International Review of Pure and Applied Mathematics
(January-June) 2017, Volume 13, No. 1, pp. 121–138

A Generalized Double Sampling Estimator for the Mean of a Finite Population

Peeyush Misra

Abstract: This article considers the problem of estimating finite population mean using auxiliary information. A generalized double sampling modified regression type estimator using auxiliary information is proposed for the purpose. The bias and mean square error expressions are found and its comparative study with some of the well known estimators is done. An empirical study is also given to judge the efficiency of the proposed generalized estimator over the others.

Keywords: Auxiliary Variable, Taylor's Series Expansion, Bias, Mean Squared Error and Efficiency.

1. Introduction

Statisticians often make use of the information available on an auxiliary variable with the variable under study for improving the efficiency of an estimator. For better understanding one may see Cochran (1977), Des Raj (1968), Murthy (1967), Mukhopadhyay (2012), Singh & Chaudhary (1997) and Sukhatme et, al. (1984). It is a well known fact that the auxiliary information in sample surveys results in substantial improvement in the precision of the estimators of the population parameters and we know that sometimes parameters of the auxiliary variables are not known in advance then double or two phase sampling technique is used. In double sampling or two-phase sampling technique, we first take a preliminary large sample of size n' (called first phase sample) from a population of size N and then a sub-sample of size n (called second phase sample) is drawn from the first phase sample of size n' using simple random sampling without replacement at both the phases. At first phase sample of size n', only the auxiliary variable X be observed but at the second phase sample of size n, the study variable Y and the auxiliary variable X both are observed.

Let $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ be the population mean of study variable y and $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ be the population mean of auxiliary variable x.

$$\sigma_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2,\ \sigma_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2 \text{ and}$$

$$\rho = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)\left(X_i - \bar{X}\right)}{\sigma_Y \sigma_X}$$

be the population correlation coefficient between y and x.

Also let $\mu_{rs} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^r \left(X_i - \bar{X}\right)^s, C_Y^2 = \frac{\sigma_Y^2}{\bar{Y}^2}, C_X^2 = \frac{\sigma_X^2}{\bar{X}^2} = \frac{\mu_{02}}{\bar{X}^2}, \rho = \frac{\mu_{11}}{\sigma_Y \sigma_X}$,

$\beta_2 = \frac{\mu_{04}}{\mu_{02}^2}, \beta_1 = \frac{\mu_{03}^2}{\mu_{02}^3}, \gamma_1 = \sqrt{\beta_1}$.

Let the first phase sample of size n' be $\left(x_1', x_2', ..., x_n'\right)$ on x and the second phase sample of size n be $\left\{\left(y_1, x_1\right), \left(y_2, x_2\right), ..., \left(y_n, x_n\right)\right\}$ on variables (y, x) with the first phase sample mean $\bar{x}' = \frac{1}{n'}\sum_{i=1}^{n'} x'_i$ estimator of population mean $\bar{X}$ and the second phase sample mean $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ respectively on y and x.

For simplicity, it is assumed that N is large enough as compared to n so that finite population correction terms may be ignored. A new generalized double sampling regression type estimator represented by $\bar{y}_g$ for estimating the population mean is proposed as

$$\bar{y}_g = g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right) \tag{1.1}$$

and $g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)$ is a bounded function such that at the point $\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)$

(i) $g\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right) = \bar{Y}$

(ii) $g_0 = \frac{\partial}{\partial \bar{y}} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = 1$

(iii) $$g_1 = \frac{\partial}{\partial b} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = 0$$

(iv) $$g_{11} = \frac{\partial^2}{\partial b^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = 0$$

(v) $$g_2 = \frac{\partial}{\partial \bar{x}} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = -\beta$$

(vi) $$g_3 = \frac{\partial}{\partial \bar{x}'} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = -g_2$$

(vii) $$g_{00} = \frac{\partial^2}{\partial \bar{y}^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = 0$$

(viii) $$g_{22} = \frac{\partial^2}{\partial \bar{x}^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)} = 0$$

(ix) $g_{02} = -g_{03}$

where, $$g_{02} = \frac{\partial^2}{\partial \bar{y} \partial \bar{x}} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)}$$

and $$g_{03} = \frac{\partial^2}{\partial \bar{y} \partial \bar{x}'} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)}$$

(x) $g_4 = -g_5$

where, $$g_4 = \frac{\partial^2}{\partial s_x^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)}$$

and $$g_5 = \frac{\partial^2}{\partial s_x'^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)}$$

(xi) $g_{04} = -\,g_{05}$

where, $g_{04} = \frac{\partial^2}{\partial \bar{y} \partial s_x^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)}$

and $g_{05} = \frac{\partial^2}{\partial \bar{y} \partial s_x'^2} g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)\Big]_{\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)}$

and b is an estimate of the change in y when x is increased by unity.

