COMPARISON OF SOME RATIO-CUM-PRODUCT ESTIMATORS

By M. P. SINGH
Indian Statistical Institute

SUMMARY. Two estimators using information on two supplementary characters are suggested for estimating the ratio of population means (or totals). Approximate expressions for their bias and mean square error are obtained and compared with those for the usual ratio estimator and the estimators suggested by the author (1965). An empirical study is included for illustration.

1. Introduction

Let Y_{ij} be the value of the i-th unit, in the population $U = (U_1, U_2, ..., U_i, ..., U_N)$ for the j-th character y_j $(j = 0, 1, 2, 3)$, of which y_0 and y_1 are the characters under investigation and y_2 and y_3 are the supplementary characters defined over U. Let $\bar{y}_j$ be the usual unbiased estimator of the corresponding population mean $\bar{Y}_j$, based on a sample of size n. More efficient estimators of $\bar{Y}_j$, using the usual estimators of the ratio and product are well known in the literature. In an earlier paper the author (1965) considered estimation of the ratio $R = \bar{Y}_0/\bar{Y}_1$ and the product $P = \bar{Y}_0.\bar{Y}_1$ themselves, the usual estimators of which are respectively $r = \bar{y}_0/\bar{y}_1$ and $p = \bar{y}_0.\bar{y}_1$, and proposed estimators which utilise information on y_2 (or y_3). These estimators for R and P are of the form

$$R^* = r(\bar{y}_2/\bar{Y}_2)^{\alpha_2} \quad \text{and} \quad P^* = p(\bar{y}_2/\bar{Y}_2)^{\beta_2} \qquad \dots \quad (1)$$

respectively, where α_2 and β_2 are constants to be determined by minimising the mean square error of the corresponding estimators. Thus the optimum values of α_2 and β_2 are respectively given by

$$\alpha_2^* = (c_1/c_2)\partial_{12} - (c_0/c_2)\partial_{02} \qquad \dots \quad (1.2)$$

and

$$\beta_2^* = -(c_1/c_2)\partial_{12} - (c_0/c_2)\partial_{02} \qquad \dots \quad (1.3)$$

where c_j denotes coefficient of variation of y_j and $\partial_{jj'}$ the correlation coefficient between y_j and $y_{j'}$ $(j \neq j' = 0, 1, 2, 3)$.

It is easily seen that the estimators R^* and P^* are more efficient than r and p respectively if the conditions $\alpha_2^* > 1/2$ and $\beta_2^* > 1/2$ are satisfied. The estimators $R_1^* (= r.\bar{y}_2/\bar{Y}_2)$ and $R_2^* (= r.\bar{Y}_3/\bar{y}_3)$ for R and $P_1^* (= p.\bar{y}_2/\bar{Y}_2)$ and $P_2^* (= p.\bar{Y}_3/\bar{y}_3)$ for P, proposed earlier by the author (1965), assume $\alpha_2 = \beta_2 = 1$ for R_1^* and P_1^* and $\alpha_3 = \beta_3 = -1$ for R_2^* and P_2^* where α_3 and β_3 have definitions similar to α_2 and β_2 given above.

In the present paper we consider estimators

$$R_c^* = r\left(\frac{\bar{y}_2}{\bar{Y}_2}\right)^{\alpha_2}\left(\frac{\bar{y}_3}{\bar{Y}_3}\right)^{\alpha_3} \qquad \dots \quad (1.4)$$

and

$$R_c^{*\prime} = w_1 r\left(\frac{\bar{y}_2}{\bar{Y}_2}\right)^{\alpha_2} + w_2 r\left(\frac{\bar{y}_3}{\bar{Y}_3}\right)^{\alpha_3} \qquad \dots \quad (1.5)$$

which utilise information on both y_2 and y_3. The weights w_1 and w_2 above are such that $w_1+w_2 = 1$ and α_1 and α_2 are constants to be suitably chosen. It is pertinent to note that the usual double ratio estimator (say, R^*_d) and an estimator (say, R^*_w), suggested earlier by the author (see, Murthy, 1967, p. 404), are special cases of R^*_c and $R^{*\prime}_c$ respectively for $\alpha_2 = 1$ and $\alpha_3 = -1$.

