Biometrika (1967), 54, 3 *and* 4, *p.* 499
Printed in Great Britain

On sampling without replacement with unequal probabilities of selection

BY M. R. SAMPFORD*
A.R.C. Unit of Statistics, University of Edinburgh

SUMMARY

A sample of n different units is to be drawn from a population or stratum in such a way that unit i has probability np_i, assumed less than 1, of appearing in the sample. A mathematical solution of this problem is given by a formula from which the required probability of selection of any possible sample can be calculated: this formula is an extension of one, due to Durbin, for $n = 2$. The required np_i can be achieved in practice in three ways: (a) by evaluating the required probabilities for all possible samples, and selecting one; (b) selecting units without replacement, with probabilities of selection that must be recalculated after each drawing; and (c) by selecting up to n units with replacement, the first drawing being made with probabilities p_i, and all subsequent ones with probabilities proportional to $p_i/(1-np_i)$, and rejecting completely any sample that does not contain n different units. Method (c) seems likely to be the most convenient in practice. The probability of the simultaneous appearance in the sample of any pair of units is relatively easily calculated, so that unbiased variance estimates can be obtained without undue labour.

1. INTRODUCTION

A sample of size n from a population or stratum of N units is to be used to estimate the population total of a quantity z, where for unit i

$$z_i = \alpha_i y_i.$$

Generally, α_i will be in some sense a measure of the size of the unit, and is here assumed known and positive for all units of the population: for subsequent use, define

$$A = \sum_{i=1}^{N} \alpha_i, \quad p_i = \alpha_i/A. \tag{1}$$

However, y_i and z_i are generally unknown, and are to be determined or estimated for the units included in the sample. For example: α_i may be the number of secondary units in primary unit i, y_i the mean of y over these secondary units, or an unbiased estimate of that mean; α_i the area of a plantation, y_i the yield per acre; α_i the payroll of a factory, y_i the productivity per head. Alternatively α_i may be an indirect measure of size, chosen as expected to be positively correlated with z_i: for example, α_i may be an eye estimate of crop yield, z_i an objective estimate; α_i the crop yield of plantation i in 1966, z_i the pre-harvest estimate, based on crop samples, of 1967 yield; α_i the population of a village at the last census, z_i the present population. The population total of z is then a weighted total of the y_i, say

$$\eta = \sum_{i=1}^{N} \alpha_i y_i. \tag{2}$$

* Present address: Department of Computational and Statistical Science, University of Liverpool.

In this situation y commonly has a much smaller coefficient of variation than z, with the result that, if units are selected with equal probabilities, the obvious estimator

$$N \times \text{sample mean of } z,$$

though unbiased, has a relatively large variance: on the other hand,

$$Y_{uw} = A\bar{y}, \tag{3}$$

a multiple of the unweighted sample mean of the y_i (e.g., total area $\times$ unweighted mean of yield/unit area), has a relatively small variance, but may be seriously biased unless α and y are effectively uncorrelated in the population. The well-known method of Hansen & Hurwitz (1943), in which Y_{uw} is made unbiased by sampling *with* replacement, with probabilities of selection p_i, is reasonably efficient for small sampling fractions, but not for large sampling fractions such as may be required in small strata. It is also well known that Y_{uw} can be made unbiased for sampling *without* replacement, provided the probability π_i that unit i is included in the sample is proportional to α_i; this is most easily seen by introducing a variable r_i which takes the value 1 if unit i is in the sample, and 0 otherwise (e.g. Cochran, 1963, pp. 27–9; Cornfield, 1944). Then, since the sample size is fixed, $\Sigma r_i = n$, so that if $\pi_i = k\alpha_i$,

$$E(r_i) = \pi_i = k\alpha_i, \quad k\sum_{i=1}^{N}\alpha_i = E\left(\sum_{i=1}^{N} r_i\right) = n,$$

whence

$$k = n\Big/\sum_{i=1}^{N}\alpha_i = \frac{n}{A}, \quad \pi_i = \frac{n\alpha_i}{A} = np_i, \quad E(\bar{y}) = \frac{1}{n}E\left(\sum_{i=1}^{N} r_i y_i\right) = \frac{1}{n}\sum_{i=1}^{N}\left(\frac{n\alpha_i}{A}\right)y_i = \frac{\eta}{A}.$$

For any method of selection providing the required probabilities π_i, the sampling variance of Y_{uw} is given by

$$V(Y_{uw}) = (A/n)^2 \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(\pi_i\pi_j - \pi_{ij})(y_i - y_j)^2, \tag{4}$$

where π_{ij} is the probability that units i and j both occur in the sample: this variance is estimated without bias by

$$v(Y_{uw}) = \left(\frac{A}{n}\right)^2 \underset{i}{S}\,\underset{j>i}{S}\left(\frac{\pi_i\pi_j}{\pi_{ij}} - 1\right)(y_i - y_j)^2, \tag{5}$$

where summation is over the $\frac{1}{2}n\,(n-1)$ pairs of units in the sample, provided $\pi_{ij} \neq 0$, $1 \leqslant i < j \leqslant N$; see Horwitz & Thompson (1952) and Yates & Grundy (1953). Since $\pi_{ij} = E(r_i r_j)$, the coefficient of $(y_i - y_j)^2$ in (4) is $-\operatorname{cov}(r_i, r_j)$: it is obviously desirable that this covariance should be negative for most, if not all, of the possible pairs i, j, to remove the possibility of negative variance estimates. (If the y_i are not true values, but estimates calculated from subsamples, additional terms are, of course, needed in (4) and (5).)

Various methods of selection have been advanced to provide the specified probabilities π_i; see, for example, Madow (1949), Grundy (1954), Fellegi (1963), Hartley & Rao (1962). With most of these methods the problem of determining the π_{ij} is almost impossibly severe for samples of size greater than 2, with the result that approximate formulae for error must be used. Durbin (1967) has shown that if, for samples of 2, the first unit is drawn with probability p_i, and the second from the remainder of the population with

$$\begin{aligned}\text{pr (selection of } j|i \text{ already selected)} &= p_{j.i} \\ &= p_j\left(\frac{1}{1-2p_i} + \frac{1}{1-2p_j}\right)\Big/\left(1 + \sum_{i=1}^{N}\frac{p_i}{1-2p_i}\right),\end{aligned} \tag{6}$$

the required probabilities, $2p_i$, of inclusion are attained, and

$$\pi_{ij} = 2p_i p_{j.i} = 2p_j p_{i.j}. \tag{7}$$

This paper shows how Durbin's result may be extended to samples of size greater than two, while still permitting the relatively easy evaluation of the π_{ij}.

