Jour. Ind. Soc. Ag. Statistics
***44** (3), 1992 : 271–285*

A Class of Estimators for Finite Population Correlation Coefficient Using Auxiliary Information

R.S. Biradar and Housila P. Singh
Vikram University, Ujjain - 456 010
(Received : September, 1991)

Summary

A class of estimators is defined for population correlation coefficient ρ utilising the information on population mean and population variance of one of the variables. Asymptotic expressions for bias and mean square error of the class are obtained. It is shown that the proposed class of estimators besides being broader is more efficient than the classes of estimators of Srivastava and Jhajj [7]. Relative efficiency of different estimators is evaluated empirically through a numerical example.

Keywords: Auxiliary information; bias; mean square error; efficient estimator.

Introduction

It is well known that in most of the survey situations use of auxiliary information results in considerable gain in efficiency over estimators which do not use such information. Improved estimators utilising auxiliary information have been developed for population mean and population variance. However, the problem of estimating finite population correlation coefficient ρ between two variables say y and x has not received much attention of survey statisticians so far. The parameter ρ as a measure of the degree of linear relationship between two variables is one of the important parameters.

The correlation coefficient ρ for a finite population of N units with variate values (x_i, y_i), i = 1, 2, . . ., N, is usually defined by

$$\rho = \frac{S_{xy}}{S_x S_y} = \frac{\mu_{11}}{\sqrt{\mu_{20}\,\mu_{02}}} \tag{1.1}$$

Where $S_{xy} = \sum_{i=1}^{N} \frac{(x_i - \overline{X})(y_i - \overline{Y})}{N-1} = \frac{N}{N-1} \mu_{11}$

$$S_x^2 = \sum_{i=1}^{N} \frac{(x_i - \overline{X})^2}{N-1} = \frac{N}{N-1} \mu_{20}$$

$$S_y^2 = \sum_{i=1}^{N} \frac{(y_i - \overline{Y})^2}{N-1} = \frac{N}{N-1} \mu_{02}$$

$$\mu_{rs} = \sum_{i=1}^{N} \frac{(x_i - \overline{X})^r (y_i - \overline{Y})^s}{N}$$

where r and s are non-negative integers.

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

In case no auxiliary (apriori) information is available, an estimator for ρ based on simple random sampling without replacement (SRSWOR) of size n, (x_i, y_i), 1, 2, ..., N, is the conventional sample correlation coefficient r given by

$$r = \frac{s_{xy}}{s_x s_y} \qquad (1.2)$$

where $$s_{xy} = \sum_{i=1}^{n} \frac{(x_i - \overline{x})(y_i - \overline{y})}{n-1}$$

$$s_x^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1}$$

$$\overline{x} = \sum_{i=1}^{n} \frac{x_i}{n}, \quad \overline{y} = \sum_{i=1}^{n} \frac{y_i}{n}$$

Gupta *et. al.* [2] studied the sample correlation coefficient r and obtained asymptotic expressions for bias and variance upto terms of order n^{-1} in terms of sums of powers of x_i and y_i and sums of their products. Srivastava and Jhajj [7] obtained asymptotic bias and variance of r based on bivariate moments of finite population. They also suggested a class of estimators of the population correlation coefficient ρ when information about the population mean and variance of an auxiliary variable is available, as

$$d_1 = r\, t(u,v) \tag{1.3}$$

where t(.) is a function of $u = \frac{\bar{x}}{\bar{X}}$ and $v = \frac{s_x^2}{S_x^2}$ and satisfies certain regularity conditions. Further they defined wider class of estimators of ρ as

$$d_2 = g(r,u,v) \tag{1.4}$$

where g(.) is a function of r, u and v such that g(ρ, 1, 1) = ρ. The common minimum mean square error (MSE) of the estimators d_1 and d_2 under SRSWOR is

$$M_o(d_1) = \rho^2 f_2 \left[KQ + \frac{\lambda_1^* A_1^*}{D^*} + \frac{GA_2^*}{D^*}\right] \tag{1.5}$$

where $K = \frac{(N-1)(Nn-N-n-1)}{(n-1)N(N-3)}$, $\quad M = \frac{N^2n-3N^2+6N-3n-3}{(n-1)N(N-3)}$

$$Q = \left[\frac{\mu_{22}}{\mu_{11}^2} + \frac{1}{4}\{\beta_2(x) + \beta_2(y) + 2l\} - (\lambda_3 + \lambda_4)\right]$$

