SOME RESULTS CONCERNING STRATEGIES OF SAMPLING ON TWO OCCASIONS

BY

ARIJIT CHAUDHURI AND RAGHUNATH ARNAB

Indian Statistical Institute, Calcutta

(Received : November, 1977)

SUMMARY

We propose a few simple modifications on Avadhani and Sukhatme's (3) strategies of sampling a finite population on two successive occasions and examine situations when they may fare better than the latter. The results are derived mostly under two customary super-population models and occassionally with a few approximations and restrictive assumptions on variate-values.

1. INTRODUCTION

We consider a finite universe U of N units labelled 1, 2,..., i,..., N supposed to remain unchanged on two successive time-periods with variate-value y_i, x_i (i=1,..., N) defined on it with means $\bar{Y}$, $\bar{X}$, our problem being to estimate $\bar{Y}$, the mean for the current occasion on utilizing information gathered on two consecutive surveys undertaken on it. Among many sampling strategies available in the literature for the purpose we shall pay attention to a couple of particularly useful ones due to Avadhani and Sukhatme [3] and suggest additional strategies as modifications on them. The details concerning their relative efficiencies are given in what follows. In studying their comparative performances we consider two well-known models (along with necessary modifications) and point out different respective situations which are favourable to the respective strategies.

2. A BRIEF REVIEW OF AVADHANI AND SUKHATME'S [3] STRATEGIES

In one of their strategies to be called strategy I (or simply, I) on the first occasion and SRSWOR sample, S_1 of size n is taken from U and on the second occasion a sub-sample S_{2m} of size m is chosen from S_1 by following the Rao, Hartley and Cochran [12] scheme (R.H.C. scheme, in brief) using the normed size measure

$$p'_i = x_i / \sum_{i \in S_1} x_i \text{ for } i \in S_1$$

and an independent SRSWOR sample S_{2u} of size $u=n-m$ is chosen from the units of U not included in S_1. The estimator used is

$$t_1=a\bar{y}_{RHC}+(1-a)\,\bar{y}_u$$

where $\bar{y}_{RHC}$ is the usual R.H.C. estimator for

$$\frac{1}{n}\sum_{i\,\varepsilon\,S_1} y_i=\bar{Y}\text{ (say)}$$

based on y_i's for $i\,\varepsilon\,S_{2m}$' $\bar{y}_u$ is the mean based on S_{2u} and 'a' is a constant to be determined so as to minimize the variance $V(t_1)$ for a given value of $\lambda=m/n=1-\mu$ (say).

In their other strategy to be called strategy II (or simply II) S_1 and S_{2u} are chosen as in I but the matched sample S_{2m} on the second occasion is sub-sampled as an SRSWOR from S_1 and the estimator employed is

$$t_2=b\,\bar{y}_R+(1-b)\,\bar{y}_u$$

where

$$\bar{y}_R=\frac{1}{n}\Big[\sum_{i\,\epsilon\,S_{2m}} y_i/\sum_{i\,\epsilon\,S_{2m}} x_i\Big]\sum_{i\,\epsilon\,S_1} x_i'$$

the usual biased ratio-estimator for $\bar{y}$ (whose bias we shall throughout ignore and whose mean square error about $\bar{Y}$ will be taken as its variance supposing the necessary approximations are valid, without mentioning the fact in what follows) and b is a constant to be determined to minimize the variance $V(T_2)$. Following Cochran [7] and Avadhani-Srivastava [2, pp. 159, 161, equation (19)], we shall assume the sample-size n to be sfficiently large so that we may ignore the the bias of $\bar{y}_R$ and its mean square error about $\bar{Y}$ will be taken approximately equal to its variance. Writing

$$S_y^2=\frac{1}{N-1}\sum_1^N (y_i-\bar{Y})^2,\quad R=Y/X=N\bar{Y}/N\bar{X},$$

$$p_i=x_i/x,\ \forall\ =1,\ \ldots,\ N$$

$$\delta=\frac{1}{N}\sum\frac{1}{p_i}(y_i-Rx_i)^2/\sum(y_i-\bar{Y})^2$$

we note from Avadhani-Sukhatme [3], that

$$V(\bar{y}_{RHC})=(1/n-1/N)\,S_y^2+(1/m-1/n)\,\frac{1}{N(N-1)}\sum p_i^{-1}(y_i-Rx_i)^2$$

