ESTIMATION OF FINITE POPULATION MEAN USING KNOWN CORRELATION COEFFICIENT BETWEEN AUXILIARY CHARACTERS

H.P. Singh, R. Tailor

1. INTRODUCTION

Let $U=\{U_1,U_2,...,U_N\}$ be a finite population of N units. Suppose two auxiliary variables X_1 and X_2 are observed on $U_i(i=1,2,...,N)$, where X_1 is positively and X_2 is negatively correlated with the study variable *Y*. A simple random sample without replacement (SRSWOR) of size *n* with *n* < *N*, is drawn from the population U to estimate $\bar{Y}=\sum_{i=1}^{N} y_i/N$, the population mean of Y, when the population means $\bar{X}_1=\sum_{i=1}^{N} x_{1i}/N$ and $\bar{X}_2=\sum_{i=1}^{N} x_{2i}/N$ of X_1 and X_2 are respectively, known. For estimating $\bar{Y}$, Singh (1967) suggested a ratio-cum-product estimator

$$\hat{\bar{Y}}_1=\bar{y}\left(\frac{\bar{X}_1}{\bar{x}_1}\right)\left(\frac{\bar{x}_2}{\bar{X}_2}\right) \tag{1}$$

where $\bar{y}=\sum_{i=1}^{n} y_i/n$, $\bar{x}_1=\sum_{i=1}^{n} x_{1i}/n$ and $\bar{x}_2=\sum_{i=1}^{n} x_{2i}/n$.

Wide applicability of the estimator $\hat{\bar{Y}}_1$ has led many authors to suggest unbiased versions of $\hat{\bar{Y}}_1$ with their properties, for instance, see Sahoo and Swain (1980), Biradar and Singh (1992-93) and Tracy *et al.* (1998). Sahai and Sahai (1985) and Singh (1987 b) have mentioned that the past association with experimental material might provide a close guess for the correlation coefficient ρ_{yx_1} between study variate Y and auxiliary character X_1 i.e. ρ_{yx_1} can be guessed quite accurately. Recently, Singh and Tailor (2003) have utilized the information on ρ_{yx_1} and suggested a modified ratio estimator for $\bar{Y}$ with its properties. Further Singh and Singh (1984) advocated that the correlation coefficient $\rho_{x_1x_2}$ between auxiliary variates X_1 and X_2 may be known in many practical situations and hence utilizing the known value of $\rho_{x_1x_2}$ suggested a class of estimators for population variance σ_y^2 of Y with its properties. This led authors to suggest modified ratio-cum-product estimator using $\rho_{x_1x_2}$ with its properties.

A jackknife version of the suggested estimator $\hat{\bar{Y}}_2$ is also given and its properties are studied. An empirical study is carried out in support of the proposed estimator.

2. SUGGESTED RATIO-CUM-PRODUCT ESTIMATOR

Assuming that the correlation coefficient $\rho_{x_1x_2}$ between auxiliary characters X_1 and X_2 is known, we define a ratio-cum-product estimator of $\bar{Y}$ as

$$\hat{\bar{Y}}_2 = \bar{y}\left(\frac{\bar{X}_1 + \rho_{x_1x_2}}{\bar{x}_1 + \rho_{x_1x_2}}\right)\left(\frac{\bar{x}_2 + \rho_{x_1x_2}}{\bar{X}_2 + \rho_{x_1x_2}}\right). \tag{2}$$

To the first degree of approximation, the bias and mean square error (MSE) of the proposed estimator $\hat{\bar{Y}}_2$ are respectively, given by

$$B(\hat{\bar{Y}}_2) = \theta \bar{Y}[\mu_1^* C_{x_1}^2(\mu_1^* - K_{yx_1}) + \mu_2^* C_{x_2}^2(K_{yx_2} - \mu_1^* K_{x_1x_2})] \tag{3}$$

and

$$MSE(\hat{\bar{Y}}_2) = \theta \bar{Y}^2[C_y^2 + \mu_1^* C_{x_1}^2(\mu_1^* - 2K_{yx_1}) + \mu_2^* C_{x_2}^2\{\mu_2^* + 2(K_{yx_2} - \mu_1^* K_{x_1x_2})\}], \tag{4}$$

