Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 7, Issue 1 (2014): 79-91

RATIO-PRODUCT-DIFFERENCE (RPD) TYPE DOUBLE SAMPLING ESTIMATORS FOR FINITE POPULATION VARIANCE USING AUXILIARY INFORMATION

Peeyush Misra and R. Karan Singh
Department of Statistics, Lucknow University, Lucknow - 226007, India

Received August 07, 2013
Modified May 30, 2014
Accepted June 03, 2014

Abstract

For estimating finite population variance using information on single auxiliary variable in the form of mean, the Ratio-Product-Difference (RPD) type double sampling estimators d_1 and d_2 and their generalized estimators as d_{1g} and d_{2g} are proposed. The bias and mean square error (MSE) of the proposed estimators are found. Theoretical comparisons with the traditional estimator are carried out. By this comparison it is shown that the proposed estimators are more efficient than the traditional one.

Key Words: Auxiliary Variable, Taylor's Series Expansion, Bias, Mean Square Error (MSE), Efficiency.

1. Introduction

In sampling theory, auxiliary information is widely used at the stages of selection and estimation, at the selection stage the auxiliary information is used by designing various sampling schemes and at the estimation stage it is used in formulating various types of estimators of different population parameters with a view of getting increased efficiency. Estimators like ratio, product, difference, regression and the classes of ratio and product type estimators for population parameters mainly population mean and variance are studied by many authors and are available in the literature. But when parameters of one or more auxiliary variables are not available in advance then the alternative is to use double sampling or two phase sampling technique where we first take a preliminary large sample of size n' (called first phase sample) on which only the auxiliary variable is observed and then from n' taking a sub-sample of size n (called second phase sample) on which both the variables are observed. In such situations the different estimators known as double sampling ratio, product, difference and regression estimators were developed. This present paper too contributes to this area.

Let a first phase simple random sample of size n' without replacement be drawn from a population of size N and a second phase simple random sample of size n without replacement be drawn from the first phase sample of size n'. At first phase sample of size n', only the auxiliary character X is observed and at the second phase sub-sample of size n, both the study variably Y and the auxiliary character X are observed.

Let $(\bar{Y},\bar{X})$ be the population means of (Y, X) respectively. ρ be the population correlation coefficient between (Y, X) and $\bar{x}'$ be the sample mean of the first phase n' sample values on the auxiliary character X. Let

$S_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i-\bar{Y})^2$, $S_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})^2$ and $\mu_{rs} = \frac{1}{N}\sum_{i=1}^{N}(Y_i-\bar{Y})^r(X_i-\bar{X})^s$

where (y_i, x_i) are the values on (Y, X) respectively for the i^{th} (i = 1, 2, . . . , N) unit of the population. Also let $(\bar{y},\bar{x})$ be the sample means of (y,x) based on second phase sample of size n.

2. The Suggested Estimators

For estimating the population variance $\sigma_Y^{\ 2}$ of the study (main) variable Y, the proposed double sampling estimators are defined as

$$d_1 = \hat{\theta} - \bar{y}^2\left\{1+\frac{k_1(\bar{x}-\bar{x}')}{\bar{x}'}\right\} \tag{2.1}$$

Let us generalize it as

$$d_{1g} = \hat{\theta} - \bar{y}^2.f(w) \tag{2.2}$$

where $w = \frac{\bar{x}}{\bar{x}'}$ and $f(w)$ is a bounded function of w such that $f(1)$ = 1 at the point w =1 satisfying the regularity conditions for the validity of Taylor's series expansion and having first two derivatives with respect to w to be bounded. And

$$d_2 = \hat{\theta}\left\{1+\frac{k_2(\bar{x}-\bar{x}')}{\bar{x}'}\right\} - \bar{y}^2, \text{ where } \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} y_i^{\ 2} \tag{2.3}$$

Let us generalize it as

$$d_{2g} = \hat{\theta}\ f(w) - \bar{y}^2 \tag{2.4}$$

where $w = \frac{\bar{x}}{\bar{x}'}$ and $f(w)$ is a bounded function of w such that $f(1)$ = 1 at the point w =1 satisfying the regularity conditions for the validity of Taylor's series expansion and having first two derivatives with respect to w to be bounded.

3. Some Theorems

Theorem I : The bias in d_1 up to terms of order $\left(\frac{1}{n}\right)$ is given by

$$\text{Bias}(d_1) = \frac{2k_1\bar{Y}}{\bar{X}}\frac{\mu_{11}}{n'} - \frac{2k_1\bar{Y}}{\bar{X}}\frac{\mu_{11}}{n} - \frac{\mu_{20}}{n} \tag{3.1}$$

Where as the mean square error of d_1 to the first degree of approximation for the optimum value of k_1 is given by

$$\text{MSE}(d_1)\min = \frac{1}{n}\left(\mu_{40}-\mu_{20}^{\ 2}\right) - \left(\frac{1}{n}-\frac{1}{n'}\right)\frac{\mu_{21}^{\ 2}}{\mu_{02}} \tag{3.2}$$

Proof : Let $\bar{y} = \bar{Y} + e_0$, $\bar{x} = \bar{X} + e_1$, $\bar{x}' = \bar{X} + e_1'$ and $\frac{1}{n}\sum_{i=1}^{n} y_i^2 = \frac{1}{N}\sum_{i=1}^{N} Y_i^2 + e_2$ or $\frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{N}\sum_{i=1}^{N} Z_i + e_2$, where $z_i = y_i^2$ & $Z_i = Y_i^2$ or $\bar{z} = \bar{Z} + e_2$ or $\hat{\theta} = \theta + e_2$.

For simplicity, we assume that the population size N is large enough as compared to the sample size so that finite population correction terms may be ignored.

