RESEARCH PAPER | STATISTICS | Volume : 6 | Issue : 1 | JANUARY 2016 | ISSN - 2249-555X

A Generalized Double Sampling Estimator of Population Mean Using Variable And Attribute Both

KEYWORDS: Auxiliary information, Bias, Mean square error and Taylor's Series Expansion.

Peeyush Misra
Department of Statistics, D.A.V.(P.G.) College, Dehradun- 248001, Uttarakhand, India

R. Karan Singh
Department of Statistics, University of Lucknow, Lucknow, Uttar Pradesh (India)

ABSTRACT *For double sampling procedure, the problem of estimating population mean by using auxiliary information in the form of attribute and variable both is considered. A generalized class of estimator is proposed and the bias and mean square error are obtained. It is shown that the proposed generalized class of estimator is superior to some of the previously studied estimators under the minimum mean square error criterion.*

1. Introduction

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' by simple random sampling without replacement scheme at both the phases.

At first phase sample of size n', only the auxiliary variable X and auxiliary attribute ϕ are observed but at the second phase sample of size n, the study variable Y, auxiliary variable X and auxiliary attribute ϕ all are observed.

Let us denote by $\overline{Y}, \overline{X}$ and P as the population mean of study variable, population mean of auxiliary variable and population mean of auxiliary attribute ϕ i.e.

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i,\ \overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \text{ and } P = \frac{1}{N}\sum_{i=1}^{N}\phi_i \quad .$$

S_Y^2, S_X^2 and S_P^2 are the population variance of study variable, population variance of auxiliary variable and population variance of auxiliary attribute and are given by

$$S_Y^2 = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2,\quad S_X^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2 \quad \text{and } S_P^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\phi_i - P\right)^2 .$$

Let $(\bar{y}, \bar{x}, p)$ based on second phase sample of size n be the sample mean estimators of population means $\left(\overline{Y}, \overline{X}, P\right)$ of (Y, X, P) respectively and $\left(\bar{x}', p'\right)$ based on the first phase n' sample values on (X, P) be the mean per unit estimator of $\left(\overline{X}, P\right)$ respectively. Also s_y^2, s_x^2 and s_p^2 are the sample variance of the study variable, auxiliary variable and auxiliary attribute respectively based on the second phase sample of size n. Hence we are given with

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i\ ,\ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ and } p = \frac{1}{n}\sum_{i=1}^{n}\phi_i$$

$s_y^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$, $s_x^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ and $s_p^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\phi_i - p\right)^2$.

For estimating population mean $\bar{Y}$ of the study (main) variable Y, a generalized double sampling estimator $\bar{y}_{gd}$ as the bounded function of $(\bar{y}, \bar{x}, \bar{x}', p, p')$ is proposed as

$$\bar{y}_{gd} = g(\bar{y}, \bar{x}, \bar{x}', p, p') \tag{1.1}$$

satisfying the validity conditions of Taylor's series expansion such that

(i) $g\left(\bar{Y}, \bar{X}, \bar{X}, P, P\right) = \bar{Y}$ (1.2)

(ii) first order partial differential coefficient of $g(\bar{y}, \bar{x}, \bar{x}', p, p')$ with respect to $\bar{y}$ at $T = \left(\bar{Y}, \bar{X}, \bar{X}, P, P\right)$ is unity, that is

$$g_0 = \left(\frac{\partial}{\partial \bar{y}} g(\bar{y}, \bar{x}, \bar{x}', p, p')\right)_T = 1 \tag{1.3}$$

(iii) $g_{00} = \left(\frac{\partial^2}{\partial \bar{y}^2} g(\bar{y}, \bar{x}, \bar{x}', p, p')\right)_T = 0$ (1.4)

(iv) $g_1 = -g_2$ (1.5)

for g_1 and g_2 being the first order partial derivatives of $g(\bar{y}, \bar{x}, \bar{x}', p, p')$ with respect to $\bar{x}$ and $\bar{x}'$ respectively at the point $T = \left(\bar{Y}, \bar{X}, \bar{X}, P, P\right)$, that is

$$g_1 = \left(\frac{\partial}{\partial \bar{x}} g(\bar{y}, \bar{x}, \bar{x}', p, p')\right)_T$$

$$g_2 = \left(\frac{\partial}{\partial \bar{x}'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

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

for $$g_{01} = \left(\frac{\partial^2}{\partial \bar{y} \partial \bar{x}} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

$$g_{02} = \left(\frac{\partial^2}{\partial \bar{y} \partial \bar{x}'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

(vi) $g_3 = -g_4$ (1.7)

for g_3 and g_4 being the first order partial derivatives of $g(\bar{y}, \bar{x}, \bar{x}', p, p')$ with respect to p and p' respectively at the point $T = (\bar{Y}, \bar{X}, \bar{X}, P, P)$, that is

$$g_3 = \left(\frac{\partial}{\partial p} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

$$g_4 = \left(\frac{\partial}{\partial p'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$ and

(vii) $g_{03} = -g_{04}$ (1.8)

for $$g_{03} = \left(\frac{\partial^2}{\partial \bar{y} \partial p} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$ and

$$g_{04} = \left(\frac{\partial^2}{\partial \bar{y} \partial p'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T .$$

It can be seen that the proposed generalized double sampling estimator is more efficient than the commonly used estimators available in the literatures as mean per unit estimator, ratio estimator, ratio estimator by Naik and Gupta (1996) and exponential ratio estimator by and Tuteja (1991).