$$\lambda_1^* = [\lambda_1 - \frac{1}{2}\{\gamma_1(x)\, C(x) + \lambda_2\}]$$

$$D^* = [f_2 C_x^2 \{f_1(K\beta_2(x) - M) - f_2 \gamma_1^2(x)\}]$$

$$A_1^* = f_2^2 [G\gamma_1(x) C_x - \lambda_1^* \{K\beta_2(x) - M\}]$$

$$A_2^* = f_2 C_x [f_2 \lambda_1^* \gamma_1(x) - f_1 GC_x]$$

$$G = \left[(K\lambda_3 - M) - \frac{1}{2}\{(K\beta_2(x) - M) + Kl + 2\beta^* \rho^2 - D\}\right]$$

$$B' = \frac{(N-1)(N-n-1)}{(n-1)N(N-3)}, \qquad D = \frac{N^2n-2nN-N^2+2N-n-1}{(n-1)N(N-3)},$$

$$f_1 = \frac{N-n}{n(N-1)}, \qquad f_2 = \frac{N-n}{n(N-2)},$$

$$\beta_2(x) = \frac{\mu_{40}}{\mu_{20}^2}, \qquad \beta_2(y) = \frac{\mu_{04}}{\mu_{02}^2},$$

$$\gamma_1(x) = \sqrt{\beta_1(x)}, \qquad \beta_1(x) = \frac{\mu_{30}^2}{\mu_{20}^3},$$

$$\lambda_1 = \frac{\mu_{21}}{\bar{X}\mu_{11}}, \qquad \lambda_2 = \frac{\mu_{12}}{\bar{X}\mu_{02}},$$

$$\lambda_3 = \frac{\mu_{31}}{\mu_{20}\,\mu_{11}}, \qquad \lambda_4 = \frac{\mu_{13}}{\mu_{02}\,\mu_{11}};$$

$$C_x = \frac{\sqrt{\mu_{20}}}{\bar{X}}, \qquad l = \frac{\mu_{22}}{\mu_{20}\,\mu_{02}}.$$

Although the classes of estimators d_1 and d_2 suggested by Srivastava and Jhajj [7] are very general, yet they fail to include the following type of estimators of ρ :

$$d_3 = r\left[w_0 + w_1 u^{\alpha_1} + w_2 v^{\alpha_2}\right], \qquad d_4 = r\left[w_0 + w_1 u + w_2 v\right],$$

$$d_5 = r\left[w_0 + w_1 u + w_2 v^{-1}\right], \qquad d_6 = r\left[w_0 + w_1 u^{-1} + w_2 v\right],$$

$$d_7 = r\left[w_0 + w_1 u^{-1} + w_2 v^{-1}\right], \qquad d_8 = r\left[w_0 + w_1 u^{\alpha_1}\right],$$

$$d_9 = r\left[w_0 + w_2 v^{\alpha_2}\right], \qquad d_{10} = r\, w_1 u^{\alpha_1},$$

$$d_{11} = r\, w_2 u^{\alpha_2} \text{ etc}$$

where w_i $(i = 0, 1, 2)$, α_1 and α_2 are suitably chosen scalars.

In this paper an attempt has been made to develop a more efficient class of estimators for ρ, of which the estimators d_i $(i = 1, 2, \ldots, 11)$ are particular cases, when the values of population mean $\bar{X}$ and population variance S_x^2 of an auxiliary variable x are known. The bias and MSE of the proposed class of estimators are derived upto terms of order n^{-1}. Conditions for an estimator to be optimum in the minimum MSE sense upto terms of order n^{-1} in the class and the corresponding minimum MSE have been obtained. An empirical study is carried out to examine the performance of the proposed class of estimators.

2. *Class of Estimators for* ρ

With $u = \frac{\bar{x}}{\bar{X}}$ and $v = \frac{s_x^2}{S_x^2}$ such that E(u) = E(v) = 1, the class of estimators for ρ as

$$\hat{\rho}_h = h(r, u, v) \text{ has been proposed} \tag{2.1}$$

where h(r, u, v) is a function of r, u and v. The function h(r, u, v) is assumed to satisfy the following conditions:

(i) h(r, u, v) is defined in a bounded closed convex subset $S \in R^3$ containing the point $p = (\rho, 1, 1)$

(ii) h(r, u, v) is continuous and bounded in S

(iii) The first and second order partial derivatives of h(r, u, v) denoted by $h_i(p)$ (i = 0, 1, 2) and $h_{ij}(p)$ (i = j = 0, 1, 2) respectively exist and are continuous and bounded in S.

