ASYMPTOTIC EFFICIENCY OF DES RAJ'S STRATEGY-I

By P. K. PATHAK

University of Illinois and *Indian Statistical Institute*

SUMMARY. The main object of this paper is to study the asymptotic efficiency of Des Raj's strategy of estimating the population total. This strategy, namely Des Raj's estimator of the population total under sampling with unequal probabilities (without replacement), was introduced in the literature by Des Raj (1956). Subsequently this strategy was found to be more precise than the standard strategy of estimating the population total under sampling with unequal probabilities (with replacement) (Roychoudhury, 1956; Sethi, 1962). In this paper an exact expression and an upper bound to the variance of Des Raj's estimator are derived. These results are then used to compare Des Raj's strategy with other similar strategies of estimating the population total. Under certain regularity conditions Des Raj's strategy is found to be more precise than other strategies.

1. Introduction and statement of the problem

In the sequel a strategy will mean an estimation procedure. For estimating the population total a strategy is thus a sampling scheme and an estimator for the population total under the sampling scheme.

One of the main problems of the theory of sampling from finite populations is to search for efficient strategies of estimating the population total. In the past few years several authors have introduced and studied the efficiency of a number of strategies of estimating the population total (see, for example, Des Raj, 1956; H'jek, 1964; Hartley and Rao, 1962; Murthy, 1957; Rao, Hartley and Cochran, 1962; Sampford, 1962; and Stevens, 1958). In 1956 Des Raj introduced a simple strategy of estimating the population total involving sampling with unequal probabilities (without replacement). Unfortunately we do not yet know much about the efficiency of this strategy, although we know a great deal about the efficiencies of other similar strategies (see Godambe, 1955; Godambe and Joshi, 1965; Hájek, 1964; Hartley and Rao, 1962; Pathak, 1964; Rao, Hartley and Cochran, 1962; Roychoudhuri, 1956; Sampford, 1962; and Sethi, 1962 in this connection). The main purpose of this paper is to study the asymptotic efficiency of Des Raj's strategy. Our line of approach is to first obtain a simple upper bound close to the variance of Des Raj's estimator. We then compare this bound with the variance expressions involved in other strategies. We begin this study in a very general manner.

For each integer k, consider a finite population $\pi_k = (U_{k1}, U_{k2}, \ldots, U_{kN_k})$ of N_k units, where U_{kj} denotes the j-th population unit of π_k. Unless otherwise stated the subscripts j and j' run from 1 through N_k. Let P_{kj} and Y_{kj} be the probability of selection and the value of a Y-characteristic of U_{kj} respectively. Let $s_k = (u_{k1}, \ldots, u_{ki}, \ldots, u_{kn_k})$ denote a sample of size n_k selected from π_k according to the above probabilities and without replacement, u_{ki} being the sample unit selected from π_k at the i-th draw. Unless otherwise stated the subscripts i and i' run from 1 through n_k.

Under this sampling scheme Des Raj (1956) introduced the following unbiased and uncorrelated estimators of the population total $Y_k = \sum_j Y_{kj}$.

$$t_{k1} = y_{k1}/p_{k1},$$

$$t_{ki} = (y_{ki}/p_{ki})(1-p_{k1} - \ldots -p_{k(i-1)})+y_{k1}+\ldots+y_{k(i-1)} \quad (i = 2, \ldots, n_k), \quad \ldots \quad (1)$$

where p_{ki} and y_{ki} respectively denote the probability of selection and the Y-characteristic value of u_{ki}.

Des Raj (1956) considered $\bar{t}_{k(d)} = n_k^{-1} \sum_i t_{ki}$ for estimating the population total Y_k. Using the fact that $t_{k1}, \ldots, t_{kn_k}$ are uncorrelated, Des Raj (1956) suggested $v(\bar{t}_{k(d)}) = [n_k(n_k-1)]^{-1} \sum_i (t_{ki}-\bar{t}_{k(d)})^2$ as an unbiased estimator of $V(\bar{t}_{k(d)})$.

Remark : It may be noted that Des Raj's estimator, $\bar{t}_{k(d)}$, is introduced here through a sequence of populations $\{\pi_k : k = 1, 2, \ldots\}$. This sequence $\{\pi_k : k = 1, 2, \ldots\}$ has, of course, nothing to do with the definition of Des Raj's estimator. In practice there is only one population. The sequence of populations considered herein is meant only for making asymptotic comparisons in subsequent sections.

Clearly Des Raj's estimator is simple to compute and possesses a simple non-negative variance estimator. But in spite of these desirable properties, Des Raj's estimator has attracted very little attention in the literature. The main reason for this is probably the lack of information about its efficiency and the non-availability of an explicit expression for $V(\bar{t}_{k(d)})$. It would, therefore, be worthwhile to derive an expression for $V(\bar{t}_{k(d)})$ in a closed form. The following theorem is an attempt in this direction.

Theorem 1 :

$$V(t_{ki}) = \frac{1}{2} \sum_{j,j'} [(Y_{kj}/P_{kj})-(Y_{kj'}/P_{kj'})]^2 \, P_{kj} \, P_{kj'} \, Q_{kjj'} \, (i-1) \quad \ldots \quad (2)$$

where $Q_{kjj'}$ $(i-1)$ denotes the probability of non-inclusion of U_{kj} and $U_{kj'}$ in the first $(i-1)$ sample units selected from π_k; $Q_{kjj'}$ $(0) = 1$.

Proof : Suppose after i units $u_{k1}, \ldots, u_{ki}$ have been selected and recorded with unequal probabilities and without replacement from π_k, the i-th unit u_{ki} is replaced to the population. Then another unit $u_{ki'}$ is drawn independently of u_{ki} from π_k according to the original sampling scheme. It is evident that when $u_{k1}, \ldots, u_{k(i-1)}$ are given, u_{ki} and $u_{ki'}$ are conditionally independent and identically distributed. Therefore

$$V(t_{ki}) = E[V(t_{ki}\,|\,u_{k1}, \ldots, u_{k(i-1)})]+V[E(t_{ki}\,|\,u_{k1}, \ldots, u_{k(i-1)})]$$

$$= E[V(t_{ki}\,|\,u_{k1}, \ldots, u_{k(i-1)})]$$

$$= E[E\{\tfrac{1}{2}(t_{ki}-t_{ki'})^2\,|\,u_{k1}, \ldots, u_{k(i-1)}\}] \quad \ldots \quad (3)$$

where $t_{ki'} = (y_{ki'}/p_{ki'})(1-p_{k1}-\ldots-p_{k(i-1)})+y_{k1}+\ldots+y_{k(i-1)}$, the estimator of the population total, Y_k, based on u_{ki}.