2. Notation

A sample of size n is to be drawn without replacement from a population of size N in such a way that the probability of inclusion π_i of unit i is proportional to p_i, where $p_1 + \ldots + p_N = 1$.

Since $\pi_i = np_i$ this requirement can be satisfied only if $np_i \leqslant 1$ for all units, and since if $np_i = 1$, unit i is automatically included, and the problem reduces to the selection of the remaining $n-1$ units from the rest of the population, it may be assumed here that $np_i < 1$. Define, for each unit,

$$\lambda_i = p_i/(1-np_i). \tag{8}$$

Denote by $S(m)$ a set of $m \leqslant N$ different units, $i_1, i_2, \ldots, i_m$, and define L_m by

$$L_0 = 1, \quad L_m = \sum_{S(m)} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_m} \quad (1 \leqslant m \leqslant N), \tag{9}$$

where summation is over all possible sets of m units drawn from the population. Further $L_m(\bar{\imath})$ and $L_m(\bar{\imath}\bar{\jmath})$ are defined similarly, but relate to the subpopulations obtained by omitting, respectively, unit i, and units i and j, from the population. Routines for calculating these functions are discussed in § 5.

3. The method

The method consists essentially of selecting the particular sample $S(n)$, consisting of units $i_1, i_2, \ldots, i_n$, with probability

$$P\{S(n)\} = K_n \sum_{u=1}^{n} p_{i_u} \prod_{\substack{v=1 \\ v \neq u}}^{n} \lambda_{i_v}$$

or, substituting $p_{i_u} = \lambda_{i_u}(1-np_{i_u})$,

$$P\{S(n)\} = nK_n \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_n} \left(1 - \sum_{u=1}^{n} p_{i_u}\right), \tag{10}$$

where K_n is a constant multiplier; it will be shown in the next section that

$$K_n = \left(\sum_{t=1}^{n} \frac{tL_{n-t}}{n^t}\right)^{-1}. \tag{11}$$

When $n = 2$, (10) reduces to

$$P(i,j) = K_2 p_i p_j \left(\frac{1}{1-2p_i} + \frac{1}{1-2p_j}\right),$$

which is equivalent to Durbin's formula (7); when $n = 3$,

$$P(i,j,k) = K_3 p_i p_j p_k \left\{\frac{1}{(1-3p_i)(1-3p_j)} + \frac{1}{(1-3p_i)(1-3p_k)} + \frac{1}{(1-3p_j)(1-3p_k)}\right\}.$$

The probabilities (10) can be achieved in practice by at least three methods.

(*a*) For small strata, particularly if the sample is being drawn by computer, the probabilities can be evaluated for the set of all possible samples, and a drawing of one sample made from this set with the required probabilities. For even moderately large strata this is impracticable.

(b) Alternatively, units may be selected without replacement, with probabilities recalculated at each drawing, according to the rule illustrated, for $n = 4$, by

1st drawing: the probability of selection of unit i is

$$q_i = c_1\lambda_i \sum_{k\neq i} \sum_{\substack{j<k\\ j\neq i}} \lambda_j\lambda_k(1-p_i-p_j-p_k)\,; \tag{12a}$$

2nd drawing: the probability of selection of unit j, if unit i is already selected, is

$$q_{j.i} = c_2\lambda_j \sum_{k\neq i,j} \lambda_k(1-p_i-p_j-p_k); \tag{12b}$$

3rd drawing: the probability of selection of unit k, if units i and j are already selected, is

$$q_{k.ij} = c_3\lambda_k(1-p_i-p_j-p_k); \tag{12c}$$

4th drawing: the probability of selection of unit l, if units i, j and k are already selected, is

$$q_{l.ijk} = c_4 p_l. \tag{12d}$$

It is easily verified by equating the sum of the q's at each stage to 1 that

$$q_{k.ij} = c_3\lambda_k c_4^{-1}, \quad q_{j.i} = c_2\lambda_j c_3^{-1}, \quad q_i = \tfrac{1}{2}c_1\lambda_i c_2^{-1},$$

so that the probability of drawing units i, j, k, l *in that order* is $\frac{1}{2}c_1\lambda_i\lambda_j\lambda_k p_l$, that of drawing i, j, k, in any order, followed by l, is $3c_1\lambda_i\lambda_j\lambda_k p_l$, and hence that of drawing i, j, k, l in any order is

$$3c_1\lambda_i\lambda_j\lambda_k\lambda_l\left(\frac{p_i}{\lambda_i}+\frac{p_j}{\lambda_j}+\frac{p_k}{\lambda_k}+\frac{p_l}{\lambda_l}\right).$$

Comparison with (10) shows that this is the required probability, and that $c_1 = \frac{1}{3}K_4$; generally, it is $K_n/(n-1)$.

This method is a generalization of one due to Brewer (1963) and is suitable for small populations or strata, particularly with large sampling fractions. I am indebted to Professor Durbin for suggesting it as an alternative to my own proposed method, which follows.

(c) For larger populations the evaluation of the successive probabilities in (b) may be rather tedious, and a rejective method, in which sampling is initially with replacement, is preferable. This method depends on the fact that

$$P\{S(n)\} = \frac{K_n\left(\sum_{i=1}^{N}\lambda_i\right)^{n-1}}{(n-1)!} \times \text{coefficient of } t_{i_1}\ldots t_{i_n} \text{ in } W(t),$$

where

$$W(t) = \left(\sum_{i=1}^{N} p_i t_i\right)\left(\sum_{i=1}^{N}\lambda_i t_i \Big/ \sum_{i=1}^{N}\lambda_i\right)^{n-1}. \tag{13}$$

Now $W(t)$ is the probability generating function for a sampling procedure in which units are selected *with* replacement, the first or other single specified drawing being made with probabilities p_i, and all others with probabilities

$$\left(\lambda_i \Big/ \sum_{i=1}^{N}\lambda_i\right):$$

the probability of a sample in which unit i occurs r_i times ($i = 1, 2, \ldots, N$) is then the coefficient in $W(t)$ of $t_1^{r_1}\ldots t_N^{r_N}$. Since by this scheme the probabilities of all samples without repeated units are proportional to those required by formula (10), these latter probabilities may be achieved by drawing a sample according to the 'with replacement' scheme, and accepting it if it contains n different units, otherwise discarding it and repeating the process until an acceptable sample is obtained. The probability of obtaining an acceptable sample at any attempt is then

$$(n-1)!\Big/\left\{K_n\left(\sum_{i=1}^{N}\lambda_i\right)^{n-1}\right\} \tag{14}$$

and the expected number of samples that must be drawn to obtain an acceptable one is the reciprocal of this quantity. In practice, of course, a sample can be discarded as soon as a duplicate unit is drawn.