$$V(\bar{y}_u)=(1/u-1/N)\ S_y^2,$$

$$\text{Cov}(\bar{y}_u', \bar{y}_{RHC})=-S_y^2/N=V_{12}\ \text{(say)}$$

and writing

$$V_1'=V(\bar{y}_{RHC})-\text{Cov}(\bar{y}_u, \bar{y}_{RHC})$$

$$V_2'=V(\bar{y}_R)-\text{Cov}(\bar{y}_u, \bar{y}_{RHC})$$

the minimum (for variation in 'a' but λ remaining fixed) value of $V(t_1)$ is

$$V_{opt}(t_1)=\frac{V_1'\ V_2'}{V_1'+V_2'}+V_{12}'$$

the value of 'a' yielding this optimum being $a=1/2=a_0$ (say),

The optimum value of μ is $\mu_{opt}=(1+\sqrt{\delta})^{-1}$ yielding the minimum value for $V_{opt}(t_1)$ as

$$V_{min}(t_1)=\left(\frac{1+\sqrt{\delta}}{2n}-\frac{1}{N}\right)S_y^2.$$

The minimum value of the variance of t_2 for appropriate choices of b and μ turns out to

$$V_{min}(t_2)=\left(\frac{1+\sqrt{1-\rho^2}}{2n}-\frac{1}{N}\right)S_y^2$$

ρ being the correlation coefficient between y and x.

In comparing I and II Avadhani and Sukhatme [3] considered the following model :

$$M_1: y_i=\beta x_i+e_i \quad \forall\ i$$

such that

$$\sum_1^N e_i=0=\sum_1^N e_i x_i$$

and

$$\bar{e_i^2}=\sigma^2 x_i^g\ \forall\ i \text{ with } 0<\sigma^2<\infty,\ g\geqslant 0$$

where $\bar{e_i^2}$ is the average of e_i^2 in the array for which x_i is fixed.

Assuming this model they found that $V_{min}(t_1) \gtreqless V_{min}(t_2)$ according as $g \gtreqless 1$ the decision about adopting I or II resting on the availability of evidence supporting the hypothesis about g lying inside or outside (0,1).

3. Modifications of Avadhani-Sukhatme's [3] Strategies and their Properties

First we consider a strategy, to be called the strategy III, (or simply III) where on the first occasion S_1 is chosen as in I and II, on the 2nd occasion a matched sample S_{2m} of size m is chosen from S_1 following a π_{ps} design with p_i' s as normed size measure and an unmatched sample S_{2u} of size u is chosen from U-S_1 as in I and II. Writing

$$\pi_i = m\ p_i' \ \forall\ i$$

and

$$\hat{\bar{Y}}_m = \frac{1}{n} \sum_{i \in S_{2m}} y_i/\pi_i'$$

the unbiased (for $\bar{Y}$) estimator employed here is

$$t_3 = \phi \hat{\bar{y}}_m + (1-\phi)\ \bar{y}_u$$

with ϕ as a constant to minimize $V(t_3)$ for fixed λ.

Denoting (generically) by E_1, V_1 the operators for expectations and variances in respect of the sampling design adopted in choosing S_1 and E_2, V_2 the operators for conditional expectations and variances with respect to the sampling design adopted in choosing the matched sub-sample S_{2m} for given S_1' we have

$$V(\hat{\bar{Y}}_m) = E_1 V_2 (\hat{\bar{Y}}_m) + V_1 E_2 (\hat{\bar{Y}}_m)$$

$$= \frac{1}{n^2} E_1 \left[\sum_{i \in S_1} y_i^2 (1/\pi_i - 1) + \sum_{i \neq j} \sum_{\in S_1} y_i y_j (\pi_{ij}/\pi_i \pi_j - 1) \right]$$

$$+ (1/n - 1/N)\ S_y^2$$

where π_{ij}'s are the inclusion-probabilities of pairs of units in sub-sampling from S_1 according to a π_{ps} design. The term inside the square bracket above for different π_{ps} designs will be different and for many of the usual π_{ps} designs it is difficult to have an elegant expression for its expectation. We shall not present a more explicit expression for it but shall note in what follows that it is still possible to study the performances of the strategy III relative to I and II and later still, supply an elegant expression (*i*) for the expectation of $V(\hat{\bar{Y}}_m)$ with respect to a super-population model (introduced in section 4) and (*ii*) for an unbiased estimator for $V(\hat{\bar{Y}}_m)$ (vide section 5, remark V and VI).