$$= E\Big[E\left\{\frac{1}{2}\left(\frac{y_{ki}}{p_{ki}} - \frac{y_{ki'}}{p_{ki'}}\right)^2 (1-p_{k1}- \ldots -p_{k(i-1)})^2\,|\,u_{k1}, \ldots, u_{k(i-1)}\right\}\Big]$$

$$= E\Big[\frac{1}{2} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj}P_{kj'} \, \alpha_{kjj'(i-1)}\Big]$$

where

$$\alpha_{kjj'}(i-1) = \begin{cases} 1 & \text{if the first } (i-1) \text{ sample units do not contain } U_{kj} \text{ and } U_{kj'}, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$V(t_{ki}) = \frac{1}{2} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj} P_{kj'} Q_{kjj'}(i-1) \qquad \dots \quad (4)$$

where $Q_{kjj'}(i-1)$ has been defined in (2). This completes the proof.

It is easily seen that $Q_{kjj'}(i)$ involved in $V(t_{ki})$ in (4) is given by the following expression.

$$Q_{kjj'}(i) = \Sigma' \; p_{k1} \frac{p_{k2}}{(1-p_{k1})} \cdot \dots \cdot \frac{p_{ki}}{(1-p_{k1}-\dots-p_{k(i-1)})} \qquad \dots \quad (5)$$

where the sum, Σ', is taken over all permutations of $p_{k1}, \dots, p_{ki}$ chosen from $P_{k1}, \dots, P_{kN_k}$ but not including P_{kj} and $P_{kj'}$.

Now it can be seen by induction over i that

$$Q_{kjj'}(i) \leqslant (1-P_{kj}-P_{kj'})^i. \qquad \dots \quad (6)$$

The following corollaries are now simple consequences of the above inequality.

Corollary 1.1 :

$$V(t_{ki}) \leqslant \frac{1}{2} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj} P_{kj'} (1-P_{kj}-P_{kj'})^{i-1}. \qquad \dots \quad (7)$$

Corollary 1.2 :

$$V(\bar{t}_{k(d)}) = \frac{1}{2n_k^2} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj} P_{kj'} [1 + \sum_{i=2}^{n_k} Q_{kjj'}(i-1)]$$

$$\leqslant \frac{1}{2n_k^2} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj} P_{kj'} \frac{[1-(1-P_{kj}-P_{kj'})^{n_k}]}{(P_{kj}+P_{kj'})} \qquad \dots \quad (8)$$

$$\leqslant \frac{1}{2n_k} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj} P_{kj'} \Big[1 - \frac{(n_k-1)}{2} (P_{kj}+P_{kj'}) + \frac{(n_k-1)(n_k-2)}{6} (P_{kj}+P_{kj'})^2\Big]. \qquad \dots \quad (9)$$

The proof of the second corollary is immediate on noting that $t_{k1}, \dots, t_{kn_k}$ are uncorrelated.

The above inequality provides us a simple upper bound to $V(\bar{t}_{k(d)})$. From (9) it is clear that if $\max\limits_{1 \leqslant j \leqslant N_k} P_{kj} = 0(N_k^{-1})$ then

$$V(\bar{t}_{k(d)}) \leqslant \frac{1}{n_k} \sum_j P_{kj} \left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2 -$$

$$\left[1+0\left(\frac{n_k}{N_k}\right)\right] \left[\frac{(n_k-1)}{2n_k}\right] \left[\left(\sum_j P_{kj}^2\right) \sum_j P_{kj} \left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2 + \sum_j P_{kj}^2 \left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2\right]. \qquad \dots \quad (10)$$

It may be of some interest to note that the above bound provides a satisfactory approximation to $V(\bar{t}_{k(d)})$ if N_k is large and n_k is relatively small compared to N_k. Since the proof of this is a little involved it has not been presented here. It is, nevertheless, easy to check that the above bound is attained when $n_k = 2$ and agrees with $V(\bar{t}_{k(d)})$ up to terms of the order $0(N_k^{-1})$ in case of equal probability sampling.

Having obtained a simple upper bound close to $V(\bar{t}_{k(d)})$, the study of the asymptotic efficiency of Des Raj's strategy now becomes a distinct possibility. Subsequent sections, in this paper, are therefore devoted to the study of the relative efficiency of Des Raj's strategy as compared to other well-known strategies of similar kind.

In the next section we compare Des Raj's (1956) strategy with Sampford's (1962) strategy.

2. Comparison with Sampford's strategy

Sampford (1962) considered the following sampling scheme in order to select a sample with n_k distinct population units from π_k: Units are drawn one by one from π_k until the sample for the first time contains (n_k+1) different population unit. The last sample unit is not recorded in the sample to insure simplicity in the estimation procedure. The sampling scheme is called inverse sampling with unequal probabilities.

Under this sampling scheme Sampford (1962) considered the following unbiased estimator of the population total

$$\bar{t}_{k(s)} = \frac{1}{r_k} \sum_{i=1}^{r_k} \frac{y_{ki}}{p_{ki}} \quad \dots \quad (11)$$

where r_k is the recorded sample size and p_{ki} and y_{ki} respectively being the probability of selection and the Y-characteristic value of the i-th sample unit.

The variance of the above estimator is given by Pathak (1964)

$$V(\bar{t}_{k(s)}) = \frac{1}{2} \sum_{j,j'} \left(\frac{Y_{kj}}{P_{kj}} - \frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj} P_{kj'} E\left[\frac{1}{r_k}\middle| u_{k1} = U_{kj}, u_{k2} = U_{kj'}\right]. \quad \dots \quad (12)$$

Now to compare Des Raj's and Sampford's strategies, we compare $V(\bar{t}_{k(d)})$ with $V(\bar{t}_{k(s)})$. At this point it is worth mentioning that Des Raj's strategy is known to be more precise than Sampford's strategy (i.e. $V(\bar{t}_{k(d)}) < V(\bar{t}_{k(s)})$) in two special cases: (a) when $n_k = 2$ and (b) when $P_{kj} = N_k^{-1}$ for all j (Pathak, 1964).

