Jour. Ind. Soc. Ag. Statistics
54(1), 2001 : 1–12

On the Use of Auxiliary Information in Successive Sampling

G.N. Singh and V.K. Singh[1]
Indian School of Mines, Dhanbad-826 004
(Received : March, 1994)

SUMMARY

In the present work, we make use of auxiliary information on the second occasion for improving the efficiency of the proposed difference-type estimator in successive sampling over two occasions. The estimator is compared with the (*i*) sample mean estimator $\bar{y}_2$ when there is no matching and (*ii*) the optimum estimator $\hat{\bar{Y}}_2$ which is a combination of the means of the matched and unmatched portion of the sample at the second occasion. Optimum replacement policy is also discussed.

Key Words : Difference-type estimator, Successive sampling, Optimum replacement policy, Optimum estimator.

1. *Introduction*

In sample surveys the use of auxiliary information for improving the estimates is well known. In case of successive sampling, it has been seen that to utilize entire information collected in the previous investigations are very advantageous. Jessen [1] was perhaps the first author who used the information collected on the previous occasion for improving the current estimate. Later on this technique was extended by others.

The method usually consists in retaining a fraction of the samples selected in the previous occasions in the repeated sampling enquiries and using the entire information available from such studies with the help of regression technique. Sen [2, 3] generalized the theory of successive sampling over two occasions to provide the optimum estimate of the current average assuming that information on p auxiliary variables $X_1, X_2, \ldots X_p$ with known population means and correlated with Y are available from the previous occasion. No doubt if the interval between successive surveys is small, the information gathered on auxiliary variable at the previous occasion may efficiently be utilized to develop precise estimators. But if the time lapsed between two successive surveys is large and population characteristics are changing rapidly over time, intuitively

1 B.H.U., Varanasi-221 005

it will not be very much fruitful to utilize such auxiliary information collected on the previous occasion except the data on the study variable used in the matched portion of the sample. However, if any kind of auxiliary information is available at the second occasion, it may be used through appropriate sampling design to improve the estimate in the unmatched portion of the sample.

The aim of the present work is to propose a difference type estimator for estimating mean of the current occasion based on successive sampling scheme on two occasions where a fraction of the previously collected sample is retained and a new sample is drawn with SRSWOR strategy from the population at the second occasion. In order to improve the estimate of the current mean based on the fresh sample, it has been assumed that the information on some auxiliary variable which is positively correlated with the main variable is available at the second occassion. The problem of optimum replacement policy has been discussed. The gain in efficiency of the proposed estimator has been obtained over sample mean estimator $\bar{y}_2$ when there is no matching and the optimum estimator $\hat{\bar{Y}}_2$ which is a combination of the means of the matched and unmatched portion of the sample at the second occasion.

2. *The Proposed Estimator*

Let a population of size N which is sampled over two occasions. Assume that the size of the population remains unchanged but values of units changes over occasions. Let a sample of size n_1 be selected using SRSWOR scheme at the first occasion. Out of this sample let n'_2 units are retained on the second occasion while a fresh sample of size n''_2 is drawn on the second occasion from the remaining $(N - n_1)$ units of the population so that the total sample size at the second occasion becomes $n_2 = n'_2 + n''_2$. We further assume that the information on an auxiliary variable X, which is positively correlated to Y is available at the second occasion. In this work we use the following notations:

$\bar{Y}_i$: The population mean of the study variable Y on the ith occasion (i = 1, 2)

S^2_{iY} : The population mean square of Y for the ith occasion

$\bar{y}_1$: The sample mean based on n_1 units drawn on the first occasion

$\bar{y}'_2$: The sample mean based on n'_2 units observed on the second occasion and common with the first occasion

$\bar{y}''_2$: The sample mean based on n''_2 units drawn afresh on the second occasion

$\bar{y}'_1$: The sample mean based on n'_2 units common to both occasions and observed on the first occasion

$\bar{X}_2$: The population mean of the auxiliary variable X on the second occasion

S^2_{2X} : The population mean square of X on the second occasion

$\bar{x}'_2$: The sample mean of X based on n'_2 units common to both occasions and observed on the second occasion

$\bar{x}''_2$: The sample mean of X based on n''_2 units drawn afresh on the second occasion

ρ_{21} : Correlation between the measurements on the same unit on the two successive occasions, if it is observed on study variable Y on both the occasions

ρ_{21XY} : Correlation between the study variable Y and the auxiliary variable X over two occasions

ρ_{2XY} : Correlation between X and Y at the second occasion

For estimating $\bar{Y}_2$ based on successive sampling, two independent estimators can be made. First, based on sample of size n''_2 drawn afresh on the second occasion and second based on the sample of size n'_2 common to both the occasions.

