ON SAMPLING SCHEMES PROVIDING UNBIASED RATIO ESTIMATORS

By P. K. Pathak[1]

Indian Statistical Institute and Michigan State University

1. Summary and introduction. The problem of finding unbiased ratio estimators of the population total of some character with the help of an auxiliary character has drawn much attention in recent years. Some references to this are given at the end. Under commonly adopted sampling schemes Hartley and Ross (1954), Goodman and Hartley (1958), Mickey (1959) and others derived certain unbiased ratio type estimators of the population total, while Lahiri (1951), Midzuno (1952), Des Raj (1954), Nanjamma, Murthy, and Sethi (1959) and others gave modifications of certain sampling schemes under which the ratio estimators of current type were unbiased. The first group of authors was primarily concerned with getting new types of unbiased ratio estimators under common sampling schemes. However, the second group was concerned with introducing little modifications in certain sampling schemes so that the usual ratio estimators of these sampling schemes become unbiased under the modified sampling schemes. Certain extensions in the latter direction are given in this paper. No comparison concerning the relative merits and demerits of the two methods of getting unbiased ratio type estimators is attempted here.

Nanjamma, Murthy, and Sethi (1959) have given a general procedure of unbiased estimation of certain type of parameters such as population total and variance etc., and have shown how a given sampling scheme can be modified to make ratio estimators of such parameters unbiased. The modification can be applied to most of the sampling schemes commonly met in practice. In this paper it is shown that if in these modified sampling schemes a sufficient statistic is available [as is usually the case in with replacement sampling schemes, reference Basu (1958) and Pathak (1962(a), (b))] and if the ratio estimator does not depend on the sufficient statistic, it can be uniformly improved by Rao-Blackwell theorem. This result has been used to derive ratio estimators better than the ratio estimators given by Nanjamma, Murthy and Sethi. It is shown that in these modified sampling schemes the improved ratio estimator is the ratio of improved estimators of the numerator to the denominator derived under the original sampling scheme (without modification). Application of this result is given to some commonly used sampling schemes.

2. Statement of the problem. We consider the general procedure of estimation given by Nanjamma, Murthy and Sethi (1959). Let $\mathfrak{X}$ denote a population of finite number of units (e.g., a population of N units $u_1, u_2, \cdots, u_N$) and A

Received 22 October 1962; revised 14 April 1963.

[1] Now on leave from the Indian Statistical Institute.

the class of sets α whose elements are in $\mathfrak{X}$. Consider the problem of estimating the population parameter

$$F = \sum_{\alpha \varepsilon A} f(\alpha) \tag{1}$$

under a given sampling scheme, where $f(\alpha)$ is some real-valued characteristic of the set α. The sampling scheme under consideration assigns probability $P(\omega)$ to a sample $\omega(\sum_{\omega \varepsilon \Omega} P(\omega) = 1, \omega \varepsilon \Omega)$ sample units of ω being the elements of A.

An unbiased estimator of the population parameter F is given by

$$\hat{F} = \sum_{\alpha \varepsilon \omega} f(\alpha)\phi(\omega, \alpha)/P(\omega) \tag{2}$$

provided each set α is contained in at least one ω and $\sum^* \phi(\omega, \alpha) = 1$ where the summation $\sum^*$ is taken over all samples ω which contain α. This is so because under these conditions

$$E[\hat{F}] = \sum_{\omega \varepsilon \Omega} \sum_{\alpha \varepsilon \omega} f(\alpha)\phi(\omega, \alpha) = \sum_{\alpha \varepsilon A} f(\alpha)\{\sum^* \phi(\omega, \alpha)\} = F.$$

It is to be noted that ultimate sampling units are the set α's.

If T is a sufficient statistic and if $\hat{F}$ does not depend on T, then an estimator better than $\hat{F}$ is given by

$$\hat{F}_T = E[\hat{F} \mid T] = \sum_{\omega \supset T} \hat{F}P(\omega)/\sum_{\omega \supset T} P(\omega) \tag{3}$$

where the summation is taken over samples giving rise to the observed T. For any convex loss function $\hat{F}_T$ has smaller risk than $\hat{F}$.