where

$$K_{yx_1} = \rho_{yx_1}(C_y / C_{x_1}),\ K_{yx_2} = \rho_{yx_2}(C_y / C_{x_2}),\ K_{x_1x_2} = \rho_{x_1x_2}(C_{x_1} / C_{x_2}),$$
$$\mu_i^* = \bar{X}_i / (\bar{X}_i + \rho_{x_1x_2}),\ i = (1,2);\ \theta = \left(\frac{1}{n} - \frac{1}{N}\right),\ C_y = S_y / \bar{Y},$$
$$C_{x_i} = S_{x_i} / \bar{X}_i, (i = 1,2);\ \rho_{y\,x_i} = S_{yx_i} / (S_y S_{x_i}), (i = 1,2);$$

$$S_y^2 = \sum_{j=1}^{N}(y_i - \bar{Y})^2 / (N-1),\ S_{x_i}^2 = \sum_{j=1}^{N}(x_{ij} - \bar{X}_i)^2 / (N-1), (i = 1,2)$$

and

$$S_{yx_i} = \sum_{j=1}^{N}(y_i - \overline{Y})(x_{ij} - \overline{X}_i)/(N-1), (i=1,2).$$

When no auxiliary information is used the estimator $\hat{\overline{Y}}_2$ reduces to the conventional unbiased estimator $\overline{y}$. If the information only on auxiliary variate X_1 is used, then the estimator $\hat{\overline{Y}}_2$ tends to the usual ratio estimator $\overline{y}_R = \overline{y}(\overline{X}_1/\overline{x}_1)$. On the other hand if the information is available on auxiliary variate X_2 only, $\hat{\overline{Y}}_2$ reduces to the usual product estimator $\overline{y}_P = \overline{y}(\overline{x}_2/\overline{X}_2)$.

It is well known that sample mean $\overline{y}$ is an unbiased estimator of $\overline{Y}$ and its variance under SRSWOR sampling scheme is given by

$$V(\overline{y}) = \theta \overline{Y}^2 C_y^2. \tag{5}$$

To the first degree of approximation, the biases and MSEs of $\overline{y}_R$, $\overline{y}_P$ and $\hat{\overline{Y}}_1$ are respectively given by

$$B(\overline{y}_R) = \theta \overline{Y} C_{x_1}^2 (1 - K_{yx_1}), \tag{6}$$

$$B(\overline{y}_P) = \theta \overline{Y} C_{x_2}^2 K_{yx_2}, \tag{7}$$

$$B(\hat{\overline{Y}}_1) = \theta \overline{Y}[C_{x_1}^2 (1 - K_{yx_1}) + C_{x_2}^2 (K_{yx_2} - K_{x_1x_2})], \tag{8}$$

$$MSE(\overline{y}_R) = \theta \overline{Y}^2 [C_y^2 + C_{x_1}^2 (1 - 2K_{yx_1})], \tag{9}$$

$$MSE(\overline{y}_P) = \theta \overline{Y}^2 [C_y^2 + C_{x_2}^2 (1 + 2K_{yx_2})], \tag{10}$$

and

$$MSE(\hat{\overline{Y}}_1) = \theta \overline{Y}^2 [C_y^2 + C_{x_1}^2 (1 - 2K_{yx_1}) + C_{x_2}^2 \{1 + 2(K_{yx_2} - K_{x_1x_2})\}]. \tag{11}$$

3. EFFICIENCY COMPARISIONS

It follows from (4), (5), (9), (10) and (11) that

(i) $MSE(\overline{y}_R) < V(\overline{y})$ if

$$K_{yx_1} > \frac{1}{2} \tag{12}$$

(ii) $MSE(\bar{y}_P) < V(\bar{y})$ if

$$K_{yx_2} < -\frac{1}{2} \tag{13}$$

(iii) $MSE(\hat{\bar{Y}}_1) < V(\bar{y})$ if

$$[C_{x_1}^2(1-2K_{yx_1}) + C_{x_2}^2\{1+2(K_{yx_2} - K_{x_1x_2})\}] < 0$$

which is always true if

$$K_{yx_1} > \frac{1}{2} \quad \text{and} \quad K_{yx_2} < \left(K_{x_1x_2} - \frac{1}{2}\right) \tag{14}$$

