Families of Estimators of Finite Population Variance using a Random Non-Response in Survey Sampling

Housila P. Singh1 · Rajesh Tailor2 · Jong-Min Kim3 · Sarjinder Singh4

1School of Studies in Statistics, Vikram University; 2School of Studies in Statistics, Vikram University
3Statistics Discipline, Division of Science and Mathematics, University of Minnesota-Morris
4Department of Mathematics, Texas A&M University-Kingsville

(Received March 6, 2012; Revised April 8, 2012; Accepted June 7, 2012)

Abstract

In this paper, a family of estimators for the finite population variance investigated by Srivastava and Jhajj (1980) is studied under two different situations of random non-response considered by Tracy and Osahan (1994). Asymptotic expressions for the biases and mean squared errors of members of the proposed family are obtained; in addition, an asymptotic optimum estimator (AOE) is also identified. Estimators suggested by Singh and Joarder (1998) are shown to be members of the proposed family. A correction to the Singh and Joarder (1998) results is also presented.

Keywords: Finite population variance, study and auxiliary variables, bias, mean squared error, random non-response.

1. Introduction

A finite population parameter can be estimated more accurately by making use of information on an auxiliary variable $x$ that is correlated with the study variable $y$. Ratio and regression methods of estimation are good examples in this context. Isaki (1983) showed that under realistic conditions efficient estimators of the finite population variance exist in the presence of auxiliary information. Let $y$ denote the character whose population variance $S_y^2 = \sum_{i=1}^{N}(y_i - \bar{Y})^2/(N-1)$ is estimated using information on an auxiliary variable $x$, where $\bar{Y} = (\sum_{i=1}^{N}y_i)/(N-1)$. Assuming that the population mean $\bar{X} = (\sum_{i=1}^{N}x_i)/(N-1)$ and variance $S_x^2 = \sum_{i=1}^{N}(x_i - \bar{X})^2/(N-1)$ of $x$ are known, Isaki (1983) proposed a ratio estimator:

$$d_1 = s_y^2 \left( \frac{S_x^2}{s_x^2} \right),$$

where $s_y^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2/(n-1)$ and $s_x^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)$ with $\bar{y} = (\sum_{i=1}^{n}y_i)/n$ and $\bar{x} = (\sum_{i=1}^{n}x_i)/n$. 

3Corresponding author: Professor, Statistics Discipline, Division of Science and Mathematics, University of Minnesota-Morris, Morris, MN 56267, USA. E-mail: jongmink@morris.umn.edu
When $\bar{X}$ and $S_x^2$ are known, Srivastava and Jhajj (1980) proposed a family of estimators for $S_y^2$ given by

$$d_2 = s_y^2 h(u, v), \quad (1.2)$$

where $u = \bar{x}/\bar{X}$, $v = s_x^2/S_x^2$ and $h(u, v)$ is a parametric function that satisfies certain conditions given in Srivastava and Jhajj (1980) and is such that $h(1, 1) = 1$. The bias and mean squared error (MSE) of $d_2$, to the first degree of approximation, are

$$B(d_2) = \theta \left( \frac{S_y^2}{2} \right) \left[ \lambda_{21} C_x h_1(1, 1) + (\lambda_{22} - 1)h_2(1, 1) + C_x^2 h_{11}(1, 1) + 2\lambda_{03} C_x h_{12}(1, 1) 
+ (\lambda_{04} - 1)h_{22}(1, 1) \right] \quad (1.3)$$

$$\text{MSE}(d_2) = \theta S_y^4 \left[ (\lambda_{40} - 1) + C_x^2 h_1(1, 1) + (\lambda_{04} - 1)h_2(1, 1) 
+ 2\lambda_{21} C_x h_1(1, 1) + 2(\lambda_{22} - 1)h_2(1, 1) + 2\lambda_{03} C_x h_1(1, 1)h_2(1, 1) \right],$$

where $\theta = (1/n - 1/N)$, $C_x = S_x/\bar{X}$, $h_i(1, 1)$, $i = 1, 2$ and $h_{ij}(u, v)$, $i, j = 1, 2$ denote the first and second order partial derivatives of $h(u, v)$,

$$\lambda_{ls} = \mu_{ls} \left( \mu_{20}^2 \mu_{02}^2 \right)^{-1}, \quad (1.5)$$

and

$$\mu_{ls} = (N - 1)^{-1} \sum_{i=1}^{N} (y_i - \bar{Y})^l (x_i - \bar{X})^s, \quad (l, s) = 0, 1, 2, 3, 4. \quad (1.6)$$

The MSE of $d_2$ is minimized for

$$h_1(1, 1) = \frac{\lambda_{03}(\lambda_{22} - 1) - \lambda_{21}(\lambda_{04} - 1)}{C_x(\lambda_{04} - \lambda_{03}^2 - 1)} = A \quad \text{and} \quad h_2(1, 1) = \frac{(\lambda_{21} \lambda_{03} - \lambda_{22} + 1)}{\lambda_{04} - \lambda_{03}^2 - 1} = B \quad \text{(say)}. \quad (1.7)$$

This gives the minimum MSE of $d_2$ as

$$\text{min.}MSE(d_2) = \theta S_y^4 [\lambda_{40} - D], \quad (1.8)$$

where

$$D = \frac{\lambda_{21}^2 + (\lambda_{21} \lambda_{03} - \lambda_{22} + 1)^2}{\lambda_{04} - \lambda_{03}^2 - 1}. \quad (1.9)$$

Note that $d_1$ is a particular case of $d_2$.

Singh and Joarder (1998) studied the properties of $d_1$ under two different situations of a random non-response considered by Tracy and Osahan (1994): (i) random non-response on both the study and auxiliary variables, and (ii) on the study variable only. In this paper, we study the effect of random non-response on the study and auxiliary variables of several families of estimators of variance. The estimators reported by Singh and Joarder (1998) are shown to be particular cases of the proposed families.
2. Notation and Expectations

Let $U = (U_1, U_2, \ldots, U_N)$ denote a population of $N$ units from which a simple random sample of size $n$ is drawn without replacement. If $r$ ($r = 0, 1, 2, \ldots, (n-2)$) denotes the number of sampling units on which information could not be obtained due to a random non-response, then the remaining $(n-r)$ units can be treated as a simple random sample from $U$. We assume $0 \leq r \leq (n-2)$, as we are interested in the problem of an unbiased estimation of the finite population variance. Singh and Joarder (1998) have given the distribution of $r$ as

$$P(r) = \frac{n-r}{nq+2p} \left( \frac{2-n}{r} \right) p^{n-2-r},$$  \hspace{1cm} (2.1)

where $p$ is the probability of non-response, $q = 1 - p$ and $\binom{n-2}{r}$ represents the total number of ways to obtain $r$ non-responses out of a possible $(n-2)$. We write $e_0 = s_x^2/S_x^2 - 1$, $e_1 = \bar{x}/\bar{X} - 1 = (u^* - 1)$, $e_2 = s_x^2/S_x^2 - 1 = (v^* - 1)$, $e_3 = \bar{x}/\bar{X} - 1 = (u - 1)$, and $e_4 = s_x^2/S_x^2 - 1 = (v - 1)$; where $\bar{x} = (n-r)^{-1} \sum_{i=1}^{n-r} x_i$, $\bar{y} = (n-r)^{-1} \sum_{i=1}^{n-r} y_i$, $s_x^2 = (n-r-1)^{-1} \sum_{i=1}^{n-r} (y_i - \bar{y})^2$ and $s_y^2$ as $(n-r-1)^{-1} \sum_{i=1}^{n-r} (x_i - \bar{x})^2$.