2. Bias and Mean Squared Error of the Proposed Estimator

In order to obtain bias and mean square error of the proposed estimator, let us denote by

$$\bar{y} = \bar{Y}(1 + e_0)$$

$$\bar{x} = \bar{X}(1 + e_1)$$

$$\bar{x}' = \bar{X}(1 + e_1')$$

$$s_{yx} = e_2 + S_{YX}$$

$$s_y^2 = e_3 + S_Y^2$$

$$s_x^2 = e_4 + S_X^2$$

$$s_x'^2 = e_4' + S_X^2 \qquad (2.1)$$

so that ignoring finite population correction, for simplicity we have

$$E(e_0) = E(e_1) = E(e_1') = E(e_2) = E(e_3) = E(e_4) = E(e_4') = 0 \qquad (2.2)$$

$$E\left(e_0^2\right) = \frac{\mu_{20}}{n\bar{Y}^2} = \frac{1}{n} C_Y^2$$

$$E\left(e_1^2\right) = \frac{\mu_{02}}{n\bar{X}^2} = \frac{1}{n} C_X^2$$

$$E\left(e_1'^2\right) = \frac{\mu_{02}}{n'\bar{X}^2} = \frac{1}{n'} C_X^2$$

$$E\left(e_3^2\right)=\left(\frac{\beta_2(y)-1}{n}\right)S_Y^4=\frac{\mu_{20}^2}{n}\left(\frac{\mu_{40}}{\mu_{20}^2}-1\right)$$

$$E\left(e_4^2\right)=\left(\frac{\beta_2(x)-1}{n}\right)S_X^4=\frac{\mu_{02}^2}{n}\left(\frac{\mu_{04}}{\mu_{02}^2}-1\right)$$

$$E\left(e_0e_1\right)=\frac{\mu_{11}}{n\overline{Y}\overline{X}}=\frac{1}{n}\rho C_Y C_X$$

$$E\left(e_0e_1'\right)=\frac{\mu_{11}}{n'\overline{Y}\overline{X}}=\frac{1}{n'}\rho C_Y C_X$$

$$E\left(e_0e_3\right)=\frac{\mu_{30}}{n\overline{Y}}$$

$$E\left(e_0e_4\right)=\frac{\mu_{12}}{n\overline{Y}}$$

$$E\left(e_0e_4'\right)=\frac{\mu_{12}}{n'\overline{Y}}$$

$$E\left(e_1e_1'\right)=\frac{\mu_{02}}{n'\overline{X}^2}=\frac{1}{n'}C_X^2$$

$$E\left(e_1e_2\right)=\frac{\mu_{12}}{n\overline{X}}$$

$$E\left(e_1e_3\right)=\frac{\mu_{21}}{n\overline{X}}$$

$$E\left(e_1e_4\right)=\frac{\mu_{03}}{n\overline{X}}$$

$$E\left(e_1e_4'\right)=\frac{\mu_{03}}{n'\overline{X}}$$

$$E\left(e_1'e_2\right)=\frac{\mu_{12}}{n'\overline{X}}$$

$$E(e_1'e_3) = \frac{\mu_{21}}{n'\overline{X}}$$

$$E(e_1'e_4) = \frac{\mu_{03}}{n'\overline{X}}$$

$$E(e_1'e_4') = \frac{\mu_{03}}{n'\overline{X}}$$

$$E(e_3e_4) = \frac{1}{n}(\mu_{22} - \mu_{20}\mu_{02})$$

$$E(e_3e_4') = \frac{1}{n'}(\mu_{22} - \mu_{20}\mu_{02})$$

$$E(e_4'^2) = \left(\frac{\beta_2(x)-1}{n'}\right)S_X^4 = \frac{\mu_{02}^2}{n'}\left(\frac{\mu_{04}}{\mu_{02}^2} - 1\right)$$

$$E(e_2e_4) = \frac{\mu_{13}}{n}$$

$$E(e_2e_3) = \frac{\mu_{21}}{n}$$

$$E(e_3e_4) = \frac{\mu_{22}}{n}$$

$$E(e_4e_4') = \left\{\frac{\beta_2(x)-1}{n'}\right\}S_X^4$$

$$E(e_2e_4') = \frac{\mu_{13}}{n'} \tag{2.3}$$

The proposed generalized double sampling regression type estimator represented by $\overline{y}_g$ for estimating the population mean given in (1.1) is

$$\overline{y}_g = g\left(\overline{y}, b, \overline{x}, \overline{x}', s_x^2, s_x'^2, s_y^2\right) \tag{2.4}$$

