAIN PRODUCT ESTIMATOR WITH LINEAR COMBINATION OF TWO AUXILIARY VARIABLES

K. B. PANDA & S.S. PARIDA
Department of Statistics, Utkal University, Bhubaneswar, Odisha, India



ABSTRACT

In this paper, we propose a new product estimator for estimating the mean of a finite population using linear combination of two auxiliary variables, each of which is negatively correlated with the study variable .We have found that the estimator fares better than the traditional multivariate product estimator using information on two auxiliary variables under the condition that holds good in practice very often. Numerical illustrations involving quantities of two real populations have been considered in favour of the proposed estimator.

Key words: Auxiliary variable, Chain product estimator, Multivariate product estimator,Mean square error.

1 INTRODUCTION

When complete information is available for an auxiliary variate x, which is negatively correlated with the study variable y, the product method of estimation is usually employed to estimate the population mean $\bar{Y}$ or the population total Y. Let us consider a population of size N whose units are labelled$1, 2, \ldots\ldots, N$, and let y_i and x_i, respectively, be the values of the study variable y and the auxiliary variable x for the ith unit. Let $\bar{Y}$and $\bar{X}$ be the respective population means, and $\bar{y}$ and $\bar{x}$ be the respective sample means of the y and the x observations. When the auxiliary variable x is negatively correlated with y, the usual product estimator due to Murthy(1964) is given by

$$\bar{y}_P = \bar{y}\frac{\bar{x}}{\bar{X}}. \tag{1.1}$$

The estimator is biased and its mean square error(MSE), to $O(n^{-1})$, is expressed as

$$MSE(\bar{y}_P) = \frac{1-f}{n}({S_y}^2+R^2{S_x}^2 + 2RS_{yx}), \tag{1.2}$$

where $f = \frac{n}{N}$, $R = \frac{\bar{Y}}{\bar{X}}$, the population ratio, S_y^2, S_x^2 are the population variances of y and x, respectively, and S_{yx} is the population covariance between y and x.

Following Kadilar and Cingi(2005), a chain product estimator is defined as

$$\bar{y}_{CP} = \bar{y}\,(\frac{\bar{x}}{\bar{X}})^{\alpha}, \quad (1.3)$$

where α is a real number. The MSE of this estimator, to $O(n^{-1})$, is given as follows:

$$MSE(\bar{y}_{CP}) = \frac{1-f}{n}({S_y}^2 + \alpha^2 R^2 {S_x}^2 + 2\alpha R S_{yx}). \quad (1.4)$$

The customary multivariate product estimator using information on two auxiliary variables x_1 and x_2 is expressed as

$$\bar{y}_{MP} = \in_1 \bar{y}\frac{\bar{x}_1}{\bar{X}_1} + \in_2 \bar{y}\frac{\bar{x}_2}{\bar{X}_2}, \quad (1.5)$$

where $\bar{x}_i$ and $\bar{X}_i (i = 1,2)$ denote, respectively, the sample and the population means of the variable $x_i (i = 1,2)$ and $\in_1$ and $\in_2$ are the weights that satisfy the condition $\in_1 + \in_2 = 1$.

The MSE of the usual multivariate product estimator is , to $O(n^{-1})$, derived as

$$MSE(\bar{y}_{MP}) = \frac{1-f}{n}(S_y^2 + \in_1^2 R_1^2 S_{x_1}^2 + \in_2^2 R_2^2 S_{x_2}^2 + 2\in_1 R_1 S_{yx_1} + 2\in_2 R_2 S_{yx_2} + 2\in_1\in_2 R_1 R_2 S_{x_1x_2}), \quad (1.6)$$

where $S_{x_i}^2 (i = 1,2)$ is the population variance of x_i, $S_{yx_i} (i = 1,2)$ is the covariance between y and x_i, $S_{x_1x_2}$ is the covariance between x_1 and x_2 and $R_i = \frac{\bar{Y}}{\bar{X}_i} (i = 1,2)$. With a view to determining the optimum weights $\in_1$ and $\in_2$, we take recourse to partial differentiation of $MSE(\bar{y}_{MP})$ with respect to $\in_1$ and $\in_2$ and equate them to 0. As a result , the following optimum weights are obtained:

$$\in_1^* = \frac{R_2^2 S_{x_2}^2 - R_1 S_{x_1y} + R_2 S_{x_2y} - R_1 R_2 S_{x_1x_2}}{R_1^2 S_{x_1}^2 + R_2^2 S_{x_2}^2 - 2R_1 R_2 S_{x_1x_2}} = 1 - \in_2^*. \quad (1.7)$$