The first estimator proposed is a difference estimator based on n_2'' observations and is given by

$$\hat{\bar{Y}}''_2 = \bar{y}''_2 + \beta_{2XY}(\bar{X}_2 - \bar{x}''_2) \tag{1}$$

while the second estimator is again a difference-type estimator based on the sample of size n_2' is given by

$$\hat{\bar{Y}}'_2 = \bar{y}'_2 + \beta_{21}(\bar{y}_1 - \bar{y}'_1) + \beta_{2XY}(\bar{X}_2 - \bar{x}'_2) \tag{2}$$

where β_{2XY} is the regression coefficient of Y on X at the second occasion and β_{21} stands for the regression coefficient of the variable Y of the second occasion on the same variate of the first occasion. We assume that both β_{21} and β_{2XY} are known.

Combining the two estimators $\hat{\bar{Y}}''_2$ and $\hat{\bar{Y}}'_2$ of $\bar{Y}_2$, we have the final estimator of $\bar{Y}_2$ as follows

$$\hat{T}_2 = \phi \hat{Y}''_2 + (1-\phi)\hat{Y}'_2 \tag{3}$$

where ϕ is an unknown constant to be determined under certain criterion.

3. *Bias of* $\hat{T}_2$

Theorem-1 : $\hat{T}_2$ is an unbiased estimator of $\overline{Y}_2$.

Proof : It is clear that $\bar{y}_1$ and $\bar{y}'_1$ are unbiased estimators of $\overline{Y}_1$. Similarly $\bar{y}''_2$ is an unbiased estimator of $\overline{Y}_2$. Also, Since $\bar{x}'_2$ and $\bar{x}''_2$ are unbiased estimators of $\overline{X}_2$, $\bar{y}'_2$ is an unbiased estimator for $\overline{Y}_2$ and β_{21} and β_{2XY} are known constants, both the estimators $\hat{\overline{Y}}''_2$ and $\hat{\overline{Y}}'_2$ are unbiased for $\overline{Y}_2$. Thus, $\hat{T}_2$ is an unbiased estimator of $\overline{Y}_2$.

4. *Variance of* $\hat{T}_2$

Theorem-2 : Variance of $\hat{T}_2$ is obtained as

$$V(\hat{T}_2) = \left[\phi^2\left(\frac{1}{n''_2}-\frac{1}{N}\right)(1-\rho^2_{2XY}) + (1-\phi^2)\left\{\left(\frac{1}{n'_2}-\frac{1}{N}\right)(1-\rho^2_{2XY}) + \left(\frac{1}{n'_2}-\frac{1}{n_1}\right)(2\rho_{21}\,\rho_{2XY}\,\rho_{21XY}-\rho^2_{21})\right\} - 2\phi(1-\phi)\frac{(1-\rho^2_{2XY})}{N}\right] S^2_{2Y} \tag{4}$$

Proof : It is clear that the variance of $\hat{T}_2$ is given by

$$V(\hat{T}_2) = \phi^2\, V(\hat{\overline{Y}}''_2) + (1-\phi^2)\, V(\hat{\overline{Y}}'_2) + 2\phi(1-\phi)\ \mathrm{COV}(\hat{\overline{Y}}''_2, \hat{\overline{Y}}'_2) \tag{5}$$

Variance and covariance terms of equation (5) can be easily derived as

$$V(\hat{\overline{Y}}''_2) = \left(\frac{1}{n''_2}-\frac{1}{N}\right) S^2_{2Y}(1-\rho^2_{2XY}) \tag{6}$$

$$V(\hat{\overline{Y}}'_2) = \left(\frac{1}{n'_2}-\frac{1}{N}\right) S^2_{2Y}(1-\rho^2_{2XY}) + \left(\frac{1}{n'_2}-\frac{1}{n_1}\right)(2\rho_{21}\,\rho_{2XY}\,\rho_{21XY}-\rho^2_{21})\, S^2_{2Y} \tag{7}$$

and

$$\mathrm{COV}(\hat{\overline{Y}}''_2, \hat{\overline{Y}}'_2) = -\frac{(1-\rho^2_{2XY})}{N} S^2_{2Y} \tag{8}$$

Now substituting the values of $V(\hat{\bar{Y}}''_2)$, $V(\hat{\bar{Y}}'_2)$ and $COV(\hat{\bar{Y}}''_2, \hat{\bar{Y}}'_2)$ from (6), (7) and (8) in (5), we get the $V(\hat{T}_2)$ as in (4).

