TECHNOMETRICS ©, VOL. 21, NO. 2, MAY 1979

A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code

M. D. McKay and R. J. Beckman
Los Alamos Scientific Laboratory
P.O. Box 1663
Los Alamos, NM 87545

W. J. Conover
Department of Mathematics
Texas Tech University
Lubbock, TX 79409

Two types of sampling plans are examined as alternatives to simple random sampling in Monte Carlo studies. These plans are shown to be improvements over simple random sampling with respect to variance for a class of estimators which includes the sample mean and the empirical distribution function.

KEY WORDS

Latin hypercube sampling
Sampling techniques
Simulation techniques
Variance reduction

1. INTRODUCTION

Numerical methods have been used for years to provide approximate solutions to fluid flow problems that defy analytical solutions because of their complexity. A mathematical model is constructed to resemble the fluid flow problem, and a computer program (called a "code"), incorporating methods of obtaining a numerical solution, is written. Then for any selection of input variables $\mathbf{X} = (X_1, \cdots, X_K)$ an output variable $Y = h(\mathbf{X})$ is produced by the computer code. If the code is accurate the output Y resembles what the actual output would be if an experiment were performed under the conditions $\mathbf{X}$. It is often impractical or impossible to perform such an experiment. Moreover, the computer codes are sometimes sufficiently complex so that a single set of input variables may require several hours of time on the fastest computers presently in existence in order to produce one output. We should mention that a single output Y is usually a graph $Y(t)$ of output as a function of time, calculated at discrete time points t, $t_0 \leq t \leq t_1$.

When modeling real world phenomena with a computer code one is often faced with the problem of what values to use for the inputs. This difficulty can arise from within the physical process itself when system parameters are not constant, but vary in some manner about nominal values. We model our uncertainty about the values of the inputs by treating them as random variables. The information desired from the code can be obtained from a study of the probability distribution of the output $Y(t)$. Consequently, we model the "numerical" experiment by $Y(t)$ as an unknown transformation $h(\mathbf{X})$ of the inputs $\mathbf{X}$, which have a *known* probability distribution $F(\mathbf{x})$ for $\mathbf{x} \in S$. Obviously several values of $\mathbf{X}$, say $\mathbf{X}_1, \cdots, \mathbf{X}_N$, must be selected as successive inputs sets in order to obtain the desired information concerning $Y(t)$. When N must be small because of the running time of the code, the input variables should be selected with great care.

The next section describes three methods of selecting (sampling) input variables. Sections 3, 4 and 5 are devoted to comparing the three methods with respect to their performance in an actual computer code.

The computer code used in this paper was developed in the Hydrodynamics Group of the Theoretical Division at the Los Alamos Scientific Laboratory, to study reactor safety [8]. The computer code is named SOLA-PLOOP and is a one-dimensional version of another code SOLA [7]. The code was used by us to model the blowdown depressurization of a straight pipe filled with water at fixed initial temperature and pressure. Input variables include: X_1, phase change rate; X_2, drag coefficient for drift velocity; X_3, number of bubbles per unit volume; and X_4, pipe roughness. The input variables are assumed to be uniformly distributed over given ranges. The output variable is pressure as a function of time, where the initial time t_0 is the time the pipe ruptures and depressurization

Received January 1977; revised May 1978

initiates, and the final time t_1 is 20 milliseconds later. The pressure is recorded at 0.1 millisecond time intervals. The code was used repeatedly so that the accuracy and precision of the three sampling methods could be compared.

2. A DESCRIPTION OF THE THREE METHODS USED FOR SELECTING THE VALUES OF INPUT VARIABLES

From the many different methods of selecting the values of input variables, we have chosen three that have considerable intuitive appeal. These are called random sampling, stratified sampling, and Latin hypercube sampling.

Random Sampling. Let the input values $\mathbf{X}_1, \cdots, \mathbf{X}_N$ be a random sample from $F(\mathbf{x})$. This method of sampling is perhaps the most obvious, and an entire body of statistical literature may be used in making inferences regarding the distribution of $Y(t)$.

Stratified Sampling. Using stratified sampling, all areas of the sample space of $\mathbf{X}$ are represented by input values. Let the sample space S of $\mathbf{X}$ be partitioned into I disjoint strata S_i. Let $p_i = P(\mathbf{X} \in S_i)$ represent the size of S_i. Obtain a random sample $X_{ij}, j = 1, \cdots, n_i$ from S_i. Then of course the n_i sum to N. If $I = 1$, we have random sampling over the entire sample space.

