SAMPLING WITH UNEQUAL PROBABILITIES AND WITHOUT REPLACEMENT[1]

By H. O. Hartley and J. N. K. Rao

Iowa State University

0. Summary. Given a population of N units, it is required to draw a random sample of n distinct units in such a way that the probability for the ith unit to be in the sample is proportional to its "size" x_i (sampling with p.p.s. without replacement). From a number of alternatives of achieving this, one well known procedure is here selected: The N units in the population are listed in a random order and their x_i are cumulated and a systematic selection of n elements from a "random start" is then made on the cumulation. The mathematical theory associated with this procedure, not available in the literature to date, is here provided: With the help of an asymptotic theory, compact expressions for the variance of the estimate of the population total are derived together with variance estimates. These formulas are applicable for moderate values of N. The reduction in variance, as compared to sampling with p.p.s. *with* replacement, is clearly demonstrated.

1. Introduction. Most survey designs incorporate as a basic sampling procedure the selection of n units at random, with equal probabilities and without replacement drawn from a population of N units. It is, however, sometimes advantageous to select units with unequal probabilities. For example, such a procedure may be found appropriate when a "measure of size" x_i is known for all the units in the population ($i = 1, 2, \cdots, N$) and it is suspected that these known sizes x_i are correlated with the characteristics y_i for which the population total Y is to be estimated. One method (though by no means the only method) of utilizing the x_i is to draw units with probabilities proportional to sizes x_i (p.p.s.), a technique frequently used in sample surveys, particularly for primary sampling units in multi-stage designs. Now the theory of sampling with unequal probabilities is equivalent to multinomial sampling provided units are drawn *with* replacement. On the other hand, it is well known from the theory of equal probability selection that sampling with replacement results in estimators which are less precise than those computed from samples selected without replacement; the proportional reduction in variance is given by n/N ("finite population correction"). It has therefore been felt for some time that a similar increase in precision should be reaped by switching to selection without replacement in unequal probability sampling. However, the theory of unequal probability sampling without replacement involves mathematical and computational difficulties and has therefore not been fully developed.

Received January 30, 1961; revised September 18, 1961.

[1] Research sponsored by the Office of Ordnance Research, U.S. Army, under Grant No. DA-ARO (D)-31-124-G93.

A general theory of unequal probability sampling without replacement was first given by Horvitz and Thompson [4]. Since then, several papers have been published on the topic, but we shall review here only the papers relevant to the particular problem considered here. The estimator of the total Y as proposed by Horvitz and Thompson is

$$\hat{Y} = \sum^{n} \frac{y_i}{\pi_i}, \tag{1}$$

with variance

$$V(\hat{Y}) = \sum_{1}^{N} \frac{y_i^2}{\pi_i} + 2 \sum_{i'>i=1}^{N} \frac{P_{ii'}}{\pi_i \pi_{i'}} y_i y_{i'} - Y^2, \tag{1.2}$$

where π_i is the probability for the ith unit to be in a sample of size n and $P_{ii'}$ is the probability for the units i and i' both to be in the sample. Now when the π_i are exactly proportional to the y_i, $V(\hat{Y})$ is zero which suggests that if we make π_i proportional to x_i, i.e., if we set

$$(n - 1)\pi_i = \sum_{i' \neq i}^{N} P_{ii'} = \gamma x_i, \tag{1.3}$$

a considerable reduction in the variance will result since "x_i" and "y_i" are correlated. Horvitz and Thompson, for the case $n = 2$, propose two methods to satisfy (1.3) approximately, but their methods have some limitations. Yates and Grundy [9] also deal with $n = 2$, and they suggest an iterative procedure to obtain revised "size measures" which satisfy (1.3) approximately, but their method becomes cumbersome when N becomes large. Moreover, this method is unmanageable when $n > 2$. Des Raj [3] employs (1.3) as a set of N equations for the $N(N - 1)/2$ probabilities $P_{ii'}$ and determines the latter by minimizing (1.2) subject to (1.3). This leads to a "linear programming problem" for the $N(N - 1)/2$ positive $P_{ii'}$. The "objective function" (the variance) involves the unknown population values y_i and these are replaced by the known sizes x_i with the assumption that

$$y_i = \alpha + \beta x_i \tag{1.4}$$

exactly. Even if this assumption is accepted, the method is clearly unmanageable for large N. Moreover, if it is assumed that the y_i of the population satisfy (1.4) exactly with unknown α and β, then, clearly the regression estimator has zero variance and, even if an error term is introduced into (1.4), the regression estimator would still be the "best estimator" so that it is of little interest to consider other estimators under such assumptions.

In this paper, we adopt a *particular* procedure of drawing a sample in such a way that (1.3) is satisfied exactly with the original sizes x_i. Although this procedure, which is described in Section 2, is well known to survey practitioners (e.g., Horvitz and Thompson [4], p. 678), no formulas for the probabilities $P_{ii'}$ in terms of the π_i are available in the literature, due to mathematical difficulties.

These difficulties are resolved in this paper and compact expressions for the variance and estimated variance of $\hat{Y}$ are obtained for moderate values of N. In Section 2, the sampling procedure is described and illustrated with the help of a numerical example. Section 3 deals with the cases $n = 2$, $N = 3$ and 4. We develop an asymptotic theory for the case $n = 2$ and N large in Section 4. In Sub-section 4.1, $P_{ii'}$ is explicitly derived in terms of the π_i. Sub-section 4.2 deals with the evaluation of the variance formulas, and a numerical example is given in Sub-section 4.3. Estimation of the variance is considered in Sub-section 4.4. Section 5 deals with the general case $n \geqq 2$, and compact expressions for the variance and estimated variance of $\hat{Y}$ are obtained, which are applicable when N is relatively large compared to n. In Section 6, a comparison with ratio estimation is made.

The extension of the theory developed here for simple sampling to more complex sample designs such as stratification, multi-stage sampling, etc., is comparatively straightforward and therefore is not presented here.

2. A particular sampling procedure.

2.1. *Description of the procedure.* It is easy to show that there is no sampling procedure whatsoever satisfying (1.3) unless $\gamma = n(n-1)/X$ and

$$np_i \leqq 1, \tag{2.1.1}$$

where $p_i = x_i/X$ and X is the population total of the x_i. Henceforth we shall only consider such sizes x_i and associated probabilities p_i which satisfy the necessary condition (2.1.1). The following sampling procedure is now considered:

a. Arrange the units in a random order and denote (without loss of generality) by $j = 1, 2, \cdots, N$ this random order and by $\Pi_j = \sum_{i=1}^{j} (np_i)$, $\Pi_0 = 0$, the progressive totals of the (np_i) in that order.

b. Select a "random start", i.e., select a "uniform variate" d with $0 \leqq d < 1$. Then the n selected units are those whose index, j, satisfies

$$\Pi_{j-1} \leqq d + k < \Pi_j \tag{2.1.2}$$

for some integer k between 0 and $n - 1$. Since $np_i \leqq 1$, every one of the n integers $k = 0, 1, \cdots, n - 1$ will select a different sampling unit j. It is easy to show that $\pi_i = np_i$ for this sampling procedure.

2.2. *Numerical example.* Consider a population of $N = 8$ units arranged in a random order and with sizes x_j shown in the second column of Table 1. A sample of $n = 3$ is to be drawn using our sampling procedure. Instead of computing the quantities np_i we scale all computations up by a factor of $X/n = 300/3 = 100$. Thus we compute the progressive sums of the x_i and these are shown in column 3 of Table 1 and correspond to the quantities $X\,\Pi_j/n$. Then we select a random integer between 1 and 100 and this corresponds to the quantity Xd/n. In our example this integer turned out to be 36 and the selection of the three units in accordance with (2.1.2) is shown in column 4. We must find the lines (j) where the column 100 Π_j passes through the levels $100d = 36$ (for $k = 0$), $100d + 100 = 136$ (for $k = 1$) and $100d + 200 = 236$ (for $k = 2$). The units

TABLE 1

Selection of $n = 3$ units from population of $N = 8$ units (p.p.s.)

Unit Number j	Size x_j	Progressive Sum $100\ \Pi_j$	Start = 36 Step = X/n = 100
1	15	15	
2	81	96	$k = 0$, $100d = 36$
3	26	122	
4	42	164	$k = 1$, $100d + 100 = 136$
5	20	184	
6	16	200	
7	45	245	$k = 2$, $100d + 200 = 236$
8	55	300	

$j = 2$, 4 and 7 are thereby selected. This procedure (either with or without the initial randomization) has been frequently used but, in the absence of a better theory, is usually treated approximately as a p.p.s. sample drawn *with* replacement.

2.3. *A cyclical analogue to the sampling procedure.* From the point of view of the mathematical treatment it is convenient to use an alternative but stochastically equivalent procedure to the steps a and b as follows:

a′. Arrange the units in a random order, denote by $j = 1, 2, \cdots, N$ this random order and form (as before) the progressive totals Π_j given by (2.1.2). Mark off on the perimeter of a circle with radius $n/(2\pi)$ arcs of lengths (np_j) in clockwise direction starting at the top (jth unit corresponds to jth arc of length np_j on the circle).

b′. Select a uniform arc s with $0 \leqq s < n$. Then the n selected units are those whose indices j satisfy

$$\Pi_{j-1} \leqq s + k < \Pi_j \tag{2.3.1}$$

for some integer k between $-(n - 1)$ and $(n - 1)$ for which also $0 \leqq s + k < n$. Precisely n units are selected by this process. It is clear that all results in Sub-section 2.1 still hold. For we know with certainty that of the above values of $s + k$ the algebraically smallest will lie between 0 and 1 and this may be identified with the variate d in step *a*.

