

ON SOME MODIFIED RATIO AND PRODUCT TYPE ESTIMATORS-REVISITED

A K P C Swain¹

Former Professor of Statistics , Utkal University,Bhubaneswar-751004, India

ABSTRACT

In this paper different modified ratio and product type estimators are considered. Their biases and mean square errors are compared and necessary conditions are derived. Further, numerical illustrations are provided with the help of some natural population to compare their biases and efficiencies.

KEYWORDS: simple random sampling, auxiliary information, transformed ratio estimators, predictive ratio estimators, predictive product estimators, dual to ratio estimator, dual to product estimator, bias, mean square error.

MSC 62D05

RESUMEN

En este trabajo se consideran diferentes estimadores del tipo razón modificado y producto. Sus sesgos y errores cuadráticos medios son comparados. Condiciones necesarias son derivadas. Además se presentan ilustraciones numéricas con la ayuda de algunas poblaciones naturales para comprobar su sesgos y eficiencias.

1. INTRODUCTION

In survey sampling a researcher often comes across supplementary information supplied by auxiliary variables, which are correlated with the study variable and he takes advantage of this to devise methods of estimation to obtain more precise estimators of finite population parameters in question. The use of auxiliary information in sample surveys dates back to Cochran(1940), who used it for estimation of yields of agricultural crops in course of his researches in agricultural sciences. Since then there have been tremendous advancements in sampling theory using auxiliary information both at design and estimation stage. The popular estimation methods that are widely discussed in sampling theory literature and used in practice are ratio, product and regression methods of estimation. All these three methods are biased estimators and bias decreases with increase in sample size. Ratio (or product) estimator is used when the auxiliary variable is highly and positively (or highly and negatively) correlated with the study variable and under certain conditions is more efficient than the simple mean per unit estimator to estimate either the population mean or population total of the study variable. In large samples ratio and product estimators are less efficient than the linear regression estimator and have equal efficiency when the regression of study variable on the auxiliary variable is a straight line passing through the origin. In practice, sometimes the regression of the study variable on the auxiliary variable is linear but does not pass through the origin. In such a situation it becomes more appropriate to search for a transformation of the auxiliary variable to estimate the population total or mean of the study variable (see Mohanty and Das, 1971), which may be asymptotically equal to the mean square error of the linear regression estimator or at least close to it compared to the classical ratio or product estimators. The simplicity in computations of ratio and product estimators have encouraged the researchers to devise various ways to modify the classical ratio and product estimators to arrive at more precise estimators and sometimes arriving at estimators having same asymptotic efficiency as that of the linear regression estimator. The use of auxiliary information in sample surveys is vividly discussed in well known classical text books such as Cochran(1977), Sukhatme and Sukhatme(1970), Sukhatme, Sukhatme, and Asok(1984), Murthy(1967) and Yates(1960) among others. Recent developments in ratio and product methods of estimation along with their variety of modified forms are lucidly described in detail by Singh, Sarjinder(2003).

Let $U = (U_1, U_2, \dots, U_N)$ be the finite population of size N . To each unit U_i ($i=1, 2, \dots, N$) in the population paired values (y_i, x_i) corresponding to study variable y and an auxiliary variable x , correlated with y are attached.

Now, define the population means of the study variable y and auxiliary variable x as

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i, \quad \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$$

Thus, the population ratio is defined as $R = \frac{\bar{Y}}{\bar{X}}$

¹akpcs@rediffmail.com

Further, define the finite population variances of y and x and their covariance as $S_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1}$, $S_x^2 = \frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N-1}$ and $S_{yx} = \frac{\sum_{i=1}^N (y_i - \bar{Y})(x_i - \bar{X})}{N-1}$, respectively.

Thus, the coefficients of variation of y and x and their coefficient of covariation are respectively defined by

$$C_y = \frac{S_y}{\bar{Y}}, C_x = \frac{S_x}{\bar{X}} \text{ and } C_{yx} = \frac{S_{yx}}{\bar{Y}\bar{X}} = \frac{\rho S_y S_x}{\bar{Y}\bar{X}} = \rho C_y C_x,$$

where ρ is the correlation coefficient between y and x .

A simple random sample 's' of size n is selected from U without replacement and the values (y_i, x_i) , $i=1, 2 \dots n$ are observed on the sampled units. Define \bar{y} and \bar{x} as sample means of y and x respectively, given by

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

$$\text{Define } \theta = \frac{1}{n} - \frac{1}{N} = \frac{1-f}{n}, f = \frac{n}{N} \text{ and } r = \frac{\bar{y}}{\bar{x}}$$

$$\text{Further define } s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2, s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2,$$

$$s_{yx} = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}), b = \frac{s_{yx}}{s_x^2}.$$

In simple random sampling without replacement the sample mean \bar{y} is an unbiased estimator of the population mean \bar{Y} and

$$V(\bar{y}) = \theta \bar{Y}^2 C_y^2 \quad (1.1)$$

If information on an auxiliary variable (x) is available, one may resort to ratio or product or regression methods of estimation to improve upon the efficiency of \bar{y} .

When y and x are positively correlated, the classical ratio estimator

$$\bar{y}_R = \frac{\bar{y}}{\bar{x}} \bar{X} \quad (1.2)$$

To $O(1/n)$ the bias and mean square error of \bar{y}_R are found (Sukhatme and Sukhatme, 1970) as

$$\text{Bias}(\bar{y}_R) = \theta \bar{Y} (C_x^2 - \rho C_y C_x) \quad (1.3)$$

$$\text{MSE}(\bar{y}_R) = \theta \bar{Y}^2 (C_y^2 + C_x^2 - 2\rho C_y C_x) \quad (1.4)$$

\bar{y}_R is more efficient than \bar{y} if

$$\rho \frac{C_y}{C_x} > \frac{1}{2} \quad (1.5)$$

As \bar{y}_R is a biased estimator it can be made unbiased by the following Hartley and Ross (1954) technique and an unbiased Hartley-Ross unbiased ratio type estimator is given by

$$\bar{y}_{HR} = \bar{r} \bar{X} + \frac{(N-1)n}{N(n-1)} (\bar{y} - \bar{r} \bar{x}). \quad (1.6)$$

The variance of \bar{y}_{HR} to $O(1/n)$ is found as (Sukhatme, Sukhatme and Asok, 1984)

$$V(\bar{y}_{HR}) = \theta (S_y^2 + \bar{R}^2 S_x^2 - 2\bar{R} \rho S_y S_x) \quad (1.7)$$

Murthy(1964) proposed product estimator to estimate the population mean \bar{Y} , when y and x are negatively correlated, given by

$$\bar{y}_P = (\bar{y} - \bar{x}) / \bar{X} \quad (1.8)$$

This estimator was previously obtained by Robson(1957) in applying Multivariate Polykays.

The bias and mean square error of \bar{y}_P are found as(Sukhatme, Sukhatme and Asok,1984)

$$Bias(\bar{y}_P) = \theta \bar{Y} C_{yx} \quad (1.9)$$

An unbiased estimator of \bar{Y} after correcting the bias of \bar{y}_P is given by

$$\bar{y}_{PU} = \bar{y}_P - \theta \frac{s_{yx}}{\bar{X}} \quad (1.10)$$

$$\text{To } O(1/n), \text{ } MSE(\bar{y}_P) = MSE(\bar{y}_{PU}) = \theta \bar{Y}^2 (C_y^2 + C_x^2 + 2\rho C_y C_x) \quad (1.11)$$

\bar{y}_P is more efficient than \bar{y} if

$$\rho \frac{C_y}{C_x} < -\frac{1}{2} \quad (1.12)$$

The classical linear regression estimator is given by

$$\bar{y}_{Reg} = \bar{y} + b(\bar{X} - \bar{x}) \quad (1.13)$$

To $O(1/n)$,

$$MSE(\bar{y}_{Reg}) = \theta \bar{Y}^2 C_y^2 (1 - \rho^2) \text{ (Sukhatme and Sukhatme,1970)} \quad (1.14)$$

On comparison of $MSE(\bar{y}_R)$ and $MSE(\bar{y}_P)$ with $MSE(\bar{y}_{Reg})$ it may be seen from (1.4),(1.11) and (1.14) that \bar{y}_{Reg} is more efficient than both \bar{y}_R and \bar{y}_P .

During last five decades the researchers have suggested a number of modified ratio type and product type estimators which are less biased and more efficient than the classical ratio estimator and product estimator respectively and in certain situations have asymptotic mean square errors equal to that of the asymptotic mean square error of the classical linear regression estimator

A number of researchers namely Srivastava(1967), Walsh(1970), Reddy(1973,1974,1978), Mohanty and Das(1971),Mohanty and Sahoo(1995),Sisodia and Dwivedi(1981)Mishra and Sahoo(1986),Singh(2003),Singh et.al(2004,2008,2010),Sen(1978),Pandey and Dubey(1988),Singh and Tailor(2003),Srivenkataramana(1978) and Upadhyaya and Singh(1999) among others have resorted to transformation in auxiliary variable to arrive at improved ratio type and product type estimators. Some of their contributions ,though not exhaustive, are mentioned below.