(iv) $MSE(\hat{\bar{Y}}_2) < V(\bar{y})$ if

$$[C_{x_1}^2\mu_1^*(\mu_1^* - 2K_{yx_1}) + C_{x_2}^2\mu_2^*\{\mu_2^* + 2(K_{yx_2} - \mu_1^* K_{x_1x_2})\}] < 0$$

which always holds if

$$K_{yx_1} > \frac{\mu_1^*}{2} \quad \text{and} \quad K_{yx_2} < \left(\mu_1^* K_{x_1x_2} - \frac{\mu_2^*}{2}\right) \tag{15}$$

(v) $MSE(\hat{\bar{Y}}_1) < MSE(\bar{y}_R)$ if

$$K_{yx_2} < K_{x_1x_2} - \frac{1}{2} \tag{16}$$

(vi) $MSE(\hat{\bar{Y}}_1) < MSE(\bar{y}_P)$ if

$$K_{yx_1} > -K_{x_2x_1} + \frac{1}{2}, \tag{17}$$

where $K_{x_2x_1} = \rho_{x_1x_2}(C_{x_2}/C_{x_1})$.

(vii) $MSE(\hat{\bar{Y}}_2) < MSE(\bar{y}_R)$ if

$$[(1-\mu_1^*)\{2K_{yx_1} - (1+\mu_1^*)\}C_{x_1}^2 + \mu_2^*\{\mu_2^* + 2(K_{yx_2} - \mu_1^* K_{x_1x_2})\}C_{x_2}^2] < 0$$

which is always true if

$$K_{yx_1} < \frac{(1+\mu_1^*)}{2} \quad \text{and} \quad K_{yx_2} < \left(\mu_1^* K_{x_1x_2} - \frac{\mu_2^*}{2}\right) \tag{18}$$

(viii) $MSE(\hat{\bar{Y}}_2) < MSE(\bar{y}_P)$ if

$$[\mu_1^*\{\mu_1^* - 2(K_{yx_1} + \mu_2^* K_{x_2x_1})\}C_{x_1}^2 - (1-\mu_2^*)\{(1+\mu_2^*) + 2K_{yx_2}\}C_{x_2}^2] < 0$$

which always holds if

$$K_{yx_1} > -\mu_2^* K_{x_2x_1} + \frac{\mu_1^*}{2} \quad \text{and} \quad K_{yx_2} > -\frac{(1+\mu_2^*)}{2} \tag{19}$$

and

(ix) $MSE(\hat{\bar{Y}}_2) < MSE(\hat{\bar{Y}}_1)$ if

$$[C_{x_1}^2(1-\mu_1^*)\{2K_{yx_1} - (1+\mu_1^*)\} + C_{x_2}^2\{2K_{x_1x_2}(1-\mu_1^*\mu_2^*) - -(1-\mu_2^*)(1+\mu_2^* + 2K_{yx_2}\}] < 0$$

which is always true if

$$K_{yx_1} < \frac{(1+\mu_1^*)}{2} \quad \text{and} \quad K_{yx_2} > \left[\frac{K_{x_1x_2}(1-\mu_1^*\mu_2^*)}{(1-\mu_2^*)} - \frac{(1+\mu_2^*)}{2}\right]. \tag{20}$$

Now combining (12), (16) and (20) we get that the proposed estimator $\hat{\bar{Y}}_2$ is more efficient than $\bar{y}$, $\bar{y}_R$ and Singh's (1967) estimator $\hat{\bar{Y}}_1$ i.e. $MSE(\hat{\bar{Y}}_2) < MSE(\hat{\bar{Y}}_1) < MSE(\bar{y}_R) < V(\bar{y})$ if

$$\frac{1}{2} < K_{yx_1} < \frac{(1+\mu_1^*)}{2} \quad \text{and} \quad \left[\frac{K_{x_1x_2}(1-\mu_1^*\mu_2^*)}{(1-\mu_2^*)} - \frac{(1+\mu_2^*)}{2}\right] < K_{yx_2} < \left(K_{x_1x_2} - \frac{1}{2}\right). \tag{21}$$

