The Annals of Statistics
1977, Vol. 5, No. 2, 414–418

CONVEX SETS OF FINITE POPULATION PLANS

By H. P. Wynn

Imperial College, London

Let P_1 be a finite population sampling plan and V a collection of subsets of units. The inclusion probabilities for members of V may be calculated. For example, if V comprises all single units and pairs of units we obtain all first and second order inclusion probabilities π_i, π_{ij}. Another plan P_2 is called equivalent to P_1 with respect to V if the corresponding inclusion probabilities for P_1 are equal to those for P_2. However, P_2 may have fewer samples with positive probability of selection, that is to say smaller "support." An upper bound is put on the minimum support size of all such P_2. For P_1 simple random sampling, some examples are given for P_2 with small support.

1. The problem. Since the introduction of unequal probability sampling by Horvitz and Thompson (1952) the emphasis in the theory has been towards working with the probabilities π_i and π_{ij} of units and pairs of units appearing in the sample. Some selected further developments appear in Yates and Grundy (1953), Durbin (1953), Grundy (1954).

Two different sampling plans can have certain inclusion probabilities equal. It is possible for one of them to have a smaller number of samples with positive probability of selection. Convexity properties of sets of sampling plans are developed to obtain upper bounds on the minimum number of such samples required. There are analogous results in the theory of continuous experimental designs (Kiefer (1961), page 303). Indeed, it is the similar incidence structure of sampling plans and experimental designs which prompts this analogy. The notation, however, will be that of sampling theory.

Let S be a population of N units labelled $1, \cdots, N$. A sample of size n is a subset u of S. Identifying a sample with its labels we can think of u as a member of the set U of all *combinations* of n distinct integers out of N. Thus different permutations are not distinguished, and sampling is without replacement. A *sampling plan* P allocates to each u in U a probability $p(u) \geqq 0$ of being chosen. Only one sample may be chosen so that

$$\sum_{u \in U} p(u) = 1 .$$

For any set of k units $v = (i_1, \cdots, i_k)$ an *inclusion probability* is defined as

$$\pi_v = \sum_{u \ni v} p(u) \tag{1}$$

where the summation is over all u such that $i_1, \cdots, i_k$ lie in u. When $k = 1$ and $k = 2$ we obtain respectively the first order and second order probabilities π_i

Received July 1974; revised July 1976.

AMS 1970 *subject classification.* Primary 62D05.

Key words and phrases. Finite populations, survey sampling, convexity, randomization.

and π_{ij} of units and pairs of units appearing in the sample. When $k = n$ we have $\pi_v = p(v)$. Thus for a given plan the sets of $p(u)$ and π_v completely determine one another, and two sampling plans are different if any of them are different. Note that we only consider here plans with fixed sample size n.

Let V be a collection of sets of units $v = (i_1, \cdots, i_k)$ ($k \leqq N$, but takes possibly several values for different $v \in V$). Two sampling plans P_1 and P_2 are said to be *equivalent with respect to* V if

$$\pi_v^{(1)} = \pi_v^{(2)}$$

for all $v \in V$, where the index refers to the sampling plan. Let $\mathscr{N}(V)$ be the number of subsets in V. Thus, when V is composed of all singletons and pairs equivalence means that all π_i and π_{ij} are equal between P_1 and P_2 and $\mathscr{N}(V) = \frac{1}{2}N(N+1)$.

Define the *support* of a sampling plan P as those u for which $p(u) > 0$, that is those u having a positive probability of being selected. We call the number of such u the *support size*. For simple random sampling the support size is the maximum possible, $\binom{N}{n}$.

Given a sampling plan P_1 we may ask the following question: what is the minimum support size of a plan P_2 such that P_1 and P_2 are equivalent with respect to a given V?

2. Convexity. Given P_1 and P_2 and corresponding $p^{(1)}(u)$ and $p^{(2)}(u)$ we use the shorthand notation

$$P = (1-\alpha)P_1 + \alpha P_2 \qquad (0 \leqq \alpha \leqq 1)$$

to denote the sampling plan with

$$p(u) = (1-\alpha)p^{(1)}(u) + \alpha p^{(2)}(u)\,,$$

for all u in U. From (1) we see that for a given v, with obvious notation

$$\pi_v = (1-\alpha)\pi_v^{(1)} + \alpha\pi_v^{(2)}\,. \tag{2}$$

For a collection V define a vector $\pi(V)$ whose entries are all the inclusion probabilities π_v for $v \in V$ in some order.