2. Bias and Mean Square Error of the Proposed Estimator

The proposed generalized double sampling estimator is given by

$$\bar{y}_{gd} = g(\bar{y}, \bar{x}, \bar{x}', p, p') \ .$$

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

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

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

$$p = P(1+e_2)$$

$$p' = P(1+e_2')$$

with $E(e_0) = E(e_1) = E(e_1') = E(e_2) = E(e_2') = 0$ (2.1)

$$E(e_0^2) = f_n \ C_Y^2$$

$$E(e_1^2) = f_n \ C_X^2$$

$$E(e_2^2) = f_n \ C_P^2$$

$$E(e_0 e_1) = f_n \ \rho_{YX} C_Y C_X$$

$$E(e_0 e_2) = f_n \ \rho_{YP} C_Y C_P$$

$$E(e_1 e_2) = f_n \ \rho_{XP} C_X C_P$$

$$E(e_1'^2) = f_{n'} \ C_X^2$$

$$E(e_2'^2) = f_{n'} \ C_P^2$$

$$E(e_0e_1') = f_{n'}\ \rho_{YX}C_YC_X$$

$$E(e_0e_2') = f_{n'}\ \rho_{YP}C_YC_P$$

$$E(e_1e_1') = f_{n'}\ C_X^2$$

$$E(e_1e_2') = f_{n'}\ \rho_{XP}C_XC_P$$

$$E(e_1'e_2) = f_{n'}\ \rho_{XP}C_XC_P$$

$$E(e_1'e_2') = f_{n'}\ \rho_{XP}C_XC_P \quad \text{and}$$

$$E(e_2e_2') = f_{n'}\ C_P^2 \qquad (2.2)$$

where $f_n = \left(\frac{1}{n} - \frac{1}{N}\right)$, $f_{n'} = \left(\frac{1}{n'} - \frac{1}{N}\right)$, $C_y{}^2 = \frac{s_y{}^2}{\overline{Y}^2}$, $C_x{}^2 = \frac{s_x{}^2}{\overline{X}^2}$, $C_p{}^2 = \frac{s_\phi{}^2}{P^2}$; ρ_{yx}, ρ_{yp} and ρ_{xp} are the correlation coefficients between (y,x), (y,p) and (x,p) respectively.

Now expanding $g(\bar{y},\bar{x},\bar{x}',p,p')$ in the third order Taylor's series about the point $T = (\overline{Y},\overline{X},\overline{X},P,P)$, we have

$$\bar{y}_{gd} = g(\overline{Y},\overline{X},\overline{X},P,P) + (\bar{y}-\overline{Y})g_0 + (\bar{x}-\overline{X})g_1$$

$$+ (\bar{x}'-\overline{X})g_2 + (p-P)g_3 + (p'-P)g_4$$

$$+ \frac{1}{2!}\Big\{(\bar{y}-\overline{Y})^2 g_{00} + (\bar{x}-\overline{X})^2 g_{11} + (\bar{x}'-\overline{X})^2 g_{22}$$

$$+ (p-P)^2 g_{33} + (p'-P)^2 g_{44} + 2(\bar{y}-\overline{Y})(\bar{x}-\overline{X})g_{01}$$

$$+ 2(\bar{y}-\overline{Y})(\bar{x}'-\overline{X})g_{02} + 2(\bar{y}-\overline{Y})(p-P)g_{03}$$

$$+ 2(\bar{y}-\overline{Y})(p'-P)g_{04} + 2(\bar{x}-\overline{X})(\bar{x}'-\overline{X})g_{12}$$

$$+ 2(\bar{x}-\overline{X})(p-P)g_{13} + 2(\bar{x}-\overline{X})(p'-P)g_{14}$$

$$+2\left(\bar{x}'-\bar{X}\right)\left(p-P\right)g_{23}+2\left(\bar{x}'-\bar{X}\right)\left(p'-P\right)g_{24}+2\left(p-P\right)\left(p'-P\right)g_{34}\}$$

$$+\frac{1}{3!}\left\{\frac{\partial}{\partial \bar{y}}+\frac{\partial}{\partial \bar{x}}+\frac{\partial}{\partial \bar{x}'}+\frac{\partial}{\partial p}+\frac{\partial}{\partial p'}\right\}^3 g\left(\bar{y}^*,\bar{x}^*,\bar{x}'^*,p^*,p'^*\right)$$

