مجله‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **4** (2007): 175-189

دوره‌ی ۴، شماره‌ی ۲، پاییز و زمستان ۱۳۸۶، صص ۱۷۵-۱۸۹

Effect of Measurement Errors on a Class of Estimators of Population Mean Using Auxiliary Information in Sample Surveys

Housila P. Singh and Namrata Karpe*

Vikram University

Abstract. We consider the problem of estimating the population mean of the study variate Y in presence of measurement errors when information on an auxiliary character X is known. A class of estimators for population mean using information on an auxiliary variate X is defined. Expressions for its asymptotic bias and mean square error are obtained. Optimum conditions are obtained for which the mean square errors of the proposed class of estimators are minimum.

Keywords. study variate; auxiliary variate; measurement errors; bias; mean squared error.

1 Introduction

The standard theory of survey sampling usually assumes that we observe the "true values" when data are collected. In practice such a supposition may not be tenable and the data may contain observational or measurement errors due to various reasons; see Cochran (1968), Sukhatme et al. (1984), Fuller (1995), Srivastava and Shalabh (2001), Sud and Srivastava (2000), Allen et al. (2003), and Singh and Karpe (2008, 2009). For more interesting details and illustrations readers are referred to Beimer et al. (2004), Groves et al. (2004)

* Corresponding author

and Lessler (1992). In many sample surveys it is recognized that errors of measurement can arise from the person being interviewed, from the interviewer, from the supervisor or leader of team of interviewer, and from the processor who transmits the information from the recorded interview on to the punched cards or tapes that will be analyzed (see Cochran, 1968, p. 638). When two sampling strategies are compared, it is customary to compare them on the basis of their mean square errors calculated on the assumption that there are no errors of measurement. But such a comparison may be misleading when measurement errors are present. The reason is that measurement errors may introduce bias and variance. Hence there is a need to take the contribution of measurement errors into account when making such a comparison (see Chandhok and Han, 1990).

It is well known that the use of auxiliary information in sample surveys which is related to the survey variable results in substantial increase in the accuracy of the estimators of population mean or total. Out of many ratio, product and regression methods of estimation are good examples in this context. When the population mean μ_X of the auxiliary variate X is known, a large number of estimator for population mean or total of the study variate Y have been proposed and their properties are studied assuming that response variable measured without error.

Let (Y, X) be the study variate and the auxiliary variate respectively. For a simple random sample of size n, let (y_i, x_i) be the pair of values instead of the true values (Y_i, X_i) on the two variates (Y, X) respectively for the ith $(i = 1, 2, \ldots, n)$ in the sample.

The observational or measurement errors are defined as

$$u_i = (y_i - Y_i)\,, \tag{1}$$

$$v_i = (x_i - X_i)\,, \tag{2}$$

which are assumed to be stochastic with mean 'zero' but possibly with different variances σ_U^2 and σ_V^2.

For the sake of simplicity in exposition, we assume that (i) u_i's and v_i's are uncorrelated although X_i's and Y_i's $(i = 1, 2, \ldots, n)$ are correlated, (ii) u_i's and v_i's are uncorrelated with true observation X_i's and Y_i's $(i = 1, 2, \ldots, n)$, (iii) the population is large enough so that finite population correction terms can be ignored.

Let (μ_Y, μ_X) and $\left(\sigma_Y^2, \sigma_X^2\right)$ be population means and variances of (Y, X), respectively. Further, let ρ be the correlation coefficient between Y and X. We also denote:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i : \text{ the mean of sample observations on } x,$$

$$s_X^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2 : \text{the sample variance for } X,$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i : \text{ the mean of sample observations on } Y,$$

$$C_Y = \frac{\sigma_Y}{\mu_Y}, \qquad C_X = \frac{\sigma_X}{\mu_X},$$

$$\gamma_{1X} = \sqrt{\beta_1(X)} = \frac{\mu_3(X)}{\sigma_x^3}, \qquad \gamma_{2X} = (\beta_2(X) - 3),$$

$$\beta_2(X) = \frac{\mu_4(X)}{\sigma_X^4}, \qquad \lambda = \frac{\mu_{12}(Y,X)}{\sigma_Y^2\sigma_X^2},$$

$$\gamma_{1V} = \sqrt{\beta_1(v)} = \frac{\mu_3(v)}{\sigma_V^3}, \qquad \gamma_{2V} = \beta_2(v) - 3,$$

$$\beta_2(v) = \frac{\mu_4(v)}{\sigma_V^4}, \qquad \mu_4(x) = E(X - \mu_x)^4,$$

$$\mu_3(x) = E(X - \mu_x)^3, \qquad \mu_3(v) = E(v^3),$$

$$\mu_4(v) = E(v^4), \qquad \mu_{12}(X,Y) = E\{(X_i - \mu_X)(Y_i - \mu_Y)^2\}.$$

Let $\hat{\sigma}_x^2 = s_X^2 - \sigma_V^2 > 0$ is unbiased estimator of population variance σ_X^2 of the auxiliary variate X.

