Jour. Ind. Soc. Ag. Statistics
***44** (3), 1992 : 267–270*

A Note on Improving Classes of Estimators in Survey Sampling

Surendra K. Srivastava
Punjabi University, Patiala
(Received : April, 1991)

Summary

Singh and Kataria [2] and Singh and Upadhyaya [1] have generalized the classes of estimators of population variance and population mean respectively, proposed earlier by Srivastava and Jhajj [3], [4] and claim that these are more efficient. Their claim is shown to be incorrect.

Keywords : Asymptotic mean square error; Auxiliary information; Finite population; Mean; Variance.

Introduction

For estimating the population variance σ_y^2 and the mean $\overline{Y}$ of a finite population utilizing the known population variance S_x^2 and the population mean $\overline{X}$ of an auxiliary variable x, Srivastava and Jhajj [3], [4] considered classes of estimators defined by

$$t_h = s_y^2\, h(u, v) \tag{1}$$

for the variance σ_y^2, and

$$\tilde{y}_t = \overline{y}\ \ t(u, v) \tag{2}$$

and $$\tilde{y}_g = g(\overline{y}, u, v) \tag{3}$$

for the mean $\overline{Y}$. Here the functions h(.,.), t(.,.) and g(.,.,.) are parametric functions satisfying certain regularity conditions and $u=\frac{s_x^2}{S_x^2}$, $v=\frac{\overline{x}}{\overline{X}}$; s_x^2 and $\overline{x}$ denote the sample variance and the sample mean of the auxiliary variable x based on a simple random sample of size n from the given finite population. By putting the conditions $h(1,1) = 1$, $t(1,1) = 1$ and $g(\bar{Y}, 1, 1) = \bar{Y}$, it was ensured that the bias of these estimators are of the order n^{-1}. The mean square errors of these estimators up to terms of order n^{-1}, were obtained and minimized with respect to the unknown parameters in the classes. The minimum mean square errors of the estimators up to terms of order n^{-1} are given by

$$\min M(t_h) = n^{-1}\sigma_y^4\left[\{\beta_2(y)-1\} - \frac{(1-1)^2 C_x^2 + K^2\{\beta_2(x)-1\} - 2(1-1)KM}{\{\beta_2(x)-1\} C_x^2 - M^2}\right] \quad (4)$$

$$\min M(\tilde{y}_t) = \min M(\tilde{y}_g) = n^{-1}\bar{Y}^2\left[C_y^2(1-\rho^2) - \frac{(\rho C_y \nu_1 - \lambda)^2}{(\beta_2 - \beta_1 - 1)}\right] \quad (5)$$

where the notations are as used in Srivastava and Jhajj (1980, 1981).

Recently Singh and Kataria [2] and earlier Singh and Upadhyaya [1] have proposed new classes of estimators for the variance and mean respectively and claim that these are more efficient than (1) and (2), (3) respectively. Their claim is not justified as the improvement in mean square errors obtained and shown by them is in the terms of order n^{-2} whereas they have taken the minimum mean square errors of (1), (2) and (3) up to terms of order n^{-1} only. Thus such a comparison is invalid and the claim about improvement false.

2. *Estimator of Variance*

Singh and Kataria [2] in their paper have proposed a class of estimators of σ_y^2 given by

$$t_s = s_y^2\, h(u,v) + \frac{\alpha}{S_y^2} \quad (6)$$

where $h(1,1) = 1$ and α is taken a constant. Following the Taylor's series expansion of the $h(u,v)$, they arrived at the following expression for the minimum mean square error of (6),

$$\min M(t_s) = \min M(t_h) - \frac{n^{-2}\,\sigma_y^4\,\{\beta_2(y)-1\}}{1+3n^{-1}\,\{\beta_2(y)-1\}} \qquad (7)$$

where min $M(t_h)$ is as given in (4), and claim that the proposed class of estimators is more efficient than the class of estimators (1) considered by Srivastava and Jhajj [3]. Now even if the expression (7) is taken as correct, the conclusion drawn by them is incorrect. The authors write that the expression for $M(t_s)$ is up to terms of order n^{-1} and the expression for min $M(t_h)$ as given in (4) is also up to terms of order n^{-1}. However, the difference min $M(t_h$– min $M(t_s)$ as given by (7) is of order n^{-2}. Hence their conclusion is not valid. In fact, if (7) is regarded as correct, the two estimators t_s and t_h have same minimum mean square errors up to terms of order n^{-1}.