where $\bar{y}^* = \bar{Y} + h\left(\bar{y}-\bar{Y}\right)$

$$\bar{x}^* = \bar{X} + h\left(\bar{x}-\bar{X}\right)$$

$$\bar{x}'^* = \bar{X} + h\left(\bar{x}'-\bar{X}\right)$$

$$p^* = P + h\left(p-P\right)$$

$$p'^* = P + h\left(p'-P\right),\ 0<h<1$$

and g_0, g_1, g_2, g_3, g_4, g_{00}, g_{01}, g_{02}, g_{03}, g_{04} are already defined in equations (1.3) to (1.8) and

$$g_{11} = \left(\frac{\partial^2}{\partial \bar{x}^2} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{22} = \left(\frac{\partial^2}{\partial \bar{x}'^2} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{33} = \left(\frac{\partial^2}{\partial p^2} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{44} = \left(\frac{\partial^2}{\partial p'^2} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{04} = \left(\frac{\partial^2}{\partial \bar{y}\partial p'} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{12} = \left(\frac{\partial^2}{\partial \bar{x}\partial \bar{x}'} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{13} = \left(\frac{\partial^2}{\partial \bar{x}\partial p} g\left(\bar{y},\bar{x},\bar{x}',p,p'\right)\right)_T$$

$$g_{14} = \left(\frac{\partial^2}{\partial \bar{x} \partial p'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

$$g_{23} = \left(\frac{\partial^2}{\partial \bar{x}' \partial p} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

$$g_{24} = \left(\frac{\partial^2}{\partial \bar{x}' \partial p'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T$$

$$g_{34} = \left(\frac{\partial^2}{\partial p \partial p'} g(\bar{y}, \bar{x}, \bar{x}', p, p') \right)_T .$$

Now, using the conditions given from (1.2) to (1.8), we have

$$\bar{y}_{gd} - \bar{Y} = \bar{Y}e_0 + \bar{X}e_1 g_1 - \bar{X}e_1' g_1 + Pe_2 g_3 - Pe_2' g_3$$

$$+ \frac{1}{2!}\Big\{ \bar{X}^2 e_1^2 g_{11} + \bar{X}^2 e_1'^2 g_{22} + P^2 e_2^2 g_{33}$$

$$+ P^2 e_2'^2 g_{44} + 2\bar{Y}\bar{X}e_0 e_1 g_{01} - 2\bar{Y}\bar{X}e_0 e_1' g_{01}$$

$$+ 2\bar{Y}Pe_0 e_2 g_{03} - 2\bar{Y}Pe_0 e_2' g_{03} + 2\bar{X}^2 e_1 e_1' g_{12}$$

$$+ 2\bar{X}Pe_1 e_2 g_{13} + 2\bar{X}Pe_1 e_2' g_{14} + 2\bar{X}Pe_1' e_2 g_{23}$$

$$+ 2\bar{X}Pe_1' e_2' g_{24} + 2P^2 e_2 e_2' g_{34} \Big\}$$

$$+ \frac{1}{3!}\left\{ \frac{\partial}{\partial \bar{y}} + \frac{\partial}{\partial \bar{x}} + \frac{\partial}{\partial \bar{x}'} + \frac{\partial}{\partial p} + \frac{\partial}{\partial p'} \right\}^3 g\left(\bar{y}^*, \bar{x}^*, \bar{x}'^*, p^*, p'^*\right) \tag{2.3}$$

Now, taking expectation on both the sides of (2.3) and using the values of the expectations given in (2.1) and (2.2), the bias in $\bar{y}_{gd}$ to the first degree of approximation is given by

Bias in $\left(\bar{y}_{gd}\right) = E\left(\bar{y}_{gd}\right) - \bar{Y}$

$$= \frac{1}{2!}\Big\{ \bar{X}^2 \left(f_n g_{11} + f_{n'} g_{22} + 2 f_{n'} g_{12} \right) C_X^2$$

$$+P^2\left(f_n g_{33}+f_{n'}g_{44}\right)C_P^2+2P^2C_XC_P\rho_{XP}f_{n'}g_{34}$$

$$+2\overline{Y}\,\overline{X}C_YC_X\,\rho_{YX}\left(f_n-f_{n'}\right)g_{01}+2\overline{Y}PC_YC_P\,\rho_{YP}\left(f_n-f_{n'}\right)g_{03}$$

$$+2\overline{X}PC_XC_P\,\rho_{XP}\left(f_ng_{13}+f_{n'}g_{14}+f_{n'}g_{23}+f_{n'}g_{24}\right)\Big\}\quad(2.4)$$