Further combining (20), (17) and (13) we obtained that the suggested estimator $\hat{\bar{Y}}_2$ is more efficient than $\bar{y}$, $\bar{y}_P$ and Singh's (1967) estimator $\hat{\bar{Y}}_1$

i.e. $MSE(\hat{\bar{Y}}_2) < MSE(\hat{\bar{Y}}_1) < MSE(\bar{y}_P) < V(\bar{y})$ if

$$\left(K_{x_2x_1} + \frac{1}{2}\right) < K_{yx_1} < \frac{(1+\mu_1^*)}{2} \quad \text{and} \quad \left[\frac{K_{x_1x_2}(1-\mu_1^*\mu_2^*)}{(1-\mu_2^*)} - \frac{(1+\mu_2^*)}{2}\right] < K_{yx_2} < -\frac{1}{2}. \tag{22}$$

It is to be noted that the suggested estimator $\hat{\bar{Y}}_2$ is biased. In some applications, bias is a major disadvantage. Keeping this in view, we have discussed the unbiasedness of the proposed estimator $\hat{\bar{Y}}_2$, and using the technique suggested by Quenouille (1956) known as 'Jack-knife' technique, proposed a family of almost unbiased estimators with its properties.

4. FAMILY OF UNBIASED ESTIMATORS OF POPULATION MEAN $\overline{Y}$ USING JACKKNIFE TECHNIQUE

Let a simple random sample of size *n* = *gm* drawn without replacement and split at random into *g* sub-samples, each of size *m*. Then we define the Jack-knife ratio-cum-product estimator for population mean $\overline{Y}$ as

$$\hat{\overline{Y}}_{2J}=\frac{1}{g}\sum_{j=1}^{g}\overline{y}_j'\left(\frac{\overline{X}_1+\rho_{x_1x_2}}{\overline{x}_{1j}'+\rho_{x_1x_2}}\right)\left(\frac{\overline{x}_{2j}'+\rho_{x_1x_2}}{\overline{X}_2+\rho_{x_1x_2}}\right) \tag{23}$$

where $\overline{y}_j'=(n\,\overline{y}-m\,\overline{y}_j)/(n-m)$ and $\overline{x}_{ij}'=(n\,\overline{x}_i-m\,\overline{x}_{ij})/(n-m)$, $i=1,2$; are the sample means based on a sample of (*n-m*) units obtained by omitting the j^{th} group and $\overline{y}_j$ and $\overline{x}_{ij}$ $(i=1,2;\quad j=1,2,...,g)$ are the sample means based on the j^{th} sub samples of size *m* = *n*/*g*.

The bias of $\hat{\overline{Y}}_{2J}$, to terms of order n^{-1}, can be easily obtained as

$$B(\hat{\overline{Y}}_{2J})=\frac{(N-n+m)}{N(n-m)}\overline{Y}[\mu_1^*C_{x_1}^2(\mu_1^*-K_{yx_1})+\mu_2^*C_{x_2}^2(K_{yx_2}-\mu_1^*K_{x_1x_2})]\,. \tag{24}$$

From (3) and (24) we have

$$\frac{B(\hat{\overline{Y}}_2)}{B(\hat{\overline{Y}}_{2J})}=\frac{(N-n)(n-m)}{n(N-n+m)} \tag{25}$$

or $$B(\hat{\overline{Y}}_2)=\frac{(N-n)(n-m)}{n(N-n+m)}B(\hat{\overline{Y}}_{2J})$$

or $$B(\hat{\overline{Y}}_2)-\frac{(N-n)(n-m)}{n(N-n+m)}B(\hat{\overline{Y}}_{2J})=0$$

or $$\lambda^*B(\hat{\overline{Y}}_2)-\delta^*\lambda^*B(\hat{\overline{Y}}_{2J})=0 \tag{26}$$

for any scalar λ^*, we have

$$\delta^*=\frac{(N-n)(n-m)}{n(N-n+m)}\,. \tag{27}$$

From (26), we have

$$\lambda^* E(\hat{\bar{Y}}_2 - \bar{Y}) - \delta^* \lambda^* E(\hat{\bar{Y}}_{2J} - \bar{Y}) = 0$$

or $$\lambda^* E(\hat{\bar{Y}}_2 - \bar{y}) - \delta^* \lambda^* E(\hat{\bar{Y}}_{2J} - \bar{y}) = 0$$

or $$E[\lambda^* \hat{\bar{Y}}_2 - \lambda^* \delta^* \hat{\bar{Y}}_{2J} - \bar{y}\{\lambda^*(1-\delta^*) - 1\}] = \bar{Y} .$$