Furthermore we define

$$a = \frac{\bar{x}}{\mu_X}, \quad b = \frac{\hat{\sigma}_x^2}{\sigma_X^2}, \quad \varepsilon_1 = a - 1, \quad \varepsilon_2 = b - 1, \quad \varepsilon_0 = \frac{\bar{y}}{\mu_Y} - 1,$$

Therefore

$$E(\varepsilon_0) = E(\varepsilon_1) = E(\varepsilon_2) = 0,$$

$$E(\varepsilon_0^2) = \frac{C_Y^2}{n}\left(1 + \frac{\sigma_U^2}{\sigma_Y^2}\right), \quad E(\varepsilon_1^2) = \frac{C_X^2}{n}\left(1 + \frac{\sigma_V^2}{\sigma_X^2}\right), \quad E(\varepsilon_0\varepsilon_1) = \frac{\rho C_Y C_X}{n},$$

and to the first degree of approximation,

$$E(\varepsilon_2^2) = \frac{1}{n}\left\{\gamma_{2X} + \gamma_{2V}\frac{\sigma_V^4}{\sigma_X^4} + 2\left(1 + \frac{\sigma_V^2}{\sigma_X^2}\right)^2\right\},$$

$$E(\varepsilon_0\varepsilon_2) = \left(\frac{1}{n}\right)\lambda C_Y,$$

$$E(\varepsilon_1\varepsilon_2) = \frac{C_X}{n}\left(\gamma_{1X} + \gamma_{1V}\frac{\sigma_V^3}{\sigma_X^3}\right).$$

The objective of this paper is to suggest a class of estimators of the population mean μ_Y of the study variate Y using information on the known population mean μ_X, population variance σ_X^2 and the known error variance σ_V^2 in the presence of measurement errors. The class of estimators considered is very large and the estimators proposed by Shalabh (1997) and Manisha and Singh (2001) are members of the suggested class of estimators. Expression for the bias and mean square error of the estimator, up to the terms of order n^{-1} has been obtained. Conditions for an estimator to be optimum in the sense of having smallest mean square error upto terms of the order n^{-1} in the class and the corresponding minimum mean square error have been obtained. The results hold for simple random sampling with replacements also. When there are no measurement errors in both the variates X and Y, the results of this paper reduce to Srivastava and Jhajj (1981).

2 The Suggested Class of Estimators

Let the error variance σ_V^2 be known a priori. Then the unbiased estimator of the population variance σ_X^2 is $\hat{\sigma}_x^2 = s_X^2 - \sigma_V^2 > 0$; see for example, Madansky (1959), Birch (1964), Barnett (1967), Schneeweiss (1976), Fuller (1987, Ch.1), Cheng and Van Ness (1991, 1994), and Srivastava and Shalabh (1996, 1997). We also assume that the measurement errors are independent of the true value of variables.

Whatever be the sample chosen, let (a, b) assume values in a bounded, closed convex subset, Q on the two dimensional real space containing the point $(1, 1)$.

Following the approach similar to that adopted by Srivastava (1971, 1980) and Srivastava and Jhajj (1980, 1981), we consider the class of estimators of the population mean μ_Y of the study variate Y defined by

$$d_h = \bar{y}h(a, b), \tag{3}$$

where $h(a, b)$ is a function of a and b such that

$$h(1, 1) = 1 \tag{4}$$

and such that it satisfies the following conditions:

(i) The function $h(a, b)$ is continuous and bounded in Q.

(ii) The first and second order partial derivatives of $h(a,b)$ exist and are continuous and bounded in Q.

