Jour. Ind. Soc. Ag. Statistics
Vol. XXXVIII, No. 3 (1986), pp. 383-387

A NOTE ON THE EFFICIENCY OF A PRODUCT-TYPE ESTIMATOR UNDER A SUPER-POPULATION MODEL

L. N. SAHOO
O. U. A. T., Bhubaneswar-751003

(Received : December, 1983)

SUMMARY

In a recent paper Srivenkataramana [5] suggested a new product-type estimator which is complementary, in a certain sense, to the traditional ratio estimator. In this note some results are derived concerning the efficiency of his estimator over the mean per unit estimator and the traditional ratio estimator under a super-population model.

Keywords : Bias; Efficiency; Finite population; Product-type estimator; Ratio estimator; Super-population model.

Introduction

Consider a finite population with N units and let x_i and y_i denote the values for two positively correlated characters x and y respectively for the ith unit in this population, $i = 1, 2, \ldots, N$. Assume that the population mean $\bar{X}$ of x is known. Let $\bar{x}$ and $\bar{y}$ be the sample means of x and y respectively based on a srswor of n $(n < N)$ units. Then the traditional ratio estimator $\bar{y}_r = \bar{y}\,(\bar{X}/\bar{x})$ is used to improve upon the simple unbiased estimator $\bar{y}$ as an estimator of the population mean $\bar{Y}$ of y when the correlation between x and y is highly positive.

Srivenkataramana [5] proposed a new product-type estimator defined by

$$\bar{y}_a = \bar{y}\,\frac{N\bar{X} - n\bar{x}}{(N-n)\bar{X}}$$

which is more efficient than $\bar{y}$ or $\bar{y}_r$ when

$$\frac{K}{2} < \rho \frac{C_y}{C_x} < \frac{1+K}{2}$$

where C_x and C_y denote the coefficient of variations of x and y respectively, ρ denotes the correlation coefficient between x and y, and $K = n/N - n$. Since, $\bar{y}_r$ is less efficient than $\bar{y}$ when $\rho(C_y/C_x) < \frac{1}{2}$, Srivenkataramana suggested that $\bar{y}_a$ is complementary to $\bar{y}_r$ in the sense that, when $\bar{y}_r$ is inferior to $\bar{y}$ it is also inferior to $\bar{y}_a$.

In the present note the efficiencies of these strategies under a super-population model concerning the relationship between y and x are compared. The results indicate that in most cases $\bar{y}_a$ is superior to $\bar{y}$ and inferior to $\bar{y}_r$.

Following Cochran [2] it is assumed that the finite population under consideration is itself a random sample from a hypothetical super-population for which

$$y_i = \beta x_i + e_i \;\forall\; i \quad \text{(I)}$$

such that $\mathcal{E}(e_i/x_i) = 0$,

$\mathcal{E}(e_i^2/x_i) = \delta x_i^g$,

and $\mathcal{E}(e_i e_j/x_i, x_j) = 0$ for $i \neq j$

where $0 < \delta < \infty$, $0 \leqslant g \leqslant 2$ and $\mathcal{E}$ denotes the expectation with respect to the distribution of e_i's over the hypothetical super-population. Further, following Durbin (1959), Tin [6] Rao and Webster [4] and others, assume that x_i's are i. d. gamma variates with parameter h.

Efficiency Comparisons

The usual expression for the bias and the approximate expression for the variance of $\bar{y}_a$ based on srswor scheme are

$$B(\bar{y}_a) = -\frac{1}{N(N-1)\bar{X}} \sum_{i=1}^{N} y_i (x_i - \bar{X})$$

and $$V(\bar{y}_a) = \frac{N-n}{Nn(N-1)} \sum_{i=1}^{N} [y_i - KRx_i - (1-K)\bar{Y}]^2,$$

where $R = \bar{Y}/\bar{X}$.

