

BIBLIOTECA DI SCIENZE STATISTICHE

SERVIZIO BIBLIOTECARIO NAZIONALE

BID PVV0850391 DID

ACQ. 483 / '02 INV. 82183

COLL. _____ GLASS 5-Coll. W/P. 2002/8

**Estimation for finite
population parameters in
two phase sampling**

G. Diana, C. Tommasi

2002.8

UNIVERSITA' DI PADOVA
BIBLIOTECA DI SCIENZE STATISTICHE
Via C. Battisti, 241 - 35121 PADOVA

**Dipartimento di Scienze Statistiche
Università degli Studi
Via C. Battisti 241-243
35121 Padova**

luglio 2002

BIBLIOTECA DI SCIENZE STATISTICHE

SEZIONE DI STATISTICA

NUMERO 1000

ANNO 1883

82/83

CLASS. 2-654/W/9 200/3

Estimation for finite
population parameters in
two phase sampling

G. CHAN, C. TOWNSEND

200/3

UNIVERSITA' DI PADOVA
BIBLIOTECA DI SCIENZE STATISTICHE
VIA E. BELTRAMI, 241 - 35131 PADOVA

Dipartimento di Scienze Statistiche
Università degli Studi
Via C. Battisti, 241
35131 Padova

Indice 2003

Estimation for finite population parameters in two phase sampling

Giancarlo Diana

Chiara Tommasi

Riassunto

In questo lavoro si propone un approccio generale per la stima di un generico parametro di una popolazione finita, nell'ambito del campionamento doppio. I molti stimatori per la media e per il rapporto di due medie presenti in letteratura, così come le poche proposte per la stima della varianza, possono essere visti come casi particolari di una classe generale di stimatori. Si fornisce, al primo ordine di approssimazione, il limite inferiore della varianza degli stimatori in detta classe e si mostra come esista almeno uno stimatore di tipo regressione a catena che raggiunge tale limite e che è perciò ottimale nella classe.

1 Introduction

The estimation of a parameter of interest when at least one auxiliary variable is available, was widely discussed. It is well known that if there are not information about the population characteristics of the auxiliary variable then a double sampling scheme may be used. Several estimators were proposed to estimate the population mean and the ratio of two means, while only few papers concern the population variance. For instance, the population mean estimation is treated, among others, by Srivastava *et al.* (1990), Singh and Singh (1991), Upadhyaya *et al.* (1992), Tracy and Singh (1999) and Singh and Espejo (2000). While Chand (1975), Srivastava *et al.* (1988,1989), Khare (1991), Singh *et al.* (1994), Singh and Singh (1994), Prasad *et al.* (1996) are some of the authors who deal with the estimation of ratio of two means. On the other hand, only Singh (1991) and Gupta *et al.* (1992) work about the population variance in a two-phase sampling.

In this paper we propose a general approach for estimating a finite population parameter in double sampling. Specifically, in Section 2 we give a general

class of estimators which includes all the previously quoted estimators. Furthermore, we provide the minimum attainable MSE for the estimators in this class, at the first order of approximation. We call it the minimum variance bound, since MSE and variance are the same at this order of approximation. We provide also a regression type estimator which reaches this minimum and we call it the "best" estimator in the class, even if any other estimator which is equivalent to the best estimator at the first order of approximation is optimum as well. All the results of Section 2 are valid whatever the parameter of interest and the used sampling scheme. The other sections specify such general results to the usually considered parameters of interest, i.e. the population mean, the ratio or product of two means and the population variance. These subjects, which are discussed exhaustively in Diana and Tommasi (2001, 2002a, 2002b), are here unified. In fact, in Section 4 and 5 we show that to estimate the ratio or product of two means and the variance is equivalent to estimate a mean, at least approximatively. Therefore, a central role is played by the mean case, which is discussed in Section 3, then the other parameter estimators follow from that one. Finally in Section 6 we show that some proposed estimators may be not optimal in the general class. Anyway, some examples of estimators which are optimal are also given.

2 A general class of estimators

Let $\mathcal{U} = \{1, \dots, i, \dots, N\}$ be a finite population and U_i be the value of an observable variable U , for the i -th population unit. Then, the population mean and variance of U are $\bar{U} = \sum_{i=1}^N U_i/N$ and $S_U^2 = V(U) = \sum_{i=1}^N (U_i - \bar{U})^2/(N-1)$.

Denoting the study variable by Y or Y_j , $j = 0, 1$, the interest is usually in estimating one of the following population parameters, $\bar{Y}$, $R = \bar{Y}_0/\bar{Y}_1$, $P = \bar{Y}_0\bar{Y}_1$ and S_Y^2 . Let T_Y be any one of such population parameters. Assuming that two auxiliary variables, X and Z , are related with Y or Y_j , $j = 0, 1$, but no information about the population parameters of X are available, then estimation of T_Y can be based on double sampling. Thus, we assume that a preliminary large sample of n' ($n' < N$) units is drawn by some sampling scheme. At this phase only X and Z are measured. As a second step, a sub-sample of size n ($n < n'$) is drawn from the first one. Without loss of generality, at this phase all the variables Y or Y_j , $j = 0, 1$, X and Z are observed.

If T_X and T_Z are $k \times 1$ and $l \times 1$ vectors of population parameters of X and Z and t_X , t'_X and t_Z , t'_Z are the corresponding first and second phase unbiased

estimates, then the proposed class of estimators is defined as

$$\hat{T}_Y = g(t_Y, t^T), \quad (1)$$

where t_Y is an approximatively unbiased estimate of T_Y , $t = (t_X^T, t_Z^T, t_X^T, t_Z^T)^T$ and function g is such that

- a. $g : \mathcal{S} \rightarrow \mathbb{R}$ where $\mathcal{S} \in \mathbb{R}^{2(k+l)+1}$ is a convex and bounded set which contains the point $(T_Y, T^T)^T$, where $T = E(t) = (T_X^T, T_Z^T, T_X^T, T_Z^T)^T$.
- b. It is a continuous and bounded function in $\mathcal{S}$.
- c. Its first and second partial derivatives are continuous and bounded in $\mathcal{S}$.
- d. $g(t_Y, T^T) = t_Y$.

Let

$$g_0 = \frac{\partial g(t_Y, t^T)}{\partial t_Y} \Big|_{(t_Y, t^T) = (T_Y, T^T)}$$

be the partial derivative of g with respect to t_Y and

$$g_t = (g_X^T, g_Z^T, g_{X'}^T, g_{Z'}^T)^T = \frac{\partial g(t_Y, t^T)}{\partial t} \Big|_{(t_Y, t^T) = (T_Y, T^T)}$$

be the $2(k+l)$ vector of the partial derivatives of g with respect to each component of t . Specifically, g_X and $g_{X'}$ are the $k \times 1$ vectors of the partial derivatives of g with respect to the elements of t_X and t_X' , while g_Z and $g_{Z'}$ are $l \times 1$ vectors defined in the same way but with respect to t_Z and t_Z' .