(iv) $h(p) = \rho h_0(p)$,

(v) $h_{01}(p) = a\rho^{-1} h_1(p)$,

(vi) $h_{02}(p) = a\rho^{-1} h_2(p)$,

(vii) $h_{11}(p) = a(\alpha_1 - 1)h_1(p)$,

(viii) $h_{22}(p) = a(\alpha_2 - 1)h_2(p)$,

(ix) $h_{00}(p) = 0$ and $h_{12}(p) = 0$

where α_i's (i = 1, 2) are suitably chosen scalars. They may take for instance values −1 or +1 according as the estimator is ratio or product type. The constant 'a' takes the value 'zero' or 'unity' according as the estimator is difference or ratio/product type. It is to be noted that the class of estimators $\hat{\rho}_h$ proposed in (2.1) reduces to the class of estimators d_2 of Srivastava and Jhajj [7] for $h_0(\rho, 1, 1) = 1$.

To illustrate the above condition we consider the following estimators.

(I) $\hat{\rho}_{h_1} = w_0 r + w_1(u - 1) + w_2(v - 1)$

where w_i's (i = 0, 1, 2) are suitable chosen constants.

We have,

$$\left.\begin{cases} h(p) = \rho w_0 \\ h_0(p) = w_0 \end{cases}\right\} \Rightarrow h(p) = \rho\, h_0(p)$$

$$h_{ij}(p) = 0 \ (i, j = 0, 1, 2)$$

Thus we see that the conditions stated under (iv) to (ix) with a=0 hold good for difference type estimator.

(II) $\hat{\rho}_{h_2} = r(w_0 + w_1 u^{\alpha_1} + w_2 v^{\alpha_2})$

where w_i's (i = 0, 1, 2) are suitably chosen constants.

We have,

$$\left.\begin{cases} h(p) = \rho \sum_{i=0}^{2} w_i \\ h_0(p) = \sum_{i=0}^{2} w_i \end{cases}\right\} \Rightarrow h(p) = \rho h_0(p)$$

$$\left.\begin{cases} h_1(p) = w_1 \alpha_1 \rho \\ h_2(p) = w_2 \alpha_2 \rho \\ h_{01}(p) = w_1 \alpha_1 \\ h_{02}(p) = w_2 \alpha_2 \\ h_{11}(p) = \alpha_1 (\alpha_1 - 1) w_1 \rho \\ h_{22}(p) = \alpha_2 (\alpha_2 - 1) w_2 \rho \\ h_{00}(p) = h_{12}(p) = 0 \end{cases}\right\} \Rightarrow \begin{cases} h_{01}(p) = a\, \rho^{-1} h_1(p) \\ h_{02}(p) = a \rho^{-1} h_2(p) \\ h_{11}(p) = a(\alpha_1 - 1)\, h_1(p) \\ h_{22}(p) = a(\alpha_2 - 1)\, h_2(p) \\ \text{with } a = +1 \end{cases}$$

Thus it is noticed that the conditions stated under (iv) to (ix) with a = + 1 are satisfied by ratio/product type of estimators.

Since there are only a finite number of samples, the expectation and MSE of the estimators $\hat{\rho}_h$ exist under conditions (i) to (iii).

Expanding h(r, u, v) about the point p in a third order Taylor's series, we have

$$\hat{\rho}_h = h(p) + (r-\rho)\, h_0(p) + (u-1)h_1(p) + (v-1)h_2(p) + \frac{1}{2}\Big[(r-\rho)^2\, h_{00}(p)$$

$$+ (u-1)^2 h_{11}(p) + (v-1)^2 h_{22}(p) + (r-\rho)(u-1)h_{01}(p) + 2(r-\rho)(v-1)h_{02}(p)$$

$$+ 2(u-1)(v-1)h_{12}(p)\Big] + \frac{1}{6}\left\{(r-\rho)\frac{\partial}{\partial r} + (u-1)\frac{\partial}{\partial u} + (v-1)\frac{\partial}{\partial v}\right\}^3 h(r^*, u^*, v^*) \quad (2.2)$$

where $r^* = \left[\rho + \theta_0\,(r-\rho)\right]$, $u^* = 1 + \theta_1(u-1)$, $v^* = 1 + \theta_2(v-1)$,

$0 < \theta_i < 1$ $(i = 0, 1, 2)$ and θ_i's may depend on (r, u, v).