Hájek (1964) has considered in some detail the problem of rejective sampling with the same selection probabilities at each drawing. He gives a rather elaborate formula for calculating selection probabilities, by the use of which a very close approximation to the required π_i can be obtained, and some simplified formulae, that give less good, but still usually adequate, approximations. These formulae resemble that for λ, with minor modifications reflecting the fact that they are designed to yield selection probabilities for use at all drawings.

It will be shown in the next section that method (c), with K_n given by (11), does in fact give the required π_i, and that the probability that the sample contains both units i and j is

$$\pi_{ij} = K_n\lambda_i\lambda_j\sum_{t=2}^{n}\frac{\{t-n(p_i+p_j)\}\,L_{n-t}(\bar{i}\,\bar{j})}{n^{t-2}}. \tag{15}$$

For a particular pair (i,j), π_{ij} is easily calculated. Since π_{ij} must be calculated for all pairs of units occurring in the sample, the computations may become somewhat tedious for large samples; however they are easily programmed for computer evaluation. Some notes on the calculations are given in § 5.

4. Theoretical results

4.1. Lemma. *For p_i, λ_i defined as in § 2, define*

$$G(n,r;k) = \sum_{\substack{S(n-r)\\ k\notin S}}\lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_{n-r}}\left(1-rp_k-\sum_{u=1}^{n-r}p_{i_u}\right), \tag{16}$$

where summation is over all sets of $(n-r)$ units not containing unit k. Then

$$G(n,r;k) = (1-np_k)\sum_{t=r}^{n}\frac{tL_{n-t}}{n^{t-r+1}}. \tag{17}$$

Proof. We have that

$$1-rp_k-\sum_{u=1}^{n-r}p_{i_u} = \frac{1}{n}\left\{r(1-np_k)+\sum_{u=1}^{n-r}(1-np_{i_u})\right\},$$

whence

$$G(n,r;k) = (1/n)\{r(1-np_k)L_{n-r}(\bar{k})+H(n,r;k)\}, \tag{18}$$

where

$$H(n,r;k) = \sum_{\substack{S(n-r)\\ k\notin S}}\lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_{n-r}}\sum_{u=1}^{n-r}(1-np_{i_u}).$$

Since $\lambda_j(1-np_j) = p_j$, each term in $H(n,r;k)$ consists of a product of $(n-r-1)$ λ's and one p, say p_t, where $t \neq k$, and λ_t does not appear in the product: for a particular product of

λ's there are $(N-n+r)$ such terms, one for each of the possible values of t. Hence, collecting the p_t and summing over all such products, we have

$$H(n,r;k) = \sum_{\substack{S(n-r-1)\\ k\notin S}} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_{n-r-1}}\left(1-p_k-\sum_{u=1}^{n-r-1} p_{i_u}\right)$$
$$= G(n,r+1;k)+rp_k L_{n-r-1}(\bar{k}).$$

Also
$$L_m(\bar{k}) = L_m - \lambda_k L_{m-1}(\bar{k}): \tag{19}$$

on substituting in (18), with $m = n-r$, the terms in $L_{n-r-1}(\bar{k})$ cancel, and

$$G(n,r;k) = (1/n)\{r(1-np_k)L_{n-r}+G(n,r+1;k)\}. \tag{20}$$

Since, from (16),
$$G(n,n;k) = L_0(\bar{k})(1-np_k) = 1-np_k = (1-np_k)L_0,$$

the result (17) follows by repeated application of (20), and the Lemma is proved.

Although in practice np_k must be less than 1, in one subsequent application of the Lemma the possibility $np_k \nless 1$, but $np_i < 1$, $i \neq k$, must be considered. The preceding proof holds if $np_k > 1$: if $np_k = 1$ the Lemma must be stated in the modified form

$$G(n,r;k) = p_k \sum_{t=r}^{n-1} \frac{tL_{n-t-1}(\bar{k})}{n^{t-r-1}}. \tag{17a}$$

For proof, observe first that $G(n,r;k)$ is validly defined by (16), which does not contain λ_k: then from (18) and the subsequent result on $H(n,r;k)$,

$$G(n,r;k) = (1/n)\{G(n,r+1;k)+rp_k L_{n-r-1}(\bar{k})\}, \tag{21}$$

and since $G(n,n;k) = 1-np_k = 0$, the result follows.

This Lemma is a special case of the result, which can be proved by a similar argument to that used above,

$$\sum_{\substack{S(m-r)\\ k\notin S}} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_{m-r}}\left(1-rp_k-\sum_{u=1}^{m-r} p_{i_u}\right) = (1-np_k)\sum_{t=r}^{m}\frac{tL_{m-t}}{n^{t-r+1}}+(n-m)\sum_{t=r}^{m}\frac{L_{m-t}(\bar{k})}{n^{t-r+1}}. \tag{22}$$

This relationship gives an alternative, but not an appreciably simpler, derivation of the π_{ij}: for the purposes of this paper, at least, it does not seem to have any other useful applications.

4·2. *Proof that the method yields the required* π_k. From (10), summing over all sets $S(n)$ containing unit k, we have that

$$\pi_k = nK_n \sum_{\substack{S(n)\\ k\in S}} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_n}\left(1-\sum_{u=1}^{n} p_{i_u}\right) \tag{23}$$

$$= nK_n\lambda_k G(n,1;k) \tag{24}$$

$$= np_k\left(K_n\sum_{t=1}^{n}\frac{tL_{n-t}}{n^t}\right), \tag{25}$$

from (17). Hence, when K_n is defined by (11), $\pi_k = np_k$ as required.

Formula (17) for $G(n,r;k)$ does not depend explicitly on the population size N. This property is useful as well as interesting, in that it permits us to modify the population without affecting G, providing due account is taken of modifications to the L_m. In particular, we can augment the population with a 'phantom unit' ϕ for which $p_\phi = 0$. Summation over all possible samples of n of the original population then becomes equivalent to summation over all samples of n not containing ϕ: hence from formula (10) we can write

$$\sum_{S(n)} P\{S(n)\} = nK_n G(n, 0; \phi),$$

and setting this equal to 1, formula (11) follows directly from (17). This manoeuvre has perhaps only curiosity value: a more useful modification of the population is illustrated in the next subsection.