Then, under model (2.1) $E(e_0) = \theta^*(\lambda_{40} - 1)$, $E(e_1^2) = \theta^* C_2^2$, $E(e_2^2) = \theta^* (\lambda_{40} - 1)$, $E(e_3) = \theta C_2$, $E(e_4) = \theta (\lambda_{40} - 1)$, $E(e_3 e_4) = \theta (\lambda_{22} - 1)$, $E(e_3^2) = \theta (\lambda_{22} - 1)$, $E(e_4^2) = \theta (\lambda_{22} - 1)$, $E(e_3 e_4) = \theta (\lambda_{22} - 1)$, $E(e_3^2) = \theta (\lambda_{22} - 1)$. Thus, the bias of $d_3$ is of the order of $n^{-1}$. The MSE of $d_3$ up to terms of order $n^{-1}$ is

$$\text{MSE}(d_3) = \theta^* S_y^4 \left[ (\lambda_{40} - 1) + C_x^2 e_2^2 (1, 1) + (\lambda_{40} - 1) e_2^2 (1, 1) \\
+ 2 \lambda_{40} C_x t_1 (1, 1) t_2 (1, 1) + 2 \lambda_{22} C_x t_1 (1, 1) + 2 (\lambda_{22} - 1) t_2 (1, 1) \right],$$  \hspace{1cm} (3.2)

where $t_1 (1, 1)$ and $t_2 (1, 1)$ denote the first order partial derivatives of $t(u^*, v^*)$. The MSE of $d_3$ in (3.2) is minimized for

$$t_1 (1, 1) = A \quad \text{and} \quad t_2 (1, 1) = B,$$  \hspace{1cm} (3.3)
where $A$ and $B$ are given in (1.7). Substitution of (3.3) into (3.2) yields the minimum MSE of $d_3$ as

$$\text{min}.\text{MSE}(d_3) = \theta S_y^4[\lambda_{40} - 1 - D],$$  

(3.4)

where $D$ is defined in (1.9). Thus we have the following theorem.

**Theorem 3.1.** Up to terms of order $n^{-1}$

$$\text{MSE}(d_3) \geq \theta S_y^4[\lambda_{40} - 1 - D]$$

with equality holding if $t_1(1, 1) = A$ and $t_2(1, 1) = B$.

Any parametric function $t(u^*, v^*)$ satisfying conditions (1) and (2) can define an estimator of $S_y^2$. Therefore in addition to the Singh and Joarder (1998) estimator of $S_y^2$

$$d_3(0) = s_y^2 \left( \frac{S_y^2}{S_x^2} \right)$$  

(3.5)

the estimators

$$d_3(1) = s_y^2 u^* \alpha + v^* \beta$$

$$d_3(2) = s_y^2 \frac{1 + \alpha(u^* - 1)}{1 + \beta(v^* - 1)}$$

$$d_3(3) = s_y^2 \{1 + \alpha(u^* - 1) + \beta(v^* - 1)\}$$

$$d_3(4) = s_y^2 \{1 - \alpha(u^* - 1) - \beta(v^* - 1)\}$$

$$d_3(5) = s_y^2 \{2 - u^* \alpha + v^* \beta\}$$

$$d_3(6) = \frac{1 + \gamma(u^* \alpha v^* \beta - 1)}{S_y^2}$$

$$d_3(7) = s_y^2 \exp\{\alpha(u^* - 1) + \beta(v^* - 1)\}$$

$$d_3(8) = s_y^2 \{\alpha u^* + (1 - \alpha)v^* \beta\}$$  

(3.6)

and so on are particular members of the proposed family $d_3$, where $\alpha, \beta, \gamma$ are constants. It is easily seen that the optimum values of the parameters $\alpha$ and $\beta$ in the above estimators are given by $A$ and $B$ in (1.7). The minimum MSE's of $d_3(j); j = 1$ to 8 are equal to (3.4). Note also that it can be shown that if we consider a wider family of estimators $d_4 = T(s_y^2, u^*, v^*)$, where the function $T(\bullet)$ satisfies $T(S_y^2, 1, 1) = S_y^2$ and $T_1(S_y^2, 1, 1) = 1$ with $T_1(\bullet)$ denoting the first partial derivative of $T(\bullet)$ with respect to $s_y^2$, the minimum MSE of $d_4$ is equal to (3.4); subsequently, it is not smaller than that of $d_3$. The difference type estimator

$$d_(4(1)) = s_y^2 u^* \alpha + (1 - \alpha)v^* \beta$$

is a member of the class $d_4$ but not of $d_3$. Putting $t_1(1, 1) = 0$ and $t_2(1, 1) = -1$ in (3.2), we obtain the MSE of the Singh and Joarder (1998) estimator $d_{3(0)}$ to the first order of approximation as

$$\text{MSE}(d_{3(0)}) = \theta S_y^4[\lambda_{40} + \lambda_{04} - 2\lambda_{22}].$$  

(3.7)

To determine an estimator of the minimum MSE of $d_3$, we make use of the following lemma given in Singh and Joarder (1998) and Singh et al. (2000).
Lemma 3.1. A maximum likelihood estimator of the probability of non-response \( p \) is given by

\[
\hat{p} = \frac{(n - 1 + r) - \sqrt{(n - 1 + r)^2 - 4rn(n - 3)/(n - 2)}}{2(n - 3)}. \tag{3.8}
\]

If \( r = 0 \) then \( \hat{p} = 0 \), and if \( r = (n - 2) \) then \( \hat{p} = 1 \); thus \( \hat{p} \) is an admissible estimator of response probability \( p \).