Unlike $\bar{y}_r$, the bias of $\bar{y}_a$ is independent of n and hence it is not a consistent estimator of $\bar{Y}$.

The expressions for the bias and variance of $\bar{y}_a$ under (I) using Rao and Webster's [4] theorem are obtained as follows:

Let $z_1, z_2, \ldots, z_n$ be independent gamma variates with parameter h. Then for $i \neq j$

$$E\left[\frac{z_i^a z_j^b}{\left(\sum_{i=1}^{n} z_i\right)^c}\right] = \frac{\Gamma(a+h)\,\Gamma(b+h)}{(\Gamma(h))^2}\;\frac{1}{\prod_{t=1}^{c}(nh+a+b-t)}, \tag{1}$$

where a, b and c are non-negative integers.

Using relation (1) under the model we get

$$\mathscr{E}B(\bar{y}_a) = -\frac{\beta h}{Nh+1} \tag{2}$$

and
$$\mathscr{E}V(\bar{y}_a) = \frac{N-n}{Nn}\Bigg[(1-K)^2\beta^2 h + \frac{\delta\,\Gamma(h+g)}{\Gamma(h)\,(Nh+g)\,(Nh+g+1)}$$
$$\{Nh(Nh+2) + Nh(2g-K-1)(1-K) + g(1+g)(1-K)^2\}\Bigg], \tag{3}$$

Again, under (I) $\bar{y}_r$ is unbiased and

$$\mathscr{E}V(\bar{y}) = \frac{N-n}{Nn}\left[\beta^2 h + \frac{\delta\,\Gamma(h+g)}{\Gamma(h)}\right], \tag{4}$$

and following Arnab (1979)

$$\mathscr{E}V(\bar{y}_r) = \frac{N-n}{Nn}\left[\frac{\delta\,\Gamma(h+g)\,Nh\,(Nh+2)}{\Gamma(h)\,(Nh+g)\,(Nh+g+1)}\right]. \tag{5}$$

From (3) and (4) we find that

$$\mathscr{E}V(\bar{y}) - \mathscr{E}V(\bar{y}_a) = \frac{1}{N}\Bigg[(2-K)\,\beta^2 h$$
$$+ \frac{\delta\,\Gamma(h+g)\,\{Nh\,(2g-K) + g\,(1+g)\,(2-K)\}}{\Gamma(h)\,(Nh+g)\,(Nh+g+1)}\Bigg]. \tag{6}$$

Hence, from (6), $\bar{y}_a$ is more efficient than $\bar{y}$ if $K \leqslant 2$ and $g > K/2$ i. e. if $n \leqslant 2/3N$ and $g > n/2(N-n)$ and $\bar{y}_a$ is less efficient than $\bar{y}$ if $K \geqslant 2$ and $g < K/2$. $\bar{y}_a$ may also be more efficient than $\bar{y}$ even when $K < 2$ and $g < K/2$. When $g \geqslant \frac{1}{2}$, $\bar{y}_a$ is more efficient than $\bar{y}$ if $K \leqslant 1$ i.e. $n \leqslant N/2$. If g and K are negligible compared to Nh then $\bar{y}_a$ is more efficient than $\bar{y}$ whenever $K < 2$. However, when $K = 2$ and $g = 1$, the two estimators are equally efficient.