Let $(r-\rho) = \rho\, e_0$, $(u-1) = e_1$ and $(v-1) = e_2$ such that $E(e_1) = E(e_2) = 0$, $E(e_o) = \left[\frac{B(r)}{\rho}\right] = RB(r)$ and in view of conditions (iv) to (ix) we obtain from (2.2).

$$\hat{\rho}_h = \left[(1+e_o)\,\rho h_o(p) + \left\{e_1 + a e_o e_1 + \frac{a(\alpha_1-1)e_1^2}{2}\, h_1(p)\right\}\right.$$

$$+ \left\{e_2 + a e_o e_2 + \frac{a(\alpha_2-1)e_1^2}{2}\right\} h_2(p)$$

$$\left. + \frac{1}{6}\left\{P e_o \frac{\partial}{\partial r} + e_1 \frac{\partial}{\partial u} + e_2 \frac{\partial}{\partial v}\right\}^3 h(r^*, u^*, v^*)\right] \quad (2.3)$$

We assume that

$$E(e_o^{q^*}\, e_1^{s^*}\, e_2^{t^*}) = \begin{cases} O(\overline{n}^{1}) & \text{for } q^* + s^* + t^* = 2 \\ O(\overline{n}^{p}),\quad p > 1 & \text{for } q^* + s^* + t^* > 2 \end{cases}$$

Where q^*, s^* and t^* are either zero or positive integers.

The bias and MSE of $\hat{\rho}_h$ up to terms of order n^{-1} are respectively,

$$B(\hat{\rho}_h) = \left[\rho\,\left\{(1 + f_2 K A^*)\, h_0(p) - 1\right\} + a\left\{\left(f_2\lambda_1^* + \frac{\alpha_1 - 1}{2}\, f_1 C_x^2\right) h_1(p)\right.\right.$$

$$\left.\left. + f_2\left(G + \frac{(\alpha_2-1)}{2}\,(K\beta_2(x) - M)\right) h_2\,(p)\right\}\right] \quad (2.4)$$

$$M(\hat{\rho}_h) = \Big[\rho^2 \{(1+f_2K(Q+2A^*))\, h_0^2(p) - 2(1+f_2KA^*)\, h_o(p) + 1\} + f_1C_x^2h_1^2(p)$$
$$+ 2\rho f_2\lambda_1^* h_o(p)h_1(p) + f_2(K\beta_2(x) - M)\, h_2^2(p) + 2\rho f_2 G h_o(p)h_2(p)$$
$$+ 2f_2\gamma_1(x)C_xh_1(p)\, h_2(p) + 2\rho\, a \Big\{\Big(f_2\lambda_1^* + \frac{(\alpha_1-1)}{2} f_1\, C_x^2\Big)\, h_1(p)$$
$$+ f_2\Big(G + \frac{\alpha_2-1}{2}(K\beta_2(x) - M)\, h_2(p)\Big)\Big\}(h_o(p) - 1)\Big] \tag{2.5}$$

where $A^* = \left[\frac{1}{4}l + \frac{3}{8}\{\beta_2(x) + \beta_2(y)\} - \frac{1}{2}(\lambda_3 + \lambda_4)\right]$

The optimum values of h_i (i = 0, 1, 2) which minimize $M(\hat{\rho}_h)$ in (2.5) are

$$h_o(p) = \frac{\Delta_o}{\Delta}, \quad h_1(p) = \rho\frac{\Delta_1}{\Delta}, \quad h_2(p) = \rho\frac{\Delta_2}{\Delta} \tag{2.6}$$

where $\Delta_o = \left[(1 + f_2KA^*)D^* - X^* - AA_1^* - BA_2^*\right]$

$$\Delta_1 = [\, f_2^2K(Q + A^*)\{A(K\beta_2(x) - M) - B\gamma_1(x)\, C_x\}$$
$$+ \{1 + f_2K(Q + 2A^*)\}\, A_1^* + Bf_2(B\lambda_1^* - AG)]$$

$$\Delta_2 = [\, f_2^2\, KC_x\, (Q + A^*)\, \{B\, f_1C_x - A\, f_2\gamma_1(x)\}$$
$$+ \{1 + f_2K(Q + 2A^*)\}\, A_2^* - Af_2(B\lambda_1^* - AG)]$$

$$\Delta = \left[\{1 + f_2K(Q + 2A^*)\}\, D^* - X^*\right]$$

$$D^* = \left[f_2C_x^2\{f_1(K\beta_2(x) - M) - f_2\, \gamma_1^2(x)\}\right]$$

$$X^* = \left[A^2\, f_2\, (KB_2(x) - M) + B^2f_1\, C_x^2 - 2ABf_2\, \gamma_1(x)\, C_x\right]$$

$$A_1^* = \left[f_2^2\{G\gamma_1(x)\, C_x - \lambda_1^*(K\beta_2(x) - M)\}\right]$$

$$A_2^* = f_2C_x\{f_2\, \lambda_1^*\, \gamma_1(x) - f_1C_xG\}$$

$$A = \left[f_2\lambda_1^* + a\left\{f_2\, \lambda_1^* + \frac{(\alpha_1-1)}{2} f_1C_x^2\right\}\right]$$

$$B = f_2\left[G + a\left\{fG + \frac{(\alpha_2-1)}{2}(K\beta_2(x) - M)\right\}\right]$$

