Ultra Scientist Vol. 24(2)A, 249-254 (2012).

An Estimator for Population Mean Using Power Transformation and Auxiliary Information

RAJ. K. GANGELE and SHAILESH K. CHOUBE

Department of Mathematics and Statistics
Dr. Sir Hari Singh Gour University Sagar, M.P. (INDIA)
Mob. : 09425635980

(Acceptance Date 28th April, 2012)

Abstract

In this paper we have suggested a family of ratio-cum product estimator when the population coefficient of variation (C_x) alongwith population means ($\overline{X}$) of the auxiliary variable x is known in advance. Its bias and variance to the first degree of approximation are obtained. For an appropriate weight 'w' and a good range of α-values, it is shown that the suggested estimator is more efficient than the set of estimators viz. usual unbiased estimator, usual ratio and product estimators, Sisodia and Dwivedi's[5] estimator, Bansal and Singh's[1] estimator, Chakraborty's[2] type estimator, Vos's[8] type estimator and Sahai and Ray[4] type estimator, some of which are in fact particular members of it.

Key words: Finite population, Study variable, Auxiliary variable, Bias, Variance.

1. Introduction

It is well that the use of auxiliary variable (x) at the estimation stage improve of the population mean ($\overline{Y}$) of the study variable (y). Out of many ratio, product and regression estimator's are good examples in this context.

Let U = {U_1, U_2, ... U_n} be a finite population of N units and y be the study variable which takes values x_i on the i^{th} unit, (I = 1, 2, ..., N). Let ($\overline{Y}$, $\overline{X}$) be the population means of y and x respectively. For estimating $\overline{Y}$, a simple random sample of size n is drawn without replacement (WOR) from the population U. When the population mean $\overline{X}$ of x is known, the classical ratio and product estimators of population mean $\overline{Y}$ determined by the intuitively plausible relations are respectively given by

$$\bar{y}_R = \bar{y}\frac{\overline{X}}{\bar{x}} \qquad (1.1)$$

and

$$\bar{y}_P = \bar{y}\frac{\bar{x}}{\bar{X}} \qquad (1.2)$$

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

Utilizing information on known population coefficient of variation (C_x) alongwith known population mean ($\bar{X}$) of auxiliary variable x, Sisodia and Dwivedi[5], Pandey and Dubey[3] suggested the following modified ratio and product estimators for as

$$\bar{y}_R^{(m)} = \bar{y}\frac{(\bar{X} + Cx)}{(\bar{x} + Cx)} \qquad (1.3)$$

and

$$\bar{y}_P^{(m)} = \bar{y}\frac{(\bar{x} + Cx)}{(\bar{X} + Cx)} \qquad (1.4)$$

By applying power transformation on $\bar{y}_R^{(m)}$ and $\bar{y}_P^{(m)}$, Bansal and Singh (1992-93) proposed a modified estimator for $\bar{Y}$ as

$$\bar{y}_B = \bar{y}\left(\frac{\bar{X} + Cx}{\bar{x} + Cx}\right)^{a}, \qquad (1.5)$$

where α is a scalar obtained from a pilot survey, past data or experience. Such an estimator for $\bar{Y}$ has been considered earlier by Srivastava[7] when $\bar{X}$ is known alone.

The variance of $\bar{y}$ is given by

$$V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2 C_y^2, \qquad (1.6)$$

To the first degree of approximation, the variances of $\bar{y}_R$, $\bar{y}_P$, $\bar{y}_R^{(m)}$ and $\bar{y}_P^{(m)}$ are respectively given by

$$V(\bar{y}_R) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + C_x^2(1 - 2C)\right], \qquad (1.7)$$

$$V(\bar{y}_P) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + C_x^2(1 + 2C)\right], \qquad (1.8)$$

$$V(\bar{y}_R^{(m)}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \theta\, C_x^2(\theta - 2C)\right], \qquad (1.9)$$

$$V(\bar{y}_P^{(m)}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \theta\, C_x^2(\theta + 2C)\right], \qquad (1.10)$$

$$V(\bar{y}_B^{(m)}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \alpha\theta\, C_x^2(\alpha\alpha + 2C)\right], \qquad (1.11)$$

where $\theta = \frac{\bar{X}}{\bar{X} + Cx}$, $C = \rho Cy/Cx$, $Cy = Sy/\bar{y}$,

$Cx = Sx/\bar{x}$,

$\rho = Sxy/(SySx)$, $Sxy = \frac{1}{(n-1)}\sum_{i=1}^{N}(x_i - \bar{X})(y_i - \bar{Y})$

$$S_y^2 = \frac{1}{(n-1)}\sum_{i=1}^{N}(y_i - \bar{Y})^2$$

$$S_x^2 = \frac{1}{(n-1)}\sum_{i=1}^{N}(x_i - \bar{X})^2$$

The objective of this paper is to suggest a family of estimators of $\overline{Y}$ and analyze its properties.