Srivastava (1967) proposed a modified ratio estimator using power transformation as

$$\hat{Y}_{SR} = \bar{y} \frac{\bar{X}^\alpha}{\bar{x}},$$

where α is a suitably chosen constant obtained by minimizing the approximate mean square error (MSE) of \hat{Y}_{SR} to $O\left(\frac{1}{n}\right)$. Srivastava, Jhajj and Sharma(1986) proposed an alternative estimator of the type $\bar{y}_{SJS} = (1-w)\bar{y} + w\left(\frac{\bar{X}}{\bar{x}}\right)$,

where $0 < w < 1$.

Sisodia and Dwivedi(1981) proposed an estimator of the form $\bar{y}_{SD}' = (1-b)\bar{y} + b\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^p$, where $0 < b < 1$

and $p > 0$.

Walsh (1970) and Reddy (1974) proposed an alternative modified ratio estimator as

$$\hat{Y}_{WR} = \frac{\bar{y}\bar{X}}{\alpha\bar{x} + (1-\alpha)\bar{X}}, \text{ where } \alpha \text{ is a real constant and suitably chosen by minimizing approximate MSE of } \hat{Y}_{WR} \text{ to } O\left(\frac{1}{n}\right).$$

Sisodia and Dwivedi (1981) and Singh and Tailor(2003), following Sen(1978) suggested a modified ratio estimator with the advance knowledge of C_x the coefficient of variation of the auxiliary variable x as $\hat{Y}_{SD} = \bar{y} \frac{\bar{X} + C_x}{\bar{x} + C_x}$, Swain(1981) used known values of the coefficients of variations of both study and auxiliary variables to estimate population ratio. Pandey and Dubey(1988) modified product estimator by using known coefficient variation of auxiliary variable. Singh and Kakran(1993), Upadhyaya and Singh(1999), Singh et al. (2004), Kadilar and Cingi(2006 a) and Tailor and Sharma(2009) used advance knowledge of the coefficient of Kurtosis $\beta_2(x)$ and C_x , the coefficient of variation of x to arrive at improved estimators of the population mean of the study variate as

$$\hat{Y}_{US} = \bar{y} \left(\frac{\bar{X}C_x + \beta_2(x)}{\bar{x}C_x + \beta_2(x)} \right)$$

Singh (2003) used known value S_x , standard deviation of x in forming a modified product estimator as

$$\hat{Y}_{GS} = \bar{y} \left(\frac{\bar{x} + S_x}{\bar{X} + S_x} \right)$$

Singh and Tailor (2003) assumed advance knowledge of the coefficient of correlation ρ to form a modified ratio estimator of the population mean of y as

$$\hat{Y}_{ST} = \bar{y} \left(\frac{\bar{X} + \rho}{\bar{x} + \rho} \right)$$

Mishra and Sahoo (1986), motivated by Mohanty and Das (1971) and Srivenkantaramana(1978) discussed a generalized modified ratio estimator of the form

$$\hat{Y}_{MS} = \bar{y} \frac{A\bar{X} + B}{A\bar{x} + B},$$

which reduces to Walsh(1970) with $A = \alpha$ and $B = (1-\alpha)\bar{X}$, for any arbitrary values of A and B .

Mohanty and Sahoo (1995) suggested a modified ratio estimator through linear transformation using known extreme values x_{\max} and x_{\min} of the maximum and minimum values respectively of x , besides the known population mean \bar{X} of x .

Other developments in sampling theory concerning ratio and product estimators are (i) dual ratio and dual product estimators (Srivenkantaramana, 1980 and Bandopadhyaya, 1980) (ii) predictive ratio type and product type estimators (Basu, 1971; Smith, 1976; Sampford, 1978; Srivastava, 1983 and Agrawal and Sthapit, 1997), aiming at exploring more efficient estimators compared to classical ones.

Prabhu-Ajgaonkar(1993) showed that the optimum estimator does not exist uniformly in the class $\bar{y}_c = \bar{y}f(u)$, where $u = \bar{x} / \bar{X}$ and $f(\cdot)$ is a parametric function which may depend on known parameters like \bar{X}, S_x^2, C_x^2 etc. satisfying certain regularity conditions such as $f(1) = 1$ and $f(u)$ is continuous and has bounded first and second derivatives for $u \neq 0$.

In this paper we make a comparative study of (i) modified ratio and product estimators, using transformation in auxiliary variable (ii) dual ratio and dual product estimators and (iii) Predictive ratio type and product type estimators theoretically and with help of numerical examples as regards their biases and efficiencies.

2. GENERALIZED MODIFIED RATIO ESTIMATORS

Consider a transformation in auxiliary variable x as

$$z_i = \alpha x_i + (1-\alpha)\bar{X},$$

where α is a constant to be suitably chosen.

The correlation coefficient ρ between y and z is same as the correlation coefficient between y and x .
Now define

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = \alpha \bar{x} + (1 - \alpha) \bar{X}$$

and

A generalized class of modified ratio estimator is defined as

$$\hat{Y}_{GR} = \bar{y} \left[\frac{\bar{Z}}{\bar{z}} \right] = \bar{y} \left[\frac{\bar{X}}{\alpha \bar{x} + (1 - \alpha) \bar{X}} \right], \quad \dots (2.1)$$

where α is chosen in an optimum manner so as to minimize MSE of \hat{Y}_{GR} to 0 $\frac{1}{n}$.

2.1. Bias and Mean square error of Generalized modified ratio estimator

Let $\bar{y} = \bar{Y}(1 + e_0)$, $\bar{x} = \bar{X}(1 + e_1)$ with $E(e_0) = 0, E(e_1) = 0, V(e_0) = \theta C_y^2, V(e_1) = \theta C_x^2$

And $Cov(e_0, e_1) = \theta C_{yx}$, where $\theta = \left(\frac{1}{n} - \frac{1}{N} \right) = \frac{1-f}{n}$, $f = \frac{n}{N}$

Expanding \hat{Y}_{GR} in a Taylor series expansion and keeping terms up to second degree in e_1 we have

$$\hat{Y}_{GR} = \bar{Y} + \bar{Y}(e_0 - \alpha e_1 + \alpha^2 e_1^2 - \alpha e_0 e_1) \quad (2.2)$$

Thus, to terms of $O(1/n)$,

$$Bias(\hat{Y}_{GR}) = \theta \bar{Y}(\alpha^2 C_x^2 - \alpha \rho C_y C_x) \quad (2.3)$$

$$MSE \hat{Y}_{GR} = \theta \bar{Y}^2 [C_y^2 + \alpha^2 C_x^2 - 2\alpha \rho C_y C_x] \quad (2.4)$$

Now, minimizing MSE \hat{Y}_{GR} , with respect to α , we have the optimum value of α is given by

$$\alpha_{opt} = \rho \frac{C_y}{C_x} = \frac{\beta}{R} \quad (2.5)$$

where β is the linear regression coefficient of y on x .

Substituting α_{opt} in the expression for MSE \hat{Y}_{GR} in (2.4), we have

$$MSE \hat{Y}_{GR_{opt}} = \theta \bar{Y}^2 [C_y^2 (1 - \rho^2)] \quad (2.6)$$

Again, substituting the optimum value of α in the expression for $Bias \hat{Y}_{GR}$ in (2.3), we get

$$Bias(\hat{Y}_{GR})_{opt} = \theta \bar{Y} \left[\left(\rho \frac{C_y}{C_x} \right)^2 C_x^2 - \left(\rho \frac{C_y}{C_x} \right) \rho C_y C_x \right] = 0 \quad (2.7)$$

Thus, with optimum α , \hat{Y}_{GR} is rewritten as

$$\hat{Y}_{GR} = \frac{\bar{y}}{\bar{x}} \bar{X} \left[\frac{R}{\beta + (R - \beta) \frac{\bar{X}}{\bar{x}}} \right] \quad (2.8)$$

As β and R are not known advance we substitute their consistent estimators sample regression coefficient b and sample ratio r respectively in the expression for \hat{Y}_{GR} in (2.8) so as to get its reduced form as

$$\hat{Y}_{GR}^* = \left(\frac{\bar{y}}{\bar{x}}\right) \bar{X} \left[\frac{r}{b + (r - b)\left(\frac{\bar{X}}{\bar{x}}\right)} \right] \quad (2.9)$$

By Taylor's series approximation we write (2.9) as

$$\hat{Y}_{GR}^* \sqcup \bar{Y} \left[1 + e_0 - \left(\frac{\beta}{R}\right)e_1 - \left(\frac{\beta}{R}\right)e_3e_1 + \left(\frac{\beta}{R}\right)e_2e_1 + \left(\frac{\beta}{R} - 1\right)\left(\frac{\beta}{R}\right)e_1^2 + \dots \right], \quad (2.10)$$

where $e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}$, $e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}}$, $e_2 = \frac{s_x^2 - S_x^2}{S_x^2}$ and $e_3 = \frac{s_{yx} - S_{yx}}{S_{yx}}$ with

$$E(e_0) = E(e_1) = E(e_2) = E(e_3) = 0.$$

Thus, to $O(1/n)$,

$$Bias(\hat{Y}_{GR}^*) = \left[\left(\frac{\beta}{R} - 1\right)\left(\frac{\beta}{R}\right) \frac{V(\bar{x})}{\bar{X}^2} - \left(\frac{\beta}{R}\right) \frac{cov(s_{yx}, \bar{x})}{S_{yx}\bar{X}} + \left(\frac{\beta}{R}\right) \frac{Cov(s_x^2, \bar{x})}{S_x^2\bar{X}} \right] \quad (2.11)$$

$$MSE(\hat{Y}_{GR}^*) = \theta C_y^2 (1 - \rho^2), \quad (2.12)$$

which is equal to asymptotic mean square error of the linear regression estimator.