Theorem 3.2. An estimator of the minimum \( \text{MSE}(d_3) \) is given by

\[
\min \hat{\text{MSE}}(d_3) = \hat{\theta}^* s^*_p \left( \lambda^*_v - 1 - D \right),
\]

where

\[
\lambda^*_v = \frac{\bar{\mu}^*_v}{(\bar{\mu}^*_v)^2 + (\bar{\mu}^*_2)^2}, \tag{3.9}
\]

\[
\bar{\mu}^*_v = (n - r - 1)^{-1} \sum_{i=1}^{n-r} (y_i - \bar{y})'(x_i - \bar{x})^*, \quad (l, s) = 0, 1, 2, 3, 4, \tag{3.10}
\]

\[
D = \lambda^*_2 + \frac{\left( \lambda^*_2 \lambda^*_v - \lambda^*_2 + 1 \right)^2}{\lambda^*_4 - \lambda^*_2} \quad \text{and} \quad \hat{\theta}^* = \left( \frac{1}{n\hat{q} + 2\hat{p}} - \frac{1}{N} \right).
\]

### Bias in the estimator

To obtain the bias of \( d_3 \), we assume that the third partial derivatives of \( t(u^*, v^*) \) also exist and are continuous and bounded. Expanding \( t(u^*, v^*) \) about \((1, 1)\) to third order and taking expectations, we obtain up to terms of order \( n^{-1} \)

\[
B(d_3) = \hat{\theta}^* \left( \frac{S^2}{2} \right) \left[ 2\lambda^*_2 C_x t_{11}(1, 1) + 2(\lambda^*_2 - 1)t_{21}(1, 1) + C^2_x t_{11}(1, 1) + 2\lambda^*_0 C_x t_{12}(1, 1) + (\lambda^*_0 - 1)t_{22}(1, 1) \right], \tag{3.11}
\]

where \( t_{ij}(1, 1), (i, j) = 1, 2 \) denote the second order partial derivatives of \( t(u^*, v^*) \). The bias and \( \text{MSE} \) of an estimator that belong to the proposed family \( d_3 \) can be easily obtained from (3.11) and (3.2).

Theorem 3.3. If \( t_{11}(1, 1) = 2t^2_{11}(1, 1), t_{12}(1, 1) = 2t_{11}(1, 1)t_{21}(1, 1) \) and \( t_{22}(1, 1) = 2t^2_{11}(1, 1) \), then the subsequent estimator that belongs to the family \( d_3 \) would be an asymptotically optimum unbiased estimator \((AOUE)\) with approximate variance formula given by (3.4).

The results of Theorem 3.3 hold true for the estimator \( d_{3(4)} \). The bias of \( d_{3(4)} \) is zero for optimum values of \( \alpha \) and \( \beta \).

### Estimators with estimated optimum parameters

In practice, optimum values \( \hat{A} \) and \( \hat{B} \) of \( t_{11}(1, 1) \) and \( t_{21}(1, 1) \) are rarely known. Consistent estimators of \( t_{11}(1, 1) \) and \( t_{21}(1, 1) \) are given by

\[
\hat{t}_{11}(1, 1) = \hat{A} \quad \text{and} \quad \hat{t}_{21}(1, 1) = \hat{B}, \tag{3.12}
\]
where $\hat{A}$ and $\hat{B}$ are determined by replacing $\lambda_d$ in (1.5) by $\bar{\lambda}_s^*$ given in (3.9), which is obtained using $\hat{\mu}_s^*$ in (3.10). To develop a family of estimators $d_5^*$ and associated MSE’s analogous to the class $d_3$ when optimum values are unknown, the regularity conditions for $d_3$ and (3.3) suggest that a function $t(u^*,v^*)$ is required such that $t(1,1) = t_1(1,1) = \partial t(\bullet)/\partial u^*|(1,1) = A$, and $t_2(1,1) = \partial t(\bullet)/\partial v^*|(1,1) = B$, which indicates that the function $t(\bullet)$ will contain not only $u^*$ and $v^*$ but $A$ and $B$ as well. Thus a function $t^*(u^*,v^*,A,B)$ is needed such that $t^*(1,1,A,B) = 1$, $t_1(1,1,A,B) = \partial t^*(\bullet)/\partial u^*|(1,1,A,B) = A$, and $t_2(1,1) = \partial t^*(\bullet)/\partial v^*|(1,1,A,B) = B$. Since $A$ and $B$ are not known, we may take $t^*(u^*,v^*,A,B) = t^*(u^*,v^*,\hat{A},\hat{B})$. Now

$$t^*(Q) = 1, \quad t_1^*(Q) = \frac{\partial t^*(\bullet)}{\partial u^*}|_Q = A, \quad t_2^*(Q) = \frac{\partial t^*(\bullet)}{\partial v^*}|_Q = B,$$

where $Q = (1,1,A,B)$; thus we may consider

$$d_5^* = s_5^*t^*\left(u^*,v^*,\hat{A},\hat{B}\right) \quad (3.13)$$

as an estimator of $S_5^*$. Performing a Taylor series expansion of $t^*(u^*,v^*,\hat{A},\hat{B})$ about $Q$ yields

$$d_5^* = s_5^*\left[1 + (u^*-1)A + (v^*-1)B + \left(\hat{A} - A\right) t_1^*(Q) + \left(\hat{B} - B\right) t_2^*(Q) + \cdots\right], \quad (3.14)$$

where $t_1^*(Q) = \partial t^*(\bullet)/\partial A|_Q = 0$ and $t_2^*(Q) = \partial t^*(\bullet)/\partial B|_Q = 0$ are the first order partial derivatives of $t^*(\bullet)$ with respect to $\hat{A}$ and $\hat{B}$ at $Q = (1,1,A,B)$ respectively. Putting $s_5^2 = S_5^2(1 + e_1, e_1 = (u^*-1), e_2 = (v^*-1)$ in (3.14) we have

$$d_5^* - S_5^2 = S_5^2\left[\epsilon_0 + e_1 A + e_2 B + e_0 e_1 A + e_0 e_2 B + \left(\hat{A} - A\right) t_1^*(Q) + \left(\hat{B} - B\right) t_2^*(Q) + \cdots\right]. \quad (3.15)$$

Squaring both sides of (3.15) and taking expectation, the first degree approximation of the MSE of $d_5^*$ is equal to the minimum MSE of $d_3$ in (3.4) if $t_2^*(Q) = 0$, $t_2^*(Q) = 0$. Therefore, replacing $\alpha$ and $\beta$ by $\hat{A}$ and $\hat{B}$ in the expression for $d_{3(j)}$, $j = 1,2,\ldots,8$ yields a set of estimators, $d_{3(j)}$, of $S_5^2$ that are members of the proposed family $d_5^*$ and that attain the same minimum MSE as given in (3.4).