Another strategy we consider will be labelled as strategy IV (or simply IV) where S_1 is drawn as earlier and a (matched) sub-sample S_{2m} of size m is drawn from S_1 by following Sen-Midzuno [15, 10] sampling scheme using the normed size measures p_i's and finally the unmatched sample S_{2u} is chosen by SRSWOR method from $U-S_1$ as in strategies I-III. The estimator for $\bar{Y}$ based on this sampling scheme to be used is

$$t_4=\psi\,\hat{\bar{Y}}_m'+(1-\psi)\,\bar{y}_u$$

where $\hat{\bar{Y}}_m'$ is formally same as $\bar{Y}_R$ in II and ψ is a constant to be chosen so as to minimize $V(t_4)$.

Formally,

$$V(\hat{\bar{Y}}_m')=E_1V_2(\hat{\bar{Y}}_m')+(1/n-1/N)\,S_y^2 \ .$$

An explicit expression for $V_2(\bar{Y}_m')$ is readily available, *i.g.* from T.J. Rao (13) but we will defer its presentation till section 4.

4. Comparison of Relative Efficiencies of the Strategies I-IV

Observing that the above-noted estimators t_i $(i=1, 2, 3, 4)$ are of the form

$$t_i=\alpha\, e_i+(1-\alpha)\, e_i'$$

with

$$E(e_i)=\bar{Y},\ i=1, \ldots, 4$$

where

$$e_i=\bar{y}_{RHC}',\ \bar{y}_R',\ \hat{\bar{y}}_m',\ \hat{\bar{Y}}_m'$$

and

$$e_i'=\bar{y}_u \quad \text{for} \quad i=1, \ldots, 4$$

and writing

$$V(e_i')=V_i^{(2)},\ \text{Cov}\ (e_i,\ e_i')=V_i^{(12)} \qquad \text{(say)}$$

we recall the well-known fact that the minimum (in respect of variation in α) value of the variance of t_i is of the form

$$V_{opt}(t_i)=\frac{w_i^{(1)}w_i^{(2)}}{w_i^{(1)}+w_i^{(2)}}+V_i^{(12)}$$

where

$$w_i^{(1)}=V_i^{(1)}-V_i^{(12)},\ w_i^{2}=V_i^{(2)}-V_i^{(12)}$$

Noting further in this case that $V_i^{(2)}=(1/u-1/N)\ S_y^2$ and $V_i^{(12)}=-S_y^2/N\ \forall i=1,\ldots,4$ so that $w_i^{(2)}$ and $V_i^{(12)}$ are fixed for $i=1,\ldots,4$ it follows that for any fixed λ.

$$V_{opt}\ (t_i)\gtreqless V_{opt}(t_j)\ \text{according as}\ w_i^{(1)}\gtreqless w_j^{(1)}$$

for $i, j=1, \ldots, 4\ (i\neq j)$.

So, the relative efficiencies of the strategies $I-IV$ are determined by the relative magnitudes of

$$V(\bar{y}_{RHC}),\ V(\bar{y}_R),\ V(\hat{\bar{Y}}_m)\ \text{and}\ V(\hat{\bar{Y}}'_m)$$

only. Again, the initial sample S_1 being chosen following the same design in each of strategies I—IV their relative efficiencies are determined by the relative magnitudes of the conditional variances

$$V_2(\bar{y}_{RHC}\mid S_1),\ V_2(\bar{y}_R\mid S_1),\ V_2(\hat{\bar{Y}}_m\mid S_1)\ \text{and}\ V(\hat{\bar{Y}}'_m\mid S_1)$$

for any given sample S_1 chosen on the first occasion. Bearing these points in mind, we state below the results concerning the relative efficiencies of the strategies I—IV one after another.