Now, squaring (2.3) on both the sides and then taking expectation, the mean square error of $\bar{y}_{gd}$ to the first degree of approximation is given by

$$MSE\ \left(\bar{y}_{gd}\right)=E\left(\bar{y}_{gd}-\overline{Y}\right)^2$$

$$=\ E\left\{\overline{Y}e_0+\overline{X}e_1g_1-\overline{X}e_1'g_1+Pe_2g_3-Pe_2'g_3\right\}^2$$

$$=\ \overline{Y}^2E\left(e_0^2\right)+\overline{X}^2g_1^2E\left(e_1^2\right)+\overline{X}^2g_1^2E\left(e_1'^2\right)-2\overline{X}^2g_1^2E\left(e_1e_1'\right)$$

$$+P^2g_3^2E\left(e_2^2\right)+P^2g_3^2E\left(e_2'^2\right)-2P^2g_3^2E\left(e_2e_2'\right)$$

$$+2\overline{Y}\,\overline{X}g_1E\left(e_0e_1\right)-2\overline{Y}\,\overline{X}g_1E\left(e_0e_1'\right)+2\overline{Y}\,Pg_3E\left(e_0e_2\right)$$

$$-2\overline{Y}\,Pg_3E\left(e_0e_2'\right)+2\,\overline{X}Pg_1g_3E\left(e_1e_2\right)+2\overline{X}Pg_1g_3E\left(e_1'e_2'\right)$$

$$-2\overline{X}\,Pg_1g_3E\left(e_1e_2'\right)-2\,\overline{X}Pg_1g_3E\left(e_1'e_2\right)$$

Using the values of the expectations given in (2.2), the mean square error is given by

$$MSE\ \left(\bar{y}_{gd}\right)=\overline{Y}^2f_nC_Y^2+\overline{X}^2f_nC_X^2g_1^2+\overline{X}^2f_{n'}C_X^2g_1^2$$

$$-2\overline{X}^2f_{n'}C_X^2g_1^2+P^2f_nC_P^2g_3^2+P^2f_{n'}C_P^2g_3^2$$

$$-2P^2f_{n'}C_P^2g_3^2+2\overline{Y}\,\overline{X}f_n\rho_{YX}C_YC_Xg_1-2\overline{Y}\,\overline{X}f_{n'}\rho_{YX}C_YC_Xg_1$$

$$+2\overline{Y}Pf_n\rho_{YP}C_YC_Pg_3-2\overline{Y}\,Pf_{n'}\rho_{YP}C_YC_Pg_3+2\overline{X}Pf_n\rho_{XP}C_XC_Pg_1g_3$$

$$+2\overline{X}Pf_{n'}\rho_{XP}C_XC_Pg_1g_3-2\overline{X}\,Pf_{n'}\rho_{XP}C_XC_Pg_1g_3-2\overline{X}Pf_{n'}\rho_{XP}C_XC_Pg_1g_3$$

$$=f_n\overline{Y}^2C_Y^2+\overline{X}^2C_X^2\left(f_n-f_{n'}\right)g_1^2$$

$$+P^2C_P^2(f_n - f_{n'})g_3^2 + 2\overline{Y}\,\overline{X}\rho_{YX}C_YC_X(f_n - f_{n'})g_1$$

$$+2\overline{Y}P\rho_{YP}C_YC_P(f_n - f_{n'})g_3 + 2\overline{X}P\rho_{XP}C_XC_P(f_n - f_{n'})g_1g_3$$

or $MSE\left(\overline{y}_{gd}\right) = f_n\overline{Y}^2C_Y^2 + (f_n - f_{n'})\Big[\overline{X}^2C_X^2g_1^2 + P^2C_P^2g_3^2 + 2\overline{Y}\,\overline{X}\rho_{YX}C_YC_Xg_1$

$$+2\overline{Y}P\rho_{YP}C_YC_Pg_3 + 2\overline{X}P\rho_{XP}C_XC_Pg_1g_3\Big] \quad . \qquad (2.5)$$

For minimizing (2.5) in two unknowns g_1 and g_3, the two normal equations after differentiating (2.5) partially with respect to g_1 and g_3 are

$$\overline{X}C_Xg_1 + P\rho_{XP}C_Pg_3 + \overline{Y}\rho_{YX}C_Y = 0 \qquad (2.6)$$

$$\overline{X}C_X\rho_{XP}g_1 + PC_Pg_3 + \overline{Y}\rho_{YP}C_Y = 0 \quad . \qquad (2.7)$$