From point (d), we have that $g(T_Y, T^T) = T_Y$ and $g_0 = 1$. Thus, expanding $\hat{T}_Y$ at the point (T_Y, T^T) in a second order Taylor's series we have

$$\begin{aligned} \hat{T}_Y &\cong t_Y + (t - T)^T g_t = t_Y + (t_X - T_X)^T g_X + (t_Z - T_Z)^T g_Z \\ &\quad + (t_X' - T_X')^T g_{X'} + (t_Z' - T_Z')^T g_{Z'}. \end{aligned} \quad (2)$$

Since the population parameters of X are unknown, we have to impose the following constraint, $g_X = -g_{X'}$. For computational convenience, it is useful to change parameterization and to rewrite equation (2) as

$$\hat{T}_Y \cong t_Y + (t_V - t_V')^T g_V + (t_Z' - T_Z)^T g_Z \quad (3)$$

where $t_V' = (t_X^T, t_Z^T)^T$ and $t_V = (t_X^T, t_Z^T)^T$ are the first and second phase sample estimates of $T_V = (T_X^T, T_Z^T)^T$, t_Z' is the first phase sample estimate

of T_Z , $g_V = (g_X^T, g_Z^T)^T$ and $\ddot{g}_Z = (g_Z + g_{Z'})$.

The first order approximation for $\text{MSE}(\hat{T}_Y)$ is given by

$$\text{MSE}(\hat{T}_Y) \cong S_{t_Y}^2 + g_V^T S_{t_Y t_V} g_V + \ddot{g}_Z^T S_{t_Z t'_Z} \ddot{g}_Z + 2 (g_V^T S_{t_Y t_V} + \ddot{g}_Z^T S_{t_Y t'_Z}), \quad (4)$$

where

$$S_{t_Y}^2 = E[(t_Y - T_Y)^2], \quad S_{t_Y t_V} = E[(t_Y - t'_Y)(t_Y - t'_Y)^T],$$

$$S_{t'_Z t'_Z} = E[(t'_Z - T_Z)(t'_Z - T_Z)^T], \quad S_{t_Y t'_Z} = E[(t_Y - T_Y)(t'_Z - T_Z)]$$

and

$$S_{t_Y t'_Z} = E[(t_Y - T_Y)(t'_Z - T_Z)].$$

Minimizing (4) with respect to g_V and $\ddot{g}_Z$ we obtain the optimum values

$$g_V^* = -S_{t_Y t_V}^{-1} S_{t_Y t_V} \quad \text{and} \quad \ddot{g}_Z^* = -S_{t'_Z t'_Z}^{-1} S_{t_Y t'_Z}, \quad (5)$$

and replacing them in (4) we get the minimum first order approximation of $\text{MSE}(\cdot)$, denoted by $\text{MSE}^*(\cdot)$. It is the minimum variance bound for all the estimators based on the auxiliary vector t and has a simple form,

$$\text{MSE}^*(\hat{T}_Y) = S_{t_Y}^2 (1 - \rho_{t_Y, t'_Z}^2 - \rho_{t_Y, t_V}^2), \quad (6)$$

where

$$\rho_{t_Y, U}^2 = \frac{S_{t_Y, U}^T S_{U, U}^{-1} S_{t_Y, U}}{S_{t_Y}^2}$$

denotes the multiple correlation coefficient between t_Y and any one random vector U .

Finally, replacing g_V^* and $\ddot{g}_Z^*$ in (3) we get an optimum estimator. We will treat such estimator in a special way since, for all the estimation problems that will be considered, it has a nice interpretation: it is a regression type estimator. We will call it "the best" estimator in the class g , but any other estimator which reaches the minimum variance bound (6) is optimum as well.

All the results of this section are valid whatever the used sampling scheme and the involved auxiliary statistics. The computations can be further developed if the sampling scheme is specified and all the sample quantities (both of interest and auxiliary) are sample means, at least at the first order of approximation. For this reason, the estimation of a finite population mean plays a central role. This case is treated in the next section and then the obtained results are used in Sections 4 and 5, where the estimation of ratio and product of two population means and the estimation of the population

variance are dealt with. It will be seen that the sample estimates of such parameters may be approximated by sample means, at least at the first order of approximation.

From now on, we assume that both the first and the second phase samples are drawn by a simple random sample without replacement (SRSWOR).

3 Estimation of a finite population mean

In this section we consider the estimation of $\bar{Y}$, the population mean of Y , thus $t_Y = \bar{y}$, the second phase sample mean of Y .

Following Diana and Tommasi (2001, 2002a, 2002b), henceforth two parameter vectors $T_X = (\bar{X}, S_X^2)^T$ and $T_Z = (\bar{Z}, S_Z^2)^T$ will be considered. Thus, denoting the first and second phase sample means and variances of any variable U , by $\bar{u}'$, $s_U'^2$ and $\bar{u}$, s_U^2 , the general class (1) is defined as

$$\bar{y}_g = g(\bar{y}, t^T), \quad (7)$$

where $t = (\bar{x}, s_X^2, \bar{z}, s_Z^2, \bar{x}', s_X'^2, \bar{z}', s_Z'^2)^T$ and equation (3) becomes

$$\bar{y}_g \cong \bar{y} + (t_V - t_V')^T g_V + (t_Z' - T_Z)^T \ddot{g}_Z, \quad (8)$$

where $t_V = (\bar{x}, s_X^2, \bar{z}, s_Z^2)^T$, $t_V' = (\bar{x}', s_X'^2, \bar{z}', s_Z'^2)^T$ and $t_Z' = (\bar{z}', s_Z'^2)^T$.