Expanding $g\left(\bar{y}, b, \bar{x}, \bar{x}', s_x^2, s_x'^2, s_y^2\right)$ about the point $\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right)$ in the third order Taylor's series, we have

$$\begin{aligned}
\bar{y}_g = {} & g\left(\bar{Y}, \beta, \bar{X}, \bar{X}, S_X^2, S_X^2, S_Y^2\right) + \left(\bar{y} - \bar{Y}\right) g_0 + \left(b - \beta\right) g_1 + \left(\bar{x} - \bar{X}\right) g_2 \\
& + \left(\bar{x}' - \bar{X}\right) g_3 + \left(s_x^2 - S_X^2\right) g_4 + \left(s_x'^2 - S_X^2\right) g_5 + \left(s_y^2 - S_Y^2\right) g_6 \\
& + \frac{1}{2!}\Big\{\left(\bar{y} - \bar{Y}\right)^2 g_{00} + \left(b - \beta\right)^2 g_{11} + \left(\bar{x} - \bar{X}\right)^2 g_{22} + \left(\bar{x}' - \bar{X}\right)^2 g_{33} \\
& + \left(s_x^2 - S_X^2\right)^2 g_{44} + \left(s_x'^2 - S_X^2\right)^2 g_{55} + \left(s_y^2 - S_Y^2\right)^2 g_{66} \\
& + 2\left(\bar{y} - \bar{Y}\right)\left(b - \beta\right) g_{01} + 2\left(\bar{y} - \bar{Y}\right)\left(\bar{x} - \bar{X}\right) g_{02} + 2\left(\bar{y} - \bar{Y}\right)\left(\bar{x}' - \bar{X}\right) g_{03} \\
& + 2\left(\bar{y} - \bar{Y}\right)\left(s_x^2 - S_X^2\right) g_{04} + 2\left(\bar{y} - \bar{Y}\right)\left(s_x'^2 - S_X^2\right) g_{05} \\
& + 2\left(\bar{y} - \bar{Y}\right)\left(s_y^2 - S_Y^2\right) g_{06} + 2\left(b - \beta\right)\left(\bar{x} - \bar{X}\right) g_{12} + 2\left(b - \beta\right)\left(\bar{x}' - \bar{X}\right) g_{13} \\
& + 2\left(b - \beta\right)\left(s_x^2 - S_X^2\right) g_{14} + 2\left(b - \beta\right)\left(s_x'^2 - S_X^2\right) g_{15} \\
& + 2\left(b - \beta\right)\left(s_y^2 - S_Y^2\right) g_{16} + 2\left(\bar{x} - \bar{X}\right)\left(\bar{x}' - \bar{X}\right) g_{23} \\
& + 2\left(\bar{x} - \bar{X}\right)\left(s_x^2 - S_X^2\right) g_{24} + 2\left(\bar{x} - \bar{X}\right)\left(s_x'^2 - S_X^2\right) g_{25} \\
& + 2\left(\bar{x} - \bar{X}\right)\left(s_y^2 - S_Y^2\right) g_{26} + 2\left(\bar{x}' - \bar{X}\right)\left(s_x^2 - S_X^2\right) g_{34} \\
& + 2\left(\bar{x}' - \bar{X}\right)\left(s_x'^2 - S_X^2\right) g_{35} + 2\left(\bar{x}' - \bar{X}\right)\left(s_y^2 - S_Y^2\right) g_{36} \\
& + 2\left(s_x^2 - S_X^2\right)\left(s_x'^2 - S_X^2\right) g_{45} + 2\left(s_x^2 - S_X^2\right)\left(s_y^2 - S_Y^2\right) g_{46} \\
& + 2\left(s_x'^2 - S_X^2\right)\left(s_y^2 - S_Y^2\right) g_{56}\Big\} + \frac{1}{3!}\Big\{\left(\bar{y} - \bar{Y}\right)\frac{\partial}{\partial \bar{y}} + \left(b - \beta\right)\frac{\partial}{\partial b} \\
& + \left(\bar{x} - \bar{X}\right)\frac{\partial}{\partial \bar{x}} + \left(\bar{x}' - \bar{X}\right)\frac{\partial}{\partial \bar{x}'} + \left(s_x^2 - S_X^2\right)\frac{\partial}{\partial s_x^2} + \left(s_x'^2 - S_X^2\right)\frac{\partial}{\partial s_x'^2}
\end{aligned}$$

$$+\left(s_y^2-S_Y^2\right)\frac{\partial}{\partial s_y^2}\Bigg\}^3 g\left(\bar{y}^*, b^*, \bar{x}^*, \bar{x}'^*, s_x^{2*}, s_x'^{2*}, s_y^{2*}\right) \qquad (2.4)$$

where g_0, g_1, g_2, g_3, g_4, g_5, g_{00}, g_{11}, g_{22}, g_{02}, g_{03}, g_{04} and g_{05} are defined earlier and