Thus, the minimum MSE of $\bar{y}_{MP}$ is given by

$$MSE_{min}(\bar{y}_{MP}) = \frac{1-f}{n}(S_y^2 + \in_1^* R_1^2 S_{x_1}^2 + \in_2^* R_2^2 S_{x_2}^2 + 2\in_1^* R_1 S_{yx_1} + 2\in_2^* R_2 S_{yx_2} + 2\in_1^*\in_2^* R_1 R_2 S_{x_1x_2}). \quad (1.8)$$

2. The New Estimator

Following Lu(2013),we propose the multivariate chain product estimator using linear combination of two auxiliary variables as

$$\bar{y}_{\alpha lcP} = \bar{y}(\frac{w_1\bar{x}_1 + w_2\bar{x}_2}{w_1\bar{X}_1 + w_2\bar{X}_2})^{\alpha}, \quad (2.1)$$

where α is , as usual, a real number and w_1 and w_2 are the weights subject to the condition $w_1 + w_2 = 1$.

The MSE of the proposed estimator, to the first degree of approximation, i.e., to $O(n^{-1})$, is derived to be

$$MSE(\bar{y}_{\alpha lcP}) = \frac{1-f}{n}(S_y^2 + \alpha^2 w_1^2 R_{lc}^2 S_{x_1}^2 + w_2^2\alpha^2 R_{lc}^2 S_{x_2}^2 + 2w_1\alpha R_{lc} S_{yx_1} + 2w_2\alpha R_{lc} S_{yx_2} + 2w_1 w_2 \alpha^2 R_{lc}^2 S_{x_1x_2}), \quad (2.2)$$

where $R_{lc} = \frac{\bar{y}}{w_1\bar{X}_1 + w_2\bar{X}_2}$.

Minimising (2.2) subject to variations in $w_1$and w_2, we get the optimum values of w_1 and $w_2$as

$$w_1^* = \frac{\propto S_{x_2}^2\bar{X}_1\bar{Y} - \alpha S_{x_1x_2}\bar{X}_2\bar{Y} + \bar{X}_1\bar{X}_2 S_{x_2y} - \bar{X}_2^2 S_{x_1y}}{\alpha\bar{X}_2\bar{Y}S_{x_1}^2 - \alpha\bar{X}_2\bar{Y}S_{x_1x_2} - \alpha\bar{X}_1\bar{Y}S_{x_1x_2} + \alpha\bar{X}_1\bar{Y}S_{x_2}^2 + \bar{X}_1\bar{X}_2 S_{x_1y} - \bar{X}_2^2 S_{x_1y} - \bar{X}_1^2 S_{x_1y} + X_1X_2 S_{x_2y}} = 1 - w_2^*$$

and the minimum value of mean square error as

$$MSE_{min}(\bar{y}_{\alpha lcP}) = \frac{1-f}{n}(S_y^2 + \alpha^2 {w_1^*}^2 {R_{lc}^*}^2 S_{x_1}^2 + \alpha^2 {w_2^*}^2 {R_{lc}^*}^2 S_{x_2}^2 + 2\alpha w_1^* R_{lc}^* S_{x_1y} + 2\alpha w_2^* R_{lc}^* S_{x_2y} + 2\alpha^2 w_1^* w_2^* {R_{lc}^*}^2 S_{x_1x_2}) \quad (2.3)$$

3. Efficiency Comparison

With a view to comparing the efficiency of the proposal estimator, we proceed as follows:

$$MSE_{min}(\bar{y}_{\alpha lcP}) < MSE_{min}(\bar{y}_{MP})$$

$$\Rightarrow \frac{1-f}{n}(S_y^2 + \alpha^2 w_1^{*^2} R_{lc}^{*^2} S_{x_1}^2 + \alpha^2 w_2^{*^2} R_{lc}^{*^2} S_{x_2}^2 + 2\alpha w_1^* R_{lc}^* S_{x_1 y} + 2\alpha w_2^* R_{lc}^* S_{x_2 y} + 2\alpha^2 w_1^* w_2^* R_{lc}^{*^2} S_{x_1 x_2})$$

$$< \frac{1-f}{n}(S_y^2 + \in_1^{*^2} R_1^2 S_{x_1}^2 + \in_2^{*^2} R_2^2 S_{x_2}^2 + 2 \in_1^* R_1 S_{yx_1} + 2 \in_2^* R_2 S_{yx_2} + 2 S_{x_1 x_2} \in_1^* \in_2^* R_1 R_2).$$