3. The cases $n = 2$; $N = 3$ and 4.

3.1. *The case $n = 2$, $N = 3$.* Clearly we have

$$P_{ii'} = 1 - \pi_{i''} = \pi_i + \pi_{i'} - 1, \tag{3.1.1}$$

where i'' is the third unit in the population and $\pi_i + \pi_{i'} + \pi_{i''} = 2$. Substituting for $P_{ii'}$ from (3.1.1) in (1.2), and after some algebra, we get

$$V(\hat{Y}) = \sum_1^3 (1 - \pi_j)((y_j/\pi_j) - Y^*)^2, \tag{3.1.2}$$

where $Y^* = \sum_1^3 (1 - \pi_j)y_j/\pi_j$. In the special case of equal $\pi_j = \frac{2}{3}$, (3.1.2) reduces to the well known variance formula in equal probability sampling without replacement.

3.2. *The case* $n = 2$, $N = 4$. We evaluate P_{12} and assume[2] without loss of generality that

$$\pi_1 \geqq \pi_2 \quad \text{and} \quad \pi_4 \geqq \pi_3 . \tag{3.2.1}$$

We distinguish two subclasses of the randomization results:

Class 1. The units $i = 1$ and $i = 2$ are adjacent on the circle.
Class 2. The units $i = 1$ and $i = 2$ are separated by one unit.

It is easy to see that in class 1 the probability that $i = 1$ and $i = 2$ are the sampled units is given by

$$P'_{12} = \begin{cases} (\pi_1 + \pi_2 - 1) & \text{if } \pi_1 + \pi_2 \geqq 1 \\ 0 & \text{if } \pi_1 + \pi_2 < 1 \end{cases} \tag{3.2.2}$$

while in class 2, using (3.2.1), it is given by

$$P''_{12} = \begin{cases} (\pi_1 + \pi_2 + \pi_3 - 1) & \text{if } \pi_1 + \pi_3 \leqq 1 \\ \pi_2 & \text{if } \pi_1 + \pi_3 > 1. \end{cases} \tag{3.2.3}$$

Since the relative frequency of sequences in the above classes 1 and 2 are in the ratio $\frac{2}{3}$ to $\frac{1}{3}$, the overall probability P_{12} is given by

$$P_{12} = \tfrac{2}{3}P'_{12} + \tfrac{1}{3}P''_{12} . \tag{3.2.4}$$

The substitution of (3.2.4) in (1.2) yields a formula for $V(\hat{Y})$, although not in a compact form. Similar results have been obtained for the case $n = 2$, $N = 5$. As N increases, the exact evaluation of $P_{ii'}$ becomes more and more laborious and, in any case, the resulting formula will be too complicated to yield a practical formula for $V(\hat{Y})$. Therefore, we develop an asymptotic theory in the next section and obtain compact expressions for $V(\hat{Y})$.

3.3. *Numerical example.* To compare the efficiency of our sampling procedure for the case $n = 2$, $N = 4$ with both the procedures of Yates and Grundy [9] and Des Raj [3] we use the three populations examined by these authors. The three populations have the same set of p_i values and are given in Table 2.

Variances of $\hat{Y}$ for the three procedures and the three populations are given in Table 3. Variances for procedures 1 and 2 are taken from Des Raj [3]. For procedure 3 the variance is obtained from (3.2.4) and (1.2). It may be noted that substitution of (3.2.4) in (1.2) gives an exact formula for the variance. The variance of $\hat{Y}$ for p.p.s. sampling *with* replacement is also shown in Table 3 for comparison.

From Table 3 it is seen that procedures 1, 2 and 3 are more efficient than

[2] The numbering of the units $i = 1, 2, 3, 4$ is here *before* randomization.

TABLE 2
Three populations of size $N = 4$

Unit Number	p_i	Population A y_i	Population B y_i	Population C y_i
1	0.1	0.5	0.8	0.2
2	0.2	1.2	1.4	0.6
3	0.3	2.1	1.8	0.9
4	0.4	3.2	2.0	0.8

TABLE 3
Comparative efficiency of four sampling procedures

Procedure	Population A		Population B		Population C	
	Var.	Eff. %	Var.	Eff. %	Var.	Eff. %
1. Des Raj	0.200	100.0	0.200	100.0	0.100	100.0
2. Yates and Grundy	0.323	61.9	0.269	74.3	0.057	175.4
3. Hartley and Rao	0.367	54.5	0.367	54.5	0.033	333.3
4. With replacement	0.500	40.0	0.500	40.0	0.125	80.0

sampling *with* replacement for all the three populations. For populations A and B the linear model (1.4) is fairly well satisfied so that Des Raj's "optimum procedure" is more efficient. For population C the model is not appropriate so that considerable loss in efficiency results from Des Raj's procedure as compared to procedures 2 and 3. No general statement can be made between procedures 2 and 3 regarding efficiency. However, it can be shown that procedures 2 and 3 have exactly the same $V(\hat{Y})$ to order $O(N^1)$ for large N. For the present example with $N = 4$, this result for "large N" does not, of course, apply.

4. The case $n = 2$ and N large. Since the case $n = 2$ is useful particularly in stratified designs and since this case has been extensively dealt with in the literature, we consider this case in detail in the following subsections.

4.1 *Evaluation of the probabilities $P_{ii'}$*. For the evaluation of $P_{ii'}$ we make the following two assumptions: (a) $\pi_i \leqq cN^{-1}$ for all i and N, and (b) $c_1N^{-2} \leqq S_{ii'}^2 \leqq c_2N^{-2}$ for all pairs (i, i') and N, where c, c_1 and c_2 are universal constants and where $S_{ii'}^2$ is the mean square of the π_j $(j \neq i, i')$. It should be noted that the upper limit in (b) is, of course, a consequence of (a). While assumption (a) is vital, the lower limit in assumption (b) could be circumvented by a special argument not given here. Under the above two assumptions it can be shown that statements on the *relative* order of magnitude of our leading term $V'(\hat{Y})$ in the variance formula $V(\hat{Y})$ (see eq. 4.2.4) to the terms of lower order of magnitude can be made. In fact it can be shown that the tth lower order term in $V(\hat{Y})$ divided by $V'(\hat{Y})$ is less than or equal to const. N^{-t} $(t = 1, 2, 3)$, where the terms for $t = 1, 2$ have been retained in equation (4.2.2). In order to simplify the argument, however, we fix an arbitrary scale of *absolute* order of magnitude by making, in what follows, the additional assumptions that $0 \leqq y_i \leqq c^*$ for all i and N and that the variance formula in sampling with replacement, $V'(\hat{Y})$, is of order $O(N^2)$ (in practice this is usually the case). In sampling without replacement $V'(\hat{Y})$ will be the leading term in the variance formula and it is necessary, in order to supply formulas for moderately large N, to evaluate the term of next lower order of magnitude in powers of N^{-1}. This term will represent the gain in precision due to sampling without replacement. The variance of $\hat{Y}$ to order $O(N^1)$ is obtained by evaluating the probabilities $P_{ii'}$ to order $O(N^{-3})$ and substituting it in the variance formula (1.2). For the benefit of smaller size populations we also find $V(\hat{Y})$ to order $O(N^0)$ by evaluating $P_{ii'}$ to order $O(N^{-4})$ and substituting it in (1.2).

We use the circular analogue to the sampling procedure described in Subsection 2.3. The total number of arrangements of the N units on the circle, namely $N!$, can be divided into $(N - 1)$ groups according to whether there are $v = 0, 1, \cdots, (N - 2)$ units "between" i and i', where "between" means that we have v units when proceeding from unit i to i' in clockwise direction. There are $N \times (N - 2)!$ arrangements in each of these $(N - 1)$ groups so that they are all represented with equal probability $1/N - 1$. Consider now the contribution to $P_{ii'}$ from a *particular* group with v units between units i and i'. For the ith unit to be in the sample, the inequalities $\Pi_{i-1} \leqq s + k < \Pi_i$ must be satisfied where k can take values -1, 0 and 1 and s is a uniform arc with $0 \leqq s < 2$. This means that either $\Pi_{i-1} \leqq s < \Pi_i$ or $\Pi_{i-1} - 1 \leqq s < \Pi_i - 1$ if $\Pi_{i-1} \geqq 1$ and $\Pi_{i-1} + 1 \leqq s < \Pi_i + 1$ if $\Pi_i \leqq 1$ must be satisfied. Therefore, to evaluate $P_{ii'}$ we have to add the contributions to $P_{ii'}$ from the first case, say $P'_{ii'}$, and from the second case, say $P''_{ii'}$. Since the length of the range of s is equal to π_i in both cases, $P'_{ii'}$ and $P''_{ii'}$ are identical. Consider now the evaluation of $P'_{ii'}$. Since

$$\Pi_{i-1} \leqq s < \Pi_i, \tag{4.1.1}$$

a positive contribution to $P'_{ii'}$ can be made only if

$$\Pi_i + T_v \leqq s + 1 < \Pi_i + T_v + \pi_{i'} \tag{4.1.2}$$

is satisfied, where $T_v = \sum^v \pi_j$. (4.1.2) can be written as

$$1 + t - \pi_i - \pi_{i'} < T_v \leqq 1 + t - \pi_i,$$

where $t = s - \Pi_{i-1}$. The uniform variate t (like s) has an ordinate density of $\frac{1}{2}$, and from (4.1.1) we have $0 \leqq t < \pi_i$. Therefore, the integrated contribution to $P'_{ii'}$ is given by

$$\begin{aligned} &\int_0^{\pi_i} \tfrac{1}{2} \Pr(1 + t - \pi_i - \pi_{i'} < T_v \leqq 1 + t - \pi_i)\, dt \\ &\qquad = \tfrac{1}{2} \int_0^{\pi_i} [F_v(1 + t - \pi_i) - F_v(1 + t - \pi_i - \pi_{i'})]\, dt, \end{aligned} \tag{4.1.3}$$

where $F_v(T)$ denotes the cumulative distribution function of the total T_v. Since the units are arranged in a random order prior to drawing the sample, T_v represents the total of v values of the π_j drawn with equal probability and without replacement from the finite population of $(N - 2)$ values of the π_j. Therefore, noting that $\sum^N \pi_j = 2$, we find that

$$E(T_v) = \frac{v(2 - \pi_i - \pi_{i'})}{N - 2} = v\bar{\pi} \quad \text{(say)}$$

and

$$V(T_v) = v[1 - v/(N - 2)]S_{ii'}^2,$$

where

$$\begin{aligned} S_{ii'}^2 &= (N - 3)^{-1} \sum_{j \neq (i,i')}^{N} (\pi_j - \bar{\pi})^2 \\ &= (N - 3)^{-1} \left\{ \sum_1^N \pi_j^2 - \pi_i^2 - \pi_{i'}^2 - [(2 - \pi_i - \pi_{i'})^2/(N - 2)] \right\}. \end{aligned}$$