Latin Hypercube Sampling. The same reasoning that led to stratified sampling, ensuring that all portions of S were sampled, could lead further. If we wish to ensure also that each of the input variables X_k has all portions of its distribution represented by input values, we can divide the range of each X_k into N strata of equal marginal probability $1/N$, and sample once from each stratum. Let this sample be $X_{kj}, j = 1, \cdots, N$. These form the X_k component, $k = 1, \cdots, K$, in $\mathbf{X}_i$, $i = 1, \cdots, N$. The components of the various X_k's are matched at random. This method of selecting input values is an extension of quota sampling [13], and can be viewed as a K-dimensional extension of Latin square sampling [11].

One advantage of the Latin hypercube sample appears when the output $Y(t)$ is dominated by only a few of the components of $\mathbf{X}$. This method ensures that each of those components is represented in a fully stratified manner, no matter which components might turn out to be important.

We mention here that the N intervals on the range of each component of $\mathbf{X}$ combine to form N^K cells which cover the sample space of $\mathbf{X}$. These cells, which are labeled by coordinates corresponding to the intervals, are used when finding the properties of the sampling plan.

2.1 Estimators

In the Appendix (Section 8), stratified sampling and Latin hypercube sampling are examined and compared to random sampling with respect to the class of estimators of the form

$$T(Y_1, \cdots, Y_N) = (1/N) \sum_{i=1}^{N} g(Y_i),$$

where $g(\cdot)$ = arbitrary function.

If $g(Y) = Y$ then T represents the sample mean which is used to estimate $E(Y)$. If $g(Y) = Y^r$ we obtain the rth sample moment. By letting $g(Y) = 1$ for $Y \leq y$, 0 otherwise, we obtain the usual empirical distribution function at the point y. Our interest is centered around these particular statistics.

Let τ denote the expected value of T when the Y_i's constitute a random sample from the distribution of $Y = h(\mathbf{X})$. We show in the Appendix that both stratified sampling and Latin hypercube sampling yield unbiased estimators of τ.

If T_R is the estimate of τ from a random sample of size N, and T_S is the estimate from a stratified sample of size N, then $\mathrm{Var}(T_S) \leq \mathrm{Var}(T_R)$ when the stratified plan uses equal probability strata with one sample per stratum (all $p_i = 1/N$ and $n_{ij} = 1$). No direct means of comparing the variance of the corresponding estimator from Latin hypercube sampling, T_L, to $\mathrm{Var}(T_S)$ has been found. However, the following theorem, proved in the Appendix, relates the variances of T_L and T_R.

Theorem. If $Y = h(X_1, \cdots, X_K)$ is monotonic in each of its arguments, and $g(Y)$ is a monotonic function of Y, then $\mathrm{Var}(T_L) \leq \mathrm{Var}(T_R)$.

2.2 The SOLA-PLOOP Example

The three sampling plans were compared using the SOLA-PLOOP computer code with $N = 16$. First a random sample consisting of 16 values of $\mathbf{X} = (X_1, X_2, X_3, X_4)$ was selected, entered as inputs, and 16 graphs of $Y(t)$ were observed as outputs. These output values were used in the estimators.

For the stratified sampling method the range of each input variable was divided at the median into two parts of equal probability. The combinations of ranges thus formed produced $2^4 = 16$ strata S_i. One observation was obtained at random from each S_i as input, and the resulting outputs were used to obtain the estimates.

To obtain the Latin hypercube sample the range of each input variable X_i was stratified into 16 intervals of equal probability, and one observation was drawn at random from each interval. These 16 values for the 4 input variables were matched at random to form 16 inputs, and thus 16 outputs from the code.

The entire process of sampling and estimating for the three selection methods was repeated 50 times in order to get some idea of the accuracies and precisions involved. The total computer time spent in running the SOLA-PLOOP code in this study was 7 hours on a CDC-6600. Some of the standard devia-

tion plots appear to be inconsistent with the theoretical results. These occasional discrepancies are believed to arise from the non-independence of the estimators over time and the small sample sizes.

