STATISTICA, anno LXIV, n. 4, 2004

A FAMILY OF ALMOST UNBIASED ESTIMATORS FOR NEGATIVELY CORRELATED VARIABLES USING JACKKNIFE TECHNIQUE

L.N. Upadhyaya, H.P. Singh, S. Singh

1. INTRODUCTION

It is a common practice to use auxiliary information at the estimation stage for increasing the efficiency of the estimators. Out of many ratio, regression and product methods of estimation are good examples in this context. If the correlation between the study character y and the auxiliary character x is (high) positive, ratio method of estimation is generally used. On the other hand, if this correlation is high but negative , product method of estimation can be employed.

Consider a finite population of N units

$$\Omega : \{u_1, u_2, ..., u_N\} \tag{1}$$

Let Y and X be the characteristics taking value y_i and x_i respectively on u_i $(i = 1, 2, ..., N)$. We denote the population mean of X by $\overline{X}$, which is assumed to be known and the population mean of Y by $\overline{Y}$, which is to be estimated. For estimating $\overline{Y}$, Srivenkataramana (1980) and Bandyopadhyay (1980) proposed a dual to product estimator as

$$\hat{\overline{Y}}_T = \overline{y}\frac{\overline{X}}{\overline{x}^*} \tag{2}$$

where $\overline{x}^* = \dfrac{N\overline{X} - n\overline{x}}{N - n}$ denote the mean of X for the non-sampled units, $\overline{y}$ and $\overline{x}$ are the sample means of y and x respectively.

Using predictive approach advocated by Basu (1971), Srivastava (1983) envisaged another estimator for $\overline{Y}$ as

$$\hat{\overline{Y}}_s = \overline{y}\frac{[n\overline{X} + (N - 2n)\overline{x}]}{(N\overline{X} - n\overline{x})} = \overline{y}\left[1 - \left(\frac{\overline{X} - \overline{x}}{\overline{x}^*}\right)\right] \tag{3}$$

Let a sample of size n be drawn without replacement from the population and let it be split into g sub-samples each of size $m = n/g$, where m is an integer. Let $(\bar{x}_j, \bar{y}_j)$, $j = 1,2,...,g$ be the unbiased estimators of $(\bar{X}, \bar{Y})$ based on j^{th} sub-sample of size m. The Jackknife estimators of the type (2) and (3), respectively, based on *j-th* sub-sample are given by

$$\hat{\bar{Y}}_{s_j} = \bar{y}_j \left(\frac{\bar{X}}{\bar{x}_j^*} \right) \tag{4}$$

and

$$\hat{\bar{Y}}_{Tj} = \bar{y}_j \left[1 - \frac{\bar{X} - \bar{x}_j}{\bar{x}_j^*} \right], \; j = 1,2,...,g \tag{5}$$

where

$$\bar{x}_j^* = \frac{N\bar{X} - n\bar{x}_j}{N - n} \quad j = 1,2,...,g\,. \tag{6}$$

The estimators discussed by Srivenkataramana (1980) and Srivastava (1983) are generally biased estimators. In this paper, an attempt has been made to reduce the biases of these estimators employing Jackknife technique developed by Quenouille (1956). Many almost unbiased product-type estimators are obtained and explicit expressions for their variances are also derived to the first degree of approximation. The meaning of almost unbiased estimators is that there is no bias up to terms of order n^{-1}. For details on the bias reduction from product type estimators, the reader can also refer to Tripathi and Singh (1992) and Singh (2003).

2. PROPOSED ESTIMATORS

Let $\hat{\bar{y}}_1 = \frac{1}{g}\sum_{j=1}^{g}\hat{\bar{Y}}_{Tj}$, $\hat{\bar{y}}_2 = \frac{1}{g}\sum_{j=1}^{g}\hat{\bar{Y}}_{sj}$, $\hat{\bar{y}}_3 = \hat{\bar{Y}}_T$, $\hat{\bar{y}}_4 = \hat{\bar{Y}}_s$ and $\hat{\bar{y}}_5 = \bar{y}$ and define the class of estimators for $\bar{Y}$ as

$$\hat{\bar{Y}}_\theta = \sum_{i=1}^{5}\theta_i \hat{\bar{y}}_i \tag{7}$$

such that $\sum_{i=1}^{5}\theta_i = 1$ and where θ_i, $i = 1,2,3,4,5$ are suitably chosen constants.