The bias and mean square error (m.s.e.) of R^*_c and $R^{*\prime}_c$ are obtained and m.s.e. of R^*_c is compared with that of R^* using information either on y_2 (or y_3). An empirical study is also included. The estimation of product $P = \bar{Y}_0.\bar{Y}_1$ and also of the population mean $\bar{Y}_0$ itself, if $\bar{Y}_1$ is known, can be dealt with in a similar manner.

2. Bias and mean square error

We shall assume that n units in the sample have been selected with equal probability and with replacement. Writing

$$\bar{y}_j = \bar{Y}_j(1+e_j), \quad j = 0, 1, 2, 3 \qquad \dots \quad (2.1)$$

where $E(e_j) = 0$ and assuming that for large values of n, $|e_j|$ is less than unity for $j = 1$ and 3, the bias and m.s.e. of R^*_c and $R^{*\prime}_c$, respectively, to order n^{-1}, are given by

$$B(R^*_c) = B(r)+Rn^{-1}\left(d_2+d_3+\alpha_2\alpha_3c_{23}+\frac{\alpha_2(\alpha_2-1)}{2}c_2^2+\frac{\alpha_3(\alpha_3-1)}{2}c_3^2\right) \qquad \dots \quad (2.2)$$

$$B(R^{*\prime}_c) = B(r)+Rn^{-1}\left(w_1\left(d_2+\frac{\alpha_2(\alpha_2-1)}{2}c_2^2\right)+w_2\left(d_3+\frac{\alpha_3(\alpha_3-1)}{2}c_3^2\right)\right) \qquad \dots \quad (2.3)$$

$$M(R^*_c) = M(r)+R^2n^{-1}(\alpha_2^2c_2^2+\alpha_3^2c_3^2+2(d_2+d_3+\alpha_2\alpha_3c_{23})) \qquad \dots \quad (2.4)$$

$$M(R^{*\prime}_c) = M(r)+R^2n^{-1}(w_1^2\alpha_2^2c_2^2+w_2^2\alpha_3^2c_3^2+2(w_1d_2+w_2d_3+w_1w_2d_2d_3c_{23})) \qquad \dots \quad (2.5)$$

where

$$d_2 = \alpha_2(c_{02}-c_{12}), \quad d_3 = \alpha_3(c_{03}-c_{13}),$$

$$c_{jj'} = c_jc_{j'}\,\partial_{jj'}, \quad \text{and} \quad B(r) = Rn^{-1}(c_1^2-c_{01})$$

and

$$M(r) = Rn^{-1}(c_0^2+c_1^2-2c_{01})$$

are bias and m.s.e. of r to order n^{-1}.

The optimum weights w_1 and w_2 may be determined by minimising (2.5), under the condition $w_1+w_2 = 1$, and we get

$$w_1 = \frac{\alpha_3^2c_3^2-d_2+d_3-\alpha_2\alpha_3c_{23}}{\alpha_2^2c_2^2+\alpha_3^2c_3^2-2\alpha_2\alpha_3c_{23}} = 1-w_2. \qquad \dots \quad (2.6)$$

The optimum values of α_2 and α_3 in (1.4) and (1.5) may be obtained by minimising the m.s.e. in (2.4) and (2.5) using optimum weights in (2.6). But in practice it would be difficult to get the exact optimum values of α_2 and α_3 as they involve many unknown parameters. However, in situations where good guessed values of c_i and

∂_{ij} ($i = 0, 1, 2$) are available, α_2^* (and corresponding α_3^*) given by (1.2) may be used as approximations to these optimum values. In that case, it is observed that R_c^* is more efficient than $R^*_{(\alpha_2^*)}$ if

$$\frac{\alpha_2^* c_2}{\alpha_3^* c_3} \partial_{23} < \frac{1}{2} \qquad \ldots \quad (2.7)$$

and that it is more efficient than $R^*_{(\alpha_3^*)}$ if

$$\frac{\alpha_3^* c_3}{\alpha_2^* c_2} \partial_{23} < \frac{1}{2} \qquad \ldots \quad (2.8)$$

where $R^*_{(\alpha_2^*)}$ and $R^*_{(\alpha_3^*)}$ are values of R^* using α_2^* (with y_2) and α_3^* (with y_3) respectively. Thus it is expected that R_c^* will improve over both $R^*_{(\alpha_2^*)}$ and $R^*_{(\alpha_3^*)}$ if the magnitude of ∂_{23} is quite small. Comparison of $R_c^{*\prime}$ with $R^*_{(\alpha_2^*)}$, $R^*_{(\alpha_3^*)}$ and R_c^* is not attempted here since that does not lead to practically usuable conclusions. However, a comparison of these estimators have been made in the next section on the basis of an empirical study.