$$g_{33}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x}'^2}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{44}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial\left(s_x^2\right)^2}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{55}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial\left(s_x'^2\right)^2}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{01}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{y}\partial b}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{12}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial b\partial \bar{x}}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{13}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial b\partial \bar{x}'}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{14}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial b\partial s_x^2}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{15}=\left(\frac{\partial^2 g\left(\bar{y},\ b, \bar{x}, \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial b\partial s_x'^2}\right)_{\left(\bar{Y}\ \ \beta, \bar{X}, \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)}$$

$$g_{23} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x} \partial \bar{x}'} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{24} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x} \partial s_x^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{25} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x} \partial s_x'^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{34} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x}' \partial s_x^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{35} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x}' \partial s_x'^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{45} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_y^2\right)}{\partial s_x^2 \partial s_y^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{06} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{y} \partial s_y^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{16} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_y^2\right)}{\partial b \partial s_y^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{26} = \left(\frac{\partial^2\ g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial \bar{x} \partial s_y^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{36} = \left(\frac{\partial^2 g\left(\bar{y},\ b,\ \bar{x},\ \bar{x}',\ s_x^2,\ s_y^2\right)}{\partial \bar{x}' \partial s_y^2} \right)_{\left(\bar{Y}\ \ \beta,\ \bar{X},\ \bar{X},\ S_X^2,\ S_X^2,\ S_Y^2\right)},$$

$$g_{46}=\left(\frac{\partial^2 g\left(\bar{y},\ b,\bar{x},\bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial s_x^2 \partial s_y^2}\right)_{\left(\bar{Y}\ \ \beta,\,\bar{X},\,\bar{X},\ S_X^2,\,S_X^2,\ S_Y^2\right)},$$

$$g_{56}=\left(\frac{\partial^2 g\left(\bar{y},\ b,\bar{x},\bar{x}',\ s_x^2,\ s_y^2\right)}{\partial s_x'^2 \partial s_y^2}\right)_{\left(\bar{Y}\ \ \beta,\,\bar{X},\,\bar{X},\ S_X^2,\,S_X^2,\ S_Y^2\right)},$$

$$g_{66}=\left(\frac{\partial^2 g\left(\bar{y},\ b,\bar{x},\bar{x}',\ s_x^2,\ s_x'^2,\ s_y^2\right)}{\partial\left(s_y^2\right)^2}\right)_{\left(\bar{Y}\ \ \beta,\,\bar{X},\,\bar{X},\ S_X^2,\,S_X^2,\ S_Y^2\right)}$$

and $\bar{y}^* = \bar{Y}+h\left(\bar{y}-\bar{Y}\right),\ b=\beta+h\left(b-\beta\right),\ \bar{x}^* = \bar{X}+h\left(\bar{x}-\bar{X}\right),$

$\bar{x}'^* = \bar{X}+h\left(\bar{x}'-\bar{X}\right),\ s_x^{2^*} = S_X^2+h\left(s_x^2-S_X^2\right),\ s_x'^{2^*} = S_X^2+h\left(s_x'^2-S_X^2\right)$ and

$s_y^{2^*} = S_Y^2+h\left(s_y^2-S_Y^2\right)$ for $0<h<1.$

Under the conditions (i) to (xi) given above, up to terms of order $O\left(1/n\right)$, the above proposed generalized double sampling estimator in terms of e_i's, $i=0,1,2,3,4$ reduces to

$$\begin{aligned}
\bar{y}_g-\bar{Y} \ = \ & \bar{Y}e_0-\beta\bar{X}e_1+\beta\bar{X}e'+e_4 g_4-e_4' g_4+e_3 g_6+\frac{1}{2!}\Big\{\bar{X}^2 e_1^2 g_{22}+\bar{X}^2 e_1'^2 g_{33} \\
& +e_4^2 g_{44}+e_4'^2 g_{55}+e_3^2 g_{66}+2\bar{Y}e_0\beta\left(\frac{e_2}{S_{YX}}-\frac{e_4}{S_X^2}\right)g_{01}+2\overline{YX}e_0 e_1 g_{02} \\
& -2\overline{YX}e_0 e_1' g_{03}+2\bar{Y}e_0 e_4 g_{04}-2\bar{Y}e_0 e_4' g_{05}+2\bar{Y}e_0 e_3 g_{06} \\
& +2\beta\left(\frac{e_2}{S_{YX}}-\frac{e_4}{S_X^2}\right)\bar{X}e_1 g_{12}+2\beta\left(\frac{e_2}{S_{YX}}-\frac{e_4}{S_X^2}\right)\bar{X}e_1' g_{13} \\
& +2\beta\left(\frac{e_2}{S_{YX}}-\frac{e_3}{S_X^2}\right)e_4 g_{14}+2\beta\left(\frac{e_2}{S_{YX}}-\frac{e_4}{S_X^2}\right)e_4' g_{15} \\
& +2\beta\left(\frac{e_2}{S_{YX}}-\frac{e_4}{S_X^2}\right)e_3 g_{16}+2\bar{X}^2 e_1 e_1' g_{23}+2\bar{X}e_1 e_4 g_{24}+2\bar{X}e_1 e_4' g_{25}
\end{aligned}$$