In order to study the bias property of $\hat{\bar{Y}}_\theta$, we have the following lemma which can be easily proved by the procedure adopted in Cochran (1963) and Sukhatme and Sukhatme (1970).

Lemma 2.1. Under SRSWOR scheme, the relative biases of the estimators $\hat{\bar{y}}_i$, $i = 1, 2, 3, 4, 5$, to the first degree of approximation, are given by

$$\left.\begin{aligned} RB(\hat{\bar{y}}_1) &= \left(\frac{g}{n} - \frac{1}{N}\right) p\lambda \\ RB(\hat{\bar{y}}_2) &= \left(\frac{g}{n} - \frac{1}{N}\right) \lambda \\ RB(\hat{\bar{y}}_3) &= \left(\frac{1}{n} - \frac{1}{N}\right) p\lambda \\ RB(\hat{\bar{y}}_4) &= \left(\frac{1}{n} - \frac{1}{N}\right) \lambda \\ RB(\hat{\bar{y}}_5) &= 0 \end{aligned}\right\} \tag{8}$$

where $RB(\hat{\bar{y}}_i) = \dfrac{B(\hat{\bar{y}}_i)}{\bar{Y}}$, $i = 1,2,3,4,5$; $\lambda = (p + k)\, C_x^2$; $p = \dfrac{n}{N - n}$; $k = \rho \dfrac{C_y}{C_x}$, $C_y = \dfrac{S_y}{\bar{Y}}$; $C_x = \dfrac{S_x}{\bar{X}}$; ρ is the correlation coefficient between x and y, $(N-1)S_y^2 = \sum_{i=1}^{N} (y_i - \bar{Y})^2$ and $(N-1)S_x^2 = \sum_{i=1}^{N} (x_i - \bar{X})^2$.

Using the results of (8) in (7), it is easy to state the following theorem.

Theorem 2.1. An estimator in the class of estimators $\hat{\bar{Y}}_\theta$ at (7) would be unbiased if and only if

$$(\theta_1 p + \theta_2)\, h + \theta_3 p + \theta_4 = 0 \tag{9}$$

where

$$h = \frac{(g - f)}{(1 - f)} \text{ and } f = \frac{n}{N}.$$

Proof. It follows from $\sum_{i=1}^{5} \theta_i = 1$ and (9) that

$$\theta_5 = 1 - [\theta_1(1 - ph) + \theta_2(1 - h) + \theta_3(1 - p)] \tag{10}$$

Using (9) and (10) in (7), we obtain a family of almost unbiased estimators for $\bar{Y}$ as

$$\hat{\bar{Y}}_\theta^{(u)} = \{1 - (\theta_1 + \theta_2)(1 - h)\}\bar{y} + \theta_1 \left\{ \frac{1}{g} \sum_{j=1}^{g} \hat{\bar{Y}}_{T_j} - h\,\hat{\bar{Y}}_T \right\} + \theta_2 \left\{ \frac{1}{g} \sum_{j=1}^{g} \hat{\bar{Y}}_{s_j} - h\,\hat{\bar{Y}}_s \right\} \tag{11}$$

Remark 2.1. If we set $\theta_2 = 0$, then $\hat{\bar{Y}}_\theta^{(u)}$ reduces to the family of almost unbiased estimators for $\bar{Y}$ as

$$\hat{\bar{Y}}_1^{(u)} = \{1 - \theta_1(1 - h)\}\bar{y} + \theta_1 \left\{ \frac{1}{g} \sum_{j=1}^{g} \hat{\bar{Y}}_{T_j} - h\,\hat{\bar{Y}}_T \right\} \tag{12}$$

while for $\theta_1 = 0$ in (11), we get another family of almost unbiased estimators as

$$\hat{\bar{Y}}_2^{(u)} = \{1 - \theta_2(1 - h)\}\bar{y} + \theta_2 \left\{ \frac{1}{g} \sum_{j=1}^{g} \hat{\bar{Y}}_{s_j} - h\,\hat{\bar{Y}}_s \right\} \tag{13}$$

The estimator $\hat{\bar{Y}}_1^{(u)}$ is based on Srivenkataramana (1980) and Bandyopadhyay (1980) estimator $\hat{\bar{Y}}_T$, while $\hat{\bar{Y}}_2^{(u)}$ is based on Srivastava (1983) estimator $\hat{\bar{Y}}_s$.