For a better interpretation of the results, it is convenient to define a random vector $V = (\tilde{X}^T, \tilde{Z}^T)^T$ where, for any variable U , $\tilde{U} = (U, \delta_U)^T$ and $\delta_U = (U - \bar{U})^2$. In addition, let us denote the population covariance between any two variables U_1 and U_2 as $\text{Cov}(U_1, U_2) = \sum_{i=1}^N (U_{1i} - \bar{U}_1)(U_{2i} - \bar{U}_2) / (N - 1)$. Replacing in equation (4) the following expected values (which are valid up to terms of order $O(n^{-1})$)

$$S_{t_Y}^2 \cong \left(\frac{1}{n} - \frac{1}{N} \right) S_Y^2, \quad S_{t_Y t_Y} \cong \left(\frac{1}{n} - \frac{1}{n'} \right) S_{Y,V},$$

$$S_{t_Z' t_Z'} \cong \left(\frac{1}{n'} - \frac{1}{N} \right) S_{\tilde{Z}, \tilde{Z}}, \quad S_{t_Y t_Z'} \cong \left(\frac{1}{n} - \frac{1}{n'} \right) S_{Y,V}$$

and

$$S_{t_Y t_Z'} \cong \left(\frac{1}{n'} - \frac{1}{N} \right) S_{Y, \tilde{Z}},$$

where

$$S_{V,V} = \begin{bmatrix} S_{\tilde{X}, \tilde{X}} & S_{\tilde{X}, \tilde{Z}} \\ S_{\tilde{Z}, \tilde{X}} & S_{\tilde{Z}, \tilde{Z}} \end{bmatrix}$$

with

$$S_{\bar{X},\bar{X}} = \begin{bmatrix} V(X) & \text{Cov}(X, \delta_X) \\ \text{Cov}(\delta_X, X) & V(\delta_X) \end{bmatrix}, \quad S_{\bar{Z},\bar{Z}} = \begin{bmatrix} V(Z) & \text{Cov}(Z, \delta_Z) \\ \text{Cov}(\delta_Z, Z) & V(\delta_Z) \end{bmatrix},$$

$$S_{\bar{X},\bar{Z}} = \begin{bmatrix} \text{Cov}(X, Z) & \text{Cov}(X, \delta_Z) \\ \text{Cov}(\delta_X, Z) & \text{Cov}(\delta_X, \delta_Z) \end{bmatrix}, \quad S_{\bar{Z},\bar{X}} = S_{\bar{X},\bar{Z}}^T,$$

and

$$S_{Y,V} = \begin{bmatrix} S_{Y,\bar{X}} \\ S_{Y,\bar{Z}} \end{bmatrix} = \begin{bmatrix} \text{Cov}(Y, X) \\ \text{Cov}(Y, \delta_X) \\ \text{Cov}(Y, Z) \\ \text{Cov}(Y, \delta_Z) \end{bmatrix},$$

the first order approximation for $\text{MSE}(\bar{y}_g)$ is

$$\begin{aligned} \text{MSE}(\bar{y}_g) &\cong \left(\frac{1}{n} - \frac{1}{N}\right) S_Y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right) g_V^T S_{VV} g_V + \left(\frac{1}{n'} - \frac{1}{N}\right) \ddot{g}_Z^T S_{\bar{Z}\bar{Z}} \ddot{g}_Z \\ &+ 2 \left[\left(\frac{1}{n} - \frac{1}{n'}\right) g_V^T S_{YV} + \left(\frac{1}{n'} - \frac{1}{N}\right) \ddot{g}_Z^T S_{Y\bar{Z}} \right]. \end{aligned} \quad (9)$$

Minimizing equation (9) with respect to g_V and $\ddot{g}_Z$, the "best" values g_V^* and $\ddot{g}_Z^*$ (ch. 5) are the partial regression coefficients of Y on V and $\bar{Z}$, respectively, that is

$$g_V^* = \begin{bmatrix} g_X^* \\ g_Z^* \end{bmatrix} = -S_{V,V}^{-1} S_{Y,V} = - \begin{bmatrix} \beta_{Y,X|\delta_X,\bar{Z}} \\ \beta_{Y,\delta_X|X,\bar{Z}} \\ \beta_{Y,Z|\bar{X},\delta_Z} \\ \beta_{Y,\delta_Z|\bar{X},Z} \end{bmatrix} \quad \text{and} \quad \ddot{g}_Z^* = -S_{\bar{Z},\bar{Z}}^{-1} S_{Y,\bar{Z}} = - \begin{bmatrix} \beta_{Y,Z|\delta_Z} \\ \beta_{Y,\delta_Z|Z} \end{bmatrix}.$$

The minimum first order approximation of $\text{MSE}(\cdot)$ is

$$\text{MSE}^*(\bar{y}_g) = S_Y^2 \left[\left(\frac{1}{n'} - \frac{1}{N}\right) (1 - \rho_{Y,\bar{Z}}^2) + \left(\frac{1}{n} - \frac{1}{n'}\right) (1 - \rho_{Y,V}^2) \right]$$

where $\rho_{\cdot,\cdot}$ denotes the correlation coefficient between the specified variables, i.e.

$$\rho_{Y,\bar{Z}}^2 = \frac{S_{Y,\bar{Z}}^T S_{\bar{Z},\bar{Z}}^{-1} S_{Y,\bar{Z}}}{S_Y^2} \quad \text{and} \quad \rho_{Y,V}^2 = \frac{S_{Y,V}^T S_{V,V}^{-1} S_{Y,V}}{S_Y^2}.$$

Up to terms of $O(n^{-1})$, $\text{MSE}^*(\bar{y}_g)$ is the minimum variance bound for all the estimators based on the auxiliary vector t , which summarizes the first and second phase sample means and variances of X and Z .

Replacing g_V^* and $\ddot{g}_Z^*$ in (8) we get the following chain regression type estimator

$$\begin{aligned} \bar{y}_{reg} &= \bar{y} - \beta_{Y,X|\delta_X,\bar{Z}} (\bar{x} - \bar{x}') - \beta_{Y,\delta_X|X,\bar{Z}} (s_X^2 - s_X'^2) - \beta_{Y,Z|\bar{X},\delta_Z} (\bar{z} - \bar{z}') \\ &- \beta_{Y,\delta_Z|\bar{X},Z} (s_Z^2 - s_Z'^2) - \beta_{Y,Z|\delta_Z} (\bar{z}' - \bar{Z}) - \beta_{Y,\delta_Z|Z} (s_Z'^2 - S_Z^2), \end{aligned} \quad (10)$$

whose bias is of order n^{-1} . Expression (10) shows that the construction of $\bar{y}_{reg}$ is based on two steps. At the first step a regression type estimator for the unknown first phase sample mean of Y is provided. Of course, this estimator is based on X and Z conditionally to the first phase sample. Then such estimator is improved through another regression type estimator which uses the first phase sample information about Z . Actually, a similar interpretation was provided by Sahoo and Sahoo (1999) in a less general context. We call $\bar{y}_{reg}$ the "best" estimator in the class g but any other estimator which is equivalent to $\bar{y}_{reg}$, at the first order of approximation, is optimum as well.