The modification of the above sampling scheme which gives an unbiased estimator of the ratio

$$R = \sum_{\alpha \varepsilon A} f(\alpha)/\sum_{\alpha \varepsilon A} g(\alpha) = F/G \tag{4}$$

[where $g(\alpha)$ (> 0) is the value of the auxiliary variate related to $f(\alpha)$ and is assumed to be known for all $\alpha \varepsilon A$] is as follows:

(i) Select a set α with probabilities proportional to $g(\alpha)$.

(ii) Select the remaining sets by the original sampling scheme (without modification).

In this modified sampling scheme, the probability of selecting a sample ω is given by

$$P'(\omega) = \sum_{\alpha \varepsilon \omega} g(\alpha)P(\omega \mid \alpha)/\sum_{\alpha \varepsilon A} g(\alpha) \tag{5}$$

where $P(\omega \mid \alpha)$ is the conditional probability of selecting the sample ω given that the set α was selected in the first draw.

Under this modification Nanjamma, Murthy and Sethi (1959) showed that

$$\hat{R} = \sum_{\alpha \varepsilon \omega} f(\alpha)P(\omega \mid \alpha)/\sum_{\alpha \varepsilon \omega} g(\alpha)P(\omega \mid \alpha) \tag{6}$$

is an unbiased estimator of the ratio $R = F/G$.

Now if T is a sufficient statistic, we have the following theorem.

THEOREM 1. *For any convex loss function, an estimator uniformly better than* $\hat{R}$ *is given by*

$$\hat{R}_T = E[\hat{R} \mid T] = \hat{F}_T/\hat{G}_T \tag{7}$$

where $\hat{F}_T$ *and* $\hat{G}_T$ *are obtained from* (3) *on putting* $\phi(\omega \mid \alpha) = P(\omega \mid \alpha)$ *in* (2).

PROOF. Obviously

$$\hat{R}_T = E[\hat{R} \mid T] = \frac{\sum_{\omega \supset T} \hat{R}P'(\omega)}{\sum_{\omega \supset T} P'(\omega)} = \frac{\sum_{\omega \supset T} \{\sum_{\alpha \varepsilon \omega} f(\alpha)P(\omega \mid \alpha)\}}{\sum_{\omega \supset T} \{\sum_{\alpha \varepsilon \omega} g(\alpha)P(\omega \mid \alpha)\}}$$

$$= \frac{\sum_{\omega \supset T} \{\sum_{\alpha \varepsilon \omega} f(\alpha)P(\omega \mid \alpha)/P(\omega)\}P(\omega)/ \sum_{\omega \supset T} P(\omega)}{\sum_{\omega \supset T} \{\sum_{\alpha \varepsilon \omega} g(\alpha)P(\omega \mid \alpha)/P(\omega)\}P(\omega)/ \sum_{\omega \supset T} P(\omega)} = \frac{\hat{F}_T}{\hat{G}_T} \tag{8}$$

where
$$\hat{F}_T = E\left[\frac{\sum_{\alpha \varepsilon \omega} f(\alpha)P(\omega \mid \alpha)}{P(\omega)} \,\middle|\, T\right]$$

and the conditional expectation is taken under the original sampling scheme and $\hat{G}_T$ has a similar meaning. Hence the theorem is proved.

The theorem shows that the improved ratio estimator is the ratio of the improved estimators of the numerator to the denominator derived under the original sampling scheme (without modification). We now apply this theorem to a simple numerical example for clarity and to some of the sampling schemes considered by Nanjamma, Murthy and Sethi (1959) and derive more efficient unbiased ratio estimators.

3. Application to some sampling schemes. Let Y be a characteristic associated with the population under study and W be an auxiliary characteristic. It is assumed that the value of the W characteristic of every population unit is known in advance and is greater than zero. We give below unbiased ratio estimators of the ratio of the population total of Y and W characteristics. Unbiased ratio estimators of population total of Y-characteristic can be obtained by multiplying these estimators by the population total of W-characteristic.