Theorem 1.

If we assume that n is so large that we may neglect the error in writing $\frac{1}{n-2}$ for $\frac{1}{n-1}$, then using the results in theorem I in Chaudhuri (4) we may assert that

$$V_2(\hat{\bar{y}}_m\mid S_1)<V_2(\bar{y}_{RHC}\mid S_1)$$

uniformly in S_1 and hence that $V_{opt}(t_3)<V_{opt}(t_1)$ provided S_{2m} is based on the modified Midzuno π_{ps} sampling scheme vide Chaudhuri [4] Mukhopadhyay [11] and Sankaranarayanan [14]. If we assume (which we shall call the assumption A) that the model M_1 holds not

only for the entire population U but for every sample S_1 of size n taken from U on the first occasion by SRSWOR method, then we get the following results.

Theorem 2.

If we assume A and neglect the error in writing

$$\bar{x}_{S_{2m}}=\frac{1}{m}\sum_{i\epsilon S_{2m}} x_i \text{ for } \bar{x}_{S_1}=\frac{1}{n}\sum_{i\epsilon S_1} x_i$$

for every sub-sample S_{2m} of S_1 for every S_1, then we have [recalling the results due to Avadhani and Srivastava (2) and theorem 3 in Chaudhuri (5)]

(*i*) $V_{opt}(t_4)=V_{opt}(t_2)$

(*ii*) $V_{opt}(t_4)=V_{opt}(t_2)\geqslant V_{opt}(t_1)$ if $g\geqslant 1$

and (*iii*) $V_{opt}(t_2)<V_{opt}(t_1)$ if $0\leqslant g<1$

Here we assume that not only n but also m is sufficiently large so that we may ignore the bias in $\bar{x}_{S_{2m}}$ and approximately Sukhatme [3] have also assumed this [vide (2) pp. 257, equation (35) onwards].

Theorem 3.

If the assumption A holds and if we ignore the error in writing $\frac{1}{n}$ for $\frac{1}{(n-1)}$ then assuming the asymptotic relationship considered by Asok and Sukhatme (1) we have

$$V_{opt}(t_3)\lesseqqgtr V_{opt}(t_1) \text{ according as } g\lesseqqgtr 1.$$

(for further clarification vide Remark III)

Theorem 4.

If the assumption A holds, then we have $V_{opt}(t_3)<V_{opt}(t_1)$ if $g\geqslant 2$, provided S_{2m} is selected following (modified) Midzuno π_{ps} sampling scheme. (The proof is given in Remark IV in section 5).

As opposed to the model M_1 and its extension inherent in assumption A one may assume the following alternative model M_2 considered earlier by Hanurav (7), Cochran (6) among others where we assume

$$M_2: Y_i=\beta x_i+e_i \qquad i=1, \ldots, N$$

where e_i's are random variables such that

$$\varepsilon(e_i)=0=\varepsilon(e_i e_j) \text{ for all } i, j\ (i\neq j)$$

$$\varepsilon(e_i^2)=\sigma^2 x_i^g,\ 0<\sigma^2<\infty g\geqslant 0,$$

ε denoting the operator for taking expectation with respect to the distribution of $e_i's$.

Assuming this model we shall compare the relative efficiencies of the strategies I—IV by comparing the minimum (with respect to α, the value of λ remaining fixed) values of $\varepsilon V(t)$ which we write as $\varepsilon_{opt} V(t)$ for $t=t_i'$ $i=1, ..., 4$. Clearly, their relative magnitudes are determined by the values of $\varepsilon V_2(t_i \mid S_1)$ for every fixed sample S_1 chosen as one may readily check.

Recalling Hanurav's [7] results and using the variance expression for $\hat{\bar{Y}}'_m$ given that T.J. Rao [12] it can be seen (vide Remark II in section 5) that

$$\varepsilon V(\hat{\bar{Y}}_m \mid S_1) \lesseqgtr \varepsilon V_2(\bar{y}_{RHC} \mid S_1) \lesseqgtr \varepsilon V_2(\hat{\bar{Y}}'_m \mid S)$$

according as $g \lesseqgtr 1$. So, we get

Theorem 5.