4·3. *Derivation of formula for* π_{kl}. From (10), summing over all sets $S(n)$ containing units k and l,

$$\pi_{kl} = nK_n \sum_{\substack{S(n)\\ k,l \in S}} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_n}\left(1-\sum_{u=1}^{n} p_{i_u}\right) \tag{26}$$

$$= K_n \lambda_k \lambda_l \Phi_{kl},$$

where

$$\Phi_{kl} = n \sum_{\substack{S(n-2)\\ k,l \notin S}} \lambda_{j_1}\lambda_{j_2}\ldots\lambda_{j_{n-2}}\left\{1-(p_k+p_l)-\sum_{u=1}^{n-2} p_{j_u}\right\}. \tag{27}$$

Now imagine a population consisting of all units of the original population except units k and l, together with a new unit α for which $p_\alpha = p_k + p_l$, $\lambda_\alpha = p_\alpha/(1-np_\alpha)$, and denote by G', L' the functions G and L defined over this modified population. Then

$$\Phi_{kl} = n \sum_{\substack{S(n-2)\\ \alpha \notin S}} \lambda_{j_1}\lambda_{j_2}\ldots\lambda_{j_{n-2}}\left(1-p_\alpha-\sum_{u=1}^{n-2} p_{j_u}\right)$$

$$= nG'(n, 2; \alpha) + np_\alpha L'_{(n-2)}(\bar{\alpha}). \tag{28}$$

Now by (19)

$$L'_{n-t} = L'_{n-t}(\bar{\alpha}) + \lambda_\alpha L'_{n-t-1}(\bar{\alpha}) \quad (t < n-1)$$

and $L'_0 = L'_0(\bar{\alpha}) = 1$; substituting from (17) and using these results, we have on some reduction

$$\Phi_{kl} = (1-np_\alpha)\left\{\sum_{t=2}^{n} \frac{tL'_{n-t}(\bar{\alpha})}{n^{t-2}} + \sum_{t=2}^{n-1} \frac{t\lambda_\alpha L'_{n-t-1}(\bar{\alpha})}{n^{t-2}}\right\} + np_\alpha L'_{n-2}(\bar{\alpha}) \tag{29}$$

$$= \sum_{t=2}^{n} \frac{(t-np_\alpha) L'_{n-t}(\bar{\alpha})}{n^{t-2}}. \tag{30}$$

(If, as is possible, $np_\alpha = 1$, (30) follows directly by substituting from (17a) into (28).) Reverting to the original population, and writing $p_\alpha = p_k + p_l$, $L'_{n-t}(\bar{\alpha}) = L_{n-t}(\bar{k}\,\bar{l})$, we obtain

$$\pi_{kl} = K_n \lambda_k \lambda_l \Phi_{kl} = K_n \lambda_k \lambda_l \sum_{t=2}^{n} \frac{\{t-n(p_k+p_l)\} L_{n-t}(\bar{k}\,\bar{l})}{n^{t-2}},$$

which is formula (15).

4·4. *Proof that* $\operatorname{cov}(r_k, r_l) < 0$. Now

$$\pi_{kl} = \lambda_k \lambda_l K_n \Phi_{kl}$$

$$= p_k p_l \Phi_{kl}/\{(1-np_k)(1-np_l)K_n^{-1}\}. \tag{31}$$

Next

$$K_n^{-1} = \sum_{t=1}^{n} \frac{tL_{n-t}}{n^t}$$

and

$$L_{n-t} = L_{n-t}(\bar{k}\,\bar{l}) + (\lambda_k + \lambda_l) L_{n-t-1}(\bar{k}\,\bar{l}) + \lambda_k \lambda_l L_{n-t-2}(\bar{k}\,\bar{l}), \tag{32}$$

whence, writing for brevity $\bar{L}_m$ in place of $L_m(\bar{k}\,\bar{l})$,

$$(1-np_k)(1-np_l)L_{n-t} = (1-np_k)(1-np_l)\bar{L}_{n-t} + \{(p_k+p_l)-2np_kp_l\}\bar{L}_{n-t-1}+p_kp_l\bar{L}_{n-t-2}, \tag{33}$$

where $\bar{L}_m = 0$ if $m < 0$. Some manipulation now shows that $(1-np_k)(1-np_l)K_n^{-1} > \Phi_{kl}/n^2$. Hence, from (31), $\pi_{kl} < n^2p_kp_l = \pi_k\pi_l$, $\pi_{kl}-\pi_k\pi_l < 0$.

4·5. *Alternative versions of probability formulae for sampling without replacement.* Formulae (12*a*)–(12*d*), and their equivalents for general n, can be expressed in terms of the functions L, $L(\bar{\imath})$, etc. Since some at least of these functions must be evaluated whichever method of selection is used, the alternative forms may be more practically useful than those given in (12).

The general version of (12*a*) is

$$q_i = c_1\lambda_i \sum_{\substack{S(n-2)\\ i\notin S}} \lambda_{j_1}\ldots\lambda_{j_{n-2}}\left(1-p_i-\sum_{u=1}^{n-2}p_{j_u}\right) \tag{34}$$

$$= c_1\lambda_i\{nG(n,1;i)-(1-np_i)L_{n-1}(\bar{\imath})\}$$

(from (18))

$$= c_1\{n\lambda_iG(n,1;i)-p_iL_{n-1}(\bar{\imath})\}, \tag{35}$$

and since (from (24), writing $\pi_i = np_i$),

$$\lambda_iG(n,1;i) = p_iK_n^{-1},$$

it follows that

$$q_i = c_1p_i\{nK_n^{-1}-L_{n-1}(\bar{\imath})\}. \tag{36}$$

If the L_m that appear in the formula for K_n^{-1} are replaced, using (19), by $L_m(\bar{\imath})$, formula (36) can be expressed as

$$q_i = c_1\lambda_i\sum_{t=2}^{n}\frac{(t-np_i)L_{n-t}(\bar{\imath})}{n^{t-1}}; \tag{37}$$

this form is less convenient for calculation than (36), but its relevance will become apparent.