**Strategy II.** Consider the situation where information on $x$ is available for all sampled units but information on $y$ could not be obtained for $r$ units. If $X$ and $S_5^*$ are known, we define a family of estimators for $S_5^*$ as

$$d_5 = s_5^2 h(u,v), \quad (3.16)$$

where $u = \bar{x}/X$, $v = s_5^2/S_5^2$ and $h(u,v)$ is a function that satisfies certain conditions similar to those for $t$ in $d_3$, and is such that $h(1,1) = 1$. To the first degree of approximation, the bias and MSE of the $d_5$ are given by

$$B(d_5) = B(d_2) \quad (3.17)$$

and

$$\text{MSE}(d_5) = \text{MSE}(d_2) + (\theta^* - \theta)(\lambda_{40} - 1)S_5^4, \quad (3.18)$$

where $B(d_2)$ and $\text{MSE}(d_2)$ are given in (1.3) and (1.4). The bias and MSE of an estimator that
belongs to the family of estimators $d_5$ can be easily obtained from (3.17) and (3.18). The MSE($d_5$) is minimum when

$$h_1(1, 1) = A \quad \text{and} \quad h_2(1, 1) = B,$$

(3.19)

where $A$ and $B$ are defined in (1.7). Thus the minimum MSE of $d_5$ is given by

$$\min \text{MSE}(d_5) = \min \text{MSE}(d_2) + (\theta^* - \theta)(\lambda_{40} - 1)S_y^4,$$

(3.20)

where $\min \text{MSE}(d_2)$ is given in (1.8). Thus, we have the following theorem:

**Theorem 3.4.** *Up to terms of order $n^{-1}$,*

$$\text{MSE}(d_5) \geq \min \text{MSE}(d_2) + (\theta^* - \theta)(\lambda_{40} - 1)S_y^4$$

*with the equality holding if $h_1(1, 1) = A$ and $h_2(1, 1) = B$.*

Any parametric function $h(u, v)$ satisfying the regularity conditions can generate an asymptotically acceptable estimator. Estimators $d_5(j)$ of $S_y^2$ that are particular members of the proposed family $d_5$ can be derived that are identical to $d_3(j)$ in (3.6) except that $u$ and $v$ would be used instead of $u^*$ and $v^*$. It is also easy to show that the $d_5(j)$ will have the same MSE given in (3.20).

Singh and Joarder (1998) suggested an estimator of $S_y^2$ as

$$d_{5(0)} = s_y^2 \frac{S_x^2}{s_x^2},$$

(3.21)

which is a particular case of the proposed family $d_5$. The MSE of $d_{5(0)}$ is obtained by setting $h_1(1, 1) = 0$ and $h_2(1, 1) = -1$ in (3.18) yielding

$$\text{MSE}(d_{5(0)}) = S_y^4 \left[ \theta(\lambda_{40} + \lambda_{22} - 2\lambda_{22}) + (\theta^* - \theta)(\lambda_{40} - 1) \right].$$

(3.22)

Note that the family of estimators in (3.16) does not include simple difference-type estimators such as

$$d_{6(1)} = s_y^2 - w_1 (\bar{x} - \bar{X}) - w_2 \left( s_x^2 - S_x^2 \right)$$

(3.23)

and

$$d_{6(2)} = s_y^2 - \phi_1 \left( s_x^2 - S_x^2 \right) - \phi_2 \left( C_x^2 - C_x^2 \right),$$

(3.24)

where $w_i$ and $\phi_i$, $i = 1, 2$ are constants and $C_x^2 = s_x^2/\bar{x}^2$. However, (3.23) and (3.24) are members of a wider family of estimators defined by

$$d_6 = H \left( s_y^2, u, v \right),$$

(3.25)

where $H(\bullet)$ is a function of $(s_y^2, u, v)$ such that

$$H \left( S_y^2, 1, 1 \right) = S_y^2 \quad \text{and} \quad H_1 \left( S_y^2, 1, 1 \right) = \frac{\partial H(\bullet)}{\partial s_y^2} \bigg|_{(s_y^2, 1, 1)} = 1.$$  

(3.26)

The minimum asymptotic MSE of $d_6$ in (3.25) is equal to (3.20). Thus, for optimum values of
Estimators with estimated optimum parameters

When the optimum values \( A \) and \( B \) in (1.7) are unknown, they may be replaced by \( \hat{A}_1 \) and \( \hat{B}_1 \), where \( \hat{A}_1 \) and \( \hat{B}_1 \) are determined by substituting \( \hat{\lambda}_{2x}^* \) in (3.9) for \( \lambda_{2x} \) and \( \hat{\lambda}_{0x} \) for \( \lambda_{0x} \), where \( \hat{\lambda}_{0x} = \bar{\mu}_0 / \hat{\mu}_{0x}^2 \) with \( \hat{\mu}_{0x} = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \). This yields the estimators

\[
d_{S}^* = s_{i}^2 h \left( u, v, \hat{A}_1, \hat{B}_1 \right),
\]

(3.27)

where \( h^*(u, v, \hat{A}_1, \hat{B}_1) \) is a function of \((u, v, \hat{A}_1, \hat{B}_1)\) such that \( h^*(Q) = 1, h_1(Q) = \partial h^*(\bullet) / \partial u|_Q = A, h_2(Q) = \partial h^*(\bullet) / \partial v|_Q = B, h_3(Q) = \partial h^*(\bullet) / \partial A_1|_Q = 0, \) and \( h_4(Q) = \partial h^*(\bullet) / \partial B_1|_Q = 0 \) for \( Q = (1, 1, A, B) \). Proceeding similarly to \( d_{S}^* \) it can be shown to the first degree of approximation that the MSE of \( d_{S}^* \) is equivalent to the minimum MSE of \( d_{S} \) given by (3.20). Estimators \( d_{S}^{(j)} \) of \( S_{y}^2 \) that are members of the proposed family \( d_{S}^* \) can be computed using estimated optimum values by simply replacing \( \alpha \) and \( \beta \) in \( d_{S}^{(j)} \) by \( \hat{A}_1 \) and \( \hat{B}_1 \). Note that it can easily be verified to the first order of approximation that the estimators \( d_{S}^{(j)} \) of \( S_{y}^2 \) that are members of the proposed family \( d_{S}^* \) obtain the same minimum MSE given by (3.20).