Remark 1. Usually the partial regression coefficients of Y on V or $\tilde{Z}$ are unknown but replacing suitable estimates of such coefficients in (10) we get a new estimator which is equivalent to $\bar{y}_{reg}$, at the first order of approximation.

4 Estimation for ratio and product of two population means

This section is devoted to the estimation of ratio or product of two population means, i.e. $R = \bar{Y}_0/\bar{Y}_1$ or $P = \bar{Y}_0\bar{Y}_1$. Only the ratio estimation is developed in an exhaustive way since the results for the product case follow in a straightforward way and are briefly given at the end of this section.

Thus, at first let us consider $T_Y = R$. In this case, $t_Y = \hat{R} = \bar{y}_0/\bar{y}_1$ and the general class (1) becomes

$$\hat{R}_g = g(\hat{R}, t^T). \quad (11)$$

Expanding $\hat{R}$ at the point $(\bar{Y}_0, \bar{Y}_1)$ in a second order Taylor's series, we have

$$\hat{R} \cong R \left(1 + \frac{\bar{y}_0}{\bar{Y}_0} - \frac{\bar{y}_1}{\bar{Y}_1} \right) = R(1 + \bar{d}),$$

where $\bar{d}$ is the second phase sample mean of $D = Y_0/\bar{Y}_0 - Y_1/\bar{Y}_1$. Thus, with the same notation of the previous section, equation (3) becomes

$$\begin{aligned} \hat{R}_g &\cong \hat{R} + (t_V - t'_V)^T g_V + (t'_Z - T_Z)^T \ddot{g}_Z \\ &\cong R(1 + \bar{d}) + (t_V - t'_V)^T g_V + (t'_Z - T_Z)^T \ddot{g}_Z. \end{aligned} \quad (12)$$

Since $\hat{R}$ can be considered as a convenient sample mean (at the first order of approximation), we can use all the results of the previous section replacing

Y with $R(1 + D)$. Specifically, S_{YV} becomes RS_{DV} , S_Y^2 becomes $R^2S_D^2$,

$$g_V^* = -RS_{V,V}^{-1} S_{D,V} = -R \begin{bmatrix} \beta_{D,X|\delta_X,\bar{Z}} \\ \beta_{D,\delta_X|X,\bar{Z}} \\ \beta_{D,Z|\bar{X},\delta_Z} \\ \beta_{D,\delta_Z|\bar{X},Z} \end{bmatrix} \quad \text{and} \quad \ddot{g}_Z^* = -RS_{\bar{Z},\bar{Z}}^{-1} S_{D,\bar{Z}} = -R \begin{bmatrix} \beta_{D,Z|\delta_Z} \\ \beta_{D,\delta_Z|Z} \end{bmatrix},$$

where

$$S_{D,V} = \begin{bmatrix} S_{D,\bar{X}} \\ S_{D,\bar{Z}} \end{bmatrix} = \begin{bmatrix} \text{Cov}(D, X) \\ \text{Cov}(D, \delta_X) \\ \text{Cov}(D, Z) \\ \text{Cov}(D, \delta_Z) \end{bmatrix}.$$

With such changes formula (9) provides the minimum first order approximation of $\text{MSE}(\hat{R}_g)$,

$$\begin{aligned} \text{MSE}^*(\hat{R}_g) &= R^2 S_D^2 \left[\left(\frac{1}{n'} - \frac{1}{N} \right) (1 - \rho_{D,\bar{Z}}^2) + \left(\frac{1}{n} - \frac{1}{n'} \right) (1 - \rho_{D,V}^2) \right] \\ &= \text{MSE}'(\hat{R}) - R^2 S_D^2 \left[\left(\frac{1}{n'} - \frac{1}{N} \right) \rho_{D,\bar{Z}}^2 + \left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{D,V}^2 \right], \end{aligned}$$

where $\text{MSE}'(\hat{R}) = (N - n)/(nN) R^2 S_D^2$ is the first order approximation of the MSE of the standard estimator $\hat{R}$ and $\rho_{D,U}^2$ is the multiple correlation coefficient between D and any one random vector U .

Replacing g_V^* and $\ddot{g}_Z^*$ in equation (12) we get the best estimator in the class (11). Unfortunately g_V^* and $\ddot{g}_Z^*$ are not known since they depend on the population ratio R and on the likely unknown regression coefficients of D on V and $\bar{Z}$. However, using suitable estimates of g_V^* and $\ddot{g}_Z^*$ in equation (12) we get a chain regression type estimator which is equivalent to the best estimator (at the first order of approximation) and thus reaches $\text{MSE}^*(\hat{R}_g)$. Let us denote it as $\hat{R}_{reg}$. For giving the expression of $\hat{R}_{reg}$ we need to introduce more notation. Let U_1 and U_2 be any two auxiliary variables (where U_2 can be a random vector) and let us denote the partial regression coefficient of Y_j on U_1 given U_2 as $\beta_{j,U_1|U_2}$, $j = 0, 1$. If $\hat{\beta}_{D,U_1|U_2}$ is a suitable estimate of $\beta_{D,U_1|U_2} = \beta_{0,U_1|U_2}/\bar{Y}_0 - \beta_{1,U_1|U_2}/\bar{Y}_1$, then $\hat{R}_{reg}$ is given by

$$\begin{aligned} \hat{R}_{reg} &= \hat{R} \left[1 - \hat{\beta}_{D,X|\delta_X,\bar{Z}} (\bar{x} - \bar{x}') - \hat{\beta}_{D,\delta_X|X,\bar{Z}} (s_X^2 - s_X'^2) - \hat{\beta}_{D,Z|\bar{X},\delta_Z} (\bar{z} - \bar{z}') \right. \\ &\quad \left. - \hat{\beta}_{D,\delta_Z|\bar{X},Z} (s_Z^2 - s_Z'^2) - \hat{\beta}_{D,Z|\delta_Z} (\bar{z}' - \bar{z}) - \hat{\beta}_{D,\delta_Z|Z} (s_Z'^2 - S_Z^2) \right]. \end{aligned}$$

Remark. If the interest is in the population product $P = \bar{Y}_0 \bar{Y}_1$ all the results of this section are still valid. It is enough to replace R and $\hat{R}$ with P and $\hat{P}$ respectively, where now $\hat{P} = \bar{y}_0 \bar{y}_1 \cong P(\bar{d}-1)$ with $D = Y_0/\bar{Y}_0 + Y_1/\bar{Y}_1$.

