SAMPLING THEORY USING EXPERIMENTAL DESIGN CONCEPTS

Jaya Srivastava, Colorado State University

and

Zhao Ouyang, Colorado State University

Abstract

In this paper, we consider the application of concepts of Statistical Experimental Design to Sampling Theory. As is well-known, because of its inherent nature, Experimental Design Theory involves a relatively heavy amount of Combinatorial Mathematics. It turns out that, over the years, relatively speaking, it is this combinatorial aspect of Design, that has found much application in Sampling. We present a brief review of the same, including some of the latest work in the field.

Introduction

The subject of *sampling using experimental design concepts* has attracted more and more attention in recent years. A very explicit connection was made by M.C. Chakrabarti (1963) who indicated that balanced incomplete block designs (BIBD's) could be used as sampling schemes. At first, it was shown that a BIBD procedure has properties similar to SRSWOR (simple random sampling without replacement). But later on it was found that a BIBD corresponds, in a sense, to *controlled sampling*, which was proposed by Goodman and Kish in 1950, and to which further contributions were made by Avadhani and Sukhatme (1965, 1968, 1973).

Consider an agricultural survey. Suppose we use SRSWOR to draw a sample of n counties from a population of N counties. It may happen that the n counties in our sample are spread out in an undesirable or inconvenient manner. As pointed out by Avadhani and Sukhatme (1973), "this may not only increase considerably the expenditure on travel, but the quality of data collected is also likely to be seriously affected by non-sampling errors, particularly non-response and investigator bias, since in such cases organizing close supervision over the field work would generally be fraught with administrative difficulties". Such a sample is considered as *non-preferred*. Hence the total set of $\binom{N}{n}$ samples can be classified into two classes: *preferred* samples and *non-preferred* samples (Goodman and Kish, 1950). Hence, our objective is to design a sampling procedure which reduces the probability of drawing a non-preferred sample as much as possible, and at the same time *resembles* SRSWOR (assuming no stratification, clustering, etc. is present, and there are no auxiliary variables).

The problem of controlled sampling was first proposed by Goodman and Kish (1950). This method involves stratified sampling and emphasizes the minimization of the probability of the selection of the non-preferred samples. But, as discussed by Avadhani and Sukhatme (1973), this method may lose precision in estimation. In their three papers (1965, 1968, 1973), Avadhani and Sukhatme discuss the problem of minimizing the chance of selection of non-preferred samples without losing efficiency relative to SRSWOR.

We recall some useful notation from Srivastava (1985). Let U denote a population with N units denoted by the integers 1, 2,...,N. Let y be the variable of interest, and let y_i ($i = 1,...,N$) be the value of y for the unit i in U. Let $Y \equiv \sum_1^N y_i$ be the *population total*. The class of all subsets of U is denoted by 2^U, and any $\omega \in 2^U$ is called a sample of U. (This includes the empty sample.) For any set K, let $|K|$ denote the number of elements in K. For any $\omega \in 2^U$, let $(\omega : n)$ be the class of all n-element subsets of ω; if $|\omega| < n$, then this class is empty. A sampling measure, denoted by $p(\cdot)$, is a probability density $\{p(\omega)\}$ defined on 2^U. For a given $p(\cdot)$, let

$$\pi_i = \sum_{\omega:\, i \in \omega} p(\omega), \quad i = 1,...,N. \tag{1}$$

Then, π_i ($i = 1,...,N$) is the probability that the unit i is included in the sample. For any non-empty sample ω, let $\bar{y}_\omega$ denote the sample mean. Consider a sampling measure p for which all inclusion probabilities π_i ($i = 1,...,N$) equal (n/N). Then, Avadhani and Sukhatme define p to be admissible if (i) $N\bar{y}_\omega$ is an unbiased estimator of Y, and (ii) $Var_p(N\bar{y}_\omega) \leq Var_{SRS}(N\bar{y}_\omega)$, where Var_p and Var_{SRS} denote the variance respectively under the measure p, and the measure q induced by SRSWOR with sample size n. (Note that, for all $\omega \in 2^U$, $q(\omega) = \{1/\binom{N}{n}\}$, if $|\omega| = n$, and $q(\omega) = 0$, otherwise.)

Let $3 \leq n \leq N-3$. The following results are given by Avadhani and Sukhatme (1973).

Theorem 1

Let $S \subset (U: n)$, and let $|S| = b$. Then the sampling measure which selects each $\omega \in S$ with probability $(1/b)$ is admissible if and only if $|\{\omega: \omega \in S, i,j \in \omega, i \neq j\}|$, are the same for all $i \neq j$, $i,j = 1,...,N$. For such a measure, $|\{\omega: \omega \in S, i \in \omega\}|$ are the same for all $i = 1,...,N$.

Under the condition of Theorem 1, let

$$\lambda = |\{\omega: \omega \in S, i,j \in \omega, i \neq j\}| \tag{2}$$

$$r = bn/N. \tag{3}$$

It is easy to see that the existence of S in Theorem 1 is equivalent to the existence of a BIBD with parameters (N, b, r, n, λ), such that N is the number of treatments, b the number of blocks, r the number of replications for each treatment, λ the number of blocks which contain any given pair of treatments, and n the block size. In fact, such a S is a BIBD with the above parameters. But, when N and n are large, such a BIBD may be hard to identify. So, the next two theorems are useful.

Theorem 2

The measure induced by the following (two-part) sampling procedure is admissible:

(i)   Split the population randomly into k subpopulations with fixed sizes N_i $(i = 1,...,k)$ such that $\sum_{i=1}^{k} N_i = N$,

(ii)   For $i = 1,...,k$, select n_i units from the ith subpopulation by using an admissible sampling measure (with inclusion probability (n_i/N_i)). The selection of the units from the different subpopulations should be done independently.

Corollary 1

The measure induced by the following procedure is admissible:

(i)   Draw a sample of size $n' > n$ from the population by SRSWOR.

(ii)   From the sample selected in (i), draw a sample of size n by using an admissible measure with inclusion probability n/n' for each unit.

In view of the above, Avadhani and Sukhatme suggest that the following steps may be followed for controlled sampling:

(i)   Let $N_1 + N_2 + ... + N_g = N$. Divide the original population randomly into g subpopulations, which have sizes $N_1, N_2,...,N_g$ respectively.

(ii)   Let $n_1 + n_2 + ... + n_g = n$. For $i = 1, 2,...,g$, select an integer n'_i such that $n_i \leq n'_i < N_i$ and also select a BIBD with parameters $(n'_i, b_i, r_i, n_i, \lambda_i)$. (It is preferred that n_i be much smaller than n'_i.) Use SRSWOR to select (independently for each i) a sample of size n'_i from the ith subpopulation of size N.

(iii)   For each sample of size n'_i $(i = 1,...,g)$ drawn in step (ii), collect the information on all the preferred subsamples of size n_i and then find a

BIBD with parameters $(n'_i, b_i, r_i, n_i, \lambda_i)$ such that the number of the blocks which correspond to the preferred subsamples of size n_i is as large as possible. Then draw one block with probability $1/b_i$ from the ith BIBD independently for $i = 1,...,g$. In this way, we get a sample of total size $n_1 + ... + n_g = n$.