**Theorem 3.5.** An estimator of the min. MSE(\( d_{S}^* \)) (or MSE(\( d_{S}^* \))) is given by

\[
\text{min.MSE}(d_{S}^*) = s_{y}^4 \left[ \theta^* (\hat{\lambda}_{40} - 1) - \theta \left( \hat{\lambda}_{21}^2 + \frac{\hat{\lambda}_{21} \hat{\lambda}_{03} - \hat{\lambda}_{22} + 1}{\hat{\lambda}_{04} - \hat{\lambda}_{03}^2 - 1} \right) \right].
\]

(3.28)

**Strategy III.** Consider a nonresponse situation identical to strategy II. However, the population parameters \( \hat{x} \) and \( S_{y}^2 \) are unknown. Here, we suggest a class of estimators of \( S_{y}^2 \) as

\[
d_{T}^* = s_{i}^2 f(w, z),
\]

(3.29)

where \( w = \bar{x}^*/\bar{x}, \) \( z = s_{i}^2 / s_{y}^2 \) and \( f(w, z) \) is a function of \((w, z)\) that satisfies certain conditions similar to \( d_{S} \) in (3.1) and is also such that \( f(1, 1) = 1 \). To the first order of approximation, the bias and MSE of \( d_{T} \) are given by

\[
B(d_{T}) = \frac{\theta^* - \theta}{2} S_{y}^4 \left[ C_{i}^{2} f_{11}(1, 1) + (\lambda_{04} - 1) f_{22}(1, 1) + 2 \lambda_{03} C_{x} f_{12}(1, 1) + 2 \lambda_{21} C_{x} f_{1}(1, 1) + 2(\lambda_{22} - 1) f_{2}(1, 1) \right]
\]

(3.30)

and

\[
\text{MSE}(d_{T}) = S_{y}^4 \left[ \theta^* (\lambda_{04} - 1) + (\theta^* - \theta) \left( C_{i}^{2} f_{11}(1, 1) + (\lambda_{04} - 1) f_{22}(1, 1) + 2 \lambda_{03} C_{x} f_{12}(1, 1) + 2 \lambda_{21} C_{x} f_{1}(1, 1) f_{2}(1, 1) + 2 \lambda_{22} - 1 \right) f_{2}(1, 1) \right],
\]

(3.31)

where \( f_{1}(1, 1), i = 1, 2 \) and \( f_{ij}(1, 1), (i, j) = 1, 2; \) denote the first and second partial derivatives of \( f(w, z) \). The MSE(\( d_{T} \)) is minimum when

\[
f_{1}(1, 1) = A \quad \text{and} \quad f_{2}(1, 1) = B.
\]

(3.32)

Substitution of (3.32) in (3.31) yields the minimum MSE of \( d_{T} \) as

\[
\text{min.MSE}(d_{T}) = \text{min.MSE}(d_{S}) + \theta S_{y}^4 D,
\]

(3.33)
where $D$ and $\text{min.MSE}(d_7)$ are given in (1.9) and (3.4). It is to be noted that the family of estimators $d_7$ in (3.29) is very large. In particular, estimators $d_{7(j)}$ of $S_y^2$ that are members of the family $d_7$ can be obtained by simply replacing $u^*$ and $v^*$ in $d_{3(j)}$ in (3.6) by $w$ and $z$. To the first order of approximation, the minimum MSE of $d_{7(j)}$ is given by (3.33). Thus, we have the following theorems.

**Theorem 3.6.** Up to terms of order $n^{-1}$,

$$\text{MSE}(d_7) \geq S_y^2 [\theta^*(\lambda_{40} - 1) - (\theta^* - \theta) D]$$

with equality holding if $f_1(1, 1) = A$ and $f_2(1, 1) = B$.

**Theorem 3.7.** If $f_{11}(1, 1) = 2f_2(1, 1)$, $f_{12}(1, 1) = 2f_1(1, 1)f_2(1, 1)$ and $f_{22}(1, 1) = 2f_2^2(1, 1)$, then an estimator of the family $d_7$ would be an asymptotically optimum unbiased estimator (AOUE) with approximate variance given by (3.33).

**Theorem 3.8.** An estimator of $\text{min.MSE}(d_7)$ is given by

$$\text{min.MSE}(d_7) = S_y^{4*} \left[ \hat{\theta}^* \left( \hat{\lambda}_{40} - 1 \right) - \left( \hat{\theta}^* - \theta \right) \left\{ \frac{\lambda_{21}^2 \lambda_{10} - \lambda_{22} + 1}{\lambda_{40} - \lambda_{22} + 1} \right\} \right].$$

**Remark 3.1.** It is to be noted that the bias and MSE of an estimator belonging to the family $d_7$ can be easily obtained from (3.30) and (3.31) respectively. To illustrate this, we consider an estimator

$$d_{7(0)} = s_y^2 \left( \frac{s_y^2}{\lambda_{40}} \right) = s_y^2 z^{-1}$$

of $S_y^2$ suggested by Singh and Joarder (1998, p.248). Putting $f_1(1, 1) = 0$, $f_2(1, 1) = -1$, $f_{11}(1, 1) = f_{12}(1, 1) = 0$, $f_{22}(1, 1) = 2$ in (3.30) and (3.31), we obtain the approximate bias and MSE of $d_{7(0)}$ up to terms of order $O(n^{-1})$ as

$$B(d_{7(0)}) = (\theta^* - \theta) S_y^2 (\lambda_{40} - \lambda_{22})$$

and

$$\text{MSE}(d_{7(0)}) = S_y^2 \left[ \theta^* (\lambda_{40} - 1) + (\theta^* - \theta) (\lambda_{40} - 2\lambda_{22} + 1) \right].$$

The expressions for the bias and MSE of $d_{7(0)}$ do not agree with those given in Singh and Joarder (1998) in Equation (3.20) and Equation (3.21), p.248. However, these authors incorrectly evaluated $E(\delta \eta)$ as $\theta(\lambda_{22} - 1)$ instead of $\theta(\lambda_{40} - 1)$ (see Singh and Joarder, 1998, p.243).

**Remark 3.2.** If we consider a wider family of estimators $d_8 = F(s_y^2, w, z)$ of $S_y^2$, where $F(s_y^2, 1, 1) = S_y^2$ and $F_1(S_y^2, 1, 1) = 1$ is the first partial derivative of $F(\bullet)$ with respect to $s_y^2$ about the point $(S_y^2, 1, 1)$, the minimum MSE of $d_8$ is identical to that of $d_7$. The difference type estimator $d_{8(1)} = s_y^2 + \alpha(w - 1) + \beta(z - 1)$ is a member of the family represented by $d_8$, but not of $d_7$.

**Estimators with estimated optimum parameters**

When $A$ and $B$ are unknown, they may be replaced by $\hat{A}_2$ and $\hat{B}_2$, where $\hat{A}_2$ and $\hat{B}_2$ are determined
by substituting \( \hat{\lambda}_a \) in (3.9) for \( \lambda_a \) in (1.5), \( \hat{\lambda}_{0s} = \hat{\mu}_{0s}/\hat{\mu}_{0s}/2 \) with \( \hat{\mu}_{0s} = (n - 1)^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^* \) for \( \lambda_{0s} \) and \( \hat{C}_a = s_a/\bar{x} \) for \( C_a \). If we define a family of estimators (based on estimated optimum values) for \( S_{y}^2 \) as

\[
d_c = s_c^2 f^* (w, z, \hat{A}, \hat{B}) \tag{3.37}
\]

then following an approach identical to that used to derive \( d_c \), we can establish a set of estimators \( d^{*}_{c(j)} \) belonging to the family \( d_c^* \) with MSE properties analogous to those for \( d^{*}_{c(j)} \).