Many other almost unbiased estimators for $\bar{Y}$ can be derived from $\hat{\bar{Y}}_\theta^{(u)}$ in (11) just by putting the suitable values of θ_1 and θ_2.

Remark 2.2. For $\theta_1 = (1-h)^{-1}$, $\hat{\bar{Y}}_1^{(u)}$ reduce to the usual almost unbiased Jackknife version of $\hat{\bar{Y}}_T$:

$$\hat{\bar{Y}}_{11}^{(u)} = \frac{(g-f)}{(g-1)} \bar{y} \frac{\bar{X}}{\bar{x}^*} - \frac{(1-f)}{(g-1)} \frac{1}{g} \sum_{j=1}^{g} \bar{y}_j \frac{\bar{X}}{\bar{x}_j^*} \tag{14}$$

and for $\theta_2 = (1-h)^{-1}$, $\hat{\bar{Y}}_2^{(u)}$ yields to the usual almost unbiased Jackknife version of $\hat{\bar{Y}}_s$ as

$$\hat{\bar{Y}}_{21}^{(u)} = \frac{(g-f)}{(g-1)} \bar{y} \frac{\{n\bar{X} + (N-2n)\bar{x}\}}{N\bar{X} - n\bar{x}} - \frac{(1-f)}{(g-1)} \frac{1}{g} \sum_{j=1}^{g} \bar{y}_j \frac{\{n\bar{X} + (N-2n)\bar{x}_j\}}{N\bar{X} - n\bar{x}_j} \tag{15}$$

For $\theta_1 = \theta_2 = 0$, the estimators $\hat{\bar{Y}}_\theta^{(u)}$, $\hat{\bar{Y}}_1^{(u)}$ and $\hat{\bar{Y}}_2^{(u)}$ reduce to the usual unbiased estimator $\bar{y}$.

3. VARIANCE EXPRESSIONS

From (11) we have

$$\mathrm{V}(\hat{\bar{Y}}_\theta^{(u)}) = [\{1-(\theta_1+\theta_2)(1-h)\}^2 V(\bar{y}) + \theta_1^2 V(d_1) + \theta_2^2 V(d_2)$$
$$+2\{1-(\theta_1+\theta_2)(1-h)\}\,\theta_1 Cov(\bar{y}, d_1) + 2\{1-(\theta_1+\theta_2)(1-h)\}\,\theta_2 Cov(\bar{y}, d_2)$$
$$+2\theta_1\theta_2 Cov(d_1, d_2)] \tag{16}$$

where

$$d_1 = \frac{1}{g}\sum_{j=1}^{g}\hat{\bar{Y}}_{T_j} - h\hat{\bar{Y}}_T \text{ and } d_2 = \frac{1}{g}\sum_{j=1}^{g}\hat{\bar{Y}}_{s_j} - h\hat{\bar{Y}}_s .$$

It is well known under SRSWOR scheme that

$$\mathrm{V}(\bar{y}) = \frac{(1-f)}{n}\bar{Y}^2 C_y^2 \tag{17}$$

Assuming that $\hat{\bar{Y}}_T \approx \frac{1}{g}\sum_{j=1}^{g}\hat{\bar{Y}}_{T_j}$ and $\hat{\bar{Y}}_s \approx \frac{1}{g}\sum_{j=1}^{g}\hat{\bar{Y}}_{s_j}$, then to the first degree of approximation

$$\mathrm{V}(d_1) = \frac{(1-f)}{n}(1-h)^2\bar{Y}^2[C_y^2 + p(p+2k)C_x^2] \tag{18}$$

$$\mathrm{V}(d_2) = \frac{(1-f)}{n}(1-h)^2\bar{Y}^2[C_y^2 + (1+2k)C_x^2] \tag{19}$$