An example of controlled sampling using BIBD will be given in the last section in this paper.

Other Works on Sampling Using Concepts of Experiment Design

In the first section, we discussed the use of BIBD in controlled sampling. It is very clear that for a BIBD with parameters (N, b, r, n, λ), N corresponds to the number of units in the population, b corresponds to the (maximum possible) number of distinct samples, and n corresponds to the size of the sample. With this interpretation, it is easy to see that the parameters r and λ in the BIBD correspond respectively to the first order and the second order inclusion probabilities. So, for some time, the use of BIBD in sampling has been discussed widely.

As early as 1963, Chakrabarti pointed out the equivalence between SRSWOR and BIBD in the sense of having the same first order and second order inclusion probabilities. It is clear that the smaller the support of (i.e., the number of distinct blocks in) the BIBD, the better is the possibility of adapting it for a given situation of controlled sampling. Thus, BIBD's with a small support size have importance in sampling theory. Because of this, the work of Hedayat and others in the field of BIBD's with small supports is useful.

In 1977, Wynn showed that for each sampling measure p_1 there is a measure p_2, which gives rise to the same first and second order inclusion probabilities as p_1, and whose support size is not greater than $N(N-1)/2$. For the case of SRSWOR, he showed that no BIBD with support size less than N can be equivalent to SRSWOR in the above sense. Hence, with the help of BIBD's we can reduce the support size from SRSWOR's $\binom{N}{n}$ to something between $\binom{N}{2}$ and N.

Besides BIBD, Fienberg and Tanur (1985) listed some parallel concepts in Design of Experiments and Sampling. These include randomization in design and random sampling, blocking in design and stratification in sampling, Latin square in design and lattice sampling, split-plot design and cluster sampling, and covariance adjustment in design and post-stratification in sampling. By using some similar parallel concepts in design and sampling, Meeden and Ghosh (1983) found some admissible strategies in sampling and Cheng and Li (1983) showed that Rao-Hartley-Cochran and Hansen-Hurwitz strategies are approximately minimax under some models. Brewer et al. (1977) discussed use of experimental design in the planning of sample surveys, and Sedransk (1967) discussed the use of experimental design in the analysis of sample surveys. But, even though experimental design and sampling have so many parallel concepts and similar structure, sampling has been developed separately from experimental design.

Smith and Snyder (1985) pointed out the main distinction between experimental design and sampling from their nature of inference. They concluded that "the differences between survey and experiments are as important as the similarities, and that each will continue to develop in its own way". An excellent discussion of experimental design and sample surveys, both with respect to their similarities and differences, was given by Fienberg and Tanur (1985).

Hedayat (1979) gave a method for finding a sampling design which has the same first and second order inclusion probabilities, but has a reduced support size than SRSWOR. (In other words, he gave a general method for obtaining BIBD's with relatively small support sizes.) Let M denote the incidence matrix of all the pairs (i, j) versus all the samples of U with size n, where $i, j \in U$. Thus, M is a $\left(\binom{N}{2} \times \binom{N}{n}\right)$ zero-one matrix. Suppose all the samples of U with size n are arranged in a list in an arbitrary but fixed order. Consider a BIBD (with block size n) in which f_k denotes the frequency of the kth sample in the above list. Let $\underline{f} = (f_1, f_2, \ldots, f\binom{N}{n})$. Consider a sampling measure p which assigns probability $(f_k / \sum_i f_i)$ to the kth sample. Then, p has the same first and second order inclusion probabilities as SRSWOR of size n iff $M\underline{f} = \lambda\underline{1}$, where λ is a positive integer and $\underline{1}$ is a column vector with all entries equal to 1. So each feasible solution of the system

$$M\underline{f} = \lambda\underline{1}, \underline{f} \geq 0 \qquad (4)$$

gives a sampling measure equivalent to SRSWOR of size n. Notice that there is always a solution for the system. So we can introduce another quantity, for example, the number of non-zero entries in $\underline{f}$, and find a feasible solution of the system to minimize the quantity. The algorithm of mathematical programming can be used to get such a solution. In other papers in combinatorics, Hedayat and others give further results.

In Hedayat and Pesotan (1983), $(R \times L)$ *triply balanced matrices* was discussed. The $(R \times L)$ triply balanced matrices arise in estimating the mean square error of nonlinear estimators in sampling. Briefly, a $(R \times L)$ triply balanced matrix is $\Delta = (\delta_{ij})$ with entries $+1$ or -1 such that $\sum_{r=1}^{R} \delta_{rh} = 0$, $\sum_{r=1}^{R} \delta_{rh}\delta_{rs} = 0$, $\sum_{r=1}^{R} \delta_{rh}\delta_{rs}\delta_{rt} = 0$, where the h, s, t are distinct and h, s, $t = 1, \ldots, L$. It was proved that a $(R \times L)$ triply balanced matrix Δ is an orthogonal array of strength 3 and 2 symbols.

In Hedayat, Rao, and Stufken (1988), balanced sampling plans excluding contiguous units are discussed. In some situations, the N units of the population are arranged in a natural order. In this case it may happen that contiguous units provides us *similar* information so that it seems more reasonable to select a sampling plan such that the contiguous unit cannot appear in the sample. Here

the term *balanced* means that the first and second order inclusion probabilities are fixed. The condition of the existence of such a sampling measure is given in this paper, and a method of constructing such a sampling measure is also proposed.

Use of t-Design

Suggested by the usefulness of BIBD with sampling, the use of t-design in sampling was proposed by Srivastava and Saleh (1985). A BIBD, which has the same inclusion probabilities (of individual units, and pairs of units) as SRSWOR, has the same moments as SRSWOR up to order two. Generalizing this, Srivastava and Saleh showed that a t-design has the same moments as SRSWOR up to order t, because it has the same inclusion probabilities as SRSWOR up to order t (i.e. every set of i units ($i = 1, \ldots, t$) has the same inclusion probability, say q_i). Also, as for the BIBD, the sample space under a t-design can be much smaller than the sample space under SRSWOR. Thus, using t-designs we can try to avoid non-preferred samples, and still maintain resemblance to SRSWOR up to moments of order t.

For later use, define $a_{i\omega}$ ($i \in U$, $\omega \in 2^U$) by

$$a_{i\omega} = 1, \text{ if } i \in \omega$$

$$= 0, \text{ otherwise.} \tag{5}$$

Let $1 \leq k \leq N$. For any sampling measure $\{p(\omega): \omega \in 2^U\}$ define

$$\pi(i_1, \ldots, i_k) = \sum_\omega p(\omega) a_{i_1\omega} a_{i_2\omega} \cdots a_{i_k\omega}, \tag{6}$$

where $i_1, i_2, \ldots, i_k \in U$.

In this section, we suppose the sample size is always equal to n, a fixed integer. We are interested in estimating the population total Y.

The following results from Srivastava and Saleh (1985) are useful in the studies on using t-design theory in sampling.