3. ESTIMATING THE MEAN

The goodness of an unbiased estimator of the mean can be measured by the size of its variance. For each sampling method, the estimator of $E(Y(t))$ is of the form

$$\bar{Y}(t) = (1/N) \sum_{i=1}^{N} Y_i(t) \tag{3.1}$$

where

$$Y_i(t) = h(\mathbf{X}_i), \qquad i = 1, \cdots, N.$$

In the case of the stratified sample, the $\mathbf{X}_i$ comes from stratum S_i, $p_i = 1/N$ and $n_i = 1$. For the Latin hypercube sample, the $\mathbf{X}_i$ is obtained in the manner described earlier. Each of the three estimators $\bar{Y}_R$, $\bar{Y}_S$, and $\bar{Y}_L$ is an unbiased estimator of $E(Y(t))$. The variances of the estimators are given in (3.2):

$$\mathrm{Var}(\bar{Y}_R(t)) = (1/N)\,\mathrm{Var}(Y(t))$$

$$\mathrm{Var}(\bar{Y}_S(t)) = \mathrm{Var}(\bar{Y}_R(t)) - (1/N^2) \sum_{i=1}^{N} (\mu_i - \mu)^2$$

$$\mathrm{Var}(\bar{Y}_L(t)) = \mathrm{Var}(\bar{Y}_R(t)) + ((N-1)/N) \cdot 1/(N^K(N-1)^K)) \sum_{R} (\mu_i - \mu)(\mu_j - \mu) \tag{3.2}$$

where $\mu = E(Y(t))$,

$\mu_i = E(Y(t) \mid \mathbf{X} \in S_i)$ in the stratified sample, or

$\mu_i = E(Y(t) \mid \mathbf{X} \in \text{cell } i)$ in the Latin hypercube sample,

and R means the restricted space of all pairs μ_i, μ_j having no cell coordinates in common.

For the SOLA-PLOOP computer code the means and standard deviations, based on 50 observations, were computed for the estimators just described. Comparative plots of the means are given in Figure 1. All of the plots of the means are comparable, demonstrating the unbiasedness of the estimators.

Comparative plots of the standard deviations of the estimators are given in Figure 2. The standard deviation of $\bar{Y}_S(t)$ is smaller than that of $\bar{Y}_R(t)$ as expected. However, $\bar{Y}_L(t)$ clearly demonstrates superiority as an estimator in this example, with a standard deviation roughly one-forth that of the random sampling estimator.

4. ESTIMATING THE VARIANCE

For each sampling method, the form of the estimator of the variance is

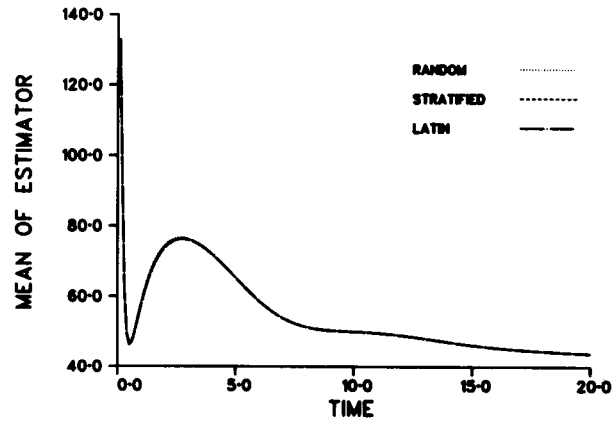

FIGURE 1. Estimating the mean: the sample mean of $\bar{Y}_R(t)$, $\bar{Y}_S(t)$, and $\bar{Y}_L(t)$.

$$S^2(t) = (1/N) \sum_{i=1}^{N} (Y_i(t) - \bar{Y}(t))^2, \tag{4.1}$$

and its expectation is

$$E(S^2(t)) = \mathrm{Var}(Y(t)) - \mathrm{Var}(\bar{Y}(t)), \tag{4.2}$$

where $\bar{Y}(t)$ is one of $\bar{Y}_R(t)$, $\bar{Y}_S(t)$, or $\bar{Y}_L(t)$.

In the case of the random sample, it is well known that $N\,S^2_R/(N-1)$ is an unbiased estimator of the variance of $Y(t)$. The bias in the case of the stratified sample is unknown. However, because $\mathrm{Var}(\bar{Y}_S(t)) \leq \mathrm{Var}(\bar{Y}_R(t))$,

$$(1 - 1/N)\,\mathrm{Var}(Y(t)) \leq E(S_s^2(t)) \leq \mathrm{Var}(Y(t)). \tag{4.3}$$

The bias in the Latin hypercube plan is also unknown, but for the SOLA-PLOOP example it was small. Variances for these estimators were not found.