Expanding the function $h(a,b)$ about the point $(1,1)$ in a second order Taylor's series, we obtain

$$d_h = \bar{y}\Big[h(1,1) + (a-1)h_1(1,1) + (b-1)h_2(1,) + \frac{1}{2}\{(a-1)^2 h_{11}(a^*,b^*) + 2(a-1)(b-1)h_{12}(a^*,b^*) + (b-1)^2 h_{22}(a^*,b^*)\}\Big],$$

where $a^* = 1+\eta(a-1)$, $b^* = 1+\eta(v-1)$, $0<\eta<1$, and

$$h_1(1,1) = \left.\frac{\partial h(\cdot)}{\partial a}\right|_{(1,1)}, \quad h_2(1,1) = \left.\frac{\partial h(\cdot)}{\partial b}\right|_{(1,1)}, \quad h_{11}(a^*,b^*) = \left.\frac{\partial^2 h(\cdot)}{\partial a^2}\right|_{(a^*,b^*)},$$

$$h_{12}(a^*,b^*) = \left.\frac{\partial^2 h(\cdot)}{\partial a \partial b}\right|_{(a^*,b^*)}, \qquad h_{22}(a^*,b^*) = \left.\frac{\partial^2 h(\cdot)}{\partial b^2}\right|_{(a^*,b^*)}.$$

Substituting for $\bar{y}$, a, and b in terms of $\varepsilon_0, \varepsilon_1$, and ε_2, where $\varepsilon_0 = (\bar{y}/\mu_y)-1$, $\varepsilon_1 = (a-1)$ and $\varepsilon_2 = (b-1)$ respectively, we have

$$d_h = \mu_y(1+\varepsilon_0)\Big[1 + \varepsilon_1 h_1(1,1) + \varepsilon_2 h_2(1,1) + \frac{1}{2}\left\{\varepsilon_1^2 h_{11}(a^*,b^*) + 2\varepsilon_1\varepsilon_2 h_{12}(a^*,b^*) + \varepsilon_2^2 h_{22}(a^*,b^*)\right\}\Big],$$

$$d_h = \mu_y\Big[1 + \varepsilon_0 + \varepsilon_1 h_1(1,1) + \varepsilon_2 h_2(1,1) + \varepsilon_0\varepsilon_1 h_1(1,1) + \varepsilon_0\varepsilon_2 h_2(1,1) + \frac{1}{2}(1+\varepsilon_0)\left\{\varepsilon_1^2 h_{11}(a^*,b^*) + 2\varepsilon_1\varepsilon_2 h_{12}(a^*,b^*) + \varepsilon_2^2 h_{22}(a^*,b^*)\right\}\Big]. \tag{5}$$

Taking expectation and noting that the second order partial derivatives of the function $h(a,b)$ are bounded, we obtain

$$E(d_h) = \mu_Y + O(n^{-1}),$$

which shows that the bias is of the order n^{-1}, and hence the contribution to the order of n^{-2}. The mean square error (MSE) of the proposed class of estimator

d_h upto the first degree of approximation is given by

$$\begin{aligned}
\text{MSE}(d_h) &= E(d_h - \mu_Y)^2 \\
&= \mu_y^2 E[\varepsilon_0 + \varepsilon_1 h_1(1,1) + \varepsilon_2 h_2(1,1)]^2 \\
&= \mu_y^2 E\left[\varepsilon_0^2 + \varepsilon_1^2 h_1^2(1,1) + \varepsilon_2^2 h_2^2(1,1) + 2\varepsilon_0\varepsilon_1 h_1(1,1) + 2\varepsilon_0\varepsilon_2 h_2(1,1)\right. \\
&\quad \left. + 2\varepsilon_1\varepsilon_2 h_1(1,1) h_2(1,1)\right] \\
&= \frac{\mu_Y^2}{n}\left[C_Y^2\left(1 + \frac{\sigma_U^2}{\sigma_Y^2}\right) + C_x^2\left(1 + \frac{\sigma_V^2}{\sigma_X^2}\right) h_1^2(1,1) + \left\{\gamma_{2X} + \gamma_{2V}\frac{\sigma_V^4}{\sigma_X^4}\right.\right. \\
&\quad \left. + 2\left(1 + \frac{\sigma_V^2}{\sigma_X^2}\right)^2\right\} h_2^2(1,1) + 2\rho C_X C_Y h_1(1,1) + 2C_X\left(\gamma_{1X}\right. \\
&\quad \left.\left. + \gamma_{1V}\frac{\sigma_V^3}{\sigma_X^3}\right) h_1(1,1) h_2(1,1) + 2\lambda C_Y h_2(1,1)\right] \\
&= \frac{\mu_Y^2}{n}\left\{C_Y^2(1 + p_Y) + C_X^2(1 + p_X) h_1^2(1,1) + A h_2^2(1,1)\right. \\
&\quad \left. + 2\rho C_X C_Y h_1(1,1) + 2\lambda C_Y h_2(1,1) + 2C_X B h_1(1,1) h_2(1,1)\right\}, \qquad (6)
\end{aligned}$$

where

$$p_Y = \frac{\sigma_U^2}{\sigma_Y^2}, \qquad p_X = \frac{\sigma_V^2}{\sigma_X^2},$$

$$A = [\gamma_{2X} + \gamma_{2V} p_X^2 + 2(1 + p_X)^2], \qquad B = [\gamma_{1X} + \gamma_{1V} p_X^{3/2}].$$