Now E (e_0) = E (e_1) = E (e_1') = E (e_2) = 0 (3.3)

$$\left.\begin{aligned} E(e_0^2) &= \frac{1}{n} S_Y^2 = \frac{\mu_{20}}{n} & E(e_0 e_1) &= \frac{1}{n}\rho\, S_Y S_X = \frac{\mu_{11}}{n} \\ E(e_1^2) &= \frac{1}{n} S_X^2 = \frac{\mu_{02}}{n} & E(e_0 e_1') &= \frac{\mu_{11}}{n'} \\ E(e_1'^2) &= \frac{\mu_{02}}{n'}, & E(e_1 e_1') &= \frac{\mu_{02}}{n'} \end{aligned}\right\} \quad (3.4)$$

and

$$\left.\begin{aligned} E(e_2^2) &= \frac{1}{n}(\mu_{40} + 4\bar{Y}\mu_{30} + 4\bar{Y}^2\mu_{20} - \mu_{20}^2) \\ E(e_0 e_2) &= \frac{1}{n}(\mu_{30} + 2\bar{Y}\mu_{20}) \\ E(e_1 e_2) &= \frac{1}{n}(\mu_{21} + 2\bar{Y}\mu_{11}) \\ E(e_1' e_2) &= \frac{1}{n'}(\mu_{21} + 2\bar{Y}\mu_{11}) \end{aligned}\right\} \quad (3.5)$$

We have

$$d_1 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2\left\{1 + \frac{k_1(\bar{x} - \bar{x}')}{\bar{x}'}\right\}$$

$$= \hat{\theta} - \bar{y}^2\left\{1 + \frac{k_1(\bar{x} - \bar{x}')}{\bar{x}'}\right\}$$

$$= (\theta + e_2) - (\bar{Y} + e_0)^2\left\{1 + \frac{k_1\left(\bar{X} + e_1 - \bar{X} - e_1'\right)}{\bar{X} + e_1'}\right\}$$

$$= (\theta + e_2) - (\bar{Y} + e_0)^2\left\{1 + \frac{k_1}{\bar{X}}\left(e_1 - e_1'\right)\left(1 + \frac{e_1'}{\bar{X}}\right)^{-1}\right\}$$

$$= (\theta + e_2) - (\bar{Y} + e_0)^2\left\{1 + \frac{k_1}{\bar{X}}\left(e_1 - e_1'\right)\left(1 - \frac{e_1'}{\bar{X}} + \frac{e_1'^2}{\bar{X}^2}\right)\right\}$$

$$= (\theta + e_2) - (\bar{Y}^2 + e_0^2 + 2\bar{Y}e_0)\left(1 + \frac{k_1 e_1}{\bar{X}} - \frac{k_1 e_1'}{\bar{X}} + \frac{k_1 e_1'^2}{\bar{X}^2} - \frac{k_1 e_1 e_1'}{\bar{X}^2}\right)$$

$$= (\theta + e_2) - \bar{Y}^2 - \frac{k_1\bar{Y}^2 e_1}{\bar{X}} + \frac{k_1\bar{Y}^2 e_1'}{\bar{X}} - \frac{k_1\bar{Y}^2 e_1'^2}{\bar{X}^2} + \frac{k_1\bar{Y}^2 e_1 e_1'}{\bar{X}^2}$$

$$- e_0{}^2 - 2\bar{Y}e_0 - \frac{2k_1\bar{Y}e_0 e_1}{\bar{X}} + \frac{2k_1\bar{Y}e_0 e_1'}{\bar{X}}$$

$$= \left(\frac{1}{N}\sum_{i=1}^{N} Y_i{}^2 - \bar{Y}^2\right) + e_2 - 2\bar{Y}e_0 - \frac{k_1\bar{Y}^2 e_1}{\bar{X}} + \frac{k_1\bar{Y}^2 e_1'}{\bar{X}} - e_0{}^2 - \frac{k_1\bar{Y}^2 e_1'^2}{\bar{X}^2} + \frac{k_1\bar{Y}^2 e_1 e_1'}{\bar{X}^2}$$

$$- \frac{2k_1\bar{Y}e_0 e_1}{\bar{X}} + \frac{2k_1\bar{Y}e_0 e_1'}{\bar{X}}$$

$$(d_1 - \sigma_Y^2) = e_2 - 2\bar{Y}e_0 - \frac{k_1\bar{Y}^2 e_1}{\bar{X}} + \frac{k_1\bar{Y}^2 e_1'}{\bar{X}} - e_0{}^2 - \frac{k_1\bar{Y}^2 e_1'^2}{\bar{X}^2} + \frac{k_1\bar{Y}^2 e_1 e_1'}{\bar{X}^2} - \frac{2k_1\bar{Y}e_0 e_1}{\bar{X}} + \frac{2k_1\bar{Y}e_0 e_1'}{\bar{X}} \quad (3.6)$$

Taking expectation on both sides of (3.6) and using values of the expectations given from (3.3) to (3.5), the bias in $d_1 (= E(d_1) - \sigma_Y^2)$ to the first degree of approximation is given by

Bias (d_1) $= E(d_1) - \sigma_Y^2$

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

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

$$= -\frac{\mu_{20}}{n} - \frac{k_1\bar{Y}^2}{\bar{X}^2}\frac{\mu_{02}}{n'} + \frac{k_1\bar{Y}^2}{\bar{X}^2}\frac{\mu_{02}}{n'} - \frac{2k_1\bar{Y}}{\bar{X}}\frac{\mu_{11}}{n} + \frac{2\bar{Y}k_1}{\bar{X}}\frac{\mu_{11}}{n'}$$

$$= \frac{2k_1\bar{Y}}{\bar{X}}\frac{\mu_{11}}{n'} - \frac{2k_1\bar{Y}}{\bar{X}}\frac{\mu_{11}}{n} - \frac{\mu_{20}}{n} \quad (3.7)$$

Now squaring (3.6) on both sides and then taking expectation, the mean square error of $d_1 = E\left(d_1 - \sigma_Y{}^2\right)^2$ to the first degree of approximation is given by

$$\text{MSE}(d_1) = E\left(e_2 - 2\bar{Y}e_0 - \frac{k_1\bar{Y}^2 e_1}{\bar{X}} + \frac{k_1\bar{Y}^2 e_1'}{\bar{X}} - e_0{}^2 - \frac{k_1\bar{Y}^2 e_1'^2}{\bar{X}^2} + \frac{k_1\bar{Y}^2 e_1 e_1'}{\bar{X}^2} - \frac{2k_1\bar{Y}e_0 e_1}{\bar{X}} + \frac{2k_1\bar{Y}e_0 e_1'}{\bar{X}}\right)^2$$

$$= E(e_2{}^2) + 4\bar{Y}^2 E(e_0{}^2) + \frac{k_1{}^2\bar{Y}^4}{\bar{X}^2} E(e_1{}^2) + \frac{k_1{}^2\bar{Y}^4}{\bar{X}^2} E(e_1'^2) - 4\bar{Y}E(e_0 e_2) - \frac{2k_1\bar{Y}^2}{\bar{X}} E(e_1 e_2)$$