Hence, the minimum MSE of $\hat{\rho}_h$ is

$$M_o(\hat{\rho}_h) = \rho^2\left[1 - (1+f_2KA^*)\frac{\Delta_o}{\Delta} - \frac{a}{\Delta}\left\{(f_2\lambda_1^* + \frac{(\alpha_1-1)}{2}f_1C_x^2)\Delta_1 + \left(G + \frac{(\alpha_2-1)}{2}(K\beta_2(x) - M)\right)\Delta_2\right\}\right] \quad (2.7)$$

Remark 2.1 : The proposed class of estimators $\hat{\rho}_h$ attains minimum MSE given in (2.7) with optimum values of h_i (i=0,1,2) given in (2.6) which are functions of unknown population parameters. Das and Tripathi [1] has shown that guess values can effectively be used in place of population parameters to obtain estimators which are efficient than the usual estimator. Srivastava and Jhajj [6], Singh and Singh [5] and Sampat [4] have shown that the class of estimators with estimated optimum values attains the same minimum MSE of the class of estimators based on optimum values. Thus the presence of unknown parameters in optimum h_i's (i = 0, 1, 2) is not going to be a drawback in obtaining efficient class of estimators.

Remark 2.2 : The MSE of Srivastava and Jhajj [7] class of estimators can be obtained directly from (2.5) by putting h(p) = ρ i.e. $h_o(p) = 1$. It is given by

$$M(\hat{\rho}_s) = \left[\rho^2 f_2KQ + f_1C_x^2h_1^2(p) + f_2(K\beta_2(x) - M)\,h_2^2(p) + 2\rho f_2\left(\lambda_1^*h_1(p) + Gh_2(p)\right) + 2f_2\gamma_1(x)C_xh_1(p)h_2(p)\right] \quad (2.8)$$

Remark 2.3 : For infinite populations, the expressions (2.4) to (2.7) up to terms of order n^{-1} are obtained by taking limit as N tends to infinity and are respectively given by

$$B^*(\hat{\rho}_h) = \left[\rho\left\{\left(1 + \frac{A^*}{n}\right)h_o(p) - 1\right\} + \frac{a}{n}\left\{\left(\lambda_1^* + \frac{\alpha_1-1}{2}C_x^2\right)h_1(p) + \left(G^* + \frac{\alpha_2-1}{2}\left(\beta_2(x) - \frac{n-3}{n-1}\right)\right)h_2(p)\right\}\right] \quad (2.9)$$

$$M^*(\hat{\rho}_h) = \Bigg[\rho^2 \left\{\left(1+\frac{Q+2A^*}{n}\right)h_o^2(p) - 2\left(1+\frac{A^*}{n}\right)h_o(p)+1\right\} + \frac{C_x^2}{n}h_1^2(p)$$

$$+\left(\frac{2\rho\lambda_1^*}{n}\right)h_o(p)\,h_1(p) + \frac{1}{n}\left(\beta_2(x) - \frac{n-3}{n-1}\right)h_2^2(p) + \left(\frac{2\rho G^*}{n}\right)h_o(p)h_2(p)$$

$$+\frac{2C_x\gamma_1(x)}{n}h_1(p)h_2(p) + \frac{2a\rho}{n}\left\{\left(\lambda_1^* + \frac{\alpha_1-1}{2}C_x^2\right)h_1(p)\right.$$

$$\left. + \left(G^* + \frac{\alpha_2-1}{2}\left(\beta_2(x) - \frac{n-3}{n-1}\right)\right)h_2(p)\right\}(h_o(p)-1)\Bigg] \qquad (2.10)$$

where $G^* = \left[\left(\lambda_3 - \frac{n-3}{n-1}\right) - \frac{1}{2}\left(\beta_2(x) - \frac{n-3}{n-1} + 1 + \frac{2\rho^2}{n-1} - 1\right)\right]$