Solving (2.6) and (2.7) for g_1 and g_3, we get the minimizing optimum values to be

$$g_1^* = \frac{\overline{Y}C_Y(\rho_{XP}\rho_{YP} - \rho_{YX})}{\overline{X}C_X(1-\rho_{XP}^2)} \quad \text{and} \qquad (2.8)$$

$$g_3^* = \frac{\overline{Y}C_Y(\rho_{YX}\rho_{XP} - \rho_{YP})}{PC_P(1-\rho_{XP}^2)} \qquad (2.9)$$

which when substituted in (2.5) gives the minimum value of mean square error of the estimator $\overline{y}_{gd}$ as

$$MSE\left(\overline{y}_{gd}\right)_{\min} = f_n\overline{Y}^2C_Y^2 - (f_n - f_{n'})\overline{Y}^2R_{Y.XP}^2C_Y^2$$

$$= f_n\overline{Y}^2C_Y^2 - \left(\frac{1}{n} - \frac{1}{n'}\right)\overline{Y}^2R_{Y.XP}^2C_Y^2 = M \text{ (say)} \qquad (2.10)$$

3. Some Particular Estimators Belonging to the Proposed Class of Estimator

Volume : 6 | Issue : 1 | JANUARY 2016 | ISSN - 2249-555X

Some particular members (estimators) belonging to the proposed class $\bar{y}_{gd}$ of estimators are

(i) $$\bar{y}_{gd(1)} = \bar{y} + k_1\left(\bar{x} - \bar{x}'\right) + k_2\left(p - p'\right) \quad (3.1)$$

(ii) $$\bar{y}_{gd(2)} = \bar{y}\left(\frac{\bar{x}}{\bar{x}'}\right)\left\{1 + k_1\left(\frac{\bar{x}}{\bar{x}'} - 1\right) + k_2\left(\frac{p}{p'} - 1\right)\right\} \quad (3.2)$$

(iii) $$\bar{y}_{gd(3)} = \bar{y}\left(\frac{\bar{x}}{\bar{x}'}\right)^{k_1}\left(\frac{p}{p'}\right)^{k_2} \quad (3.3)$$

(iv) $$\bar{y}_{gd(4)} = \bar{y}\ \exp\left\{k_1\left(\bar{x} - \bar{x}'\right) + k_2\left(p - p'\right)\right\} \quad (3.4)$$

which may be easily shown to satisfy the conditions of $\bar{y}_{gd}$ and thus belongs to the class $\bar{y}_{gd}$.

Let us consider the estimator $\bar{y}_{gd(1)}$ given by

$$\bar{y}_{gd(1)} = \bar{y} + k_1\left(\bar{x} - \bar{x}'\right) + k_2\left(p - p'\right)$$

In order to obtain the bias and mean square error of $\bar{y}_{gd(1)}$, we have

$$\bar{y}_{gd(1)} = \bar{Y}(1 + e_0) + k_1\left\{\bar{X}\left(1 + e_1\right) - \bar{X}\left(1 + e_1'\right)\right\} + k_2\left\{P\left(1 + e_2\right) - P\left(1 + e_2'\right)\right\}$$

or $$\bar{y}_{gd(1)} - \bar{Y} = \bar{Y}\,e_0 + k_1\,\bar{X}\left(e_1 - e_1'\right) + k_2\,P\left(e_2 - e_2'\right) \quad (3.5)$$

Now, taking expectation on both the sides of (3.5), the bias in $\bar{y}_{gd(1)}$ is given by

$$\begin{aligned}
\text{Bias in } \bar{y}_{gd(1)} &= E\left(\bar{y}_{gd(1)}\right) - \bar{Y} \\
&= E\left(\bar{Y}\,e_0 + k_1\,\bar{X}\left(e_1 - e_1'\right) + k_2\,P\left(e_2 - e_2'\right)\right) \\
&= \bar{Y}\,E\left(e_0\right) + k_1\,\bar{X}\left\{E\left(e_1\right) - E\left(e_1'\right)\right\} + k_2\,P\left\{E\left(e_2\right) - E\left(e_2'\right)\right\}
\end{aligned}$$

Volume : 6 | Issue : 1 | JANUARY 2016 | ISSN - 2249-555X

using values of the expectations given in (2.1), we have

$$\text{Bias in } \bar{y}_{gd(1)} = 0 \tag{3.6}$$

Now, squaring (3.5) on both the sides and then taking expectation, the mean square error in $\bar{y}_{gd(1)}$ is given by

$$MSE\left(\bar{y}_{gd(1)}\right) = E\left(\bar{y}_{gd(1)} - \bar{Y}\right)^2$$

$$= E\left\{\bar{Y} e_0 + k_1 \bar{X}\left(e_1 - e_1'\right) + k_2 P\left(e_2 - e_2'\right)\right\}^2$$

$$= E\left\{\bar{Y} e_0 + \bar{X} e_1 k_1 - \bar{X} e_1' k_1 + P e_2 k_2 - P e_2' k_2\right\}^2$$

$$= \bar{Y}^2 E\left(e_0^2\right) + \bar{X}^2 k_1^2 E\left(e_1^2\right) + \bar{X}^2 k_1^2 E\left(e_1'^2\right) - 2\bar{X}^2 k_1^2 E\left(e_1 e_1'\right)$$