$$+2\bar{X}e_1e_3g_{26}+2\bar{X}e_1'e_4g_{34}+2\bar{X}e_1'e_4'g_{35}+2\bar{X}e_1'e_3g_{36}+2e_4e_4'g_{45}$$
$$+2e_3e_4g_{46}+2e_3e_4'g_{56}\Big\}\quad(2.5)$$

Taking expectation on both the sides of (2.5), the bias of $\bar{y}_g$ up to terms of order $O(1/n)$ is given by

$$\text{Bias}\left(\bar{y}_g\right) = \left\{E\left(\bar{y}_g\right)-\bar{Y}\right\}$$
$$= \frac{1}{2}\Bigg[\frac{\mu_{02}}{n}g_{22}+\frac{\mu_{02}}{n'}g_{33}+\left\{\frac{\beta_2(x)-1}{n}\right\}S_X^4g_{44}+\left\{\frac{\beta_2(x)-1}{n'}\right\}S_X^4g_{55}$$
$$+\left\{\frac{\beta_2(y)-1}{n}\right\}S_Y^4g_{66}+2\beta\left(\frac{\mu_{21}}{nS_{YX}}-\frac{\mu_{12}}{nS_X^2}\right)g_{01}+2\frac{\mu_{11}}{n}g_{02}$$
$$-2\frac{\mu_{11}}{n'}g_{03}+2\frac{\mu_{12}}{n}g_{04}-2\frac{\mu_{12}}{n'}g_{05}+2\frac{\mu_{30}}{n}g_{06}$$
$$+2\beta\left(\frac{\mu_{12}}{nS_{YX}}-\frac{\mu_{03}}{nS_X^2}\right)g_{12}+2\beta\left(\frac{\mu_{12}}{n'S_{YX}}-\frac{\mu_{03}}{n'S_X^2}\right)g_{13}$$
$$+2\beta\left(\frac{\mu_{13}}{nS_{YX}}-\frac{\mu_{22}}{nS_X^2}\right)g_{14}+2\beta\left(\frac{\mu_{13}}{n'S_{YX}}-\left\{\frac{\beta_2(x)-1}{n'}\right\}S_X^2\right)g_{15}$$
$$+2\beta\left(\frac{\mu_{21}}{nS_{YX}}-\frac{\mu_{22}}{nS_X^2}\right)g_{16}+2\frac{\mu_{02}}{n'}g_{23}+2\frac{\mu_{03}}{n}g_{24}+2\frac{\mu_{03}}{n'}g_{25}$$
$$+2\frac{\mu_{21}}{n}g_{26}+2\frac{\mu_{03}}{n'}g_{34}+2\frac{\mu_{03}}{n'}g_{35}+2\frac{\mu_{21}}{n'}g_{36}$$
$$+2\left\{\frac{\beta_2(x)-1}{n'}\right\}S_X^4g_{45}+2\frac{\mu_{22}}{n}g_{46}+\frac{2}{n'}\left(\mu_{22}-\mu_{20}\mu_{02}\right)g_{56}\Bigg]\quad(2.6)$$

Now squaring both sides of (2.5) and taking expectation, the mean square error of $\bar{y}_g$ up to terms of order $O(1/n)$ is given by

$$\text{MSE}\left(\bar{y}_g\right) = \left\{E\left(\bar{y}_g\right)-\bar{Y}\right\}^2$$
$$= \bar{Y}^2E\left(e_0^2\right)+\beta^2\bar{X}^2E\left(e_1^2\right)+\beta^2\bar{X}^2E\left(e_1'^2\right)+g_4^2E\left(e_4^2\right)+g_4^2E\left(e_4'^2\right)$$

$$+g_6^2 E\left(e_3^2\right) - 2\beta\overline{Y}\overline{X}E\left(e_0 e_1\right) + 2\beta\overline{Y}\overline{X}E\left(e_0 e_1'\right) + 2\overline{Y}g_4 E\left(e_0 e_4\right)$$

$$-2\overline{Y}g_4 E\left(e_0 e_4'\right) + 2\overline{Y}g_6 E\left(e_0 e_3\right) - 2\beta^2 \overline{X}^2 E\left(e_1 e_1'\right) - 2\beta\overline{X}g_4 E\left(e_1 e_4\right)$$