The optimum values of h_i $(i = 0, 1, 2)$ which minimize $M^*(\hat{\rho}_h)$ are given by

$$h_0^*(p) = \frac{\delta_0}{\delta}, \qquad h_1^*(p) = \frac{\rho\delta_1}{\delta}, \quad h_2^*(p) = \frac{\rho\delta_2}{\delta} \qquad (2.11)$$

where $\delta_o = \left[(1+A^*/n)\,D^{**} - X^{**} - A^{**}A_1^{**} - B^{**}A_2^{**}\right]$

$$\delta_1 = \left[\frac{Q+A^*}{n^2}\left\{A^{**}\left(\beta_2(x) - \frac{n-3}{n-1}\right) - B^{**}\gamma_1(x)\,C_x\right\}\right.$$

$$\left. + \left\{1+\frac{Q+2A^{**}}{n}\right\}A_1^{**} + \frac{B^{**}}{n}(B^{**}\lambda_1^* - A^{**}G^*)\right]$$

$$\delta_2 = \left[\frac{C_x(Q+A^*)}{n^2}(B^{**}C_x - A^{**}\gamma_1(x)) + \left\{1+\frac{Q+2A^*}{n}\right\}A_2^{**} - \frac{A^{**}}{n}(B^{**}\lambda_1^* - A^{**}G^*)\right]$$

$$\delta = \left[\left\{1+\frac{Q+2A^*}{n}\right\}D^{**} - X^{**}\right]$$

$$D^{**} = \frac{C_x^2}{n^2}\left[\left(\beta_2(x) - \frac{n-3}{n-1}\right) - \gamma_1^2(x)\right]$$

$$X^{**} = \frac{1}{n}\left[A^{**2}\left(\beta_2(x) - \frac{n-3}{n-1}\right) + B^{**2}C_x^2 - 2A^{**}B^{**}C_x\gamma_1(x)\right]$$

$$A_1^{**} = \frac{1}{n^2}\left[G^* C_x \gamma_1(x) - \lambda_1^* \left(\beta_2(x) - \frac{n-3}{n-1}\right)\right]$$

$$A_2^{**} = \frac{1}{n^2}\left\{ \lambda_1^* \gamma_1(x) - G^* C_x \right\} C_x$$

$$A^{**} = \frac{1}{n}\left[\lambda_1^* + a\left(\lambda_1^* + \frac{\alpha_1 - 1}{2} C_x^2\right)\right]$$

$$B^{**} = \frac{1}{n}\left[G^* + a\left\{G^* + \frac{\alpha_2 - 1}{2}\left(\beta_2(x) - \frac{n-3}{n-1}\right)\right\}\right]$$

Hence, the minimize MSE of $\hat{\rho}_h$ is

$$M_o^*(\hat{\rho}_h) = \rho^2\left[1 - (1 + A^*/n)\frac{\delta_o}{\delta} - \frac{a}{n\delta}\left\{\left(\lambda_1 + \frac{\alpha_1 - 1}{2} C_x^2\right)\delta_1 + \left(G^* + \frac{\alpha_2 - 1}{2}(\beta_2(x) - \frac{n-3}{n-1})\right)\delta_2\right\}\right] \tag{2.12}$$

If we further assume that the infinite population is bivariate normal and n is large (i.e., [(n–3)/(n–3)] = 1) the MSE expression in (2.10) reduces to

$$\begin{aligned} M^{**}(\hat{\rho}_h) = \Bigg[\rho^2\Bigg\{&\left(1 + \frac{(1-\rho^2)(1-2\rho^2)}{n\rho^2}\right)h_o^2(p) - 2\left(1 - \frac{(1-\rho^2)}{2n}\right)h_o(p) + 1\Bigg\} \\ &+ \frac{C_x^2}{n} h_1^2(p) + 2h_2^2(p) + \frac{2\rho(1-\rho^2)}{n} h_o(p)\, h_2(p) \\ &+ \left(\frac{2a\rho}{n}\right)\left\{\frac{(\alpha_1 - 1)}{2} C_x^2 h_1(p)(h_o(p) - 1) + (\alpha_2 - \rho^2) h_2(p)(h_o(p) - 1)\right\}\Bigg] \end{aligned} \tag{2.13}$$

The optimum values of h_i's (i = 0, 1, 2) which minimize $M^{**}(\hat{\rho}_h)$ in (2.13) are

$$h_o'(p) = \frac{\delta_o^*}{\delta^*}, \quad h_1'(p) = \rho\frac{\delta_1^*}{\delta^*}, \quad h_2'(p) = \rho\frac{\delta_2^*}{\delta^*} \tag{2.14}$$

where

$$\delta_o^* = \left[\frac{2}{n^2}\left\{1 - \frac{(1-\rho^2)}{2n}\right\} C_x^2 - U - RT\right], \quad \delta_1^* = \frac{S}{n}\left[\frac{(1-\rho^2)(2-3\rho^2)}{n\rho^2} - R(1-\rho^2)\right]$$