Hence, the general family of almost unbiased ratio-cum-product estimators of $\bar{Y}$ as

$$\hat{\bar{Y}}_{2u} = [\bar{y}\{1 - \lambda^*(1-\delta^*)\} + \lambda^* \hat{\bar{Y}}_2 - \lambda^* \delta^* \hat{\bar{Y}}_{2J}] \tag{28}$$

see Singh (1987 a).

Remark 4.1. For $\lambda^* = 0$, $\hat{\bar{Y}}_{2u}$ yields the usual unbiased estimator $\bar{y}$ while $\lambda^* = (1-\delta^*)^{-1}$, gives an almost unbiased estimator for $\bar{Y}$ as

$$\begin{aligned}\hat{\bar{Y}}^*_{2u} = {} & \frac{(N-n+m)}{N} g\,\bar{y} \left(\frac{\bar{X}_1 + \rho_{x_1x_2}}{\bar{x}_1 + \rho_{x_1x_2}} \right) \left(\frac{\bar{x}_2 + \rho_{x_1x_2}}{\bar{X}_2 + \rho_{x_1x_2}} \right) \\ & - \frac{(N-n)(g-1)}{Ng} \sum_{j=1}^{g} \bar{y}'_j \left(\frac{\bar{X}_1 + \rho_{x_1x_2}}{\bar{x}'_{1j} + \rho_{x_1x_2}} \right) \left(\frac{\bar{x}'_{2j} + \rho_{x_1x_2}}{\bar{X}_2 + \rho_{x_1x_2}} \right)\end{aligned} \tag{29}$$

which is Jack-knifed version of the proposed estimator $\hat{\bar{Y}}_2$.

Many other almost unbiased estimator from (28) can be generated by putting suitable values of λ^*.

5. SEARCH OF AN OPTIMUM ESTIMATOR IN FAMILY $\hat{\bar{Y}}_{2u}$ AT (28)

The family of almost unbiased estimator $\hat{\bar{Y}}_{2u}$ at (28) can be expressed as

$$\hat{\bar{Y}}_{2u} = \bar{y} - \lambda^* \bar{y}_1 , \tag{30}$$

where $\bar{y}_1 = [(1-\delta^*)\bar{y} - \bar{y}_2]$ and $\bar{y}_2 = \hat{\bar{Y}}_2 - \delta^* \hat{\bar{Y}}_{2J}$. The variance of $\hat{\bar{Y}}_{2u}$ is given by

$$V(\hat{\bar{Y}}_{2u}) = V(\bar{y}) + \lambda^{*2} V(\bar{y}_1) - 2\lambda^* Cov(\bar{y}, \bar{y}_1) \tag{31}$$

which is minimized for

$$\lambda^* = Cov(\bar{y}, \bar{y}_1)/V(\bar{y}_1). \tag{32}$$

Substitution of (32) in (31) yields minimum variance of $\hat{\bar{Y}}_{2u}$ as

$$\min.V(\hat{\bar{Y}}_{2u}) = V(\bar{y}) - \frac{\{Cov(\bar{y}, \bar{y}_1)\}^2}{V(\bar{y}_1)}$$
$$= V(\bar{y})(1-\rho_{01}^2), \tag{33}$$

where ρ_{01} is the correlation coefficient between $\bar{y}$ and $\bar{y}_1$.
From (33) it is immediate that

$$\min.V(\hat{\bar{Y}}_{2u}) < V(\bar{y}).$$

To obtain the explicit expression of the variance of $\hat{\bar{Y}}_{2u}$, we write the following results to terms of order n^{-1}, as

$$MSE(\hat{\bar{Y}}_{2J}) = Cov(\hat{\bar{Y}}_2, \hat{\bar{Y}}_{2J}) = MSE(\hat{\bar{Y}}_2) \tag{34}$$

and

$$Cov(\bar{y}, \hat{\bar{Y}}_2) = Cov(\bar{y}, \hat{\bar{Y}}_{2J}) = \theta \bar{Y}^2 [C_y^2 - \mu_1^* \rho_{yx_1} C_y C_{x_1} + \mu_2^* \rho_{yx_2} C_y C_{x_2}] \tag{35}$$

where $MSE(\hat{\bar{Y}}_2)$ is given by (4).