Under bivariate normality of (y, x) with parameters $(\bar{Y}, \bar{X}, \sigma_y^2, \sigma_x^2, \rho\sigma_y\sigma_x)$ the bias of \hat{Y}_{GR}^* is

$$\frac{\beta}{R} \left(\frac{\beta}{R} - 1 \right) \frac{V(\bar{x})}{\bar{X}^2} = \theta \bar{Y} (\rho^2 C_y^2 - \rho C_y C_x),$$

but the bias \bar{Y}_{Reg} vanishes to $O(1/n)$.

3. GENERALIZED MODIFIED PRODUCT ESTIMATOR

When y and x are negatively correlated, we may define a generalized product estimator as

$$\hat{Y}_{GP} = \frac{\bar{y}}{\bar{X}} [\alpha \bar{x} + (1 - \alpha) \bar{X}] \quad (3.1)$$

Linearizing \hat{Y}_{GP} we have $\bar{Y}_{GP} \cong \bar{Y}(1 + e_0 + \alpha e_1 + \beta e_0 e_1)$, where $e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}$ and $e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}}$

$$E(\hat{Y}_{GP}) = \bar{Y} + \theta \bar{Y} (\alpha \rho C_y C_x) \quad (3.2)$$

$$MSE(\hat{Y}_{GP}) = \theta \bar{Y}^2 (C_y^2 + \alpha^2 C_x^2 + 2\alpha \rho C_y C_x) \quad (3.3)$$

Optimizing (3.3) with respect to α , we have

$$\alpha_{opt} = -\rho \frac{C_y}{C_x} = -\frac{\beta}{R} \quad (3.4)$$

Substituting α_{opt} in the expression for \hat{Y}_{GP} , we have

$$\hat{Y}_{GP} = \bar{y} \left[1 - \frac{\beta}{R} \left(\frac{\bar{x} - \bar{X}}{\bar{X}} \right) \right] \quad (3.5)$$

with bias and mean square error given by

$$Bias(\hat{Y}_{GP})_{opt} = -\theta \bar{Y} \rho \frac{C_y}{C_x} \rho C_y C_x = -\theta \bar{Y} \rho^2 C_y^2 \quad (3.6)$$

$$MSE(\hat{Y}_{GP})_{opt} = \theta \bar{Y}^2 C_y^2 (1 - \rho^2) \quad (3.7)$$

As β and R are not known in practice, we substitute their consistent estimates b and r respectively in the expression for \hat{Y}_{GP} to get

$$\hat{Y}_{GP}^* = \bar{y} \left[1 - \frac{b}{r} \left(\frac{\bar{x} - \bar{X}}{\bar{X}} \right) \right] \quad (3.8)$$

Linearizing \hat{Y}_{GP}^* by Taylor series approximation we have

$$\hat{Y}_{GP}^* = \bar{Y} \left[1 + e_0 - \frac{\beta}{R} (e_1 + e_1 e_3 + e_1^2 - e_1 e_2 - e_0 e_1) \right],$$

where

$$e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}, e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}}, e_2 = \frac{s_x^2 - S_x^2}{S_x^2}, e_3 = \frac{s_{yx} - S_{yx}}{S_{yx}}$$

$$\begin{aligned} Bias(\hat{Y}_{GP}^*) &= -\frac{\beta}{R} \bar{Y} \left[\frac{Cov(s_{yx}, \bar{x})}{S_{yx} \bar{X}} + \frac{V(\bar{x})}{\bar{X}^2} - \frac{Cov(\bar{y}, \bar{x})}{\bar{Y} \bar{X}} - \frac{Cov(s_x^2, \bar{x})}{S_x^2 \bar{X}} \right] \\ &= \beta \left[\frac{Cov(s_x^2, \bar{x})}{S_x^2} - \frac{Cov(s_{yx}, \bar{x})}{S_x^2} - \frac{V(\bar{x})}{\bar{X}} + \frac{Cov(\bar{y}, \bar{x})}{\bar{X}} \right] \end{aligned} \quad (3.9)$$

$$MSE(\hat{Y}_{GP}^*) = \theta \bar{Y}^2 C_y^2 (1 - \rho^2) \quad (3.10)$$

Under bivariate normality of (y,x) with parameters $(\bar{Y}, \bar{X}, \sigma_y^2, \sigma_x^2, \rho \sigma_y \sigma_x)$ the bias of \hat{Y}_{GP}^* to $O(1/n)$ is

$$Bias(\hat{Y}_{GP}^*) = \beta \left(\frac{Cov(\bar{y}, \bar{x})}{\bar{Y}} - \frac{Var(\bar{x})}{\bar{X}} \right) = \theta \bar{Y} (\rho^2 C_y^2 - \rho C_y C_x),$$

where as the bias of \bar{y}_{Reg} is zero.

Comments: As ρ is assumed to be negative, under bivariate normality the $Bias(\hat{Y}_{GP}^*)$ is positive.

4. DUAL TO RATIO AND PRODUCT ESTIMATORS :

Srivenkataraman (1980) proposed a new product type estimator which is complementary in certain sense to the classical ratio estimator. He calls the new product type estimator as dual to ratio estimator. The advantage of the dual ratio estimator lies in the fact that the bias of \hat{Y}_{DR} is an exact expression, which can be estimated from the sample to provide an unbiased dual to ratio estimator. Further its mean square error can be derived without any approximation.

4.1. Dual to Ratio Estimator :

Make a transformation $x_i^* = \frac{N\bar{X} - x_i}{N - n}$ ($i = 1, 2, \dots, N$). We have now $\bar{x}^* = \frac{N\bar{X} - n\bar{x}}{N - n}$

If y is positively correlated with x , y is negatively correlated with the transformed variable x^* . This made Srivenkataramana(1980) to define a product estimator as dual to ratio estimator as

$$\hat{Y}_{DR} = \bar{y} \left(\frac{\bar{x}^*}{\bar{X}} \right) \quad (4.1)$$

The bias of \hat{Y}_{DR} is found as

$$Bias(\hat{Y}_{DR}) = -\theta \bar{Y} \gamma k C_x^2 \quad (4.2)$$

$$\text{where } k = \frac{C_{yx}}{C_x^2} = \rho \frac{C_y}{C_x} \text{ and } \gamma = \frac{n}{N - n} = \frac{f}{1 - f}$$

Further, as seen from (4.2) \hat{Y}_{DR} is an inconsistent estimator of \bar{Y} . To $O(1/n)$,

$$MSE(\hat{Y}_{DR}) = \theta \bar{Y}^2 (C_y^2 + \gamma^2 C_x^2 - 2\gamma \rho C_y C_x) \quad (4.3)$$

When $\rho > 0$, k positive and hence bias is always negative.

\hat{Y}_{DR} is to be preferred to \bar{y}_R as regards efficiency, when $k < \frac{1}{2}(1 + \gamma)$, assuming $(1 - \gamma) > 0$ or $n < \frac{N}{2}$.

4.2. Dual to Product estimator:

The dual to product estimator is defined as

$$\hat{Y}_{DP} = \frac{\bar{y}}{\bar{x}^*} \bar{X} = \frac{\bar{y} \bar{X} (N - n)}{N \bar{X} - n \bar{x}} = \frac{\bar{y} \bar{X} (1 - f)}{\bar{X} - f \bar{x}} \quad (4.4)$$

To $O(1/n)$,

$$Bias(\hat{Y}_{DP}) = \theta \bar{Y} \gamma (k + 1) C_x^2 \quad (4.5)$$

and

$$MSE(\hat{Y}_{DP}) = \theta \bar{Y}^2 (C_y^2 + \gamma^2 C_x^2 + 2\gamma \rho C_y C_x) \quad (4.6)$$

\hat{Y}_{DP} is to be preferred over \bar{y}_P if $k > -\frac{1}{2}(1 + \gamma)$, assuming $(1 - \gamma) > 0$, k being negative because $\rho < 0$.