The general version of (12*b*) is

$$q_{j.i} = c_2\lambda_j\sum_{\substack{S(n-3)\\ i,j\notin S}}\lambda_{k_1}\ldots\lambda_{k_{n-3}}\left\{1-(p_i+p_j)-\sum_{u=1}^{n-3}p_{k_u}\right\} \tag{38}$$

$$= c_2\lambda_j\{nG'(n,2;\alpha)-2(1-np_\alpha)L'_{n-2}(\bar{\alpha})\},$$

from (18), where α, G' refer again to the imaginary population of § 4·3. Hence, from (28) and (30)

$$q_{j.i} = c_2\lambda_j\{\Phi_{kl}-(2-np_\alpha)L'_{n-2}(\bar{\alpha})\}$$

$$= c_2\lambda_j\sum_{t=3}^{n}\frac{\{t-n(p_i+p_j)\}L_{n-t}(\bar{\imath}\,\bar{\jmath})}{n^{t-2}}. \tag{39}$$

The $q_{j.i}$ must be evaluated for all j other than i, and so the use of formula (39) will involve some additional calculation beyond that for variance estimation, which requires the $L(\bar{\imath}\,\bar{\jmath})$ only for pairs i, j in the sample. Comparison of (37) and (39) indicates immediately the general form for these probabilities: however, the additional calculation required for sample sizes greater than 4, or possibly 5, is likely to outweigh the convenience of sampling without subsequent rejection, so further formulae in the series are unlikely to be of great practical use.

5. Some notes on computation

The major part of the computational load consists of the evaluation of the sums of products L_m and $L_m(\bar{i}\,\bar{j})$ occurring in formulae (11) and (15), and perhaps of the $L_m(\bar{i})$ in formula (36). These may be derived systematically in two ways. The easier method, particularly for use with desk calculators, is to start from

$$L_0 = L_0(\bar{i}) = L_0(\bar{i}\,\bar{j}) = 1,$$

and determine the higher terms successively, using the relations

$$L_m = \frac{1}{m}\sum_{r=1}^{m}(-1)^{r-1}R_r L_{m-r}, \tag{40}$$

where $R_r = \lambda_1^r + \ldots + \lambda_N^r$ and

$$L_m(\bar{i}\,\bar{j}) = L_m - (\lambda_i + \lambda_j) L_{m-1}(\bar{i}\,\bar{j}) - \lambda_i\lambda_j L_{m-2}(\bar{i}\,\bar{j}). \tag{41}$$

The $L_m(\bar{i})$, if required, may be calculated from formula (19), and if this is done, the obvious extension of (19),

$$L_m(\bar{i}\,\bar{j}) = L_m(\bar{i}) - \lambda_j L_{m-1}(\bar{i}\,\bar{j}) \tag{42}$$

can be used instead of (41). The computations involved in using (40) and (41) are illustrated in Part 1 of Table 2 for a sample of 5 units (1 3 4 6 7) from the population of 10 defined in Table 1.

Table 1. *Example: details of imaginary population and preliminary calculations*

Unit i	α_i	y_i	p_i	$\sum_1^i p_j$	$1-5p_i$	$\lambda_i = \dfrac{p_i}{1-5p_i}$	$\phi_i = \dfrac{\lambda_i}{\Sigma\lambda_i}$	$\sum_1^i \phi_j$	λ_i^2
1	18	16	0·18	0·18	0·10	1·8000	0·4945	0·4945	3·2400
2	14	14	0·14	0·32	0·30	0·4667	0·1282	0·6227	0·2178
3	13	18	0·13	0·45	0·35	0·3714	0·1020	0·7247	0·1379
4	11	11	0·11	0·56	0·45	0·2444	0·0672	0·7919	0·0597
5	10	10	0·10	0·66	0·50	0·2000	0·0549	0·8468	0·0400
6	10	17	0·10	0·76	0·50	0·2000	0·0549	0·9017	0·0400
7	8	18	0·08	0·84	0·60	0·1333	0·0366	0·9383	0·0178
8	7	13	0·07	0·91	0·65	0·1077	0·0296	0·9679	0·0116
9	5	10	0·05	0·96	0·75	0·0667	0·0183	0·9862	0·0044
10	4	12	0·04	1·00	0·80	0·0500	0·0137	0·9999	0·0025
Total	100	—	1·00	—	5·00	3·6402	0·9999	—	3·7717

A possible objection to this approach is that, since individual terms in (40) may appreciably exceed the final result in magnitude, the uncritical use of (40) in computer programs may result in substantial round-off errors; also, errors accumulate in the repeated use of (41) or (42). These dangers are largely avoided by the use of the tabular method illustrated (again using the population of Table 1) in Part 2 of Table 2. For sample size n, values $L_0, L_1, \ldots, L_{n-1}$ are required: the table therefore has n columns, numbered 0 to $n-1$, and $N+1$ rows, numbered 0 to N. Column 0 has 1 in row 0 and the λ's for all the units of the population in the remaining rows: the precise order of the λ's is immaterial, but if the sample has already been drawn by the rejective method (*c*) it is convenient to enter the λ's corresponding to the sampled units in rows $1-n$. Denote by $\lambda_{(i)}$ the entry in row i. Then the entries I_{1j} in column 1 are cumulative sums of the λ's, from the bottom, shifted up by one row: thus

$$I_{1j} = \sum_{i=j+1}^{N} \lambda_{(i)} \quad (j = 0, 1, \ldots, N-1).$$

The entries in each subsequent column are cumulative sums of products of the entries in the previous column and the corresponding $\lambda_{(i)}$, again from the bottom and shifted up by one row: thus

$$I_{kj} = \sum_{i=j+1}^{N-k+1} \lambda_{(i)} I_{(k-1)i} \quad (j = 0, 1, \ldots, N-k).$$

For example, in Table 2

$$I_{18} = \lambda_{(10)} + \lambda_{(9)} = 0{\cdot}0500 + 0{\cdot}0667 = 0{\cdot}1167$$

and

$$\begin{aligned} I_{27} &= (\lambda_{(9)} \times I_{19}) + (\lambda_{(8)} \times I_{18}) \\ &= (0{\cdot}0667 \times 0{\cdot}0500) + (0{\cdot}1077 \times 0{\cdot}1167) = 0{\cdot}0159. \end{aligned}$$