Now using the results from (4), (5) and (35) into (31) we get the variance of $\hat{\bar{Y}}_{2u}$ to the terms of order n^{-1} as

$$V(\hat{\bar{Y}}_{2u}) = \theta \bar{Y}^2 [C_y^2 + \lambda^{*2}(1-\delta^*)^2(\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\rho_{x_1x_2} C_{x_1} C_{x_2} \mu_1^* \mu_2^*)$$
$$-2\lambda^*(1-\delta^*)(\mu_1^* \rho_{yx_1} C_y C_{x_1} - \mu_2^* \rho_{yx_2} C_y C_{x_2})] \tag{36}$$

which is minimized for

$$\lambda^* = \frac{(\mu_1^* \rho_{yx_1} C_y C_{x_1} - \mu_2^* \rho_{yx_2} C_y C_{x_2})}{(1-\delta^*)(\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\mu_1^* \mu_2^* \rho_{x_1x_2} C_{x_1} C_{x_2})} = \lambda_{opt}^*. \tag{37}$$

Substitution of λ_{opt}^* in $\hat{\bar{Y}}_{2u}$ yields the optimum estimator $\hat{\bar{Y}}_{2u(opt)}$ (say). Thus the resulting minimum variance of $\hat{\bar{Y}}_{2u}$ is given by

$$\min.V(\hat{\bar{Y}}_{2u}) = \theta \bar{Y}^2 C_y^2 \left[1 - \frac{(\mu_1^* \rho_{yx_1} C_{x_1} - \mu_2^* \rho_{yx_2} C_{x_2})^2}{(\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\mu_1^* \mu_2^* \rho_{x_1x_2} C_{x_1} C_{x_2})} \right] = V(\hat{\bar{Y}}_{2u(opt)}) . \tag{38}$$

From (4), (11) and (38) we have

$$V(\bar{y}) - \min.V(\hat{\bar{Y}}_{2u}) = \theta \bar{Y}^2 C_y^2 \left[\frac{(\mu_1^* \rho_{yx_1} C_{x_1} - \mu_2^* \rho_{yx_2} C_{x_2})^2}{(\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\mu_1^* \mu_2^* \rho_{x_1x_2} C_{x_1} C_{x_2})} \right] \geq 0 \tag{39}$$

and

$$MSE(\hat{\bar{Y}}_2) - \min.V(\hat{\bar{Y}}_{2u}) =$$
$$= \theta \bar{Y}^2 \left[\frac{(\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\mu_1^* \mu_2^* \rho_{x_1x_2} C_{x_1} C_{x_2} - \rho_{yx_1} C_y C_{x_1} \mu_1^* + \rho_{yx_2} C_y C_{x_2} \mu_2^*)^2}{(\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\mu_1^* \mu_2^* \rho_{x_1x_2} C_{x_1} C_{x_2})} \right] \geq 0. \tag{40}$$

Thus from (39) and (40) we have the following inequalities:

$$\min.V(\hat{\bar{Y}}_{2u}) \leq V(\bar{y}) \tag{41}$$

and

$$\min.V(\hat{\bar{Y}}_{2u}) \leq MSE(\hat{\bar{Y}}_2) \tag{42}$$

which follows that $\hat{\bar{Y}}_{2u}$ with $\lambda^* = \lambda^*_{opt}$ is more efficient than $\bar{y}$ and $\hat{\bar{Y}}_2$.