5 Estimation of a finite population variance

In this section the estimation of the population variance, $S_Y^2 = \sum_{i=1}^N (Y_i - \bar{Y})^2 / (N - 1)$, is considered. Thus $t_Y = s_Y^2 = \sum_{i=1}^n (Y_i - \bar{y})^2 / (n - 1)$ and the general class (1) becomes

$$\hat{S}_Y^2 = g(s_Y^2, t^T). \quad (13)$$

Setting $\delta_Y = (Y - \bar{Y})^2$, s_Y^2 is the sample mean of δ_Y except for terms of order n^{-1} . In fact

$$s_Y^2 = \frac{n}{n-1} [\tilde{s}_Y^2 - (\bar{y} - \bar{Y})^2] \cong \tilde{s}_Y^2,$$

where $\tilde{s}_Y^2 = \sum_{i=1}^n \delta_{Yi} / n$. Thus, with the same notation for $t_V, t'_V, t_{\bar{Z}}, t'_{\bar{Z}}, g_V$ and $\tilde{g}_Z$ used in Sections 3 and 4, equation (3) is now

$$\begin{aligned} \hat{S}_Y^2 &\cong s_Y^2 + (t_V - t'_V)^T g_V + (t'_{\bar{Z}} - T_{\bar{Z}})^T \tilde{g}_Z \\ &\cong \tilde{s}_Y^2 + (t_V - t'_V)^T g_V + (t'_{\bar{Z}} - T_{\bar{Z}})^T \tilde{g}_Z. \end{aligned} \quad (14)$$

Replacing Y with δ_Y , we can use again all the results of Section 3. Now, S_{YV} becomes $S_{\delta_Y V}$, S_Y^2 becomes $S_{\delta_Y}^2$,

$$g_V^* = -S_{V,V}^{-1} S_{\delta_Y, V} = - \begin{bmatrix} \beta_{\delta_Y, X | \delta_X, \bar{Z}} \\ \beta_{\delta_Y, \delta_X | X, \bar{Z}} \\ \beta_{\delta_Y, Z | \bar{X}, \delta_Z} \\ \beta_{\delta_Y, \delta_Z | \bar{X}, Z} \end{bmatrix} \quad \text{and} \quad \tilde{g}_Z^* = -S_{\bar{Z}, \bar{Z}}^{-1} S_{\delta_Y, \bar{Z}} = - \begin{bmatrix} \beta_{\delta_Y, Z | \delta_Z} \\ \beta_{\delta_Y, \delta_Z | Z} \end{bmatrix},$$

where

$$S_{\delta_Y, V} = \begin{bmatrix} S_{\delta_Y, \bar{X}} \\ S_{\delta_Y, \bar{Z}} \end{bmatrix} = \begin{bmatrix} \text{Cov}(\delta_Y, X) \\ \text{Cov}(\delta_Y, \delta_X) \\ \text{Cov}(\delta_Y, Z) \\ \text{Cov}(\delta_Y, \delta_Z) \end{bmatrix}.$$

With such changes, from (9) we have the minimum first order approximation of $\text{MSE}(\hat{S}_Y^2)$,

$$\text{MSE}^*(\hat{S}_Y^2) = S_{\delta_Y}^2 \left[\left(\frac{1}{n'} - \frac{1}{N} \right) (1 - \rho_{\delta_Y, \bar{Z}}^2) + \left(\frac{1}{n} - \frac{1}{n'} \right) (1 - \rho_{\delta_Y, V}^2) \right]$$

where $\rho_{\delta_Y, U}^2$ is the multiple correlation coefficient between δ_Y and any one random vector U .

Replacing g_V^* and $\tilde{g}_Z^*$ in equation (14) we get the best estimator in the class (13),

$$\begin{aligned} \hat{S}_{reg, Y}^2 &= s_Y^2 - \beta_{\delta_Y, X | \delta_X, \bar{Z}} (\bar{x} - \bar{x}') - \beta_{\delta_Y, \delta_X | X, \bar{Z}} (s_X^2 - s_X'^2) - \beta_{\delta_Y, Z | \bar{X}, \delta_Z} (\bar{z} - \bar{z}') \\ &\quad - \beta_{\delta_Y, \delta_Z | \bar{X}, Z} (s_Z^2 - s_Z'^2) - \beta_{\delta_Y, Z | \delta_Z} (\bar{z}' - \bar{Z}) - \beta_{\delta_Y, \delta_Z | Z} (s_Z'^2 - S_Z^2), \end{aligned}$$

As in the previous cases, when g_V^* and $\ddot{g}_Z^*$ are unknown we can replace them with suitable estimates, without efficiency lost.

6 A partial use of the auxiliary variables

In the previous sections the whole auxiliary vector $t = (\bar{x}, s_X^2, \bar{z}, s_Z^2, \bar{x}', s_X'^2, \bar{z}', s_Z'^2)^T$ was assumed to be known. Sometimes only some of those auxiliary quantities are available (or some information are ignored). For instance, if we have only the first phase information about Z , t becomes $t_6 = (\bar{x}, s_X^2, \bar{x}', s_X'^2, \bar{z}', s_Z'^2)^T$. For the case of the population mean estimation, $g_V^* = g_X^* = -(\beta_{Y,X|\delta_X}, \beta_{Y,\delta_X|X})^T$ and thus the best estimator is

$$\bar{y}_6 = \bar{y} - \beta_{Y,X|\delta_X}(\bar{x} - \bar{x}') - \beta_{Y,\delta_X|X}(s_X^2 - s_X'^2) - \beta_{Y,Z|\delta_Z}(\bar{z}' - \bar{Z}) - \beta_{Y,\delta_Z|Z}(s_Z'^2 - S_Z^2)$$

and

$$\text{MSE}^*(\bar{y}_6) = S_Y^2 \left[\left(\frac{1}{n'} - \frac{1}{N} \right) (1 - \rho_{Y,\bar{z}}^2) + \left(\frac{1}{n} - \frac{1}{n'} \right) (1 - \rho_{Y,\bar{x}}^2) \right].$$

In the same way, all the results of Sections 3, 4 and 5 can be easily adapted to any specific choice of vector t .

Choosing different auxiliary quantities we may have several general classes of estimators. However, at least when $\hat{T}_Y$ is $\bar{y}_g$, $\hat{R}_g$ or $\hat{S}_Y^2$, the best choice is to use the whole vector t since the more the used auxiliary information the more efficient the estimates. This follows from the expression of $\text{MSE}^*(\hat{T}_Y)$ which depends on the multiple correlation coefficients. In general, if $t = (t_-, t_m)^T$ is a $m \times 1$ random vector, where t_- is the $(m-1) \times 1$ vector got removing the last variable t_m from t , then the square multiple correlation coefficient of any one variable U with t is

$$\rho_{U,t}^2 = \rho_{U,t_-}^2 + \rho_{U,(t_m \cdot t_-)}^2,$$

where $\rho_{U,(t_m \cdot t_-)}^2$ is the semi-partial correlation coefficient between U and t_m given t_- .