The entry in row 0 and column m is then L_m, and the entry in row 2 and column m is $L_m(\bar{k}\,\bar{l})$, if λ_k and λ_l appear in rows 1 and 2 of column 0. To obtain each of the remaining $L_m(\bar{i}\,\bar{j})$ some recalculation is necessary: provided that the sampled units are at the top of the table, at most rows $2-(n-1)$ need be recalculated. Some of these recalculations, together with the results of the rest, are shown in Part 2 of Table 2. Each recalculation is preceded by a bracket showing the revised order of units in the rows 1–5 of the table. For example, any arrangement having λ_1 and λ_6 in rows 1 and 2 of column 0 can be used to obtain $L_m(\bar{1}\,\bar{6})$. The order [1 6 3 4 7] differs as little as possible from the original order [1 3 4 6 7]: since the λ_i in

Table 2. *Calculation of L_m and $L_m(\bar{i}\,\bar{j})$ for the sample* [1 3 4 6 7] *from the population of Table* 1

Part 1

(a) Calculation of L_m from formula (40)

r	R_r	L_r	
0	—	1	
1	3·6402	3·6402	$(= R_1)$
2	3·7717	4·7397	$(= R_1L_1 - R_2)$
3	6·0195	3·1811	$(= R_1L_2 - R_2L_1 + R_3)$
4	10·5713	1·2610	$(= R_1L_3 - R_2L_2 + R_3L_1 - R_4)$

(b) Tabulations for calculation of $L_m(\bar{i}\,\bar{j})$ from formula (41)

i, j	$\lambda_i + \lambda_j$	$\lambda_i\lambda_j$	$L_0(\bar{i}\,\bar{j})$	$L_1(\bar{i}\,\bar{j})$	$L_2(\bar{i}\,\bar{j})$	$L_3(\bar{i}\,\bar{j})$
1, 3	2·1714	0·6685	1	1·4688	0·8818	0·2845
1, 4	2·0444	0·4399	1	1·5958	1·0373	0·3585
1, 6	2·0000	0·3600	1	1·6402	1·0993	0·3920
1, 7	1·9333	0·2399	1	1·7069	1·1999	0·4518
3, 4	0·6158	0·0908	1	3·0244	2·7865	1·1906
3, 6	0·5714	0·0743	1	3·0688	2·9119	1·2892
3, 7	0·5047	0·0495	1	3·1355	3·1077	1·4574
4, 6	0·4444	0·0489	1	3·1958	3·2706	1·5714
4, 7	0·3777	0·0326	1	3·2625	3·4749	1·7623
6, 7	0·3333	0·0267	1	3·3069	3·6108	1·8893

Table 2 (*cont.*)

PART 2

(*a*) *Calculation of* L_m *and* $L_m(\bar{1}\,\bar{3})$ *by tabular method*

(*For details of calculations, see* §5)

Row	(Unit)	Column (0)	(1)	(2)	(3)	(4)
0	—	1	3·6402	4·7396	3·1809	1·2608
1	1	1·8000	1·8402	1·4273	0·6118	—
2	3	0·3714	1·4688	0·8818	0·2843	—
3	4	0·2444	1·2244	0·5825	0·1419	—
4	6	0·2000	1·0244	0·3776	0·0664	—
5	7	0·1333	0·8911	0·2589	0·0319	—
6	2	0·4667	0·4244	0·0608	0·0035	—
7	5	0·2000	0·2244	0·0159	0·0004	—
8	8	0·1077	0·1167	0·0033	—	—
9	9	0·0667	0·0500	—	—	—
10	10	0·0500	—	—	—	—

(*b*) *Recalculations of rows 2–4 for the remaining* $L_m(\bar{\imath}\,\bar{\jmath})$

[1 4 3 6 7]						
2	—	—	1·5958	1·0372	0·3582	$= L_m(\bar{1}\,\bar{4})$
3	3	0·3714	1·2244	0·5825	0·1419	
[1 6 3 4 7]						
2	—	—	1·6402	1·0992	0·3919	$= L_m(\bar{1}\,\bar{6})$
3	3	0·3714	1·2688	0·6280	0·1587	
4	4	0·2444	1·0244	0·3776	0·0664	
[1 7 3 4 6]						
2	—	—	1·7069	1·1998	0·4519	$= L_m(\bar{1}\,\bar{7})$
3	3	0·3714	1·3355	0·7038	0·1905	
4	4	0·2444	1·0911	0·4371	0·0837	
5	6	0·2000	0·8911	0·2589	0·0319	
[3 4 6 7 1]						
2	—	—	3·0244	2·7865	1·1905	$= L_m(\bar{3}\,\bar{4})$
3	6	0·2000	2·8244	2·2216	0·7462	
4	7	0·1333	2·6911	1·8629	0·4979	
5	1	1·8000	0·8911	0·2589	0·0319	
[3 6 4 7 1]	—	—	3·0688	2·9119	1·2892	$= L_m(\bar{3}\,\bar{6})$
[3 7 4 6 1]	—	—	3·1355	3·1077	1·4573	$= L_m(\bar{3}\,\bar{7})$
[4 6 7 3 1]	—	—	3·1958	3·2706	1·5714	$= L_m(\bar{4}\,\bar{6})$
[4 7 6 3 1]	—	—	3·2625	3·4749	1·7623	$= L_m(\bar{4}\,\bar{7})$
[6 7 4 3 1]	—	—	3·3069	3·6109	1·8894	$= L_m(\bar{6}\,\bar{7})$

rows 5–10 are unchanged, row 4 is also unchanged, except that λ_4 appears in column 0. Row 3 is recalculated from row 4; e.g.,

$$1{\cdot}0244 + 0{\cdot}2444 = 1{\cdot}2688; \quad 0{\cdot}3776 + 0{\cdot}2444 \times 1{\cdot}0244 = 0{\cdot}6280;$$

finally, row 2 is recalculated from the revised row 3, with λ_3 in column 0. Once all the $L_m(\bar{1}\,\bar{\jmath})$ have been obtained, λ_1 is put into row 5, where it remains, and subsequent recalculations start from the revised table for [3 4 6 7 1], instead of from the original table. Note that, since $L_{n-1}(\bar{\imath}\,\bar{\jmath})$ is not required, the individual cumulative totals need not be written down in column $(n-1)$ of the original table.

Minor discrepancies between the results in the two parts of Table 2 are due to rounding-off: agreement is better if one further decimal place is retained throughout the intermediate calculations. The values obtained in Part 2 are the more accurate, and are therefore used in subsequent calculations.

6. EXAMPLES

A sample of 5 units is to be selected from an imaginary population of 10, with probabilities proportional to the α_i shown in the first column of Table 1. With $n = 5$, the 'without-replacement' version of the method requires a good deal of recalculation of probabilities, so that the 'with-replacement' version, with rejection of unsatisfactory samples, is used here for the main example. (For completeness, an illustration of the 'without-replacement' method is given at the end of the section.)