$$+ \frac{2k_1\bar{Y}^2}{\bar{X}} E(e_1' e_2) + \frac{4k_1\bar{Y}^3}{\bar{X}} E(e_0 e_1) - \frac{4k_1\bar{Y}^3}{\bar{X}} E(e_0 e_1') - \frac{2k_1{}^2\bar{Y}^4}{\bar{X}^2} E(e_1 e_1')$$

$$= \frac{1}{n}(\mu_{40} + 4\bar{Y}\mu_{30} + 4\bar{Y}^2\mu_{20} - \mu_{20}{}^2) + 4\bar{Y}^2\frac{\mu_{20}}{n} - 4\bar{Y}\frac{1}{n}(\mu_{30} + 2\bar{Y}\mu_{20})$$

$$+\frac{k_1^2\bar{Y}^4}{\bar{X}^2}\frac{\mu_{02}}{n}+\frac{k_1^2\bar{Y}^4}{\bar{X}^2}\frac{\mu_{02}}{n'}-\frac{2k_1\bar{Y}^2}{\bar{X}}\frac{1}{n}(\mu_{21}+2\bar{Y}\mu_{11})+\frac{2k_1\bar{Y}^2}{\bar{X}}\frac{1}{n'}(\mu_{21}+2\bar{Y}\mu_{11})$$

$$+\frac{4k_1\bar{Y}^3}{\bar{X}}\frac{\mu_{11}}{n}-\frac{4k_1\bar{Y}^3}{\bar{X}}\frac{\mu_{11}}{n'}-\frac{2k_1^2\bar{Y}^4}{\bar{X}^2}\frac{\mu_{02}}{n'}$$

or MSE $(d_1) = \frac{1}{n}(\mu_{40}-\mu_{20}^2)+\frac{k_1^2\bar{Y}^4}{\bar{X}^2}\frac{\mu_{02}}{n}-\frac{k_1^2\bar{Y}^4}{\bar{X}^2}\frac{\mu_{02}}{n'}-\frac{2k_1\bar{Y}^2}{\bar{X}}\frac{\mu_{21}}{n}+\frac{2k_1\bar{Y}^2}{\bar{X}}\frac{\mu_{21}}{n'}$

$$=\frac{1}{n}(\mu_{40}-\mu_{20}^2)+\left(\frac{1}{n}-\frac{1}{n'}\right)\left(\frac{k_1^2\bar{Y}^4}{\bar{X}^2}\mu_{02}-\frac{2k_1\bar{Y}^2}{\bar{X}}\mu_{21}\right) \quad (3.8)$$

The optimum value of k_1 minimizing mean square error of d_1 is given by

$$k_1^* = \frac{\bar{X}}{\bar{Y}^2}\frac{\mu_{21}}{\mu_{02}} \quad (3.9)$$

which when substituted in (3.8) gives the minimum value of mean square error as

$$\text{MSE}(d_1)\min = \frac{1}{n}\left(\mu_{40}-\mu_{20}^2\right)-\left(\frac{1}{n}-\frac{1}{n'}\right)\frac{\mu_{21}^2}{\mu_{02}} \quad (3.10)$$

showing that mean square error of the proposed estimator d_1 is less than that of usual conventional unbiased estimator $s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$ of population variance σ_Y^2.

Theorem II: The bias in d_{1g} up to terms of order $\left(\frac{1}{n}\right)$ is given by

$$\text{Bias}(d_{1g}) = \frac{\bar{Y}^2}{2\bar{X}^2}f''(1)\frac{\mu_{02}}{n'}-\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)\frac{\mu_{02}}{n}+\frac{2\bar{Y}}{\bar{X}}f'(1)\frac{\mu_{11}}{n'}-\frac{2\bar{Y}}{\bar{X}}f'(1)\frac{\mu_{11}}{n}+\frac{\bar{Y}^2}{\bar{X}^2}f''(1)\frac{\mu_{02}}{n'}-\frac{\mu_{20}}{n} \quad (3.11)$$

Where as the mean square error of d_{1g} to the first degree of approximation for the optimum value of $f'(1)$ is given by

$$\text{MSE}(d_{1g})\min = \frac{1}{n}\left(\mu_{40}-\mu_{20}^2\right)-\left(\frac{1}{n}-\frac{1}{n'}\right)\frac{\mu_{21}^2}{\mu_{02}} \quad (3.12)$$

Proof : Let us consider (2.2)

$$d_{1g} = \hat{\theta}-\bar{y}^2.f(w)$$

For $f'(1)$, $f''(1)$ and $f'''(1)$ to be first, second and third order derivatives of $f(w)$ at the point w =1 respectively and $w^* = 1+h(w-1)$, $0 < h < 1$, expanding $f(w)$ in d_{1g} in third order Taylor's series, we have

$$d_{1g}=\hat{\theta}-\bar{y}^2\left\{f(1)+(w-1)f'(1)+\frac{(w-1)^2}{2!}f''(1)+\frac{(w-1)^3}{3!}f'''(w)\right\}$$

$$=\hat{\theta}-\bar{y}^2\left\{1+\left(\frac{e_1-e_1'}{\bar{X}+e_1'}\right)f'(1)+\frac{1}{2!}\left(\frac{e_1-e_1'}{\bar{X}+e_1'}\right)^2 f''(1)+\frac{1}{3!}\left(\frac{e_1-e_1'}{\bar{X}+e_1'}\right)^3 f'''(w^*)\right\}$$

$$=\hat{\theta}-\bar{y}^2\left\{1+\frac{1}{\bar{X}}(e_1-e_1')\left(1+\frac{e_1'}{\bar{X}}\right)^{-1}f'(1)+\frac{1}{2!}\frac{1}{\bar{X}^2}(e_1-e_1')^2\left(1+\frac{e_1'}{\bar{X}}\right)^{-2}f''(1)\right\}$$

$$=\hat{\theta}-\bar{y}^2\left\{1+\frac{1}{\bar{X}}(e_1-e_1')\left(1-\frac{e_1'}{\bar{X}}+\frac{e_1'^2}{\bar{X}^2}\right)f'(1)+\frac{1}{2!}\frac{1}{\bar{X}^2}(e_1-e_1')^2\left(1-\frac{2e_1'}{\bar{X}}+\frac{3e_1'^2}{\bar{X}^2}\right)f''(1)\right\}$$

$$=(\theta+e_2)-(\bar{Y}+e_0)^2\left\{1+\left(\frac{e_1}{\bar{X}}-\frac{e_1e_1'}{\bar{X}^2}-\frac{e_1'}{\bar{X}}+\frac{e_1'^2}{\bar{X}^2}\right)f'(1)+\frac{1}{2}\left(\frac{e_1^2}{\bar{X}^2}-\frac{e_1'^2}{\bar{X}^2}-\frac{2e_1e_1'}{\bar{X}^2}\right)f''(1)\right\}$$