If the model M_2 holds, then it follows that

$$\varepsilon_{opt} V(t_3) \lesseqgtr \varepsilon_{opt} V(t_1) \lesseqgtr \varepsilon_{opt} V(t_4)$$

according as $g \lesseqgtr 1$. In Practice, it being well-known that most frequently $g > 1$ of the strategy III should be preferred in most of the situations.

Theorem 6.

If the model M_2 holds, then

$$\varepsilon_{opt} V(t_2) > \varepsilon_{opt} V(t_3) \text{ if } g \geqslant 1$$

[vide theorem 4 in Chaudhuri (5)]

Theorem 7.

If M_2 holds, and if we are justified to neglect the error in writing $\bar{x}_{S_{2m}}$ for $\bar{x}_{S'}$ for every $S_{2m} \subset S_1$ for each S_1 then

$$\varepsilon_{opt} V^{(t_4)} \lesseqgtr \varepsilon_{opt} V^{(t_3)} \text{ according as } g \lesseqgtr 1$$

[recalling theorem 5 in Chudhuri (5)]

5. A Few Remarks

Remark I. Writing $V_{opt}(t \mid \lambda)$, $\varepsilon_{opt} V(t \mid \lambda)$, the minimum [with respect to α values of $V(t)$ and $\varepsilon V(t)$ for a given λ and $V_{min}(t)$, $\varepsilon_{min} V(t)$] their minimum values over variation in λ, it readily follows that

$$V_{opt}(t_i \mid \lambda) \gtrless V_{opt}(t_j \mid \lambda) \Rightarrow V_{min}(t_i) \gtrless V_{min}(t_j)$$

and

$$\varepsilon_{opt} (V(t_i \mid \lambda) \lesseqgtr \varepsilon_{opt} V(t_t \mid \lambda) \Rightarrow \varepsilon_{min} V(t_i) \lesseqgtr \varepsilon_{min} V(t_j)$$

So, the results relating to the comparative efficiencies of the strategies I—IV given earlier have the obvious extended interpretation.

Remark II. Writing the variances of the usual estimators for a finite population total based on a single-sample selected on one occasion only according to the strategies due to Rao, Hartley and Cochran, Midzuno-Sen and Horvitz-Thompson (using a π_{ps} design) as V_1, V_3, V_4 respectively and their expected values assuming M_2 as E_1, E_3, E_4, respectively it is known that [vide Hanurav (8) and Rao (12)]

$$\left.\begin{array}{ll} (i) & E_3 \lesseqgtr E_1 \\ (ii) & E_3 \lesseqgtr E_4 \end{array}\right\} \text{according as } g \lesseqgtr 1.$$

Also it is known, [vide Chaudhuri and Arnab (6)]

$$E_3 < E_1 < E_4 \quad \text{if } g > 1$$

$$E_3 > E_1 > E_4 \quad \text{if } g < 1$$

and

$$E_3 = E_1 = E_4 \quad \text{if } g = 1$$

Remark III. Recently, Asok and Sukhatme [1] have considered an asymptotic approximation for the variance of the Horvitz-Thompson estimator based on Sampford's π_{ps} sampling scheme to get the approximate variance expression V_3' (say) as

$$V_3' = \frac{1}{n} \Sigma p_i t_i^2 \{1-(n-1) p_i\}$$

where

$$t_i = \frac{y_i}{p_i} - Y,$$

y_i's are the values of the variate under study,

$$Y = \Sigma y_i, \quad p_i = \frac{x_i}{x, x_i}$$

being a size measure for the i—th unit ($i=1, \ldots, N$) and $x = \Sigma X_i$ then, it can be seen under the model M_1 that

$$\varepsilon (V_1) - \varepsilon (V_3') = \varepsilon \frac{n-1}{n} \Sigma p_i t_i^2 \left(p_i - \frac{1}{N-1}\right)$$

$$\simeq \frac{n-1}{n} \varepsilon \Sigma p_i t_i^2 \left(p_i - \frac{1}{N}\right)$$

$$=\sigma^2\frac{n-1}{n} N \text{ Cov } (x_i^{g-1}, x_i)$$

implying that $\varepsilon(V_1) \lesseqgtr \varepsilon(V_3')$ according as $g \lesseqgtr 1$.