$$+ P^2 k_2^2 E\left(e_2^2\right) + P^2 k_2^2 E\left(e_2'^2\right) - 2P^2 k_2^2 E\left(e_2 e_2'\right)$$

$$+ 2\bar{Y}\bar{X} k_1 E\left(e_0 e_1\right) - 2\bar{Y}\bar{X} k_1 E\left(e_0 e_1'\right) + 2\bar{Y} P k_2 E\left(e_0 e_2\right)$$

$$- 2\bar{Y} P k_2 E\left(e_0 e_2'\right) + 2\bar{X} P k_1 k_2 E\left(e_1 e_2\right) + 2\bar{X} P k_1 k_2 E\left(e_1' e_2'\right)$$

$$- 2\bar{X} P k_1 k_2 E\left(e_1 e_2'\right) - 2\bar{X} P k_1 k_2 E\left(e_1' e_2\right)$$

using values of the expectations given in (2.2), the mean square error in $\bar{y}_{gd(1)}$ is given by

$$MSE\left(\bar{y}_{gd(1)}\right) = \bar{Y}^2 f_n C_Y^2 + \bar{X}^2 f_n C_X^2 k_1^2 + \bar{X}^2 f_{n'} C_X^2 k_1^2$$

$$- 2\bar{X}^2 f_{n'} C_X^2 k_1^2 + P^2 f_n C_P^2 k_2^2 + P^2 f_{n'} C_P^2 k_2^2$$

$$- 2P^2 f_{n'} C_P^2 k_2^2 + 2\bar{Y}\bar{X} f_n \rho_{YX} C_Y C_X k_1 - 2\bar{Y}\bar{X} f_{n'} \rho_{YX} C_Y C_X k_1$$

$$+ 2\bar{Y} P f_n \rho_{YP} C_Y C_P k_2 - 2\bar{Y} P f_{n'} \rho_{YP} C_Y C_P k_2 + 2\bar{X} P f_n \rho_{XP} C_X C_P k_1 k_2$$

$$+ 2\bar{X} P f_{n'} \rho_{XP} C_X C_P k_1 k_2 - 2\bar{X} P f_{n'} \rho_{XP} C_X C_P k_1 k_2 - 2\bar{X} P f_{n'} \rho_{XP} C_X C_P k_1 k_2$$

Volume : 6 | Issue : 1 | JANUARY 2016 | ISSN - 2249-555X

$$= f_n \bar{Y}^2 C_Y^2 + \bar{X}^2 C_X^2 (f_n - f_{n'}) k_1^2$$

$$+ P^2 C_P^2 (f_n - f_{n'}) k_2^2 + 2\bar{Y}\bar{X}\rho_{YX} C_Y C_X (f_n - f_{n'}) k_1$$

$$+ 2\bar{Y} P \rho_{YP} C_Y C_P (f_n - f_{n'}) k_2 + 2\bar{X} P \rho_{XP} C_X C_P (f_n - f_{n'}) k_1 k_2$$

or $MSE\left(\bar{y}_{gd(1)}\right) = f_n \bar{Y}^2 C_Y^2 + (f_n - f_{n'}) \left[\bar{X}^2 C_X^2 k_1^2 + P^2 C_P^2 k_2^2 + 2\bar{Y}\bar{X}\rho_{YX} C_Y C_X k_1\right.$

$$\left. + 2\bar{Y} P \rho_{YP} C_Y C_P k_2 + 2\bar{X} P \rho_{XP} C_X C_P k_1 k_2 \right] \quad . \qquad (3.7)$$

The minimum value of mean square error is obtained if the optimum values of k_1 and k_2 are

$$k_1^* = \frac{\bar{Y} C_Y (\rho_{XP}\rho_{YP} - \rho_{YX})}{\bar{X} C_X (1 - \rho_{XP}^2)} \qquad (3.8)$$

$$k_2^* = \frac{\bar{Y} C_Y (\rho_{YX}\rho_{XP} - \rho_{YP})}{P C_P (1 - \rho_{XP}^2)} \qquad (3.9)$$

and the minimum mean square error of $\bar{y}_{gd(1)}$ under the optimizing values of the characterizing scalars is given by

$$MSE\left(\bar{y}_{gd(1)}\right)_{\min} = f_n \bar{Y}^2 C_Y{}^2 - \left(\frac{1}{n} - \frac{1}{n'}\right) \bar{Y}^2 R^2{}_{Y.XP} C_Y{}^2 \qquad (3.10)$$

which is the same as the minimum mean square error of the proposed generalized class of estimator $\bar{y}_{gd}$ and this can also be verified for the estimators $\bar{y}_{gd(2)}$, $\bar{y}_{gd(3)}$ and $\bar{y}_{gd(4)}$ on the similar lines.