Finally, adding the two (identical) contributions to $P'_{ii'}$ and $P''_{ii'}$ given by (4.1.3), multiplying by the factor $(N - 1)^{-1}$ which represents the (constant) probability of a random arrangement of N units in which v units lie "between" units i and i', and summing over v, we obtain

$$\begin{aligned} P_{ii'} &= (N - 1)^{-1} \sum_{v=0}^{N-2} \\ &\cdot \int_0^{\pi_i} [F_v(1 + t - \pi_i) - F_v(1 + t - \pi_i - \pi_{i'})]\, dt. \end{aligned} \tag{4.1.4}$$

In order to obtain usable results, we now evaluate an approximation to (4.1.4) by expanding $F_v(T)$ in an Edgeworth series of which the cumulative normal integral is the leading term. Edgeworth series representation of a cumulative distribution function $F(x)$ is (see, e.g., Kendall and Stuart [5], p. 158) given by

$$F(x) = \exp\left\{ \sum_{j=3}^{\infty} D^j \frac{k_j}{j!} (-1)^j \right\} P(x), \tag{4.1.5}$$

where

$$P(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp\{-\tfrac{1}{2}y^2\}\, dy;$$

D^j is the jth order derivative operating on $P(x)$ and k_j denotes standardized cumulants. In what follows we assume, without loss of generality, that $i = 1$ and $i' = 2$. In our case, (4.1.5) is applied to the standardized variate

$$z_v = \frac{T_v - [v(2 - \pi_1 - \pi_2)/(N - 2)]}{S_{12}\{v[1 - v/(N - 2)]\}^{\frac{1}{2}}}$$

in place of x so that $F(x)$ is the finite proportion $F_v(z)$ (say) of values z_v with $z_v \leqq z$. This function is therefore a step function with a finite number of discontinuities and the complete series representation (4.1.5) will yield this step function for almost all values of z while at the points of discontinuity the right hand side of (4.1.5) will yield $\Pr(z_v < z) + \frac{1}{2}\Pr(z_v = z)$. We, therefore, have

$$F_v(z) = P(z) - (k_3/6)D^3P(z) + R(v), \tag{4.1.6}$$

where the k_j are the cumulants of z_v and

$$R(v) = \exp\left\{\sum_j \frac{k_j(-1)^j}{j!} D^j\right\} P(z) - \left\{1 - \frac{k_3}{3!} D^3\right\} P(z) \tag{4.1.7}$$

and therefore is a double infinite series each term involving a power product of the cumulants k_j and an associated high-order derivative $D^rP(z)$, the term with the least order differential is seen to be $(k_4/4!)\, D^4P(z)$. In order to express the cumulant k_3 in terms of the standardized cumulant K_3 (say) of the finite population of the π_j, we make use of the results given by Wishart [8]. In terms of the variable z_v we find, after some substitution, that

$$k_3 = \left[v^{-\frac{1}{2}}\left(1 - \frac{v}{N-2}\right)^{\frac{1}{2}} - \left\{v^{\frac{1}{2}}\left(1 - \frac{v}{N-2}\right)^{-\frac{1}{2}} \Big/ (N-2)\right\}\right] K_3 . \tag{4.1.8}$$

Substituting (4.1.6) in (4.1.4) we obtain

$$\begin{aligned} P_{12} &= (N-1)^{-1} \sum_{v=0}^{N-2} \\ &\cdot \int_0^{\pi_1} \left\{P(z_1) - P(z_2) - \frac{k_3}{6}[P^{(3)}(z_1) - P^{(3)}(z_2)]\right\} dt + \rho. \end{aligned} \tag{4.1.9}$$

where $D^jP(z) = P^{(j)}(z)$,

$$\begin{aligned} z_1 &= \{t + 1 - \pi_1 - v(2-\pi_1-\pi_2)(N-2)^{-1}\} \Big/ S_{12}\, v^{\frac{1}{2}}\left(1 - \frac{v}{N-2}\right)^{\frac{1}{2}}, \\ z_2 &= \{t + 1 - \pi_1 - \pi_2 - v(2-\pi_1-\pi_2)(N-2)^{-1}\} \Big/ S_{12}\, v^{\frac{1}{2}}\left(1 - \frac{v}{N-2}\right)^{\frac{1}{2}}, \end{aligned} \tag{4.1.10}$$

and the remainder term ρ is given by

$$\rho = (N-1)^{-1} \sum_{v=0}^{N-2} \int_0^{\pi_1} [R(z_1) - R(z_2)]\, dt \tag{4.1.11}$$

while k_3 is given by (4.1.8). We now apply the Euler-Maclaurin formula

$$\begin{aligned} \int_a^b g^{(1)}(t)\, dt &= g(b) - g(a) = (b-a)g^{(1)}\left(\frac{a+b}{2}\right) \\ &+ \frac{(b-a)^3}{24} g^{(3)}\left(\frac{a+b}{2}\right) + \frac{(b-a)^5}{1920} g^{(5)}(\bar{t}), \end{aligned} \tag{4.1.12}$$

here formulated for a general function $g(x)$ satisfying the required continuity conditions, where $g^{(r)}(x)$ denotes the rth order derivative with regard to x and $\bar{t}$ is such that $a \leqq \bar{t} \leqq b$. This formula is first applied to the differences

$$P(z_1) - P(z_2)$$

and $P^{(3)}(z_1) - P^{(3)}(z_2)$ in (4.1.9) thereby giving rise to four terms involving respectively $P^{(1)}$, $P^{(3)}$, $P^{(4)}$ and $P^{(6)}$. The integration over t of these four terms is then performed by again using (4.1.12). Retaining only the relevant terms in (4.1.9) we obtain

$$(4.1.13)\quad \begin{aligned} P_{12} = (N-1)^{-1} \int_0^{N-2} \Bigg[& \frac{\pi_1 \pi_2}{S_{12}} v_1^{-\frac{1}{2}} P^{(1)}(v_2) + \frac{\pi_1 \pi_2^3 v_1^{-\frac{3}{2}}}{24 S_{12}^3} P^{(3)}(v_2) \\ & + \frac{\pi_1^3 \pi_2 v_1^{-\frac{1}{2}}}{24 S_{12}} P^{(3)}(v_2) - \frac{k_3 \pi_1 \pi_2 v_1^{-\frac{1}{2}}}{6 S_{12}} P^{(4)}(v_2) \Bigg] dv + \rho + \omega + \rho', \end{aligned}$$

where $v_1 = v[1 - v/(N-2)]$,

$$(4.1.14)\quad v_2 = \left[1 - \tfrac{1}{2}(\pi_1 + \pi_2) - \frac{v(2 - \pi_1 - \pi_2)}{N-2}\right] \Big/ S_{12} v^{\frac{1}{2}} \left(1 - \frac{v}{N-2}\right)^{\frac{1}{2}},$$

and ρ is given by (4.1.11) while ω represents the aggregated remainder terms in the application of formula (4.1.12) and ρ' denotes the remainder when approximating $\sum_v$ by $\int dv$. The discussion of the remainder terms ρ, ω and ρ' is given in Appendix I where it is shown that these remainder terms do not contribute to P_{12} to the order of approximation desired, namely $O(N^{-4})$.