LEMMA. *As P varies over all without replacement sampling plans of sample size n the vector $\pi(V)$ forms a closed convex set in a space of dimension $\mathscr{N}(V)$.*

PROOF. The dimension is merely the number of entries in the vector $\pi(V)$ which is $\mathscr{N}(V)$ by definition. For $u \in U$ let $\pi(\{u\})$ be the vector $\pi(V)$ for the special plan which selects u with probability unity. The entries of $\pi(\{u\})$ are unity for $v \subset u$ and zero for any $v \not\subset u$. From (2) for a general P

$$\pi(V) = \textstyle\sum_{u \in U} p(u)\pi(\{u\})\,, \tag{3}$$

and as P varies we obtain the convex hull of all the vectors $\pi(\{u\})$.

Now Caratheodory's theorem (see Rockafellar (1970), page 151) says that a vector in R^M lying in the convex hull of a set of vectors can be written as a convex combination of no more than $M+1$ of them. Any $\pi(V)$ is in the convex hull of the vectors $\pi(\{u\})$ and thus we obtain

THEOREM 1. *For any sampling plan P_1 and a collection V of $\mathcal{N}(V)$ sets of units there is a sampling plan P_2 with support size no greater than $\mathcal{N}(V)+1$ such that P_2 is equivalent to P_1 with respect to V.*

Equivalence with respect to certain sets in V may imply equivalence with respect to other sets. In more general terms there may be known linear constraints on the inclusion probabilities for sets in V. For example, if V comprises all π_i and π_{ij} then we have the $N+1$ distinct constraints $\sum_{i(\neq j)} \pi_{ij} = (n-1)\pi_j$ and $\sum_i \pi_i = n$. This gives the

COROLLARY. *For a sampling plan P_1 there is a sampling plan P_2 with support size no greater than $\frac{1}{2}N(N-1)$ with the same π_i and π_{ij} as P_1.*

3. Simple random sampling. Simple random sampling (SRS) puts $p(u) = \binom{N}{n}^{-1}$ for all u in U. Consider a balanced incomplete block design (BIBD). Imagine the blocks as samples and treatments as units. The blocks define a possible support for a sampling plan. If for each sample so defined $p(u) = 1/b$ where b is the number of blocks we obtain a sampling plan associated with the BIBD. The following result is given by Chakrabarti (1963) and discussed also in Avadhani and Sukhatme (1973). It is also implicit in the work of Youden (1956) on constrained randomization.

THEOREM 2. *A sampling plan with $p(u)$ uniform over the samples u in the support is equivalent to* SRS *with respect to all first and second order inclusion probabilities if and only if it is associated with a* BIBD, *with $N = t$ and $n = k$, which has distinct blocks.*

The following theorem covers the extreme case when the support size is $\leqq N$. It shows that if the support size is N then we can dispense with the uniform $p(u)$ condition in Theorem 2.

THEOREM 3. *There is a sampling plan P with support size N equivalent to* SRS *with respect to all π_i and π_{ij} if and only if there exists a symmetric* BIBD *with $t = b = N$. No such plan exists with support size less than N.*

PROOF. The sufficiency of the existence of a symmetric BIBD follows from Theorem 2.

Let P have support size N. Label the samples in the support $u^{(1)}, \cdots, u^{(N)}$. Let Q be the incidence matrix with (i, j) entry 1 if unit i is in $u^{(j)}$ and zero otherwise. Let $D = \operatorname{diag}(p(u^{(1)}), \cdots, p(u^{(N)}))$. Then if Π is the matrix with diagonal element π_i and off-diagonal elements π_{ij}

$$\Pi = QDQ^T .$$

Now consider maximising $\det(\Pi)$ over all plans with sample size n. Let $\lambda_1 \geqq \cdots \geqq \lambda_N$ be the ordered eigenvalues of Π so that $\det(\Pi) = \lambda_1 \lambda_2 \cdots \lambda_N$.

The conditions on π_i and π_{ij} give

(i) $\operatorname{trace}(\Pi) = \sum \lambda_i = n$

and

(ii) $q^T \Pi q = n^2$, where q is the $N \times 1$ vector of ones.

