

A Generalized Double Sampling Estimator for Finite Population Variance

S. Ali Manzar Rizvi and S. Athar Hasan Rizvi*

(rizvi_athar@hotmail.com)

Department of Statistics, Shia P. G. College,
Lucknow-226020, INDIA.

ABSTRACT

A generalized double sampling estimator for the estimation of finite population variance is proposed, its bias and mean squared error are found, and its comparison with the usual estimator of the finite population variance is made. A subclass of estimators depending upon optimum values for which the subclass attains the minimum mean square error is investigated and further a subclass of estimators depending upon estimated optimum value is also searched and its properties are studied.

KEY-WORDS: Optimum estimators, Bias and mean square error, Estimated optimum and Efficiency.

I. INTRODUCTION

For a first phase large simple random sample of size n' from a population of size N , let the auxiliary character x be observed to find an estimate of population mean \bar{X} of x , and further, let the characters y, x be observed on the second phase simple random sample of size n from the first phase sample of size n' . Let (\bar{Y}, \bar{X}) be the population means of the characters (y, x) respectively, \bar{x}' be the sample mean of n' first phase sample values on x and (\bar{y}, \bar{x}) be the sample means of n second phase sample values on (y, x) respectively. Also let ρ be the correlation coefficient between (y, x) and $s_y^2 = \frac{1}{(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2$ based on the second phase sample observations (y_1, y_2, \dots, y_n) , be the conventional estimator of the finite population variance σ_y^2 .

For estimating σ_y^2 , the proposed generalized double sampling estimator is

$$\hat{\sigma}_d^2 = g(s_y^2, \bar{x}, \bar{x}') \quad (1.1)$$

where $g(s_y^2, \bar{x}, \bar{x}')$ satisfying the validity conditions of Taylor's series expansion is a bounded function of $(s_y^2, \bar{x}, \bar{x}')$ such that at the point $Q = (\sigma_y^2, \bar{X}, \bar{X}')$,

$$(i) \quad g(\sigma_y^2, \bar{X}, \bar{X}') = \sigma_y^2 \quad ; \quad (1.2)$$

(ii) first order partial derivative of $g(s_y^2, \bar{x}, \bar{x}')$ with respect to s_y^2 at the point Q is unity, that is,

$$g_0 = \left. \frac{\partial g(s_y^2, \bar{x}, \bar{x}')}{\partial s_y^2} \right]_Q = 1 \quad , \quad (1.3)$$

$$(iii) \quad g_1 = -g_2 \quad (1.4)$$

for first order partial derivatives

* Author for Correspondence

$$g_1 = \left. \frac{\partial g(s_y^2, \bar{x}, \bar{x}')}{\partial \bar{x}} \right]_Q, \quad g_2 = \left. \frac{\partial g(s_y^2, \bar{x}, \bar{x}')}{\partial \bar{x}'} \right]_Q;$$

(iv) second order partial derivative

$$g_{00} = \left. \frac{\partial^2 g(s_y^2, \bar{x}, \bar{x}')}{\partial (s_y^2)^2} \right]_Q = 0 \tag{1.5}$$

(v) $g_{01} = -g_{02}$ (1.6)

for $g_{01} = \left. \frac{\partial^2 g(s_y^2, \bar{x}, \bar{x}')}{\partial s_y^2 \partial \bar{x}} \right]_Q, \quad g_{02} = \left. \frac{\partial^2 g(s_y^2, \bar{x}, \bar{x}')}{\partial s_y^2 \partial \bar{x}'} \right]_Q.$

II. BIAS AND MEAN SQUARE ERROR

Let

$$\mu_{rs} = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^r (X_i - \bar{X})^s$$

where (Y_i, X_i) are the values on the characters (y, x) respectively for the i^{th} ($i = 1, 2, \dots, N$) unit of the population.