We now evaluate the remaining terms in (4.1.13). The first term which involves

$$P^{(1)}(v_2) = (2\pi)^{-\frac{1}{2}} \exp\{-v_2^2/2\}$$

will be called A. We now make the transformation

$$(4.1.15)\quad u = v - \tfrac{1}{2}(N-2)$$

so that $v_1 = \frac{1}{4}(N-2)\{1 - [4u^2/(N-2)]\}$. Expanding the exponential in $P^{(1)}(v_2)$ as well as the term $v_1^{-\frac{1}{2}}$ in powers of u, we obtain

$$(4.1.16)\quad \begin{aligned} A = \frac{(N-2)}{(N-1)} \frac{\pi_1 \pi_2}{(2 - \pi_1 - \pi_2)} (2\pi)^{-\frac{1}{2}} \int_{-h}^{h} e^{-p^2/2} \exp\{ & -\tfrac{1}{2}h^{-2}p^4 \\ - \tfrac{1}{2}h^{-4}p^6 + \text{higher terms}\}(1 + \tfrac{1}{2}h^{-2}p^2 + \tfrac{3}{8}h^{-4}p^4 & + \text{higher terms})\, dp, \end{aligned}$$

where

$$(4.1.17)\quad h = (2 - \pi_1 - \pi_2)(N-2)^{-\frac{1}{2}} S_{12}^{-1},$$

and the variable of integration has been changed to

$$(4.1.18)\quad p = 2\,uh(N-2)^{-1}.$$

Expanding the term exp { } in (4.1.16) and multiplying by the series inside

(4.1.16) we reach

$$A = \frac{(N-2)}{(N-1)} \frac{\pi_1 \pi_2}{(2-\pi_1-\pi_2)} (2\pi)^{-\frac{1}{2}} \int_{-h}^{h} e^{-p^2/2} \tag{4.1.19}$$
$$\cdot [1 + \tfrac{1}{2}h^{-2}(p^2 - p^4) + \tfrac{1}{8}h^{-4}(3p^4 - 6p^6 + p^8) + \text{higher terms}]\, dp.$$

Since h is $O(N^{\frac{1}{2}})$, we can (for large N) replace the integration limits in (4.1.19) by $-\infty$ and $+\infty$ apart from errors which are $O(e^{-N}N^a)$. Using the standardized normal moments we therefore obtain

$$A \doteq \frac{(N-2)}{(N-1)} \frac{\pi_1 \pi_2}{(2-\pi_1-\pi_2)} (1 - h^{-2} + 3h^{-4}) \tag{4.1.20}$$

to order $O(N^{-4})$. By a similar argument it can be shown that the first of the two terms in (4.1.13) which involves $P^{(3)}(v_2)$ and which we call B, can be reduced to

$$B = \frac{\pi_1 \pi_2^3 S_{12}^{-2} (2\pi)^{-\frac{1}{2}}}{6(N-1)(2-\pi_1-\pi_2)} \tag{4.1.21}$$
$$\cdot \int_{-h}^{h} e^{-p^2/2}[(p^2 - 1) - \tfrac{1}{2}h^{-2}(p^6 - 6p^4 + 3p^2) + \text{higher terms}]\, dp$$

while the other term involving $P^{(3)}(v_2)$, which we call C, reduces to

$$C = \frac{\pi_2 \pi_1^3 (N-2)(2\pi)^{-\frac{1}{2}}}{24(N-1)(2-\pi_1-\pi_2)} \tag{4.1.22}$$
$$\cdot \int_{-h}^{h} e^{-p^2/2}[(p^2 - 1) - \tfrac{1}{2}h^{-2}(p^6 - 4p^4 + p^2) + \text{higher terms}]\, dp.$$

The integral of the terms retained in (4.1.21) is zero while C can be seen to be $O(N^{-5})$ so that B and C do not contribute to P_{12} to order $O(N^{-4})$. Finally, consider the term D (say) involving k_3 in (4.1.13). Substituting for k_3 from (4.1.18), we obtain

$$D = \frac{2(N-2)^{\frac{1}{2}} \pi_1 \pi_2 K_3}{3(N-1)(2-\pi_1-\pi_2)}$$
$$\cdot (2\pi)^{-\frac{1}{2}} \int_{-h}^{h} e^{-p^2/2}[\tfrac{1}{2}h^{-1}(p^4 - 3p^2) + \tfrac{1}{4}h^{-3}(8p^6 - 9p^4 - p^8) + \text{higher terms}]\, dp,$$

and the evaluation of the terms retained yields

$$D \doteq -\frac{2K_3 S_{12}^3 (N-2)^2 \pi_1 \pi_2}{(N-1)(2-\pi_1-\pi_2)^4} \tag{4.1.23}$$

which is of order $O(N^{-4})$ since

$$K_3 S_{12}^3 = (N-2)^{-1} \sum_3^N \left(\pi_j - \frac{2-\pi_1-\pi_2}{N-2}\right)^3. \tag{4.1.24}$$

Adding the expressions A and D we obtain for the probability P_{12} an approximation to order $O(N^{-4})$ given by

$$(4.1.25) \quad P_{12} \doteq \frac{(N-2)\pi_1 \pi_2}{(N-1)(2-\pi_1-\pi_2)} [1 - h^{-2} + 3h^{-4} - 2K_3 h^{-3}(N-2)^{\frac{1}{2}}],$$

where h is given by (4.1.17) and K_3 by (4.1.24). Since the last two terms in (4.1.25) are $O(N^{-4})$, we obtain to order $O(N^{-3})$ the simplified formula

$$(4.1.26) \qquad P_{12} \doteq \frac{(N-2)\pi_1 \pi_2}{(N-1)(2-\pi_1-\pi_2)} (1 - h^{-2}).$$

We now apply two checks to verify, independently of the arguments given in Appendix I, that the remainder terms ρ, ω and ρ' do not contribute to P_{12} to order $O(N^{-4})$, and that all terms to order $O(N^{-4})$ are retained in (4.1.25) above. The first check is the special case when all π_j are equal to $2/N$ so that $S_{12} = 0$ and $h^{-1} = 0$, and (4.1.25) reduces to $2/N(N-1)$ which is the correct probability for units 1 and 2 to be both in a sample of size 2. However, this check tests only the leading term in (4.1.25) since $h^{-1} = 0$ so that the coefficients of the remaining terms in (4.1.25) are not affected by the check. A more searching check which takes account of all the terms in (4.1.25) is provided by testing the order to which the equation

$$(4.1.27) \qquad \sum_{i' \neq i}^{N} P_{ii'} = (n-1)\pi_i$$

is satisfied. It is shown in Appendix II that (4.1.27) is in fact satisfied to order $O(N^{-3})$ if (4.1.25) is substituted in (4.1.27) which confirms that (4.1.25) is correct to order $O(N^{-4})$.

4.2. *Variance formulas to orders* $O(N^1)$ *and* $O(N^0)$. In Appendix II we have simplified (4.1.25) to the following expression:

$$(4.2.1) \qquad \begin{aligned} P_{ii'} \doteq{} & \tfrac{1}{2}\pi_i \pi_{i'}[1 + \tfrac{1}{2}(\pi_i + \pi_{i'}) + \tfrac{1}{4}(\pi_i^2 + \pi_{i'}^2 + 2\pi_i \pi_{i'})] \\ & - \tfrac{1}{8}\pi_i \pi_{i'} \sum_1^N \pi_j^2[1 + \tfrac{3}{2}(\pi_i + \pi_{i'})] + \tfrac{1}{8}(\pi_i^3 \pi_{i'} + \pi_i \pi_{i'}^3) \\ & + \tfrac{3}{32}\pi_i \pi_{i'} \left(\sum_1^N \pi_j^2\right)^2 - \tfrac{1}{8}\pi_i \pi_{i'} \left(\sum_1^N \pi_j^3\right) \end{aligned}$$

to $O(N^{-4})$, where the subscripts 1 and 2 are replaced by i and i' respectively. Substituting (4.2.1) in the variance formula (1.2) we obtain after considerable rearrangement of terms

$$(4.2.2) \qquad \begin{aligned} V(\hat{Y}) \doteq{} & \sum_1^N \pi_i \left(1 - \frac{\pi_i}{2}\right)\left(\frac{y_i}{\pi_i} - \frac{Y}{2}\right)^2 \\ & - \tfrac{1}{2}\sum_1^N \left(\pi_i^3 - \tfrac{1}{4}\pi_i^2 \sum_1^N \pi_j^2\right)\left(\frac{y_i}{\pi_i} - \frac{Y}{2}\right)^2 + \tfrac{1}{4}\left(\sum_1^N \pi_j y_j - \tfrac{1}{2}Y \sum_1^N \pi_j^2\right)^2 \end{aligned}$$

correct to $O(N^0)$. If only terms to $O(N^1)$ are retained in (4.2.2), we obtain

$$V(\hat{Y}) \doteq \sum_1^N \pi_i \left(1 - \frac{\pi_i}{2}\right)\left(\frac{y_i}{\pi_i} - \frac{Y}{2}\right)^2 \tag{4.2.3}$$

correct to $O(N^1)$. For the special case of equal probabilities $\pi_i = 2/N$, (4.2.2) reduces to the well known variance formula in equal probability sampling with-

TABLE 4

Data for population of size $N = 20$

$i =$	1	2	3	4	5	6	7	8	9	10
y_i	19	9	17	14	21	22	27	35	20	15
x_i	18	9	14	12	24	25	23	24	17	14
$i =$	11	12	13	14	15	16	17	18	19	20
y_i	18	37	12	47	27	25	25	13	19	12
x_i	18	40	12	30	27	26	21	9	19	12

out replacement to $O(N^0)$. In sampling *with* replacement the variance of $\hat{Y}$ is

$$V'(\hat{Y}) = \sum_1^N \pi_i \left(\frac{y_i}{\pi_i} - \frac{Y}{2}\right)^2 \tag{4.2.4}$$

which is of order $O(N^2)$. (4.2.3) when compared with (4.2.4) clearly shows the characteristic reduction in the variance through the "finite population corrections" $(1 - \frac{1}{2}\pi_i)$.

4.3. *Numerical example.* Horvitz and Thompson [4] give an example of $N = 20$ blocks in Ames, Iowa. The data are reproduced in Table 4 below where y_i denotes the number of households in ith block and x_i denotes the eye-estimated number of households in ith block. The probability π_i for the ith unit to be in the sample is taken proportional to the eye-estimated number of households x_i, i.e., $\pi_i = 2x_i/X$. In Table 5 below we give the evaluation of the variance of $\hat{Y}$ from formulas (4.2.2), (4.2.3) and (4.2.4). Also shown in Table 5 are the evaluations of the variances of alternative estimators which are described by Horvitz and Thompson.