Again using the SOLA-PLOOP example, means and standard deviations (based on 50 observations) were computed. The mean plots are given in Figure 3. They indicate that all three estimators are in relative agreement concerning the quantities they are estimating. In terms of standard deviations of the estimators, Figure 4 shows that, although stratified sampling yields about the same precision as does random sam-

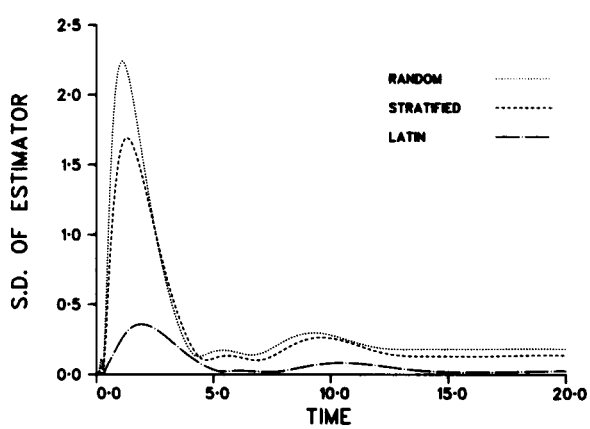

FIGURE 2. Estimating the mean: the standard deviation of $\bar{Y}_R(t)$, $\bar{Y}_S(t)$, and $\bar{Y}_L(t)$.

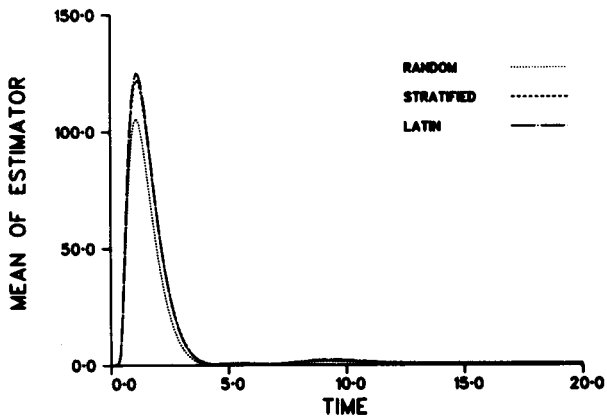

FIGURE 3. Estimating the variance: the sample mean of $S_R^2(t)$, $S_S^2(t)$, and $S_L^2(t)$.

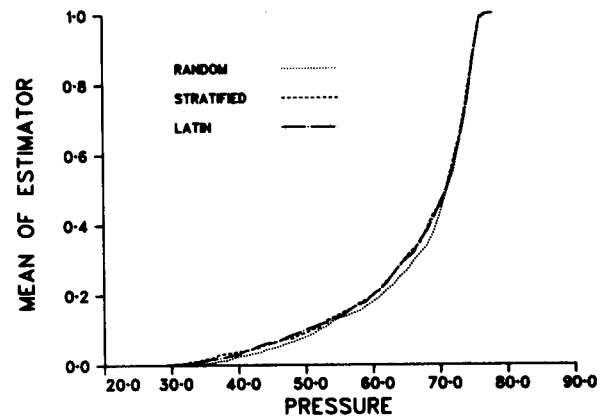

FIGURE 5. Estimating the CDF: the sample mean of $G_R(y, t)$, $G_S(y, t)$, and $G_L(y, t)$ at $t = 1.4$.

pling, Latin hypercube furnishes a clearly better estimator.

5. ESTIMATING THE DISTRIBUTION FUNCTION

The distribution function, $D(y, t)$, of $Y(t) = h(\mathbf{X})$ may be estimated by the empirical distribution function. The empirical distribution function can be written as

$$G(y, t) = (1/N) \sum_{i=1}^{N} u(y - Y_i(t)), \qquad (5.1)$$

where $u(z) = 1$ for $z > 0$ and is zero otherwise. Since equation (5.1) is of the form of the estimators in Section 2.1, the expected value of $G(y, t)$ under the three sampling plans is the same, and under random sampling, the expected value of $G(y, t)$ is $D(y, t)$.

The variances of the three estimators are given in (5.2). D_i again refers to either stratum i or cell i, as appropriate, and R represents the same restricted space as it did in (3.2).

$$\mathrm{Var}(G_R(y, t)) = (1/N)\, D(y, t)(1 - D(y, t))$$

$$\mathrm{Var}(G_S(y, t)) = \mathrm{Var}(G_R(y, t)) - (1/N^2) \sum_{i=1}^{N} (D_i(y, t) - D(y, t))^2$$

$$\mathrm{Var}(G_L(y, t)) = \mathrm{Var}(G_R(y, t)) + ((N-1)/N \cdot 1/N^K(N-1)^K) \sum_{R} (D_i(y, t) - D(y, t)) \cdot (D_j(y, t) - D(y, t)). \qquad (5.2)$$

As with the cases of the mean and variance estimators, the distribution function estimators were compared for the three sampling plans. Figures 5 and 6 give the means and standard deviations of the estimators at $t = 1.4$ ms. This time point was chosen to correspond to the time of maximum variance in the distribution of $Y(t)$. Again the estimates obtained from a Latin hypercube sample appear to be more precise in general than the other two types of estimates.