$$=(\theta+e_2)-\left(\bar{Y}^2+e_0^2+2\bar{Y}e_0\right)\cdot\left\{1+\left(\frac{e_1}{\bar{X}}-\frac{e_1'}{\bar{X}}+\frac{e_1'^2}{\bar{X}^2}-\frac{e_1e_1'}{\bar{X}^2}\right)f'(1)\right.$$
$$\left.+\left(\frac{e_1^2}{2\bar{X}^2}-\frac{e_1'^2}{2\bar{X}^2}-\frac{e_1e_1'}{\bar{X}^2}\right)f''(1)\right\}$$

$$=(\theta+e_2)-\bar{Y}^2-\bar{Y}^2\left(\frac{e_1}{\bar{X}}-\frac{e_1'}{\bar{X}}+\frac{e_1'^2}{\bar{X}^2}-\frac{e_1e_1'}{\bar{X}^2}\right)f'(1)-\bar{Y}^2\left(\frac{e_1^2}{2\bar{X}^2}-\frac{e_1'^2}{2\bar{X}^2}-\frac{e_1e_1'}{\bar{X}^2}\right)f''(1)$$
$$-e_0^2-2\bar{Y}e_0-2\bar{Y}e_0\left(\frac{e_1}{\bar{X}}-\frac{e_1'}{\bar{X}}\right)f'(1)$$

$$=\left(\frac{1}{N}\sum_{i=1}^{N}Y_i^2-\bar{Y}^2\right)+e_2-2\bar{Y}e_0-\frac{\bar{Y}^2}{\bar{X}}f'(1)e_1$$
$$+\frac{\bar{Y}^2}{\bar{X}}f'(1)e_1'-e_0^2-\frac{\bar{Y}^2}{\bar{X}^2}f'(1)e_1'^2-\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)e_1^2$$
$$+\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)e_1'^2+\frac{\bar{Y}^2}{\bar{X}^2}f'(1)e_1e_1'+\frac{\bar{Y}^2}{\bar{X}^2}f''(1)e_1e_1'$$
$$-\frac{2\bar{Y}e_0e_1}{\bar{X}}f'(1)+\frac{2\bar{Y}}{\bar{X}}f'(1)e_0e_1'$$

or

$$d_{1g}-\sigma_Y^2=e_2-2\bar{Y}e_0-\frac{\bar{Y}^2}{\bar{X}}f'(1)e_1+\frac{\bar{Y}^2}{\bar{X}^2}f'(1)e_1'-e_0^2-\frac{\bar{Y}^2}{\bar{X}^2}f'(1)e_1'^2-\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)e_1^2$$
$$+\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)e_1'^2+\frac{\bar{Y}^2}{\bar{X}^2}f'(1)e_1e_1'+\frac{\bar{Y}^2}{\bar{X}^2}f''(1)e_1'e_1'-\frac{2\bar{Y}}{\bar{X}}f'(1)e_0e_1+\frac{2\bar{Y}}{\bar{X}}f'(1)e_0e_1' \qquad (3.13)$$

Taking expectation on both sides of (3.13), the bias in d_{1g} $(=E(d_{1g})-\sigma_Y{}^2)$ up to terms of order $\left(\frac{1}{n}\right)$ is given by

Bias (d_{1g}) $=E(d_{1g})-\sigma_Y{}^2$

$$=E(e_2)-2\bar{Y}E(e_0)-\frac{\bar{Y}^2}{\bar{X}}f'(1)E(e_1)+\frac{\bar{Y}^2}{\bar{X}}f'(1)E(e_1')-E(e_0{}^2)-\frac{\bar{Y}^2}{\bar{X}^2}f'(1)E(e_1'^2)$$
$$-\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)E(e_1{}^2)+\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)E(e_1'^2)+\frac{\bar{Y}^2}{\bar{X}^2}f'(1)E(e_1e_1')+\frac{\bar{Y}^2}{\bar{X}^2}f''(1)E(e_1e_1')$$
$$-\frac{2\bar{Y}}{\bar{X}}f'(1)E(e_0e_1)+\frac{2\bar{Y}}{\bar{X}}f'(1)E(e_0e_1')$$
$$=-\frac{\mu_{20}}{n}-\frac{\bar{Y}^2}{\bar{X}^2}f'(1)\frac{\mu_{02}}{n'}-\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)\frac{\mu_{02}}{n}+\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)\frac{\mu_{02}}{n'}+\frac{\bar{Y}^2}{\bar{X}^2}f'(1)\frac{\mu_{02}}{n'}$$
$$+\frac{\bar{Y}^2}{\bar{X}^2}f''(1)\frac{\mu_{02}}{n'}-\frac{2\bar{Y}}{\bar{X}}f'(1)\frac{\mu_{11}}{n}+\frac{2\bar{Y}}{\bar{X}}f'(1)\frac{\mu_{11}}{n'}$$
$$\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)\frac{\mu_{02}}{n'}-\frac{\bar{Y}^2}{2\bar{X}^2}f''(1)\frac{\mu_{02}}{n}+\frac{2\bar{Y}}{\bar{X}}f'(1)\frac{\mu_{11}}{n'}$$
$$-\frac{2\bar{Y}}{\bar{X}}f'(1)\frac{\mu_{11}}{n}+\frac{\bar{Y}^2}{\bar{X}^2}f''(1)\frac{\mu_{02}}{n'}-\frac{\mu_{20}}{n} \qquad (3.14)$$