4. Efficiency Comparison with the Available Estimators

For comparing the efficiency of the proposed generalized double sampling estimator, let us consider the following

(i) Double Sampling Ratio Estimator

$$\hat{\bar{y}}_1 = \bar{y}.\frac{\bar{x}'}{x}$$

with $MSE\left(\hat{\bar{y}}_1\right) = f_n \bar{Y}^2 \left(C_Y^2 + C_X^2 - 2\rho_{YX} C_Y C_X\right) - f_{n'} \bar{Y}^2 \left(C_X{}^2 - 2\rho_{YX} C_Y C_X\right)$ (4.1)

from (4.1) and (2.10), we have

$$MSE\left(\hat{\bar{y}}_1\right) - M = f_n \bar{Y}^2 \left(C_Y^2 + C_X^2 - 2\rho_{YX} C_Y C_X\right) - f_{n'} \bar{Y}^2 \left(C_X{}^2 - 2\rho_{YX} C_Y C_X\right)$$

$$- f_n \bar{Y}^2 C_Y{}^2 + \left(f_n - f_{n'}\right) \bar{Y}^2 R^2{}_{Y.XP}\, C_Y{}^2$$

$$= f_n \bar{Y}^2 \left[C_X^2 - 2\rho_{YX} C_Y C_X + R_{Y.XP}^2\, C_Y^2\right]$$

$$- f_{n'} \bar{Y}^2 \left[C_X^2 - 2\rho_{YX} C_Y C_X + R_{Y.XP}^2\, C_Y^2\right]$$

$$= f_n \bar{Y}^2 \left[C_X^2 - 2\rho_{YX} C_Y C_X + \rho_{YX}^2 C_Y^2 - \rho_{YX}^2 C_Y^2 + R_{Y.XP}^2\, C_Y^2\right]$$

$$- f_{n'} \bar{Y}^2 \left[C_X^2 - 2\rho_{YX} C_Y C_X + \rho_{YX}^2 C_Y^2 - \rho_{YX}^2 C_Y^2 + R_{Y.XP}^2\, C_Y^2\right]$$

$$= \bar{Y}^2 \left[\left(C_X - \rho_{YX} C_Y\right)^2 + \left(R_{Y.XP}^2 - \rho_{YX}^2\right) C_Y^2\right]\left(f_n - f_{n'}\right)$$

$$\geq 0 \text{ since } R_{Y.XP}^2 \geq \rho_{YX}^2 \quad . \qquad (4.2)$$

Showing that the proposed generalized double sampling estimator is more efficient than the ratio estimator.

(ii) Double Sampling Ratio Estimator by Naik and Gupta (1996)

$$\hat{\bar{y}}_2 = \bar{y}.\frac{p'}{p}$$

with $MSE\left(\hat{\bar{y}}_2\right) = f_n \bar{Y}^2 \left(C_Y^2 + C_P^2 - 2\rho_{YP} C_Y C_P\right) - f_{n'} \bar{Y}^2 \left(C_P^2 - 2\rho_{YP} C_Y C_P\right)$ (4.3) from (4.3) and (2.10), we have

$$MSE\left(\hat{\bar{y}}_2\right)-M = f_n \bar{Y}^2\left(C_Y^2 + C_P^2 - 2\rho_{YP}C_Y C_P\right) - f_{n'}\bar{Y}^2\left(C_P{}^2 - 2\rho_{YP}C_Y C_P\right)$$

$$- f_n \bar{Y}^2 C_Y{}^2 + \left(f_n - f_{n'}\right)\bar{Y}^2 R^2{}_{Y.XP}\, C_Y{}^2$$

$$= f_n \bar{Y}^2\left[C_P^2 - 2\rho_{YP}C_Y C_P + R_{Y.XP}^2\, C_Y^2\right]$$

$$- f_{n'} \bar{Y}^2\left[C_P^2 - 2\rho_{YP}C_Y C_P + R_{Y.XP}^2\, C_Y^2\right]$$

$$= f_n \bar{Y}^2\left[C_P^2 - 2\rho_{YP}C_Y C_P + \rho_{YP}^2 C_Y^2 - \rho_{YP}^2 C_Y^2 + R_{Y.XP}^2\, C_Y^2\right]$$

$$- f_{n'} \bar{Y}^2\left[C_P^2 - 2\rho_{YP}C_Y C_P + \rho_{YP}^2 C_Y^2 - \rho_{YP}^2 C_Y^2 + R_{Y.XP}^2\, C_Y^2\right]$$

$$= \bar{Y}^2\left[\left(C_P - \rho_{YP}C_Y\right)^2 + \left(R_{Y.XP}^2 - \rho_{YP}^2\right) C_Y^2\right]\left(f_n - f_{n'}\right)$$

$$\geq 0 \text{ since } R_{Y.XP}^2 \geq \rho_{YP}^2 \quad . \qquad (4.4)$$

Showing that the proposed generalized double sampling estimator is more efficient than the estimator given by Naik and Gupta (1996).