If information on \( x \) is available for all \( n \) units, Singh and Joarder (1998) suggested a family of estimators of \( S_{y}^2 \) as

\[
d_0 = s_0^2 S_0^2 + \alpha \left( \frac{s_0^2}{S_0^2} - 1 \right), \tag{4.1}
\]

where \( \alpha \) is a suitably chosen constant such that the MSE of \( d_0 \) is minimum. To obtain the bias and MSE of \( d_0 \), we express (4.1) as

\[
d_0 - S_y^2 = S_y^2 \{ \epsilon_0 - \epsilon_4 - \epsilon_0 \epsilon_4 + \epsilon_4^2 + \cdots \} + \alpha \epsilon_2. \tag{4.2}
\]

Taking the expectation of (4.2), we obtain a first order approximation of bias as

\[
B(d_0) = \theta S_y^2 (\lambda_{04} - \lambda_{22}). \tag{4.3}
\]

This is same result obtained by Singh and Joarder (1998). To develop an expression for the MSE, we square both sides of (4.2), neglecting terms with \( \epsilon \)'s having power greater than two. Then taking the expectation of both sides and using the results given in Section 2, the MSE of \( d_0 \) to terms of order \( n^{-1} \), is

\[
\text{MSE}(d_0) = S_y^2 [\theta (\lambda_{40} + \lambda_{04} - \lambda_{22}) + (\theta^* - \theta) (\lambda_{04} - 1)]
\]

\[
+ \alpha^2 \theta (\lambda_{04} - 1) + 2 \alpha S_y^2 \theta^* (\lambda_{22} - 1) - \theta (\lambda_{04} - 1) \]

which is minimized for

\[
\alpha = - \frac{S_y^2 [\theta^* (\lambda_{22} - 1) - \theta (\lambda_{04} - 1)]}{\theta^* (\lambda_{04} - 1)}. \tag{4.5}
\]

Thus the resulting minimum MSE of \( d_0 \) is given by

\[
\text{min.MSE}(d_0) = \text{MSE}(d_{5(0)}) - \frac{S_y^2 [\theta^* (\lambda_{22} - 1) - \theta (\lambda_{04} - 1)]^2}{\theta^* (\lambda_{04} - 1)}, \tag{4.6}
\]

where \( \text{MSE}(d_{5(0)}) \) is given in (3.22).

**Theorem 4.1.** An estimator of the min.MSE(\( d_0 \)) is given by

\[
\text{min.MSE}(d_0) = \frac{S_y^4}{\text{MSE}(d_{5(0)})} \left\{ \frac{\theta^* (\hat{\lambda}_{22} - 1) - \theta (\hat{\lambda}_{04} - 1)}{\theta^* (\hat{\lambda}_{04} - 1)} \right\}^2. \tag{4.7}
\]

where

\[
\text{MSE}(d_{5(0)}) = S_y^4 \left[ \theta \left( \hat{\lambda}_{40} + \hat{\lambda}_{04} - 2 \hat{\lambda}_{22} \right) + \left( \theta^* - \theta \right) (\hat{\lambda}_{04} - 1) \right].
\]
REMARK 4.1. Note that the expressions in (4.4), (4.5), (4.6) and (4.7) are correct while the expressions obtained by Singh and Joarder (1998, equations (3.12), (3.13), (3.14) and (3.15); pp. 246–247) are incorrect.

REMARK 4.2. If the optimum value of $\alpha$ is not known, it can be replaced with a consistent estimator $\hat{\alpha}$ which is determined by replacing $S_y^2$, $\theta^*$, $\lambda_{22}$ and $\lambda_{04}$ in (4.5) by $s_y^2$, $\hat{\theta}^*$, $\hat{\lambda}_{22}$ and $\hat{\lambda}_{04}$. Substituting $\hat{\alpha}$ into (4.1) yields the estimator $d_9^*$ for $S_y^2$. Note that it can be easily shown to a first order approximation that

$$\text{MSE}(d_9^*) = \min \text{MSE}(d_9).$$  \hspace{1cm} (4.8)

4.1. A general family of estimators

A generalized version of $d_9$ is proposed as

$$d_{10} = s_y^2 \frac{S_y^2}{s_y^2} g(u^*, v^*) = s_y^2 v^{-1} g(u^*, v^*),$$  \hspace{1cm} (4.9)

where $g(u^*, v^*)$ is a function of $(u^*, v^*)$ that satisfies certain conditions similar to those for $t$ in $d_3$ and is such that $g(1,1) = 1$. To the first degree of approximation, the bias and MSE of $d_{10}$ are given by

$$B(d_{10}) = B(d_{3(0)}) + \frac{s_y^2}{2} \left[ \theta^* \left\{ C_y^2 g_{11}(1,1) + (\lambda_{04} - 1) g_{22}(1,1) + 2\lambda_{03} C_x g_{12}(1,1) + 2\lambda_{23} C_x g_{11}(1,1) \right. 
+ 2(\lambda_{22} - 1) g_{2}(1,1) \right] - 2\theta \left\{ \lambda_{03} C_x g_{1}(1,1) + (\lambda_{04} - 1) g_{2}(1,1) \right\}.$$  \hspace{1cm} (4.10)

and

$$\text{MSE}(d_{10}) = \text{MSE}(d_{3(0)}) + s_y^4 \left[ \theta^* \left\{ C_y^2 g_{11}(1,1) + (\lambda_{04} - 1) g_{22}(1,1) + 2\lambda_{03} C_x g_{12}(1,1) + 2\lambda_{23} C_x g_{11}(1,1) \right. 
+ 2(\lambda_{22} - 1) g_{2}(1,1) \right] - 2\theta \left\{ \lambda_{03} C_x g_{1}(1,1) + (\lambda_{04} - 1) g_{2}(1,1) \right\}.$$  \hspace{1cm} (4.11)

The MSE($d_{10}$) is minimised for

$$g_1(1,1) = A \quad \text{and} \quad g_2(1,1) = \frac{\theta}{\theta^*} + B.$$  \hspace{1cm} (4.12)

Thus the resulting (minimum) MSE of $d_{10}$ is given by

$$\min \text{MSE}(d_{10}) = \min \text{MSE}(d_3) + \frac{\theta(\theta^* - \theta)}{\theta^*} s_y^4 (\lambda_{04} - 1).$$  \hspace{1cm} (4.13)

Thus we have the following theorems.