It has been seen that the dual to product estimator \hat{Y}_{DP} is in a ratio form and also a biased estimator, which can be made unbiased by following the technique of Hartley and Ross (1954). To this end, make transformation

$$z_i = \frac{N\bar{X} - nx_i}{N - n} \quad (4.7)$$

Then the sample mean of z is

$$\bar{z} = \frac{N\bar{X} - n\bar{x}}{N - n} \quad (4.8)$$

and the population mean of z is

$$\bar{Z} = \bar{X} \quad (4.9)$$

Further, $V(\bar{z}) = \left(\frac{f}{1-f}\right)^2 \theta S_x^2 = \gamma^2 \theta S_x^2$ and $Cov(\bar{y}, \bar{z}) = -\frac{f}{1-f} \theta S_{yx} = -\gamma \theta S_{yx}$

Define

$$\bar{r}_1 = \frac{1}{n} \sum_{i=1}^n \frac{y_i(N-n)}{N\bar{X} - nx_i} \quad (4.10)$$

Using the technique of Hartley and Ross(1954) we have an unbiased dual to product estimator, given by (see Singh, Sarjinder, 2003)

$$\hat{Y}_{DPHR} = \bar{r}_1 \bar{Z} + \frac{(N-1)n}{N(n-1)} (\bar{y} - \bar{r}_1 \bar{z}) \quad (4.11)$$

The variance of \hat{Y}_{DPHR} to $O(1/n)$ is given by

$$V(\hat{Y}_{DPHR}) = \theta(S_y^2 + \bar{R}_1^2 \gamma^2 S_x^2 - 2\bar{R}_1 \gamma S_{yx}), \quad (4.12)$$

where $\bar{R}_1 = \frac{1}{N} \sum_{i=1}^N \frac{y_i(N-n)}{N\bar{X} - n\bar{x}}$ and $\gamma = \frac{n}{N-n}$. Now,

$$V(\bar{y}_{PU}) - V(\hat{Y}_{DPHR}) = \theta \left[(R + \beta)^2 - (\bar{R}_1 \gamma + \beta)^2 \right] S_x^2$$

and

$$V(\hat{Y}_{DP}) - V(\hat{Y}_{DPHR}) = \theta \left[(R\gamma + \beta)^2 - (\bar{R}_1 \gamma + \beta)^2 \right] S_x^2$$

Thus, \hat{Y}_{DPHR} would be more efficient than \bar{y}_{PU} if $|R + \beta| > |\bar{R}_1 \gamma + \beta|$ and \hat{Y}_{DPHR} would be more efficient than \hat{Y}_{DP} if $|R\gamma + \beta| > |\bar{R}_1 \gamma + \beta|$.

4.3. Numerical Illustration:

Consider a simple random sample of size $n=15$ drawn from a finite population of size $N=50$ with x =age in years and y =daily hours of sleep.

x : 80 75 70 50 55 50 45 40 35 65 60 52 25 72 35

y : 4 4 5 6 7 6 8 7 8 7 7 7 8 5 8

Given $\bar{X} = 55$. Now,

$$\gamma = \frac{n}{N-n} = \frac{15}{55} = 0.4286, \hat{S}_y^2 = 1.9809, \hat{S}_x^2 = 266.4952, \hat{S}_{yx} = -19.8238, \hat{R} = 0.1199, \hat{R}_1 = 0.1156$$

Estimate	Estimate of Variance	Percent Relative Efficiency
\bar{y}	$\theta(1.9809)$	100
\hat{Y}_{PU}	$\theta(1.0583)$	187
\hat{Y}_{DPHR}	$\theta(0.6714)$	295
\hat{Y}_{DP}	$\theta(0.6472)$	306

Discussions: Numerical illustration shows that there might occur situations when the Hartley-Ross type unbiased dual to product estimator will be more efficient than the mean per unit estimator and product type estimator. However for the present illustration this is less efficient than dual to product estimator.

4.4. Generalized Class of Dual to Product –cum Dual to Ratio Estimator

A generalized class of dual to product-cum-dual to ratio estimators may be suggested by using transformation as

$$\hat{Y}_{DGR} = \bar{y} \left[\alpha \left(\frac{\bar{X}}{\bar{x}^*} \right)^g + (1-\alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right)^h \right]^\delta, \quad (4.13)$$

where $\bar{x}^* = \frac{N\bar{X} - n\bar{x}}{N-n}$, $\bar{y} \left(\frac{\bar{X}}{\bar{x}^*} \right)$ is the dual to product estimator and $\bar{y} \left(\frac{\bar{x}^*}{\bar{X}} \right)$ is the dual to ratio estimator ;

g, h and δ are the free parameters and α is chosen so as to minimize the approximate mean square error of \hat{Y}_{DGR} to $O(1/n)$.

Now $V(\bar{x}^*) = \theta \bar{Y}^2 C_{x^*}^2$ and $Cov(\bar{y}, \bar{x}^*) = \theta \bar{Y}^2 C_{yx^*}$, where $C_{x^*}^2 = \gamma^2 C_x^2$ and $C_{yx^*} = -\gamma C_{yx}$

Define $k^* = \frac{C_{yx^*}}{C_{x^*}^2} = -\frac{k}{\gamma}$, where $k = \frac{C_{yx}}{C_x^2}$

Let $\bar{y} = \bar{Y}(1 + e_0)$ and $\bar{x}^* = \bar{X}(1 + e_1^*)$

Assuming $|e_1^*| < 1$ for all samples we may expand (4.13) in binomial series and keeping terms up to second degree in e_1^* , we have

$$\hat{Y}_{DGR} = \bar{Y}(1 + e_0 + \delta A e_1^* + \delta A e_0 e_1^* + \delta B e_1^{*2} + \frac{\delta(\delta-1)}{2} A^2 e_1^{*2} + \dots), \quad (4.14)$$

where $A = (1-\alpha)h - \alpha g$ and $B = \left[\alpha \frac{g(g+1)}{2} + (1-\alpha) \frac{h(h-1)}{2} \right]$

To $O(1/n)$, $MSE(\hat{Y}_{DGR}) = \theta \bar{Y}^2 \left[C_y^2 + \delta^2 A^2 C_{x^*}^2 + 2\delta A C_{yx^*} \right]$

Minimizing with respect to α , we have

$$\alpha_{opt} = \frac{k^* + \delta h}{\delta(g+h)} = \frac{\gamma \delta h - k}{\gamma \delta(g+h)} \quad (4.15)$$

Substituting the optimum value of α in $MSE(\hat{Y}_{DGR})$, the optimum mean square error of \hat{Y}_{DGR} is given by

$$MSE(\hat{Y}_{DGR})_{opt} = \theta \bar{Y}^2 C_y^2 (1 - \rho^2) \quad (4.16)$$

Thus to $O(1/n)$, $MSE(\hat{Y}_{DGR}) = MSE(\hat{Y}_{GR}) = \theta \bar{Y}^2 C_y^2 (1 - \rho^2)$, which is independent of free parameters g, h and δ . To $O(1/n)$

$$\begin{aligned}
Bias(\hat{Y}_{DGR}) &= \bar{Y}E \left[\delta A e_0 e_1^* + \delta B e_1^{*2} + \frac{\delta(\delta-1)}{2} A^2 e_1^{*2} \right] \\
&= \bar{Y}\theta \left[\delta AC_{yx}^* + \delta BC_x^{*2} + \frac{\delta(\delta-1)}{2} A^2 C_x^{*2} \right]
\end{aligned} \tag{4.17}$$

Substituting the optimum value of α in (4.17) we have

$$\begin{aligned}
Bias(\hat{Y}_{DGR})_{opt} &= \theta \bar{Y} \left[\left\{ \frac{\delta h \gamma - k}{\gamma(g+h)} \left(\frac{g(g+1)}{2} - \frac{h(h-1)}{2} \right) + \frac{h(h-1)\delta}{2} \right\} \gamma^2 C_x^2 + \frac{\delta-1}{2\delta} k^2 C_x^2 - k C_{yx} \right]
\end{aligned} \tag{4.18}$$

However, the bias of \hat{Y}_{DGR} would vary as per different choices of g, h, γ and δ . Some special cases of \hat{Y}_{DGR} are shown in Table-1

Table 1. Special cases of generalized class of dual to product-cum-dual to ratio estimators.

Sl.No.	g	h	δ	Estimator	α_{opt}	Bias with α_{opt} excepting the multiplier $\theta \bar{Y}$
1	1	1	1	$t_1^* = \bar{y} \left[\alpha \left(\frac{\bar{X}}{\bar{x}^*} \right) + (1-\alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right) \right]$	$\frac{\gamma-k}{2\gamma}$	$\frac{(\gamma-k)\gamma}{2} C_x^2 - k C_{yx}$
2	1	0	1	$t_2^* = \bar{y} \left[\alpha \left(\frac{\bar{X}}{\bar{x}^*} \right) + (1-\alpha) \right]$	$-\frac{k}{\gamma}$	$-(\gamma k C_x^2 + k C_{yx})$
3	0	1	1	$t_3^* = \bar{y} \left[\alpha + (1-\alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right) \right]$	$\frac{\gamma-k}{\gamma}$	$-k C_{yx}$
4	1	1	-1	$t_4^* = \bar{y} / \left[\alpha \left(\frac{\bar{X}}{\bar{x}^*} \right) + (1-\alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right) \right]$	$\frac{\gamma+k}{2\gamma}$	$-\frac{(k+\gamma)\gamma}{2} C_x^2$
5	1	0	-1	$t_5^* = \bar{y} / \left[\alpha \left(\frac{\bar{X}}{\bar{x}^*} \right) + (1-\alpha) \right]$	$\frac{k}{\gamma}$	$-k \gamma C_x^2$
6	0	1	-1	$t_6^* = \bar{y} / \left[\alpha + (1-\alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right) \right]$	$\frac{\gamma+k}{\gamma}$	0
7	-	-	-	$t_7^* = \bar{y} \left(\frac{\bar{X}}{\bar{x}^*} \right)^\alpha$	$-\frac{k}{\gamma}$	$\frac{k(k-\gamma)}{2} C_x^2 - k C_{yx}$
8	-	-	-	$t_8^* = \bar{y} \left(\frac{\bar{x}^*}{\bar{X}} \right)^\alpha$	$\frac{k}{\gamma}$	$\frac{k(k-\gamma)}{2} C_x^2 - k C_{yx}$

Discussion:

- (i) All estimators belonging to the proposed class of estimators in (3.1) taking arbitrary values for g, h, δ and choosing α suitably so as to minimize the approximate mean square error of the class have optimum asymptotic mean square equal to that of the classical linear regression estimator.
- (ii) Among the considered special cases the class only

$t_6^* = \bar{y} / \left[\alpha + (1 - \alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right) \right]$ is almost unbiased, that is, having first order bias equal to zero and should be preferred over other members in the class under consideration including

$$t_1^* = \bar{y} \left[\alpha \left(\frac{\bar{X}}{\bar{x}^*} \right) + (1 - \alpha) \left(\frac{\bar{x}^*}{\bar{X}} \right) \right]$$

considered by Choudhury and Singh(2012).