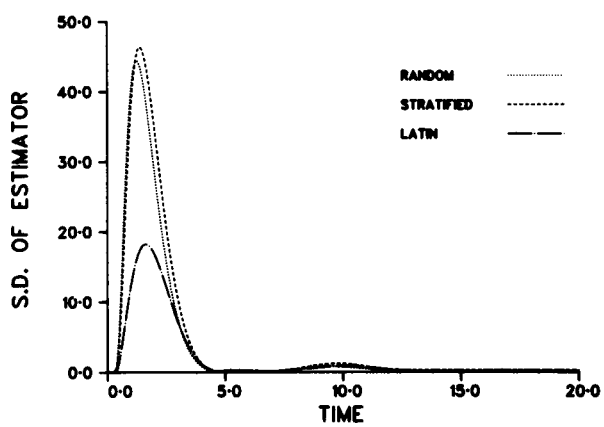

FIGURE 4. Estimating the variance: the standard deviation of $S_R^2(t)$, $S_S^2(t)$, and $S_L^2(t)$.

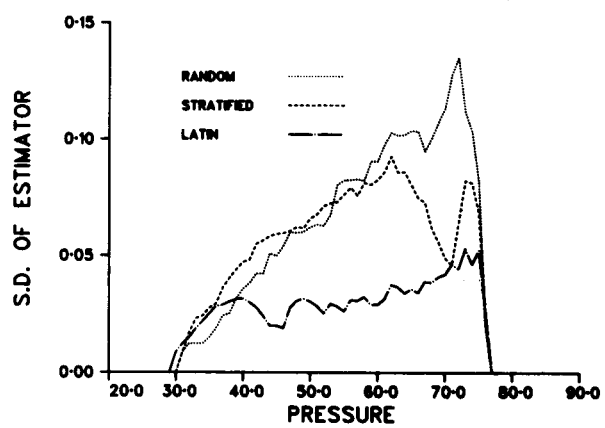

FIGURE 6. Estimating the CDF: the standard deviation of $G_R(y, t)$, $G_S(y, t)$, and $G_L(y, t)$ at $t = 1.4$.

6. DISCUSSION AND CONCLUSIONS

We have presented three sampling plans and associated estimators of the mean, the variance, and the population distribution function of the output of a computer code when the inputs are treated as random variables. The first method is simple random sampling. The second method involves stratified sampling and improves upon the first method. The third method is called here Latin hypercube sampling. It is an extension of quota sampling [13], and it is a first cousin to the "random balance" design discussed by Satterthwaite [12], Budne [2], Youden, et al [15], Anscombe [1], and to the highly fractionalized factorial designs discussed by Enrenfeld and Zacks [5, 6], Dempster [3, 4], and Zacks [16, 17], and to lattice sampling as discussed by Jessen [9]. This third method improves upon simple random sampling when certain monotonicity conditions hold, and it appears to be a good method to use for selecting values of input variables.

7. ACKNOWLEDGMENTS

We extend a special thanks to Ronald K. Lohrding, for his early suggestions related to this work and for his continuing support and encouragement. We also thank our colleagues Larry Bruckner, Ben Duran, C. Phive, and Tom Boullion for their discussions concerning various aspects of the problem, and Dave Whiteman for assistance with the computer.

This paper was prepared under the support of the Analysis Development Branch, Division of Reactor Safety Research, Nuclear Regulatory Commission.

8. APPENDIX

In the sections that follow we present some general results about stratified sampling and Latin hypercube sampling in order to make comparisons with simple random sampling. We move from the general case of stratified sampling to stratified sampling with proportional allocation, and then to proportional allocations with one observation per stratum. We examine Latin hypercube sampling for the equal marginal probability strata case only.

8.1 *Type I Estimators*

Let $\mathbf{X}$ denote a K variate random variable with probability density function (*pdf*) $f(\mathbf{x})$ for $\mathbf{x} \in S$. Let Y denote a univariate transformation of $\mathbf{X}$ given by $Y = h(\mathbf{X})$. In the context of this paper we assume

$\mathbf{X} \sim f(\mathbf{x}),\ \mathbf{x} \in S$	KNOWN *pdf*
$Y = h(\mathbf{X})$	UNKNOWN but observable transformation of $\mathbf{X}$.