$$\delta_2^* = \left[\frac{R(1-\rho^2)(2-3\rho^2)C_x^2}{2\,n^2\,\rho^2} + T\left\{1 + \frac{(1-\rho^2)(1-2\rho^2)}{n\rho^2}\right\} + \frac{(1-\rho^2)S^2}{n}\right],$$

$$\delta^* = \left[\frac{2}{n^2}\left\{1 - \frac{(1-\rho^2)(1-2\rho^2)}{n\rho^2}\right\}C_x^2 - U\right];\ T = \frac{(1-\rho^2)C_x^2}{n^2},$$

$$R = \frac{1}{n}\left[(1-\rho^2) + a(\alpha_2 - \rho^2)\right], \qquad U = \frac{1}{n}(2S^2 + R^2\,C_x^2),$$

$$S = \frac{a(\alpha_1 - 1)\,C_x^2}{2n}.$$

Hence, the minimize MSE of $\hat{\rho}_h$ is

$$M_0^{**}(\hat{\rho}_h) = \rho^2\left[1 - \left\{1 - \frac{(1-\rho^2)}{n}\right\}\frac{\delta_o^*}{\delta^*} - \frac{a}{n\delta^*}\left\{\frac{(\alpha_1-1)C_x^2\delta_1^*}{2} + (\alpha_2 - \rho^2)\delta_2^*\right\}\right];$$

(2.15)

3. *Illustration of Results for a Particular Estimator*

Consider the estimator d_4 listed in section 1 for illustration. The estimator d_4 is

$$d_4 = r\,[w_0 + w_1 u + w_2 v]$$

For this estimator first order partial derivatives about the point $(\rho, 1, 1)$ are

$$h_o(p) = w_o + w_1 + w_2 \qquad h_1(p) = \rho\,w_1 \qquad h_2(p) = \rho\,w_2 \tag{3.1}$$

Substituting the values of h_i (i=0, 1, 2) in (2.9) and (2.10) and putting a = 1 as d_4 is ratio/product type estimator the bias and MSE of d_4 for infinite populations up to terms up to terms of order n^{-1} are obtained

$$B^*(d_4) = \rho\left[\left(1 + \frac{A^*}{n}\right)\left(\sum_{i=0}^{2} w_i\right) + \frac{\lambda_1^* w_1}{n} + \frac{G^* w_2}{n} - 1\right] \tag{3.2}$$

$$M^{*}(d_4)=\rho^2\left[\left\{1+\frac{(Q+2A^{*})}{n}\right\}\left(\sum_{i=0}^{2} w_i\right)^2-2\left(1+\frac{A^{*}}{n}\right)\left(\sum_{i=0}^{2} w_i\right)+\frac{w_1^2C_x^2}{n}\right.$$

$$+\frac{2\lambda_1^{*}w_1}{n}\left(\sum_{i=0}^{2} w_i\right)+\frac{1}{n}\left\{\beta_2(x)-\frac{n-3}{n-1}\right\}w_2^2+\frac{2G^{*}}{n}\left(\sum_{i=0}^{2} w_i\right)w_2+\frac{2\gamma_1(x)}{n}C_xw_1w_2$$

$$\left.+\frac{2a}{n}\left\{\lambda_1^{*}w_1\left(\sum_{i=0}^{2} w_i-1\right)+G^{*}w_2\left(\sum_{i=0}^{2} w_i-1\right)\right\}\right] \tag{3.3}$$

The MSE of $M^{*}(d_4)$ is minimized for

$$w_o=\frac{\delta_o-(\delta_1+\delta_2)}{\delta},\qquad w_1=\frac{\delta_1}{\delta},\qquad w_2=\frac{\delta_2}{\delta} \tag{3.4}$$

where δ and δ_i's (i=0, 1, 2) are as defined in section 2, with A^{**} and B^{**} redefined as

$$A^{**}=\frac{2\lambda_1^{*}}{n},\qquad B^{**}=\frac{2G^{*}}{n}$$

Hence, the minimum MSE of d_4 is

$$M_o^{*}(d_4)=\rho^2\left[1-\left(1+\frac{A^{*}}{n}\right)\left(\frac{\delta_o}{\delta}\right)-\left(\frac{\lambda_1^{*}}{n}\right)\left(\frac{\delta_1}{\delta}\right)-\left(\frac{G^{*}}{n}\right)\left(\frac{\delta_2}{\delta}\right)\right] \tag{3.5}$$