Further, let

$$e_1 = \bar{x} - \bar{X}, \quad e_1' = \bar{x}' - \bar{X}, \quad e_2 = s_y^2 - S_y^2 \text{ so that (for large } N)$$

$$E(e_1) = E(e_1') = E(e_2) = 0 \text{ and}$$

$$E(e_1^2) = \frac{\mu_{02}}{n}, \quad E(e_1'^2) = \frac{\mu_{02}}{n'}, \quad E(e_2^2) = \frac{\mu_{20}^2}{n} \left(\frac{\mu_{40}}{\mu_{20}^2} - 1 \right) = \frac{\mu_{20}^2}{n} (\beta_{2y} - 1),$$

$$E(e_1 e_1') = \frac{\mu_{02}}{n'}, \quad E(e_1 e_2) = \frac{\mu_{21}}{n}, \quad E(e_1' e_2) = \frac{\mu_{21}}{n'}.$$

Expanding $g(s_y^2, \bar{x}, \bar{x}')$ about the point $Q = (s_y^2, \bar{X}, \bar{X}')$ in third order Taylor's series, we have

$$\begin{aligned} \hat{\sigma}_d^2 &= g(s_y^2, \bar{X}, \bar{X}') + (s_y^2 - S_y^2)g_0 + (\bar{x} - \bar{X})g_1 + (\bar{x}' - \bar{X}')g_2 \\ &\quad + \frac{1}{2!} \left\{ (s_y^2 - S_y^2)^2 g_{00} + (\bar{x} - \bar{X})^2 g_{11} + (\bar{x}' - \bar{X}')^2 g_{22} \right. \\ &\quad + 2(s_y^2 - S_y^2)(\bar{x} - \bar{X})g_{01} + 2(s_y^2 - S_y^2)(\bar{x}' - \bar{X}')g_{02} \\ &\quad \left. + 2(\bar{x} - \bar{X})(\bar{x}' - \bar{X}')g_{12} \right\} + \frac{1}{3!} \left\{ (s_y^2 - S_y^2) \frac{\partial}{\partial s_y^2} + (\bar{x} - \bar{X}) \frac{\partial}{\partial \bar{x}} \right. \\ &\quad \left. + (\bar{x}' - \bar{X}') \frac{\partial}{\partial \bar{x}'} \right\}^3 g(s_{y*}^2, \bar{x}_*, \bar{x}'_*) \end{aligned} \tag{2.1}$$

where second order partial derivatives g_{11} , g_{22} , and g_{12} are given by

$$g_{11} = \left. \frac{\partial^2 g(s_y^2, \bar{x}, \bar{x}')}{\partial \bar{x}^2} \right]_0, \quad g_{22} = \left. \frac{\partial^2 g(s_y^2, \bar{x}, \bar{x}')}{\partial \bar{x}'^2} \right]_0, \quad g_{12} = \left. \frac{\partial^2 g(s_y^2, \bar{x}, \bar{x}')}{\partial \bar{x} \partial \bar{x}'} \right]_0,$$

$g_0, g_1, g_2, g_{00}, g_{01}$ and g_{02} are already defined and

$$s_{y*}^2 = \sigma_y^2 + \theta(s_y^2 - S_y^2), \quad \bar{x}_* = \bar{X} + \theta(\bar{x} - \bar{X}),$$

$$\bar{x}'_* = \bar{X} + \theta(\bar{x}' - \bar{X}) \quad \text{for } 0 < \theta < 1.$$

Employing regularity conditions $g(\sigma_y^2, \bar{X}, \bar{X}) = \sigma_y^2, g_0 = 1, g_{00} = 0, g_1 = -g_2$ and $g_{01} = -g_{02}$ from (1.2) to (1.6) in (2.1), we have