When λ^* does not coincide with λ^*_{opt} then from (5) and (36) we note that $V(\hat{\bar{Y}}_{2u}) \leq V(\bar{y})$ if

$$\left. \begin{array}{ll} \textit{either} & 0 < \lambda^* < 2\lambda^*_{opt} \\ \textit{or} & 2\lambda^*_{opt} < \lambda^* < 0 \end{array} \right\} \tag{43}$$

It is observed from (11) and (36) that $MSE(\hat{\bar{Y}}_{2u}) < MSE(\hat{\bar{Y}}_1)$ if

$$\frac{B - \sqrt{(B^2 - AC)}}{(1-\delta^*)A} < \lambda^* < \frac{B + \sqrt{(B^2 - AC)}}{(1-\delta^*)A} \ , \tag{44}$$

$$A = (\mu_1^{*2} C_{x_1}^2 + \mu_2^{*2} C_{x_2}^2 - 2\mu_1^* \mu_2^* \rho_{x_1x_2} C_{x_1} C_{x_2}),$$

$$B = (\mu_1^* \rho_{yx_1} C_y C_{x_1} - \mu_2^* \rho_{yx_2} C_y C_{x_2}),$$

$$C = [C_{x_1}^2 (1 - 2K_{yx_1}) + C_{x_2}^2 \{1 + 2(K_{yx_2} - K_{x_1 x_1})\}].$$

We also note from (4) and (36) that the estimator $\hat{\bar{Y}}_{2u}$ is better than $\hat{\bar{Y}}_2 (or \hat{\bar{Y}}_{2u}^*)$ if

$$\left.\begin{array}{ll} \textit{either} & \dfrac{1}{(1-\delta^*)} < \lambda^* < \left[2\lambda_{opt}^* - \dfrac{1}{(1-\delta^*)}\right] \\ \textit{or} & \left[2\lambda_{opt}^* - \dfrac{1}{(1-\delta^*)}\right] < \lambda^* < \dfrac{1}{(1-\delta^*)} \end{array}\right\} \quad (45)$$

The optimum value λ_{opt}^* of λ^* can be obtained quite accurately through past data or experience.

6. EMPIRICAL STUDY

To observe the relative performance of different estimators of $\bar{Y}$, we consider a natural population data set given in Steel and Torrie (1960, p.282). The population description is given below:

y : Log of leaf burn in sec.
x_1 : Potassiam percentage
x_2 : Clorine percentage.

The required population values are:

$\bar{Y} = 0.6860$, $C_y = 0.4803$, $\rho_{yx_1} = 0.1794$, N=30 ,
$\bar{X}_1 = 4.6537$, $C_{x_1} = 0.2295$, $\rho_{yx_2} = -0.4996$, n=6,
$\bar{X}_1 = 0.8077$, $C_{x_2} = 0.7493$, $\rho_{x_1 x_2} = 0.4074$, g=2.

The percentage relative efficiencies (PREs) of various estimators of $\bar{Y}$ with respect to $\bar{y}$ have been computed and presented in Table 1.

TABLE 1

Percent relative efficiencies of different estimators of $\bar{Y}$ with respect to $\bar{y}$

Estimator	$\bar{y}$	$\bar{y}_R$	$\bar{y}_P$	$\hat{\bar{Y}}_1$	$\hat{\bar{Y}}_2(\hat{\bar{Y}}_{2u}^*)$	$\hat{\bar{Y}}_{2u}$ with $\lambda_{opt}^* = 1.19751$
PRE $(\bullet, \bar{y})$	100.00	94.62	53.33	75.50	142.18	165.88

Table 1 clearly indicates that the suggested estimators $\hat{\bar{Y}}_2 (or\, \hat{\bar{Y}}^*_{2u})$ and $\hat{\bar{Y}}_{2u}$ with $\lambda^* = \lambda^*_{opt}$, are more efficient than usual unbiased estimator $\bar{y}$, ratio estimator $\bar{y}_R$, product estimator $\bar{y}_P$, and Singh's (1967) ratio-cum-product estimator $\hat{\bar{Y}}_1$ with considerable gain in efficiency.