$$\mathrm{Cov}(d_1, d_2) = \frac{(1-f)}{n}(1-h)^2\bar{Y}^2[C_y^2 + \{p+(1+p)k\}C_x^2] \tag{20}$$

$$\mathrm{Cov}(\bar{y},\ d_1) = \frac{(1-f)}{n}(1-h)\bar{Y}^2[C_y^2 + pkC_x^2] \tag{21}$$

and

$$\mathrm{Cov}(\bar{y},\ d_2)=\frac{(1-f)}{n}(1-h)\bar{Y}^2[C_y^2+kC_x^2] \tag{22}$$

Putting (17)-(22) in (16), we get the variance of $\hat{\bar{Y}}_\theta^{(u)}$ to the first degree of approximation as

$$\mathrm{V}(\hat{\bar{Y}}_\theta^{(u)})=\frac{(1-f)}{n}\bar{Y}^2[C_y^2+(p\theta_1+\theta_2)(1-h)C_x^2\{(p\theta_1+\theta_2)(1-h)\}+2k] \tag{23}$$

which is minimised for

$$p\theta_1+\theta_2=-(1-h)^{-1}k \tag{24}$$

Substitution of (24) in (23) yields the minimum variance of $\hat{\bar{Y}}_\theta^{(u)}$ as

$$\mathrm{minV}(\hat{\bar{Y}}_\theta^{(u)})=\frac{(1-f)}{n}\bar{Y}^2C_y^2[1-\rho^2] \tag{25}$$

Thus we proved the following theorem:

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

$$\mathrm{V}(\hat{\bar{Y}}_\theta^{(u)})\geq\frac{(1-f)}{n}S_y^2[1-\rho^2] \tag{26}$$

with the equality sign holds if and only if

$$p\theta_1+\theta_2=-(1-h)^{-1}k\,.$$

It is interesting to remark that the lower bound of the variance at (26) is the variance of the usual biased linear regression estimator, which depicts that the estimators belonging to the class $\hat{\bar{Y}}_\theta^{(u)}$ are asymptotically no more efficient than the linear regression estimator. We also note from (17) and (25) that the minimum variance of $\hat{\bar{Y}}_\theta^{(u)}$ is no longer more than $\frac{(1-f)}{n}S_y^2$, the variance of the usual unbiased estimator $\bar{y}$, since the quantity $[1-\rho^2]$ of (17) is no more greater than one.

Remark 3.1. Setting $\theta_2=0$ in (24), we get the optimum value of θ_1 as

$$\theta_1=-(1-h)^{-1}k=\theta_{1opt}\ \text{(say)} \tag{27}$$

for which the variance of $\hat{\bar{Y}}_1^{(u)}$ in (12) is least and equal to $minV(\hat{\bar{Y}}_\theta)$ in (25). Thus the substitution of (27) in (12) yields the 'asymptotically optimum almost unbiased estimator' (AOAUE) in $\hat{\bar{Y}}_1^{(u)}$ as

$$\hat{\bar{Y}}_{1opt}^{(u)} = (1+k)\bar{y} - \frac{k}{(1-h)}\frac{1}{g}\sum_{j=1}^{g}\bar{y}_j\frac{\bar{X}}{\bar{x}_j^*} + \frac{kh}{(1-h)}\bar{y}\frac{\bar{X}}{\bar{x}^*} \tag{28}$$

with the variance given at (25).

Remark 3.2. For putting $\theta_1 = 0$ in (24), we get the optimum value of θ_2 as

$$\theta_2 = -(1-h)^{-1}k = \theta_{2opt} \text{ (say)} \tag{29}$$

for which the variance of $\hat{\bar{Y}}_2^{(u)}$ is minimum. For the optimum value of $\theta_2 = \theta_{2opt}$ in $\hat{\bar{Y}}_2^{(u)}$, we get the AOAUE in $\hat{\bar{Y}}_2^{(u)}$ as

$$\hat{\bar{Y}}_{2opt}^{(u)} = (1+k)\bar{y} - \frac{k}{(1-h)}\frac{1}{g}\sum_{j=1}^{g}\bar{y}_j\frac{n\bar{X}+(N-n)\bar{x}_j}{N\bar{X}-n\bar{x}_j^*} + \frac{kh}{(1-h)}\bar{y}\frac{n\bar{X}+(N-2n)\bar{x}_j}{N\bar{X}-n\bar{x}} \tag{30}$$

with the variance given at (25).