$$\begin{aligned} \hat{\sigma}_d^2 &= \sigma_y^2 + (s_y^2 - S_y^2) + (\bar{x} - \bar{X})g_1 - (\bar{x}' - \bar{X})g_1 + \frac{1}{2!} \left\{ (\bar{x} - \bar{X})^2 g_{11} \right. \\ &\quad + (\bar{x}' - \bar{X})^2 g_{22} + 2(s_y^2 - S_y^2)(\bar{x} - \bar{X})g_{01} - 2(s_y^2 - S_y^2)(\bar{x}' - \bar{X})g_{01} \\ &\quad \left. + 2(\bar{x} - \bar{X})(\bar{x}' - \bar{X})g_{12} \right\} + \frac{1}{3!} \left\{ (s_y^2 - S_y^2) \frac{\partial}{\partial s_y^2} + (\bar{x} - \bar{X}) \frac{\partial}{\partial \bar{x}} \right. \\ &\quad \left. + (\bar{x}' - \bar{X}) \frac{\partial}{\partial \bar{x}'} \right\}^3 g(s_{y*}^2, \bar{x}_*, \bar{x}'_*) \end{aligned}$$

or

$$\begin{aligned} \hat{\sigma}_d^2 - \sigma_y^2 &= (s_y^2 - S_y^2) + \{(\bar{x} - \bar{X}) - (\bar{x}' - \bar{X})\}g_1 + \frac{1}{2!} [(\bar{x} - \bar{X})^2 g_{11} \\ &\quad + (\bar{x}' - \bar{X})^2 g_{22} + 2\{(s_y^2 - S_y^2)(\bar{x} - \bar{X}) - (s_y^2 - S_y^2)(\bar{x}' - \bar{X})\}g_{01} \\ &\quad + 2(\bar{x} - \bar{X})(\bar{x}' - \bar{X})g_{12}] + \frac{1}{3!} \left\{ (s_y^2 - S_y^2) \frac{\partial}{\partial s_y^2} + (\bar{x} - \bar{X}) \frac{\partial}{\partial \bar{x}} \right. \\ &\quad \left. + (\bar{x}' - \bar{X}) \frac{\partial}{\partial \bar{x}'} \right\}^3 g(s_{y*}^2, \bar{x}_*, \bar{x}'_*) \\ &= e_2 + (e_1 - e_1')g_1 + \frac{1}{2!} \{e_1^2 g_{11} + e_1'^2 g_{22} + 2(e_1 e_2 - e_1' e_2')g_{01} \\ &\quad + 2e_1 e_1' g_{12}\} + \frac{1}{3!} \left\{ (s_y^2 - S_y^2) \frac{\partial}{\partial s_y^2} + (\bar{x} - \bar{X}) \frac{\partial}{\partial \bar{x}} \right. \\ &\quad \left. + (\bar{x}' - \bar{X}) \frac{\partial}{\partial \bar{x}'} \right\}^3 g(s_{y*}^2, \bar{x}_*, \bar{x}'_*) \end{aligned} \tag{2.2}$$

Taking expectation on both the sides of (2.2), to the first degree of approximation, we have

$$E(\hat{\sigma}_d^2 - \sigma_y^2) = E\left\{e_2 + (e_1 - e'_1)g_1 + \frac{e_1^2}{2}g_{11} + \frac{e_1'^2}{2}g_{22} + (e_1e_2 - e'_1e_2)g_{01} + e_1e'_1g_{12}\right\}$$

or

$$\begin{aligned} Bias(\hat{\sigma}_d^2) &= \frac{\mu_{02}}{2n}g_{11} + \frac{\mu_{02}}{2n'}g_{22} + \left(\frac{\mu_{21}}{n} - \frac{\mu_{21}}{n'}\right)g_{01} + \frac{\mu_{02}}{n'}g_{12} \\ &= \frac{\mu_{02}}{2}\left(\frac{g_{11}}{n} + \frac{g_{22}}{n'} + 2\frac{g_{12}}{n'}\right) + \left(\frac{1}{n} - \frac{1}{n'}\right)\mu_{21}g_{01}. \end{aligned} \tag{2.3}$$