(iii) Finally, the condition under which the first order bias of the proposed class of estimators vanishes is

$$\frac{\gamma(\delta\gamma h - k)}{g + h} - \frac{g(g + 1)}{2} - \frac{h(h - 1)}{2} + \frac{h(h - 1)}{2} \delta\gamma^2 = k^2 \frac{\delta + 1}{2\delta}$$

4.5. Numerical Illustration

In Table-2 the first order biases of $t_1^*, t_2^*, \dots, t_7^*, t_8^*$ have been computed for illustration for simple random sample of size $n = 20$, selected from Population -1(Kadilar and Cingi-2006b) and Population -2(Murthy-1967). The measured variables(y,x) for Population-1 are(Level of apple production,Number of apples) and for Population-2 are (Output,Fixed capital).The required population parameters for both the populations are given in Table-2.

Table 2. Comparison of biases of estimators

Population	N	C_y	C_x	ρ	$k = \frac{\rho C_y}{C_x}$	Absolute Bias excepting the multiplier $\theta\bar{Y}$
1	106	4.18	2.02	0.82	0.96	$t_1^* : 24.8857$ $t_2^* : 13.3584$ $t_3^* : 11.7480$ $t_4^* : 0.9156$ $t_5^* : 1.6104$ $t_6^* : 0$ $t_7^* : 6.6792$ $t_8^* : 6.6792$
2	80	0.354 2	0.75 07	0.941	0.4440	$t_1^* : 0.2430$ $t_2^* : 0.1945$ $t_3^* : 0.1111$ $t_4^* : 0.0730$ $t_5^* : 0.0834$ $t_6^* : 0$ $t_7^* : 0.0972$ $t_8^* : 0.0972$

Comments: The numerical illustrations for the populations under consideration show that t_6^* is least biased (being zero), followed by t_4^* and t_5^* .

5. PREDICTIVE RATIO AND PRODUCT ESTIMATORS

Basu(1971) proposed a model free approach of prediction where the unobserved observations of the population outside the observed part in the sample are predicted with the help of auxiliary information to compute the estimate of the population total or population mean of the study variable and was well supported by Smith(1976). Sampford(1978) showed that the classical ratio estimator is model free predictor having the property of internal congruency. Basu's predictive approach is different from the approach of Royall(1970,1971) where the distributional form or an assumed model is sought for to provide a link between the observed part and unobserved part of the population.

Now, write the population total $Y = \sum_{i=1}^n y_i + \sum_{i=n+1}^N y_i$ where the

first part is observed and the second part is not observed. To estimate Y , we have to predict the second part which is unobserved. Thus, we write

$$\hat{Y} = \sum_{i=1}^n y_i + \sum_{i=n+1}^N \hat{Y}_i,$$

where \hat{Y}_i is the predictor of $y_i (i = n+1, \dots, N)$

5.1 Predictive Ratio type Estimators

Case I:

It may be reasonable to predict $y_i (i = n+1, \dots, N)$ by $\hat{Y}_i = \frac{\bar{y}}{\bar{x}} x_i (i = n+1, \dots, N)$, where $\frac{\bar{y}}{\bar{x}}$ acts as a slope of the regression line of y on x passing through the origin.

The estimator of Y thus is

$$\hat{Y}_R = \sum_{i=1}^n y_i + \frac{\bar{y}}{\bar{x}} \sum_{i=n+1}^N x_i = N \frac{\bar{y}}{\bar{x}} \bar{X}$$

and the predictive estimator of \bar{Y} is $\bar{y}_R = \frac{\bar{y}}{\bar{x}} \bar{X}$, which is the classical ratio estimator of the population mean \bar{Y} .

According to Sampford(1978) the idea that any estimator taken as predictive representation of an estimator should be of identical form to the estimator itself is a simple and appealing one. Such estimators are called by him as 'internally congruent' and this property is seen to be satisfied in case of the ratio estimator discussed above.

Case II

Agrawal and Sthapit (1997) considered \bar{y}_R as an alternative intuitive predictor of $y_i (i = n+1, \dots, N)$ in the unobserved part of the population and their predictive estimator takes the form $\hat{\bar{Y}} = f\bar{y} + (1-f)\bar{y}_R = \bar{y}_R^{(1)}$, say, to $O(1/n)$,

$$Bias(\bar{y}_R^{(1)}) = \theta \bar{Y} (1-f)(C_x^2 - C_{yx}) = \theta \bar{Y} \lambda (C_x^2 - C_{yx}),$$

where $\lambda = (1-f)$

$$MSE(\bar{y}_R^{(1)}) = \theta \bar{Y}^2 (C_y^2 + \lambda^2 C_x^2 - 2\lambda C_{yx})$$

For second iteration $\bar{y}_R^{(1)}$ is taken as an intuitive predictor of y_i in unobserved part of the population. As a result we have

$$\bar{y}_R^{(2)} = (1-\lambda^2)\bar{y} + \lambda^2 \bar{y}_R$$

Proceeding further we have for the m th. order iteration $\bar{y}_R^{(m)} = (1-\lambda^m)\bar{y} + \lambda^m \bar{y}_R$. It may be noted that for

$m=0$, $\bar{y}_R^{(m)} = \bar{y}_R$ to $O(1/n)$ the bias and mean square error of $\bar{y}_R^{(m)}$ are

$$Bias(\bar{y}_R^{(m)}) = \lambda^m \frac{1-f}{n} \bar{Y} (C_x^2 - C_{yx})$$

$$MSE(\bar{y}_R^{(m)}) = \theta \bar{Y}^2 (C_y^2 + \lambda^{2m} C_x^2 - 2\lambda^m \rho C_y C_x).$$

m is determined optimally by minimizing $MSE(\bar{y}_R^{(m)})$ with respect to λ^m . Thus, $\lambda_{opt}^m = \frac{C_{yx}}{C_x^2} = \rho \frac{C_y}{C_x} =$

k , say and the optimum mean square or variance of $\bar{y}_R^{(m)}$ is

$$MSE(\bar{y}_R^{(m)})_{opt} = \frac{1-f}{n} \bar{Y}^2 C_y^2 (1-\rho^2)$$

With optimum λ^m , the optimum bias of $\bar{y}_R^{(m)}$ reduces to

$$Bias(\bar{y}_R^{(m)})_{opt} = \frac{1-f}{n} \bar{Y} k (1-k) C_x^2$$

$\bar{y}_R^{(m)}$ will be more efficient than \bar{y}_R if

$$\rho \frac{C_y}{C_x} < \frac{1}{2} (1 + \lambda^m)$$

Again $\bar{y}_R^{(m)}$ will be more efficient than \bar{y} if

$$\rho \frac{C_y}{C_x} > \frac{1}{2} \lambda^m$$

Thus, $\bar{y}_R^{(m)}$ will be more efficient than both \bar{y}_R and \bar{y} if

$$\frac{1}{2} \lambda^m < \rho \frac{C_y}{C_x} < \frac{1}{2} (1 + \lambda^m)$$

These results are due to Agrawal and Sthapit(1997).

Case III

If we predict y_i in the unobserved part as $\hat{Y}_i = \frac{\bar{y}}{\bar{X}} x_i$

$$\bar{y}_R^* = f\bar{y} + \frac{\bar{y}}{\bar{X}} \left(\frac{N\bar{X} - n\bar{x}}{N} \right) = f\bar{y} + \frac{\bar{y}}{\bar{X}} (\bar{X} - f\bar{x})$$

The bias and mean square error of \bar{y}_R^* are found as

$$Bias(\bar{y}_R^*) = -\theta \bar{Y} f \rho C_y C_x$$

$$MSE(\bar{y}_R^*) = \theta \bar{Y}^2 (C_y^2 + f^2 C_x^2 - 2f C_{yx})$$

Case IV

If we predict y_i in the unobserved part as $\hat{Y}_i = \bar{r} x_i$ we have a predictive estimator of \bar{Y} as

$$\bar{y}_{HR}^{**} = f\bar{y} + \bar{r} (\bar{X} - f\bar{x})$$

To $O(1/n)$,

$$MSE(\bar{y}_{HR}^{**}) = \theta f^2 (S_y^2 + \bar{R}^2 S_x^2 - 2\bar{R} \rho S_y S_x)$$

It may be seen that \bar{y}_{HR}^{**} is more efficient than \bar{y}_{HR} as $f^2 < 1$.