7. CONCLUDING REMARKS

Usually information regarding correlation coefficient $\rho_{x_1 x_2}$ between the two auxiliary variates X_1 and X_2 is known or can made known to the experimenter through past studies or with the familiarity of experimental material. When $\rho_{x_1 x_2}$ is known an improved version $\hat{\bar{Y}}_{2u}$ of Singh's (1967) estimator $\hat{\bar{Y}}_1$ is suggested with its properties. Using 'Jack-knife' technique envisaged by Quenouille (1956), a family of unbiased estimators $\hat{\bar{Y}}_{2u}$ is also proposed. A large number of unbiased estimators can be generated from $\hat{\bar{Y}}_{2u}$. Asymptotically optimum estimator (AOE) in the family of estimators $\hat{\bar{Y}}_{2u}$ is identified with its variance formula. It is shown that the suggested family of estimators $\hat{\bar{Y}}_{2u}$ is more efficient than $\bar{y}$ and $\hat{\bar{Y}}_2$ at optimum conditions. Empirical study also suggests that the suggested estimators $\hat{\bar{Y}}_2 (or\, \hat{\bar{Y}}^*_{2u})$ and $\hat{\bar{Y}}_{2u}$ with $\lambda^* = \lambda^*_{opt}$ are better than $\bar{y}$, $\bar{y}_R$, $\bar{y}_P$ and Singh's (1967) estimator $\hat{\bar{Y}}_1$. Thus we conclude that the proposed estimators $\hat{\bar{Y}}_2 (or\, \hat{\bar{Y}}^*_{2u})$ and $\hat{\bar{Y}}_{2u}$ are to be preferred in practice.

School of Studies in Statistics
Vikram University, Ujjain, India HOUSILA P. SINGH

Department of Educational Surveys and Data Processing
New Delhi, India RAJESH TAILOR

ACKNOWLEDGEMENTS

Authors are thankful to the referee for his valuable suggestions regarding improvement of the paper.

REFERENCES

R.S. BIRADAR, H.P. SINGH (1992-93), *Almost unbiased ratio-cum-product estimators*, "Aligarh Journal of Statistics", 12, pp. 13-19.

M.H. QUENOUILLE (1956), *Notes on bias in estimation*, "Biometrika", 43, pp. 353-360.

A. SAHAI, A. SAHAI (1985), *On efficient use of auxiliary information*, "Journal of Statistical Planning and Inference", 12, pp. 203-212.

L.N. SAHOO, A.K.P.C. SWAIN (1980), *Unbiased ratio-cum-product estimator*, "Sankhya", C, 42, pp. 56-62.

H.P. SINGH (1987 a), *Class of almost unbiased ratio and product-type estimators for finite population mean applying Quenouilles method*, "Journal of Indian Society of Agriculture Statistics", 39, pp. 280-288.

H.P. SINGH (1987 b), *On the estimation of population mean when the correlation coefficient is known in two phase sampling*, "Assam Statistical Review", 1, pp. 17-21.

H.P. SINGH, R. TAILOR (2003), *Use of known correlation coefficient in estimating the finite population mean*, "Statistics in Transition", 6, pp. 555-560.

M.P. SINGH (1967), *Ratio-cum-product method of estimation*, "Metrika", 12, pp. 34-42.

R.K. SINGH, G. SINGH (1984), *A class of estimators for population variance using information on two auxiliary variates*, "Aligarh Journal of Statistics", (3-4), pp. 43-49.

D.S. TRACY, H.P. SINGH, R. SINGH (1998), *A class of almost unbiased estimators for finite population mean using two auxiliary variables*, "Biometrical Journal", 40, pp. 753-766.

R.G.D. STEEL, J.H. TORRIE (1960), *Principles and Procedures of Statistics*, Mc Graw Hill Book Co.

RIASSUNTO

Stima della media di un popolazione finita con coefficiente di correlazione tra caratteri ausiliari noto

Il contributo propone uno stimatore *ratio-cum-product* modificato della media di una popolazione finita di una variabile oggetto di studio Y sfruttando il coefficiente di correlazione noto tra due caratteri ausiliari X_1 e X_2. Si ottiene uno stimatore *ratio-cum-product* quasi corretto attraverso la tecnica Jacknife del tipo previsto da Quenille (1956). In seguito vengono esaminati con un esempio numerico i meriti dello stimatore proposto.

SUMMARY

Estimation of finite population mean using known correlation coefficient between auxiliary characters

This paper proposes a modified ratio-cum-product estimator of finite population mean of the study variate Y using known correlation coefficient between two auxiliary characters X_1 and X_2. An almost unbiased ratio-cum-product estimator has also been obtained by using Jackknife technique envisaged by Quenouille (1956). The merits of the proposed estimator are examined through a numerical illustration.