$$\Rightarrow S_{x_1}^2 \left(\alpha^2 w_1^{*^2} R_{lc}^{*^2} - \in_1^{*^2} R_1^2\right) + S_{x_2}^2 \left(\alpha^2 w_2^{*^2} R_{lc}^{*^2} - \in_2^{*^2} R_2^2\right) + 2 S_{x_1 y}(\alpha w_1^* R_{lc}^* - \in_1^* R_1)$$

$$+ 2 S_{x_2 y}(2\alpha w_2^* R_{lc}^* - \in_2^* R_2) + 2 S_{x_1 x_2}(\alpha^2 w_1^* w_2^* R_{lc}^{*^2} - \in_1^* \in_2^* R_1 R_2) < 0$$

4. Empirical Investigation

Example 1: We refer to Example 8.1(Highway data) given in Weisberg(1980, p.179). The sample quantities given therein have been taken as the corresponding population quantities which are as follows:

N=39, $\bar{Y}$ = 3.933, $\bar{X}_1 = 55, \bar{X}_2 = 12.884$, $S^2_{y=3.94,}$ $S_{x_1}^2 = 34.211$, $S_{x_2}^2 = 57.907$, $S_{yx_1} = -0.621$, $S_{yx_2} = -6.852$, $S_{x_1x_2} = 0.186$, where the variables are described as

$Y = RATE = 1973\ accident\ rate\ per\ million\ vehicle\ miles$

$X_1 = SLIM = Speed\ limit(\ in\ 1973, before\ the\ 55\ mph\ limits)$

$X_2 = LEN = Length\ of\ the\ segment\ in\ miles$

In addition to the above information, we choose the sample size to be 6. The Table given below presents the $MSEs$ of the competing estimators.

Table 1: Efficiency of the competing estimators

$\boldsymbol{Estimator}$	$\boldsymbol{MSE}$
$\bar{y}_{MP}$	0.76
$\bar{y}_{alcP}$	0.70

The above Table clearly points to the fact that the newly proposed chain product estimator comes out to be more efficient than the existing multivariate product estimator, the gain in efficiency being around 9 per cent.

Example 2: We refer to an investigation carried out by the Biometry Research Unit of the Indian Statistical Institute with regard to multivariate investigations of blood chemistry and morphology. The details of this investigation are given by Das (1966). To illustrate the performance of the competing estimators, we consider, out of the 32 variables, "Alkaline Phosphatase" as the study variable and 'Age" and "Weight" as the two auxiliary variables negatively correlated with the study variable**.** The following data correspond to urban (rural upbringing) vegetarian males (Group C):

N=69, $\bar{Y}$ = 9.81, $\bar{X}_1 = 27.49$, $\bar{X}_2 = 49.23$, $S_{y,}^2 = 49.23$, $S_{x_1}^2 = 59.44$, $S_{x_2}^2 = 54.46, S_{yx_1} = -0.1300$, $S_{yx_2} = -0.2152$, $S_{x_1x_2} = 0.2610$, where the variables are described as

$\boldsymbol{Y = Alkaline\ Phosphatase}$

$\boldsymbol{X_1 = Age}$

$\boldsymbol{X_2 = Weight}$

In addition to the above information, we choose the sample size to be 7, i.e.,$n = 7$. We present below the MSEs of the competing estimators:

Table 2: Efficiency of the competing estimators

$\boldsymbol{Estimator}$	$\boldsymbol{MSE}$
$\bar{y}_{MP}$	3.55
$\bar{y}_{alcP}$	3.19

The above findings once again establish the superiority of the proposed estimator over the multivariate product estimator indicating 11% gain in efficiency approximately.

REFERENCES

[1]. Das, B.C.(1966): Multivariate investigations of blood chemistry and morphology.

[2]. Proceedings of the symposium on Human Adaptability to Environments and Physical Fitness, Defence Institute of Physiology and Allied Sciences, Madras.

[3]. Kadilar, C. and Cingi, H.(2005): A new estimator using two auxiliary variables, Appl. Math. Comp., 162;901-908.

[4]. Lu, Jingli(2013): The chain ratio estimator and regression estimator with linear combination of two auxiliary variables, PLOS ONE, www.plosone.org, 1-4.

[5]. Murthy, M.N.(1964): Product method of estimation, Sankhya,26, 69-74.

[6]. Weisberg, S.(1980): Applied Linear Regression, john Wiley & Sons, USA