Lemma 1

Let $2 \leq k \leq n$. Suppose $i_1, \ldots, i_k$ are distinct elements of U. Then we have

$$\sum_{i_k=1}^{N} \pi(i_1, \ldots, i_k) = (n - k + 1)\pi(i_1, \ldots, i_{k-1}),$$

$$i_k \neq i_1, \ldots, i_{k-1} \tag{7}$$

$$\sum_{i=1}^{N} \pi(i) = n. \tag{8}$$

This lemma says that for $2 \leq k \leq n$, the inclusion probabilities of order $j (1 \leq j \leq k - 1)$ are determined by the inclusion probabilities of order k.

Theorem 3

Suppose there are two different sampling measures on 2^U. Let t be a positive integer. Then these two sampling measures give the same inclusion probabilities of order t if and only if these two sampling measures give the same values of $E(\overline{y}_\omega^k)$, $k = 1, \ldots, t$, for all possible values of $(y_1, \ldots, y_n)$.

Let

$$\psi_q(\omega) = \sum_{i \in \omega} (y_i - \overline{y}_\omega)^9 \tag{9}$$

$$s_\omega^2 = \frac{1}{n-1} \sum_{i \in \omega} (y_i - \overline{y}_\omega)^2 = \frac{1}{n-1} \psi_2(\omega). \tag{10}$$

Then, we have

Theorem 4

Consider two sampling measures on 2^U. Consider the following four conditions:

(i)   For all possible values of $\underline{y} = (y_1, \ldots, y_n)'$, $E(\overline{y}_\omega)$ is the same under these two sampling measures,

(ii)   For all possible values of $\underline{y}$, $E(\overline{y}_\omega^2)$, or $E(s_\omega^2)$, or $V(\overline{y}_\omega)$ is the same under these two sampling measures,

(iii)   For all possible values of $\underline{y}$, $cov(\overline{y}_\omega, s_\omega^2)$ is the same under these two sampling measures,

(iv)   For all possible values of $\underline{y}$, $V(s_\omega^2)$ is the same under these two sampling measures.

Let t be an integer such that $1 \leq t \leq 4$. Then the above conditions (i), (ii), up to (t) are true if and only if these two sampling measures have the same inclusion probabilities of order t.

One can generalize Theorem 4 to higher order. But the most important case is order 4. In this case, we can characterize the mean and the variance of a linear estimator, and characterize the variance of a quadratic estimator of the variance of the linear estimator.

Now consider a t-design $D(N, n, t, b)$ where N is the number of varieties, n the block size, b the number of blocks (which may or may not be distinct), and where every combination of t varieties $(t \leq u)$ occurs in $b\binom{u}{t} / \binom{N}{n}$ blocks.

Consider a sampling measure (called a *t-design sampling measure*) which selects each block of $D(N, n, t, b)$ with probability $1/b$. When $b = \binom{N}{n}$ and each block in $D\left(N, n, t, \binom{N}{n}\right)$ is distinct, this sampling measure becomes SRSWOR. In this case SRSWOR is a t-design $D\left(N, n, t, \binom{N}{n}\right)$, where t can take any value from 1 to n.

For the t-design sampling measure mentioned above, for distinct $i_1, \ldots, i_t \in U$, we have

$$\pi(i_1, \ldots, i_t) = \frac{1}{b} \, b \binom{n}{t} \bigg/ \binom{N}{t} = \binom{n}{t} \bigg/ \binom{N}{t}. \tag{11}$$

Hence we have the following theorem.

Theorem 5

SRSWOR (with sample size n) and the t-design sampling measure have the same inclusion probabilities of order t and hence have the same moments up to order t.

For any t-design $D(N, n, t, b)$, the number of distinct blocks is not greater than $\binom{N}{n}$, and usually is much less than $\binom{N}{n}$. This makes a t-design useful in controlled sampling. In fact, a BIBD is a 2-design. Because we need to estimate $V(\bar{y}_\omega)$, we need to consider up to the fourth moments; the first two moments are not enough. In view of this, Srivastava and Saleh assert that it would be much better to use 4-designs rather than BIBD's, since the former gives rise to the same moments as SRSWOR up to order 4.

Connection with Arrays

The theory of factorial designs constitutes a major part of the whole subject of experimental design. Furthermore, the modern theory of factorial designs is largely built around the concept of arrays. Indeed, arrays constitute a very important tool in all of design theory, since for example, BIBD's, PBIBD's and t-designs, etc. may (through their *incidence matrices*) be studied in terms of arrays. Because of this, in this section, we discuss the application of arrays in sampling theory. An array is a matrix whose elements come from a finite set. Suppose the finite set has m elements in it. Without loss of generality, we use the integers $0, 1, \ldots, m-1$ to denote the elements of the finite set. In this case, an array is a matrix whose elements belong to the set $\{0, 1, \ldots, m-1\}$. When $m = 2$, such an array becomes $(0, 1)$ matrix which is of special importance.

A special case of a $(0, 1)$ matrix is the incidence matrix of a class of subsets of a given finite set. The rows of an incidence matrix correspond to the elements of the given finite set and the columns correspond to the subsets of the given finite set. In sampling, an incidence matrix is Ω_U which is a $(N \times 2^N)$ $(0, 1)$-matrix such that its columns correspond to the elements of 2^U, and rows to

the elements of U. In order to simplify the discussion, and without loss of generality, we assume that the ith row of Ω_U corresponds to the element i of U, and the jth column of Ω_U corresponds to the jth element of 2^U such that the elements of 2^U are arranged in the following standard order:

(i)   If $\omega_1, \omega_2 \in 2^U$ and $|\omega_1| < |\omega_2|$, then ω_1 precedes ω_2;

(ii)   If $|\omega_1| = |\omega_2|$ but there exists a $k \in U$ such that $|\{1,...,k\} \cap \omega_1| > |\{1,...,k\} \cap \omega_2|$ and $|\{1,...,\ell\} \cap \omega_1| = |\{1,...,\ell\} \cap \omega_2|$ for $0 \leq \ell < k$, then ω_1 precedes ω_2.

In this way, the elements of 2^U are arranged as

$$\{\omega(0), \omega(1),...,\omega(2^N - 1)\} \tag{12}$$

and $i \in \omega(j)$ if and only if the ith coordinate of the jth column of Ω_U is equal to 1. For $N = 3$, Ω_U is equal to

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}. \tag{13}$$

Now, given any sampling measure $\{p(\omega): \omega \in 2^U\}$, we can rewrite it as a vector which is called the vector form of sampling measure

$$\underline{p}' = \Big(p(\omega(0)),\ p(\omega(1)),...,p(\omega(2^N - 1))\Big). \tag{14}$$

Combining Ω_U and $\underline{p}'$, we have a matrix $\pi_U(p)$ where

$$\pi_U(p) = \left[\frac{\Omega_U}{\underline{p}'} \right]. \tag{15}$$

Thus, $\pi_U(p)$ presents a sampling measure in a matrix form. Now suppose all the $p(\omega(j))$ are rational numbers and $p(\omega(j)) = v_j/v$ such that v_j is a non-negative integer, $v = \Sigma v_j$ (where the sum runs over all j), also suppose there is no common factor other than 1 among the v_j ($j = 0, 1,...,2^n-1$). Suppose

$$\Omega_U = [\underline{c}_0,\ \underline{c}_1,...,\underline{c}_{2^n-1}]. \tag{16}$$

Now we introduce another matrix $\Delta_U(p)$ such that

$$\Delta_U(p) = \left[\underline{c}_0 1'_{v_0} \Big| \underline{c}_1 1'_{v_1} \Big| \cdots \Big| \underline{c}_{2^n-1} 1'_{v_{2^n-1}} \right] \tag{17}$$

where $1'_k$ is the $(1 \times k)$ vector containing 1 *everywhere*, and where if for any j, we have $v_j = 0$, then the columns $(\underline{c}_j)$ do not appear in $\Delta_U(p)$. Now, drawing a column from $\Delta_U(p)$ with probability $(1/v)$ is equivalent to drawing a column from Ω_U with probability measure $\{p(\omega(j)) = v_j/v, \ j = 0, 1,\ldots,2^N-1\}$. So, the matrix $\Delta_U(p)$ represents the sampling measure in the form of an array and it is called *sampling array* in Srivastava (1988), wherein the following result is proved.