5. *Minimum Variance of* $\hat{T}_2$

Theorem-3 : Optimum variance of $\hat{T}_2$ is obtained as

$$V(\hat{T}_2)_{opt} = \frac{1}{\left[A + \frac{n''_2}{n_2}\left(1 - \frac{n'_2}{n_1}\right)B\right]^2}\left[\left(\frac{n''_2}{n_2}\right)^2 \left\{A + \left(1 - \frac{n'_2}{n_1}\right)B\right\}^2 \right.$$

$$\left\{\left(\frac{1}{n''_2} - \frac{1}{N}\right)A\right\} + A^2\left(1 - \frac{n''_2}{n_2}\right)^2 \left\{\left(\frac{1}{n'_2} - \frac{1}{N}\right)A + \left(\frac{1}{n'_2} - \frac{1}{n_1}\right)B\right\}$$

$$\left. -2\frac{n''_2}{n_2}\left\{A + \left(1 - \frac{n'_2}{n_1}\right)B\right\}\frac{A^2}{N}\left(1 - \frac{n''_2}{n_2}\right)\right] S^2_{2Y} \tag{9}$$

where

$$A = (1 - \rho^2_{2XY}) \text{ and}$$

$$B = \rho_{21}\,(2\,\rho_{2XY}\,\rho_{21XY} - \rho_{21})$$

Proof : Since $V(\hat{T}_2)$ is a function of ϕ, it can be minimized with respect to ϕ. We get the optimum value of ϕ as

$$\hat{\phi} = \frac{\frac{1}{n'_2}A + \left(\frac{1}{n'_2} - \frac{1}{n_1}\right)B}{\left(\frac{1}{n'_2} + \frac{1}{n''_2}\right)A + \left(\frac{1}{n'_2} - \frac{1}{n_1}\right)B} \tag{10}$$

Making some manipulations the value of $\hat{\phi}$ can be reduced to

$$\hat{\phi} = \frac{\frac{n''_2}{n_2}\left[A + \left(1 - \frac{n'_2}{n_1}\right)B\right]}{\left[A + \frac{n''_2}{n_2}\left(1 - \frac{n'_2}{n_1}\right)B\right]} \tag{11}$$

Putting the value of $\hat{\phi}$ from (11) in (4) we get the $V(\hat{T}_2)_{opt}$ as given in (9).

6. *A Particular Case*

Sukhatme *et al.* [4] have discussed the following estimator of $\bar{Y}_2$ in successive sampling on two occasions as

$$\hat{\bar{Y}}_2 = \phi_2 \bar{y}''_2 + (1 - \phi_2) \bar{y}'_{12} \tag{12}$$

where ϕ_2 is an unknown constant and

$$\bar{y}'_{12} = \bar{y}'_2 + \beta_{21} (\bar{y}_1 - \bar{y}'_1) \tag{13}$$

Obviously $\hat{\bar{Y}}_2$ does not involve the information on auxiliary variable X available at the second occasion. The optimum value of ϕ_2 and corresponding variance of $\hat{\bar{Y}}_2$ are given by

$$\hat{\phi}_2 = \frac{\frac{n''_2}{n_2}\left[1 - \left(1 - \frac{n'_2}{n_1}\right)\rho_{21}^2\right]}{\left[1 - \frac{n''_2}{n_2}\left(1 - \frac{n'_2}{n_1}\right)\rho_{21}^2\right]} \tag{14}$$

and

$$V(\hat{\bar{Y}}_2)_{opt} = \frac{\left[1 - \left(1 - \frac{n'_2}{n_1}\right)\rho_{21}^2\right]}{\left[1 - \frac{n''_2}{n_2}\left(1 - \frac{n'_2}{n_1}\right)\rho_{21}^2\right]} \frac{S_{2Y}^2}{n_2} - \frac{S_{2Y}^2}{N} \tag{15}$$