Under certain regularity conditions we show that Des Raj's strategy is more precise than Sampford's strategy. This is achieved by finding two numbers v_1 and v_2 which satisfy the chain of inequalities: $V(\bar{t}_{k(d)}) \leqslant v_2 \leqslant v_1 \leqslant V(\bar{t}_{k(s)})$. The required v_2 is obtained from (10). The following theorem furnishes v_1.

Theorem 2: *Let*

(i) $N_k \to \infty$ as $k \to \infty$,

(ii) $\max_j P_{kj} = 0(N_k^{-1})$ *and*

(iii) $\overline{\lim}\; n_k < \infty$.

Then

$$V(\bar{t}_{k(s)}) \geqslant \frac{1}{n_k}\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2-$$

$$[1+0(N_k^{-1})]\left[\frac{(n_k-1)}{2(n_{k+1})}\right]\left[\left(\sum_j P_{kj}^2\right)\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2+\frac{2}{n_k}\sum_j P_{kj}^2\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2\right] \quad \ldots (13)$$

Proof: From (12) we have

$$V(\bar{t}_{k(s)}) = \frac{1}{2}\sum_{j,j'}\left(\frac{Y_{kj}}{P_{kj}}-\frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj}P_{kj'}\,E\left[\frac{1}{r_k}\middle| u_{k1}=U_{kj}, u_{k2}=U_{kj'}\right]$$

$$\geqslant \frac{1}{2}\sum_{j,j'}\left(\frac{Y_{kj}}{P_{kj}}-\frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj}P_{kj'}[n_k^{-1}P(r_k=n_k\,|\,u_{k1}=U_{kj}, u_{k2}=U_{kj'})+$$

$$(n_{k+1})^{-1}P(r_k=n_k+1\,|\,u_{k1}=U_{kj}, u_{k2}=U_{kj'})]. \quad \ldots (14)$$

Clearly

$$P(r_k=n_k\,|\,u_{k1}=U_{kj}, u_{k2}=U_{kj'}) = \Sigma' p_{ki_1}p_{ki_2}\cdots p_{ki_{(n_k-1)}} \quad \ldots (15)$$

where Σ' stands for the summation over all permutations of $p_{ki_1}, p_{ki_2}, \ldots, p_{ki_{(n_k-1)}}$ chosen from $P_{k1}, \ldots, P_{kN_k}$ but not including P_{kj} and $P_{kj'}$.

From Lemma 7.1, given in the appendix, it follows that

$$P(r_k=n_k\,|\,u_{k1}=U_{kj}, u_{k2}=U_{kj'}) \geqslant T_{k1}^{n_k-1}-\binom{n_k-1}{2}T_{k2}T_{k1}^{n_k-3} \quad \ldots (16)$$

where $$T_{km} = \left(\sum_{l=1}^{N_k} P_{kl}^m\right)-P_{kj}^m-P_{kj'}^m. \qquad (m=1, 2, \ldots).$$

Similarly

$$P(r_k=n_k+1\,|\,u_{k1}=U_{kj},\ u_{k2}=U_{kj'})$$

$$=(n_k-1)(P_{kj}+P_{kj'})\Sigma'\ p_{ki_1}\cdots p_{ki_{(n_k-1)}}+\binom{n_k-1}{2}\Sigma' p_{ki_1}^2 p_{ki_2}p_{ki_{(n_k-1)}} \quad \ldots (17)$$

where Σ' has a meaning similar to that described in (15).

Again it follows from Lemma 7.1 that

$$P(r_k=n_k+1\,|\,u_{k1}=U_{kj}, u_{k2}=U_{kj'})$$

$$\geqslant (n_k-1)(P_{kj}+P_{kj'})\left[T_{k1}^{n_k-1}-\binom{n_k-1}{2}T_{k2}T_{k1}^{n_k-3}\right]+$$

$$\binom{n_k-1}{2}\left[T_{k2}T_{k1}^{n_k-2}-\binom{n_k-2}{1}T_{k3}T_{k1}^{n_k-3}-\binom{n_k-3}{2}T_{k2}^2T_{k1}^{n_k-4}\right]. \quad \ldots (18)$$

Now from (14), (16) and (18), it follows that

$$V(\bar{l}_{k(s)}) \geqslant \frac{1}{2}\sum_{j,j'}\left(\frac{Y_{kj}}{P_{kj}}-\frac{Y_{kj'}}{P_{kj'}}\right)^2 P_{kj}P_{kj'}\Big[T_{k1}^{n_k-1}-\binom{n_k-1}{2}T_{k2}T_{k1}^{n_k-3}+$$

$$\frac{1}{(n_k+1)}\left((n_k-1)(P_{kj}+P_{kj'})T_{k1}^{n_k-1}+\binom{n_k-1}{2}T_{k2}T_{k1}^{n_k-2}\right)(1+0(N_k^{-1}))\Big] \quad \ldots \quad (19)$$

because $\overline{\lim}\ n_k < \infty$ and $\max_j P_{kj} = 0(N_k^{-1})$ imply that $T_{km} = 0(N_k^{-m+1})$.

Some simplification on expanding powers of $T_{k1} = (1-P_{kj}-P_{kj'})$ in powers of $(P_{kj}+P_{kj'})$ leads to the following inequality.

$$V(\bar{l}_{k(s)}) \geqslant \frac{1}{n_k}\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2-$$

$$[1+0(N_k^{-1})]\left[\frac{(n_k-1)}{2(n_k+1)}\right]\left[\left(\sum_j P_{kj}^2\right)\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2-\frac{2}{n_k}\sum_j P_{kj}^2\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2\right]. \quad \ldots \quad (20)$$

This completes the proof.

Remark : A weakness in the derivation of the above mentioned lower bound to $V(\bar{l}_{k(s)})$ is the assumption $\overline{\lim}\ n_k < \infty$ since asymptotically inverse sampling becomes equivalent to with replacement sampling when (i) $\overline{\lim}\ n_k < \infty$, (ii) $\max P_{kj} = 0(N_k^{-1})$ and (iii) $N_k \to \infty$. But in spite of this we feel that Theorem 2 provides a better approximation to $V(\bar{l}_{k(s)})$ than does the approximation of inverse sampling by with replacement sampling. The reason perhaps is the following : In order to get an asymptotic expression for $V(\bar{l}_{k(s)})$ if we approximate inverse sampling by with replacement sampling, we would approximate $E(r_k^{-1} \mid u_{k1} = U_{kj},\ u_{k2} = U_{kj'})$ by $n_k^{-1}P(r_k = n_k \mid u_{k1} = U_{kj}, u_{k2} = U_{kj'}) \simeq n_k^{-1}$, whereas in the approximation used in Theorem 2 we go a step further and obtain a lower bound to $V(\bar{l}_{k(s)})$ by using the inequality $E(r_k^{-1} \mid u_{k1} = U_{kj}, u_{k2} = U_{kj'}) \geqslant n_k^{-1}P(r_k = n_k \mid u_{k1} = U_{kj}, u_{k2} = U_{kj'})+(n_k+1)^{-1}P(r_k = n_k+1 \mid u_{k1} = U_{kj}, u_{k2} = U_{kj'})$. It is, therefore, hoped that generally our approximation would be more accurate.

