ON THE EFFICIENCY OF A DUAL TO RATIO-CUM-PRODUCT ESTIMATOR IN SAMPLE SURVEYS

Housila P. Singh and Rajesh Singh
School of Studies in Statistics, Vikram University, Ujjain MP, India

Mariano Ruiz Espejo* and M. Delgado Pineda
Departamento de Matemáticas Fundamentales, Universidad Nacional de Educación a Distancia, Facultad de Ciencias, Madrid, Spain

and

Saralees Nadarajah
Department of Mathematics, University of South Florida, Tampa FL, USA

[Received 31 July 2002. Read 12 July 2005. Published 21 July 2005.]

Abstract

This article suggests a dual to ratio-cum-product estimator for finite population mean that is complementary, in a certain sense, to the commonly used ratio-cum-product estimator reported by Singh [4]. The expressions for bias and mean square error of the proposed estimator have been obtained to the first degree of approximation. This study also generalises the work of Srivenkataramana [5] and Bandyopadhyay [1]. To illustrate the results an empirical study is carried out.

1. Introduction

Let U be a finite population consisting of N units $u_1, u_2, ..., u_N$. The units of this finite population are identifiable in the sense that they are uniquely labelled from 1 to N and the label of each unit is known. Let y and (x, z) denote the study variate and auxiliary variates taking the values y_i and (x_i, z_i), respectively, on the unit u_i $(i = 1, 2, ..., N)$, where x is positively correlated with y and z is negatively correlated with y. We wish to estimate the population mean $\bar{Y} = (1/N)\sum_{i=1}^{N} y_i$ of y, assuming that the population means $\left(\bar{X}, \bar{Z}\right)$ of (x, z) are known. Assume that a simple random sample of size n is drawn without replacement from U. The usual ratio and product estimators for $\bar{Y}$ are

$$\bar{y}_R = \bar{y}(\bar{X}/\bar{x})$$

and

$$\bar{y}_P = \bar{y}(\bar{z}/\bar{Z}),$$

*Corresponding author, e-mail: ruizespejo@terra.es

respectively, where $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$, $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$ and $\bar{z} = (1/n)\sum_{i=1}^{n} z_i$ are the sample means of y, x and z respectively.

Singh [4] improved the ratio and product methods of estimation and suggested the 'ratio-cum-product' estimator for $\bar{Y}$ as

$$\bar{y}_{RP} = \bar{y}\frac{\bar{X}}{\bar{x}}\frac{\bar{z}}{\bar{Z}}.$$

The variance of $\bar{y}$ under simple random sampling without replacement (SRSWOR) design is

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

and to the first degree of approximation, the mean square errors (MSEs) of $\bar{y}_R$, $\bar{y}_P$ and $\bar{y}_{RP}$ are, respectively, given by

$$MSE(\bar{y}_R) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2 + C_x^2(1-2K_{yx})\right], \tag{2}$$

$$MSE(\bar{y}_P) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2 + C_z^2(1+2K_{yz})\right], \tag{3}$$

$$MSE(\bar{y}_{RP}) = \frac{1-f}{n}\bar{Y}^2\left[C_y^2 + C_z^2(1-2K_{yz}) + C_x^2(1-2K)\right], \tag{4}$$

where

$$f = n/N,\ \ C_y = S_y/\bar{Y},\ \ C_x = S_x/\bar{X},\ \ C_z = S_z/\bar{Z},$$

$$K_{yx} = \rho_{yx}C_y/C_x,\ \ K_{yz} = \rho_{yz}C_y/C_z,\ \ K_{zx} = \rho_{xz}C_z/C_x,$$

$$K = K_{yx} + K_{zx},$$

$$S_y^2 = \sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2/(N-1),$$

$$S_x^2 = \sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2/(N-1),\ \ S_z^2 = \sum_{i=1}^{N}\left(z_i - \bar{Z}\right)^2/(N-1),$$

$$\rho_{yx} = S_{yx}/\left(S_yS_x\right),\ \ \rho_{yz} = S_{yz}/\left(S_yS_z\right),\ \ \rho_{xz} = S_{xz}/\left(S_xS_z\right),$$

$$S_{yx} = \sum_{i=1}^{N}\left(y_i - \bar{Y}\right)\left(x_i - \bar{X}\right)/(N-1),$$

$$S_{yz} = \sum_{i=1}^{N}\left(y_i - \bar{Y}\right)\left(z_i - \bar{Z}\right)/\left(N-1\right)$$

and

$$S_{xz} = \sum_{i=1}^{N} \left(x_i - \bar{X}\right) \left(z_i - \bar{Z}\right) / (N-1).$$

In this article, using a simple transformation, a dual to usual ratio-cum-product estimator is proposed and its properties are studied. Numerical illustrations are given to show the performance of the constructed estimator over other estimators.