In defining $\hat{T}_2$ we have assumed that an auxiliary variable X, correlated with Y, is available at the current occasion which can be used to improve the estimation of $\bar{Y}_2$. However , if we assume that even though X is available but it is uncorrelated with Y, that is, $\rho_{2XY} = 0$, our estimator $\hat{\bar{Y}}''_2$ and $\hat{\bar{Y}}'_2$ reduce to $\bar{y}''_2$ and $\bar{y}'_{12}$ respectively and ultimate estimator $\hat{T}_2$ reduces to $\hat{\bar{Y}}_2$. Thus, $\hat{\bar{Y}}_2$ can be viewed as a particular case of $\hat{T}_2$ when $\rho_{2XY} = 0$. It can be seen that in this case A = 1 and B = $-\rho_{21}^2$. Finally , it can be checked that $\hat{\phi}$ and $V(\hat{T}_2)_{opt}$ reduce to $\hat{\phi}_2$ and $V(\hat{\bar{Y}}_2)_{opt}$ respectively.

7. *Case of Equal Sample Size at Both the Occasions*

If we assume that samples of equal size have been drawn at both the occasions. Thus, assume that $n_1 = n_2 = n$ (say). Further, let $\frac{n'_2}{n} = \lambda$ and

$\frac{n''_2}{n} = \mu = 1 - \lambda$. Clearly, μ is the fraction of the sample which has been replaced by a new sample on the second occasion. Substituting the value of $\frac{n'_2}{n}$ and $\frac{n''_2}{n}$ in (11) and (9) in terms of μ we have

$$\hat{\phi} = \frac{\mu(A + \mu B)}{(A + \mu^2 B)} \tag{16}$$

and

$$V(\hat{T}_2)_{opt} = \frac{A(A + \mu B)}{(A + \mu^2 B)} \frac{S^2_{2Y}}{n} - \frac{AS^2_{2Y}}{N} \tag{17}$$

It should be noted that if there is a complete matching, that is, if $\mu = 0$ then $\hat{T}_2 = \hat{\bar{Y}}'_2$, since $\phi = 0$, and

$$V(\hat{T}_2)_{opt} = V(\hat{\bar{Y}}'_2)_{opt} = \left(\frac{1}{n} - \frac{1}{N}\right) S^2_{2Y}(1 - \rho^2_{2XY}) \tag{18}$$

Similarly, when there is no matching and a new sample is selected at the second occasion then $\mu = 1$ and $\hat{T}_2 = \hat{\bar{Y}}''_2$ and the variance reduces to

$$V(\hat{T}_2)_{opt} = V(\hat{\bar{Y}}''_2)_{opt} = \left(\frac{1}{n} - \frac{1}{N}\right) S^2_{2Y}(1 - \rho^2_{2XY}) \tag{19}$$

Thus is both the cases $V(\hat{T}_2)_{opt}$ has the same value which is actually the variance of the difference estimator with $\beta_{2XY} = \rho_{2XY} S_{2Y}/S_{2x}$. This gives an implication that there must be an optimum choice of μ, other than extreme values zero and one, such that $V(\hat{T}_2)_{opt}$ will be smaller than the quantity given in (18) or (19). Thus, for making current estimates, neither the case of "complete matching" nor the case of "no matching" is better, it is always preferable to replace the sample partially.

8. Replacement Policy

As stated in the previous section, an optimum value of μ should be determined so as to know what fraction of the sample on the first occasion should be replaced so that $\bar{Y}_2$ may be estimated with maximum precision. For this, we minimize $V(\hat{T}_2)_{opt}$ in (17) with respect to μ to get the optimum value of μ as

$$\hat{\mu} = \frac{-A \pm [A(A+B)]^{1/2}}{B} \tag{20}$$

Therefore, real values of $\hat{\mu}$ exist if $A(A+B) \geq 0$. Since $A = 1 - \rho^2_{2XY}$; $0 \leq A \leq 1$, so in order to $\hat{\mu}$ to be real, $A + B \geq 0$, that is,

$$1 - \rho^2_{2XY} - \rho^2_{21} + 2\rho_{21}\,\rho_{2XY}\,\rho_{21XY} \geq 0 \tag{21}$$