Using the 'cumulative sum' technique, and drawing successive 4-digit random numbers from the left-most 4 columns of p. vi of Fisher and Yates's random number table [1963; Table XXXIII] the following samples are obtained, using the Σp_j column in Table 1 for identifying the first unit drawn, and the $\Sigma \phi_j$ column for the remainder: unacceptable samples are discarded as soon as they prove so. Unit numbers are shown in parentheses:

(i)	2519 (2);	2302 (1);	5585 (2).		
(ii)	6845 (6);	6931 (3);	3731 (1);	6642 (3).	
(iii)	3365 (3);	7632 (4);	4333 (1);	2831 (1).	
(iv)	9719 (10);	8280 (5);	0368 (1);	6516 (3);	2465 (1).
(v)	0272 (1);	7916 (4);	0475 (1).		
(vi)	4064 (3);	0627 (1);	6240 (3).		
(vii)	0098 (1);	5064 (2);	3854 (1).		
(viii)	4686 (4);	9072 (7);	6621 (3);	8705 (6);	4690 (1).

The final sample, obtained at the eighth attempt, therefore consists of units

1, 3, 4, 6, and 7.

Calculation of the L_m and the $L_m(\bar{\imath}\,\bar{\jmath})$ for this sample is shown in Table 2. Substituting in formula (11), we have

$$K_5 = \left(\sum_{t=1}^{5} \frac{tL_{5-t}}{5^t}\right)^{-1},$$
$$= \left\{\frac{1\cdot2608}{5} + \frac{2(3\cdot1809)}{5^2} + \frac{3(4\cdot7396)}{5^3} + \frac{4(3\cdot6402)}{5^4} + \frac{5(1)}{5^5}\right\}^{-1}$$
$$= 1\cdot5497.$$

Hence, from formula (13), the probability of drawing an acceptable sample is

$$4!/(1\cdot5497 \times 3\cdot6402^4) = 0\cdot0882,$$

so that the expected number of attempts required for an acceptable sample is

$$1/0\cdot0882 = 11\cdot34.$$

Also, from Table 2, Part 1 (b), or Part 2 (a),

$$L_0(\bar{1}\,\bar{3}) = 1, \quad L_1(\bar{1}\,\bar{3}) = 1\cdot4688, \quad L_2(\bar{1}\,\bar{3}) = 0\cdot8818, \quad L_3(\bar{1}\,\bar{3}) = 0\cdot2843,$$

whence, using formula (15), we have

$$\pi_{13} = K_5\lambda_1\lambda_3 \sum_{t=2}^{5} \frac{\{t-5(p_1+p_3)\}\,L_{5-t}(\bar{1}\,\bar{3})}{5^{t-2}}$$

$$= 1.0360 \sum_{t=2}^{5} \frac{(t-1\cdot 55)\,L_{5-t}(\bar{1}\,\bar{3})}{5^{t-2}}$$

$$= 0\cdot 5752,$$

$$\pi_{13}-\pi_1\pi_3 = 0\cdot 5752-(0\cdot 90)\,(0\cdot 65) = -0\cdot 0098$$

and the coefficient of $(y_1-y_3)^2$ in formula (5) is

$$(\pi_1\pi_3-\pi_{13})/\pi_{13} = 0\cdot 0098/0\cdot 5752 = 0\cdot 0170.$$

The population total estimated from the sample [1 3 4 6 7] is

$$Y_{uw} = \tfrac{100}{5}(16+18+11+17+18) = 1600.$$

The true value is $\Sigma\alpha_i y_i = 1442$. The estimated variance, using coefficients given in Table 3 is

$$\begin{aligned} v(Y_{uw}) = {} & (\tfrac{100}{5})^2\times\{0\cdot 0170\times(-2)^2+0\cdot 0228\times(5)^2+0\cdot 0258\times(-1)^2 \\ & +0\cdot 0318\times(-2)^2+0\cdot 0931\times(7)^2+0\cdot 1055\times(1)^2+0\cdot 1301\times(0)^2 \\ & +0\cdot 1426\times(-6)^2+0\cdot 1756\times(-7)^2+0\cdot 1987\times(-1)^2\} \\ = {} & 7758, \end{aligned}$$

whence s.e.$(Y_{uw}) = 88\cdot 1$. From formula (4), the true variance is 9072.

As already remarked, the quantity $\pi_i\pi_j-\pi_{ij}$, the coefficient of $(y_i-y_j)^2$ in formula (4), is minus the covariance of the random variates r_i, r_j, where $r_i = 1$ if unit i is in the sample and $= 0$ otherwise. These coefficients, together with the corresponding coefficients $(\pi_i\pi_j-\pi_{ij})/\pi_{ij}$ for the Yates–Grundy variance estimator (5), are shown in Table 3, for all pairs of units in the population.

Example of selection without replacement. Again, it is required to select a sample of 5 from the population of Table 1. From formulae (36), (40), and the obvious extensions of (40) and (12), and using the value, already obtained, of $1\cdot 5497$ for K_5, we have

$$q_i = c_1 p_i\{3\cdot 2264-L_4(\bar{i})\},$$

$$q_{j.i} = c_2\lambda_j \sum_{t=3}^{5} \frac{\{t-5(p_i+p_j)\}\,L_{5-t}(\bar{i}\,\bar{j})}{5^{t-2}},$$

$$q_{k.ij} = c_3\lambda_k \sum_{t=4}^{5} \frac{\{t-5(p_i+p_j+p_k)\}\,L_{5-t}(\bar{i}\,\bar{j}\,\bar{k})}{5^{t-3}},$$

$$q_{l.ijk} = c_4\lambda_l\{1-(p_i+p_j+p_k+p_l)\},$$

$$q_{m.ijkl} = c_5 p_m = p_m/\{1-(p_i+p_j+p_k+p_l)\},$$

where $c_1 = \frac{1}{4}K_5 = 0\cdot 3874$, and c_2, c_3, c_4 must be derived progressively from the formulae

$$c_2 = \tfrac{1}{3}c_1\lambda_1/q_i, \quad c_3 = \tfrac{1}{2}c_2\lambda_j/q_{j.i}, \quad c_4 = c_3\lambda_k/q_{k.ij}.$$