TABLE 1—RELATIVE EFFICIENCIES OF THE STRATEGIES

g	β	$n = 10$		$n = 20$		$n = 30$		$n = 40$	
		E_1	E_2	E_1	E_2	E_1	E_2	E_1	E_2
	0.5	121.95	61.09	159.93	80.13	199.56	100.0	99.58	49.89
0.0	1.0	140.45	28.15	249.94	50.09	498.96	100.0	99.83	20.00
	1.5	147.92	14.82	307.64	30.82	997.93	100.0	99.92	10.01
	0.5	110.54	81.84	124.77	91.81	135.90	100.0	99.69	73.35
0.5	1.0	126.96	52.12	179.33	73.61	243.62	100.0	99.82	40.98
	1.5	137.93	32.59	234.12	55.33	423.14	100.0	99.90	23.61
	0.5	104.24	92.46	109.26	96.91	112.73	100.0	100.00	88.70
1.0	1.0	113.70	75.64	133.52	88.83	150.31	100.0	100.00	66.53
	1.5	123.60	58.04	166.06	77.98	212.94	100.0	100.00	46.96
	0.5	101.60	97.07	103.40	98.79	104.66	100.0	100.39	95.92
1.5	1.0	105.59	89.96	112.43	95.78	117.38	100.0	100.35	85.49
	1.5	111.11	80.17	126.33	91.16	138.59	100.0	100.30	72.37
	0.5	100.65	98.66	101.40	99.39	102.02	100.0	100.82	98.82
2.0	1.0	102.09	96.11	104.48	98.37	106.21	100.0	100.78	94.89
	1.5	104.31	92.15	109.48	96.71	113.20	100.0	100.73	88.98

Again, from (3) and (5) we have

$$\mathscr{E}V(\bar{y}_a) - \mathscr{E}V(\bar{y}_r) = \frac{N-n}{Nn}\left[(1-K)^2 \beta^2 h + \frac{(1-K)\,\delta\sqrt{h+g}\,\{Nh(2g-K-1)+g(1+g)(1-K)\}}{\sqrt{h}\,(Nh+g)\,(Nh+g+1)}\right]. \tag{7}$$

So, in situations where $K < 1$ and $g \geqslant 1 + K/2$, i.e. $n < N/2$ and $g \geqslant N/2(N-n)$, $\bar{y}_r$ will be more efficient than $\bar{y}_a$. If $g \geqslant 1$, $\bar{y}_r$ is more efficient than $\bar{y}_a$ when $K < 1$. If g and K are negligible compared to Nh, $\bar{y}_r$ will be more efficient than $\bar{y}_a$. In situations where $K = 1$, $\bar{y}_r$ and $\bar{y}_a$ are equally efficient.

Remark. The conditions for the superiority of different estimators derived above are only sufficient conditions. The necessary conditions are, however, difficult to get.

Numerical Values of Ralative Efficiencies of the Strategies

Defining $E_1 = 100\ \mathscr{E}V(\bar{y})/\mathscr{E}V(\bar{y}_a)$ and $E_2 = 100\ \mathscr{E}V(\bar{y}_r)/\mathscr{E}V(\bar{y}_a)$ we present below in Table 1 the values of the relative efficiencies of $\bar{y}_a$ with respect to $\bar{y}$ and $\bar{y}_r$ for a few combinations of the parametric values under the model (I). Values are given for $N = 60$, $\delta = 2.0$ and $h = 8.0$.

From the above results we may conclude that, for the given model, $\bar{y}_a$ is not better than $\bar{y}_r$ which contradict the complementarity hypothesis of Srivenkataramana.

Acknowledgement

The author is grateful to the referees for some helpful suggestions on the earlier draft of the paper.

REFERENCES

[1] Arnab, R. (1979) : On the relative efficiencies of some sampling strategies under a super-population model, *Jour. Ind. Soc. Agri. Statist.*, **31** : 89-96.

[2] Cochran, W. G. (1963) : Sampling Techniques (Second Edition), Wiley & Sons : New York.

[3] Durbin, J. (1959) : A note on the application of Quenouille's method of bias reduction of ratios, *Biometrika*, **46** : 477-480.

[4] Rao, J. N. K. and Webster, J. T. (1966) : On two methods of bias reduction in the estimation of ratios, *Biometrika*, **53** : 571-577.

[5] Srivenkataramana, T. (1980) : A dual to ratio estimator in sample surveys, *Biometrika*, **67** : 199-204.

[6] Tin, M. (1965): Comparison of some ratio estimators, *Jour. Amer. Statist. Assoc.*, **60** : 294-307.