Using this argument we have

$$\text{MSE}^*(\hat{T}_{Y2}) \geq \text{MSE}^*(\hat{T}_{Y3}) \geq \dots \geq \text{MSE}^*(\hat{T}_{Y8}),$$

where $\hat{T}_{Y2}$ is $\hat{T}_Y$ based on the simplest auxiliary vector $t_2 = (\bar{x}, \bar{x}')^T$, $\hat{T}_{Y3}$ on $t_3 = (\bar{x}, \bar{x}', \bar{z}')^T$ and so on up to $\hat{T}_{Y8}$ where the whole vector of auxiliary quantities is considered. However, when competitive estimators are based on auxiliary vectors of the same size but involving different statistics, this

general result is useless and specific comparisons are necessary. For instance, we could be interested in the convenience of using information about Z in terms of $\bar{z}'$ or $s_Z'^2$. If we use auxiliary vector t_3 then the best estimator for the population mean is

$$\bar{y}_{3reg} = \bar{y} - \beta_{YX}(\bar{x} - \bar{x}') - \beta_{YZ}(\bar{z}' - \bar{Z})$$

and

$$\text{MSE}^*(\bar{y}_3) = S_Y^2 \left[\left(\frac{1}{n'} - \frac{1}{N} \right) (1 - \rho_{Y,Z}^2) + \left(\frac{1}{n} - \frac{1}{n'} \right) (1 - \rho_{Y,X}^2) \right]. \quad (15)$$

Otherwise, using $t'_3 = (\bar{x}, \bar{x}', s_Z'^2)^T$, we get

$$\bar{y}'_{3reg} = \bar{y} - \beta_{YX}(\bar{x} - \bar{x}') - \beta_{Y\delta_Z}(s_Z'^2 - S_Z^2)$$

and

$$\text{MSE}^*(\bar{y}'_3) = S_Y^2 \left[\left(\frac{1}{n'} - \frac{1}{N} \right) (1 - \rho_{Y,\delta_Z}^2) + \left(\frac{1}{n} - \frac{1}{n'} \right) (1 - \rho_{Y,X}^2) \right]. \quad (16)$$

Comparing (15) with (16) we have that the knowledge of $\bar{z}'$ is preferable to $s_Z'^2$ if $\rho_{YZ}^2 > \rho_{Y,\delta_Z}^2$, which is usually true when Y and Z are strongly related. For the mean case, as well as for ratio and product of two means, many different estimators based on a two phase sampling were proposed. On the other hand few proposals are available for the variance estimation. It is easily proved that only some of these estimators are equivalent to the best estimator while others are less efficient. A simple way to verify if a specific estimator is optimum is through the comparison of its second order Taylor's expansion with equation (3). If these developments are essentially the same, then the proposed estimator is optimum as well. For instance Singh and Espejo (2000) propose the following class of estimators for the population mean,

$$\bar{y}_G = \bar{y} + \hat{\beta}_{YX}[G(\bar{x}', u, v, w, \psi) - \bar{x}],$$

where $u = \bar{z}'/\bar{Z}$, $v = \bar{z}/\bar{Z}$, $w = s_Z'^2/S_Z^2$, $\psi = s_Z^2/S_Z^2$ and G is a general function with some regularity properties. This estimator is not optimal, in fact its second order Taylor's expansion is given by

$$\bar{y} + \beta_{YX}(\bar{x} - \bar{x}') + \alpha_1(s_Z^2 - s_Z'^2) + \alpha_2(\bar{z} - \bar{z}') + \alpha_3(s_Z'^2 - S_Z^2) + \alpha_4(\bar{z}' - \bar{Z})$$

with suitable values for α_i , $i = 1, \dots, 4$, got minimizing $\text{MSE}(\bar{y}_G)$ (up to terms of $O(n^{-1})$). The best estimator when the auxiliary vector is $t'_6 = (\bar{x}, \bar{x}', \bar{z}', \bar{z}, s_Z'^2, s_Z^2)^T$ follows from (8) setting $t_V = (\bar{x}, \bar{z}, s_Z^2)^T$, $t'_V = (\bar{x}', \bar{z}', s_Z'^2)^T$,

$g_V = g_V^* = -(\beta_{Y,X|\bar{Z}}, \beta_{Y,Z|X,\delta_Z}, \beta_{Y,\delta_Z|X,Z})^T$ and $\ddot{g}_Z = \ddot{g}_Z^* = -(\beta_{Y,Z|\delta_Z}, \beta_{Y,\delta_Z|Z})^T$.
It is given by

$$\begin{aligned}\bar{y}'_{\text{GREG}} &= \bar{y} + \beta_{Y,X|\bar{Z}}(\bar{x} - \bar{x}') - \beta_{Y,Z|X,\delta_Z}(\bar{z} - \bar{z}') - \beta_{Y,\delta_Z|X,Z}(s_Z^2 - s_Z'^2) \\ &\quad - \beta_{Y,Z|\delta_Z}(\bar{z}' - \bar{Z}) - \beta_{Y,\delta_Z|Z}(s_Z'^2 - S_Z^2).\end{aligned}$$

The difference in the coefficient of the term $(\bar{x} - \bar{x}')$ shows that $\bar{y}_G$ cannot be optimal. Another example is given by the following class of estimators

Table 1: *Classes of estimators for the population mean whose optimal estimator is equivalent to the best one.*