Theorem 6

For any vector form of sampling measure $\underline{p}$ and $\epsilon > 0$, there exists a vector form of sampling measure $\underline{p}^*$ whose elements are rational such that $(\underline{p} - \underline{p}^*)'(\underline{p} - \underline{p}^*) < \epsilon$. (Note that every sampling measure can be expressed in the vector form.)

Although this theorem seems simple it has an important interpretation in that we can replace a sampling measure by a rational sampling measure as closely as we want. On the other hand, by using a rational sampling measure we get a sampling array. So the above theorem connects sampling theory to the theory of arrays in a fundamental manner, and hence to factorial and other experimental designs.

Now consider the problem of estimating the population total Y by a general linear estimator $\hat{Y}_G$ (G means general), where

$$\hat{Y}_G = \sum_{i \in \omega} c_{i\omega}y_i = \sum_{i=1}^{N} c_{i\omega}a_{i\omega}y_i, \tag{18}$$

and where $c_{i\omega}$ are known real numbers which depend on i and ω for all $i \in U$, $\omega \in 2^U$. Define

$$\phi_{ic} = \sum_{\omega} c_{i\omega}a_{i\omega}p(\omega) \tag{19}$$

$$\phi^0_{iic} = \sum_{\omega} c^2_{i\omega}a_{i\omega}p(\omega), \ \phi^0_{ijc} = \sum_{\omega} c_{i\omega}c_{j\omega}a_{i\omega}a_{j\omega}p(\omega) \tag{20}$$

$$\underline{\phi}_c = (\phi_{1c},\ldots,\phi_{Nc})', \ \Phi^0_c = (\phi^0_{ijc})_{N \times N} \tag{21}$$

$$\Phi_c = \Phi^0_c - (J_{N1}\phi'_c + \underline{\phi}_c J_{1N}) + J_{NN} \tag{22}$$

where J_{mn} is a m $\times$ n matrix which elements are equal to 1. It is easy to check that

$$\Phi_c = \sum_{\omega} p(\omega)\underline{U}_{c_\omega}\underline{U}'_{c_\omega} \tag{23}$$

where

$$\underline{U}_{c_\omega} = (c_{1\omega}a_{1\omega}-1, \ c_{2\omega}a_{2\omega}-1,\ldots,c_{N\omega}a_{N\omega}-1). \tag{24}$$

We have the following theorem:

Theorem 7

The mean square error of $\hat{Y}_G$ as an estimator of Y denoted by $MSE(\hat{Y}_G)$ is

$$MSE(\hat{Y}_G) = \underline{Y}'\Phi_c\underline{Y}. \tag{25}$$

Notice that the matrix Φ_c is known when the sampling measure and the estimator $\hat{Y}_G$ are selected. The matrix Φ_c in sampling theory is similar to the information matrix in the theory of experimental design.

A General Estimator

In this section, we consider an estimator proposed in Srivastava (1985). There is an interesting history relevant here. First, in 1985, Srivastava observed the connection between combinatorial arrays and sampling theory, discussed in the last section. This appeared to open up a quite new theoretical field, in which variable sample size appeared to be inherent. Thus, there seemed to be a need for a general estimator in which sample size was not necessarily fixed. Now, most estimators in sampling theory relate to fixed size. In many ways, the most general estimator (which, among other things, allows variable sample size) existing in 1985 was the Horvitz-Thompson estimator. But this is entirely dependent on the sampling measure, which is of course decided upon *before* the sample is drawn. In an attempt to be able to utilize the new knowledge (*independent* of the sample, but obtained during the course of actual sampling) the concepts of the *sample weight function* (discussed below), and the estimator of this section, were discovered. This estimator is extremely general, in that most of the known estimators turn out to be its special cases.

The most important concept in this estimator is the introduction of the *sample weight function* r, defined on 2^U, such that for all $\omega \in 2^U$, $r(\omega)$ is a finite real number. For every $K \subset U$, and $k \in (1, 2, \ldots, N)$, let

$$(K{:}\ k) = \{\omega{:}\ w \subset K, |\omega| = k\}.$$

Clearly, if $\underline{i} \in (U{:}\ k)$, then $\underline{i}$ is a k-tuple, with k distinct elements from U. From here on, $\sum_{\underline{i}}$ will denote the sum over all $\underline{i} \in (U{:}\ k)$, $\sum_{\omega\underline{i}}$ will denote the sum over all $\omega \in 2^U$ such that $\underline{i} \subset \omega$, and $\sum_{\underline{i}\omega}$ will denote the sum over all $\underline{i} \in (w{:}\ k)$. Note that the last sum could be empty. In this section, we always look upon $\underline{i} = (i_1, \ldots, i_k)$ as an unordered set $\{i_1, \ldots, i_k\}$. Let

$$\pi_r(\underline{i}) = \sum_{\omega\underline{i}} p(\omega) r(\omega). \tag{26}$$

For $k \in (1, 2, \ldots, N{-}1)$, $t \in (1, 2, \ldots, N)$, and $\underline{i} \in (U{:}\ k)$, let $T_r(\underline{i}, t)$ be the class of all unordered sets $\underline{j} = (j_0, \ldots, j_{t-1})$ such that $\underline{j} \in (U{:}\ t)$ and $\underline{i} \subset \underline{j}$

where $\underline{i} = (i_1, \dots, i_k)$, and $\pi_r(\underline{j}) \neq 0$. Let

$$\nu_r(\underline{i}, t) = |T_r(\underline{i}, t)|. \tag{27}$$

Also, let $\alpha(\underline{i}, t)$ be real numbers which satisfy the following two conditions:

$$\alpha(\underline{i}, t) = 0, \text{ if } \nu_r(\underline{i}, t) = 0, \text{ and} \tag{28}$$

$$\sum_{t=1}^{N} \alpha(\underline{i}, t) = 1. \tag{29}$$

For all $\omega \in 2^U$, $\underline{i} \in (U: k)$, define

$$\beta_r(\underline{i}, \omega) = r(\omega) \sum_{t=1}^{N} \alpha(\underline{i}, t) \big[v_r(\underline{i}, t) \big] \Big\{ \sum{}^* \big[\pi_r(\underline{j}) \big] \Big\} \tag{30}$$

where $a^- = a^{-1}$ if $a \neq 0$ and $a^- = 0$ if $a = 0$, and Σ^* runs over all $\underline{j} \in (U: t)$ such that $\underline{i} = (i_1, \dots, i_k) \subset \underline{j}$ and $\underline{j} \subset \omega$. Now, consider the estimation of the following *symmetric linear population function* $Q(\psi)$ where

$$Q(\psi) = \sum_{\underline{i}} \psi(\underline{i}), \tag{31}$$

where ψ, defined over $(U: k)$, is such that for all $\underline{i} \in (U: k)$, $\psi(\underline{i})$ is a real number. Notice when $\underline{i} \subset \omega$ and ω is selected, $\psi(\underline{i})$ can be calculated. Thus, once a sample ω is drawn, we can compute $\hat{Q}^{sr}(\psi)$ where

$$\hat{Q}^{sr}(\psi) = \sum_{\underline{i}\omega} \psi(\underline{i}) \beta_r(\underline{i}, \omega). \tag{32}$$