An example. For illustration, let us consider the simplest case of simple random sampling (with replacement). In this sampling scheme a well-known estimator of the population mean is the average of n sample units and another estimator which has smaller variance than this estimator is the average of distinct sample units (Basu, 1958). This follows from the fact that the set of distinct units is a sufficient statistic and that the conditional expectation of the average of n sample units given the set of distinct units is the average of distinct units.

Now consider the modification of this sampling scheme which makes the ratio of the sample average of the characteristic under study to the sample average of the auxiliary characteristic an unbiased estimator of the corresponding ratio of the population totals of these characteristics. The modification consists in selecting the first sample unit with probability proportional to the size of the auxiliary characteristic (with replacement) and then the remaining $(n-1)$ sampling units by equal probability. Since under the original sampling scheme the conditional expectation of the average of n sample units given the set of distinct units is the average of the distinct sample units, it follows as a direct consequence of the above theorem that an estimator of the population ratio which has smaller variance than the ratio of the averages of n sample units of two characteristic is given by the ratio of the averages of the distinct sample units. The number of distinct units in a sample is always less than or equal to n.

To render the exposition clear, consider the following artificial population of five units.

j	Y_j	W_j	
1	63	37	
2	121	87	$S_y^2 = (N-1)^{-1} \sum (Y_j - \bar{Y})^2$
3	201	163	$= 139522$
4	254	208	$S_w^2 = (N-1)^{-1} \sum (W_j - \bar{W})^2$
5	367	329	$= 12804.2$
Total	$Y = 1006$	$W = 824$	

This population is a random sample of size 5 selected from 196 United States cities (Mickey, 1959); Y_j denotes the number of inhabitants in these cities in 1930 and W_j the number of inhabitants in 1920. It is assumed that the number of inhabitants in different cities in 1920 is known in advance. If a random sample of size $n = 3$ is drawn by simple random sampling (with replacement), then the variance of the average number of inhabitants in 1930 in three selected cities (including duplicates if some city is included more than once) is given by $(4/15) S_y^2 = 3720.58667$, and the variance of the average number of inhabitants in 1930 in the distinct cities selected in the sample is given by $[(1^2 + 2^2 + 3^2 + 4^2)/5^3] \cdot S_y^2 = 3348.528$. The latter estimator is a uniformly better estimator than the former. The usual ratio estimator, ratio of the sample averages of Y and W characteristics, of the ratio of the total number of inhabitants in 5 cities in 1930 to the total number of inhabitants in 1920 is a biased estimator. In the modified sampling scheme one unit is selected first (with replacement) with probabilities proportional to the number of inhabitants in 1920 and then the remaining two sample units are drawn by equal probability (with replacement). In this sampling scheme the ratio of the average number of inhabitants in 1930 to the number of inhabitants in 1920 in three sampled cities is unbiased for the ratio of the number of inhabitants in five cities in 1930 and 1920. The variance of this estimator is 0.005223. However, if one takes the ratio of the average number of inhabitants in 1930 and 1920 in the distinct cities included in the sample, then by Theorem 1 this is a better estimator of the ratio of the average number of inhabitants in 1930 and 1920 in five cities. The variance of this estimator is equal to .004833. The relative gain in precision by using the better estimator is roughly about 7%.

Sampling with unequal probabilities. Consider a population of N units and an unequal probability selection method; let P_j be the probability of selection of the jth population unit ($\sum P_j = 1$). Unless otherwise stated, summation over j runs from 1 to N and over i from 1 to ν. In sampling with unequal probabilities a sample of size n is drawn with replacement from the population according to the above probabilities. The modification of this sampling scheme which provides an unbiased ratio estimator is as follows:

(1) Draw one unit from the population with ppw and replace it; ppw is an abbreviation for "probabilities proportional to W."

(2) Draw the remaining $(n - 1)$ sample units from the whole population in the usual manner, i.e. with unequal probabilities with replacement, P_j being the probability of selection associated with the jth population unit.