Now squaring (3.13) on both sides and then taking expectation, the mean square error of d_{1g} to the first degree of approximation is given by

$$E(d_{1g}-\sigma_Y{}^2)^2=E\left(e_2-2\bar{Y}e_0-\frac{\bar{Y}^2}{\bar{X}}f'(1)e_1+\frac{\bar{Y}^2}{\bar{X}}f'(1)e_1'\right)^2$$
$$=E\left(e_2{}^2\right)+4\bar{Y}^2E\left(e_0{}^2\right)+\frac{\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2E\left(e_1{}^2\right)+\frac{\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2E\left(e_1'^2\right)-4\bar{Y}E(e_0e_2)-\frac{2\bar{Y}^2}{\bar{X}}f'(1)E(e_1e_2)$$
$$+\frac{2\bar{Y}^2}{\bar{X}}f'(1)E\left(e_1'e_2\right)+\frac{4\bar{Y}^3}{\bar{X}}f'(1)E(e_0e_1)-\frac{4\bar{Y}^3}{\bar{X}}f'(1)E(e_0e_1')-\frac{2\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2E(e_1e_1')$$
$$=\frac{1}{n}\left(\mu_{40}+4\bar{Y}\mu_{30}+4\bar{Y}^2\mu_{20}-\mu_{20}{}^2\right)+4\bar{Y}^2\frac{\mu_{20}}{n}+\frac{\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n}+\frac{\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n'}$$
$$-4\bar{Y}\frac{1}{n}(\mu_{30}+2\bar{Y}\mu_{20})-\frac{2\bar{Y}^2}{\bar{X}}f'(1)\frac{1}{n}\left(\mu_{21}+2\bar{Y}\mu_{11}\right)+\frac{2\bar{Y}^2}{\bar{X}}f'(1)\frac{1}{n'}\left(\mu_{21}+2\bar{Y}\mu_{11}\right)$$
$$+\frac{4\bar{Y}^3}{\bar{X}}f'(1)\frac{\mu_{11}}{n}-\frac{4\bar{Y}^3}{\bar{X}}f'(1)\frac{\mu_{11}}{n'}-\frac{2\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n'}$$

$$\text{MSE}(d_{1g})=\frac{1}{n}\left(\mu_{40}-\mu_{20}{}^2\right)+\frac{\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n}-\frac{\bar{Y}^4}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n'}-\frac{2\bar{Y}^2}{\bar{X}}f'(1)\frac{\mu_{21}}{n}+\frac{2\bar{Y}^2}{\bar{X}}f'(1)\frac{\mu_{21}}{n'} \qquad (3.15)$$

The optimum value of $f'(1)$ minimizing the mean square error of d_{1g} is given by

$$f'(1)^* = \frac{\overline{X}}{\overline{Y}^2}\frac{\mu_{21}}{\mu_{02}} \tag{3.16}$$

which when substituted in (3.15) gives the minimum value of mean square error as

$$\text{MSE}\,(d_{1g})\text{ min} = \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{21}{}^2}{\mu_{02}} \tag{3.17}$$

showing that mean square error of the proposed estimator d_{1g} is less than that of usual conventional unbiased estimator $s_y{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$ of population variance $\sigma_Y{}^2$.

Theorem III: The bias in d_2 up to terms of order $\left(\frac{1}{n}\right)$ is given by

$$\text{Bias}\,(d_2) = \frac{k_2}{n\overline{X}}\left(\mu_{21} + 2\overline{Y}\mu_{11}\right) - \frac{k_2}{n'\overline{X}}\left(\mu_{21} + 2\overline{Y}\mu_{11}\right) - \frac{\mu_{20}}{n} \tag{3.18}$$

Where as the mean square error of d_2 to the first degree of approximation for the optimum value of k_2 is given by

$$\text{MSE}\,(d_2)\text{ min} = \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{21}{}^2}{\mu_{02}} \tag{3.19}$$

Proof : Now consider the proposed estimator in (2.3)

$$d_2 = \hat{\theta}\left\{1 + \frac{k_2\left(\bar{x} - \bar{x}'\right)}{\bar{x}'}\right\} - \bar{y}^2$$

$$= (\theta + e_2)\left\{1 + \frac{k_2\left(\overline{X} + e_1 - \overline{X} - e_1{}'\right)}{\overline{X} + e_1{}'}\right\} - (\overline{Y} + e_0)^2$$

$$= (\theta + e_2)\left\{1 + \frac{k_2}{\overline{X}}\left(e_1 - e_1{}'\right)\left(1 + \frac{e_1'}{\overline{X}}\right)^{-1}\right\} - (\overline{Y} + e_0)^2$$

$$= (\theta + e_2)\left\{1 + \frac{k_2}{\overline{X}}\left(e_1 - e_1{}'\right)\left(1 - \frac{e_1'}{\overline{X}} + \frac{e_1'^2}{\overline{X}^2}\right)\right\} - (\overline{Y} + e_0)^2$$

$$= (\theta + e_2)\left(1 - \frac{k_2 e_1'}{\overline{X}} + \frac{k_2 e_1'^2}{\overline{X}^2} + \frac{k_2 e_1}{\overline{X}} - \frac{k_2 e_1 e_1'}{\overline{X}^2}\right) - (\overline{Y} + e_0)^2$$

$$= \theta - \frac{k_2\theta e_1'}{\overline{X}} + \frac{k_2\theta e_1}{\overline{X}} + \frac{k_2\theta e_1'^2}{\overline{X}^2} - \frac{k_2\theta\, e_1 e_1'}{\overline{X}^2} + e_2 - \frac{k_2 e_1' e_2}{\overline{X}} + \frac{k_2 e_1 e_2}{\overline{X}} - \overline{Y}^2 - e_0{}^2 - 2\overline{Y}e_0$$

$$= \left(\frac{1}{N}\sum_{i=1}^{N} Y_i{}^2 - \overline{Y}^2\right) + e_2 - 2\overline{Y}e_0 - \frac{k_2\theta\, e_1'}{\overline{X}} + \frac{k_2\,\theta\, e_1}{\overline{X}} - e_0{}^2 + \frac{k_2\theta e_1'^2}{\overline{X}^2} + \frac{k_2 e_1 e_2}{\overline{X}} - \frac{k_2 e_1' e_2}{\overline{X}} - \frac{k_2\theta\, e_1 e_1'}{\overline{X}^2}$$

$$(d_2 - \sigma_Y^2) = e_2 - 2\overline{Y}e_0 - \frac{k_2\theta\, e_1'}{\overline{X}} + \frac{k_2\theta\, e_1}{\overline{X}} - e_0{}^2 + \frac{k_2\theta\, e_1'^2}{\overline{X}^2} + \frac{k_2 e_1 e_2}{\overline{X}} - \frac{k_2 e_1' e_2}{\overline{X}} - \frac{k_2\theta\, e_1 e_1'}{\overline{X}^2} \tag{3.20}$$

Taking expectation on both sides of (3.20), the bias in $d_2 (= E(d_2) - \sigma_Y^2)$ to the first degree of approximation is given by