The following theorem asserts that Des Raj's strategy is, in a certain sense, more precise than Sampford's strategy under some regularity conditions.

Theorem 3 : *If*

(i) $N_k \to \infty$ *as* $k \to \infty$,

(ii) $1 < \underline{\lim}\ n_k \leqslant \overline{\lim}\ n_k < \infty$,

(iii) $\max_j P_{kj} = 0(N_k^{-1})$ *and*

(iv) $\underline{\lim} \sum_j P_{kj}(Y_{kj}/P_{kj}-Y_k)^2 > 0$,

then

$$\underline{\lim}\ N_k[V(\bar{l}_{k(s)})-V(\bar{l}_{k(d)})] > 0. \quad \ldots \quad (21)$$

Proof: Under the above assumptions, we have from (10) and (20)

$$\underline{\lim}\, N_k[V(\bar{t}_{k(s)})-V(\bar{t}_{k(d)})] = \underline{\lim}\ [1+0(N_k^{-1})]$$

$$\left[\frac{(n_k-1)N_k}{2n_k(n_k+1)}\right]\left[\left(\sum_j P_{kj}^2\right)\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2+(n_k-1)\sum_j P_{kj}^2\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2\right]$$

$$\geqslant \underline{\lim}\ \frac{(n_k-1)}{2n_k(n_k+1)}\ \sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}}-Y_k\right)^2 > 0 \qquad \text{... (22)}$$

by assumptions (ii) and (iv). This completes the proof.

A few points about the regularity conditions of the above theorem are worth noting. It is easily seen that in pps sampling schemes the assumption (iii) is satisfied if the sizes, X_{kj}'s, on which P_{kj}'s are based, are uniformly bounded for all k and j and $\underline{\lim}\, N_k^{-1} \sum_j X_{kj} > 0$. The assumption (iv) is presumably needed to insure that the sampling distribution of the variate (Y/P) in the population π_k does not tend to a one point distribution as $k \to \infty$. Perhaps all unbiased estimators of Y_k will be equally efficient in the limit if (iv) were not satisfied. The assumption (iv) thus seems necessary in order to rule out such a triviality. It is believed that the assumption $\overline{\lim}\ n_k < \infty$ of (ii) could perhaps be replaced by an assumption of the kind $\overline{\lim}\ (n_k/N_k) = 0$. Unfortunately a rigorous proof of the validity of Theorem 3 under the relaxed condition, $\overline{\lim}\ (n_k | N_k) = 0$, is known to us only in the simplest case of sampling with equal probabilities (Pathak, 1964). Further work on these lines would, therefore, be interesting and rewarding.

We now compare Des Raj's strategy with the Rao-Hartley-Cochran strategy.

3. Comparison with the Rao-Hartley-Cochran strategy

The Rao-Hartley-Cochran strategy (1962) of estimating the population total involves the following sampling procedure: To select a sample of size n_k from a population of size $N_k = n_k q_k + r_k (0 < r_k \leqslant n_k)$, the population is first divided at random into n_k groups of sizes $N_{k1}, \ldots, N_{kn_k}$ such that $N_{k1} = N_{k2} = \ldots = N_{kr_k} = q_k+1$, $N_{k(r_k+1)} = \ldots = N_{kn_k} = q_k$. A sample of size one is then selected from each of the groups with probabilities proportional to the P_{kj}.

The conditional probability of selecting the j-th unit given that it is in the i-th group G_i, is given by $P_{kj}/(\sum_{G_i} P_{kj})$ where the summation $\sum_{G_i}$ is taken over the units in G_i.

As an estimator of the population total, Rao, Hartley and Cochran (1962) consider

$$\bar{t}_{k(rhc)} = \sum_i \frac{y_{ki}}{p_{ki}} \left(\sum_{G_i} P_{kj}\right) \qquad \text{... (23)}$$

where Σ stands for the summation over n_k units selected from the n_k groups.

The sampling variance of the above estimator is given by

$$V(\bar{t}_{k(rhc)}) = \left[\frac{1}{n_k} - \frac{(n_k-1)}{n_k(N_k-1)} + \frac{r_k(n_k-r_k)}{n_k N_k(N_k-1)}\right]\left[\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2\right]$$

$$= \frac{1}{n_k}\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2 - \frac{(n_k-1)}{n_k N_k}\left[1+0\left(\frac{n_k}{N_k}\right)\right]\left[\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2\right]. \quad \ldots \quad (24)$$

A comparison of (10) and (24) leads to the following theorem.

Theorem 4 : *Let*

(i) $N_k \to \infty$ *as* $k \to \infty$,

(ii) $\underline{\lim}\ n_k > 1$ *and* $\overline{\lim}\ (n_k|N_k) = 0$,

(iii) $\max_j P_{kj} = 0(N_k^{-1})$,

(iv) $\underline{\lim}\ (N_k \sum_j P_{kj}^2 - 2) > 0$ *and*

(v) $\underline{\lim} \sum_j P_{kj}(Y_{kj}/P_{kj} - Y_k)^2 > 0$.

Then

$$\underline{\lim}\ N_k[V(\bar{t}_{k(rhc)}) - V(\bar{t}_{k(d)})] > 0. \quad \ldots \quad (25)$$

Proof : Under the above assumptions, the following inequality is easily derived from (10) and (24).

$$\underline{\lim}\ N_k[V(\bar{t}_{k(rhc)}) - V(\bar{t}_{k(d)})] \geqslant \underline{\lim}\left[\frac{(n_k-1)}{2n_k}\right]\left[N_k \sum_j P_{kj}^2 - 2 + 0(n_k|N_k)\right]$$

$$\left[\sum_j P_{kj}\left(\frac{Y_{kj}}{P_{kj}} - Y_k\right)^2\right] > 0 \quad \text{by } (v). \quad \ldots \quad (26)$$

This completes the proof.