But (ii) gives $\lambda_1 = \sup_{\|x\|=1} x^T \Pi x \geqq n^2/N$. For given λ_1, $\det(\Pi)$ is maximized subject to (i) when $\lambda_2 = \cdots = \lambda_N = (n - \lambda_1)/(N - 1)$. Then as a function of λ_1, $\det(\Pi)$ is maximized subject to $\lambda_1 \geqq n^2/N$ when λ_1 attains the boundary value n^2/N. This is certainly achieved when Π has the form for SRS. In this case, then, $\det(QDQ^T)$ is maximized. But Q is $N \times N$ and thus $\det(QDQ^T) = (\det(Q))^2 \det(D)$. Moreover $\det(D) = p(u^{(1)})p(u^{(2)}) \cdots p(u^{(N)})$ and the maximum of this subject to $\sum p(u) = 1$ occurs when $p(u^{(1)}) = \cdots = p(u^{(N)})$. By Theorem 2 this must correspond to a symmetric BIBD.

When the support of P is less than N, rank $(\Pi) < N$ whereas $\det(\Pi)$ for SRS is positive, giving a contradiction.

EXAMPLE 1. When no BIBD exists satisfying $b \leqq t(t-1)/2$, it is clear from Theorem 2 that the minimum support size plan cannot be uniform. Many such examples exist. The smallest $t = N$ for which no such BIBD exists is for $t = 8$, $n = k = 3$. The bound from the corollary to Theorem 1 is 28 whereas the smallest BIBD with these parameters is the irreducible one corresponding exactly to SRS. The support size of the latter is $b = \binom{8}{3} = 56$.

After some inspection the following plan was found with $N = 8$, $n = 3$, support size 24, unequal $p(u)$ and equivalent to SRS with respect to all π_i and π_{ij}. The sets of 3 numbers below represent samples in the support. Those with the same $p(u)$, expressed as multiples of 1/56 are grouped together. Note that by writing down each sample the number of times indicated we in fact obtain a BIBD with $t = 8$, $k = 3$ but only 24 *distinct* blocks. Such a design does not appear to have been given before.

u		$p(u) \times 56$
125	456	4
236	167	
347	278	
148	358	
245	157	2
136	268	
247	357	
138	468	
123	567	1
124	568	
134	578	
234	678	

EXAMPLE 2. Advadhani and Sukhatme (1973) discuss various methods of reducing the support size. A simple method is to take a convex combination of a stratified sampling plan and a cluster sampling plan in which clusters and strata are identical.

Suppose it is possible to divide the population S into L disjoint sets $S_1, \cdots, S_L$ of equal size M, so that $N = LM$. Suppose also that there are integers l, m and n such that $mL = Ml = n$. Consider two plans of sample size n, P_1 which is stratified random sampling with SRS of sample size m in each *stratum* S_r ($r = 1, \cdots, L$) and P_2 which is cluster sampling which selects l out of L entire *clusters* $S_1, \cdots, S_L$ according to SRS. The sampling plan in the notation of Section 2:

$$P = \alpha P_1 + (1 - \alpha)P_2$$

where $\alpha = (N - L)/(N - 1)$, is equivalent to SRS with respect to all π_i and π_{ij}. This is seen by equating π_{ij} across and within the S_r.

Note that P does not imply taking both a stratified and cluster sample but one or the other with fixed probability. When $N = 1$ or L we clearly revert to P_1 or P_2. This confirms a vague notion that the cluster and stratified sampling "cancel each other out."

REFERENCES

AVADHANI, M. S. and SUKHATME, B. V. (1973). Controlled sampling with equal probabilities and without replacement. *Internat. Statist. Rev.* **41** 175-182.

CHAKRABARTI, M. C. (1963). On the use of incidence matrices of designs in sampling from finite populations. *J. Indian Statist. Soc.* **1** 78-85.

DURBIN, J. (1953). Some results in sampling theory when the units are selected with unequal probabilities. *J. Roy. Statist. Soc. B* **15** 262-269.

GRUNDY, P. M. (1954). A method of sampling with probability proportional to size. *J. Roy. Statist. Soc. B* **16** 236-238.

HORVITZ, D. G. and THOMPSON, D. J. (1952). A generalisation of sampling without replacement. *J. Amer. Statist. Soc.* **47** 663-685.

KIEFER, J. (1961). Optimum designs in regression problems, II. *Ann. Math. Statist.* **32** 298-325.

ROCKAFELLAR, R. T. (1970). *Convex Analysis.* Princeton Univ. Press.

YATES, F. and GRUNDY, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *J. Roy. Statist. Soc. B* **15** 253-261.

YOUDEN, W. J. (1956). Randomization and experimentation. *Ann. Math. Statist.* **27** 1185-1186.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720