2. *Proposed Class of Estimator :*

Motivated by Sisodia and Dwivedi[5,6] we suggest a family of estimators of $\overline{Y}$ as

$$\bar{y}_w = (1-w)\bar{y} + w\bar{y}\left(\frac{\overline{X}+Cx}{\bar{x}+Cx}\right)^{\alpha} \quad (2.1)$$

where w and α are some appropriate constant. For $\alpha > 0$, $\bar{y}_w$ is ratio-type estimator and for $\alpha < 0$, $\bar{y}_w$ is product type estimator, we observe that the estimator reduces to a set of known estimator:

$$\bar{y}_w = \left\{\bar{y},\ \bar{y}_R^{(m)}, \bar{y}_P^{(m)},\ \bar{y}_B^{(m)},\ \bar{y}_w^{(1)},\ \bar{y}_w^{(2)}, \bar{y}_w^{(3)}\right\} \text{ as follows:}$$

(i) Usual unbiased estimator

$$\bar{y}_w \to \bar{y} \text{ for } \alpha = 0 \text{ (or } w = 0)$$

(ii) Sisodia and Dwivedi's[5] estimator

$$\bar{y}_w \to \bar{y}_R^{(m)} \text{ for } \alpha = 1,\ w = 1$$

(iii) Pandey and Dubey's[3] estimator

$$\bar{y}_w \to \bar{y}_P^{(m)} \quad \text{for } \alpha = -1,\ w = 1$$

(iv) Bansal and Singh's[1] estimator

$$\bar{y}_w \to \bar{y}_B \qquad \text{for } w = 1$$

(v) Chakrabort's[2] and Vos's[8] type estimators

$$\bar{y}_w^{(1)} = (1-w)\bar{y} + w\bar{y}\left(\frac{\overline{X}+Cx}{\bar{x}+Cx}\right), \text{for } \alpha=1$$

(vi) Vos's[8] -type estimator

$$\bar{y}_w^{(2)} = (1-w)\bar{y} + w\bar{y}\left(\frac{\bar{x}+Cx}{\overline{X}+Cx}\right), \text{ for } \alpha=1$$

(vii) Sahai and Ray[4] -type estimator

$$\bar{y}_w^{(3)} = \bar{y}\left[2 - \left(\frac{\bar{x}+Cx}{\overline{X}+Cx}\right)^{\alpha}\right], \text{ for } w=-1,$$

$\alpha = -\alpha$

The estimator is a biased one. The bias and variances of $\bar{y}_w$ are derived in section 3. The estimator is compared with other estimators in section 4.

3. *Bias and variance :*

To the first degree of approximation, the bias and variance of $\bar{y}_w$ are respectively given by

$$B(\bar{y}_w) = w\,\alpha\,\theta\,\overline{Y}\left(\frac{1}{n'}-\frac{1}{N}\right)\left\{\frac{(\alpha+1)}{2}\theta - C\right\}Cx^2 \quad (3.1)$$

and

$$V(\bar{y}_w) = \left(\frac{1}{n}-\frac{1}{N}\right)\overline{Y}^2\,[Cy^2 + w\,\alpha\theta\,C_x^2\,(w\,\alpha\,\theta - 2C)] \quad (3.2)$$

If follows from (3.1) that bias of $\bar{y}_w$ are as follows:

$$B(\bar{y}_w)=0 \Rightarrow \frac{(\alpha+1)\theta}{2} - C = 0 \Rightarrow a = \frac{(2C-\theta)}{\theta} \tag{3.3}$$

The optimum values of w and α which minimizes the variance of $\bar{y}_w$ are as follows:

$$\left.\begin{aligned} w_0 &= \frac{C}{\alpha\theta} \\ \alpha_0 &= \frac{C}{w\theta} \end{aligned}\right\} \tag{3.4}$$

We note from (3.4) that the optimum values of w and α can not uniquely be determined. However, the optimum values of α can be obtained for a given value of w which can be appropriately fixed in advance as being weight. Thus, if we substitute the optimum value of α for a given value of w in (3.2), we get

$$\min.V(\bar{y}_w) = \left(\frac{1}{n}-\frac{1}{N}\right) S_y^2(1-\rho^2) \tag{3.5}$$

which is the approximate variance of the usual regression estimator $\bar{y}_{1r} = \bar{y} + \hat{\beta}(\bar{X}-\bar{x})$, $\hat{\beta} = \frac{S_{xy}}{S_x^2}$ is the sample estimate of population regression coefficient β of y on x. The expression for min.($\bar{y}_w$) in (3.5) implies that the estimator $\bar{y}_w$ will always be more efficient than the set of estimators $\bar{y}_w'$ at α_0. Nevertheless, we shall further determine a range of α-values for which is estimator $\bar{y}_w$ will have lesser variance than that of $\bar{y}_w'$.