The class of estimators to be considered are those of the form

$$T(u_1, \cdots, u_N) = (1/N) \sum_{i=1}^{N} g(u_i), \qquad (8.1)$$

where $g(\cdot)$ is an arbitrary, known function. In particular we use $g(u) = u^r$ to estimate moments, and $g(u) = 1$ for $u \geq 0$, $= 0$ elsewhere, to estimate the distribution function.

The sampling schemes described in the following sections will be compared to random sampling with respect to T. The symbol T_R denotes $T(Y_1, \cdots, Y_N)$ when the arguments $Y_1, \cdots, Y_N$ constitute a random sample of Y. The mean and variance of T_R are denoted by τ and θ^2/N. The statistic T given by (8.1) will be evaluated at arguments arising from stratified sampling to form T_S, and at arguments arising from Latin hypercube sampling to form T_L. The associated means and variances will be compared to those for random sampling.

8.2 *Stratified Sampling*

Let the range space, S, of $\mathbf{X}$ be partitioned into I disjoint subsets S_i of size $p_i = P(\mathbf{X} \in S_i)$, with

$$\sum_{i=1}^{I} p_i = 1.$$

Let X_{ij}, $j = 1, \cdots, n_i$, be a random sample from stratum S_i. That is, let $X_{ij} \sim iid\, f(\mathbf{x})/p_i, j = 1, \cdots, n_i$, for $\mathbf{x} \in S_i$, but with zero density elsewhere. The corresponding values of Y are denoted by $Y_{ij} = h(X_{ij})$, and the strata means and variances of $g(Y)$ are denoted by

$$\mu_i = E(g(Y_{ij})) = \int_{S_i} g(y)(1/p_i)f(\mathbf{x})d\mathbf{x}$$

$$\sigma_i^2 = \mathrm{Var}(g(Y_{ij})) = \int_{S_i} (g(y)-\mu_i)^2(1/p_i)f(\mathbf{x})d\mathbf{x}.$$

It is easy to see that if we use the general form

$$T_S = \sum_{i=1}^{I} (p_i/n_i) \sum_{j=1}^{n_i} g(Y_{ij}),$$

that T_S is an unbaised estimator of τ with variance given by

$$\mathrm{Var}(T_S) = \sum_{i=1}^{I} (p_i^2/n_i)\sigma_i^2. \qquad (8.2)$$

The following results can be found in Tocher [14].

Stratified Sampling with Proportional Allocation. If the probability sizes, p_i, of the strata and the sample sizes, n_i, are chosen so that $n_i = p_i N$, proportional allocation is achieved. In this case (8.2) becomes

$$\mathrm{Var}(T_S) = \mathrm{Var}(T_R) - (1/N) \sum_{i=1}^{I} p_i(\mu_i - \tau)^2. \qquad (8.3)$$

Thus, we see that stratified sampling with proportional allocation offers an improvement over random sampling, and that the variance reduction is a function of the differences between the strata means μ_i and the overall mean τ.

Proportional Allocation with One Sample per Stratum. Any stratified plan which employs subsampling, $n_i > 1$, can be improved by further stratification. When all $n_i = 1$, (8.3) becomes

$$\mathrm{Var}(T_S) = \mathrm{Var}(T_R) - (1/N^2) \sum_{i=1}^{N} (\mu_i - \tau)^2. \quad (8.4)$$

8.3 Latin Hypercube Sampling

In stratified sampling the range space S of $\mathbf{X}$ can be arbitrarily partitioned to form strata. In Latin hypercube sampling the partitions are constructed in a specific manner using partitions of the ranges of each component of $\mathbf{X}$. We will only consider the case where the components of $\mathbf{X}$ are independent.