The MSE of d_h at (6) is minimum when

$$\begin{cases} h_1(1,1) = \dfrac{(\lambda B - \rho A) C_Y}{\{A(1 + p_X) - B^2\} C_X} = \delta_0 \text{ (say)} \\[2ex] h_2(1,1) = \dfrac{\{\rho B - \lambda(1 + p_X)\} C_Y}{\{A(1 + p_X) - B^2\}} = \gamma_0 \text{ (say)} \end{cases} \qquad (7)$$

Thus the resulting minimum mean squared error (MSE) of the class of estimators d_h is given by

$$\text{min.MSE}(d_h) = \frac{\sigma_y^2}{n}\left\{(1 + p_Y) - \frac{\rho^2 A + \lambda^2(1 + p_X) - 2\lambda\rho B}{A(1 + p_X) - B^2}\right\}. \qquad (8)$$

Thus we state the following theorem.

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

$$\mathrm{MSE}(d_h) \geqslant \frac{\sigma_y^2}{n}\left\{(1+p_Y) - \frac{\rho^2 A + \lambda^2(1+p_X) - 2\lambda\rho B}{A(1+p_X) - B^2}\right\},$$

with equality holding if $h_1(1,1) = \delta_0$ (say) and $h_2(1,1) = \gamma_0$ (say), where δ_0 and γ_0 are given by (7).

The minimum MSE of d_h at (8) re-expressed as

$$\min.\mathrm{MSE}(d_h) = \frac{\sigma_Y^2}{n}\left\{(1+p_Y) - \frac{\rho^2}{1+p_X} - \frac{\{\rho B - \lambda(1+p_X)\}^2}{(1+p_X)\{A(1+p_X) - B^2\}}\right\} \quad (9)$$

$$= \frac{\sigma_Y^2}{n}\left\{(1+p_Y) - \frac{\lambda^2}{A} - \frac{(\lambda B - \rho A)^2}{A\{A\,(1+p_X) - B^2\}}\right\} \quad (10)$$

In (9) (or (10)), the first term on the right hand side gives the asymptotic MSE of class d_h when no information is used. The second term in (9) gives the reduction in asymptotic MSE when population mean of the auxiliary variate X is only used and the third term provides the amount of reduction when population mean μ_X is used along with σ_X^2 and σ_V^2. Further we note that the second reduction in asymptotic MSE when population variance σ_X^2 and the error variance σ_V^2 are used.

The class of estimators d_h in (3) is very large if the parameters in function $h(a,b)$ are so chosen that they satisfy (7), the resulting estimator will have MSE given by (9) (or (10)).

The following functions, for example, yield some simple estimators of the class:

(i) $h(a,b) = a^{\alpha} b^{\beta}$,

(ii) $h(a,b) = \alpha_1 a^{\alpha} + \alpha_2 b^{\beta}, \quad \alpha_1 + \alpha_2 = 1$,

(iii) $h(a,b) = \{1 - \alpha\,(a-1) - \beta\,(b-1)\}$,

(iv) $h(a,b) = \frac{1+\alpha_1(a-1)}{1+\beta(b-1)}$,

(v) $h(a,b) = \{1 - \alpha_1\,(a-1) - \beta\,(b-1)\}^{-1}$,

(vi) $h(a,b) = \exp\{\alpha\,(a-1) - \beta\,(b-1)\}$,

(vii) $h(a,b) = \alpha\, a + (1-\alpha)\, b^{\beta}$,

(viii) $h(a,b) = (1-\alpha)\ a + \alpha b^{\beta}$, etc. where $\alpha_i (i = 1, 2)$, α, and β are suitably chosen constants. The optimum values of these constants are determined from the conditions (7) and with these optimum values the asymptotic MSE is given by (8).

In the special case when (X, Y) follows a bivariate normal and $u_i \sim N(0, \sigma_U^2)$ and $v_i \sim N(0, \sigma_V^2)$, one would have

$$h_1(1,1) = -\frac{\rho C_Y}{(1+p_X)C_X} = -K(1+p_X)^{-1} \tag{11}$$

and

$$h_2(1,1) = 0 \tag{12}$$

and thus the resulting minimum MSE is given by

$$\text{min.MSE}(d_h) = \frac{\sigma_Y^2}{n}\left\{(1+p_Y) - \frac{\rho^2}{1+p_X}\right\}, \tag{13}$$

where $K = \rho(C_Y/C_X)$.