Squaring both the sides of (2.2) and taking expectation to the first degree of approximation, we have

$$E(\hat{\sigma}_d^2 - \sigma_y^2)^2 = E\left\{e_2^2 + (e_1^2 + e_1'^2 - 2e_1e_1')g_1^2 + 2(e_1e_2 - e_1'e_2)g_1\right\}$$

or

$$\begin{aligned} MSE(\hat{\sigma}_d^2) &= \frac{\mu_{20}^2(\beta_{2y} - 1)}{n} + \left(\frac{\mu_{02}}{n} + \frac{\mu_{02}}{n'} - 2\frac{\mu_{02}}{n'}\right)g_1^2 + 2\left(\frac{\mu_{21}}{n} - \frac{\mu_{21}}{n'}\right)g_1 \\ &= MSE(s_y^2) + \left(\frac{1}{n} - \frac{1}{n'}\right)\mu_{02}g_1^2 + 2\left(\frac{1}{n} - \frac{1}{n'}\right)\mu_{21}g_1 \end{aligned}$$

or

$$MSE(\hat{\sigma}_d^2) = MSE(s_y^2) + \left(\frac{1}{n} - \frac{1}{n'}\right)(\mu_{02}g_1^2 + 2\mu_{21}g_1) . \tag{2.4}$$

III. OPTIMUM AND ESTIMATED OPTIMUM VALUES

From (2.4), we see that the value of g_1 for which $MSE(\hat{\sigma}_d^2)$ is minimized, is given by

$$g_1 = -\frac{\mu_{21}}{\mu_{02}} = -\frac{\sigma_y^2}{\bar{X}} \left(\frac{\mu_{21}/\sigma_y^2 \bar{X}}{C_x^2} \right) = D \tag{3.1}$$

and the minimum mean square error is

$$MSE(\hat{\sigma}_d^2)_{min} = MSE(s_y^2) - \left(\frac{1}{n} - \frac{1}{n'}\right) \frac{\mu_{21}^2}{\mu_{02}} . \tag{3.2}$$

Practically, the optimum value $D = -\frac{\mu_{21}}{\mu_{02}}$ may not be available always, hence the alternative is to replace the parameters involved in the optimum value by their unbiased or consistent estimators and thus get the estimated optimum value. Replacing $\sigma_y^2, \bar{X}, \mu_{21}$ and μ_{02} by

$$s_y^2, \bar{x}, m_{21} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 (x_i - \bar{x}) \text{ and } m_{02} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \text{ respectively, we get the}$$

estimated optimum value to be

$$\hat{D} = -\frac{s_y^2}{\bar{x}} \hat{\lambda} \tag{3.3}$$

where $\hat{\lambda} = \frac{m_{21}/s_y^2 \bar{x}}{\hat{C}_x^2}, \hat{C}_x^2 = \frac{m_{02}}{\bar{x}^2} .$

The condition $g_1 = D$ in (3.1) to attain the minimum mean square error in (3.2) shows that our estimator $\hat{\sigma}_d^2$ as a function $g(s_y^2, \bar{x}, \bar{x}')$ should involve D also, that is, our estimator should be of the

form $g(s_y^2, \bar{x}, \bar{x}', D)$ satisfying the condition $g_1 = D$ along with the conditions (1.2) to (1.6) to attain the minimum mean square error in (3.2). As D is unknown in $g(s_y^2, \bar{x}, \bar{x}', D)$, we replace it by \hat{D} given in (3.3) and thus get the function $g^*(s_y^2, \bar{x}, \bar{x}', \hat{D})$ as an estimator depending on estimated optimum value.

Now expanding $g^*(s_y^2, \bar{x}, \bar{x}', \hat{D})$ about the point $Q^* = (\sigma_y^2, \bar{X}, \bar{X}, D)$ in Taylor's series, we have

$$g^*(s_y^2, \bar{x}, \bar{x}', \hat{D}) = g^*(\sigma_y^2, \bar{X}, \bar{X}, D) + (s_y^2 - \sigma_y^2)g_0^* + (\bar{x} - \bar{X})g_1^* + (\bar{x}' - \bar{X})g_2^* + (\hat{D} - D)g_3^* + \dots \tag{3.4}$$

where $g^*(\sigma_y^2, \bar{X}, \bar{X}, D) = \sigma_y^2$, $g_0^* = \left. \frac{\partial g^*}{\partial s_y^2} \right]_{Q^*} = 1$;

$$g_1^* = \left. \frac{\partial g^*}{\partial \bar{x}} \right]_{Q^*}, \quad g_2^* = \left. \frac{\partial g^*}{\partial \bar{x}'} \right]_{Q^*} \quad \text{with } g_1^* = -g_2^*,$$

and $g_3^* = \left. \frac{\partial g^*}{\partial \hat{D}} \right]_{Q^*}$.