5.2 Predictive Product type Estimators

When the study variable y is negatively correlated with the auxiliary variable x , the usual estimator is the product estimator \bar{y}_P (Murthy, 1964). Taking \bar{y}_P as the predictor of y_i ($i = n+1, \dots, N$), Agrawal and Sthapit (1997) gave the product based predictive estimator of \bar{Y} as

$$\hat{\bar{Y}} = f\bar{y} + (1-f)\bar{y}_P = \bar{y}_P^{(1)}$$

$$Bias(\bar{y}_P^{(1)}) = \theta\bar{Y}(1-f)\rho C_y C_x = \theta\bar{Y}\lambda\rho C_y C_x$$

$$MSE(\bar{y}_P^{(1)}) = \theta\bar{Y}^2(C_y^2 + \lambda^2 C_x^2 + 2\lambda C_{yx})$$

Further, m order iterated predictive product based estimator of \bar{Y} is written as

$$\bar{y}_P^{(m)} = (1-\lambda^m)\bar{y} + \lambda^m\bar{y}_P$$

$$Bias(\bar{y}_P^{(m)}) = \lambda^m Bias(\bar{y}_P)$$

$$MSE(\bar{y}_P^{(m)}) = \frac{1-f}{n}\bar{Y}^2(C_y^2 + \lambda^{2m}C_x^2 + 2\lambda^m C_{yx})$$

Minimizing $MSE(\bar{y}_P^{(m)})$ with respect to λ^m , the optimum value of m is obtained from $\lambda^m = -\rho \frac{C_y}{C_x}$.

Thus,

$$Bias(\bar{y}_P^{(m)})_{opt} = -k Bias(\bar{y}_P)$$

$$MSE(\bar{y}_P^{(m)})_{opt} = \theta\bar{Y}^2 C_y^2 (1-\rho^2)$$

$\bar{y}_P^{(m)}$ will be preferred to

$$(i) \bar{y}_P \text{ if } \rho \frac{C_y}{C_x} > -\frac{1}{2}(1+\lambda^m)$$

$$(ii) \bar{y} \text{ if } \rho \frac{C_y}{C_x} < -\frac{1}{2}\lambda^m$$

$$(iii) \text{ both } \bar{y}_P \text{ and } \bar{y} \text{ if } -\frac{1}{2}(1+\lambda^m) < \rho \frac{C_y}{C_x} < -\frac{1}{2}\lambda^m$$

These results are due to Agrawal and Sthapit (1997)

Case II

Following predictive approach of Basu (1971), Srivastava (1983) suggested prediction of y_i in the unobserved part as

$$\hat{Y}_i = \frac{\bar{y}\bar{x}}{(N\bar{X} - n\bar{x}) / (N - n)}$$

Thus, his estimator takes the form

$$\bar{y}_P' = \frac{n}{N}\bar{y} + \frac{N-n}{N} \frac{\bar{y}\bar{x}}{(N\bar{X} - n\bar{x}) / (N - n)} = \frac{n}{N}\bar{y} + \frac{(N-n)^2}{N} \frac{\bar{y}\bar{x}}{(N\bar{X} - n\bar{x})}$$

The bias and mean square error of \bar{y}_P' are found as

$$Bias(\bar{y}_P') = \theta\bar{Y}(\rho C_y C_x + \frac{n}{N-n} C_x^2)$$

$$MSE(\bar{y}_P') = \theta \bar{Y}^2 (C_y^2 + C_x^2 + 2C_{yx})$$

If we predict y_i in the unobserved part as $\hat{Y}_i = \frac{\bar{y}\bar{x}}{(N\bar{X} - n\bar{x})^2 / (N - n)^2} x_i$, the resultant estimator is also the same as \bar{y}_P' .

Case III

Further, if we predict y_i in the unobserved part as $\hat{Y}_i = \bar{y} \frac{\bar{x}}{\bar{X}^2} x_i$, the resultant estimator is

$$\bar{y}_{P1}' = \frac{n}{N} \bar{y} + \bar{y} \frac{\bar{x}}{\bar{X}^2} \left(\frac{N\bar{X} - n\bar{x}}{N} \right)$$

The bias and mean square of \bar{y}_{P1}' are found as

$$Bias(\bar{y}_{P1}') = \theta \bar{Y} (1 - 2f) \rho C_y C_x - f C_x^2$$

$$MSE(\bar{y}_{P1}') = \theta \bar{Y}^2 C_y^2 + (1 - 2f)^2 C_x^2 + 2(1 - 2f) \rho C_y C_x$$

5.3 Generalized Predictive Ratio type Estimator

Case I

The predictor \hat{Y}_i for y_i in the unobserved part of the population is formed as

$$\hat{Y}_i = \left[\frac{\bar{y}}{\alpha \bar{x} + (1 - \alpha) \bar{X}} \right] x_i$$

Thus we write the estimator of population mean \bar{Y} as,

$$\hat{\bar{Y}}_{GPR} = f\bar{y} + \frac{\bar{y}(\bar{X} - f\bar{x})}{\alpha \bar{x} + (1 - \alpha) \bar{X}},$$

where α is chosen so as to minimize the $MSE(\hat{\bar{Y}}_{GPR})$.

5.3.1 Bias and Mean square Error of $\hat{\bar{Y}}_{GPR}$

$$\bar{y} = \bar{Y}(1 + e_0), \quad \bar{x} = \bar{X}(1 + e_1) \text{ with } E(e_0) = E(e_1) = 0,$$

$$V(e_0) = \theta C_y^2, V(e_1) = \theta C_x^2, Cov(e_0, e_1) = \theta \rho C_y C_x;$$

We may now expand $\hat{\bar{Y}}_{GPR}$ in Taylor's series and keeping second degree terms in the expansion we have $\hat{\bar{Y}}_{GPR} \cong \bar{Y}(1 - \alpha e_2 - f e_2 + f \alpha e_2 + e_1 - \alpha e_1 e_2 + f \alpha e_1 e_2 - f e_1 e_2 + \alpha^2 e_2^2 - f \alpha^2 e_2^2 + f \alpha e_2^2)$. Taking expectation of both the left and right hand sides we have to terms of order $O(1/n)$

$$E(\hat{\bar{Y}}_{GPR}) = \bar{Y} + \theta \bar{Y} \left[(\alpha^2 - f \alpha^2 + f \alpha) C_x^2 - (\alpha - f \alpha + f) C_{yx} \right]$$

Further to $O(1/n)$,

$$MSE(\hat{\bar{Y}}_{GPR}) = \theta \bar{Y}^2 \left[C_y^2 + (\alpha + f - f \alpha)^2 C_x^2 - 2(\alpha + f - f \alpha) C_{yx} \right]$$

Minimizing (2.5) with respect to α gives $\alpha_{opt} = \frac{k - f}{1 - f}$, Substituting the optimum value of α in the expression for

mean square error we have $MSE(\hat{\bar{Y}}_{GPR})_{opt} = \theta \bar{Y}^2 C_y^2 (1 - \rho^2)$, which is the asymptotic mean square error of linear

regression estimator . Substituting the optimum value of α in the approximate expression for bias to $O(1/n)$, we have

$$Bias(\hat{Y}_{GPR})_{opt} = \theta \bar{Y} \frac{f}{1-f} k(k-1) C_x^2,$$

Case II

Alternatively, we may predict y_i in the unobserved part by

$$\hat{Y}_i = \frac{\bar{y}\bar{X}}{\alpha\bar{x} + (1-\alpha)\bar{X}},$$

where α is a suitably chosen constant obtained by minimizing the approximate mean square of predictive estimator of \bar{Y} given by

$$\hat{Y}_{GPR}^* = f\bar{y} + (1-f) \frac{\bar{y}\bar{X}}{\alpha\bar{x} + (1-\alpha)\bar{X}}$$

Minimizing the mean square of \hat{Y}_{GPR}^* to $O(1/n)$ with respect to α gives $\alpha_{opt} = \frac{k}{1-f}$

Bias and mean square error of \hat{Y}_{GPR}^* with optimum α to $O(1/n)$ are found as

$$Bias(\hat{Y}_{GPR}^*) = \theta \bar{Y} \left(\frac{f}{1-f} \right) k^2 C_x^2, \quad MSE(\hat{Y}_{GPR}^*) = \theta \bar{Y}^2 C_y^2 (1-\rho^2)$$

5.4 Generalized Predictive Product Type Estimator

The suggested predictive product type estimator is suggested as

$$\hat{Y}_{GPP}^* = f\bar{y} + (1-f) \frac{\bar{y}}{\bar{X}} [\alpha\bar{x} + (1-\alpha)\bar{X}]$$