Remark IV. Following Chaudhuri [4] one can see that using the same normed size measures p_i's for a Rao-Hartley-Cochran estimator and Horvitz-Thompson estimator (based on π_{ps} design following modified Midzuno's sampling scheme) in estimating a finite population total one has

$$V\text{(HTE}-V\text{(RHC Estimator)}$$

$$\leqslant \frac{n-1}{(N-1)\, n\, (N-2)}\left[\Sigma\left(\frac{y_i}{p_i}-Y\right)^2 - N\,\Sigma\, p_i\left(\frac{y_i}{p_i}-Y\right)^2\right]$$

$$=\frac{n-1}{n}\,\frac{1}{(N-1)}\,\frac{\sigma^2 x}{(N-2)}\,[X\,\Sigma\, x_i^{g-2} - N\,\Sigma\, x_{k_i}^{g-1}]$$

$$\leqslant 0 \text{ if } g \geqslant 2 \qquad (\text{provided } M_1 \text{ holds})$$

Remark V. If M_2 holds, then for any π_{ps} design we have

$$\epsilon V(\hat{Y}'_m) = \epsilon[E_1\, V_2\, (\hat{Y}'_m) + V_1\, E_2\, (\hat{Y}'_m)]$$

$$=\frac{\sigma^2}{n^2}\, E_1 \sum_{i\epsilon S_1} x_i^g \left(\frac{1}{\pi_i}-1\right)$$

$$+\left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{\sigma^2}{N}\,\Sigma\, x_i^g + \beta^2 S_x^2\right)$$

[where $(N-1)\, S_x^2 = \sum_1^N (x_i-\bar{x})^2$]

$$=\frac{\alpha^2}{n^2} E_1 \left\{\frac{1}{m}\sum_{i\epsilon S_1} x_i^g + \frac{1}{m}\sum\sum_{i\neq j\epsilon S_1} x_i^{g-1}\, x_j\right.$$

$$\left. - \sum_{i\epsilon S_1} x_i^g\right\} + \left(\frac{1}{n}-\frac{1}{N}\right)\left\{\frac{\sigma^2}{N}\sum_1^N x_i^g + \beta^2 S_x^2\right\}$$

(remembering $\pi_i = m\, x_i / \sum_{i\epsilon S_1} x_i$)

$$=\frac{\alpha^2}{nN}\left[\frac{1}{m}\left\{\sum_1^N x_i^g + \frac{n-1}{N-1}\sum\sum_{i\neq j=1} x_i^{g-1}\, x_j\right\} - \sum^N x_i^g\right]$$

$$+\left(\frac{1}{n}-\frac{1}{N}\right)\left\{\frac{\sigma^2}{N}\sum_1^N x_i^g + \beta^2\, S_x^2\right\}.$$

Remark VI. An expression for the unbiased estimator of $V(\hat{\bar{Y}}_m)$ is given by

$$v(\hat{\bar{Y}}_m)=\left[1-\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\right]^{-1}\left[\frac{1}{n^2}\sum_{i<j\epsilon S_{2m}}\sum\frac{\pi_i\pi_j-\pi_{ij}}{\pi_{ij}}\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\left\{\frac{1}{n}\sum_{i\epsilon S_{2m}}\frac{y_i^2}{\pi_i}-\left(\sum_{i\epsilon S_{2m}}\frac{y_i}{n\pi_i}\right)^2\right\}\right]$$

Proof.