3. An empirical study

In this study we compare different ratio-cum-product estimators among themselves and with the usual ratio estimator. The population under consideration is same which was used by the author (1965). That is, the data for all 61 blocks of Ahmedabad City, ward No. 1 (khadia I) taken from 1951 Population Census will be considered. The characters y_0, y_1, y_2 and y_3 are females employed, female population, educated females and females in services respectively. The purpose is to estimate the ratio (R) of females employed to total female population. For this population, we have

$\bar{Y}_0 = 7{\cdot}46$	$c_0^2 = 0{\cdot}5046$	$\partial_{01} = 0{\cdot}0388$
$\bar{Y}_1 = 265{\cdot}54$	$c_1^2 = 0{\cdot}0379$	$\partial_{02} = -0{\cdot}2070$
$\bar{Y}_2 = 179{\cdot}00$	$c_2^2 = 0{\cdot}0633$	$\partial_{03} = 0{\cdot}7737$
$\bar{Y}_3 = 5{\cdot}31$	$c_3^2 = 0{\cdot}5737$	$\partial_{12} = 0{\cdot}7373$

$$\partial_{13} = -0{\cdot}0474 \quad \text{and} \quad \partial_{23} = -0{\cdot}0033.$$

Using the relation (1.2), we get

$$\alpha_2^* = 1.1718 \text{ and } \alpha_3^* = 0.7379$$

which on substitution in (2·6) gives the optimum weights

$$w_1 = 0{\cdot}2124 \text{ and } w_2 = 0{\cdot}7876.$$

Further, on using $\alpha_2 = 1$ and $\alpha_3 = -1$, these weights are

$$w_1 = 0{\cdot}3493 \quad \text{and} \quad w_2 = 0{\cdot}6507.$$

Now the m.s.e. of the estimators suggested here can easily be calculated by using the above weights. Table 1 gives the efficiency of the ratio-cum-product estimators for estimating the ratio R.

TABLE 1. RELATIVE EFFICIENCY OF THE ESTIMATORS

estimators	% efficiency	estimators	% efficiency
r	100	$R_c^{*'}$	265
R_1^*	118	$R_w^{*'}$	309
$R^*_{(\alpha_2^*)}$	119	R_d^*	301
R_2^*	206	R_c^*	391
$R^*_{(\alpha_3^*)}$	243		

From Table 1 it is observed that the gains in efficiency of $R^*_{(\alpha_3^*)}$ and R_c^* over R_2^* and R_d^* (double ratio estimator, Rao (1957) and Keyfitz, see Yates, 1960) are about 18% and 30% respectively, but $R^*_{(\alpha_2^*)}$ has virtually same efficiency as R_1^* as α_2^* is very close to unity. $R_c^{*'}$ using w_1 and w_2 is the least efficient among the estimators using both y_2 and y_3. Efficiencies of $R_w^{*'}$ and R_d^* are about same. Thus R_1^* or R_2^* (using y_2 or y_3) and R_d^* (using both y_2 and y_3), which do not depend on α_2 or α_3, may be preferred in practice to the corresponding estimators which use optimum weights unless very good guessed values of α_2^* and α_3^*, are available.

The author is grateful to the referee for very useful suggestions.

REFERENCES

MURTHY, M. N. (1967) : *Sampling Theory and Methods*, Statistical Publishing Society, Calcuttta

RAO, J. N. K. (1957) : Double ratio estimate in forest surveys. *Jour. Ind. Soc. Agr. Stat.*, **9**, 191-204.

SINGH, M. P. (1965) : On the estimation of ratio and product of the population parameters. *Sankhyā*, Series B, **27**, 321-328.

——— (1967) : Ratio cum product method of estimation. *Metrika*, **12**(1), 34-43.

YATES, F. (1960) : *Sampling Methods for Censuses and Surveys*, 3rd Edn. Griffin, London.

Paper received : October, 1967.
Revised : April, 1968.