It is clear from (26) that asymptotically Des Raj's strategy would fare better than the Rao-Hartley-Cochran strategy if $\sum_j P_{kj}^2 \geqslant (2+\epsilon_k)N_k^{-1}$ and $\underline{\lim}\ \epsilon_k > 0$. In pps sampling schemes this condition reduces to the following condition : $[\text{c.v.}(X)]^2 = N_k^{-1} \sum_j (X_{kj} - \bar{X}_k)^2/(\bar{X}_k)^2 \geqslant 1+\epsilon_k$, where X_{kj} is the size of the j-th population unit and $P_{kj} = X_{kj}\Big/\left(\sum_j X_{kj}\right)$. This means that the assumption (iv) in Theorem 4 holds if the coefficient of variation of the size is bounded away from 1. Since many large populations commonly met in practice satisfy this condition, Des Raj's strategy would be found to be more precise than the Rao-Hartley-Cochran strategy in many cases of sampling from large populations.

Another interesting comparison concerning the relative performance of the two strategies can be made by considering the following super-population model :

$$Y_{kj} = m_k P_{kj} + e_{kj} \qquad \dots \quad (27)$$

where we assume that (i) e_{kj}'s are mutually independent and (ii) $\mathcal{E}(e_{kj}) = 0$ and $\mathcal{E}(e_{kj}^2) = a_k P_{kj}^{g_k}$ $(g_k > 0)$ in arrays where P_{kj} is fixed.

Under the above model when $\max_j P_{kj} = 0(N_k^{-1})$, we have

$$\mathcal{E}(V(\bar{T}_{k(rhc)})] = \frac{a_k}{n_k} \sum_j P_{kj}^{g_k-1} - \frac{a_k}{n_k} \left[1+0\left(\frac{n_k}{N_k}\right)\right]\left[\sum_j P_{kj}^{g_k} - \frac{(n_k-1)}{N_k} \sum_j P_{kj}^{g_k-1}\right] \qquad \dots \quad (28)$$

and

$$\mathcal{E}[V(\bar{T}_{k(d)})] \leqslant \frac{a_k}{n_k} \sum_j P_{kj}^{g_k-1} - \frac{a_k}{n_k}\left[1+0\left(\frac{n_k}{N_k}\right)\right]\left[\frac{(n_k+1)}{2} \sum_j P_{kj}^{g_k} + \frac{(n_k-1)}{2}\left(\sum_j P_{kj}^2\right)\left(\sum_j P_{kj}^{g_k-1}\right)\right] \qquad \dots \quad (29)$$

where $\mathcal{E}$ stands for the expectation under the super-population model (27).

The following theorem is easily derived from (28) and (29).

Theorem 5 : *Suppose that*

(i) $N_k \to \infty$ *as* $k \to \infty$,

(ii) $\underline{\lim}\ n_k > 1$ *and* $\overline{\lim}\ (n_k | N_k) = 0$,

(iii) $\max_j P_{kj} = 0(N_k^{-1})$,

(iv) $\underline{\lim}\ N_k^{-1}\left[\sum_j (N_k P_{kj} - 1)^2\right] > 0$ *and*

(v) *the super-population model* (27) *holds with* $g_k \geqslant 1$.

Then

$$\underline{\lim}\ N_k^{g_k-1} \{\mathcal{E}[V(\bar{T}_{k(rhc)})] - \mathcal{E}[V(\bar{T}_{k(d)})]\} > 0. \qquad \dots \quad (30)$$

Proof : From (28) and (29), we get

$$[V(\bar{T}_{k(rhc)})] - [V(\bar{T}_{k(d)})]$$

$$= 0\left(n_k N_k^{-g_k}\right) + \frac{a_k(n_k-1)}{2n_k}\left[\left(\sum_j P_{kj}^2 - N_k^{-1}\right)\sum_j P_{kj}^{g_k-1} + \left(\sum_j P_{kj}^{g_k} - N_k^{-1} \sum_j P_{kj}^{g_k-1}\right)\right]. \qquad \dots \quad (31)$$

Since

$$\sum_j P_{kj}^{g_k} - N_k^{-1} \sum_j P_{kj}^{g_k-1} = \sum_j P_{kj}^{g_k-1} (P_{kj} - N_k^{-1}) \geqslant 0$$

for $g_k \geqslant 1$, it follows that (30) is valid if

$$\underline{\lim}\ N_k^{g_k-1} \left[\left(\sum_j P_{kj}^2 - N_k^{-1}\right)\sum_j P_{kj}^{g_k-1}\right] > 0. \qquad \dots \quad (32)$$

0

Now since $$N_k^{g_k-1} \sum_j P_{kj}^{g_k-1} \geqslant N_k,$$

$$\underline{\lim}\ N_k^{g_k-1}\left[\left(\sum_j P_{kj}^2 - N_k^{-1}\right)\sum_j P_{kj}^{g_k-1}\right] \geqslant \underline{\lim}\ N_k^{-1}\left(\sum_j N_k^2 P_{kj}^2 - N_k\right)$$

$$= \underline{\lim}\ N_k^{-1}\left[\sum_j (N_k P_{kj}-1)^2\right] > 0 \text{ by virtue of } (v). \qquad \ldots \quad (33)$$

This completes the proof of the stated theorem.

It is important to note that the assumption (iv) in Theorem 6 ensures that the frequency distribution of N_kP_k in the population π_k does not tend to the one point distribution, $\varepsilon(x-1)$, as $k \to \infty$, where $\varepsilon(x-1) = 0$ for $x < 1$ and $= 1$ for $x \geqslant 1$. In pps sampling schemes this assumption is equivalent to the assumption that the coefficient of variation of the size is bounded away from zero. This assumption is clearly satisfied in nearly every situation where pps sampling is used. Hence for large populations Des Raj's strategy will fare better than the Rao-Hartley-Cochran strategy if (i) the sample size is small compared to the population size, (ii) the super-population model (27) holds with $g \geqslant 1$ and (iii) the essential range of the auxiliary variate, on which the selection probabilities are based, is bounded and contains at least two points which are not too close to each other.

It is, of course, no less important to note that the above theorem does not hold if the assumption (iv) is not satisfied. Actually sampling with unequal probabilities (without replacement) becomes asymptotically equivalent to simple random sampling (without replacement) when (iv) does not hold. It is easy to check that in this case the Rao-Hartley-Cochran strategy is more precise than Des Raj's strategy.