Let the ranges of each of the K components of $\mathbf{X}$ be partitioned into N intervals of probablitity size $1/N$. The Cartesian product of these intervals partitions S into N^K cells each of probability size N^{-K}. Each cell can be labeled by a set of K cell coordinates $\mathbf{m}_i = (m_{i1}, m_{i2}, \cdots, m_{iK})$ where m_{ij} is the interval number of component X_j represented in cell i. A Latin hypercube sample of size N is obtained from a random selection N of the cells $\mathbf{m}_1, \cdots, \mathbf{m}_N$, with the condition that for each j the set $\{m_{ij}\}_{i=1}^{N}$ is a permutation of the integers $1, \cdots, N$. One random observation is made in each cell. The density function of $\mathbf{X}$ given $\mathbf{X} \in$ cell i is $N^K f(x)$ if $\mathbf{x} \in$ cell i, zero otherwise. The marginal (unconditional) distribution of $Y_i(t)$ is easily seen to be the same as that for a randomly drawn $\mathbf{X}$ as follows:

$$\begin{aligned} P(Y_i \le y) &= \sum_{\text{all cells } q} P(Y_i \le y \mid \mathbf{X} \in \text{cell } q) P(\mathbf{X} \in \text{cell } q) \\ &= \sum \int_{\substack{\text{cell } q \\ h(\mathbf{x}) \le y}} N^K f(\mathbf{x}) d\mathbf{x}\, (1/N^K) \\ &= \int_{h(\mathbf{x}) \le y} f(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

From this we have T_L as an unbiased estimator of τ.

To arrive at a form for the variance of T_L we introduce indicator variables w_i, with

$$w_i = \begin{cases} 1 & \text{if cell } i \text{ is in the sample} \\ 0 & \text{if not.} \end{cases}$$

The estimator can now be written as

$$T_L = (1/N) \sum_{i=1}^{N^K} w_i g(Y_i), \quad (8.5)$$

where $Y_i = h(\mathbf{X}_i)$ and $\mathbf{X}_i \in$ cell i. The variance of T_L is given by

$$\begin{aligned} \mathrm{Var}(T_L) &= (1/N^2) \sum_{i=1}^{N^K} \mathrm{Var}(w_i g(Y_i)) \\ &\quad + (1/N^2) \sum_{i=1}^{N^K} \sum_{\substack{j=1 \\ j \ne i}}^{N^K} \mathrm{Cov}(w_i g(Y_i), w_j g(Y_j)). \quad (8.6) \end{aligned}$$

The following results about the w_i are immediate:

1. $P(w_i = 1) = (1/N^{K-1}) = E(w_i) = E(w_i^2)$
   $\mathrm{Var}(w_i) = (1/N^{K-1})(1 - 1/N^{K-1})$.
2. If w_i and w_j correspond to cells having no cell coordinates in common, then
$$\begin{aligned} E(w_i w_j) &= E(w_i w_j \mid w_j = 0) P(w_j = 0) \\ &\quad + E(w_i w_j \mid w_j = 1) P(w_j = 1) \\ &= 1/(N(N-1))^{K-1} \end{aligned}$$
3. If w_i and w_j correspond to cells having at least one common cell coordinate, then
$$E(w_i w_j) = 0.$$

Now

$$\begin{aligned} \mathrm{Var}(w_i g(Y_i)) &= E(w_i^2)\, \mathrm{Var}\, g(Y_i) \\ &\quad + E^2(g(Y_i))\, \mathrm{Var}(w_i) \quad (8.7) \end{aligned}$$

so that

$$\begin{aligned} \sum_{i=1}^{N^K} \mathrm{Var}(w_i g(Y_i)) &= N^{-K+1} \sum_{i=1}^{N^K} E(g(Y_i) - \mu_i)^2 \\ &\quad + (N^{-K+1} - N^{-2K+2}) \sum_{i=1}^{N^K} \mu_i^2 \quad (8.8) \end{aligned}$$

where $\mu_i = E\{g(Y) \mid \mathbf{X} \in \text{cell } i\}$. Since

$$E(g(Y_i) - \mu_i)^2 = N^K \int_{\text{cell } i} (g(y) - \tau)^2 f(x) dx + (\mu_i - \tau)^2 \quad (8.9)$$

we have

$$\begin{aligned} \sum_i \mathrm{Var}(w_i g(Y_i)) &= N\, \mathrm{Var}(Y) - N^{-K+1} \sum_i (\mu_i - \tau)^2 \\ &\quad + (N^{-K+1} - N^{-2K+2}) \sum_i \mu_i^2. \quad (8.10) \end{aligned}$$