$$+2\beta\overline{X}g_4 E\left(e_1 e_4'\right) - 2\beta\overline{X}g_6 E\left(e_1 e_3\right) + 2\beta\overline{X}g_4 E\left(e_1' e_4\right)$$

$$-2\beta\overline{X}g_4 E\left(e_1' e_4'\right) + 2\beta\overline{X}g_6 E\left(e_1' e_3\right) - 2g_4^2 E\left(e_4 e_4'\right)$$

$$+2g_4 g_6 E\left(e_3 e_4\right) - 2g_4 g_6 E\left(e_3 e_4'\right)$$

using values of the expectation given in (2.2) and (2.3), we have

$$\text{MSE}\left(\overline{y}_g\right) = \frac{\mu_{20}}{n} + \beta^2 \mu_{02}\left(\frac{1}{n} - \frac{1}{n'}\right) - 2\beta\mu_{11}\left(\frac{1}{n} - \frac{1}{n'}\right)$$

$$+\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\left\{\beta_2(x) - 1\right\} S_X^4\right] g_4^2 + \left[\frac{1}{n}\left\{\beta_2(y) - 1\right\} S_Y^4\right] g_6^2$$

$$-2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\left(\beta\mu_{03} - \mu_{12}\right)\right] g_4 - 2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\beta\mu_{21} - \frac{\mu_{30}}{n}\right] g_6$$

$$+2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\left(\mu_{22} - \mu_{20}\mu_{02}\right)\right] g_4 g_6$$

or $\text{MSE}\left(\overline{y}_g\right) = \frac{\mu_{20}}{n} + \beta^2 \mu_{02}\left(\frac{1}{n} - \frac{1}{n'}\right) - 2\beta\mu_{11}\left(\frac{1}{n} - \frac{1}{n'}\right) + \delta_1 g_4^2$

$$+\delta_2 g_6^2 - 2\delta_3 g_4 - 2\delta_4 g_6 + 2\delta_5 g_4 g_6 \qquad (2.7)$$

where, $\delta_1 = \left[\left(\frac{1}{n} - \frac{1}{n'}\right)\left\{\beta_2(x) - 1\right\} S_X^4\right]$

$$\delta_2 = \left[\frac{1}{n}\left\{\beta_2(y) - 1\right\} S_Y^4\right]$$

$$\delta_3 = \left[\left(\frac{1}{n} - \frac{1}{n'}\right)\left(\beta\mu_{03} - \mu_{12}\right)\right]$$

$$\delta_4 = \left[\left(\frac{1}{n} - \frac{1}{n'}\right)\beta\mu_{21} - \frac{\mu_{30}}{n}\right]$$

$$\delta_5 = \left[\left(\frac{1}{n} - \frac{1}{n'}\right)(\mu_{22} - \mu_{20}\mu_{02})\right]$$

which attains the minimum for the optimum values

$$g_4 = \frac{(\delta_2\delta_3 - \delta_4\delta_5)}{(\delta_1\delta_2 - \delta_5^2)} \tag{2.8}$$

and

$$g_6 = \frac{(\delta_1\delta_4 - \delta_3\delta_5)}{(\delta_1\delta_2 - \delta_5^2)} \tag{2.9}$$

Substituting the values of g_4 and g_6 given by (2.8) and (2.9) in (2.7), we get the minimum mean square error of $\bar{y}_g$ to be

$$\begin{aligned} \text{MSE}(\bar{y}_g)_{\min} &= \frac{\mu_{20}}{n} + \beta^2\mu_{02}\left(\frac{1}{n} - \frac{1}{n'}\right) - 2\beta\mu_{11}\left(\frac{1}{n} - \frac{1}{n'}\right) \\ &\quad - \left\{\frac{(\delta_2\delta_3^2 + \delta_1\delta_4^2 - 2\delta_3\delta_4\delta_5)}{(\delta_1\delta_2 - \delta_5^2)}\right\} \end{aligned} \tag{2.10}$$

3. Efficiency Comparison

(i) **General estimator of mean in case of SRSWOR:** The general estimator of mean in case of SRSWOR is $\hat{\bar{y}}_{wor} = \bar{y}$ with

$$MSE(\hat{\bar{y}}) = \frac{\mu_{20}}{n} \tag{3.1}$$

It is clear that the proposed estimator is more efficient than the estimator $\hat{\bar{y}}_{wor}$ based on simple random sampling when no auxiliary information is used.