4. Comparison with the Horvitz-Thompson estimator

In this section we compare Des Raj's strategy with the well-known Horvitz-Thompson estimator (1952).

For a given scheme of sampling from π_k, the Horvitz-Thompson estimator of the population total Y_k is given by

$$t_{k(ht)} = \sum_i \frac{y_{ki}}{\alpha_{ki}} \qquad \ldots \quad (34)$$

where α_{ki} is the probability of inclusion of the i-th sample unit and the sum, $\sum_i$, is taken over the distinct population units included in the sample. It is assumed that every population unit has non-zero probability of being included in the sample.

The sampling variance of the above estimator is given by

$$V(t_{k(ht)}) = \sum_i \frac{Y_{kj}^2}{\alpha_{kj}} + \sum_{j \neq j'} \frac{Y_{kj}}{\alpha_{kj}} \frac{Y_{kj'}}{\alpha_{kj'}} \alpha_{kjj'} - Y_k^2 \qquad \ldots \quad (35)$$

where $\alpha_{kjj'}$ denotes the probability of inclusion of the j-th and j'-th population units in the sample.

By the effective size of a sample we mean the number of distinct units in the sample. For sampling schemes of effective size n_k with $\alpha_{kj} = n_k P_{kj}$ $(j = 1, ..., N_k)$, the expected variance of the Horvitz-Thompson estimator is given by

$$\underset{-}{\mathcal{E}}\,[V(\bar{t}_{k(ht)})] = \frac{a_k}{n_k} \sum_j P_{kj}^{g_k-1} - a_k \sum_j P_{kj}^{g_k} \qquad \text{... (36)}$$

where the expectation, $\mathcal{E}$, is taken under the super-population model (27).

Since it is not very meaningful (because of cost and other considerations) to compare Des Raj's strategy with the Horvitz-Thomspon estimator obtained under sampling schemes of arbitrary nature, we compare Des Raj's strategy with the Horvitz-Thompson estimator obtained under sampling schemes of effective size n_k with $\alpha_{kj} = n_k P_{kj}$ $(j = 1, ..., N_k)$. To further simplify this comparison we assume the super-population model (27).

Theorem 6 : *Suppose that*

(i) $N_k \to \infty$ *as* $k \to \infty$,

(ii) $\underline{\lim}\ n_k > 1$ *and* $\overline{\lim}\ (n_k | N_k) = 0$,

(iii) $\max_j P_{kj} = 0(N_k^{-1})$,

(iv) *the super-population model* (26) *holds with* $1 \leqslant g_k \leqslant 2$ *and* $\overline{\lim}\ g_k < 2$, *and*

(v) *as* $k \to \infty$, *the frequency distribution of* $N_k\ P_{kj}$ $(j = 1, ..., N_k)$ *converges weakly to a distribution function with essential range containing at least two points different from* 0. *Then the following inequality holds.*

$$\underline{\lim}\ N_k^{g_k-1} \{[V(\bar{t}_{k(ht)})] - [V(\bar{t}_{k(d)})]\} > 0. \qquad \text{... (37)}$$

Proof : To prove the theorem, we need the following lemma.

Lemma 4.1 : *For each* k $(k = 1, 2, ...)$, *let* $F_k(\cdot)$ *denote a distribution function defined on the real line. Let* $\{g_k : k = 1, 2, ...,\}$ *be a sequence of real numbers such that* $0 \leqslant g_k \leqslant 2$ *and* $\lim g_k = g$. *Suppose that* $F_k(\cdot)$ *converges weakly to a distrubution function* $F(\cdot)$ *as* $k \to \infty$. *Then*

$$\lim_{k\to\infty} \int_a^b x^{g_k} dF_k = \int_a^b x^g dF \qquad \text{... (38)}$$

where $b > a > 0$.

Proof : Fix an $\varepsilon > 0$. Clearly .

$$\left| \int_a^b x^{g_k} dF_k - \int_a^b x^g dF \right| \leqslant \left| \int_a^b x^g dF_k - \int_a^b x^g dF \right| + \int_a^b \left| x^{g_k} - x^g \right| dF_k. \qquad \text{... (39)}$$

Now by virtue of the Helly-Bray lemma (Loève, 1963) it follows that there exists a k_1 such that for $k \geqslant k_1$ the first term on the right side of (39) is less than $\epsilon/2$. Also since $\lim g_k = g$, it is clear that there exists a k_2 such that for $k \geqslant k_2$ $\sup_{x \in [a,b]} |x^{g_k} - x^g| < \epsilon/2$. Hence for $k \geqslant \max(k_1, k_2)$

$$\left| \int_a^b x^{g_k} dF_k - \int_a^b x^g dF \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \qquad \ldots \quad (40)$$

This completes the proof of the lemma.

Proof of Theorem 6 : From (29) and (36) we get

$$\mathcal{E}[V(\bar{t}_{k(hl)})] - \mathcal{E}[V(\bar{t}_{k(d)})] = 0\left(n_k N_k^{-g_k}\right) + \frac{a_k(n_k-1)}{n_k}\left[\left(\sum_j P_{kj}^2\right)\left(\sum_j P_{kj}^{g_k-1}\right) - \sum_j P_{kj}^{g_k}\right] \ldots \quad (41)$$

It is thus evident that (37) holds if

$$\underline{\lim}\ N_k^{g_k-1}\left[\left(\sum_j P_{kj}^2\right)\left(\sum_j P_{kj}^{g_k-1}\right) - \sum_j P_{kj}^{g_k}\right] > 0. \qquad \ldots \quad (42)$$

Let the left side of (42) be l. We now show that $l > 0$. Without loss of generality we can and do replace "$\underline{\lim}$" by "lim" on the left side of (42) and assume that $\lim g_k = g < 2$. [This is permissible because we can always select a subsequence $\{k_n : n = 1, 2, \ldots\}$ from the sequence of natural numbers $\{k : k = 1, 2, \ldots\}$ so that $\lim_n g_{k_n} = g$ and

$$\lim_n N_{k_n}^{g_{k_n}-1}\left[\left(\sum_j P_{k_n j}^2\right)\left(\sum_j P_{k_n j}^{g_{k_n}-1}\right) - \sum_j P_{k_n j}^{g_{k_n}}\right] = l\Big].$$

It now follows from (42) that

$$\lim\ \left[\left(\int_0^b x^2 dF_k\right)\left(\int_0^b x^{g_k-1} dF_k\right) - \int_0^b x^{g_k} dF_k\right] = l \qquad \ldots \quad (43)$$

where F_k denotes the frequency distribution of $N_k P_{kj}$ $(j = 1, 2, \ldots, N_k)$ and b is chosen so that $0 \leqslant N_k P_{kj} \leqslant b$ for all j and k; such a b exists by virtue of (iii).