Noting the conditions $g^*(\sigma_y^2, \bar{X}, \bar{X}, D) = \sigma_y^2$, $g_0^* = 1$, $g_1^* = -g_2^*$ in (3.4), we see that

$$g^*(s_y^2, \bar{x}, \bar{x}', \hat{D}) - \sigma_y^2 = (s_y^2 - \sigma_y^2) + \{(\bar{x} - \bar{X}) - (\bar{x}' - \bar{X})\}g_1^* + (\hat{D} - D)g_3^* + \dots \tag{3.5}$$

Squaring both the sides of (3.5) and taking expectation, we see that the mean square error $E[g^*(s_y^2, \bar{x}, \bar{x}', \hat{D}) - \sigma_y^2]^2$ to the first degree of approximation, becomes equal to $MSE(\hat{\sigma}_d^2)_{min}$ given by (3.2) if $g_3^* = 0$, and thus the estimator taken as a function $\hat{\sigma}_{de}^2 = g^*(s_y^2, \bar{x}, \bar{x}', \hat{D})$ depending upon estimated optimum values, attains the minimum mean square error given by (3.2) if

$$\left. \begin{aligned} &g^*(\sigma_y^2, \bar{X}, \bar{X}, D) = \sigma_y^2, \quad \left. \frac{\partial g^*}{\partial s_y^2} \right]_{Q^*} = 1, \quad g_1^* = -g_2^*, \\ &\left. \frac{\partial^2 g^*}{\partial (s_y^2)^2} \right]_{Q^*} = 0, \quad g_{01}^* = \left. \frac{\partial^2 g^*}{\partial s_y^2 \partial \bar{x}} \right]_{Q^*} = -\left. \frac{\partial^2 g^*}{\partial s_y^2 \partial \bar{x}'} \right]_{Q^*} = -g_{02}^* \\ &\text{and } g_3^* = 0. \end{aligned} \right\} \tag{3.6}$$

Depending upon estimated optimum value \hat{D} , some particular members of the class $\hat{\sigma}_d^2$ attaining the minimum mean square error in (3.2) and satisfying the conditions in (3.6), are given in Concluding Remarks.

IV. CONCLUDING REMARKS

(a) Some members satisfying the regularity conditions from (1.2) to (1.6) and belonging to the class of estimators $\hat{\sigma}_d^2$ are

$$(i) \quad \hat{\sigma}_{d_1}^2 = s_y^2 \left(\frac{\bar{x}}{\bar{x}'} \right),$$

$$(ii) \quad \hat{\sigma}_{d_2}^2 = s_y^2 \left(\frac{\bar{x}}{\bar{x}'} \right)^{k_1},$$

$$(iii) \quad \hat{\sigma}_{d_3}^2 = s_y^2 + k_2(\bar{x} - \bar{x}')$$

and (iv) $\hat{\sigma}_{d_4}^2 = ks_y^2 + (1 - k)s_y^2 \left(\frac{\bar{x}'}{\bar{x}} \right)$

where k_1, k_2, k are the characterizing scalars to be chosen suitably. For the estimators $s_y^2 \left(\frac{\bar{x}}{\bar{x}'} \right)^{k_1}$,