4. *Theoretical Comparision of $\hat{\rho}_h$ with other Estimators*

From (1.5) and (2.7) for i = 1, 2, it can be proved that

$$M_0(d_i)-M_o(\hat{\rho}_h)=\frac{\rho^2F^{*2}}{\Delta D^{*}}$$

$$\geq 0 \tag{4.1}$$

where

$$F^{*}=\left[f_2K(Q+A^{*})D^{*}+f_2(\lambda_1^{*}A_1^{*}+GA_2^{*})\right.$$

$$+ a\left\{\left(f_2\,\lambda_1^* + \frac{\alpha_1 - 1}{2}\,f_1\,C_x^2\right)A_1^* + f_2\left(G + \frac{\alpha_2 - 1}{2}(K\beta_2(x) - M)\right)A_2^*\right\}\Bigg]$$

which implies that

$$M_o(\hat{\rho}_h) \le M_o(d_i) \quad \forall \quad i = 1, 2 \tag{4.2}$$

Hence, the proposed estimator $\hat{\rho}_h$ is superior to Srivastava and Jhajj [7] estimator. Srivastava and Jhajj [7] has shown that $M_0(d_i) \le M(r)$ where M(r) is the MSE of sample correlation coefficient r. Hence, it follows that

$$M_o(\hat{\rho}_h) \le M_o(d_i) \le M(r) \qquad i=1, 2 \tag{4.3}$$

5. *Empirical Study*

The relative efficiency (RE) of the proposed and alternative estimators has been evaluated using the numerical example given by Johnston [3, p. 171]. The varaites considered are

y : Percentage of lives affected by disease

x : Mean January temperature

For the given data, treating as a population of size 10, the following parameters were computed;

$$\bar{Y} = 52, \quad \bar{X} = 42, \qquad C_y = 0.1482, \; C_x = 0.1237$$

$$\beta_2(y) = 1.8857, \; \beta_2(x) = 1.8593, \; l = 1.3606, \; \gamma_1(x) = 0.2352$$

$$\lambda_1^* = -0.0144, \; \lambda_3 = 1.6887, \; \lambda_4 = 1.8471, \; \rho = 0.7966$$

Sample size was taken as n = 4.

The RE of different estimators with respect to the conventional estimator r was computed using the following formula

$$\frac{MSE(r)}{MSE\ of\ the\ estimator\ compared} \times 100 \tag{5.1}$$

The RE of the proposed class of estimators $\hat{\rho}_h$ was computed for a=0 and also for four pairs of (α_1, α_2) values viz :

(−1, −1), (−, +1), (+1, −1) and (+1, +1) with a = 1.

The results are summarized in Table 5.1.

Table 5.1. Relative Efficiency (RE) of different Estimators over r

		Estimator		RE (%)
1.		di (i = 1, 2)		123.28
2.		$\hat{\rho}_h$ (present study)		
	(i)	a = 0		126.56
	(ii)	a = 1		
		$\alpha_1 = -1$	$\alpha_2 = -1$	137.70
		$\alpha_1 = -1$	$\alpha_2 = +1$	124.85
		$\alpha_1 = +1$	$\alpha_2 = -1$	139.40
		$\alpha_1 = +1$	$\alpha_2 = +1$	125.30

The proposed class of estimators is most efficient for $\alpha_1 = +1$ and $\alpha_2 = -1$. It is obvious from the table 5.1 that the proposed class of estimators is more efficient than the conventional estimator, r and also the estimators d_i (i = 1, 2) suggested by Srivastava and Jhajj [7] irrespective of the values of α_1 and α_2.

REFERENCES

[1] Das, A.K. Tripathi, T.P., 1978. Use of auxiliary information in estimating the finite population variance. *Sankhya, C*, **40**, 139-148.

[2] Gupta, J.P., Singh, R. and Lal, B. 1978. On the estimating of the finite population correlation coefficient-I.

[3] Johnston, J., 1972. Econometric methods. Mc Graw-Hill, Kogakusha Ltd., Tokyo.

[4] Sampath, S., 1980. On the optimal choice of unknowns in ratio-type estimators. *Jour. Ind. Soc. Ag. Statistics*, **41**, 166-172.

[5] Singh, R.K. and Singh, G., 1984. A class of estimators with estimated optimum values in sample surveys. Statist. *Prob. letters*, **2**, 319-321.

[6] Srivastava, S.K. and Jhajj, H.S., 1983. A class of estimators of the population mean multi-auxiliary information. Cal. Stat. *Assoc. Bull.*, **32**, 47-56.

[7] Srivastava, S.K. and Jhajj, H.S., 1986. On the estimation of finite correlation coefficient. *Jour. Ind. Ag. Statistics*, **38**, 82-91.