A comparison of the variances in Table 5 shows that sampling with p.p.s. is vastly superior to sampling with equal probabilities. It must not be forgotten, however, that there are other devices of decreasing the variance in the latter case with the help of the known x_i values, e.g., ratio and regression methods of estimation. Among the procedures of p.p.s. sampling there is little to choose in this example except that about 7% (235/3241) is gained in precision through our procedure of sampling without replacement as compared to sampling with replacement. It is of interest to exhibit the nature of convergence of the various approximations to $V(\hat{Y})$ by regarding the variance formula for sampling with replacement as an approximation to $O(N^2)$ as set out in Table 6. The con-

TABLE 5

Variances of various estimators of the total of the y_i-population shown in Table 4

Sampling Procedure	Form of Estimator	Numerical Value of Variance of Estimator
Equal probability sampling without replacement	$N\bar{y}$	16,219
p.p.s. with replacement	$\sum^n y_i/\pi_i$	3,241-Eq. (4.2.4)
* Horvitz and Thompson procedure 1	$\sum^n y_i/P_i$	3,095
Horvitz and Thompson procedure 2	$\sum^n y_i/P_i$	3,075
Present procedure; sampling without replacement and p.p.s.	$\sum^n y_i/\pi_i$	3,025-Eq. (4.2.3) 3,007-Eq. (4.2.2)

* For a description of these procedures and definition of the P_i see Horvitz and Thompson [4], Tables 2 and 3.

TABLE 6

Approximations to $V(\hat{Y})$ for population of Table 4

Order of Approximation	Formula Used	$V(\hat{Y})$	Difference
$O(N^2)$	Eq. (4.2.4)	3,241	216
$O(N^1)$	Eq. (4.2.3)	3,025	18
$O(N^0)$	Eq. (4.2.2)		

vergence in this example appears to be satisfactory although the population size ($N = 20$) is much smaller than those encountered in survey work.

4.4. *Estimate of the variance.* Horvitz and Thompson [4] give an unbiased estimate of $V(\hat{Y})$ that seems to be unsatisfactory since it may take negative values. Yates and Grundy [9] propose an alternative unbiased estimator of variance which is believed to take negative values less often, and is given by

$$v(\hat{Y}) = \sum_{i<i'}^{n} \frac{\pi_i \pi_{i'} - P_{ii'}}{P_{ii'}} \left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2, \tag{4.4.1}$$

where $v(\hat{Y})$ denotes the estimator of $V(\hat{Y})$. For the case $n = 2$, (4.4.1) reduces to

$$v(\hat{Y}) = \frac{\pi_i \pi_{i'} - P_{ii'}}{P_{ii'}} \left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2, \tag{4.4.2}$$

where the units numbered i and i' are included in the sample of size 2. Substituting for $P_{ii'}$ from (4.2.1) in (4.4.2), we find after some simplification

$$\begin{aligned} v(\hat{Y}) \doteq \Bigg[1 - (\pi_i + \pi_{i'}) + \tfrac{1}{2}\sum_1^N \pi_j^2 - \tfrac{1}{2}(\pi_i^2 + \pi_{i'}^2) - \tfrac{1}{4}\left(\sum_1^N \pi_j^2\right)^2 \\ + \tfrac{1}{4}(\pi_i + \pi_{i'})\sum_1^N \pi_j^2 + \tfrac{1}{2}\sum_1^N \pi_j^3\Bigg]\left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2 \end{aligned} \tag{4.4.3}$$

to order $O(N^0)$. If only terms to order $O(N^1)$ are retained in (4.4.3), we get the simplified formula

$$v(\hat{Y}) \doteq \left[1 - (\pi_i + \pi_{i'}) + \tfrac{1}{2}\sum_1^N \pi_j^2\right]\left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2. \tag{4.4.4}$$

For the special case of equal probabilities $\pi_j = 2/N$, both (4.4.3) and (4.4.4) reduce to the well known formula for estimator of variance in equal probability sampling without replacement.

In this connection, it may be worth while to point out an important aspect of Yates and Grundy estimator of variance for the case $n = 2$. Narain [7] has shown that a necessary condition for a sampling procedure without replacement, for which $\pi_j = np_j$, to be more efficient than sampling with replacement is

$$P_{ii'} \leqq \frac{2(n-1)}{n}\pi_i\pi_{i'} \tag{4.4.5}$$

for all i and i'. For the case $n = 2$, (4.4.5) reduces to

$$P_{ii'} \leqq \pi_i\pi_{i'}. \tag{4.4.6}$$

Therefore, using (4.4.6), it immediately follows that Yates and Grundy's estimator of variance (4.4.2) is always positive. That is, if there is a sampling procedure without replacement for which the variance is smaller than the variance when sampling with replacement independent of the y_j's, which is the case we are interested in, then Yates and Grundy's estimator of variance is always positive. It may be noted that this result is true only for $n = 2$, since conditions (4.4.5) are not sufficient to show that (4.4.1) is always positive when $n > 2$. (4.4.1) is always positive if conditions (4.4.6) for all i and i' $(i \neq i')$ are satisfied. However, conditions (4.4.5) do not imply (4.4.6) except when $n = 2$.

(4.2.3) and (4.2.4) imply that, to $O(N^{-2})$, our *particular* sampling procedure without replacement is more precise than sampling with replacement independent of the y_j's, so that (4.4.6) follows as a necessary condition to $O(N^{-3})$. Conditions (4.4.6) can, of course, be directly verified for our sampling procedure using (4.2.1).

5. The general case $n \geqq 2$ and N large. A striking feature of our sampling procedure is that it permits an evaluation of $P_{ii'}$, and hence of $V(\hat{Y})$ and $v(\hat{Y})$, for the case $n > 2$. It is interesting to note that most of the published literature on this topic does not have anything to offer for $n > 2$, and deals only with the case $n = 2$, due to difficulties in evaluating $P_{ii'}$.

The methods of attack for $n > 2$ are similar to those used for $n = 2$. However, the former case presents certain new features. Therefore, in this section we give a brief derivation of $P_{ii'}$ with details of the new features.

Extending the arguments given in Sub-section 4.1, it can be shown that

$$
\begin{aligned}
P_{ii'} = (N-1)^{-1}\Bigg\{ &\sum_{v=0}^{N-n} \int_0^{\pi_i} [F_v(1+t-\pi_i) - F_v(1+t-\pi_i-\pi_{i'})]\,dt + \cdots \\
(5.1) \qquad &+ \sum_{v=m}^{N-n+m} \int_0^{\pi_i} [F_v(m+1+t-\pi_i) - F_v(m+1+t-\pi_i-\pi_{i'})]\,dt + \cdots \\
&+ \sum_{v=n-2}^{N-2} \int_0^{\pi_i} [F_v(n-1+t-\pi_i) - F_v(n-1+t-\pi_i-\pi_{i'})]\,dt \Bigg\},
\end{aligned}
$$

where, as before, $F_v(T)$ denotes the cumulative distribution function of the total T_v of the v values π_j,

$$E(T_v) = v(n - \pi_i - \pi_{i'})/(N-2),$$

and

$$V(T_v) = v\{1 - [v/(N-2)]\} S_{ii'}^2 .$$

It may be noted that (5.1) reduces to (4.1.4) when $n = 2$. Now it will be shown that each of the $(n-1)$ integrals summed over v in (5.1) contributes identically to $P_{ii'}$ to order $O(N^{-4})$, making the assumptions mentioned earlier in Subsection 4.1. Consider the mth integral summed over v ($m = 0, 1, \cdots, n-2$) in (5.1), say $P_{ii'}^{(m)}$, given by

$$
\begin{aligned}
&P_{ii'}^{(m)} \\
&= (N-1)^{-1} \sum_{v=m}^{N-n+m} \int_0^{\pi_i} [F_v(m+1+t-\pi_i) - F_v(m+1+t-\pi_i-\pi_{i'})]\,dt,
\end{aligned}
$$

and let $i = 1$ and $i' = 2$ without loss of generality. Proceeding now exactly as in the case of $n = 2$, by expanding $F_v(T)$ in an Edgeworth series, applying the Euler-Maclaurin formula (4.1.12), and approximating $\sum_v$ by $\int dv$, we find

$$
\begin{aligned}
P_{12}^{(m)} = (N-1)^{-1} \int_m^{N-n+m} \Bigg[&\frac{\pi_1\pi_2}{S_{12}} v_1^{-1/2} P^{(1)}(v_{2m}) + \frac{\pi_1\pi_2^3}{24 S_{12}^3} v_1^{-3/2} P^{(3)}(v_{2m}) \\
(5.2) \qquad &+ \frac{\pi_1^3\pi_2}{24 S_{12}} v_1^{-1/2} P^{(3)}(v_{2m}) - \frac{k_3\pi_1\pi_2}{6 S_{12}} v_1^{-1/2} P^{(4)}(v_{2m}) \Bigg] dv + \rho_m + \omega_m + \rho_m'
\end{aligned}
$$

where

$$v_1 = v\{1 - [v/(N-2)]\},$$

$$v_2 = \left[m + 1 - \frac{\pi_1 + \pi_2}{2} - \frac{v(n - \pi_1 - \pi_2)}{N-2} \right] \Big/ (v_1 S_{12}),$$

ρ_m, ω_m and ρ_m' are the remainder terms similar to ρ, ω and ρ' for the case $n = 2$, $P^{(r)}(x)$ denotes the rth order derivative of the cumulative normal distribution $P(x)$, and k_3 is the standardized cumulant of the total T_v given by (4.1.8). It can be shown that ρ_m, ω_m and ρ_m' do not contribute to $P_{12}^{(m)}$ to $O(N^{-4})$, using arguments similar to those employed for ρ, ω and ρ' in Appendix I. We now evaluate the remaining terms in (5.2). The first term which involves $P^{(1)}(v_{2m})$ will be called A_m. We make the transformation

$$v - c_m = u, \tag{5.3}$$

where

$$\frac{c_m}{N-2} = \frac{2(m+1) - \pi_1 - \pi_2}{2(n - \pi_1 - \pi_2)}.$$

Then, in terms of u, we have

$$v_1 = \left(c_m - \frac{c_m^2}{N-2}\right)\left[1 + \frac{u\left(1 - \frac{2c_m}{N-2}\right)}{\left(c_m - \frac{c_m^2}{N-2}\right)} - \frac{u^2}{(N-2)\left(c_m - \frac{c_m^2}{N-2}\right)}\right].$$