$s_y^2 + k_2(\bar{x} - \bar{x}')$ and $ks_y^2 + (1 - k)s_y^2 \left(\frac{\bar{x}'}{\bar{x}} \right)$, we see that the values of g_1 are $k_1 \frac{\sigma_y^2}{\bar{X}}$, k_2

and $-(1 - k) \frac{\sigma_y^2}{\bar{X}}$ respectively, and equating these values of g_1 to the optimum value D in (3.1)

gives respectively the optimum values of k_1, k_2, k to be $\frac{\bar{X}}{\sigma_y^2} D$, D and $1 + \frac{\bar{X}}{\sigma_y^2} D$ wherein

the unknown parameters are involved. Hence, replacing the unknown values in the optimum values of k_1, k_2, k by their unbiased or consistent estimators, we get the estimated optimum values

$$\hat{k}_1, \hat{k}_2, \hat{k} \text{ respectively for the estimators } \hat{\sigma}_{d_2}^2, \hat{\sigma}_{d_3}^2 \text{ and } \hat{\sigma}_{d_4}^2 \text{ to be } \hat{k}_1 = \left(\frac{\bar{x}}{s_y^2} \right) \hat{D}, \hat{k}_2 = \hat{D}$$

and $\hat{k} = 1 + \left(\frac{\bar{x}}{s_y^2} \right) \hat{D}$ for which the estimators $s_y^2 \left(\frac{\bar{x}}{\bar{x}'} \right)^{\hat{k}_1}$, $s_y^2 + \hat{k}_2(\bar{x} - \bar{x}')$ and

$\hat{k}s_y^2 + (1 - \hat{k})s_y^2 \left(\frac{\bar{x}'}{\bar{x}} \right)$ satisfying the conditions in (3.6), attain the minimum mean square error given in (3.2).

(b) It may be further mentioned here that, if we directly proceed to find the mean square error to the first

degree of approximation for the estimators $s_y^2 \left(\frac{\bar{x}}{\bar{x}'} \right)^{\hat{k}_1}$, $s_y^2 + \hat{k}_2(\bar{x} - \bar{x}')$ and

$\hat{k}s_y^2 + (1 - \hat{k})s_y^2 \left(\frac{\bar{x}'}{\bar{x}} \right)$, we will find the same minimum mean square error as obtained in (3.2) for

the generalized double sampling estimator $\hat{\sigma}_{de}^2$ depending upon estimated optimum value \hat{D} .

(c) For $n' = N$, single sampling results may be easily found from those obtained here for double sampling.

REFERENCES

- [1]. **COCHRAN, W. G.** (1977). Sampling Techniques (3rd Edition), John Wiley & Sons, Inc., New York.
- [2]. **MURTHY, M. N.** (1967). Sampling Theory and Methods, Statistical Publishing Society, India.
- [3]. **RIZVI, S. A. H. and SINGH, R. KARAN** (2016). An Alternative Estimator of Finite Population Variance Using Auxiliary Information, Chapter 21, Recent Advances in Applied Statistics and its Applications. pp. 173-176 Lambert Academic Publishing, Republic of MOLDOVA, CHISINAU, EUROPE.
- [4]. **RIZVI, S. A. H., and RIZVI, S. A. M.** (2009). Some Estimators of Population Variance Using Auxiliary Information, Journal of Applied Research, Vol. I, No.1, June-2009, pp. 54-61, INDIA.
- [5]. **RIZVI, S. A. H., RIZVI, S. A. M., SINGH, R. and SINGH, R. KARAN** (2015). Difference and Regression Type Estimators for the Estimation of Finite Population Variance Using Auxiliary information (2015), International Journal of Statistics and Systems (IJSS), vol. 10, No. 1, pp. 111-115, INDIA.
- [6]. **SINGH, R. KARAN, ZAIDI, S. M. H and RIZVI, S. A. H** (1995). Estimation of Finite Population Variance Using Auxiliary Information, Journal of Statistical Studies., Vol. 15, pp. 17-28, DHAKA.
- [7]. **SINGH, R. KARAN, ZAIDI, S. M. H and RIZVI, S. A. H** (1996). Some Estimators of Finite Population Variance Using Information on Two Auxiliary Variables, Microelectron. Reliab., Vol. 36, No. 5, pp. 667-670, Elsevier Science Ltd., GREAT BRITAIN.
- [8]. **SUKHATME, P. V., SUKHATME, B. V., SUKHATME, S. and ASOK, C.** (1984). Sampling theory of Surveys with Application, 3rd Ed., Ames, Iowa (USA) and Indian Society of Agricultural Statistics, New Delhi, (India).