Theorem 4.2. Up to terms of order $n^{-1}$,

$$\text{MSE}(d_{10}) \geq \min \text{MSE}(d_3) + \frac{\theta(\theta^* - \theta)}{\theta^*} s_y^4 (\lambda_{04} - 1).$$  \hspace{1cm} (4.14)

Theorem 4.3. An estimator of the min. MSE($d_{10}$) is given by

$$\min \text{MSE}(d_{10}) = \min \text{MSE}(d_3) + \frac{\theta(\theta^* - \theta)}{\theta^*} s_y^4 (\lambda_{04} - 1),$$  \hspace{1cm} (4.15)
Remark 4.5.

It can easily be shown for a first order of approximation that

\[ \text{MSE}(d_{10}) = s_t^2 v^{-1}(u^*)^2 \]

where \( \text{min.MSE}(d_3) \) is given in Theorem 3.2.

Some particular members of the family of estimators \( d_{10} \) are \( d_{10(1)} = s_t^2 v^{-1}(u^*)^2, \)

\[ d_{10(2)} = s_t^2 v^{-1}(1 + \alpha(u - 1)) (1 + \beta(v - 1)) \]

\[ d_{10(3)} = s_t^2 v^{-1}(1 + \alpha(u - 1) + \beta(v - 1)) \]

\[ d_{10(4)} = s_t^2 v^{-1}(1 - \alpha(u - 1) - \beta(v - 1))^{-1} \]

\[ d_{10(5)} = s_t^2 v^{-1}(2 - (u^*)^2) \]

\[ d_{10(6)} = s_t^2 v^{-1}(1 + \gamma(u^*)^2 \beta) \]

\[ d_{10(7)} = s_t^2 v^{-1} \exp(\alpha(u - 1) + \beta(v - 1)) \]

\[ d_{10(8)} = s_t^2 v^{-1} v^* \]

and so on.

Remark 4.3. A family wider than (4.9) is defined by

\[ d_{11} = G\left[ s_t^2, v, u^*, v^* \right] \]

where \( G(\bullet) \) is a function of \( (s_t^2, v, u^*, v^*) \) such that \( G(S_y^2, 1, 1, 1) = S_y^2, G_1(S_y^2, 1, 1, 1) = 1, \)

\( G_2(S_y^2, 1, 1, 1) = -1 \) and \( G_22(S_y^2, 1, 1, 1) = 2 \) where \( G_1(\bullet), G_2(\bullet) \) and \( G_22(\bullet) \) are the first and the second partial derivatives of \( G(s_t^2, u^*, v^*) \). The MSE of \( d_{11} \) is the same as that of \( d_{10} \) given by (4.13). Note that the estimator \( d_9 \) is a special member of \( d_{11} \) but not of \( d_{10} \).

Remark 4.4. Let

\[ \hat{g}_1(1, 1) = \hat{A}_3 = \hat{A}_2 \quad \text{and} \quad \hat{g}_2(1, 1) = \hat{B}_3 = \frac{\theta}{\theta^*} + \hat{B}_2 \]

be the consistent estimators of the optimum values of \( g_1(1, 1) \) and \( g_2(1, 1) \) in (4.12). Then, we define a family of estimators (based on estimated optimum values) of \( S_y^2 \) as

\[ d_{10} = s_t^2 v^{-1} g^* (u^*, v^*, \hat{A}_3, \hat{B}_3) \]

where \( g^*(u^*, v^*, \hat{A}_3, \hat{B}_3) \) is a function of \( (u^*, v^*, \hat{A}_3, \hat{B}_3) \) such that \( g^*(S) = 1, g_1^*(S) = \partial g^*(\bullet)/\partial u^*|_S = A, g_2^*(S) = \partial g^*(\bullet)/\partial v^*|_S = B + \theta/\theta^*, g_3^*(S) = \partial g^*(\bullet)/\partial \hat{A}_3|_S = 0, \) and \( g_4^*(S) = \partial g^*(\bullet)/\partial \hat{B}_3|_S = 0 \)

for \( S = (1, 1, \hat{A}_3, \hat{B}_3) \). It can be easily proved for a first order approximation that

\[ \text{MSE}(d_{10}) = \text{min.MSE}(d_{10}) \]

Estimators \( d_{10}^{(1)} \) of \( S_y^2 \) can be obtained by replacing \( \alpha, \beta, \gamma, u \) and \( v \) with \( \hat{A}_3, \hat{B}_3, \gamma, u^* \) and \( v^* \). It can be easily shown that the MSE of \( d_{10}^{(1)} \) is equal to that of \( d_{10} \)

Remark 4.5. One may also suggest the following families of estimators of \( S_y^2 \) as

\[ d_{11} = s_t^2 D(u^*, v^*, u, v) \]

and

\[ d_{12} = P(s_t^2, u^*, v^*, u, v) \]

where \( D(\bullet) \) and \( P(\bullet) \) are the functions of \( (u^*, v^*, u, v) \) and \( (s_t^2, u^*, v^*, u, v) \) such that \( D(1, 1, 1, 1) = 1, P(S_y^2, 1, 1, 1, 1) = S_y^2, P_3(S_y^2, 1, 1, 1, 1) = 1 \) and \( D(\bullet) \) and \( P(\bullet) \) satisfy certain regularity conditions. It can easily be shown for a first order of approximation that

\[ \text{min.MSE}(d_{11}) = \text{min.MSE}(d_{12}) = \text{min.MSE}(d_3), \]

where \( \text{min.MSE}(d_3) \) is given by (3.4).
5. Efficiency Comparison

It is well known that

\[ V \left( s_y^2 \right) = \theta^* S_y^2 (\lambda_{40} - 1). \]  \hfill (5.1)

From (3.4), (3.7) and (5.1) we have

\[ V \left( s_y^2 \right) - \text{min.MSE}(d_3) = \theta^* S_y^2 \left[ \frac{\lambda_{21} + (\lambda_{21} - \lambda_{22} + 1)^2}{\lambda_{04} - \lambda_{03} - 1} \right] > 0 \]  \hfill (5.2)

and

\[ \text{MSE}(d_{3(0)}) - \text{min.MSE}(d_3) = \theta^* S_y^2 \left[ \frac{(\lambda_{04} - \lambda_{22})^2}{\lambda_{04} - 1} + \Theta \right] > 0, \]  \hfill (5.3)

where \( \Theta = \{(\lambda_{04} - 1)\lambda_{21} - (\lambda_{22} - 1)\lambda_{03}\}^2/(\lambda_{04} - \lambda_{03} - 1) \). It follows from (5.2) and (5.3) that the proposed family \( d_3 \) or \( d_5 \) is more efficient than the conditional unbiased estimator \( s_y^2 \) and the Singh and Joarder (1998) estimator \( d_{3(0)} \). Similarly, from (3.20), (3.22) and (5.1) we can show that

\[ V \left( s_y^2 \right) - \text{min.MSE}(d_4) = \frac{\theta}{\theta^*} \left[ V \left( s_y^2 \right) - \text{min.MSE}(d_3) \right] > 0, \]  \hfill (5.4)

\[ \text{MSE} \left( d_{3(1)} \right) - \text{min.MSE}(d_3) = \frac{\theta}{\theta^*} \left[ \text{MSE} \left( d_{3(1)} \right) - \text{min.MSE}(d_3) \right] > 0. \]  \hfill (5.5)