It is observed that the estimators in $\bar{y}_w'$ are particular member of $\bar{y}_w$, therefore, from (3.2) we get at once variance of $\bar{y}_w'$ as follows:

V($\bar{y}_w'$) =V[$\bar{y}_w$; according to different values of α and w given in from (i) to (vii)]

$$= \left[V(\bar{y}),\ V(\bar{y}_R^{(m)}), V(\bar{y}_P^{(m)}), V(\bar{y}_B^{(m)}), V(\bar{y}_w^{(1)}),\ V(\bar{y}_w^{(2)}), V(\bar{y}_w^{(3)})\right]$$

4. Comparison of with Other Comptetors :

It follows from (1.6), (1.9), (1.10) and (3.2) that the proposed estimator $\bar{y}_w$ is more efficient than $\bar{y}$, $y_R^{(m)}$, $y_P^{(m)}$ and y_n if

$$|\alpha-\alpha_0| < |\alpha_0| \tag{4.1}$$

$$|\alpha-\alpha_0| < \left|\alpha_0 - \frac{1}{w}\right| \tag{4.2}$$

$$|\alpha-\alpha_0| < \left|\alpha_0 + \frac{1}{w}\right| \tag{4.3}$$

and

$$|\alpha-\alpha_0| < \left|\frac{\alpha_0(w-1)}{(w+1)}\right| \tag{4.4}$$

hold true respectively.

To compare $\bar{y}_w$ with $y_w^{(1)}$, $y_w^{(2)}$ and $y_w^{(3)}$, we write the variances of $y_w^{(1)}$, $y_w^{(2)}$ and $y_w^{(3)}$

upto first order of approximation, respectively as

$$V(y_w^{(1)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \overline{Y}^2 \, [C_y^2 + w\theta C_x^2 \, (w\theta - 2C)] \quad (4.5)$$

$$V(y_w^{(2)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \overline{Y}^2 \, [C_y^2 + w\theta \, C_x^2 \, (w\theta - 2C)] \quad (4.6)$$

and $$V(y_w^{(3)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \overline{Y}^2 \, [C_y^2 + \alpha\,\theta\, C_x^2 \, (\alpha\theta - 2C)] \quad (4.7)$$

From (3.2), (4.5), (4.6) and (4.7) we note that the suggested estimator $\overline{y}_w$ is more precise than $y_w^{(1)}$, $y_w^{(2)}$ and $y_w^{(3)}$ if

$$|\alpha - \alpha_0| < |\alpha_0 - 1| \quad (4.8)$$

$$|\alpha - \alpha_0| < |\alpha_0 + 1| \quad (4.9)$$

and

$$|\alpha - \alpha_0| < \left|\frac{\alpha_0 (w-1)}{(w+1)}\right| \quad (4.10)$$

hold good respectively.

From (3.2) and (1.7) we noted that $V(\overline{y}_w) < V(\overline{y}_R)$ if

$$|\alpha - \alpha_0| < \left|\alpha_0 - \frac{1}{w\theta}\right| \quad (4.11)$$

Further from (3.2) and (1.8), we note that the suggested estimator $\overline{y}_w$ is more efficient than classical product estimator $\overline{y}_P$ if

$$|\alpha - \alpha_0| < \left|\alpha_0 + \frac{1}{w\theta}\right| \quad (4.12)$$

For a given value of w, the difference between α and the optimum value α_0 in inequalities (4.1)-(4.4), (4.8), (4.9), (4.10), (4.11) and (4.12) provide quite a good range of α-values in order to have the estimator $\overline{y}_w$ more efficient that $\overline{y}_w^{'}$.

It is to be noted that the optimum value α_0 of α is a function of C, θ and w, it may be positive or negative depending upon the sign of correlation coefficient 'ρ'. In practice, with a good guess of 'C' obtained through pilot survey, past data or experience gathered in due course of time, and an appropriate weight of w, an optimum value of α fairly very close to its true value α_0 ca be obtained. It is pointed out in Sahai and Ray[4] that when good (specific) gives of C is not available we may still have more information about the range of C, e.g. $C_1 \leq C \leq C_2$ which is more realistic than a specific gives about C. However, in case a good given value of 'C' is not known, the optimum values of α for different value of C, θ and w can be computed.

References

1. Bansal, M.L. and Sukhjinder, S., An improved estimator of population mean using power transformation. *Alig. Jour. Statist. 12 & 13,* 21-24 (1992-93).
2. Chakraborty, R.P., Some ratio type estimators. *Jour. Ind. Soc. Agn. Statist., 31*, 1, 48-62

(1979).
3. Pandey, B. N. and Dubey, V., Modified product estimator using coefficient of variation of auxiliary variate. *Assam. Statist. Rev., 2*, 2, 64-66 (1988).
4. Sahai, A. and Ray, S.K., An efficient estimator using auxiliary information Metrika, *27*, 271-275 (1980).
5. Sisodia, B.V.S. and Dwivedi, V.K., A ratio-cum-product-type estimator. *Biometrical Journal, 23,* 133-139 (1981a).
6. Sisodia, B.V.S. and Dwivedi, V.K., A class of ratio cum product-type estimator in double sampling *Biometrical Journal, 24,* 4, 419-424 (1982).
7. Srivastava S.K., An estimator using auxiliary information in sample surveys, "Calcutta Statistical Association Bulletin", 16, pp. 121-132 (1967).
8. Vos, J.W.E., Mixing of direct, ratio and product method of estimation. *Statistica Neerlandica, 34(4),* 209-218 (1980).