(iii) Double Sampling Exponential Ratio Estimator by Bahl and Tuteja(1991)

$$\hat{\bar{y}}_3 = \bar{y}\exp\left(\frac{\bar{\bar{X}} - \bar{x}}{\bar{\bar{X}} + \bar{x}}\right)$$

with $MSE\left(\hat{\bar{y}}_3\right) = f_n \bar{Y}^2\left(C_Y^2 + \frac{1}{4}C_X^2 - \rho_{YX}C_Y C_X\right) - f_{n'} \bar{Y}^2\left(\frac{1}{4}C_X^2 - \rho_{YX}C_Y C_X\right)$ (4.5)

from (4.5) and (2.10), we have

$$MSE\left(\hat{\bar{y}}_3\right)-M = f_n \bar{Y}^2\left(C_Y^2 + \frac{1}{4}C_X^2 - \rho_{YX}C_Y C_X\right) - f_{n'} \bar{Y}^2\left(\frac{1}{4}C_X^2 - \rho_{YX}C_Y C_X\right)$$

$$- f_n \bar{Y}^2 C_Y{}^2 + \left(f_n - f_{n'}\right)\bar{Y}^2 R^2{}_{Y.XP}\, C_Y{}^2$$

$$= f_n \bar{Y}^2\left[\frac{C_X^2}{4} - \rho_{YX}C_Y C_X + \rho_{YX}^2 C_Y^2 - \rho_{YX}^2 C_Y^2 + R_{Y.XP}^2\, C_Y^2\right]$$

$$- f_{n'} \bar{Y}^2 \left[\frac{C_X^2}{4} - \rho_{YX} C_Y C_X + \rho_{YX}^2 C_Y^2 - \rho_{YX}^2 C_Y^2 + R_{Y.XP}^2 C_Y^2 \right]$$

$$= \bar{Y}^2 \left[\left(\frac{C_X}{2} - \rho_{YX} C_Y \right)^2 + \left(R_{Y.XP}^2 - \rho_{YX}^2 \right) C_Y^2 \right] \left(f_n - f_{n'} \right)$$

$$\geq 0 \text{ since } R_{Y.XP}^2 \geq \rho_{YX}^2 \quad . \qquad (4.6)$$

Showing that the proposed generalized double sampling estimator is more efficient than the estimator given by Bahl and Tuteja (1991).

5. Empirical Study

For comparing efficiency of the proposed generalized class of estimator, let us consider the data given in [William G. Cochran (1977), Sampling Techniques, 3rd Edition, John Wiley and Sons, New York, at Page No. 34] we have

Y = Food Cost

X = Family Income

ϕ = Family size of more than 3

$\bar{Y} = 27.49$	$C_X = 0.146$
$\bar{X} = 72.55$	$C_Y = 0.369$
$P = 0.52$	$C_P = 0.985$
$n' = 22$	$\rho_{YX} = 0.2521$
$n = 16$	$\rho_{YP} = 0.388$
$N = 33$	$\rho_{XP} = -0.153$.

Volume : 6 | Issue : 1 | JANUARY 2016 | ISSN - 2249-555X

Table 5.1: PRE of the Proposed Estimator over the Estimators Described Above

PRE of the Proposed Estimator over the Estimators	PRE
PRE of the Proposed Estimator $\bar{y}_{gd}$ over the Estimator $\hat{\bar{y}}_1$	112.63
PRE of the Proposed Estimator $\bar{y}_{gd}$ over the Estimator $\hat{\bar{y}}_2$	423.61
PRE of the Proposed Estimator $\bar{y}_{gd}$ over the Estimator $\hat{\bar{y}}_3$	111.55

6. Conclusion

The comparative study of the proposed generalized double sampling estimator establishes its superiority in the sense of having minimum mean square error over mean per unit estimator, ratio estimator, ratio estimator by Naik and Gupta (1996) and exponential ratio estimator by Bahl and Tuteja (1991).

REFERENCE 1. Cochran, W.G. (1977) : Sampling Techniques, 3rd edition, John Wiley and Sons, New York. 2. Bahl, S. and Tuteja, T.K. (1991) : Ratio and Product type exponential estimator, Information and Optimization Sciences, XII(I), 159-163. 3. Naik, V.D. and Gupta, P.C. (1996) : A note on estimation of mean with known population proportion of an auxiliary character. Journal of Indian Society of Agricultural Statistics, 48(2), 151-158. 4. Srivastava, S.K. (1967) : An estimator using auxiliary information, Calcutta Statistical Association Bulletin, 16, 121-132. 5. Sukhatme, P.V. and Sukhatme, B.V. (1970) : Sampling theory of surveys with applications, 3rd revised edition, IOWA State University Press, Ames, U.S.A.