Consider a sample drawn by the above modified sampling scheme and arrange the sample units in ascending order of their unit-indices and record the observed sample as

$$S = [(x_{(1)}, \lambda_{(1)}), \cdots, (x_{(\nu)}, \lambda_{(\nu)})] \tag{9}$$

where $x_{(i)} = [y_{(i)}, w_{(i)}, p_{(i)}, u_{(i)}]$ is the ith order statistic $\{y_{(i)}, w_{(i)}, p_{(i)}, u_{(i)}$ are its Y-variate value, W-variate value, probability of selection and unit index respectively$\}$ and $\lambda_{(i)}$ is the number of times $x_{(i)}$ is included in the sample. It is to be noted here that if the ith sample unit is the jth population unit, its unit index is j.

The probability of getting such a particular sample S is given by

$$P(S) = \frac{n! \prod_{i=1}^{\nu} p_{(i)}^{\lambda_{(i)}}}{\prod_{i=1}^{\nu} \lambda_{(i)}!} \cdot \frac{1}{W} \cdot \left(n^{-1} \sum_{i=1}^{\nu} \lambda_{(i)} w_{(i)}/p_{(i)}\right) \tag{10}$$

where $W = \sum W_j$ is the population total of W-characteristic.

In this sampling scheme an unbiased estimator of the ratio

$$R = \sum_j Y_j / \sum_j W_j = Y/W \tag{11}$$

is given by

$$\hat{R} = (n^{-1} \sum \lambda_{(i)} y_{(i)}/p_{(i)})/(n^{-1} \sum \lambda_{(i)} w_{(i)}/p_{(i)}). \tag{12}$$

To get an estimator better than $\hat{R}$, we record the "order-statistic"

$$T = [x_{(1)}, x_{(2)}, \cdots, x_{(\nu)}]. \tag{13}$$

T is a sufficient statistic. The author (1962 (b)) has shown that in the original sampling scheme

$$E[n^{-1} \sum \lambda_{(i)}(y_{(i)}/p_{(i)}) \mid T] = \sum c_{(i)} y_{(i)} \tag{14}$$

where

$$c_{(i)} = \frac{[(p_{(1)} + \cdots + p_{(\nu)})^{n-1} - \sum_1^i (p_{(1)} + p_{(2)} + \cdots + p_{(\nu-1)})^{n-1} + \cdots (-1)^{\nu-1} p_{(i)}^{n-1}]}{[(p_{(1)} + \cdots + p_{(\nu)})^n - \sum_1 (p_{(1)} + p_{(2)} + \cdots + p_{(\nu-1)})^n + \cdots (-1)^{\nu-1} \sum_1 p_{(1)}^n]}$$

the summations $\sum_1$ and $\sum_1^i$ stand for all combinations of p's and all combinations of p's containing $p_{(i)}$ (chosen out of $p_{(1)}, p_{(2)}, \cdots, p_{(\nu)}$) respectively.

It, therefore, follows from Theorem 1 that under the modified sampling scheme

$$\hat{R}_\nu = \sum c_{(i)} y_{(i)} / \sum c_{(i)} w_{(i)} \tag{15}$$

is uniformly better than $\hat{R}$.

In simple random sampling (with replacement) when $P_j = N^{-1}$, $(j = 1, \cdots, N)$, $\hat{R} = (\sum \lambda_{(i)} y_{(i)} / \sum \lambda_{(i)} w_{(i)})$ and the improved ratio estimator $\hat{R}_\nu = (\sum y_{(i)} / \sum w_{(i)})$.

For estimating the variance of $\hat{R}$ or of $\hat{R}_\nu$, we require an unbiased estimator of R^2. Nanjamma, Murthy and Sethi (1959) gave the following estimator of R^2.

$$r^2 = \frac{\sum_{i=1}^{\nu} \lambda_{(i)} (\lambda_{(i)} - 1) \{ y_{(i)}^2 / p_{(i)}^2 \} + \sum_{i \neq i'=1}^{\nu} \lambda_{(i)} \lambda_{(i')} \{ y_{(i)} y_{(i')} / p_{(i)} p_{(i')} \}}{(n-1) W (\sum \lambda_{(i)} \{ w_{(i)} / p_{(i)} \})}. \tag{16}$$