Thus it follows from (11) that with the knowledge of the parameters (K, p_X) one would be able to generate a large number of estimators which has smaller asymptotic MSE than conventional unbiased estimator $\bar{y}$. The value of the parameter K is known to the experimenter, as it is stable overtime; see Murthy (1967, p.325), Walsh (1970), and Reddy (1973, 1974).

If there is no measurement errors in both the variables X and Y, then the minimum MSE in (9) reduces to:

$$\begin{aligned} \text{min.MSE}(d_h) &= \frac{\sigma_Y^2}{n}\left\{1 - \rho^2 - \frac{(\rho\gamma_{1X} - \lambda)^2}{\gamma_{2X} - \gamma_{1X}^2 + 2}\right\} \qquad (14) \\ &= \frac{\sigma_Y^2}{n}\left\{1 - \frac{\rho^2}{\gamma_{2X}+2} - \frac{\{\lambda\gamma_{1X} - \rho(\gamma_{2X}+2)\}^2}{(\gamma_{2X}+2)(\gamma_{2X} - \gamma_{1X}^2 + 2)}\right\}, \end{aligned}$$

which is same as obtained by Srivastava and Jhajj (1981).

In case of bivariate normal population, the minimum MSE of d_h at (14) reduces to

$$\text{min.MSE}(d_h) = \frac{\sigma_Y^2}{n}(1-\rho^2), \tag{15}$$

which equals to the approximate variance of the usual regression estimator. From (9) (or (10)) and (14) (or (15)) it is observed that sampling variability increases when measurement errors are present. Thus we conclude that the minimum MSE of the class of estimators reported by Srivastava and Jhajj (1981) increases in presence of measurement errors. This fact is also true in case of bivariate normal population too; see expressions (13) and (15).

It may be noted that the minimum mean square error (8) (or (9) or (10)) is attained only when the optimum values of the parameters, which are functions of the unknown population values λ, ρ, C_Y, $C_X, \gamma_{1X}, \gamma_{1V}, \gamma_{2X}$, and γ_{2V} are used. When the optimum values of $h_1(1,1)$ and $h_2(1,1)$ yields solution in the form of constants equated to some parametric function then it may be possible to use the optimum values by using past experience or the past data regarding parameters or by estimating the parameters involved in the optimum value of constant using the sample values (see Reddy, 1978). It has been shown that upto the terms of order n^{-1}, the minimum value of the mean squared error of the estimator is unchanged if we estimate the optimum values of the constants by using the sample values (see Srivastava and Jhajj, 1983). Further it may be noted that even if the values of the parameters used in the estimator are not exactly equal to their optimum values as given by (7) but if they are close enough, the resulting estimator will be better than the conventional estimator $\bar{y}$, as has been illustrated by Das and Tripathi (1978, sec. 3) in case of estimation of population variance σ_Y^2 more discussion on a similar point may be found in Srivastava (1966).

3 Bias of the Proposed Class of Estimators d_h

To obtain the bias of the estimator d_h, we will have to strengthen the conditions on $h(a,b)$ of section 2 supposing that its third order partial derivatives also exist and are continuous and bounded. Expanding $h(a,b)$ about the point $(1,1)$ in a third order Taylor's series, we have

$$\begin{aligned}
d_h &= \bar{y}\Bigg[h(1,1) + (a-1)h_1(1,1) + (b-1)h_2(1,1) + \frac{1}{6}\left\{(a-1)\frac{\partial}{\partial a}\right.\\
&\quad \left. + (b-1)\frac{\partial}{\partial b}\right\}^3 h(a^*,b^*) + \frac{1}{2}\{(a-1)^2 h_{11}(1,1) + 2(a-1)(b-1)h_{12}(1,1)\\
&\quad + (b-1)^2 h_{22}(1,1)\}\Bigg]\\
&= \mu_Y(1+\varepsilon_0)\Bigg[1 + \varepsilon_1 h_1(1,1) + \varepsilon_2 h_2(1,1) + \frac{1}{2}\{\epsilon_1^2 h_{11}(1,1) + 2\varepsilon_1\varepsilon_2 h_{12}(1,1)\\
&\quad + \varepsilon_2^2 h_{22}(1,1)\} + \frac{1}{6}\left\{\varepsilon_1\frac{\partial}{\partial a} + \varepsilon_2\frac{\partial}{\partial b}\right\}^3 h(a^*,b^*)\Bigg]\\
&= \mu_Y\Bigg[1 + \varepsilon_0 + \varepsilon_1 h_1(1,1) + \varepsilon_2 h_2(1,1) + \varepsilon_0\varepsilon_1 h_1(1,1) + \varepsilon_0\varepsilon_2 h_2(1,1)\\
&\quad + (1+\varepsilon_0)\left\{\frac{1}{2}\left(\varepsilon_1^2 h_{11}(1,1) + 2\varepsilon_1\varepsilon_2 h_{12}(1,1) + \varepsilon_2^2 h_{22}(1,1)\right)\right.
\end{aligned}$$