Expanding now the exponential in $P^{(1)}(v_{2m})$ as well as the term $v_1^{-\frac{1}{2}}$ in powers of u (justification of the expansion can be easily shown), and changing the variable of integration u to p, where

$$p = uh(N-2)^{-\frac{1}{2}}\left(c_m - \frac{c_m^2}{N-2}\right)^{-\frac{1}{2}} \tag{5.4}$$

where

$$h = (n - \pi_1 - \pi_2)(N-2)^{-\frac{1}{2}}S_{12}^{-1}, \tag{5.5}$$

we find, after considerable simplification, that

$$\begin{aligned} A_m = \frac{(N-2)\pi_1\pi_2(2\pi)^{-\frac{1}{2}}}{(N-1)(n-\pi_1-\pi_2)} \int e^{-p^2/2}\Big[1 &+ \frac{h^{-2}}{2}(p^2 - p^4) \\ &+ \frac{h_{1m}}{8}(3p^2 - 6p^4 + p^6) + \frac{h^{-4}}{8}(3p^4 - 6p^6 + p^8) \\ &+ \frac{h_{1m}^2 h^{-2}}{16}(15p^4 - 45p^6 + 15p^8 - p^{10}) \\ &+ \frac{h_{1m}^4}{384}(105p^4 - 420p^6 + 210p^8 - 28p^{10} + p^{12}) + \text{higher terms}\Big]\,dp, \end{aligned} \tag{5.6}$$

where

$$h_{1m} = \frac{(N-2)S_{12}\left(1 - \frac{2c_m}{N-2}\right)}{(n - \pi_1 - \pi_2)\left(c_m - \frac{c_m^2}{N-2}\right)^{\frac{1}{2}}}.$$

The limits of integration in (5.6) are

$$h(m - c_m)(N-2)^{-\frac{1}{2}}\{c_m - [c_m^2/(N-2)]\}^{-\frac{1}{2}}$$

and $h(N - n + m - c_m)(N-2)^{-\frac{1}{2}}\{c_m - [c_m^2/(N-2)]\}^{-\frac{1}{2}}$, which are $-O(N^{\frac{1}{2}})$ and $O(N^{\frac{1}{2}})$ respectively. Therefore, the integration limits can be replaced by $-\infty$ and $+\infty$ apart from errors which are $O(e^{-N}N^a)$. The main feature here is the appearance of a "non-centrality type" term h_{1m} which is zero for the case $n = 2$. Using now the first twelve standardized normal moments, we find that the coefficients of all terms involving h_{1m} are zero, and that

$$A_m \doteq \frac{(N-2)\pi_1\pi_2}{(N-1)(n-\pi_1-\pi_2)}(1 - h^{-2} + 3h^{-4}) \tag{5.7}$$

to $O(N^{-4})$. Since h does not depend on m, A_m also does not depend on m.

By a similar analysis, it can be shown that the first of the two terms in (5.2) which involves $P^{(3)}(v_{2m})$ and which we call B_m, can be reduced to

$$B_m = \frac{(N-2)\pi_1\pi_2^3 S_{12}^{-2}}{24(N-1)(n-\pi_1-\pi_2)}(2\pi)^{-\frac{1}{2}}\left(c_m - \frac{c_m^2}{N-2}\right)^{-1}$$

$$\tag{5.8} \cdot \int e^{-p^2/2}\left[(p^2-1) + \frac{h_{1m}}{2}(p^5 - 6p^3 + 3p) - \frac{h^{-2}}{2}(p^6 - 6p^4 + 3p^2) + \frac{h_{1m}^2}{8}(p^8 - 15p^6 + 45p^4 - 15p^2) + \text{higher terms}\right]dp$$

while the other term involving $P^{(3)}(v_{2m})$, which we call C_m, reduces to

$$C_m = \frac{(N-2)\pi_1^3\pi_2(2\pi)^{-\frac{1}{2}}}{24(N-1)(n-\pi_1-\pi_2)}$$

$$\tag{5.9} \cdot \int e^{-p^2/2}\left[(p^2-1) + \frac{h_{1m}}{2}(p^5 - 4p^3 + p) - \frac{h^{-2}}{2}(p^6 - 4p^4 + p^2) + \frac{h_{1m}^2}{8}(p^8 - 11p^6 + 21p^4 - 3p^2) + \text{higher terms}\right]dp.$$

Using now the standardized normal moments, we find that the integral of the terms retained in B_m is zero and C_m is of order $O(N^{-5})$, so that B_m and C_m do not contribute to $P_{12}^{(m)}$ to order $O(N^{-4})$.

Finally, the term D_m (say), involving k_3, in (5.2) is reduced to

$$D_m = -\frac{(N-2)\pi_1\pi_2K_3(2\pi)^{-\frac{1}{2}}}{6(N-1)(n-\pi_1-\pi_2)}\int e^{-p^2/2}\left\{\frac{\left(1 - \frac{2c_m}{N-2}\right)}{\left(c_m - \frac{c_m^2}{N-2}\right)}\right\}\Bigg[(p^3 - 3p)$$

$$+ \frac{h_{1m}}{2}(p^6 - 6p^4 + 3p^2) - \frac{h^{-2}}{2}(p^7 - 6p^5 + 3p^3)$$

$$+ \frac{h_{1m}^2}{8}(p^9 - 13p^7 + 33p^5 - 9p^3)$$

$$- \frac{h_{1m}h^{-2}}{4}(p^{10} - 13p^8 + 33p^6 - 9p^4)$$

$$+ \frac{h_{1m}^3}{48}(p^{12} - 24p^{10} + 150p^8 - 240p^6 + 45p^4)\Bigg]$$

$$- h^{-1}(N-2)^{\frac{1}{2}}\left(c_m - \frac{c_m^2}{N-2}\right)$$

$$\cdot\left[\frac{1}{c_m^2} + \frac{1}{(N-2)^2\left(1 - \frac{c_m}{N-2}\right)^2}\right]\Bigg[(p^4 - 3p^2)$$

$$+ \frac{h_{1m}}{2}(p^7 - 6p^5 + 3p^4) - \frac{h^{-2}}{2}(p^8 - 6p^6 + 3p^4) \tag{5.10}$$

$$+ \frac{h_{1m}^2}{8}(p^{10} - 13p^8 + 33p^6 - 9p^4)\Bigg]$$

$$+ h^{-2}(N-2)\left(c_m - \frac{c_m^2}{N-2}\right)^{\frac{3}{2}}$$

$$\cdot\left[\frac{1}{c_m^3} - \frac{1}{(N-2)^3\left(1 - \frac{c_m}{N-2}\right)^3}\right]\Bigg[(p^5 - 3p^3)$$

$$+ \frac{h_{1m}}{2}(p^8 - 6p^6 + 3p^4)\Bigg] - h^{-3}(N-2)^{\frac{3}{2}}\left(c_m - \frac{c_m^2}{N-2}\right)^2$$

$$\cdot\left[\frac{1}{c_m^4} + \frac{1}{(N-2)^4\left(1 - \frac{c_m}{N-2}\right)^4}\right](p^6 - 3p^4)$$

$$+ \text{higher terms}\Bigg\}\,dp.$$

Using now the standardized normal moments, the integral of the terms retained in (5.10) reduces to

$$D_m \doteq \frac{-(N-2)\pi_1\pi_2 K_3}{6(N-1)(n - \pi_1 - \pi_2)} h^{-3}(N-2)^{\frac{3}{2}}$$

$$\cdot\left\{-\frac{12\left(1 - \frac{2c_m}{N-2}\right)^2}{(N-2)\left(c_m - \frac{c_m^2}{N-2}\right)}\right.$$

$$- 6\left(1 - \frac{2c_m}{N-2}\right)^2\left[\frac{1}{c_m^2} + \frac{1}{(N-2)^2\left(1 - \frac{c_m}{N-2}\right)^2}\right] \tag{5.11}$$

$$\left. + 6c_m^{-2}\left(1 - \frac{c_m}{N-2}\right)^2 + \frac{6c_m^2}{(N-2)^4\left(1 - \frac{c_m}{N-2}\right)^2}\right\}.$$

Further simplification of (5.11) results in

$$D_m \doteq -\frac{2(N-2)\pi_1\pi_2 K_3}{(N-1)(n-\pi_1-\pi_2)} h^{-3}(N-2)^{-\frac{1}{2}} \tag{5.12}$$

which is of order $O(N^{-4})$ and does not depend on m. Adding now the expressions (5.7) and (5.12) for A_m and D_m respectively, we obtain for $P_{12}^{(m)}$ an approximation to order $O(N^{-4})$ given by

$$P_{12}^{(m)} \doteq \frac{(N-2)\pi_1\pi_2}{(N-1)(n-\pi_1-\pi_2)} [1 - h^{-2} + 3h^{-4} - 2K_3 h^{-3}(N-2)^{-\frac{1}{2}}]. \tag{5.13}$$

Since (5.13) does not depend on m, it follows that

$$P_{12} = \sum_{m=0}^{n-2} P_{12}^{(m)} \doteq (n-1) \frac{(N-2)\pi_1\pi_2}{(N-1)(n-\pi_1-\pi_2)} [1 - h^{-2} + 3h^{-4} - 2K_3 h^{-3}(N-2)^{-\frac{1}{2}}] \tag{5.14}$$

to order $O(N^{-4})$. For the special case $n = 2$, (5.14) reduces to (4.1.25). The two checks mentioned in Section 4 for the case $n = 2$ are also satisfied by (5.14). Expressing h and K_3 in terms of the π_j we obtain from (5.14):

$$\begin{aligned} P_{ii'} \doteq & \frac{(n-1)}{n}\pi_i\pi_{i'} + \frac{(n-1)}{n^2}(\pi_i^2\pi_{i'} + \pi_i\pi_{i'}^2) \\ & - \frac{(n-1)}{n^3}\pi_i\pi_{i'}\sum_1^N \pi_j^2 + \frac{2(n-1)}{n^3}(\pi_i^3\pi_{i'} + \pi_i\pi_{i'}^3 + \pi_i^2\pi_{i'}^2) \\ & - \frac{3(n-1)}{n^4}(\pi_i^2\pi_{i'} + \pi_i\pi_{i'}^2)\sum_1^N \pi_j^2 + \frac{3(n-1)}{n^5}\pi_i\pi_{i'}\left(\sum_1^N \pi_j^2\right)^2 \\ & - \frac{2(n-1)}{n^4}\pi_i\pi_{i'}\sum_1^N \pi_j^3 \end{aligned} \tag{5.15}$$

to $O(N^{-4})$, where the subscripts 1 and 2 are replaced by i and i' respectively.