It can now be proved in a manner similar to the above [Pathak 1962(b)] that an estimator uniformly better than r^2 is given by

$$r_\nu^2 = \frac{\sum_{i=1}^{\nu} c_{(i,i)} y_{(i)}^2 + \sum_{i \neq i'=1}^{\nu} c_{(i,i')} y_{(i)} y_{(i')}}{W (\sum c_{(i)} w_{(i)})} \tag{17}$$

where

$$c_{(i,i)} = \frac{[(p_{(1)} + \cdots + p_{(\nu)})^{n-2} - \sum_1^i (p_{(1)} + \cdots + p_{(\nu-1)})^{n-2} + \cdots (-1)^{\nu-1} p_{(i)}^{n-2}]}{[(p_{(1)} + \cdots + p_{(\nu)})^n - \sum_1 (p_{(1)} + \cdots + p_{(\nu-1)})^n + \cdots (-1)^{\nu-1} \sum_1 p_{(1)}^n]}$$

and

$$c_{(i,i')} = \frac{[(p_{(1)} + \cdots + p_{(\nu)})^{n-2} - \sum_1^{ii'} (p_{(1)} + \cdots + p_{(\nu-1)})^{n-2} + \cdots (-1)^{\nu-2} (p_{(i)} + p_{(i')})^{n-2}]}{[(p_{(1)} + \cdots + p_{(\nu)})^n - \sum_1 (p_{(1)} + \cdots + p_{(\nu-1)})^n + \cdots (-1)^{\nu-1} \sum_1 p_{(1)}^n]}$$

the summations $\sum_1$ and $\sum_1^i$ have been defined in (14) and the summation $\sum^{ii'}$ stands for all combinations of p's containing $p_{(i)}$ and $p_{(i')}$.

In simple random sampling (with replacement) r_ν^2 can be expressed in a simple form as given below.

$$(18) \qquad r_\nu^2 = \frac{1}{\overline{W}\left(\frac{1}{\nu}\sum w_{(i)}\right)} \cdot \left\{\frac{1}{\nu}\sum y_{(i)}^2 + \left(1 - \frac{c_\nu(n-1)}{c_\nu(n)}\right) s_d^2\right\}$$

where
$$s_d^2 = (\nu-1)^{-1} \cdot \sum (y_{(i)} - \nu^{-1}\sum y_{(i)})^2 \qquad \text{if } \nu > 1$$
$$= 0 \qquad \text{otherwise}$$

and
$$c_\nu(n) = \nu^n - \binom{\nu}{1}(\nu-1)^n + \cdots (-1)^{\nu-1}\binom{\nu}{\nu-1}.$$

An unbiased ratio estimator of $V(\hat{R}_\nu)$ is, therefore, given by

$$(19) \qquad v(R_\nu) = (R_\nu)^2 - r_\nu^2.$$

The estimators derived above are quite tedious to compute in practice unless the sample size is small. Better estimators that are easy to compute in practice, but less efficient than $\hat{R}_\nu$ and r_ν^2 can be derived on the above lines by using another sufficient statistic considered by the author (1962(b)) in a paper on sampling with unequal probabilities.

Applications to stratified sampling with unequal probabilities (with replacement) and multi-stage sampling schemes are given below. Reader may refer to the papers by the author (1963) and by Nanjamma, Murthy and Sethi (1959) for details.

Stratified sampling. Here we consider the modification of stratified sampling with unequal probabilities (with replacement).

Let K be the number of strata, N_h and n_h be the number of units in the population and the sample respectively for the hth stratum ($h = 1, \cdots, K$). Let P_{hj} be the probability of selection associated with the jth unit of the hth stratum ($\sum_{j=1}^{N_h} P_{hj} = 1$, $h = 1, \cdots, K$) and its W and Y characteristic values being W_{hj} and Y_{hj} respectively.

In stratified sampling with unequal probabilities, n_h units are drawn independently from the hth stratum with unequal probabilities ($h = 1, \cdots, K$). The modification of this scheme which provides an unbiased ratio estimator is as follows:

(1) Draw one unit, say the jth unit from the hth stratum, from the whole population with ppw and replace it.