Bias (d_2) = $E(d_2) - \sigma_Y^2$

$$= E(e_2) - 2\bar{Y}E(e_0) - \frac{k_2\theta}{\bar{X}}E(e_1') + \frac{k_2\theta}{\bar{X}}E(e_1) - E(e_0{}^2) + \frac{k_2\theta}{\bar{X}^2}E(e_1'^2) + \frac{k_2}{\bar{X}}E(e_1 e_2) - \frac{k_2}{\bar{X}}E(e_1' e_2) - \frac{k_2\theta}{\bar{X}^2}E(e_1 e_1')$$

$$= -\frac{\mu_{20}}{n} + \frac{k_2\theta}{\bar{X}^2}\frac{\mu_{02}}{n'} + \frac{k_2}{\bar{X}}\frac{1}{n}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{k_2}{\bar{X}}\frac{1}{n'}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{k_2\theta}{\bar{X}^2}\frac{\mu_{02}}{n'}$$

$$= \frac{k_2}{n\bar{X}}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{k_2}{n'\bar{X}}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{\mu_{20}}{n} \quad (3.21)$$

Now squaring (3.20) on both sides and then taking expectation, the mean square error of $d_2 \left(= E(d_2 - \sigma_Y^2)^2\right)$ to the first degree of approximation is given by

MSE $(d_2) = E(d_2 - \sigma_Y^2)^2$

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

$$= E(e_2{}^2) + 4\bar{Y}^2 E(e_0{}^2) + \frac{k_2{}^2\theta^2}{\bar{X}^2}E(e_1'^2) + \frac{k_2{}^2\theta^2}{\bar{X}^2}E(e_1{}^2) - 4\bar{Y}E(e_0 e_2) - \frac{2k_2\theta}{\bar{X}}E(e_1' e_2)$$

$$+ \frac{2k_2\theta}{\bar{X}}E(e_1 e_2) + \frac{4k_2\bar{Y}\theta}{\bar{X}}E(e_0 e_1') - \frac{4k_2\bar{Y}\theta}{\bar{X}}E(e_0 e_1) - \frac{2k_2{}^2\theta^2}{\bar{X}^2}E(e_1 e_1')$$

or MSE (d_2) = $E(d_2 - \sigma_Y{}^2)$

$$= \frac{1}{n}(\mu_{40} - \mu_{20}{}^2) + \frac{k_2{}^2\theta^2}{\bar{X}^2}\frac{\mu_{02}}{n'} + \frac{k_2{}^2\theta^2}{\bar{X}^2}\frac{\mu_{02}}{n} - \frac{2k_2\theta}{\bar{X}}\frac{1}{n'}\left[\mu_{21} + 2\bar{Y}\mu_{11}\right]$$

$$+ \frac{2k_2\theta}{\bar{X}}\frac{1}{n}(\mu_{21} + 2\bar{Y}\mu_{11}) + \frac{4k_2\bar{Y}\theta}{\bar{X}}\frac{\mu_{11}}{n'} - \frac{4k_2\bar{Y}\theta}{\bar{X}}\frac{\mu_{11}}{n} - \frac{2k_2{}^2\theta^2}{\bar{X}^2}\frac{\mu_{02}}{n'}$$

$$= \frac{1}{n}(\mu_{40} - \mu_{20}{}^2) + \frac{k_2{}^2\theta^2}{\bar{X}^2}\frac{\mu_{02}}{n} - \frac{k_2{}^2\theta^2}{\bar{X}^2}\frac{\mu_{02}}{n'} - \frac{2k_2\theta}{\bar{X}}\frac{\mu_{21}}{n'} + \frac{2k_2\theta}{\bar{X}}\frac{\mu_{21}}{n}$$

$$= \frac{1}{n}(\mu_{40} - \mu_{20}{}^2) + \frac{k_2{}^2\theta^2}{\bar{X}^2}\mu_{02}\left(\frac{1}{n} - \frac{1}{n'}\right) + \frac{2k_2\theta}{\bar{X}}\mu_{21}\left(\frac{1}{n} - \frac{1}{n'}\right)$$

$$= \frac{1}{n}(\mu_{40} - \mu_{20}{}^2) + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(\frac{k_2{}^2\theta^2}{\bar{X}^2}\mu_{02} + \frac{2k_2\theta}{\bar{X}}\mu_{21}\right) \quad (3.22)$$

The optimum value of k_2 minimizing the mean square error of d_2 is given by

$$k_2{}^* = -\frac{\bar{X}}{\theta}\frac{\mu_{21}}{\mu_{02}} \quad (3.23)$$

which when substituted in (3.22) gives the minimum value of mean square error as

$$\text{MSE}(d_2)\min = \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{21}{}^2}{\mu_{02}} \quad (3.24)$$

showing that the mean square error of proposed estimator d_2 is less than that of usual conventional unbiased estimator $s_y{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ of population variance $\sigma_Y{}^2$.

Theorem IV: The bias in d_{2g} up to terms of order $\left(\frac{1}{n}\right)$ is given by

$$\text{Bias}(d_{2g}) = \frac{\theta}{2\bar{X}^2} f''(1)\frac{\mu_{02}}{n} - \frac{\theta}{2\bar{X}^2} f''(1)\frac{\mu_{02}}{n'} - \frac{\theta}{\bar{X}^2} f''(1)\frac{\mu_{02}}{n'} + \frac{1}{\bar{X}} f'(1)\frac{1}{n}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{1}{\bar{X}} f'(1)\frac{1}{n'}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{\mu_{20}}{n} \qquad (3.25)$$

Where as the mean square error of d_{2g} to the first degree of approximation for the optimum value of $f'(1)$ is given by

$$\text{MSE}(d_{2g})\text{min} = \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{21}{}^2}{\mu_{02}}. \qquad (3.26)$$

Proof: Let us consider (2.4)

$$d_{2g} = \hat{\theta}\ f(w)\ - \bar{y}^2$$

For $f'(1)$, $f''(1)$ and $f'''(1)$ to be first, second and third order derivatives of $f(w)$ at the point $w = 1$ respectively and $w^* = 1 + h\,(w - 1)$, $0 < h < 1$, expanding $f(w)$ in d_{2g} in third order Taylor's series, we have

$$d_{2g} = \hat{\theta}\left\{ f(1) + (w-1)f'(1) + \frac{(w-1)^2}{2!} f''(1) + \frac{(w-1)^3}{3!} f'''(w^*)\right\} - \bar{y}^2$$

$$= \hat{\theta}\left\{1 + \left(\frac{e_1 - e_1'}{\bar{X} + e_1'}\right) f'(1) + \frac{1}{2!}\left(\frac{e_1 - e_1'}{\bar{X} + e_1'}\right)^2 f''(1) + \frac{1}{3!}\left(\frac{e_1 - e_1'}{\bar{X} + e_1'}\right)^3 f'''(u^*)\right\} - \bar{y}^2$$