Table 3. Biases and Mean Square Errors of Modified Ratio Type Estimators

Estimator	Bias	Mean Square Error
$\bar{y}_R = \bar{y}(\bar{X} / \bar{x})$	$\theta \bar{Y} (1-k) C_x^2$	$\theta \bar{Y}^2 (C_y^2 + C_x^2 - 2\rho C_y C_x)$
$\hat{Y}_{DR} = \bar{y}(\bar{x}^* / \bar{X})$	$-\theta \bar{Y} \gamma k C_x^2$	$\theta \bar{Y}^2 (C_y^2 + \gamma^2 C_x^2 - 2\gamma \rho C_y C_x)$
$\bar{y}_R^{(1)} = f\bar{y} + (1-f)\bar{y}_R$	$\theta \bar{Y} (1-f)(1-k)$	$\theta \bar{Y}^2 (C_y^2 + \lambda^2 C_x^2 - 2\lambda \rho C_y C_x)$
$\bar{y}_R^* = f\bar{y} + \frac{\bar{y}}{\bar{X}} (\bar{X} - f\bar{x})$	$-\theta \bar{Y} f k C_x^2$	$\theta \bar{Y}^2 (C_y^2 + f^2 C_x^2 - 2f \rho C_y C_x)$
$\hat{Y}_{GPR} = f\bar{y} + \frac{\bar{y}(\bar{X} - f\bar{x})}{\alpha\bar{x} + (1-\alpha)\bar{X}}$	$\theta \bar{Y} \gamma k (k-1) C_x^2$	$\theta \bar{Y}^2 C_y^2 (1-\rho^2)$
$\hat{Y}_{GPR}^* = f\bar{y} + (1-f) \frac{\bar{y}\bar{X}}{\alpha\bar{x} + (1-\alpha)\bar{X}}$	$\theta \bar{Y} \gamma k^2 C_x^2$	$\theta \bar{Y}^2 C_y^2 (1-\rho^2)$

Linearizing \hat{Y}_{GPP}^* we have

$$\hat{Y}_{GPP}^* = f\bar{y} + (1-f)\bar{y} \left[\frac{\alpha\bar{X}(1+e_1) + (1-\alpha)\bar{X}}{\bar{X}} \right]$$

$$= \bar{y} + \bar{y}(1-f)\alpha e_1 = \bar{Y}(1+e_0) + \bar{Y}(1+e_0)(1-f)\alpha e_1$$

To $O(1/n)$,

$$MSE(\hat{Y}_{GPP}^*) = \theta\bar{Y}^2(C_y^2 + (1-f)^2\alpha^2C_x^2 + 2(1-f)\alpha C_{yx})$$

Minimization gives

$$\alpha_{opt} = -\frac{k}{1-f}$$

Thus,

$$Bias(\hat{Y}_{GPP}^*)_{opt} = -\theta\bar{Y}k^2C_x^2, \quad MSE(\hat{Y}_{GPP}^*)_{opt} = \theta\bar{Y}^2C_y^2(1-\rho^2)$$

Comments:

- (i) \hat{Y}_{DR} will be more efficient than both \bar{y} and \bar{y}_R if $\frac{1}{2}\gamma < k < \frac{1}{2}(1+\gamma)$
- (ii) $\bar{y}_R^{(1)}$ will be more efficient than both \bar{y} and \bar{y}_R if $\frac{1}{2}\lambda < k < \frac{1}{2}(1+\lambda)$
- (iii) \bar{y}_R^* will be more efficient than both \bar{y} and \bar{y}_R if $\frac{1}{2}f < k < \frac{1}{2}(1+f)$
- (iv) $\bar{y}_R^{(1)}$ will be more efficient than \hat{Y}_{DR} if $k > \frac{1}{2}(\lambda + \gamma)$ provided $\lambda > \gamma$.
- (v) $\bar{y}_R^{(1)}$ will be more efficient than \bar{y}_R^* if $k > \frac{1}{2}(\lambda + f)$ provided $\lambda > f$ i.e. $f < \frac{1}{2}$.

Table 4. Comparison of Modified Ratio Type Estimators

Estimator	Preferred to	Condition
\hat{Y}_{DR}	\bar{y}	$k > \frac{1}{2}\gamma$
\hat{Y}_{DR}	\bar{y}_R	$k < \frac{1}{2}(1+\gamma)$
$\bar{y}_R^{(1)}$	\bar{y}	$k > \frac{1}{2}\lambda$
$\bar{y}_R^{(1)}$	\bar{y}_R	$k < \frac{1}{2}(1+\lambda)$
\bar{y}_R^*	\bar{y}	$k > \frac{1}{2}f$
\bar{y}_R^*	\bar{y}_R	$k < \frac{1}{2}(1+f)$

Table 5. Biases and Mean Square Errors of Modified Product type Estimators

Estimator	Bias	Mean Square Error
$\bar{y}_P = \bar{y}(\frac{\bar{x}}{\bar{X}})$	$\theta\bar{Y}kC_x^2$	$\theta\bar{Y}^2(C_y^2 + C_x^2 + 2\rho C_y C_x)$
$\hat{\bar{Y}}_{DP} = \bar{y}(\frac{\bar{X}}{\bar{x}^*})$	$\theta\bar{Y}\gamma(k+1)C_x^2$	$\theta\bar{Y}^2(C_y^2 + \gamma^2 C_x^2 + 2\gamma\rho C_y C_x)$
$\bar{y}_P^{(1)} = f\bar{y} + (1-f)$	$\theta\bar{Y}\lambda kC_x^2$	$\theta\bar{Y}^2(C_y^2 + \lambda^2 C_x^2 + 2\lambda\rho C_y C_x)$
$\bar{y}_P' = f\bar{y} + (1-f)^2$	$\theta\bar{Y}(\gamma+k)C_x^2$	$\theta\bar{Y}^2(C_y^2 + C_x^2 + 2\rho C_y C_x)$
$\bar{y}_{P1}' = f\bar{y} + \bar{y}(\frac{\bar{x}}{\bar{X}^2})$	$\theta\bar{Y}(\lambda-f)k - j$	$\theta\bar{Y}^2(C_y^2 + (\lambda-f)^2 C_x^2 + 2(\lambda-f)\rho C_y C_x)$

Table 6. Comparison of Modified Product Type Estimators

Estimator	Preferred to	Condition
\bar{y}_P	\bar{y}	$k < -1/2$
$\hat{\bar{Y}}_{DP}$	\bar{y}	$k < -(1/2)\gamma$
$\bar{y}_P^{(1)}$	\bar{y}	$k < -(1/2)\lambda$
\bar{y}_P'	\bar{y}	$k < -(1/2)$
\bar{y}_{P1}'	\bar{y}	$k < -(1/2)(\lambda-f)$
$\hat{\bar{Y}}_{DP}$	\bar{y}_P	$k > -\frac{1}{2}(1+\gamma)$
$\bar{y}_P^{(1)}$	\bar{y}_P	$k > -\frac{1}{2}(1+\lambda)$
\bar{y}_P'	\bar{y}_P	equally efficient with \bar{y}_P
\bar{y}_{P1}'	\bar{y}_P	$k > -\lambda$
\bar{y}_{P1}'	\bar{y}_{DP}	$k > -\frac{1}{2}(\gamma+\lambda)$

Example 1. The data taken from Cochran(1977) used by Agrawal and Sthapit(1997) , relate to complete enumeration of 256 peach orchards where y=estimated peach production in an orchard and x= number of peach trees in the orchard. The summary computations are

$$N = 256, \bar{Y} = 56.47, \bar{X} = 44.45, S_y = 80.06, S_x = 62.46,$$

$$C_y = 1.42, C_x = 1.41, \rho = 0.887, k = \rho(C_y / C_x) = 0.8933$$

Table 7. Biases and Mean Square Errors of Modified Ratio type Estimators $n=16$

Estimator	Absolute Bias	Mean Square Error(MSE)	Relative Bias(RB)	Relative Efficiency(RE)
\bar{y}	0	376.7348	0	100
\bar{y}_R	0.701849	84.5543	0.0763	445.55
$\hat{\bar{Y}}_{DR}$	0.391748	334.1418	0.0214	112.75
$\bar{y}_R^{(1)}$	0.657984	81.0510	0.0731	464.80
\bar{y}_R^*	0.367246	336.7091	0.0200	111.89
$\hat{\bar{Y}}_{GPR}$	0.41800	80.3315	0.0046	468.98
$\hat{\bar{Y}}_{GPR}^*$	0.349949	80.3315	0.0390	468.98

Table 8. Biases and Mean Square Errors of Modified Ratio type Estimators $n=32$

Estimator	Absolute Bias	Mean Square Error(MSE)	Relative Bias(RB)	Relative Efficiency(RE)
\bar{y}	0	175.7967	0	100
\bar{y}_R	0.327506	39.4558	0.0521	445.55
$\hat{\bar{Y}}_{DR}$	0.391708	135.0946	0.0337	130.13
$\bar{y}_R^{(1)}$	0.286568	37.5405	0.0468	468.29
\bar{y}_R^*	0.342738	147.6427	0.0283	119.07
$\hat{\bar{Y}}_{GPR}$	0.041795	37.4853	0.006826	468.98
$\hat{\bar{Y}}_{GPR}^*$	0.349913	37.4853	0.057152	468.98

Example 2 : The following computations relate to data collected as to the duration of sleep (in minutes) and age of 30 people aged 50 years and over living in a small village of the United States of America(Singh,Sarjinder(2003).

y: Duration of sleep in minutes)

x: Age of subjects (greater or equal to 50 years)

$$N=30, \quad \bar{Y}=384.2, \quad \bar{X}=67.267, \quad C_y^2=0.0243, \quad C_x^2=0.0188, \quad \rho=-0.8552, \quad k=-0.9723$$