(2) Draw the remaining $(n_h - 1)$ units from the hth stratum and $n_{h'}$ units from h'th stratum ($h' \neq h = 1, 2, \cdots, K$) in the usual way, i.e. by stratified sampling with unequal probabilities.

Under this sampling scheme an unbiased estimator of the ratio

$$(20) \qquad R = \sum_{h=1}^{K}\sum_{j=1}^{N_h} Y_{hj} \Big/ \sum_{h=1}^{K}\sum_{j=1}^{N_h} W_{hj}$$

is given by

$$\hat{R} = \left(\sum_{h=1}^{K} n_h^{-1} \sum_{i=1}^{\nu_h} \lambda_{(hi)} y_{(hi)}/p_{(hi)}\right) \Big/ \left(\sum_{h=1}^{K} n_h^{-1} \sum_{i=1}^{\nu_h} \lambda_{(hi)} y_{(hi)}/p_{(hi)}\right) \tag{21}$$

where

$$n_h^{-1} \sum_{i=1}^{\nu_h} \lambda_{(hi)} y_{(hi)}/p_{(hi)} \text{ and } n_h^{-1} \sum_{i=1}^{\nu_h} \lambda_{(hi)} w_{(hi)}/p_{(hi)}$$

are unbiased estimators of the population totals of Y and W characteristics of the hth stratum respectively under the usual stratified sampling with unequal probabilities and are essentially of the same form as $n^{-1} \sum \lambda_{(i)} y_{(i)}/p_{(i)}$ given in (12).

We state without going into details that an estimator better than $\hat{R}$ is obtained by replacing

$$n_h^{-1} \sum_{h=1}^{\nu_h} \lambda_{(hi)} y_{(hi)}/p_{(hi)}$$

and

$$n_h^{-1} \sum_{h=1}^{\nu_h} \lambda_{(hi)} w_{(hi)}/p_{(hi)}$$

by $\sum_{i=1}^{\nu_h} c_{(hi)} y_{(hi)}$ and $\sum_{i=1}^{\nu_h} c_{(hi)} w_{(hi)}$ respectively, where $c_{(hi)}$ has similar meaning as $c_{(i)}$ defined in (14).

Two-stage sampling. In conclusion, we consider the modification of two-stage sampling scheme where the first-stage units are drawn with unequal probabilities (with replacement) and the second-stage units are drawn by simple random sampling (without replacement). Similar procedure can be followed for other multi-stage sampling schemes.

Consider a population of N first-stage units. Suppose that the jth first-stage unit, X_j, consists of M_j second-stage units. Let Y_{jh} and W_{jh} be the values of Y and W characteristics of the hth second-stage unit of $X_j (h = 1, \cdots, M_j)$. Let P_j be the probability of selection of the jth first-stage unit ($\sum P_j = 1$) and call

$$Z_{jh} = M_j Y_{jh}/P_j \text{ and } V_{jh} = M_j W_{jh}/P_j \tag{22}$$

the z-value and the v-value respectively of the hth second-stage unit of X_j.

In the usual two-stage sampling (unequal probabilities for first-stage and simple random sampling for second-stage) a sample of n first-stage units is drawn from the above population according to above probabilities and if the jth first-stage population unit is included λ_j times in the sample, λ_j independent sub-samples of m_j units each are selected therefrom by simple random sampling.

The modification of this scheme is as follows:

(1) Draw one second-stage unit from the whole population of second-stage units with ppw, say X_{jh}, and then select $(m_j - 1)$ second stage units from the

remaining $(M_j - 1)$ second-stage units of the jth first-stage unit by simple random sampling. This can also be achieved in a different manner by first selecting a first-stage unit with probabilities proportional to the total W characteristic of first-stage units and then selecting a second-stage unit with ppw from the selected first-stage unit.

(2) Draw the remaining $(n - 1)$ first-stage sample units and their subsamples in the usual manner (without modification).