The derivation of the expressions for the mean square error and hence its minimum value is also incorrect. The conclusion drawn from equation (2.2) of Singh and Kataria (1990) that the bias is of order n^{-1} is incorrect. The leading term in the bias of t_s is $\alpha\sigma_y^{-2}$ which is not of order n^{-1}. The expression (2.3) for $M(t_s)$ will involve not only α, h_1 and h_2 but also h_{11}, h_{12} and h_{22}, the second order partial derivatives of $h(u,v)$ at the point (1,1). Up to terms of order n^{-1} it is given by

$$M(t_s) = E\left[\frac{\alpha^2}{\sigma_y^4} + \frac{2\alpha}{\sigma_y^2}(\sigma_y^2 - \alpha)\,\varepsilon + 2\alpha\,h_1\eta + 2\alpha\,h_2\,\delta\right.$$
$$+\left\{(\sigma_y^2-\alpha)^2 + \frac{2\alpha^2}{\sigma_y^2}\right\}\varepsilon^2 + \left(\sigma_y^4\,h_1^2+\alpha\,h_{11}\right)\eta^2$$
$$+\left(\sigma_y^4\,h_2^2 + \alpha\,h_{22}\right)\delta^2 + 2\left(\sigma_y^4 - \alpha\sigma_y^2+\alpha\right)h_1\,\varepsilon\,\eta$$
$$\left.\left[+\ 2\left(\sigma_y^4-\alpha\sigma_y^2+\alpha\right)h_2\,\varepsilon\,\delta + 2\left(\sigma_y^4\,h_1h_2 + \alpha\,h_{12}\right)\eta\,\delta\right]\right. \qquad (8)$$

And now this should be minimized with respect to α, h_1, h_2, h_{11}, h_{12} and h_{22} to obtain the min $M(t_s)$. It can easily be seen that the value of α that minimizes (8) is of order n^{-1}. And if such a value of α is used in the estimator (6), the second term on its right hand side is of order n^{-1}. In all the discussion terms of order n^{-1} are assumed negligible as compared to those of order unity.

3. *Estimator of Mean*

For estimating the mean $\bar{Y}$, Singh and Upadhyaya [1] have considered the class of estimators

$$\tilde{y}_h = h(\bar{y}, u, v)$$

generalizing the estimator (3) of Srivastava and Jhajj [4] by changing the condition $g(\bar{Y}, 1, 1) = \bar{Y}$ to $h(\bar{Y}, 1, 1) = \bar{Y}\, h_1(\bar{Y}, 1, 1)$. Here again the discussion made in the preceding section holds.

The conclusion that the bias of $\tilde{y}_h$ is of order n^{-1} is not true unless $h_1(\bar{Y}, 1, 1) = 1 + 0(n^{-1})$, which has not been assumed. The expression for $M(\tilde{y}_h)$ as given in (2.3) of Singh and Upadhyaya [1] is not correct and will also involve second order partial derivatives of $h(y,u,v)$ at the point $(\bar{Y},1,1)$. Hence the optimum values as given at (2.4) are obviously incorrect. Even if (2.5) is arrived at by imposing some very restrictive conditions on the function $h(\bar{y},u,v)$, the expression (2.5) and (1.4) of their paper are equivalent up to terms of order n^{-1}. The difference min $M(\tilde{y}_g)$ – min $M(\tilde{y}_h)$ as given in (2.6) of Singh and Upadhyaya [1] is of order n^{-2} and hence their claim about the increase in efficiency is incorrect. There is an error (perhaps printing) in the expression (2.6); it should have a multiplier $\frac{1}{n\Delta}$.

REFERENCES

[1] Singh, H.P. and Upadhyaya, L.N., 1986. On a class of estimators of the population mean in survey sampling using auxiliary information. *J. Ind. Soc. Agri. Stat.*, **38**, 100-104.

[2] Singh, S. and Kataria, P., 1990. An estimator of finite population variance. *J. Ind. Soc. Agric. Stat.*, **42**, 186-188.

[3] Srivastava, S.K. and Jhajj, H.S., 1980. A class of estimators using auxiliary information for estimating finite population variance. *Sankhya*, *C*, **42**, 87-96.

[4] Srivastava, S.K. and Jhajj, H.S., 1981. A class of estimator of the population mean in survey sampling using auxiliary information. *Biometrika*, **68**, 341-343.