Here in $\hat{Q}^{sr}$, s means that we are estimating a symmetric function, and r means that the sample weight function r is being used.

Theorem 8

The statistic $\hat{Q}^{sr}(\psi)$ is an unbiased estimator of $Q(\psi)$, if and only if for every $\underline{i} \in (U: k)$ with $\psi(\underline{i}) \neq 0$, there exists a t such that $1 \leq t \leq N$, $\nu_r(\underline{i}, t) \neq 0$.

For the case of $\pi_r(\underline{i}) \neq 0$ for all $\underline{i} \in (U: k)$, let $\alpha(\underline{i}, k) = 1$ and $\alpha(\underline{i}, t) = 0$ for all $t \neq k$. Then we have

$$\beta_r(\underline{i}, \omega) = r(\omega) \big[\pi_r(\underline{i}) \big]^{-1} \text{ and} \tag{33}$$

$$\hat{Q}^{sr}(\psi) = r(\omega) \sum_{\underline{i}\omega} \psi(\underline{i})/\pi_r(\underline{i}). \qquad (34)$$

By using Theorem 8, it can be checked that (34) is unbiased for $Q(\psi)$.

The variance of $\hat{Q}^{sr}(\psi)$ and an unbiased estimator of the variance of $\hat{Q}^{sr}(\psi)$ were also obtained in Srivastava (1985). Now we turn to an estimator of the population total.

Let $k = 1$, $\underline{i} = 1$, $\psi(\underline{i}) = y_i$. Then $Q(\psi) = \sum_{i=1}^{N} y_i = Y$. Then, (34) gives

$$\hat{Q}^{sr}(\psi) = r(\omega) \sum_{i \in \omega} y_i/\pi_r(i) \equiv \hat{Y}_{sr1} \text{ say.} \qquad (35)$$

By Theorem 8, if $\pi_r(i) \neq 0$ for all $i \in U$, then $\hat{Y}_{sr1}$ is an unbiased estimator of Y. When

$$r(\omega) = 1, \text{ for all } \omega \in 2^U, \qquad (36)$$

$\pi_r(i)$'s become π_i's where π_i is the probability such that the unit i is included in the sample. At this time, $\hat{Y}_{sr1}$ becomes the well known Horvitz-Thompson estimator $\hat{Y}_{HT}$ where

$$\hat{Y}_{HT} = \sum_{i\omega} y_i/\pi_i. \qquad (37)$$

The variance of $\hat{Y}_{sr1}$ is given in the following theorem.

Theorem 9

Suppose $\pi_r(i) \neq 0$, $i = 1,...,N$. Then

$$Var(\hat{Y}_{sr1}) = \sum_{i=1}^{N} y_i^2 \left[\frac{\pi_{r2}(i) - \left(\pi_r(i)\right)^2}{\left(\pi_r(i)\right)^2} \right] + \sum_{i \neq j,=1}^{N} y_i y_j \left[\frac{\pi_{r2}(i,j) - \pi_r(i)\pi_r(j)}{\pi_r(i)\pi_r(j)} \right]$$

$$(38)$$

where

$$\pi_{r2}(i) = \sum_{\omega i} p(\omega)[r(\omega)]^2; \ i = 1,...,N \qquad (39)$$

$$\pi_{r2}(i,j) = \sum_{\omega ij} p(\omega)[r(\omega)]^2, \ i \neq , \ i,j = 1,...,N. \qquad (40)$$

It is easy to see that

$$\pi_{r^2}(i)/(\pi_r(i))^2 \geq 1/\pi_i, \ i = 1,\ldots,N. \tag{41}$$

So the term in y_i^2 in $Var(\hat{Y}_{sr1})$ is always larger than the correspondent term for $\hat{Y}_{HT}$. But we can choose $r(\omega)$ such that the cross product terms of $\hat{Y}_{sr1}$ are small so that $Var(\hat{Y}_{sr1})$ is small. Examples are given in Srivastava (1985).

Balanced Array Sampling

We have defined arrays in the fourth section. Let $K(a \times b)$ and $\underline{k}(a \times 1)$ be a matrix and a vector with elements from σ_s, where σ_s is a finite set whose elements are $(0, 1,\ldots,s-1)$. The symbol $\lambda(\cdot,\cdot)$ is defined as a counting operator, such that $\lambda(\underline{k}, K)$ is equal to the number of times $\underline{k}$ occurs as a column of K. Let ψ_s be the permutation group over σ_s. For $\psi \in \psi_s$, and $j \in \sigma_s$, let $\psi(j)$ be the image of j when the permutation ψ is applied. Similarly, we define $\psi(\underline{k}) = (\psi(k_1),\ldots,\psi(k_a))$ if $\underline{k} = (k_1,\ldots,k_a)$ is a $(a \times 1)$ array over σ_s.

Definition 1

Let K be a $(a \times b)$ array over σ_s. Then K is a *balanced array* (B-array, or BA) of strength t if and only if

$$\lambda(\underline{k}_0, K_0) = \lambda(\psi(\underline{k}_0), K_0) \tag{42}$$

where $\underline{k}_0$ is any $(t \times 1)$ array over σ_s, K_0 is any $(t \times b)$ subarray of K and ψ is any permutation in ψ_s.

Balanced arrays play an important role in factorial experimental design and coding theory. For $\underline{i} = (i_1,\ldots,i_k) \in (U\colon k)$, define

$$\pi(i_1,\ldots, i_k) = \pi(\underline{i}) = \sum_{\omega \underline{i}} p(\omega). \tag{43}$$

When $k = 1$ or 2, the following customary notations will be used instead of $\pi(i)$, $\pi(i,j)$

$$\pi_i = \pi(i), \ \pi_{ij} = \pi(i,j). \tag{44}$$

Definition 2

Let $p(\cdot) = \{p(\omega)\colon \omega \in 2^U\}$ be a sampling measure. Then $p(\cdot)$ corresponds to *balanced array sampling with strength* k iff $\pi(i_1,\ldots,i_g)$ is fixed, for all possible $(i_1,\ldots,i_g) \in (U\colon g)$. Here, $g = 0, 1,\ldots,k$.