2. The suggested estimator

Let $x_i^* = (1+g)\bar{X} - gx_i$ and $z_i^* = (1+g)\bar{Z} - gz_i$, $i = 1, 2, ..., N$, where $g = n/(N-n)$. Then clearly

$$\bar{x}^* = (1+g)\bar{X} - g\bar{x}, \;\; \bar{z}^* = (1+g)\bar{Z} - g\bar{z}$$

are also unbiased estimators for $\bar{X}$ and $\bar{Z}$ respectively and $Corr(\bar{y}, \bar{x}^*) = -\rho_{yx}$ and $Corr(\bar{y}, \bar{z}^*) = -\rho_{yz}$. With this notation, we suggest an estimator for $\bar{Y}$ as

$$\bar{y}_{RP}^* = \bar{y} \frac{\bar{x}^*}{\bar{X}} \frac{\bar{Z}}{\bar{z}^*}.$$

The *MSE* of $\bar{y}_{RP}^*$ to the first degree of approximation is given by

$$MSE\left(\bar{y}_{RP}^*\right) = \frac{1-f}{n} \bar{Y}^2 \left[C_y^2 + gC_z^2(g + 2K_{yz}) + gC_x^2\left(g - 2gK_{zx} - 2K_{yx}\right)\right]. \tag{5}$$

Remark 2.1. When information on the auxiliary variable z is not used (or variable z takes the value 'unity'), the estimator $\bar{y}_{RP}^*$ reduces to the 'dual to ratio' estimator

$$\bar{y}_R^* = \bar{y} \frac{\bar{x}^*}{\bar{X}}.$$

It reduces to the 'dual to product' estimator

$$\bar{y}_P^* = \bar{y} \frac{\bar{Z}}{\bar{z}^*}$$

if the information on the auxiliary variate x is not used. The estimators $\bar{y}_R^*$ and $\bar{y}_P^*$ are due to Srivenkataramana [5] and Bandyopadhyay [1], respectively.

The *MSE*s of $\bar{y}_R^*$ and $\bar{y}_P^*$, to the first degree of approximation, are, respectively, given by

$$MSE\left(\bar{y}_R^*\right) = \frac{1-f}{n} \bar{Y}^2 \left[C_y^2 + gC_x^2(g - 2K_{yx})\right] \tag{6}$$

and

$$MSE\left(\bar{y}_P^*\right) = \frac{1-f}{n} \bar{Y}^2 \left[C_y^2 + gC_z^2(g + 2K_{yz})\right]. \tag{7}$$

3. Efficiency comparisons

It follows from (1), (4) and (5) that $\bar{y}^*_{RP}$ is more efficient than the usual unbiased estimator $\bar{y}$ or Singh's [4] estimator $\bar{y}_{RP}$ when

$$\frac{1}{2}g < A < \frac{1}{2}(1+g), \tag{8}$$

regarded as a condition on A, where

$$A = \frac{K_{yx}C_x^2 - K_{yz}C_z^2}{C_x^2 + C_z^2 - 2K_{zx}C_x^2} \ (>0).$$

While establishing the condition (8) we assume that $1-g>0$, that is $N>2n$, which may be taken as a typical survey situation.

We note from (1) and (4) that $\bar{y}_{RP}$ dominates over the usual unbiased estimator $\bar{y}$ if

$$A > \frac{1}{2}. \tag{9}$$

Since $g = n/(N-n)$ is small (see Srivenkataramana [5], for instance), it follows from (8) and (9) that for the most part $\bar{y}^*_{RP}$ is superior, in terms of mean square error, to $\bar{y}$ just when $\bar{y}_{RP}$ is inferior to $\bar{y}$. In this sense $\bar{y}^*_{RP}$ and $\bar{y}_{RP}$ are complementary. In general, (8) specifies an interval of length $\frac{1}{2}$ for A. This interval moves to the right from $(0,\frac{1}{2})$ to $(\frac{1}{2},1)$ as sample size n is increased from 0 to $N/2$.

