Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER Applied Mathematics and Computation 162 (2005) 901–908
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

A new estimator using two auxiliary variables

Cem Kadilar *, Hulya Cingi

Department of Statistics, Hacettepe University, Beytepe, Ankara 06532, Turkey

Abstract

Utilizing the estimator in Abu-Dayyeh et al. [Appl. Math. Comput. 139 (2003) 287], we suggest an estimator using two auxiliary variables in simple random sampling. We obtain mean square error (MSE) equation of the proposed estimator and theoretically compare it with the MSE of the traditional estimator using two auxiliary variables. By this comparison, we show the condition that the proposed estimator is more efficient than the traditional one. In addition, we support this theoretical result by an application with original data.

Keywords: Ratio estimator; Regression estimator; Auxiliary variable; Mean square error; Efficiency

1. Introduction

Suppose that an auxiliary variate x_i, correlated with variate of interest y_i, is obtained for each unit in the sample which is drawn by simple random sampling and that the population mean $\overline{X}$ of the x_i is known. The regression estimate of $\overline{Y}$, the population mean of the y_i, is

$$\bar{y}_{\text{reg}\,1} = \bar{y} + b(\overline{X} - \bar{x}),$$

where b is an estimate of the change in y when x is increased by unity, $\bar{x}$ and $\bar{y}$ are the sample means of the x_i and y_i, respectively. MSE of this estimate is

$$\text{MSE}(\bar{y}_{\text{reg}\,1}) = \frac{1-f}{n} S_y^2 (1 - \rho_{yx}^2),$$

* Corresponding author.
E-mail addresses: kadilar@hacettepe.edu.tr (C. Kadilar), hcingi@hacettepe.edu.tr (H. Cingi).

0096-3003/$ - see front matter
doi:10.1016/j.amc.2003.12.130

where $f = \frac{n}{N}$, n is the sample size, N is the population size, S_y^2 and S_x^2 are the population variances of the y_i and x_i respectively, $\rho_{yx} = \frac{S_{yx}}{S_y S_x}$ is the population correlation coefficient between y_i and x_i, S_{yx} is the population covariance between y_i and x_i [2].

When there are two auxiliary variates as X_{1i} and X_{2i}, the regression estimate of $\overline{Y}$ will be

$$\bar{y}_{\mathrm{reg}\,2} = \bar{y} + b_1(\overline{X}_1 - \bar{x}_1) + b_2(\overline{X}_2 - \bar{x}_2), \tag{1.1}$$

where $b_1 = \frac{s_{yx_1}}{s_{x_1}^2}$ and $b_2 = \frac{s_{yx_2}}{s_{x_2}^2}$. Here, $s_{x_1}^2$ and $s_{x_2}^2$ are the sample variances of the x_{1i} and x_{2i}, respectively, s_{yx_1} and s_{yx_2} are the sample covariances between y_i and x_{1i} and between y_i and x_{2i}, respectively.

MSE of this estimate can be found as

$$\mathrm{MSE}(\bar{y}_{\mathrm{reg}\,2}) \cong \frac{1-f}{n} S_y^2 \left(1 - \rho_{yx_1}^2 - \rho_{yx_2}^2 + 2\rho_{yx_1}\rho_{yx_2}\rho_{x_1x_2}\right). \tag{1.2}$$

(see Appendix A).

2. The suggested estimator

Abu-Dayyeh et al. [1] proposed the following estimator of the population mean assuming that the population means $\overline{X}_1$ and $\overline{X}_2$ of the auxiliary variables were known:

$$\bar{y}_{\mathrm{ratio}} = \bar{y}\left(\frac{\overline{X}_1}{\bar{x}_1}\right)^{\alpha_1}\left(\frac{\overline{X}_2}{\bar{x}_2}\right)^{\alpha_2}, \tag{2.1}$$

where α_1 and α_2 were real numbers.

We suggest using the ratio estimator given in (2.1) instead of $\bar{y}$ in (1.1). By this way, we obtain the following estimator:

$$\bar{y}_{\mathrm{pr}} = \bar{y}\left(\frac{\overline{X}_1}{\bar{x}_1}\right)^{\alpha_1}\left(\frac{\overline{X}_2}{\bar{x}_2}\right)^{\alpha_2} + b_1(\overline{X}_1 - \bar{x}_1) + b_2(\overline{X}_2 - \bar{x}_2). \tag{2.2}$$

MSE of this estimator can be found using Taylor series method defined as

$$h(\bar{x}_1, \bar{x}_2, \bar{y}) \cong h(\overline{X}_1, \overline{X}_2, \overline{Y}) + \left.\frac{\partial h(c,d,e)}{\partial c}\right|_{\overline{X}_1,\overline{X}_2,\overline{Y}} (\bar{x}_1 - \overline{X}_1) + \left.\frac{\partial h(c,d,e)}{\partial d}\right|_{\overline{X}_1,\overline{X}_2,\overline{Y}} (\bar{x}_2 - \overline{X}_2) + \left.\frac{\partial h(c,d.e)}{\partial e}\right|_{\overline{X}_1,\overline{X}_2,\overline{Y}} (\bar{y} - \overline{Y})$$

(see [3]).