$$+\frac{1}{6}\left(\varepsilon_1\frac{\partial}{\partial a}+\varepsilon_2\frac{\partial}{\partial b}\right)^3 h(a^*,b^*)\Big\}\Big]$$

or

$$(d_h-\mu_Y)\cong\mu_Y\Big[\varepsilon_0+\varepsilon_1 h_1(1,1)+\varepsilon_2 h_2(1,1)+\varepsilon_0\varepsilon_1 h_1(1,1)+\varepsilon_0\varepsilon_2 h_2(1,1)$$
$$+\frac{1}{2}\left\{\varepsilon_1^2 h_{11}(1,1)+2\varepsilon_1\varepsilon_2 h_{12}(1,1)+\varepsilon_2^2 h_{22}(1,1)\right\}\Big].$$

Taking expectation of both sides of the above expression we get the bias of d_h to the first degree of approximation as

$$\begin{aligned}B(d_h)&=\frac{\mu_Y}{2n}\{C_X^2(1+p_X)h_{11}(1,1)+2C_XBh_{12}(1,1)+Ah_{22}(1,1)\\&\quad+2\rho C_YC_Xh_1(1,1)+2\lambda C_Yh_2(1,1)\} \qquad (16)\\&=\frac{\mu_Y}{2n}\{C_X^2(1+p_X)h_{11}(1,1)+C_XBh_{12}(1,1)+2\rho C_YC_Xh_1(1,1)\\&\quad+Ah_{22}(1,1)+C_XBh_{12}(1,1)+2\lambda C_Yh_2(1,1)\}. \qquad (17)\end{aligned}$$

Now let

$$\begin{aligned}h_{11}(1,1)&=2h_1^2(1,1),\\h_{22}(1,1)&=2h_2^2(1,1),\\h_{12}(1,1)&=2h_1(1,1)h_2(1,1).\end{aligned}$$

Then

$$\begin{aligned}B(d_h)&=\frac{\mu_Y}{n}[h_1(1,1)\{C_X^2(1+p_X)h_1(1,1)+C_XBh_2(1,1)+\rho C_YC_X\}\\&\quad+h_2(1,1)\{Ah_2(1,1)+C_XBh_1(1,1)+\lambda C_Y\}]. \qquad (18)\end{aligned}$$

The bias of d_h in (18) is zero if

$$\begin{cases}C_X(1+p_X)h_1(1,1)+Bh_2(1,1)=-\rho C_Y,\\Ah_2(1,1)+C_XBh_1(1,1)=-\lambda C_Y.\end{cases} \qquad (19)$$

Solving (19) we get

$$h_1(1,1)=\frac{C_y(\lambda B-\rho A)}{C_x\{A(1+p_X)-B^2\}}=\delta_0, \qquad (20)$$

$$h_2(1,1)=\frac{C_y\{\rho B-\lambda(1+p_X)\}}{A(1+p_X)-B^2}=\gamma_0, \qquad (21)$$

where δ_0 and γ_0 are given by (7). Thus we state the following theorem.

Theorem 2. *Under the conditions*

$$h_{11}(1,1) = 2h_1^2(1,1),$$

$$h_{22}(1,1) = 2h_2^2(1,1),$$

$$h_{12}(1,1) = 2h_1(1,1)h_2(1,1),$$

along with $h(1,1) = 1$*, the function* $h(a,b)$ *will generate unbiased as well as minimum variance bound estimator for the population mean* μ_Y*.*