Let $0 < a < 1$. Then since

$$\int_0^b x^2 dF_k \geqslant \left(\int_0^b x dF_k\right)^2 = 1,$$

$$\left(\int_0^b x^2 dF_k\right)\left(\int_0^b x^{g_k-1} dF_k\right) - \int_0^b x^{g_k} dF_k$$

$$\geqslant \left(\int_a^b x^2 dF_k\right)\left(\int_a^b x^{g_k-1} dF_k\right) - \int_a^b x^{g_k} dF_k + \int_0^a x^{g_k-1}(1-x) dF_k$$

$$\geqslant \left(\int_a^b x^2 dF_k\right)\left(\int_a^b x^{g_k-1} dF_k\right) - \int_a^b x^{g_k} dF_k. \qquad \ldots \quad (44)$$

Since F_k converges weakly to F as $k \to \infty$, by Lemma 4.1 it follows that

$$\lim \left[\left(\int_0^b x^2 dF_k \right) \left(\int_0^b x^{g_k - 1} dF_k \right) - \int_0^b x^{g_k} dF_k \right]$$

$$\geqslant \left(\int_a^b x^2 dF \right) \left(\int_a^b x^{g-1} dF \right) - \int_a^b x^g dF. \qquad \ldots \quad (45)$$

Further since (45) holds for all positive $a < 1$, we have

$$\lim \left[\left(\int_0^b x^2 dF_k \right) \left(\int_0^b x^{g_k - 1} dF_k \right) - \int_0^b x^{g_k} dF_k \right]$$

$$\geqslant \left(\int_{(0,b]} x^2 dF \right) \left(\int_{(0,b]} x^{g-1} dF \right) - \left(\int_{(0,b]} x^g dF \right)$$

$$\geqslant - \int_{(0,b]} x^{g-2} \left(x - \int_{(0,b]} x d\mu \right) d\mu \qquad \ldots \quad (46)$$

where $\mu(A) = \int_A x dF$ refers to a probability distribution on the Borel sets of the real line. By (v) it follows that $\mu(\cdot)$ is a proper probability distribution. Hence the right side of (46) is greater than zero for $1 \leqslant g < 2$. This completes the proof.

The above theorem clearly shows that under certain regularity conditions Des Raj's strategy is asymptotically more efficient than the Horvitz-Thompson estimator based on sampling schemes of fixed effective size. As regards the regularity conditions involved in this theorem, we suppose that they are quite realistic. The assumption (i) through (iv) of this theorem have already appeared in the preceding sections and we have seen that these are likely to be met in many situations commonly met in practice. The assumption (v) is new. In pps sampling schemes (v) holds if as $k \to \infty$, the frequency distribution of the size converges weakly to a distribution function F (say), where $\int x dF < \infty$ and the essential range of F contains at least two points that are different from 0. Consequently for large populations Des Raj's strategy would be more precise than the Horvitz-Thompson estimator based on pps sampling schemes of fixed effective size if (i) the sample size is relatively small compared to the population size, (ii) the super-population model (27) holds with $1 \leqslant g < 2$ and (iii) the essential range of the frequency distribution of the auxiliary variate, on which the selection probabilities are based, is bounded and contains at least two points of appreciable magnitude that are not too close to each other. It is worth noting that the Horvitz-Thompson estimator may be more precise than Des Raj's strategy when the regularity conditions of Theorem 6 are not met. For example it follows from theorems of Godambe-Joshi (1965) and Godambe (1955) that the Horvitz-Thompson estimator is more precise than Des Raj's strategy when (i) the Horvitz-Thompson is based on pps sampling schemes of fixed effective size and (ii) the super-population model (27) holds with $g = 2$; the latter condition clearly violates the assumption (iv) in Theorem 6. Also in simple random sampling (without replacement) the Horvitz-Thompson estimator is more precise than Des Raj's estimator; in this case the assumption (v) in Theorem 6 is violated.

Some of the well-known strategies which involve the Horvitz-Thompson estimator are (i) Hajek's (1964) strategy of rejective sampling, (ii) Hartley and Rao's (1965) strategy of randomised pps systematic sampling and (iii) Stevens's (1958) strategy of pps sampling. Further these strategies employ pps sampling schemes of fixed effective size. Consequently it follows from Theorem 6 that under the regularity conditions of the theorem Des Raj's strategy is more precise than the strategies introduced by Hájek (1964), Hartley and Rao (1965), and Stevens (1958)

MURTHY'S STRATEGY OR THE SDR STRATEGY

It is most appropriate to end this paper on a slightly different note concerning Murthy's strategy or the symmetrized Des Raj strategy (SDR strategy). The SDR strategy involves sampling with unequal probabilities (without replacement) and the following estimator of the population total (Murthy, 1957; Pathak, 1961)

$$\bar{t}_{k(m)} = \sum_{i} c_{ki} y_{ki} \qquad \dots \quad (47)$$

where the coefficient c_{ki} is a certain function of $p_{k1}, \dots, p_{kn_k}$.

Owing to the unwieldy nature of the c_{ki}, it is very difficult to calculate $\bar{t}_{k(m)}$ unless $n_k = 2$. For the special case $n_k = 2$, we have

$$\bar{t}_{k(m)} = \frac{1}{(2-p_{k1}-p_{k2})} \left[\left(1-p_{k2}\right) \frac{y_{k1}}{p_{k1}} + (1-p_{k1}) \frac{y_{k2}}{p_{k2}} \right]. \qquad \dots \quad (48)$$

It is known that for any convex loss function, Murthy's strategy leads to smaller risk function than Des Raj's strategy (Pathak, 1961). It would, therefore, be of particular interest to know if the SDR strategy is significantly better than Des Raj's strategy. In a subsequent paper entitled, "The efficiency of the symmetrized Des Raj strategy", we show that Murthy's and Des Raj's strategies are asymptotically equally efficient. Hence, unless $n_k = 2$, we are just as well off by relying on a slightly less efficient Des Raj's strategy. For $n_k = 2$, the Murthy's strategy should, however, be preferred to Des Raj's strategy because in this case the former strategy is not only more efficient but also equally simple to adopt.