The various calculations are shown in Table 4. The $L_4(\bar{i})$ in column (1) are obtained using formula (19), with values of λ_i from Table 1 and L_{m-t} from Table 2, Part 2. Using the same

Table 3. *Values of* $(\pi_i\pi_j-\pi_{ij})$ *(above diagonal) and* $(\pi_i\pi_j-\pi_{ij})/\pi_{ij}$ *(below diagonal) for the population of Table* 1

$i \backslash j$	1	2	3	4	5	6	7	8	9	10
1	—	0·0088	0·0098	0·0110	0·0113	0·0113	0·0111	0·0106	0·0087	0·0074
2	0·0142	—	·0247	·0277	·0283	·0283	·0274	·0259	·0211	·0178
3	·0170	·0575	—	·0304	·0310	·0310	·0299	·0283	·0229	·0194
4	·0228	·0775	·0931	—	·0343	·0343	·0329	·0309	·0249	·0210
5	·0258	·0878	·1055	·1426	—	·0348	·0332	·0312	·0250	·0210
6	·0258	·0878	·1055	·1426	·1616	—	·0332	·0312	·0250	·0210
7	·0318	·1084	·1301	·1756	·1987	·1987	—	·0294	·0234	·0196
8	·0347	·1183	·1420	·1915	·2166	·2166	·2654	—	·0217	·0183
9	·0401	·1370	·1644	·2214	·2501	·2501	·3063	·3330	—	·0145
10	·0428	·1458	·1750	·2355	·2660	·2660	·3256	·3540	·4080	—

Table 4. *Calculation of successive selection probabilities for example of sampling without replacement*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
i	$L_4(i)$	q_i	$L_1(\bar{2}\ i)$	$L_2(\bar{2}\ i)$	$q_{i.2}$	$L_1(\bar{1}\ \bar{2}\ i)$	$q_{i.12}$	$q_{i.126}$	$q_{i.1267}$
1	0·1594	0·2139	1·3735	0·7862	0·2768	—	—	—	—
2	0·4860	0·1486	—	—	—	—	—	—	—
3	0·5657	0·1340	2·8021	2·2178	0·1594	1·0021	0·2257	0·2975	0·2600
4	0·7169	0·1069	2·9291	2·5426	0·1231	1·1291	0·1714	0·2044	0·2200
5	0·7867	0·0945	2·9735	2·6638	0·1073	1·1735	0·1482	0·1709	0·2000
6	0·7867	0·0945	2·9735	2·6638	0·1073	1·1735	0·1482	—	—
7	0·9127	0·0711	3·0402	2·8532	0·0794	1·2402	0·1083	0·1186	—
8	0·9688	0·0612	3·0658	2·9283	0·0671	1·2658	0·0910	0·0978	0·1400
9	1·0688	0·0418	3·1068	3·0513	0·0450	1·3068	0·0605	0·0629	0·1000
10	1·1132	0·0327	3·1235	3·1023	0·0349	1·3235	0·0468	0·0480	0·0800

sequence of random numbers as before, the first is 2519: since $q_1 = 0{\cdot}2139$, $q_1+q_2 = 0{\cdot}3625$, the first unit selected is 2, whence $c_2 = \frac{1}{3}\times 0{\cdot}3874\lambda_2/q_2 = 0{\cdot}4056$. Values of $L_1(\bar{2}\ \bar{j})$ and $L_2(\bar{2}\ \bar{j})$ are needed for all $j \neq 2$. These are calculated using formula (41), and are shown in columns (3) and (4) of the table: with $L_0(\bar{2}\ \bar{j}) = 1$, they lead to the values of $q_{j.2}$ shown in column 5. The second random number is 2302, so since $q_{1.2} > 0{\cdot}2302$, unit 1 is selected, whence $c_3 = \frac{1}{2}\times 0{\cdot}4056\lambda_1/q_{1.2} = 1{\cdot}3188$. Values of $L_0(\bar{1}\ \bar{2}\ \bar{k})$ $(= 1)$ and $L_1(\bar{1}\ \bar{2}\ \bar{k})$, shown in column (6), lead to the probabilities for the third drawing, column (7). The third random number is 5585, corresponding to unit 6, since

$$q_{3\cdot12}+q_{4\cdot12}+q_{5\cdot12} = 0{\cdot}5453, \quad 0{\cdot}5453+q_{6\cdot12} = 0{\cdot}6935 > 0{\cdot}5585;$$

hence $c_4 = 1{\cdot}3188\lambda_6/q_{6\cdot12} = 1{\cdot}7798$. Column (8) shows probabilities for the fourth drawing: the random number is 6845, leading to the selection of unit 7. Finally, probabilities for the fifth drawing are shown in column (9); the random number is 6931, leading to the selection of unit 8, completing the sample 1, 2, 6, 7, 8. The additional calculation involved in Table 4 is substantial, and the method of sampling with replacement and discarding unsatisfactory samples seems preferable here.

REFERENCES

BREWER, K. R. W. (1963). A model of systematic sampling with unequal probabilities. *Aust. J Statist.* **5**, 5–13.

COCHRAN, W. G. (1963). *Sampling Techniques*, 2nd edition. New York: Wiley.

CORNFIELD, J. (1944). On samples from finite populations. *J. Am. Statist. Ass.* **39**, 236–9.

DURBIN, J. (1967). Estimation of sampling error in multi-stage surveys. *Appl. Statist.* **16**, 152–64.

FELLEGI, I. (1963). Sampling with varying probabilities without replacement: rotating and non-rotating samples. *J. Am. Statist. Ass.* **58**, 183–201.

FISHER, R. A. & YATES, F. (1963). *Statistical Tables for biological, agricultural and medical Research*, 6th edition. Edinburgh: Oliver and Boyd.

GRUNDY, P. M. (1954). A method of sampling with probability exactly proportional to size. *J.R. Statist. Soc.* B **16**, 236–8.

HÁJEK, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Ann. Math. Statist.* **35**, 1491–523.

HANSEN, M. H. & HURWITZ, W. N. (1943). On the theory of sampling from finite populations. *Ann. Math. Statist.* **14**, 333–62.

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

HORVITZ, D. G. & THOMPSON, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Am. Statist. Ass.* **47**, 663–85.

MADOW, W. G. (1949). On the theory of systematic sampling. II. *Ann. Math. Statist.* **20**, 333–54.

YATES, F. & GRUNDY, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *J.R. Statist. Soc.* B **15**, 235–61.

[*Received February* 1967. *Revised June* 1967]