(ii) **Usual double sampling regression estimator:** The usual double sampling regression estimator is $\bar{y}_{ld} = \bar{y} + b(\bar{x}' - \bar{x})$ with

$$MSE(\bar{y}_{ld}) = \frac{\mu_{20}}{n} - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{11}^2}{\mu_{02}} \tag{3.2}$$

It is clear that the proposed estimator is more efficient than the usual double sampling regression estimator where the auxiliary information already is in use.

4. Empirical Study

To illustrate the performance of the proposed estimator, let us consider the following data

Population I: Cochran (1977, Page Number- 181)

y : Paralytic Polio Cases 'placebo' group

x : Paralytic Polio Cases in not inoculated group

μ_{02} = 71.8650173, μ_{20} = 9.889273356, μ_{11} = 19.4349481, μ_{12} = 346.3174191,

μ_{03} = 1453.077703, μ_{40} = 424.1846721, μ_{21} = 94.21286383,

μ_{22} = 3029.312542, μ_{30} = 47.34479951, μ_{04} = 46132.5679, $\bar{y}$ = 2.588235294,

$\bar{x}$ = 8.370588235, S_x = 8.477323711, S_y = 3.144721507, ρ =0.729025009,

$\beta_2(y)$= 4.337367369, $\beta_2(x)$= 8.932490454, C_X = 1.012751251,

C_Y =1.215006037, β = 0.270436839, n = 34, n' =50 (say).

$MSE\left(\hat{\bar{y}}_{wor}\right)$= 0.290860981, $MSE\left(\bar{y}_{ld}\right)$ = 0.241393443 and $MSE\left(\bar{y}_g\right)$min = 0.075121792.

PRE of the proposed estimator $\bar{y}_g$ over $\hat{\bar{y}}_{wor}$ = 387.1858936.

PRE of the proposed estimator $\bar{y}_g$ over $\bar{y}_{ld}$ = 321.336109.

Population II: Mukhopadhyay (2012, Page Number - 104)

y : Quality of raw materials (in lakhs of bales)

x : Number of labourers (in thousands)

μ_{02} = 9704.4475, μ_{20} = 90.95, μ_{11} = 612.725, μ_{12} = 93756.3475,

μ_{03} = 988621.5173, μ_{40} = 35456.4125, μ_{21} = 11087.635, μ_{22} = 2893630.349,

μ_{30} = 1058.55, μ_{04} = 341222548.2, $\bar{y}$ = 41.5, $\bar{x}$ = 441.95, S_x = 98.51115419,

S_y = 9.536770942, ρ =0.652197067, $\beta_2(y)$= 4.286367314,

$\beta_2(x)$= 3.623231573, C_X = 0.22290113, C_Y =0.229801709,

β =0.063138576, n = 20, n' =35 (say).

$MSE\left(\hat{\bar{y}}_{wor}\right)$= 4.5475, $MSE\left(\bar{y}_{ld}\right)$= 3.718501766 and $MSE\left(\bar{y}_g\right)$min = 2.588480264.

PRE of the proposed estimator $\bar{y}_g$ over $\hat{\bar{y}}_{wor}$ = 175.6822357.

PRE of the proposed estimator $\bar{y}_g$ over $\bar{y}_{ld}$ = 143.6557898.

Population III: Murthy (1967, Page Number - 398)

y : Number of absentees

x : Number of workers

μ_{02} = 1299.318551, μ_{20} = 42.13412655, μ_{11} = 154.6041103,

μ_{12} = 5086.694392, μ_{03} = 32025.12931, μ_{40} = 11608.18508,

μ_{21} = 1328.325745, μ_{22} = 148328.4069, μ_{30} = 425.9735118,

μ_{04} = 4409987.245, $\bar{y}$ = 9.651162791, $\bar{x}$ = 79.46511628, S_x = 36.04606151,

S_y = 6.491080538, ρ = 0.660763765, $\beta_2(y)$= 6.53877409,

$\beta_2(x)$= 2.612197776, C_X = 0.453608617, C_X = 0.672569791,

β = 0.118988612, n = 43, n' =50 (say).

$MSE\left(\hat{\bar{y}}_{wor}\right)$= 0.979863408, $MSE\left(\bar{y}_{ld}\right)$= 0.919969037 and $MSE\left(\bar{y}_g\right)$min = 0.525977442.

PRE of the proposed estimator $\bar{y}_g$ over $\hat{\bar{y}}_{wor}$ = 186.2938083.

PRE of the proposed estimator $\bar{y}_g$ over $\bar{y}_{ld}$ = 174.9065573.