Substituting (5.15) in the variance formula (1.2) we obtain

$$\begin{aligned} V(\hat{Y}) \doteq & \sum_1^N \pi_i \left[1 - \frac{(n-1)}{n}\pi_i\right]\left(\frac{y_i}{\pi_i} - \frac{Y}{n}\right)^2 \\ & - \frac{(n-1)}{n^2}\sum_1^N \left(2\pi_i^3 - \frac{\pi_i^2}{n}\sum_1^N \pi_j^2\right)\left(\frac{y_i}{\pi_i} - \frac{Y}{n}\right)^2 \\ & + \frac{2(n-1)}{n^3}\left(\sum_1^N \pi_j y_j - \frac{Y}{n}\sum_1^N \pi_j^2\right)^2 \end{aligned} \tag{5.16}$$

correct to $O(N^0)$. If only terms to $O(N^1)$ are retained in (5.16), we obtain

$$V(\hat{Y}) \doteq \sum_1^N \pi_i \left[1 - \frac{(n-1)}{n}\pi_i\right]\left(\frac{y_i}{\pi_i} - \frac{Y}{n}\right)^2 \tag{5.17}$$

correct to $O(N^1)$. The variance of $\hat{Y}$ in sampling *with* replacement is

$$(5.18) \qquad V'(\hat{Y}) = \sum_1^N \pi_i \left(\frac{y_i}{\pi_i} - \frac{Y}{n}\right)^2$$

which is of order $O(N^2)$. (5.17) when compared with (5.18) clearly shows the characteristic reduction in the variance through the "finite population corrections" $\{1 - [(n - 1)/n]\pi_j\}$.

Substituting (5.15) in Yates and Grundy's estimator of variance (4.4.1), we find

$$(5.19) \qquad \begin{aligned} v(\hat{Y}) \doteq (n-1)^{-1} \sum_{i<i'}^{n} \Bigg[1 &- (\pi_i + \pi_{i'}) + \left(\sum_1^N \pi_j^2 \Big/ n\right) \\ &- \frac{1}{n}(\pi_i^2 + \pi_{i'}^2) - \frac{2}{n^3}\left(\sum_1^N \pi_j^2\right)^2 + \frac{1}{n^2}(\pi_i + \pi_{i'}) \sum_1^N \pi_j^2 \\ &+ \frac{2}{n^2} \sum_1^N \pi_j^3\Bigg] \left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2 \end{aligned}$$

correct to $O(N^0)$. If only terms to $O(N^1)$ are retained in (5.19), we obtain the simplified formula

$$(5.20) \qquad v(\hat{Y}) \doteq (n-1)^{-1} \sum_{i<i'}^{n} \left[1 - (\pi_i + \pi_{i'}) + \left(\sum_1^N \pi_j^2 \Big/ n\right)\right] \left(\frac{y_i}{\pi_i} - \frac{y_{i'}}{\pi_{i'}}\right)^2$$

correct to $O(N^1)$. For the special case of equal probabilities $\pi_i = n/N$, both (5.19) and (5.20) reduce to the familiar formula for estimator of variance in equal probability sampling without replacement.

6. A comparison with ratio method of estimation. Cochran [2] makes a comparison of the variance of $\hat{Y}$ in p.p.s. sampling *with* replacement with the variance of the ratio estimate, $\hat{Y}_R = (\bar{y}/\bar{x})X$, in equal probability sampling without the finite population correction factor. Since a compact expression for the variance of $\hat{Y}$ in p.p.s. sampling without replacement is obtained in this paper, it will be of interest to compare this variance with the variance of the ratio estimate not ignoring the finite population correction factor. (5.17) can be written as

$$(6.1) \qquad V(\hat{Y}) \doteq \frac{1}{n} \sum_1^N \frac{1}{p_i}(y_i - Yp_i)^2 - \frac{(n-1)}{n} \sum_1^N (y_i - Yp_i)^2$$

to order $O(N^1)$. The variance of the ratio estimate $\hat{Y}_R$ for large samples is

$$(6.2) \qquad V(\hat{Y}_R) = \frac{N^2}{n(N-1)}\left(1 - \frac{n}{N}\right) \sum_1^N (y_i - Yp_i)^2.$$

Since $N - 1 = N[1 - (1/N)]$, expanding $(N - 1)^{-1}$ binomially we obtain from (6.2),

$$(6.3) \qquad V(\hat{Y}_R) \doteq \frac{N}{n} \sum_1^N (y_i - Yp_i)^2 - \frac{(n-1)}{n} \sum_1^N (y_i - Yp_i)^2$$

to $O(N^1)$. Since the correction factors in (6.1) and (6.3) are exactly the same, the comparison of $V(\hat{Y})$ with $V(\hat{Y}_R)$ reduces to the comparison of the variance of $\hat{Y}$ in p.p.s. sampling *with* replacement with the variance of the ratio estimate without the correction factor. Cochran, assuming the model $y_i = Yp_i + e_i$ where $E(e_i \mid p_i) = 0$ and $E(e_i^2 \mid p_i) = ap_i^g$, $g > 0$, $a > 0$, has shown that the variance of $\hat{Y}$ in p.p.s. sampling *with* replacement is smaller or greater than the variance of the ratio estimate without the correction factor according as $g > 1$ or $g < 1$ respectively. If $g = 1$, the variances are identical. It is also stated that in practice g usually lies between 1 and 2, so that the p.p.s. estimate is generally more precise.

APPENDIX I

The order of magnitude of the remainder terms ρ', ω and ρ

ρ' denotes the remainder when approximating $\sum_v$ by $\int dv$. It involves the terminal differentials of the integrands at the end points of integration $v = 0$ and $v = N - 2$ which become zero since v_2 is infinite at these points and since the integrands involve the term $\exp(-v_2^2/2)$. ρ' also involves the remainder term of the Euler-Maclaurin formula which is of the form

$$(N - 2)B_{2m}f^{(2m)}[(N - 2)\theta_N]/(2m)!,$$

where B_{2m} is the Bernoulli number, $f^{(2m)}$ is the $(2m)$th derivative with regard to v of any of the integrand functions involved in (4.1.13) and $0 < \theta < 1$ while the order of the remainder term, $2m$, is at our disposal.

The relation between the v argument of $(N - 2)\theta_N$ and the corresponding v_2 argument, from (4.1.14), is given by

$$v_2 = [1 - \tfrac{1}{2}(\pi_1 + \pi_2)](1 - 2\theta_N)[\theta_N(1 - \theta_N)]^{-\frac{1}{2}}S_{12}^{-\frac{1}{2}}(N - 2)^{-\frac{1}{2}}.$$

We now separate the values of θ_N between 0 and 1 into two groups. In the first group, θ_N is equal to $\frac{1}{2}$ or the leading term of the difference between θ_N and $\frac{1}{2}$ is proportional to N^{-r_N} with $r_N > 0$. The remaining values of θ_N fall in the second group. It is easily seen that for the values of θ_N in the second group v_2 is $O(N^s)$ with $s > 0$ since S_{12} is of order $O(N^{-1})$, and the argument to be used for the remainder term in case (b) below also applies to the values of θ_N in this group. For values of θ_N in the first group, either v_2 is zero or is of order $O(N^{\frac{1}{2}-r_N})$. We now distinguish the two cases (a) $r_N \geqq \frac{1}{2}$ and (b) $r_N < \frac{1}{2}$. Consider first the case (a). Introducing the variable u given by (4.1.15) in (4.1.14), we find

$$\begin{aligned} v_2 &= \text{const.}\ (N - 2)^{-\frac{1}{2}}S_{12}^{-\frac{1}{2}}u[1 - (2u/N - 2)^2]^{-\frac{1}{2}} \\ &= \text{const.}\ (N - 2)^{-\frac{1}{2}}S_{12}^{-\frac{1}{2}} \sum_{i=0}^{\infty} \binom{-\frac{1}{2}}{i} (-1)^i(2u/N - 2)^{2i+1}. \end{aligned}$$

Therefore, by repeated differentiation, we have that for the largest value of $|u|$,

$$\frac{d^t v_2}{dv^t} = \frac{d^t v_2}{du^t} = O(N^{\frac{1}{2}-t}).$$

The repeated differentiation of $f(v_2)$ with regard to v or u will now be seen to have a leading term of the form $(d^t f/dv_2^t)(dv_2/du)^t$ which is of order $O(N^{-t/2})$. Therefore, from the Leibnitz formula of differentiation of a product which is of the form $v_1^{-b}f(v_2)$, $b > 0$, as evident from (4.1.13), it is seen that the remainder term is $O(N^{-k})$ with $k > 4$ provided $2m$ is chosen sufficiently large. In case (b), the remainder term goes down as $O(e^{-bN^s}N^a)$ where $s = 1 - 2r_N > 0$, and hence is smaller than $O(N^{-4})$. Therefore, the remainder term ρ' does not contribute to P_{12} to $O(N^{-4})$.