Now, it is obvious that if $B < 0$, some times two values of $\hat{\mu}$ are possible but if $B > 0$, only one value of $\hat{\mu}$ will be available. In order to choose values of $\hat{\mu}$, it should be remembered that $0 \leq \hat{\mu} \leq 1$. All other values of $\hat{\mu}$ are inadmissible. Substituting the value of $\hat{\mu}$ from (20) in (17) we have

$$V(\hat{T}_2)_{opt} = \frac{B\,[A(A+B)]^{1/2}}{A + B \pm [A(A+B)]^{1/2}}\,\frac{S^2_{2Y}}{2n} - \frac{S^2_{2Y}}{N} \tag{22}$$

9. Efficiency Comparisons

The relative gains in efficiency of $\hat{T}_2$ with respect to (*i*) sample mean estimator $\bar{y}_2$ when there is no matching and (*ii*) $\bar{Y}_2$ given in (12) when no auxiliary information is used on the second occasion, have been obtained for known values of ρ_{21}, ρ_{2XY} and ρ_{21XY}. The variance of $\bar{y}_2$ and $\bar{Y}_2$ are respectively given by

$$V(\bar{y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right) S^2_{2Y} \tag{23}$$

and

$$V(\hat{\bar{Y}}_2)_{opt} = [1 + (1 - \rho^2_{21})^{1/2}]\,\frac{S^2_{2Y}}{2n} - \frac{S^2_{2Y}}{N} \tag{24}$$

Tables 1 to 5 show the optimum values of μ, relative gain in efficiency G_1 of T_2 with respect to $\bar{y}_2$ and relative gain in efficiency G_2 of T_2 with respect to $\hat{\bar{Y}}_2$ where

$$G_1 = \left[\frac{V(\bar{y}_2) - V(\hat{T}_2)}{V(\hat{T}_2)}\right] \times 100 \tag{25}$$

and

$$G_2 = \left[\frac{V(\hat{\bar{Y}}_2) - V(\hat{T}_2)}{V(\hat{T}_2)}\right] \times 100 \tag{26}$$

The values of correlations ρ_{21}, ρ_{2XY} and ρ_{21XY} selected are $+0.2, +0.4, +0.6, +0.8$ and $+0.9$. We have taken $N = 4000$ and $n = 40$. Dashes in the tables indicate the cases where combination of ρ_{21}, ρ_{2XY} and ρ_{21XY} is inadmissible, that is, $A + B < 0$ so that $\hat{\mu}$ is not real.

In this case, for fixed values of ρ_{21XY} and ρ_{2XY}, $\hat{\mu}$ increases as ρ_{21} increases. Therefore, larger the value of ρ_{21}, the larger is the fraction to be replaced. This result is an agreement with the statement made in Sukhatme *et al.* [4]. For fixed ρ_{21} and ρ_{2XY}, $\hat{\mu}$ decreases as ρ_{21XY} increases, thus, a high correlation between the study and auxiliary variable Y and X on first and second occasions guarantees smaller replacement of the sample at the second occassion. As far .as the relation between ρ_{2XY} and $\hat{\mu}$ is concerned, it is apparent that for fixed values of ρ_{21} and ρ_{21XY} there is an inverse relationship between ρ_{2XY} and $\hat{\mu}$ except for a few combinations of ρ_{21} and ρ_{21XY}. It is clear that a smaller fraction of the sample should be replaced if the correlation between Y and X at the second occasion is high. We get a value of $\hat{\mu}$ as high as 0.789 in Table 1 when $\rho_{21} = 0.9$, $\rho_{2XY} = 0.6$ and $\rho_{21XY} = 0.2$ and as low as 0.321 in Table 5 where $\rho_{21} = 0.8$, $\rho_{2XY} = 0.9$ and $\rho_{21XY} = 0.9$. Thus, using

Table 1. Optimum μ and percentage relative efficiencies of $\hat{T}_2$ with respect to $\bar{y}_2$ and $\hat{\bar{Y}}_2$ ($\rho_{21XY} = 0.2$)