Example 1. To illustrate we consider an estimator for μ_Y as

$$d_1 = \bar{y}\{1 + \alpha_1(a-1) + \alpha_2(b-1)\}^{-1},$$

where (α_1, α_2) are suitable chosen constants. Here

$$h(a,b) = \{1 + \alpha_1 (a-1) + \alpha_2 (b-1)\}^{-1}. \tag{22}$$

For this function we have

$$h_1(1,1) = -\alpha_1, \qquad h_2(1,1) = -\alpha_2,$$

$$h_{11}(1,1) = 2\alpha_1^2, \qquad h_{22}(1,1) = 2\alpha_2^2, \qquad h_{12}(1,1) = 2\alpha_1\alpha_2,$$

which imply that

$$h_{11}(1,1) = 2h_1^2(1,1), \quad h_{22}(1,1) = 2h_2^2(1,1), \quad h_{12}(1,1) = 2h_1(1,1)h_2(1,1).$$

Thus the function $h(a,b)$ given by (22) satisfies all the conditions stated in Theorem 2. Thus to the first degree of approximation the bias and MSE of d_1 are respectively given by

$$\begin{aligned} B(d_1) = \frac{\mu_Y}{n}\{&C_X^2(1+p_X)\alpha_1^2 + 2C_X B\alpha_1\alpha_2 - \rho C_Y C_X \alpha_1 \\ &+ A\alpha_2^2 - \alpha_2\lambda C_Y\}, \end{aligned} \tag{23}$$

$$\begin{aligned} \mathrm{MSE}(d_1) = \frac{\mu_Y^2}{n}\{&C_Y^2(1+p_Y) + C_X^2(1+p_X)\alpha_1^2 + A\alpha_2^2 + 2C_X B\alpha_1\alpha_2 \\ &- 2\rho C_Y C_X\alpha_1 - 2\alpha_2\lambda C_Y\}. \end{aligned} \tag{24}$$

The MSE (d_1) is minimized for

$$\begin{cases} \alpha_1 = -\delta_0, \\ \alpha_2 = -\gamma_0. \end{cases} \tag{25}$$

Substitution of (25) in (23) yields

$$B(d_1) = 0.$$

Putting (25) in (24) we get the minimum MSE of d_1 as

$$\text{min.MSE}\,(d_1) = \text{min.MSE}\,(d_h)\,,$$

where min.MSE$\,(d_h)$ is given by (8). Thus the estimator d_1 is almost unbiased at optimum values of α_1 and α_2 has the least MSE (or variance) equals to the min.MSE$\,(d_h)$ given by (8).

4 A Wider Class of Estimators

The class of estimators d_h at (3) does not include the difference type estimator

$$d = \bar{y} - \lambda_1(\bar{x} - \mu_X) - \lambda_2\left(s_X^2 - \sigma_X^2\right),$$

where λ_1 and λ_2 are suitably chosen scalars. However, it is easily shown that if we consider a class of estimators wider than (3), defined by

$$d_H = H(\bar{y}, a, b) \tag{26}$$

of population mean μ_Y, where $H(\cdot)$ is a function of $\bar{y}$, a, and b such that

$$H\left(\mu_Y, 1, 1\right) = \mu_Y, \tag{27}$$

the minimum asymptotic mean square error of d_h is equal to (8) and is not reduced.

Remark 1. Shalabh (1997), Manisha and Singh (2001), and Singh and Karpe (2008) have respectively suggested the following estimators:

$$t_R = \bar{y}\left(\frac{\mu_X}{\bar{x}}\right),$$
$$t_{R1} = \theta\bar{y}\left(\frac{\mu_X}{\bar{x}}\right) + (1-\theta)\bar{y},$$

and

$$t_{RP} = w\bar{y}\left(\frac{\mu_X}{\bar{x}}\right) + (1-w)\bar{y}\left(\frac{\bar{x}}{\mu_X}\right),$$

where θ and w are suitably chosen constants.

It can be seen that the estimators t_R, t_{R1}, and t_{RP} are particular members of the suggested classes of estimators $d_h(d_H)$ given by (3) (or (26)). The mean squared errors and biases of t_R, t_{R1}, and t_{RP} can be obtained first by putting the suitable values of the derivatives in (6) and (16) respectively.

Acknowledgment

Authors are thankful to the Associate Editor and the learned referees for their valuable suggestions regarding improvement of the paper.

References

Allen, J., Singh, H.P., and Smarandache, F. (2003). A family of estimators of population mean using multiauxiliary information in presence of measurement errors. *Int. J. Social. Economics* **30**, 837-849.

Biermer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S. (2004). *Measurement Errors in Surveys,* Wiley, New York.

Barnett, V.D. (1967). A note on linear structural relation ships when both residuals variances are known. *Biometrika* **54**, 670-672.