Remark 3.3. The estimators $\hat{\bar{Y}}_{\theta opt}^{(u)}$, $\hat{\bar{Y}}_{1opt}^{(u)}$ and $\hat{\bar{Y}}_{2opt}^{(u)}$ can be used in practice when k is known. The value of k can be obtained from some earlier survey or pilot study or the expertise gathered in due course of time, for instance see Reddy (1974, 1978), Sahai and Sahai (1985) and Murthy (1967, pp. 96-99).

4. SIMULATION STUDY

In the present investigation of simulation study, we focused to find the exact results based on finite populations. We generated a pair of N independent random numbers y_i^* and x_i^* (say), $i = 1, 2, ..., N$, from a subroutine VNORM with PHI=*0.7*, seed(y) = 8987878 and seed(*x*) = 2348789 following Bratley, Fox and Schrage (1983). For fixed $S_Y = 30$ and $S_X = 25$, we generated transformed variables,

$$y_i = 100.0 + \sqrt{S_Y^2(1-\rho^2)}\; y_i^* + \rho S_y x_i^* \tag{31}$$

and

$$x_i = 90 + S_X x_i^* \tag{32}$$

for different values of the correlation coefficient ρ. From the generated population we computed population means $\bar{Y}$ and $\bar{X}$. In Table 4.1, we selected all possible samples of size $n=5$ from the population of size $N=20$ for a given value of ρ, which results in $\binom{N}{n}=\binom{20}{5}=15504$ samples. From the k^{th} ($k=1,2,\ldots,15504$) sample, we obtained three estimates

$$\hat{\bar{Y}}_1|_k=\hat{\bar{Y}}_{lr}=\bar{y}+\hat{\beta}(\bar{X}-\bar{x}),\text{ with }\hat{\beta}=\frac{s_{xy}}{s_x^2} \tag{33}$$

$$\hat{\bar{Y}}_2|_k=\hat{\bar{Y}}_{11},\text{ for } g=n \tag{34}$$

and

$$\hat{\bar{Y}}_3|_k=\hat{\bar{Y}}_{21},\text{ for } g=n \tag{35}$$

We used an estimate of the lower bound of the variance for each of these estimates as

$$\hat{V}_k(\hat{\bar{Y}}_{h|k})=\frac{1-f}{n}s_y^2(1-r^2),\text{ for } h=1,2,3. \tag{36}$$

where $r=\frac{s_{sy}}{s_x s_y}$. Then the 95% coverage was obtained by counting how many times the true population mean $\bar{Y}$ falls in the closed interval with limits given by

$$\hat{\bar{Y}}_{h|k}\mp t_{\frac{\alpha}{2}}(df=n-2)\sqrt{\hat{V}_k(\hat{\bar{Y}}_{h|k})} \tag{37}$$

out of all possible 15504 samples. The coverage so obtained has been presented in Table 4.1. It is interesting to note that if the correlation is negative and high, the proposed estimators are found to perform much better than regression estimator.

In Table 4.1 we increased our population size to $N=25$ and sample size was kept same $n=5$. The results based on all possible 53130 samples have been presented. It is remarkable here that the results presented in these Tables are exact and hence can be reproduced at any time.

In Table 4.1, we increased the sample size by one unit, that is $n=6$ by keeping $N=25$, which results in substantial change in total number of samples given by 177100. On the basis of simulation, one can conclude that it is worth to use the proposed estimators if the correlation is negative and high. It is to be noted that although the coverage by the proposed estimator remain less in certain cases for negative high correlation, but keep in mind it is unbiased estimator at the equal level of precision of the regression estimator.