Thus, if $p(\cdot)$ corresponds to balanced array sampling with strength k, then there exists $\theta_1,\ldots,\theta_k$ such that

$$\pi_r(i_1,\ldots,i_g) = \theta_g, \tag{45}$$

for $(i_1,\ldots,i_g) \in (U: g)$, and $g = 0, 1,\ldots,k$.

Theorem 10

Suppose, the measure $p(\cdot)$ corresponds to BA sampling with strength k. Then there exists a sampling measure $p^*(\cdot)$ whose sampling array is $\Delta_U(p^*)$ such that $\Delta_U(p^*)$ is a B-array of strength k, and $p^*(\cdot)$ is arbitrarily close to $p(\cdot)$. (In the sense of Theorem 6.)

Theorem 11

Suppose $\Delta_U^*(p)$ is $(N \times v)$ B-array of strength k. Let $p(\cdot)$ be a sampling measure such that it gives a probability $(1/v)$ to each column of $\Delta_U^*(p)$ for being selected. Then $p(\cdot)$ corresponds to balanced array sampling with strength k.

Let δ_1 and δ_2 be the mean and the variance of the sample size under the measure $p(\cdot)$, i.e.,

$$\delta_2 = \sum_\omega p(\omega)|\omega| \tag{46}$$

$$\delta_2 = \sum_\omega p(\omega)(|\omega| - \delta_1)^2. \tag{47}$$

Then we have the following theorem.

Theorem 12

Consider BA sampling whose inclusion probability is given by (45). Then

$$\hat{Y}_{HT} = \theta_1^{-1}|\omega|\bar{y}_\omega \tag{48}$$

$$V(\hat{Y}_{HT}) = S^2\left[\frac{N}{\delta_1}\left\{(N - \delta_1) - \frac{\delta_2}{\delta_1}\right\}\right] + \frac{N^2\delta_2}{\delta_1^2}\,\bar{Y}^2 \tag{49}$$

$$= S^2 N^2\left(\frac{1}{\delta_1} - \frac{1}{N}\right) + \frac{\delta_2}{\delta_1}\,[N^2\bar{Y}^2 - S^2]$$

where $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i,\ S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$, are respectively of the population mean and variance. (The significance of this result lies in the fact that if we have some idea of the value of $\bar{Y}$, we can reduce the variance below that of SRSWOR. This may happen, for example, in recursive sampling.)

Definition 3

Let $p(\cdot)$ be a sampling measure. Then, $p(\cdot)$ corresponds to *proportional array sampling with strength k* (or, briefly, *proportional sampling*) iff for all integer g such that $1 \leq g \leq k$, and all $(i_1,\ldots,i_g) \in (U\colon g)$ we have

$$\pi(i_1,\ldots,i_g) = \pi(i_1) \ldots \pi(i_g). \tag{50}$$

Notice that when π_i is fixed, say θ, for all $i \in U$, then the *proportional array sampling with strength k* is also *balanced array sampling with strength k*. In this case we call it *balanced proportional sampling with strength k*.

In order to construct a $p(\cdot)$ which corresponds to proportional array sampling with strength k, we need the definition of orthogonal array (OA).

Definition 4

Let K be a $(a \times b)$ array over σ_s. Then K is an orthogonal array of strength t if and only if

$$\lambda(\underline{k}_0, K_0) = b \times s^{-t} \tag{51}$$

where $\underline{k}_0$ is any $(t \times 1)$ array over σ_s, K_0 is any $(t \times b)$ subarray of K. It is easy to see that an OA with strength t is a BA with strength t.

Let $L(N \times b) = (\underline{\ell}_1,\ldots,\underline{\ell}_N)'$ be an OA of strength k over σ_s where s is a prime number. Let s_i be an integer satisfying $1 \leq s_i \leq s$, $i = 1,\ldots,N$. In $\underline{\ell}'_i$, replace the $(s_i - 1)$ symbols $\{2, 3,\ldots,s_i\}$ by 1, leave the original 1 unchanged, and replace the other symbols (if any) by 0. Notice when $s_i = s$, then the symbol s_i is the same as symbol 0. Let $L(N \times b)$ be the array obtained by the above replacement.

Theorem 13

Consider a sampling measure $p(\cdot)$ such that it has $L(N \times b)$ as a sampling array. Then $p(\cdot)$ corresponds to proportional sampling of strength k, such that the inclusion probability of unit i is equal to s_i/s, for $i = 1,\ldots,N$.

Theorem 14

We have

$$var(\hat{Y}_{HT}) = \sum_{i=1}^{N} y_i^2\left(\frac{1}{\pi_i} - 1\right) \tag{52}$$

for proportional sampling and

$$var(\hat{Y}_{HT}) = \left(\frac{1}{\delta_1} - 1\right)\left[(N - 1)s^2 + N\bar{Y}^2\right]. \tag{53}$$

for balanced proportional sampling.

We can use BA sampling with strength 4 to imitate SRSWOR up to the 4th moments. Notice that the binomial sampling referred to in the literature, is a balanced proportional sampling with strength N. It is clear that it should be adequate enough to use balanced proportional sampling with strength 4 instead of using binomial sampling.

Weight Balanced Sampling

Now we introduce an estimator of Y called $\hat{Y}_{s2}$ which is a special case of $\hat{Y}_{s1}$ when

$$r(\omega) = |\omega|^{-1}, \text{ for all } \omega \in 2^U, \omega \neq \phi. \tag{54}$$

Let

$$\pi'_i = \sum_\omega \frac{p(\omega)a_{i\omega}}{|\omega|}, \; i = 1,...,N \tag{55}$$

$$\pi''_i = \sum_\omega \frac{p(\omega)a_{i\omega}}{|\omega|^2}, \; i = 1,...,N \tag{56}$$

$$\pi''_{ij} = \sum_\omega \frac{p(\omega)a_{i\omega}a_{j\omega}}{|\omega|^2}, \; i \neq j, \; i,j = 1,...,N. \tag{57}$$

where we assume that empty samples are not allowed.

Theorem 15

Suppose $\pi_i > 0$ for $i = 1,...,N$. Then

$$\hat{Y}_{s2} = |\mathbf{w}|^{-1} \sum_{i \in \omega} y_i/\pi'_i, \tag{58}$$

$$E(\hat{Y}_{s2}) = Y, \tag{59}$$

$$var(\hat{Y}_{s2}) = \sum_{i=1}^N y_i^2 \left(\frac{\pi''_i}{(\pi'_i)^2} - 1 \right) + \sum_{i \neq j} y_i y_j \left(\frac{\pi''_{ij}}{\pi'_i \pi'_j} - 1 \right). \tag{60}$$

Notice that when the sample size is fixed, $\hat{Y}_{s2} = \hat{Y}_{HT}$.

Definition 5

A sampling measure $p(\cdot)$ corresponds to *weight-balanced* (WB) sampling, if and only if $(\pi''_i/(\pi'_i)^2)$ and $(\pi''_{ij}/\pi'_i\pi'_j)$ are constants for $i \in U$ and $i \neq j$, $i,j \in U$ respectively.