$$= \hat{\theta}\left\{1 + \frac{1}{\bar{X}}(e_1 - e_1')\left(1 + \frac{e_1'}{\bar{X}}\right)^{-1} f'(1) + \frac{1}{2!}\frac{1}{\bar{X}^2}(e_1 - e_1')^2\left(1 + \frac{e_1'}{\bar{X}}\right)^{-2} f''(1)\right\} - \bar{y}^2$$

$$= \hat{\theta}\left\{1 + \frac{1}{\bar{X}}(e_1 - e_1')\left(1 - \frac{e_1'}{\bar{X}} + \frac{e_1'^2}{\bar{X}^2}\right) f'(1)\right.$$

$$\left. + \frac{1}{2!}\frac{1}{\bar{X}^2}(e_1 - e_1')^2\left(1 - \frac{2e_1'}{\bar{X}} + \frac{3e'}{\bar{X}^2}\right) f''(1)\right\} - \bar{y}^2$$

$$= (\theta + e_2)\left\{1 + \left(\frac{e_1}{\bar{X}} - \frac{e_1'}{\bar{X}} + \frac{e_1'^2}{\bar{X}^2} - \frac{e_1 e_1'}{\bar{X}^2}\right) f'(1)\right.$$

$$\left. + \left(\frac{e_1{}^2}{2\bar{X}^2} - \frac{e_1'^2}{2\bar{X}^2} - \frac{e_1 e_1'}{\bar{X}^2}\right) f''(1)\right\} - \left(\bar{Y} + e_0\right)^2$$

$$= \theta + \theta\left(\frac{e_1}{\bar{X}} - \frac{e_1'}{\bar{X}} + \frac{e_1'^2}{\bar{X}^2} - \frac{e_1 e_1'}{\bar{X}^2}\right) f'(1) + \theta\left(\frac{e_1^{\,2}}{2\bar{X}^2} - \frac{e_1'^2}{2\bar{X}^2} - \frac{e_1 e_1'}{\bar{X}^2}\right) f''(1) + e_2$$

$$+ \left(\frac{e_1 e_2}{\bar{X}} - \frac{e_1' e_2}{\bar{X}}\right) f'(1) - \bar{Y}^2 - e_0^{\,2} - 2\bar{Y} e_0$$

$$= \left(\frac{1}{N}\sum_{i=1}^{N} Y_i^{\,2} - \bar{Y}^2\right) + e_2 - 2\bar{Y}e_0 + \frac{\theta\ e_1}{\bar{X}} f'(1) - \frac{\theta\ e_1'}{\bar{X}} f'(1) - e_0^{\,2} + \frac{\theta\ e_1'^2}{\bar{X}^2} f'(1) + \frac{\theta\ e_1^{\,2}}{2\bar{X}^2} f''(1)$$

$$- \frac{\theta\ e_1'^2}{2\bar{X}^2} f''(1) - \frac{\theta\ e_1 e_1'}{\bar{X}^2} f'(1) - \frac{\theta\ e_1 e_1'}{\bar{X}^2} f''(1) + \frac{e_1 e_2}{\bar{X}} f'(1) - \frac{e_1' e_2}{\bar{X}} f'(1)$$

$$d_{2g} - \sigma_Y^{\,2} = e_2 - 2\bar{Y}e_0 + \frac{\theta\ e_1}{\bar{X}} f'(1) - \frac{\theta\ e_1'}{\bar{X}} f'(1) - e_0^{\,2} + \frac{\theta\ e_1'^2}{\bar{X}^2} f'(1) + \frac{\theta\ e_1^{\,2}}{2\bar{X}^2} f''(1)$$

$$- \frac{\theta\ e_1'^2}{2\bar{X}^2} f''(1) - \frac{\theta\ e_1 e_1'}{\bar{X}^2} f'(1) - \frac{\theta\ e_1 e_1'}{\bar{X}^2} f''(1) + \frac{e_1 e_2}{\bar{X}} f'(1) - \frac{e_1' e_2}{\bar{X}} f'(1) \tag{3.27}$$

Taking expectation on both sides of (3.27), the bias in $d_{2g} = (E(d_{2g}) - \sigma_Y^{\,2})$ to the order $\left(\frac{1}{n}\right)$ is given by

Bias (d_{2g}) $= E(d_{2g}) - \sigma_Y^{\,2}$

$$= E(e_2) - 2\bar{Y}E(e_0) - \frac{\theta}{\bar{X}} f'(1)E(e_1) - \frac{\theta}{\bar{X}} f'(1)E(e_1') - E(e_0^{\,2}) + \frac{\theta}{\bar{X}^2} f'(1)E(e_1'^2)$$

$$+ \frac{\theta}{2\bar{X}^2} f''(1)E(e_1^{\,2}) - \frac{\theta}{2\bar{X}^2} f''(1)E(e_1'^2) - \frac{\theta}{\bar{X}^2} f'(1)E(e_1 e_1') - \frac{\theta}{\bar{X}^2} f''(1)E(e_1 e_1')$$

$$+ \frac{1}{\bar{X}} f'(1)E(e_1 e_2) - \frac{1}{\bar{X}} f'(1)E(e_1' e_2)$$

Using values of the expectations given from (3.3) to (3.5), we have

Bias (d_{2g}) $= E(d_{2g}) - \sigma_Y^{\,2}$

$$= -\frac{\mu_{20}}{n} + \frac{\theta}{\bar{X}^2} f'(1)\frac{\mu_{02}}{n'} + \frac{\theta}{2\bar{X}^2} f''(1)\frac{\mu_{02}}{n} - \frac{\theta}{2\bar{X}^2} f''(1)\frac{\mu_{02}}{n'} - \frac{\theta}{\bar{X}^2} f'(1)\frac{\mu_{02}}{n'} - \frac{\theta}{\bar{X}^2} f''(1)\frac{\mu_{02}}{n'}$$

$$+ \frac{1}{\bar{X}} f'(1)\frac{1}{n}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{1}{\bar{X}} f'(1)\frac{1}{n'}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right)$$

or Bias (d_{2g}) = $E(d_{2g}) - \sigma_Y^{\,2}$

$$= \frac{\theta}{2\bar{X}^2} f''(1)\frac{\mu_{02}}{n} - \frac{\theta}{2\bar{X}^2} f''(1)\frac{\mu_{02}}{n'} - \frac{\theta}{\bar{X}^2} f''(1)\frac{\mu_{02}}{n'} + \frac{1}{\bar{X}} f'(1)\frac{1}{n}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right)$$

$$- \frac{1}{\bar{X}} f'(1)\frac{1}{n'}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{\mu_{20}}{n} \tag{3.28}$$