TABLE 4.1

The 95% coverage by three estimators for different values of N, n and different values of correlation coefficients

	N=20 and n=5			N=25 and n=5			N=25 and n=6		
ρ	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{11}$	$\hat{\bar{Y}}_{21}$	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{11}$	$\hat{\bar{Y}}_{21}$	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{11}$	$\hat{\bar{Y}}_{21}$
-0.9	0.9209	0.8978	0.9414	0.9045	0.8609	0.9318	0.9174	0.8780	0.9344
-0.8	0.8834	0.9204	0.9358	0.8940	0.9248	0.9231	0.9056	0.9332	0.9191
-0.7	0.9405	0.9558	0.8940	0.8709	0.8552	0.9197	0.8823	0.8777	0.9173
-0.6	0.8816	0.9070	0.9106	0.9126	0.8890	0.9413	0.9183	0.8970	0.9369
-0.5	0.9172	0.9277	0.9449	0.9267	0.9312	0.9501	0.9312	0.9370	0.9474
-0.4	0.9145	0.9401	0.9433	0.9038	0.9486	0.8599	0.9031	0.9375	0.8523
-0.3	0.9178	0.9518	0.7794	0.8941	0.9207	0.8568	0.9015	0.9269	0.8491
-0.2	0.8881	0.8773	0.8227	0.8765	0.9002	0.8458	0.8850	0.9046	0.8536
-0.1	0.8809	0.9171	0.8393	0.8833	0.9216	0.8937	0.9013	0.9274	0.8943
0.0	0.8614	0.8900	0.8249	0.8996	0.9323	0.9141	0.9057	0.9384	0.9141
0.1	0.8690	0.9116	0.8888	0.8871	0.9273	0.8681	0.8927	0.9248	0.8683
0.2	0.9056	0.9112	0.8678	0.8917	0.9259	0.8570	0.9048	0.9244	0.8544
0.3	0.9004	0.9207	0.8578	0.9014	0.8942	0.7989	0.9142	0.8840	0.7783
0.4	0.9132	0.9469	0.9082	0.8818	0.8526	0.7443	0.9000	0.8496	0.7307
0.5	0.9014	0.8650	0.7502	0.8908	0.8899	0.7694	0.8983	0.8797	0.7498
0.6	0.9192	0.8537	0.7323	0.8790	0.8868	0.7372	0.8890	0.8673	0.7122
0.7	0.9107	0.7294	0.5742	0.8918	0.9146	0.8185	0.9006	0.9035	0.8038
0.8	0.9273	0.7283	0.5545	0.9025	0.7575	0.5540	0.9082	0.7143	0.5265
0.9	0.8806	0.7063	0.5293	0.9202	0.5193	0.3354	0.9181	0.4703	0.3164

In the next section we consider a simulation study based on real data as suggested by one of the reviewer.

5. SIMULATION STUDY BASED ON REAL DATA

In this section, we consider the problem of estimation of sleeping hours with the help of known age of the persons living in a particular locality or town. The sleeping hours generally decreases as the people becomes older. Such a data collected from N=30 persons is listed in Singh and Mangat (1996), pag. 187. A summary of the complete data is given below:

TABLE 5.1

Summary of parameters of the population of N=30 units

Parameters	Age (x)	Sleeping Hours (y)
Mean	66.93	6.38
Standard Error	1.76	0.19
Median	66.50	6.50
Mode	56.00	7.00
Standard Deviation	9.61	1.04
Sample Variance	92.41	1.08
Kurtosis	-0.77	-0.52
Skewness	0.30	-0.12
Range	36.00	4.00
Minimum	51.00	4.50
Maximum	87.00	8.50
Sum	2008.00	191.30
Count	30.00	30.00

The correlation between age and sleeping hours in this population is -0.8877. We selected all possible samples each of size $n = 6$ units from the population consisting of $N = 30$ units which results in total of 593775 samples. The 95% coverage based on this simulation is reported in Table 5.2. We also repeated the experiment by selecting all possible samples each of size n=7 units which results in total 2035800 samples.

TABLE 5.2

The 95% coverage by three estimators for N=30 and different values of n

n	$\hat{\bar{Y}}_{lr}$	$\hat{\bar{Y}}_{11}$	$\hat{\bar{Y}}_{21}$
6	0.9075	0.8172	0.9421
7	0.9119	0.8264	0.9438

The results based on real data shows that the proposed estimator $\hat{\bar{Y}}_{21}$ may perform better than the linear regression as well as the unbiased estimator $\hat{\bar{Y}}_{11}$. The empirical study was carried out in FORTRAN-77 using PENTIUM-120.