Let

$$\pi_i''/(\pi_i')^2 = \beta_1, \text{ for all } i \tag{61}$$

$$\pi_{ij}''/\pi_i'\pi_j' = \beta_2, \text{ for all } i \neq j, \ i,j \in U. \tag{62}$$

We have the following corollary of Theorem 15.

Corollary 1

Under WB sampling, we have

$$V(\hat{Y}_{sr1}) = (N-1)S^2(\beta_1 - \beta_2) + N\bar{Y}^2[(\beta_1 - \beta_2) + N(\beta_2 - 1)]. \tag{63}$$

Definition 6

A sampling measure $p(\cdot)$ corresponds to *strongly weight-balanced* (SWB) sampling if and only if π_i', π_i'', π_{ij}'' are constants for $i \in U$ and $i \neq j$, $i,j \in U$ respectively.

Let

$$\pi_i' = \beta_3 \quad i \in U, \text{ and} \tag{64}$$

$$\beta_0 = \sum_\omega p(\omega)/|\omega| = \sum_{j=1}^{N} \pi_j'' \tag{65}$$

Theorem 16

For SWB sampling, we have

$$var(\hat{Y}_{s2}) = N^2 S^2\left(\beta_0 - \frac{1}{N}\right). \tag{66}$$

Suppose $q(n) > 0$, $n = 1,\ldots,N$ and $\sum_{n=1}^{N} q(n) = 1$. Suppose we draw a sample in this way: firstly select the sample size n with probability $q(n)$, then use SRSWOR to draw a sample of size n. Then use $N\bar{y}_\omega$ to estimate the population total Y. In this way, we select a particular sample of size n with probability $q(n)/\binom{N}{n}$. We have

$$var(N\bar{y}_\omega) = \sum_{n=1}^{N} q(n)\left[(N\bar{y}_\omega - Y)^2 \big| |\omega| = n\right] \tag{67}$$

$$= \sum_{n=1}^{N} q(n)N^2 S^2\left(\frac{1}{n} - \frac{1}{N}\right)$$

$$= N^2 S^2\left(\beta_0 - \frac{1}{N}\right).$$

Hence, the technique of using $\hat{Y}_{s2}$ to estimate the population total in SWB sampling is a technique which imitates SRSWOR.

An estimator $\hat{Y}_G$ is said to be location invariant if and only if

$$\hat{Y}_G \text{ (given that } \underline{y} = \underline{y}^*) = -y_0 N + \hat{Y}_G \text{ (given that } \underline{y} = \underline{y}^* + y_0 J_{1N}) \qquad (68)$$

for all real y_0, when $\underline{y} = (y_1,\ldots,y_N)'$. It is easy to see that $\hat{Y}_G = \left(\sum_{i=1}^{N} C_{i\omega} a_{i\omega} y_i\right)$ is location invariant iff $\sum_{i=1}^{N} c_{i\omega} a_{i\omega} = N$ for all $\omega \in 2^U$.

Theorem 17

Under SWB sampling, $\hat{Y}_{s2}$ is location invariant.

The material in this section comes from Srivastava (1987), where examples of WB are given. From an unpublished paper of Srivastava and Ouyang (1988), we know that $\hat{Y}_{s2}$ is an admissible linear estimator of Y, and has a variance formula which is similar to the Yates-Grundy variance formula for $Var(\hat{Y}_{HT})$ when the sample size is fixed.

An Example of Controlled Sampling and BA Sampling

Now we discuss an example given by Avadhani and Sukhatme (1973) in controlled sampling. Let $N = 7$, and suppose these seven units are located as in the diagram below:
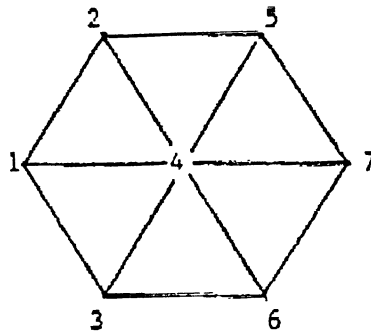
Here, any two units which are connected by a line are considered as neighbors. We are going to get a sample of size 3 from these 7 units. In order to reduce the travel cost, we hope the sample we get consists of neighboring units. So a sample $\omega = \{i_1, i_2, i_3\}$ is considered preferred if and only if after a suitable permutation, there is a line between i_1 and i_2 and also there is a line between i_2 and i_3. So the

total number of preferred samples is 21, and the total number of possible samples is $\binom{7}{3} = 35$.

Consider a BIBD with parameter $N = 7$, $k = 3$, $b = 7$, $r = 3$, $v = 1$:

$$T = \begin{bmatrix} 2 & 2 & 3 & 1 & 3 & 1 & 1 \\ 5 & 4 & 6 & 4 & 4 & 2 & 5 \\ 7 & 6 & 7 & 7 & 5 & 3 & 6 \end{bmatrix} \tag{69}$$

Now only the block correspond to column 7 is not preferred. Hence if we use probability 1/7 to draw a column from T, we reduce the probability of drawing a non-preferred sample greatly, and at the same time we have the same first two moments as SRSWOR. But this technique does not avoid the nonpreferred samples totally. To avoid the nonpreferred samples totally, consider a balanced array approach as follows. We have a list of 16 samples: {147}, {246}, {543}; {125}, {257}, {576}, {763}, {631}, {321}, {15}, {27}, {56}, {73}, {61}, {32}; {5}. With probability (1/11) we draw any one of the first three samples, and with probability (1/22) we draw any one of the remaining samples. Hence we avoid the nonpreferred samples. But we use some subsamples of the preferred samples.

The problem of controlled sampling may be approached through the concepts of array sampling as follows.

(i)   Decide the preferred and nonpreferred samples.

(ii)   Decide whether fixed sample size should be used or not.

(iii)   Consider using BA sampling or WB sampling.

(iv)   Suppose BA sampling is used. Then we need to find a BA whose columns consist of the preferred samples. If we fail to get such a BA, then consider subsamples of these samples. Sometimes we have to change the decision in step (ii) to consider using some non-preferred samples in this step (with minimal probability).

References

Avadhani, M. S. and Sukhatme, B. V. (1963): Controlled simple random sampling, *J. India Soc. Agricultural Stat.* 17, 34-42.

Avadhani, M. S. and Sukhatme, B. V. (1967): Controlled sampling with varying probability with and without replacement, *Austral. J. Stat.* 9(1), 8-15.

Avadhani, M. S. and Sukhatme, B. V. (1973): Controlled sampling with equal probabilities and without replacement, *International Stat. Review* 41(2), 175-182.

Brewer, K. R. W., Foreman, E. K., Mellor, R. W. and Trewin, D. J. (1977): Use of experimental design and population modelling in survey sampling, *Bulletin of the International Stat. Inst.* 3, 173-190.

Brewer, K. R. W. and Hanif, M. (1983): *Sampling with Unequal Probabilities*, Springer-Verlag, New York, Heidelberg, Berlin.

Brunk, M. E. and Federer, W. T. (1953): Experimental designs and probability sampling in marking research, *JASA* 48, 440-452.

Cassel, C. M., Sarndal, C. E. and Wretman, J. H. (1977): *Foundations of Inference in Survey Sampling*, Wiley, New York.