Now squaring (3.27) on both sides and then taking expectation, the mean square error of d_{2g} to the first degree of approximation is given by

$$E(d_{2g}-\sigma_Y{}^2)^2 = E\left(e_2 - 2\bar{Y}e_0 + \frac{\theta\ e_1}{\bar{X}} f'(1) - \frac{\theta\ e_1'}{\bar{X}} f'(1)\right)^2$$

$$= E\left(e_2{}^2\right) + 4\bar{Y}^2 E\left(e_0{}^2\right) + \frac{\theta^2}{\bar{X}^2}\{f'(1)\}^2 E\left(e_1{}^2\right) + \frac{\theta^2}{\bar{X}^2}\{f'(1)\}^2 E\left(e_1'^2\right) - 4\bar{Y}E(e_0e_2)$$

$$+\frac{2\theta}{\bar{X}} f'(1)E(e_1e_2) - \frac{2\theta}{\bar{X}} f'(1)E\left(e_1'e_2\right) - \frac{4\bar{Y}\theta}{\bar{X}} f'(1)E(e_0e_1)$$

$$+\frac{4\bar{Y}\theta}{\bar{X}} f'(1)E(e_0e_1') - \frac{2\theta^2}{\bar{X}^2}\{f'(1)\}^2 E(e_1e_1')$$

Using values of the expectations given from (3.3) to (3.5), the mean square error of d_{2g} is given by

MSE(d_{2g})

$$= \frac{1}{n}\left(\mu_{40} + 4\bar{Y}\mu_{30} + 4\bar{Y}^2\mu_{20} - \mu_{20}{}^2\right) - 4\bar{Y}^2\frac{\mu_{20}}{n} - 4\bar{Y}\frac{1}{n}(\mu_{30} + 2\bar{Y}\mu_{20}) + \frac{\theta^2}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n}$$

$$+\frac{\theta^2}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n'} + \frac{2\theta}{\bar{X}} f'(1)\frac{1}{n}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{2\theta}{\bar{X}} f'(1)\frac{1}{n'}\left(\mu_{21} + 2\bar{Y}\mu_{11}\right) - \frac{4\bar{Y}\theta}{\bar{X}} f'(1)\frac{\mu_{11}}{n}$$

$$+\frac{4\bar{Y}\theta}{\bar{X}} f'(1)\frac{\mu_{11}}{n'} - \frac{2\theta^2}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n'}$$

$$= \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) + \frac{\theta^2}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n} - \frac{\theta^2}{\bar{X}^2}\{f'(1)\}^2\frac{\mu_{02}}{n'} + \frac{2\theta}{\bar{X}} f'(1)\frac{\mu_{21}}{n} - \frac{2\theta}{\bar{X}} f'(1)\frac{\mu_{21}}{n'} \tag{3.29}$$

The optimum value of $f'(1)$ minimizing mean square error is given by

$$f'(1)^* = -\frac{\bar{X}}{\theta}\frac{\mu_{21}}{\mu_{02}} \tag{3.30}$$

which when substituted in (3.29) gives the minimum value of mean square error as

$$\text{MSE }(d_{2g})\text{min} = \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{21}{}^2}{\mu_{02}} \tag{3.31}$$

showing that mean square error of the proposed double sampling estimator d_{2g} is less than that of usual conventional unbiased estimator $s_y{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ of population variance $\sigma_Y{}^2$.

4. Efficiency Comparison with the Traditional Estimator

As we know that the mean square error of usual conventional unbiased estimator $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ of population variance $\sigma_Y{}^2$ is $\frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right)$ and

$$\text{MSE}(d_1)\text{min} = \text{MSE}(d_{1g})\text{min} = \text{MSE}(d_2)\text{min} = \text{MSE}(d_{2g})\text{min} = \frac{1}{n}\left(\mu_{40} - \mu_{20}{}^2\right) - \left(\frac{1}{n} - \frac{1}{n'}\right)\frac{\mu_{21}{}^2}{\mu_{02}} \tag{4.1}$$

showing that the proposed estimators have less mean square error than the usual conventional unbiased estimator $s_y^{\,2} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ of population variance $\sigma_Y^{\,2}$.

5. Empirical Study

For comparing efficiency of the proposed estimator, Let us consider the data given in Cochran (1977) dealing with Paralytic Polio cases 'Placebo' Y group and Paralytic Polio cases in not inoculated group X. We have calculated the required values of μ_{rs} and a comparison is made.

For $n = 34$ and $n' = 50$ (say), we have

μ_{20} = 9.8894, μ_{02} = 7.1865882×10^7

μ_{40} = 421.96088, μ_{21} = 93.464705×10^3

Mean Square Error of usual conventional unbiased estimator = 9.534136697 and Mean Square Error of the proposed estimators = 8.390090538. The percent relative efficiency (PRE) of the proposed estimators over the usual conventional unbiased estimator is 113.63568, Showing that the proposed estimators are more efficient than the usual conventional unbiased estimator.

6. Conclusion

We have derived new double sampling estimators and their generalized estimators of population variance using auxiliary information in the form of mean, the bias and mean square error equations are obtained. Using these equations, MSE of proposed estimators are compared with the traditional estimator in theory and it is shown that the proposed estimators have smaller MSE than the traditional one. For the practical justification of the results, an empirical study is also included. It may be noted here that when the optimum value is replaced by the estimated optimum value depending on sample values, the resulting estimators based on the estimated optimum value attains the same minimum mean square error to the first degree of approximation as that of the estimators depending upon optimum value. The details are here omitted because of derivation being straight forward.

Acknowledgment

The authors are thankful to the referees and Editor for their valuable comments and suggestions regarding improvement of the paper.

References

1. Cochran, W. C. (1977). Sampling Techniques, 3rd Edition, John Wiley and Sons, New York.
2. Isaki, C.T.(1983). Variance estimation using auxiliary information, Journal of American Statistical Association, 78, p. 117–123.
3. Singh, R. K., Zaidi, S. M. H. and Rizvi, S. A. H (1995). Estimation of finite population variance using auxiliary information, Journal of Statistical Studies, Vol. 15, p. 17-28.
4. Srivastava, S. K. and Jha, J. J. H. S. (1980). A class of estimators using auxiliary information for estimating finite population variance, Sankhya, C 42, p. 87-96.