We note from (5.4) and (5.5) that the proposed family \( d_3 \) (or \( d_5 \)) is more efficient than \( s_y^2 \) and \( d_{3(0)} \), and from (3.33), (3.36) and (5.1) that

\[ V \left( s_y^2 \right) - \text{min.MSE}(d_7) = \frac{\theta - \theta^*}{\theta^*} \left[ V \left( s_y^2 \right) - \text{min.MSE}(d_3) \right] > 0 \]  \hfill (5.6)

and

\[ \text{MSE} \left( d_{7(0)} \right) - \text{min.MSE}(d_7) = \frac{\theta - \theta^*}{\theta^*} \left[ \text{MSE} \left( d_{3(1)} \right) - \text{min.MSE}(d_3) \right] > 0, \]  \hfill (5.7)

which shows that the suggested family \( d_7 \) is better than \( s_y^2 \) and \( d_{7(0)} \). Finally, using (3.4), (3.7), (3.20), (3.22), (3.33), (4.6) and (4.15) we have

\[ \text{min.MSE}(d_5) - \text{min.MSE}(d_3) = \frac{\theta - \theta^*}{\theta^*} \left[ \text{MSE}(d_{3(1)}) - \text{min.MSE}(d_3) \right] > 0, \]  \hfill (5.8)

\[ \text{min.MSE}(d_7) - \text{min.MSE}(d_3) = \frac{\theta - \theta^*}{\theta^*} \left[ V \left( s_y^2 \right) - \text{min.MSE}(d_3) \right] > 0, \]  \hfill (5.9)

\[ \text{min.MSE}(d_{10}) - \text{min.MSE}(d_3) = \frac{(\theta - \theta^*)^2}{\theta^*} S_y^2 (\lambda_{04} - 1) > 0, \]  \hfill (5.10)

\[ \text{min.MSE}(d_9) - \text{min.MSE}(d_{10}) = \frac{\theta^*}{\lambda_{04} - 1} S_y^2 \Theta > 0, \]  \hfill (5.11)

\[ \text{MSE} \left( d_{5(0)} \right) - \text{min.MSE}(d_9) = \frac{S_y^2}{(\lambda_{04} - 1)\theta^*} (\theta^* (\lambda_{22} - 1) - \theta (\lambda_{04} - 1))^2 > 0. \]  \hfill (5.12)

From (5.8) to (5.12), we have the following inequalities:

\[ \text{min.MSE}(d_3) \leq \text{min.MSE}(d_5), \]  \hfill (5.13)

\[ \text{min.MSE}(d_3) \leq \text{min.MSE}(d_7), \]  \hfill (5.14)

\[ \text{min.MSE}(d_3) \leq \text{min.MSE}(d_{10}) \leq \text{min.MSE}(d_9) \leq \text{min.MSE}(d_{5(0)}). \]  \hfill (5.15)

which implies that the proposed family \( d_3 \) (or \( d_5 \)) is the best among all the estimators discussed in this paper.
6. A Numerical Illustration

Suppose that a bank selected a simple random sample of twenty states without replacement from the USA during 1997 and collected information (in thousands) on real \((y)\) and nonreal estate farm loans \((x)\). The selected states are CA, CT, FL, IL, ME, MS, MO, NE, NJ, NM, ND, OK, SC, TN, TX, UT, VA, WA, WV and WI. For detail of the data set, please see population-1 on page 111 in Singh (2003). However, assume the information on the real estate farm loans was not available for four states ME, ND, TX and VA. Let us apply the ratio type estimator \( \hat{\nu} = \frac{s_y^2}{s_x^2} \) for estimating the finite population variance of the real estate farm loans in the United States and construct a 75% confidence interval. From the Table 3.1 for \( n = 20 \) and \( r = 4 \) we have \( \hat{\nu} = 0.233555 \) and \( \hat{\lambda} = 0.766445 \).

From the responding units in the sample we have \( \bar{x}_{n-r} = 1123.62 \), \( \bar{y}_{n-r} = 735.77 \), \( \hat{\lambda}_{20} = s_y^2 = 432778.2 \), \( \hat{\lambda}_{40} = \mu_{40} = \frac{s_y^2}{s_x^2} = 3.8922 \times 10^{11} \) and \( \hat{\mu}_{40} = 1148.986 \), \( \hat{\mu}_{50} = 1690181.84 \), \( \hat{\mu}_{02} = 7.46411 \times 10^{12} \), \( \hat{\lambda}_{04} = \mu_{04}/\hat{\mu}_{02} = 2.6128 \), \( \hat{\lambda}_{04} = \mu_{04}/\hat{\mu}_{02} = 2.0781 \), and \( \hat{\lambda}_{22} = \hat{\mu}_{22}/(\hat{\mu}_{02} \hat{\mu}_{02}) = 1.1775 \).

Thus, an estimate of the finite population variance of the real estate farm loans is given by

\[
\hat{\nu} = \frac{s_y^2}{s_x^2} = 450706.50
\]

and an estimate of the MSE(\(\hat{\nu}\)) is given by

\[
\text{MSE}(\hat{\nu}) = \left[ \left( \frac{1}{nq + 2p} - \frac{1}{N} \right) \left( \hat{\lambda}_{40} - 1 \right) + \left( \frac{1}{nq + 2p} - \frac{1}{n} \right) \left( \hat{\lambda}_{04} - 2\hat{\lambda}_{22} - 1 \right) \right] s_y^4 = 3.4938 \times 10^{10}.
\]

The \((1 - \alpha)100\%\) confidence interval for the finite population variance is given by

\[
\hat{\nu} \pm t_{2.5 \%, \df = n-r-2} \sqrt{\text{MSE}(\hat{\nu})}.
\]

Therefore, the 75% confidence interval of the finite population variance of the real estate farm loans is given by:

\[
\hat{\nu} \pm t_{0.25 \%, \df = 20-4-2} \sqrt{\text{MSE}(\hat{\nu})} \quad \text{or} \quad \text{MSE}(181545.88, 719867.11). \]

Note that if we divide the original dataset by 100, then the 75% CI estimate for the finite population variance will be given by \([18.15, 71.98]\).

Acknowledgements

We are grateful to the Editor, the Associate Editor and two learned referees for valuable comments on the original manuscript.

References