$$Ev(\hat{\bar{Y}}_m)=\left[1-\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\right]^{-1}\left[\frac{1}{n^2}E_1\sum_{i<j\epsilon S_1}\sum(\pi_i\pi_j-\pi_{ij})\cdot\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\cdot\left\{\frac{1}{N}\Sigma y_i^2-V\left(\sum_{i\epsilon S_{2m}}\frac{y_i}{n\pi_i}\right)-\bar{Y}^2\right\}\right]$$

$$=\left[1-\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\right]^{-1}\left[\frac{1}{n^2}E_1\sum_{i<j\epsilon S_1}\sum(\pi_i\pi_j-\pi_{ij})\cdot\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\frac{\Sigma(y_i-\bar{Y})^2}{N}-\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}V\left(\sum_{i\epsilon S_{2m}}\frac{y_i}{n\pi_i}\right)\right]$$

$$=\left[1-\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}\right]^{-1}\left[\frac{1}{n^2}E_1\sum_{i<j\epsilon S_1}\sum(\pi_i\pi_j-\pi_{ij})\cdot\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2+\left(\frac{1}{n}-\frac{1}{N}\right)S_y^2-\left(\frac{1}{n}-\frac{1}{N}\right)\frac{N}{N-1}V\left(\sum_{i\epsilon S_{2m}}\frac{y_i}{n\pi_i}\right)\right]$$

$$=V(\hat{\bar{Y}}_m).$$

Acknowledgement. The authors are grateful to the referee for his helpful suggestions.

REFERENCES

[1] Asok, G. and Sukhatme, B.V. (1976) : On Sampford's procedure of unequal probability sampling without replacement. *J. Amer. Statist. Assoc.* 71, pp. 912-918.

[2] Avadhani, M.S. and Srivastava, A.K. (1972) : A comparision on Midzuno-Sen scheme with pps sampling without replacement and its application to successive sampling. *Ann. Inst. Stat. Math.*, 24, pp. 153-164.

[3] Avadhani, M.S. and Sukhatme, B.V. (1970) : A comparison of two sampling procedures with an application to successive sampling. Jour. Roy. Statist. Soc (C), Applied statistics, **19**, pp. 231-259.

[4] Chaudhuri, Arijit (1974) : On some properties of the sampling scheme due to Midzuno, *Cal. Stat. Assoc.Bull.* **23**, p. 1-19.

[5] —— (1977) : On some properties of the Horvitz-Thompson estimator based on Midzuno's π_{ps} sampling scheme. *Jour. Ind. Soc. Agri. Statist.* **29**, pp. 42-52.

[6] Chaudhuri, Arijit and Arnab, Raghunath (1979) : On the relative efficiencies of sampling strategies under a superpopulation model. *sankhya*, Ser C, **41** p. 40-43.

[7] Cochran, W.G. (1963) : *Sampling Techniques*, 2nd Edition, John Wiley and Son Ltd., New York.

[8] Hanurav, T.V. (1967) : Optimum utilisation of auxiliary information : π_{ps} sampling of two units from a stratum. *J. Roy. Statist. Soc.* Ser B, 29, pp 374-390.

[9] Horvitz, D.G. and Thompson, D.J. (1952) : A generalization of sampling without replacement from a finite universe. *J. Amer. Stati st. Asso.* 47, pp. 663-685.

[10] Midzuno, H. (1952) : On the sampling system with probability proportional to sum of sizes. *Ann. Inst. Statist. Math*, 15, pp. 67-73.

[11] Mukhopadhyay, Parimal (1974) : π_{ps} sampling schemes to base H.T.E. *Cal. Statist. Assoc. Bull*, 20, pp. 20-24.

[12] Rao, J.N.K., Hartley, H.O. and Cochran, W.G. (1962) : On a simple procedure of unequal probability sampling without replacement, *Jour. Roy. Statist. Soc. B*, 24, pp. 482-491.

[13] Rao, T.J. (1967) : On the choice of a strategy for the ratio method of estimator, *J. Roy, Statist. Soc*, Ser B, pp. 392-397.

[14] Sankaranarayan, K. (1969) : An IPSS sampling scheme using Lahiri's method of selection. *Jour. Ind. Soc. Agr. Statist.* **21**, pp. 58-66.

[15] Sen, A.R. (1953) : On the estimation of the variance in sampling with varying prababilities, *J. Ind. Soc. Agr. Statist.* 5, pp. 119-127.