ρ_{21}		ρ_{2XY}				
		0.2	0.4	0.6	0.8	0.9
0.2	$\hat{\mu}$	0.503	0.501	0.498	0.492	0.481
	G_1	4.830	19.340	55.760	173.250	405.650
	G_2	3.760	18.120	54.170	170.460	400.490
0.4	$\hat{\mu}$	0.518	0.515	0.513	0.512	0.511
	G_1	7.930	22.700	60.410	184.310	438.010
	G_2	3.380	17.500	53.640	172.340	415.320
0.6	$\hat{\mu}$	0.548	0.547	0.551	0.578	0.670
	G_1	14.480	30.370	72.460	221.580	607.930
	G_2	2.920	17.200	55.040	189.100	536.420
0.8	$\hat{\mu}$	0.613	0.615	0.646	—	—
	G_1	27.910	46.870	102.510	—	—
	G_2	2.070	17.200	61.600	—	—
0.9	$\hat{\mu}$	0.675	0.687	0.789	—	—
	G_1	41.180	64.250	147.880	—	—
	G_2	0.960	17.450	77.260	—	—

Note : Dash indicates that values of correlations are inadmissible.

Table 2. Optimum μ and percentage relative efficiencies of $\hat{T}_2$ with respect to $\bar{y}_2$ and $\hat{\bar{Y}}_2$ ($\rho_{21XY} = 0.4$)

ρ_{21}		ρ_{2XY}				
		0.2	0.4	0.6	0.8	0.9
0.2	$\hat{\mu}$	0.501	0.496	0.490	0.473	0.446
	G_1	4.390	18.200	52.940	162.460	368.590
	G_2	3.320	16.990	51.380	159.780	363.810
0.4	$\hat{\mu}$	0.513	0.505	0.494	0.470	0.436
	G_1	6.940	20.220	54.330	161.230	358.330
	G_2	2.430	15.150	47.820	150.210	339.000
0.6	$\hat{\mu}$	0.540	0.528	0.515	0.492	0.460
	G_1	12.610	25.750	60.960	173.250	383.740
	G_2	1.240	13.050	44.700	145.650	334.870
0.8	$\hat{\mu}$	0.594	0.576	0.564	0.555	0.551
	G_1	24.010	37.300	76.320	208.510	480.770
	G_2	–1.040	9.560	40.700	146.180	363.440
0.9	$\hat{\mu}$	0.644	0.619	0.610	0.628	0.723
	G_1	34.510	47.760	90.990	249.960	664.070
	G_2	–3.820	56.610	36.580	150.260	446.380

Table 3. Optimum μ and percentage relative efficiencies of $\hat{T}_2$ with respect to $\bar{y}_2$ and $\hat{\bar{Y}}_2$ ($\rho_{21XY} = 0.6$)

ρ_{21}		ρ_{2XY}				
		0.2	0.4	0.6	0.8	0.9
0.2	$\hat{\mu}$	0.499	0.492	0.481	0.456	0.419
	G_1	3.950	17.110	50.310	153.160	340.090
	G_2	2.890	15.910	48.710	150.570	335.600
0.4	$\hat{\mu}$	0.509	0.495	0.477	0.440	0.391
	G_1	5.980	17.920	49.060	144.050	310.380
	G_2	1.510	12.950	42.780	133.760	293.080
0.6	$\hat{\mu}$	0.532	0.511	0.487	0.442	0.387
	G_1	10.870	21.740	52.050	144.990	306.100
	G_2	–0.330	9.440	36.690	120.250	265.080
0.8	$\hat{\mu}$	0.578	0.545	0.573	0.462	0.404
	G_1	20.590	29.960	60.410	156.500	324.290
	G_2	–3.770	3.700	28.000	104.680	238.570
0.9	$\hat{\mu}$	0.618	0.574	0.536	0.483	0.424
	G_1	29.110	36.910	67.750	167.980	345.130
	G_2	–7.670	–2.090	19.960	91.630	218.310

Table 4. Optimum μ and percentage relative efficiencies of $\hat{T}_2$ with respect to $\bar{y}_2$ and $\hat{\bar{Y}}_2$ ($\rho_{21XY} = 0.8$)