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

Cheng, C. S. and Li, K. C. (1983): A minimax approach to sample surveys, *Annals of Statistics* 11, 552-563.

Cochran, W. G. (1977): *Sampling Techniques*, 3rd edition, Wiley, New York.

Cox, D. R. (1956): A note on weighted randomization, *Annals of Math. Stat.* 27, 1144-1151.

Cox, D. R. (1958): *Planning of Experiments*, Wiley, New York.

Cumberland, W. G. and Royal, R. M. (1981): Prediction models and unequal probability sampling, *J. Roy. Stat. Soc.*, Ser. B, 43(3), 353-367.

Deming, W. E. (1953): On the distinction between enumerative and analytic survey, *JASA* 48, 244-255.

Dharmadhikari, S. W. (1982): Connectedness and zero variance in sampling designs, *Stat. and Prob.*, 221-225.

Dwyer, P. S. (1972): Moment functions of sample moment functions, in *Symmetric Functions in Statistics*, ed. Derrick S. Tracy, University of Windsor, Windsor, Ontario, 11-51.

Fellegi, I. P. (1963): Sampling with varying probabilities without replacement: Rotating and rotating samples, *JASA* 58, 183-201.

Kalton, G. (1983): Models in the practice of survey sampling, *International Stat. Review* 51, 175-188.

Kiefer, J. (1961): Optimum designs in regression problems. II, *Annals of Math. Stat.* 321, 298-325.

Kempthorne, O. (1952): *The Design and Analysis of Experiments*, Wiley, New York.

Kempthorne, O. (1955): The randomization theory of experimental inference, *JASA* 50, 946-967.

Kish, L. (1965): *Survey Sampling*, Wiley, New York.

Kruskal, W. H. and Mosteller, F. (1980): Representative sampling IV: The history of the concept in statistics. 1815-1939, *International Stat. Review* 48, 169-195.

Lacayo, H., Pereina, C. A. de, Proschan, F. and Sarndal, C. F. (1982): Optimal sample depends on optimality criterion, *Scand. J. Stat.* 9, 47-48.

Little, R. J. A. (1983): Estimating a finite population mean from unequal probability sample, *JASA* 78(3), 596-604.

Lin, T. P. and Thompson, M. E. (1983): Journal of quadratic finite population: The batch approach, *Annals of Stat.* II(1), 275-285.

Mahalanobis, P. C. (1944): On large-scale sample surveys, *Philosophical Transactions of the Roy. Soc.*, London, 231(B), 329-451.

Meeden, G. and Ghosh, M. (1983): Choosing between experiments: Applications to finite population sampling, *Annals of Stat.* 11, 296-305.

Mikhail, N. N. and Ali, M. M. (1981): Unbiased estimates of the generalized variance for finite population, *J. Indian Stat. Assoc.* 19, 85-92.

Mukhopadhyay, P. (1975): An optimum sampling design to the HT method of estimating a population total, *Metrika* 22, 119-127.

Mukhopadhyay, P. (1982): Optimal strategies for estimating the variance of a finite population under a superpopulation model, *Metrika* 29, 143-158.

Murthy, M. N.: Sampling theory and methods, *Stat. Pub. Soc.*, Calcutta.

Padmawar, V. R. (1981): A note on the comparison of certain sampling strategies, *J. Roy. Stat. Soc.*, Ser. B, 43(3), 321-326.

Pathak, P. K. (1964): On inverse sampling with unequal probabilities, *Biometrika* 51, 185-193.

Pereira, C. A. B. and Rodriques, J. (1983): Robust linear prediction in finite populations, *International Stat. Review* 51, 293-300.

Prasad, N. G. N. and Srivenkataramana, T (1980): A modification to the Horvitz-Thompson estimator under the Midzuno sampling scheme, *Biometrika* 67, 709-711.

Raj, D. (1958): On the relation accuracy of some sampling techniques, *JASA* 53, 98-101.

Raj, D. (1972): *Sampling Theory*, McGraw-Hill Book Company, New York.

Ramakrisknan, M. K. (1975a): Choice of an optimum sampling strategy—I, *Annals of Stat.* 3(3), 669-679.

Ramakrishnan, M. K. (1975b): A generalization of the Yates-Grundy variance estimator, *Sankhya*, Ser. C, 37, 204-206.

Rao, C. R. (1975): Some problems of sample surveys, *Adv. Appd. Prob.* 7 (Supplement), 50-61.

Robinson, P M. (1982): On the convergence of the Horvitz-Thompson estimator, *Austral. J. Stat.* 24, 234-238.

Rosen, B. (1972): Asymptotic theory for successive sampling with varying probabilities without replacement I and II, *Annals of Math. Stat.* 43, 7, 373-392, 748-776.

Roy, J. (1957): A note on estimation of variance components in multistage sampling with varying probabilities, *Sankhya* 17, 367-372.

Rubin, D. B. (1978): Bayesian inference for causal effects: The role of randomization, *Annals of Stat.* 6, 34-58.

Sedransk, J. (1967): Designing some multi-factor analytical studies, *JASA* 62, 1121-1139.

Srivastava, J. (1985): On a general theory of sampling, using experimental design. Concept I: Estimation, *Bulletin of International Stat. Inst.*, Vol. 51, Book 2, 10.3-1 - 10.3-16.

Srivastava, J. (1988): On a general theory of sampling, using experimental design. Concepts II: Relation with arrays, in *Probability and Statistics*, ed. J. N. Srivastava, North-Holland.

Srivastava, J. and Ouyang, Z. (1988a): Studies on the general estimator in sampling theory, based on the sample weight function, unpublished paper.

Srivastava, J. and Ouyang, Z. (1988b): Optimal properties of balanced array sampling and weight balanced sampling, unpublished paper.

Srivastava, J. and Saleh, F. (1985): Need of t-designs in sampling theory, *Utilitas Mathematica* 28, 5-17.

Stenger, H. and Gabler, S. (1981): On the completeness of the class of fixed size sampling strategies, *Annals of Stat.* 9, 229-232.

Strauss, I. (1982): On the admissibility of estimators for the finite population variance, *Metrika* 29, 195-202.

Sukhatme, P. V. and Sukhatme, B. V. (1976): Sampling theory of surveys with applications, New Dehli, *Indian Soc. of Agricultural Stat.*

Tepping, B. J., Hurvitz, W. N. and Deming, W. E. (1943): On the efficiency of deep stratification in block sampling, *JASA* 38, 93-100.

Vardeman, S. and Meeden, G. (1983): Admissible estimators in finite population sampling employing types of prior information, *JSPI* 7, 329-341.

Wilk, M. B. and Kempthorne, O. (1956): Some aspects of the analysis of factorial experiments in a completely randomized design, *Annals of Math. Stat.* 27, 950-985.

Wynn, H. P. (1976): Optimum designs for finite populations sampling, in *Statistical Decision Theory and Related Topics*, eds. S. S. Gupta and D. S. Moore, Academic Press, New York.

Wynn, H. P. (1977a): Minimax purposive survey sampling design, *JASA* 72(359), 655-657.

Wynn, H. P. (1977b): Convex set of finite population plans, *Annals of Stat.* 5, 414-418.