We note from (5) and (6) that $\bar{y}^*_{RP}$ will beat $\bar{y}^*_R$ if

$$\text{either} \ \ K_{yz} < -g\left(\frac{1}{2} - K_{xz}\right); \ \ K_{xz} < \frac{1}{2};$$

$$\text{or} \ \ K_{yz} > -g\left(\frac{1}{2} - K_{xz}\right); \ \ K_{xz} > \frac{1}{2};$$

where $K_{xz} = \rho_{xz}C_x/C_z$. It follows from (5) and (7) that the estimator $\bar{y}^*_{RP}$ would be better than $\bar{y}^*_P$ if

$$\text{either} \ \ K_{yx} > g\left(\frac{1}{2} - K_{zx}\right); \ \ K_{zx} < \frac{1}{2};$$

$$\text{or} \ \ K_{yz} < g\left(\frac{1}{2} - K_{zx}\right); \ \ K_{zx} > \frac{1}{2}.$$

It is observed from (2) and (5) that $MSE(\bar{y}^*_{RP}) < MSE(\bar{y}_R)$ if

$$C_x^2(1-g)(1+g-2K_{yz}) < gC_z^2\left[g + 2\left(K_{yz} - gK_{xz}\right)\right]$$

with $K_{yz} > \frac{1}{2}(1+g)$ and $K_{yz} > -g\left(\frac{1}{2} - K_{xz}\right)$. Further, we note from (3) and (5)

TABLE 1—Description of population II.

x	y	z
49	35	200
40	35	212
41	38	211
46	40	212
52	40	203
59	42	194
53	44	194
61	46	188
55	50	196
64	50	190

that the estimator $\bar{y}_{RP}^{*}$ is more efficient than $\bar{y}_P$ if

$$C_z^2(1-g)(1+g+2K_{yz}) < C_x^2 g\,[g\,(1-2K_{zx}) - 2K_{yx}]$$

with $K_{yz} > -\frac{1}{2}(1+g)$ and $K_{yx} < g\left(\frac{1}{2} - K_{zx}\right)$.

4. Empirical study

In this section we illustrate the performance of the constructed estimator $\bar{y}_{RP}^{*}$ over various other estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}_R^{*}$, $\bar{y}_P$, $\bar{y}_P^{*}$ and $\bar{y}_{RP}$ through two populations of natural data.

(1) Population I (Source: Singh [4, p. 377]; a detailed description can be seen in Singh [3])
y: Number of females employed
x: Number of females in service
z: Number of educated females

$\bar{Y} = 7.46$, $\bar{X} = 5.31$, $\bar{Z} = 179.00$, $C_y^2 = 0.5046$, $C_x^2 = 0.5737$, $C_z^2 = 0.0633$, $\rho_{yx} = 0.7737$, $\rho_{yz} = -0.2070$, $\rho_{xz} = -0.0033$, $N = 61$ and $n = 20$.

(2) Population II (Source: Johnston [2, p. 171])
y: Percentage of hives affected by disease
x: Mean January temperature
z: Date of flowering of a particular summer species (number of days from January 1)

$\bar{Y} = 52$, $\bar{X} = 42$, $\bar{Z} = 200$, $C_y^2 = 0.0244$, $C_x^2 = 0.0170$, $C_z^2 = 0.0021$, $\rho_{yx} = 0.80$, $\rho_{yz} = -0.94$, $\rho_{xz} = -0.73$, $N = 10$ and $n = 4$.

A detailed description of these variables is shown in Table 1.

TABLE 2—*PRE*s $(\cdot, \bar{y})$ of various estimators of $\bar{Y}$ with respect to $\bar{y}$.

Estimator	*Population I*	*Population II*
$\bar{y}$	100.00	100.00
$\bar{y}_R$	204.96	276.85
$\bar{y}_R^*$	215.00	238.47
$\bar{y}_P$	102.17	187.00
$\bar{y}_P^*$	104.34	149.13
$\bar{y}_{RP}$	213.09	394.52
$\bar{y}_{RP}^*$	235.68	401.87

The percent relative efficiencies (*PRE*s) of the different estimators with respect to the usual unbiased estimator $\bar{y}$ computed by the formula

$$PRE\,(\cdot, \bar{y}) = \frac{V\,(\bar{y})}{MSE\,(\cdot)} \times 100$$

and are presented in Table 2.

Table 2 shows clearly that the proposed dual to ratio-cum-product estimator $\bar{y}_{RP}^*$ is more efficient than all the estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}_R^*$, $\bar{y}_P$, $\bar{y}_P^*$ and $\bar{y}_{RP}$.

REFERENCES

[1] S. Bandyopadhyay, Improved ratio and product estimators, *Sankhyā Series C* **42** (1980), 45−9.
[2] J. Johnston, *Econometric methods*, (2nd edn), McGraw-Hill, Tokyo, 1972.
[3] M.P. Singh, On the estimation of ratio and product of the population parameters, *Sankhyā Series B* **27** (1965), 321−8.
[4] M.P. Singh, Comparison of some ratio-cum-product estimators, *Sankhyā Series B* **31** (1969), 375−8.
[5] T. Srivenkataramana, A dual to ratio estimator in sample surveys, *Biometrika* **67** (1980), 199−204.