6. ACKNOWLEDGEMENT

I am much indebted to Professor C. R. Rao for his encouragement and interest in the present work; also to my colleagues, J. Sethuraman and S. W. Dharmadhikari, and to the referees for their many helpful suggestions.

Appendix

Lemma 7.1 : *Let* $\sum_j x_j^r = S_r \quad (r = 1, 2, 3).$

Let $x_j \geqslant 0$. *Then*

$$\Sigma' x_{i_1} x_{i_2} \dots x_{i_m} \geqslant s_1^n - \binom{n}{2} S_2 S_1^{n-2} \qquad \dots (49)$$

and

$$\Sigma' x_{i_1}^2 x_{i_2} \dots x_{i_n} \geqslant S_2 S_1^{n-1} - \binom{n-1}{1} S_3 S_1^{n-2} - \binom{n-1}{2} S_2^2 S_1^{n-3} \qquad \dots (50)$$

where the summation, Σ', *runs over all possible permutations of* $x_1, x_2, \dots, x_N$ *taken* n *at a time.*

Proof of (49) : The proof is based on induction. Clearly (49) is true for $n = 2$. Let (49) be true for a given n. Then by virtue of our assumption it follows that

$$\Sigma' x_{i_1} x_{i_2} \dots x_{i_{\overline{n+1}}} \geqslant \sum_{i_1=1}^{N} x_{i_1} \left[\left(S_1 - x_{i_1}\right)^n - \binom{n}{2} \left(S_2 - x_{i_1}^2\right) \left(S_1 - x_{i_1}\right)^{n-2} \right]$$

$$\geqslant \sum_{i_1=1}^{N} x_{i_1} \left[\left(S_1 - x_{i_1}\right)^n - \binom{n}{2} S_2 S_1^{n-2} \right] \qquad \dots (51)$$

since $\quad \left(S_2 - x_{i_1}^2\right)\left(S_1 - x_{i_1}\right)^{n-2} \leqslant S_2 S_1^{n-2}.$

$$= \sum_{i_1=1}^{N} x_{i_1} \left[S_1^n - \binom{n}{1} x_{i_1} S_1^{n-1} + \binom{n}{2} x_{i_1}^2 \theta^{n-2} - \binom{n}{2} S_2 S_1^{n-2} \right]$$

where $S_1 - x_{i_1} \leqslant \theta \leqslant S_1$.

$$\geqslant \sum_{i_1=1}^{N} x_{i_1} \left[S_1^n - \binom{n}{1} x_{i_1} S_1^{n-1} - \binom{n}{2} S_2 S_1^{n-2} \right]$$

$$= S_1^n \sum_{i_1=1}^{N} x_{i_1} - \binom{n}{1} S_1^{n-1} \sum_{i_1=1}^{N} x_{i_1}^2 - \binom{n}{2} S_2 S_1^{n-2} \sum_{i_1=1}^{N} x_{i_1}$$

$$= S_1^{n+1} - \binom{n+1}{2} S_2 S_1^{n-1}.$$

Hence the lemma holds for $(n+1)$. This completes the proof.

Proof of (50) : From (49) it follows that

$$\Sigma' x_{i_1}^2 x_{i_2} \dots x_{i_n} \geqslant \sum_{i_1=1}^{N} x_{i_1}^2 \left[\left(S_1 - x_{i_1}\right)^{n-1} - \binom{n-1}{2} S_2 S_1^{n-3} \right]$$

$$\geqslant \sum_{i_1=1} x_{i_1}^2 \left[S_1^{n-1} - \binom{n-1}{1} x_{i_1} S_1^{n-2} - \binom{n-1}{2} S_2 S_1^{n-3} \right. \qquad \dots (52)$$

by an argument similar to that of (51)

$$= S_1^{n-1} \sum_{i_1=1}^{N} x_{i_1}^2 - \binom{n-1}{1} S_1^{n-2} \sum_{i_1=1}^{N} x_{i_1}^3 - \binom{n-1}{2} S_2 S_1^{n-3} \sum_{i_1=1}^{N} x_{i_1}^2$$

$$= S_2 S_1^{n-2} - \binom{n-1}{1} S_3 S_1^{n-2} - \binom{n-1}{2} S_2^2 S_1^{n-3}.$$

This completes the proof.

References

DES, RAJ (1956): Some estimators in sampling with varying probabilities without replacement. *J. Amer. Stat. Assoc.*, **51**, 269-84.

GODAMBE, V. P. (1955): A unified theory of sampling from finite populations. *J. Roy. Stat. Soc.*, B, **7**, 269-77.

GODAMBE, V. P. and JOSHI, V. M. (1965): Admissibility and Bayes estimation in sampling finite populations—I. *Ann. Math. Stat.*, **36**, 1707-22.

HAJEK, J. (1964): Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Ann. Math. Stat.*, **33**, 1491-1523.

HARTLEY, H. O. and RAO, J. N. K. (1962): Sampling with unequal probabilities and without replacement. *Ann. Math. Stat.*, **33**, 350-74.

HORVITZ, D. G. and THOMPSON, D. J. (1952): A generalization of sampling without replacement from a finite universe. *J. Amer. Stat. Assoc.*, **47**, 663-85.

LOEVE, MICHEL (1963): *Probability Theory*, 3rd ed. Van Nostrand, New York.

MURTHY, M. N. (1957): Ordered and unordered estimators in sampling without replacement. *Sankhyā*, **18**, 379-90.

PATHAK, P. K. (1961): Use of 'order-statistic' in sampling without replacement. *Sankhyā*, Series A, **23**, 409-14.

——— (1964): On inverse sampling with unequal probabilities. *Biometrika*, **51**, 185-93.

RAO, J. N. K., HARTLEY, H. O. and COCHRAN, W. G. (1962): On a simple procedure of unequal probability sampling without replacement. *J. Roy. Stat. Soc.*, B, **24**, 482-91.

ROYCHOWDHURY, D. K. (1956): Unpublished thesis submitted for the Associateship of the Indian Statistical Institute.

SAMPFORD, M. R. (1962): Methods of cluster sampling with and without replacement for clusters of unequal sizes. *Biometrika*, **49**, 27-40.

SETHI, V. K. (1962): Some consequences of an interpretation of varying probability sampling. *Sankhyā*, Series B, **24**, 215-22.

STEVENS, W. L. (1958): Sampling without replacement with probability proportional to size. *J. Roy. Stat. Soc.*, B, **20**, 393-97.

Paper received : February, 1967.