Birch, M.W. (1964). A note on the maximum likelihood estimation of a linear structural relationship. *J. Amer. Statist. Assoc.* **59**, 1175-1178.

Chandhok, P.K. and Chein, P. Han. (1990). On the efficiency of the ratio estimator under mid zuno scheme with measurement errors. *J. Indian Statist. Assoc.* **28**, 31-39.

Cheng, C-L. and Van Ness, J.W. (1991). On the unreplicated ultra structural model. *Biometrika* **78**, 442-445.

Cheng, C.L. and Van Ness, J.W. (1994). On estimating linear relationships when both variables are subject to errors. *J. Roy. Statist. Soc. Ser. B* **56**, 167-183.

Cochran, W.G. (1968). Errors of measurement in statistics. *Technometrics* **10**, 637-666.

Das, A.K. and Tripathi, T.P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhyā Ser. C* **4**, 139-148.

Fuller, W.A. (1987). *Measurement Errors Models.* Wiley, New York.

Fuller, W.A. (1995). Estimation in the presence of measurement error. *International Statist. Review* **63**, 121-147.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to errors. *J. Amer. Stat. Assoc.* **54**, 173-205.

Manisha, and Singh, R.K. (2001). An estimation of population mean in the presence of measurement errors. *J. Indian Soc. Agricultural Statist.* **54**, 13-18.

Murthy, M.N. (1967). *Sampling Theory and Methods.* Statistical Publishing Society, Calcutta.

Reddy, V.N. (1973). On ratio and product methods of estimation. *Sankhyā Ser. B* **35**, 307-316.

Reddy, V.N. (1974). On transformed ratio method of estimation. *Sankhyā Ser. C* **36**, 59-70.

Reddy, V.N. (1978). A study on the use of prior knowledge on certain population parameters in estimation. *Sankhyā Ser. C* **40**, 29-37.

Schneeweines, H. (1976). Consistent estimation of a regression with errors in the variables. *Metrika* **23**, 101-105.

Shalabh (1997). Ratio method of estimation in the presence of measurement errors. *J. Indian Soc. Agricultural Statist.* **52**, 150-155.

Singh, H.P. and Karpe, N. (2008). Ratio-product estimator for population mean in presence of measurement errors. *J. Appl. Statist. Sci.* **16**, 49-64.

Singh, H.P. and Karpe, N. (2009). A class of estimators using auxiliary information for estimating finite population variance in presence of measurements errors. *Comm. Statist. Theory Methods* **38**, 734-741.

Srivastava, A.K. and Shalabh (1996). Properties of a consistent estimation procedure in ultra structural model when reliability ratio is known. *Microelectron. Reliab.* **36**, 1249-1252.

Srivastava, A.K. and Shalabh (1997). Asymptotic efficiency properties of least square estimators in ultra structural model. *Test* **6**, 419-431.

Srivastava, A.K. and Shalabh (2001). Effect of measurement errors on the regression method of estimation in survey sampling. *J. Statist. Res.* **35**, 35-44.

Srivastava, S.K. (1966). A note on Olkin's multivariate ratio estimator. *J. Indian Statist. Assoc.* **4**, 202-208.

Srivastava, S.K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *J. Amer. Statist. Assoc.* **66**, 404-407.

Srivastava, S.K. (1980). A class of estimators using auxiliary information in sample surveys. *Canad. J. Statist.* **8**, 253-254.

Srivastava, S.K. and Jhajj, H.S. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhyā Ser. C* **42**, 87-96.

Srivastava, S.K. and Jhajj, H.S. (1981). A class of estimating of the population mean in survey sampling using auxiliary information. *Biometrika* **68**, 341-343.

Srivastava, S.K. and Jhajj, H.S. (1983). A class of estimators of the population mean using multi auxiliary information. *Bull. Calcutta Statist. Assoc.* **32**, 47-56.

Sud, C. and Srivastava, S.K. (2000). Estimation of population mean in repeat surveys in the presence of measurement errors. *J. Indian Soc. Agricultural Statist.* **53**, 125-33.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S., and Asok, C. (1984). *Sampling Theory of Surveys With Applications.* Iowa State University Press, Ames, IA.

Walsh, J.E. (1970). Generalization of ratio estimate for population total. *Sankhyā Ser. A* **32**, 99-106.

Housila P. Singh
School of Studies in Statistics,
Vikram University,
Ujjain, 456010,
India.
e-mail: *hpsujn@rediffmail.com*

Namrata Karpe
School of Studies in Statistics,
Vikram University,
Ujjain, 456010,
India.
e-mail: *namratakarpe@yahoo.com*