Vector t	Classes of estimators	Author(s)
$t_4 = (\bar{x}, \bar{x}', \bar{z}, \bar{z}')^T$	$\bar{y}_{GK} = \bar{y} + d_1(\bar{x} - \bar{x}') + d_2(\bar{z} - \bar{Z}) + d_3(\bar{z}' - \bar{Z})$	Mishra, Rout '97
	$\bar{y}_{ARA} = f\left(\bar{y}, \frac{\bar{x}}{\bar{x}'}, \frac{\bar{z}}{\bar{z}'}, \frac{\bar{z}'}{\bar{Z}}\right)$	Ahmed <i>et al.</i> '98
$t'_4 = (\bar{x}, \bar{x}', \bar{z}', s_Z'^2)^T$	$\bar{y}_{TS}^* = \bar{y} + \hat{\beta}_{YX} \left[\bar{x}' f\left(\frac{\bar{z}'}{\bar{Z}}, \frac{s_Z'^2}{S_Z^2}\right) - \bar{x} \right]$	Tracy, Singh '99
	$\bar{y}_{TS} = \bar{y} + \hat{\beta}_{YX} \left[f\left(\bar{x}', \frac{\bar{z}'}{\bar{Z}}, \frac{s_Z'^2}{S_Z^2}\right) - \bar{x} \right]$	Tracy, Singh '99
$t_3 = (\bar{x}, \bar{x}', \bar{z}')^T$	$\bar{y}_{SKS} = \bar{y} \left(\frac{\bar{x}'}{\bar{x}} \right)^{\alpha_1} \left(\frac{\bar{z}'}{\bar{Z}} \right)^{\alpha_2}$	Srivastava <i>et al.</i> '90
	$\bar{y}_{SS} = f_1(\bar{y}, \bar{x}, f_2(\bar{x}', \bar{z}'))$	Sahoo, Sahoo '93
	$\bar{y}_{SSM} = \bar{y} + \hat{\beta}_{YX} [\bar{x}' + d(\bar{Z} - \bar{z}') - \bar{x}]$	Sahoo <i>et al.</i> '94
	$\bar{y}_{SS} = f_1(f_2(\bar{y}, \bar{x}), \bar{x}', \bar{z}')$	Sahoo, Sahoo '99
$t_2 = (\bar{x}, \bar{x}')^T$	$\bar{y}_{SG} = \bar{y} \left(d_1 + d_2 \frac{\bar{x}'}{\bar{x}} + d_3 \left(\frac{\bar{x}'}{\bar{x}} \right)^2 \right), \sum_{i=1}^3 d_i = 1$	Singh, Gangele '99

for ratio of two population means, proposed by Prasad *et al.* (1996)

$$\hat{R}_\omega = \hat{R} \frac{\bar{x}}{\bar{x}'} \left[\omega + (1 - \omega) \frac{\bar{z}'}{\bar{Z}} \right].$$

No estimator in that class can be optimal in the general class defined by vector $t_3 = (\bar{x}, \bar{x}', \bar{z}')^T$, since it is not equivalent to the best estimator. As a matter of fact, the second order Taylor's expansion of $\hat{R}_\omega$,

$$\hat{R}_\omega = \hat{R} + R \frac{\omega}{\bar{X}} (\bar{x} - \bar{x}') + R \frac{1 - \omega}{\bar{Z}} (\bar{z}' - \bar{z}),$$

depends on only one unknown parameter while equation (12) depends on two parameters and thus minimizing the corresponding $\text{MSE}^*(\cdot)$'s, with respect to the unknown parameters, a lowest value will be got in the latter case.

Table 2: *Classes of estimators for ratio of two population means, whose optimal estimator is equivalent to the best one.*

Vector t	Classes of (or) estimators	Author(s)
$t_4 = (\bar{x}, \bar{x}', \bar{z}, \bar{z}')^T$	$\hat{R}_A = f\left(\hat{R}, \frac{\bar{x}}{\bar{z}}, \frac{\bar{x}'}{\bar{z}'}, \frac{\bar{z}}{\bar{z}'}\right)$	Ahmed '97
	$\hat{R}_{SS} = \frac{\bar{y}_0 - \beta_{0X Z} \{\bar{x} - [\bar{x}' - \beta'_{XZ}(\bar{z}' - \bar{Z})]\} - \beta_{0Z X}(\bar{z} - \bar{Z})}{\bar{y}_1 - \beta_{1X Z} \{\bar{x} - [\bar{x}' - \beta'_{XZ}(\bar{z}' - \bar{Z})]\} - \beta_{1Z X}(\bar{z} - \bar{Z})}$	Singh, Singh '97
$t_4'' = (\bar{x}, s_X^2, \bar{x}', s_X'^2)^T$	$\hat{R}_{SS} = f\left(\hat{R}, \frac{\bar{x}}{\bar{z}}, \frac{s_X^2}{s_X'^2}\right)$	Singh, Singh '94
	$\hat{R}_{SS}^* = \hat{R} f\left(\frac{\bar{x}}{\bar{z}}, \frac{s_X^2}{s_X'^2}\right)$	Singh, Singh '94
$t_3 = (\bar{x}, \bar{x}', \bar{z}')^T$	$\hat{R}_{SRKS} = \hat{R} \left(\frac{\bar{x}}{\bar{z}}\right)^{\alpha_1} \left(\frac{\bar{z}}{\bar{z}'}\right)^{\alpha_2}$	Srivastava et al. '88
	$\hat{R}_S = f\left(\hat{R}, \frac{\bar{x}}{\bar{z}}, \frac{\bar{z}}{\bar{z}'}\right)$	Singh et al. '94
	$\hat{R}_{PSS} = \hat{R} \frac{\bar{x} + \alpha_1(\bar{x} - \bar{x}')}{\bar{z}} \frac{\bar{z}' - \alpha_2(\bar{z}' - \bar{Z})}{\bar{z}}$	Prasad et al. '96
	$\hat{R}_{SS} = \frac{\bar{y}_0 - \beta_{0X}[\bar{x} - \bar{x}'(\bar{Z}/\bar{x}')^{\hat{\alpha}}]}{\bar{y}_1 - \beta_{1X}[\bar{x} - \bar{x}'(\bar{Z}/\bar{x}')^{\hat{\alpha}}]}, \hat{\alpha} = \frac{\bar{Z}(\hat{\beta}_{0Z} - \hat{R}\hat{\beta}_{1Z})}{\bar{x}'(\hat{\beta}_{0X} - \hat{R}\hat{\beta}_{1X})}$	Singh, Singh '97
	$\hat{R}_{KS} = \hat{R} \left[\alpha_0 \left(\frac{\bar{x}}{\bar{z}}\right)^{\alpha_1} + (1 - \alpha_0) \left(\frac{\bar{z}}{\bar{z}'}\right)^{\alpha_2} \right]$	Khare, Srivastava '98
	$\hat{R}_K = \hat{R} f\left(\frac{\bar{x}}{\bar{z}}\right)$	Khare '91
$t_2 = (\bar{x}, \bar{x}')^T$	$\hat{R}_{SS} = \frac{\bar{y}_0 + \beta_{0X}(\bar{x}' - \bar{x})}{\bar{y}_1 + \beta_{1X}(\bar{x}' - \bar{x})}$	Singh, Singh '97

Table 3: *Classes of estimators for the population variance, whose optimal estimator is equivalent to the best one.*