We consider next the magnitude of the remainder ω arising from the application of the Euler-Maclaurin formula (4.1.12) to the differences $P(z_1) - P(z_2)$ and $P^{(3)}(z_1) - P^{(3)}(z_2)$. The first of these is of the form $c(z_1 - z_2)^5 P^{(5)}(\bar{z})$ where $z_2 < \bar{z} < z_1$. Now from (4.1.10), (4.1.15) and (4.1.18) we have

$$z_1 - z_2 = 2\pi_2(N-2)^{-\frac{1}{2}}S_{12}^{-\frac{1}{2}}[1 + O(p^2N^{-1})],$$

from which it appears that the leading term of $z_1 - z_2$ is of order $O(N^{-\frac{1}{2}})$ and that $\bar{z} = v_2 + O(N^{-\frac{1}{2}})$. Therefore, a contribution ω_1 (say) to ω is given by

$$\omega_1 = (N-1)^{-1}\sum_v \int_0^{\pi_1} (z_1 - z_2)^5 P^{(5)}(\bar{z})\,dt$$
$$= (N-1)^{-1}\pi_1[2\pi_2(N-2)^{-\frac{1}{2}}S_{12}^{-\frac{1}{2}}]^5 \sum_v [1 + O(p^2N^{-1})][P^{(5)}(v_2) + O(N^{-\frac{1}{2}}])).$$

Now because of the properties of the normal differentials, $\sum_v P^{(5)}(v_2)$ is $O(e^{-N}N^a)$ and, since $dv/dv_2 = O(N^{\frac{1}{2}})$ and $dp/dv_2 = O(N^0)$, all terms in ω_1 are seen to be of smaller order than $O(N^{-4})$. Similar arguments apply to the difference $P^{(3)}(z_1) - P^{(3)}(z_2)$ as well as to the remainder terms arising from applying Euler-Maclaurin formula to the integration $\int_0^{\pi_1} dt$ of the retained terms in $P^{(1)}$, $P^{(3)}$, $P^{(4)}$ and $P^{(6)}$.

Finally, we turn to the discussion of the remainder ρ given by the sum (4.1.11) with $R(v)$ given by the double expansion (4.1.7) involving the cumulants k_r and higher differentials of the standardized normal $P(z)$. It is necessary here to show that k_r $(r \geqq 4)$ is of order $O(N^{-c})$ with $c > \frac{1}{2}$ provided that

$$0 < q \leqq v/N \leqq 1 - q < 1$$

as $N \to \infty$. We have shown that k_r is of order $O(N^{-\frac{1}{2}r+1})$ when the K_r are bounded and v/N satisfies the above condition.[3] However, we shall not give this discussion here. Instead, we use a formula recently employed by Barton and David [1] who show that, if the set of finite populations of size N are assumed to be random samples from the *same* infinite "super" population with cumulants κ_r, the rth cumulant

$$k_r = [(N - v/N)^{\frac{1}{2}r}v^{-\frac{1}{2}r+1} + (-1)^r(v/N)^{\frac{1}{2}r}(N-v)^{-\frac{1}{2}r+1}]\kappa_r$$

which shows that k_r is of order $O(N^{-\frac{1}{2}r+1})$ provided we make the above assumption which, although more restrictive than that by Madow, is adequate for our

[3] This is also the assumption made by Madow [6].

purposes. We should point out that Barton and David do *not* make the above assumption since they use the above formula only as a step to proving another formula which does not involve the κ_r (see their equation (1)). We now split the sum in (4.1.11) into two parts, viz., the central part where v/N is satisfying the condition mentioned above and the remainder (tail) sum. It is easy to see that the latter sum is of order $O(e^{-N}N^a)$ because of the asymptotic properties of the standardized normal differentials. For the former sum, since k_r is $O(N^{-\frac{1}{2}r+1})$, it can be shown by an analysis similar to the term D involving k_3 that for each term of (4.1.7) when substituted in (4.1.11) there results a term of order smaller than $O(N^{-4})$. The inversion of the double summation in (4.1.7) and the limiting process $N \rightarrow \infty$ is not discussed here.

APPENDIX II

Verification of the order of magnitude of the probabilities $P_{ii'}$

We have from (4.1.25) that $P_{ii'}$ is given by

$$P_{ii'} \doteq \frac{(N-2)\pi_i\pi_{i'}}{(N-1)(2-\pi_i-\pi_{i'})}[1 - h^{-2} + 3h^{-4} - 2K_3h^{-3}(N-2)^{\frac{1}{2}}]$$

to $O(N^{-4})$, where the subscripts 1 and 2 are replaced by i and i' respectively. We now show that (4.1.25) is in fact correct to $O(N^{-4})$ by verifying that $\sum_{i' \neq i}^{N} P_{ii'} = \pi_i$ to an order $(N-1)O(N^{-4}) = O(N^{-3})$. Using (4.1.17) for h and (4.1.24) for K_3, the above expression for $P_{ii'}$ can be put in the form

$$P_{ii'} \doteq \frac{(N-2)\pi_i\pi_{i'}}{(N-1)(2-\pi_i-\pi_{i'})}\left[1 - \left(1 + \frac{1}{N-3} + \frac{6}{N-3}\right) \cdot \frac{\sum_1^N \pi_j^2 - \pi_i^2 - \pi_{i'}^2}{(2-\pi_i-\pi_{i'})^2} + \frac{3\left(\sum_1^N \pi_j^2 - \pi_i - \pi_{i'}\right)^2}{(2-\pi_i-\pi_{i'})^4} + \frac{1}{N-3} + \frac{3}{(N-3)^2} - \frac{4}{(N-2)^2} - \frac{2\sum_1^N \pi_j^3}{(2-\pi_i-\pi_{i'})^3} + \frac{6\sum_1^N \pi_j^2}{(N-2)(2-\pi_i-\pi_{i'}{}^2)}\right]$$

which, to $O(N^{-4})$, reduces to

$$P_{ii'} \doteq \pi_i\pi_{i'}\left[\frac{1}{(2-\pi_i-\pi_{i'})} - \frac{\sum_1^N \pi_j^2 - \pi_i^2 - \pi_{i'}^2}{(2-\pi_i-\pi_{i'})^3} + \frac{3\left(\sum_1^N \pi_j^2\right)^2}{(2-\pi_i-\pi_{i'})^5} - \frac{2\sum_1^N \pi_j^3}{(2-\pi_i-\pi_{i'})^4}\right].$$

Expanding all denominators binomially and retaining all terms to $O(N^{-4})$, we obtain

$$P_{ii'} \doteq \frac{\pi_i\pi_{i'}}{2}\left[1 + \tfrac{1}{2}(\pi_i + \pi_{i'}) + \tfrac{1}{4}(\pi_i^2 + \pi_{i'}^2 + 2\pi_i\pi_{i'})\right] - \frac{\pi_i\pi_{i'}}{8}\left(\sum_1^N \pi_j^2\right)\left[1 + \tfrac{3}{2}(\pi_i + \pi_{i'})\right] + \tfrac{1}{8}(\pi_i^3\pi_{i'} + \pi_i\pi_{i'}^3) + \frac{3\pi_i\pi_{i'}}{32}\left(\sum_1^N \pi_j^2\right)^2 - \frac{\pi_i\pi_{i'}}{8}\left(\sum_1^N \pi_j^3\right).$$

Summing the above expression for $P_{ii'}$ over i' from 1 to N except $i' = i$ and noting that $\sum_1^N \pi_j = 2$, we obtain

$$\sum_{i'\neq i}^{N} P_{ii'} \doteq \tfrac{1}{2}\pi_i(2 - \pi_i) + \tfrac{1}{4}\pi_i^2(2 - \pi_i) + \tfrac{1}{4}\pi_i\left(\sum_1^N \pi_j^2 - \pi_i^2\right) + \tfrac{1}{2}\pi_i^3 - \tfrac{1}{8}\pi_i^2\sum_1^N \pi_j^2 - \tfrac{1}{8}\pi_i(2 - \pi_i)\sum_1^N \pi_j^2$$

which, to $O(N^{-3})$, reduces to π_i thereby providing the announced check.

REFERENCES

[1] Barton, D. E. and David, F. N. (1961). The central sampling moments of the mean in samples from a finite population (Aty's formulae and Madow's central limit). *Biometrika* **48** 199–201.

[2] Cochran, W. G. (1953). *Sampling Techniques*. Wiley, New York.

[3] Des Raj (1956). A note on the determination of optimum probabilities in sampling without replacement. *Sankhyā* **17** 197–200.

[4] Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Stat. Assn.* **47** 663–685.

[5] Kendall, M. G. and Stuart, A. (1958). *The Advanced Theory of Statistics*, Rev. ed. Hafner, New York.

[6] Madow, W. G. (1948). On the limiting distributions of estimates based on samples from finite universes. *Ann. Math. Statist.* **19** 535–545.

[7] Narain, R. D. (1951). On sampling without replacement with varying probabilities. *J. Ind. Soc. Agric. Statist.* **3** 169–174.

[8] Wishart, John (1952). Moment coefficients of the k-statistics in samples from a finite population. *Biometrika* **39** 1–13.

[9] Yates, F. and Grundy, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *J. Roy. Statist. Soc., Ser. B* **15** 235–261.