Table 9. Biases and Mean Square Errors of Modified Product type Estimators $n=5$

Estimator	Absolute Bias	Mean Square Error(MSE)	Relative Bias(RB)	Relative Efficiency(RE)
\bar{y}	0	478.1357	0	100
\bar{y}_P	0.936132	128.7131	0.0825	371.47
$\hat{\bar{Y}}_{DP}$	0.006668	321.4210	0.00037	148.76
$\bar{y}_P^{(1)}$	0.748920	139.4112	0.063429	342.97
\bar{y}_P'	0.696408	128.7131	0.0614	371.47
\bar{y}_{P1}'	0.754254	179.7026	0.0638	266.04
$\hat{\bar{Y}}_{GPP}^*$	0.910219	128.4430	0.08031	372.26

Table 10. Biases and Mean square Errors of Modified Product type Estimators

<i>n=10</i>				
Estimator	Absolute Bias	Mean Square Error(MSE)	Relative Bias(RB)	Percent Relative Efficiency(RE)
\bar{y}	0	239.1395	0	100
\bar{y}_P	0.468207	64.3762	0.0584	371.47
$\hat{\bar{Y}}_{DP}$	0.006668	105.5228	0.00065	226.62
$\bar{y}_P^{(1)}$	0.312159	81.5153	0.0346	293.37
\bar{y}_P'	0.227486	64.3762	0.0284	371.47
\bar{y}_{P1}'	0.252414	139.7635	0.0214	171.10
$\hat{\bar{Y}}_{GPP}^*$	0.455246	64.2409	0.056799	372.25

RECEIVED MAY, 2012
REVISED NOVEMBER, 2012

REFERENCES

- [1] AGRAWAL , M.C. and STHAPIT, A.B. (1997):Hierarchic predictive ratio-based and product-based estimators and their efficiency. **Journal of Applied Statistics**, 24, 1,97-104.
- [2] BASU,D. (1971):An essay on the logical foundations of survey sampling. In: V P Godambe and D.A. Sprott (Eds) , **Foundations of Statistical Inference** , 203-33 , Holt,Rinehart and Winston, Toronto.
- [3] BANDOPADHYAYA,S.(1980):Improved ratio and product estimators. **Sankhya**, 42,45-49.
- [4] CHOUDHURY,S. and SINGH,B.K.(2012):An efficient class of dual to product-cum-dual to ratio estimators of finite population mean in sample surveys. **Global Journal of Science Frontier of Research Mathematics &Decision Sciences**, 12, 25-33.
- [5] COCHRAN,W.G.(1940):The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, **The Journal of Agricultural Science**, 30, 262-275.
- [6] COCHRAN, W.G.(1977): **Sampling Techniques**. John Wiley, New York.
- [7] HARTLEY,H.O. and ROSS,A. (1954): Unbiased ratio estimators. **Nature**, 174, 270-271.
- [8] KADILAR,C and CINGI,H.(2006a): Improvement in estimating the population mean in simple random sampling. **Applied Mathematics Letters**, 19, 75-79.
- [9] KADILAR,C. and CINGI,H.(2006b):Ratio estimators for the population variance in simple and stratified random sampling. **Applied Mathematics and Computations**, 173, 1047-1059.
- [10] MISHRA, G. and SAHOO, L.N. (1986): A Class of transformed ratio and product estimators in sample surveys. **Journ. Stat. Research**, 20, 81-85.
- [11] MOHANTY, S. and DAS, M.N. (1971): Use of transformation in sampling. **Journ. Ind. Soc. Ag. Stat.**, 23, 83-87.
- [12] MOHANTY, S. and SAHOO, J. (1995): A note on improving the ratio method of estimation through linear transformation using certain known population parameters. **Sankhya B**, 57, 93-102.
- [13] MURTHY, M.N. (1964): Product method of estimation. **Sankhya A**, 26, 69-74.
- [14] MURTHY,M.N. (1967):**Sampling Theory and Methods,Statistical** .Publishing Society, Calcutta.
- [15] PANDEY, B.N. and DUBEY, V. (1988): Modified product estimator using coefficient of variation of auxiliary variate. **Assam Stat. Review**, 2, 64-66.

- [16] PRABHU-AJGAONKAR,S.G.(1993):Non-existence of an optimum estimator in a class of ratio estimators. **Statistical Papers**, 34, 161-165.
- [17] REDDY, V.N. (1973): Ratio and product method of estimation. **Sankhya, B**, 35, 307-316.
- [18] REDDY, V.N.(1978) A study on the use of prior knowledge on certain population parameters in estimation. **Sankhya C**, 1, 29-37.
- [19] REDDY, V.N. (1974): On a transformed ratio method of estimation. **Sankhya,C**, 36, 59-70.
- [20] ROBSON,D.S.(1957): Applicationf multivariate polykays to the theory of unbiased ratio type estimation. **J.Amer.stat. Assoc.**,52, 411-422.
- [21] ROYALL, R.M.(1970):On finite population sampling theory under certain linear regression models. **Biometrika**, 57, 377-387.
- [22] ROYALL, R.M.(1971): Linear regression models in finite populations sampling theory. **Foundations of Statistical Inference**(V P Godambe and D.A. Sprott eds.), 259-274.Holt, Rinehart&Winston , Toronto.
- [23] SEN, A.R. (1978): Estimation of population mean when the coefficient of variation is known. **Comm. Statist: Theory and Methods**, A(7), 657-672.
- [24] SAMPFORD,M.R.(1978):Predictive estimation and Internal Congruency , **Contributions to survey sampling and Applied statistics**, 29-39. Academic Press, Inc., New York,
- [25] SINGH ,H.P. and KAKRAN,M.S.(1993): A modified ratio estimator using coefficient of kurtosis of an auxiliary character(**unpublished manuscript**)
- [26] SINGH,G.N.(2003): On the improvement of product method of estimation in sample surveys. **JISAS**, 56, 267-275.
- [27] SINGH, H.P. and TAILOR, R. (2003): Use of known correlation coefficient in estimating the finite population mean. **Statistics in Transition**, 6, 555-560.
- [28] SINGH, H.P., TAILOR, R. and KAKRAN, M.S. (2004): An estimator of population mean using power transformation. **Journ. Ind. Soc. Ag. Stat.**, 58, 223-230.
- [29] SINGH,H.P.,TAILOR,R., and TAILOR, R. (2010): On ratio and product methods with certain known population parameters of auxiliary variable in sample surveys. **Statistics and Operations Research Trans.**, 34, 157-180.
- [30] SINGH, H.P., TAILOR, R., SINGH, S. and KIM, JONG-MIN (2008): A modified estimator of population means using power transformation. **Statistical papers**, 49, 37-58.
- [31] SINGH,S.(2003):**Advanced sampling theory with applications**,vol. 1,Kluwer Academic Publishers, The Netherlands.
- [32] SISODIA, B.V.S. and DIWEDI, V.K. (1981): A modified ratio estimator using coefficient of auxiliary variable. **Journ. Ind. Soc. Ag. Stat.**, 33, 2, 13-18.
- [33] SMITH, T.M.F.(1976):The foundations of survey sampling : a review. **JRSS A**, 139, 183-204.
- [34] SRIVASTAVA, S.K. (1967): An estimator using auxiliary information in sample surveys. **Cal. Stat. Assoc. Bull.**, 16, 121-132.
- [35] SRIVASTAVA, S.K.(1983): Predictive estimation of finite population mean using product estimator. **Metrika**, 30, 93-99.
- [36] SRIVASTAVA,S.K. ,JHAJJ,H.S. and SHARMA,M.K.(1986): Comparison of some estimators of K and B in finite populations. **Journ. Ind. Soc. Ag. Statistics**, XXX, 230-236.
- [37] SRIVENKATARAMANA, T. (1978): Change of Origin and scale in ratio and difference methods of estimation in sampling. **Canadian Journ. Stat.**, 6, 1, 79-86.

- [38] SRIVENKATARAMANA, T. (1980): A dual to ratio estimator in sample surveys. **Biometrika**, 67, 199-204.
- [39] SUKHATME, P.V. , SUKHATME, B.V. and ASOK, C. (1984): **Sampling theory of surveys with applications**. Ind. Soc. Ag. Stat., New Delhi.
- [40] SUKHATME , P.V. and SUKHATME, B.V.(1970): **Sampling theory of surveys with Applications**. Asia Publishing House, New Delhi.
- [41] SWAIN, A.K.P.C.(1981): On the estimation of Population Ratio. **Comunicaciones en estadística y computo , Colegio de Post graduados** , Centro de Estadística y Calculo, Chapingo, Mexico. 1 (10).
- [42] TAILOR, R. and SHARMA, B. (2009): A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of Kurtosis. **Statistics in Transition**, 10, 15-24.
- [43] UPADHYAYA, L.N. and SINGH, H.P. (1999): Use of transformed auxiliary variable in estimating the finite population mean. **Biometrical Journal**, 41, 627-636.
- [44] WALSH, J.E. (1970): Generalization of ratio estimator for population total. **Sankhya A**, 32, 99-106.
- [45] YATES, F.(1960): **Sampling methods in censuses and surveys**. Charles Griffin and Co., London: .