Vector t	Classes of estimators	Author(s)
$t_6 = (\bar{x}, s_X^2, \bar{x}', s_X'^2, \bar{z}', s_Z'^2)^T$	$\hat{S}_{GSM}^2 = s_Y^2 \left(\frac{\bar{x}}{\bar{z}}\right)^\alpha \left(\frac{s_X^2}{s_X'^2}\right)^\beta \left(\frac{\bar{z}'}{\bar{z}}\right)^\gamma \left(\frac{s_Z'^2}{s_Z^2}\right)^\phi$	Gupta et al. '92
$t_4'' = (\bar{x}, s_X^2, \bar{x}', s_X'^2)^T$	$\hat{S}_S^2 = s_Y^2 h\left(\frac{\bar{x}}{\bar{z}}, \frac{s_X^2}{s_X'^2}\right)$	Singh '91
$t_3 = (\bar{x}, \bar{x}', \bar{z}')^T$	$\hat{S}_{GSM}^2 = s_Y^2 \left(\frac{\bar{x}}{\bar{z}}\right)^\alpha \left(\frac{\bar{z}'}{\bar{z}}\right)^\beta$	Gupta et al. '92
$t_3' = (s_X^2, s_X'^2, s_Z'^2)^T$	$\hat{S}_{GSM}^2 = s_Y^2 \left(\frac{s_X^2}{s_X'^2}\right)^\alpha \left(\frac{s_Z'^2}{s_Z^2}\right)^\beta$	Gupta et al. '92

On the other hand, the same argument proves that other proposed estimators are equivalent to the best ones and so they are optimal. Some examples for mean, ratio of two means and variance, are provided in Table 1, Table 2 and Table 3, respectively.

References

- Ahmed M.S. (1997), The General Class of Chain Estimators for the Ratio of two Means Using Double Sampling, *Commun. Statist. Theory & Meth.*, 26(9), 2247–2254.
- Ahmed M.S., Rahman M.S., Ahmed R. (1998), General Class of Chain Estimators for a Finite Population Mean Using Double Sampling, *Journal of Applied Statistical Science*, 4, 185–90.
- Chand L. (1975), *Some Ratio-Type Estimators based on Two or More Auxiliary Variates*, (Unpublished Ph.D. Dissertation, Iowa State University, Ames, IOWA, USA).
- Diana G., Tommasi C. (2001), *Estimation for Finite Population Variance in Double Sampling*, Working Paper #2001.22, University of Padova, Padova, Italy.
- Diana G., Tommasi C. (2002a), *Optimal Estimation for Finite Population Mean in Two Phase Sampling*, Working Paper #2002.3, University of Padova, Padova, Italy.
- Diana G., Tommasi C. (2002b), *Estimation for Ratio of Two Population Means in Double Sampling*, Working Paper #2002.5, University of Padova, Padova, Italy.
- Gupta R.K., Singh S., Mangat N.S. (1992). Some chain ratio type estimators for estimating finite population variance, *Aligarh J. Statist.*, 12-13, 65-69.
- Khare B.B. (1991), Determination of Sample Sizes for a Class of Two-Phase Sampling Estimators for Ratio and Product of two Population Means Using Auxiliary Character, *Metron*, 49, 185-197.
- Khare B.B., Srivastava S.R. (1998), Combined Generalised Chain Estimators for Ratio and Product of two Population Means Using Auxiliary Characters, *Metron*, 56, 109–116.
- Mishra G., Rout K. (1997), A Regression Estimator in Two-Phase Sampling in Presence of Two Auxiliary Variables, *Metron*, 55:1-2, 177–186.
- Prasad B., Singh R.S., Singh H.P. (1996), Some Chain Ratio-Type Estimators for Ratio of two Population Means Using Two Auxiliary Characters in Two Phase Sampling, *Metron*, 54, 95–113.
- Sahoo J., Sahoo L.N. (1993), A Class of Estimators in Two-Phase Sampling Using Two Auxiliary Variables, *J. Indian Statist. Assoc.*, 31, 107–114.

- Sahoo J., Sahoo L.N. (1999), An Alternative Class of Estimators in Double Sampling Procedures, *Cal. Statist. Assoc. Bull.*, 49, 193-194.
- Sahoo J., Sahoo L.N., Mohanty S. (1994), An Alternative Approach to Estimation in Two-Phase Sampling Using Two Auxiliary Variables, *Biom. J.*, 36, 293-298.
- Singh S. (1991), Estimation of finite population variance using double sampling, *Aligarh J. Statist.*, 11, 53-56.
- Singh H.P., Espejo Ruiz M. (2000), An Improved Class of Chain Regression Estimator in Two-Phase Sampling, *Statistics & Decisions*, 18, 205-218.
- Singh H.P., Gangele R.H. (1999), Classes of Almost Unbiased Ratio and Product-Type Estimators in Two-Phase Sampling, *Statistica*, 1, 109-124.
- Singh V.K., Singh G.N. (1991), Chain Type Regression Estimators with Two Auxiliary Variables under Double Sampling Scheme, *Metron*, 49, 279-289.
- Singh V.K., Singh H. P. (1994), Estimation of Ratio and Product of two Finite Population Means in Two Phase Sampling, *Journ. Statist. Plann. Infer.*, 41, 163-171.
- Singh V.P., Singh H. P. (1997), Chain Estimators for Population Ratio in Double Sampling, *Aligarh J. Statist.*, 17-18, 85-100.
- Singh V.K., Singh Hari P., Singh Housila P., Shukla D. (1994), A General Class of Chain Estimators for Ratio and Product of two Means in Finite Population, *Commun. Statist. Theory & Meth.*, 23(5), 1341-1355.
- Srivastava R., Khare B.B., Srivastava S.R. (1990), A Generalized Chain Ratio Estimator for Mean of Finite Population, *J. Indian Soc. Agric. Statist.*, 42, 108-117.
- Srivastava S. Rani, Khare B.B., Srivastava S.R. (1988), On Generalized Chain Estimators Ratio and Product of two Population Means Using Auxiliary Characters, *Assam. Stat. Review*, 2(1), 21-29.
- Srivastava S. Rani, Srivastava S.R., Khare B.B. (1989), Chain Ratio-Type Estimators for Ratio of two Population Means Using Auxiliary Characters, *Commun. Statist. Theory & Meth.*, 18(10), 3917-3926.
- Tracy D.S., Singh H.P. (1999), A General Class of Chain Regression Estimators in Two-Phase Sampling, *J. Appl. Statist. Sci.*, 8, 205-216.
- Upadhyaya L.N., Dubey S.P., Singh H.P. (1992), A Class of Ratio-in-Regression Estimators Using Two Auxiliary Variables in Double Sampling, *Journ. Scient. Res.*, 42, 127-134.