Population IV: Singh and Chaudhary (1997, Page Number - 176)

y : Total number of guava trees

x : Area under guava orchard (in acres)

μ_{02} = 12.50056686, μ_{20} = 187123.9172, μ_{11} = 1377.39858,

μ_{12} = 4835.465464, μ_{03} = 37.09863123, μ_{40} = 1.48935E+11,

μ_{21} = 712662.4414, μ_{22} = 8747904.451, μ_{30} = 100476814.5,

μ_{04} = 540.1635491, $\bar{y}$ = 746.9230769, $\bar{x}$ = 5.661538462, S_x = 3.535614072,

S_y = 432.5782209, ρ = 0.900596235, $\beta_2(y)$ = 4.253426603,

$\beta_2(x)$ = 3.456733187, C_X = 0.624497051, C_Y = 0.579146949,

β = 110.1868895, n = 13, n' =30 (say).

$MSE\left(\hat{\bar{y}}_{wor}\right)$ = 14394.14747, $MSE\left(\bar{y}_{ld}\right)$ = 7778.476942 and $MSE\left(\bar{y}_g\right)$min = 4262.795512.

PRE of the proposed estimator $\bar{y}_g$ over $\hat{\bar{y}}_{wor}$ = 337.6691993.

PRE of the proposed estimator $\bar{y}_g$ over $\bar{y}_{ld}$ = 182.473612.

Population V: Singh and Chaudhary (1997, Page Number: 154-155)

y : Number of milch animals in survey

x : Number of milch animals in census

μ_{02} = 431.5847751, μ_{20} = 270.9134948, μ_{11} = 247.3944637,

μ_{12} = 3119.839406, μ_{03} = 5789.778954, μ_{40} = 154027.4827,

μ_{21} = 2422.297374, μ_{22} = 210594.3138, μ_{30} = 2273.46265,

μ_{04} = 508642.4447, $\bar{y}$ = 1133.294118, $\bar{x}$ = 1140.058824, S_x = 20.77461853,

S_y = 16.45945002, ρ = 0.723505104, $\beta_2(y)$ = 2.098635139,

$\beta_2(x)$ = 2.730740091, C_X = 0.018222409, C_Y = 0.014523547,

β = 0.573223334, n = 17, n' =30 (say).

$MSE\left(\hat{\bar{y}}_{wor}\right)$ = 15.93609, $MSE\left(\bar{y}_{ld}\right)$ = 12.32127 and $MSE\left(\bar{y}_g\right)$min = 9.857806144.

PRE of the proposed estimator $\bar{y}_g$ over $\hat{\bar{y}}_{wor}$ = 161.6595792.

PRE of the proposed estimator $\bar{y}_g$ over $\bar{y}_{ld}$ = 124.9899313.

5. Conclusions

From (2.10) it is clear that the proposed double sampling generalized estimator is more efficient than the estimator $\hat{\bar{y}}_{wor}$ based on simple random sampling when no auxiliary information is used and is also more efficient than the usual double sampling regression estimator $\bar{y}_{ld}$ of mean where the auxiliary information already is in use.

From (2.8) and (2.9), the mean squared error of the estimator $\bar{y}_g$ is minimized for the optimum values

$$g_4 = \frac{(\delta_2\delta_3 - \delta_4\delta_5)}{(\delta_1\delta_2 - \delta_5^2)} \tag{5.1}$$

$$g_6 = \frac{(\delta_1\delta_4 - \delta_3\delta_5)}{(\delta_1\delta_2 - \delta_5^2)} \tag{5.2}$$

The optimum values involving some unknown parameters may not be known in advance for practical purposes; hence the alternative is to replace the unknown parameters of the optimum values by their unbiased estimators giving estimators depending upon estimated optimum values.

Acknowledgement

The author wish to thank to the learned referees for comments that greatly improved the manuscript. This research was supported by Northern Regional College Bureua (NRCB), Delhi, University Grants Commission, Ministry of HRD, Government of India through a Minor Research project No.: [F.No. 8-4(14)/2015 (MRP/NRCB)].

References

[1] Cochran, W.G. (1977): Sampling Techniques, 3rd edition, John Wiley and Sons, New York.

[2] Des Raj (1968): Sampling Theory, McGraw- Hill, New York.

[3] Murthy, M (1967): Sampling Theory and Methods, 1st edition, Calcutta Statistical Publishing Society, Kolkata, India.

[4] Mukhopadhyay, Parimal (2012): Theory and Methods of Survey and Sampling, 2nd edition, PHI Learning Private Limited, New Delhi, India.

[5] Singh, Daroga and Chaudhary, F. S. (1997): Theory and Analysis of Sampling Survey Designs, New Age International Publishers, New Delhi, India.

[6] Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S. And Asok, C. (1984): Sampling Theory of Surveys with Applications, 3rd Edition, Ames, Iowa (USA) and Indian Society of Agricultural Statistics, New Delhi, India.

Peeyush Misra
Department of Statistics,
D.A.V.(P.G.) College,
Dehradun- 248001,
Uttarakhand (India)
Email: dr.pmisra.dav@gmail.com