CONCLUSION

The present investigation provides a valuable message for the survey statisticians to deal with a situation where negative correlation exists between study and auxiliary variables. A lot of efforts have been made to improve ratio estimator which works for positive correlation, but only limited thought have been given for negatively correlated variables. The negatively correlated variables have too much role in medical and social sciences. There are several medical or social science related variables which decreases as the people grow up. For example, as the people become old the following variables have negative correlation with the age: (a) duration of sleeping hours; (b) hearing tendency; (c) eye sight (d) number of hairs on the head; (e) number of love affairs; (f) working hours capacity, and (g) amount of blood donation etc.

Department of Applied Mathematics L. N. UPADHYAYA
Indian School of Mines

School of Studies in Statistics HOUSILA P. SINGH
Vikram University

Department of Statistics SARJINDER SINGH
St. Cloud State University

ACKNOWLEDGEMENTS

The authors are heartily thankful to the Executive Editor Professor Stefania Mignani and a referee for their valuable comments and encouragement to bring the original manuscript in the present form. The opinion and results discussed in this paper are of authors' and not necessary of their institute(s).

REFERENCES

S. BANDYOPADHYAY (1980), *Improved ratio and product estimators*, "Sankhya", 42, C, 45-49.

D. BASU (1971), *An essay on the logical foundations of survey sampling*, Part I. Foundations of Statistical Inference, V.P. Godambe and D.A. Sprott editors, New York, 1971, 203-233.

P. BRATLEY, B. L. FOX, L.E. SCHRAGE (1983), *A Guide to Simulation*, Springer-Verlag, New York.

W.G. COCHRAN (1963), *Sampling Techniques*, John Wiley and Sons: New York.

M.N. MURTHY (1967), *Sampling theory and methods*, Statistical Publishing Society, Calcutta.

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

V.N. REDDY (1974), *A transformed ratio method of estimation*, "Sankhya", C, 36, 59-70.

V.N. REDDY (1978), *A study on the prior knowledge on certain population parameters*, "Sankhya", C, 40, 29-37.

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

R. SINGH, N. S. MANGAT (1996), *Elements of survey sampling*, Kluwer Academic Publishers, The Netherlands.

S. SINGH (2003), *Advanced sampling theory with applications: how Michael "selected" Amy*. pp. 1-1247, Kluwer Academic Press, The Netherlands.

S.K. SRIVASTAVA (1983), *Predictive estimation of finite population mean using product estimator*, "Metrika", 30, 93-99.

T. SRIVENKATARAMANA (1980), *A dual to ratio estimator in sample surveys*, "Biometrika", 67, 199-204.

P.V. SUKHATME, and B.V. SUKHATME (1970), *Sampling theory of surveys with applications*, second edition, Iowa State University Press.

T.P. TRIPATHI, H.P SINGH (1992), *A class of unbiased product-type estimators for the mean suitable for positive and negative correlation situations*, "Communications in statistics - Theory and methods", 21, 507-518.

RIASSUNTO

Una classe di stimatori quasi corretti per variabili aleatorie negativamente correlate basati sul metodo Jackknife

Utilizzando il metodo Jackknife viene definita una classe di stimatori quasi corretti per $\bar{Y}$, la media di popolazione della variabile di studio *Y*. Ne vengono inoltre analizzate le proprietà statistiche nel campionamento casuale semplice senza ripetizione. Attraverso una ricerca empirica viene valutata la performance della soluzione proposta rispetto allo stimatore di regressione.

SUMMARY

A family of almost unbiased estimators for negatively correlated variables using Jackknife technique

Using Jackknife technique a family of almost unbiased estimators for $\bar{Y}$, the population mean of the study variable Y, is suggested and its properties analysed under simple random sampling and without replacement (SRSWOR) scheme. An empirical investigation has been done to show the performance of the proposed unbiased strategies over the biased regression estimator.