ρ_{21}		ρ_{2XY}				
		0.2	0.4	0.6	0.8	0.9
0.2	$\hat{\mu}$	0.497	0.488	0.473	0.442	0.397
	G_1	3.510	16.050	47.850	144.990	317.120
	G_2	2.460	14.870	46.350	142.500	312.870
0.4	$\hat{\mu}$	0.504	0.486	0.463	0.416	0.359
	G_1	5.060	15.800	44.440	130.480	276.780
	G_2	0.630	10.910	38.380	120.760	260.880
0.6	$\hat{\mu}$	0.524	0.496	0.464	0.406	0.344
	G_1	9.230	18.200	44.800	125.360	260.440
	G_2	−1.810	6.260	30.180	102.590	224.030
0.8	$\hat{\mu}$	0.564	0.521	0.477	0.410	0.342
	G_1	17.550	24.020	49.060	127.500	259.080
	G_2	−6.200	−1.040	18.950	81.540	186.540
0.9	$\hat{\mu}$	0.597	0.541	0.490	0.417	0.346
	G_1	24.590	28.850	53.050	131.440	263.580
	G_2	−10.910	−7.860	9.450	65.500	159.990

Table 5. Optimum μ and percentage relative efficiencies of $\hat{T}_2$ with respect to $\bar{y}_2$ and $\hat{\bar{Y}}_2$ ($\rho_{21XY} = 0.9$)

ρ_{21}		ρ_{2XY}				
		0.2	0.4	0.6	0.8	0.9
0.2	$\hat{\mu}$	0.496	0.485	0.470	0.435	0.388
	G_1	3.300	15.540	46.680	141.270	307.160
	G_2	2.250	14.360	45.180	138.810	303.000
0.4	$\hat{\mu}$	0.502	0.482	0.456	0.405	0.346
	G_1	4.610	14.790	42.320	124.660	263.230
	G_2	0.200	9.950	36.320	115.180	247.910
0.6	$\hat{\mu}$	0.520	0.490	0.454	0.392	0.327
	G_1	8.440	16.580	41.640	117.460	243.420
	G_2	−2.510	4.800	27.340	95.490	208.730
0.8	$\hat{\mu}$	0.557	0.510	0.463	0.391	0.321
	G_1	16.140	21.430	44.440	116.850	237.230
	G_2	−7.320	−3.100	15.260	73.040	169.110
0.9	$\hat{\mu}$	0.587	0.527	0.472	0.395	0.323
	G_1	22.570	25.490	47.360	118.860	238.330
	G_2	−1.230	−10.270	5.380	56.510	141.940

an auxiliary variable highly correlated with the study variable, the optimum replacement fraction can be brought down to about one-third which is not possible without using X. When $\rho_{2XY} = \rho_{21XY} = 0$, the minimum replacement fraction is one-half.

From Tables 1 to 5 it is clear that the proposed estimator $\hat{T}_2$ is more efficient than $\bar{y}_2$ everywhere. It is also efficient than $\hat{\bar{Y}}_2$ except for some combinations of correlations. A general consensus appears that in order to increase the gain in efficiency of $\hat{T}_2$ over $\bar{y}_2$ and $\hat{\bar{Y}}_2$, ρ_{2XY} should be as large as possible, ρ_{21} and ρ_{21XY} should be smaller.

The results obtained from tables and discussions made above are encouraging enough as these depend on a one percent sample (f = 0.01) which is very small for practical purposes. If a comparatively large sample is taken, it is expected to have more gain in efficiency of $\hat{T}_2$ over $\bar{y}_2$ and $\bar{Y}_2$.

10. Concluding Remarks

Recapitulating what have been presented and discussed above, it can be said that using the information on an auxiliary variable, available on the current occasion, the estimation procedure of the mean of current occasion may certainly be improved. If a highly correlated auxiliary variable is used, relatively only a smaller fraction of sample on the second occasion is desired to be replaced by a fresh sample which, in turn, would reduce the total cost of the survey.

ACKNOWLEDGEMENTS

The authors thank the referees for their constructive suggestions and valuable comments.

REFERENCES

[1] Jessen, R.J. (1942). Statistical investigation of a sample survey for obtaining farm facts. Iowa Agricultural Experiment Station Road Bulletin No. 304, Ames, Iowa, USA, 1-104.

[2] Sen, A.R. (1971). Successive sampling with two auxiliary variables. *Sankhya*, **33, B,** 371-378.

[3] Sen, A.R. (1973). Some theory of sampling on successive occasions. *Aust. Jour. of Statistics,* **15,** 105-110.

[4] Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). *Sampling Theory of Surveys with Applications.* Iowa State University Press, Ames, Iowa, USA, 3rd Revised Edition.