In this modified sampling scheme, and unbiased estimator of the ratio

$$R = \sum_{j=1}^{N} \sum_{h=1}^{M_j} Y_{jh} / \sum_{j=1}^{N} \sum_{h=1}^{M} W_{jh} \tag{23}$$

is given by

$$\hat{R} = n^{-1} \sum_{i=1}^{n} \bar{z}_i / n^{-1} \sum_{i=1}^{n} \bar{v}_i \tag{24}$$

where $\bar{z}_i$ and $\bar{v}_i$ are averages of the z-values and the v-values of the second-stage sample units selected from the ith first-stage sample unit.

In this sampling scheme a sufficient statistic is given by

$$T^* = [\{x_{(i)}, \lambda_{(i)}; x_{(i1)}, \cdots x_{(i\nu_{(i)})}\} \; i = 1, \cdots, \nu] \tag{25}$$

where $x_{(1)}, \cdots, x_{(\nu)}$ are ν distinct first-stage units observed in the sample, $\lambda_{(i)}$ is the number of times $x_{(i)}$ is repeated and $x_{(i1)}, \cdots, x_{(i\nu_{(i)})}$ are $\nu_{(i)}$ distinct second-stage sample units of $x_{(i)}$.

The author (1963) has proved that under the original sampling scheme

$$E[n^{-1} \sum_{i=1}^{n} \bar{z}_i \mid T^*] = n^{-1} \sum_{i=1}^{\nu} \lambda_{(i)} \bar{z}_{\nu_{(i)}} \tag{26}$$

where $\bar{z}_{\nu_{(i)}}$ is the average of the z-values of $\nu_{(i)}$ distinct second-stage units $x_{(i1)}, \cdots, x_{(i\nu_{(i)})} (i = 1, \cdots, \nu)$.

Therefore, it follows as a consequence of Theorem 1 that an estimator uniformly better than $\hat{R}$ is given by

$$\hat{R}_\nu = \sum_{i=1}^{\nu} \lambda_{(i)} \bar{z}_{\nu_{(i)}} / \sum_{i=1}^{\nu} \lambda_{(i)} \bar{v}_{\nu_{(i)}} \tag{27}$$

where $\bar{v}_{\nu_{(i)}}$ has a meaning similar to $\bar{z}_{\nu_{(i)}}$ as defined in (26).

Acknowledgment. The author wishes to express his sincere thanks to Professor D. Basu of the Indian Statistical Institute for his encouragement and guidance in the preparation of the paper, and to a referee for his helpful criticisms.

REFERENCES

[1] Basu, D. (1958). On sampling with and without replacement. *Sankhyā* **20** 287–294.

[2] Des Raj (1954). Ratio estimation in sampling with equal and unequal probabilities. *J. Indian Soc. Agric. Statist.* **6** 127–138.

[3] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

[4] GOODMAN, L. A. and HARTLEY, H. O. (1958). The precision of unbiased ratio-type estimators. *J. Amer. Statist. Assoc.* **53** 491–508.
[5] HARTLEY, H. O. and ROSS, A. (1954). Unbiased ratio estimators. *Nature* **174** 270–271.
[6] LAHIRI, D. B. (1951). A method of sample selection providing unbiased ratio estimators. *Bull. Inst. Internat. Statist.* **33** 2, 133–140.
[7] MIDZUNO, H. (1952). On the sampling system with probability proportionate to sum of sizes. *Ann. Inst. Statist. Math.* **3** 99–107.
[8] MICKEY, M. R. (1959). Some finite population unbiased ratio and regression estimators. *J. Amer. Statist. Assoc.* **54** 594–612.
[9] NANJAMMA, N. S., MURTHY, M. N. and SETHI, V. K. (1959). Some sampling systems providing unbiased ratio estimators. *Sankhyā* **21** 299–314.
[10] PATHAK, P. K. (1962) (a). On simple random sampling with replacement. *Sankhyā* Ser. A **24** 287–302.
[11] PATHAK, P. K. (1962) (b). On sampling with unequal probabilities. *Sankhyā* Ser. A **24** 315–326.
[12] PATHAK, P. K. (1963). Sufficiency in sampling theory. to appear in *Ann. Math. Statist.*
[13] SEN, A. R. (1952). Present status of probability sampling and its use in estimation of characteristics (abstract). *Econometrica* **20** 103.