Furthermore

$$\begin{aligned} \sum_{\substack{i=1 \\ i \ne j}}^{N^K} \sum_{j=1}^{N^K} \mathrm{Cov}(w_i g(Y_i), w_j g(Y_j)) &= \sum\sum_{i \ne j} \mu_i \mu_j E\{w_i w_j\} \\ &\quad - N^{-2K+2} \sum\sum_{i \ne j} \mu_i \mu_j \quad (8.11) \end{aligned}$$

which combines with (8.10) to give

$$\begin{aligned} \mathrm{Var}(T_L) &= (1/N)\mathrm{Var}(Y) - N^{-K-1} \sum_i (\mu_i - \tau)^2 \\ &\quad + (N^{-K-1} - N^{-2K}) \sum_i \mu_i^2 \\ &\quad + (N-1)^{-K+1} N^{K-1} \sum\sum_R \mu_i \mu_j \\ &\quad - N^{-2K} \sum\sum_{i \ne j} \mu_i \mu_j \quad (8.12) \end{aligned}$$

where R means the restricted space of $N^K(N-1)^K$ pairs (μ_i, μ_j) corresponding to cells having no cell coordinates in common. After some algebra, and

with $\sum \mu_i = N^K \tau$, the final form for $\mathrm{Var}(T_L)$ becomes

$$\mathrm{Var}(T_L) = \mathrm{Var}(T_R) + (N-1)/N[N^{-K}(N-1)^{-K} \cdot \sum_R (\mu_i - \tau)(\mu_j - \tau)]. \quad (8.13)$$

Note that $\mathrm{Var}(T_L) \leq \mathrm{Var}(T_R)$ if and only if

$$N^{-K}(N-1)^{-K} \sum_R (\mu_i - \tau)(\mu_j - \tau) \leq 0, \quad (8.14)$$

which is equivalent to saying that the covariance between cells having no cell coordinates in common is negative. A sufficient condition for (8.14) to hold is given by the following theorem.

THEOREM. If $Y = h(X_1, \cdots, X_K)$ is monotonic in each of its arguments, and if $g(Y)$ is a monotonic function of Y, then $\mathrm{Var}(T_L) \leq \mathrm{Var}(T_R)$.

PROOF. The proof employs a theorem by Lehmann [10]. Two functions $r(x_1, \cdots, x_K)$ and $s(y_1, \cdots, y_K)$ are said to be *concordant* in each argument if r and s either increase or decrease together as a function of $x_i = y_i$, with all $x_j, j \neq i$ and $y_j, j \neq i$ held fixed, for each i. Also, a pair of random variables (X, Y) are said to be *negatively quadrant dependent* if $P(X \leq x, Y \leq y) \leq P(X \leq x)P(Y \leq y)$ for all x, y. Lehmann's theorem states that if (*i*) $(X_1, Y_1), (X_2, Y_2), \cdots, (X_K, Y_K)$ are independent, (*ii*) (X_i, Y_i) is negatively quadrant dependent for all i, and (*iii*) $X = r(X_1, \cdots, X_K)$ and $Y = s(Y_1, \cdots, Y_K)$ are concordant in each argument, then (X, Y) is negatively quadrant dependent.

We earlier described a stage-wise process for selecting cells for a Latin hypercube sample, where a cell was labeled by cell coordinates $m_1, \cdots, m_K$. Two cells $(l_1, \cdots, l_K)$ and $(m_1, \cdots, m_K)$ with no coordinates in common may be selected as follows. Randomly select two integers (R_{11}, R_{21}) without replacement from the first N integers $1, \cdots, N$. Let $l_1 = R_{11}$ and $m_1 = R_{21}$. Repeat the procedure to obtain $(R_{12}, R_{22}), (R_{13}, R_{23}), \cdots, (R_{1K}, R_{2K})$ and let $l_k = R_{1k}$ and $m_k = R_{2k}$. Thus two cells are randomly selected and $l_k \neq m_k$ for $k = 1, \cdots, K$.

Note that the pairs (R_{1k}, R_{2k}), $k = 1, \cdots, K$, are mutually independent. Also note that because $P(R_{1k} \leq x, R_{2k} \leq y) = [xy - \min(x, y)]/(n(n-1)) \leq P(R_{1k} \leq x)P(R_{2k} \leq y)$, where $[\cdot]$ represents the "greatestinteger" function, each pair (R_{1k}, R_{2k}) is negatively quadrant dependent.

Let μ_1 be the expected value of $g(Y)$ within the cell designated by $(l_1, \cdots, l_K)$, and let μ_2 be similarily defined for $(m_1, \cdots, m_K)$. Then $\mu_1 = \mu(R_{11}, R_{12}, \cdots, R_{1K})$ and $\mu_2 = \mu(R_{21}, R_{22}, \cdots, R_{2K})$ are concordant in each argument under the assumptions of the theorem. Lehmann's theorem then yields that μ_1 and μ_2 are negatively quadrant dependent. Therefore,

$$P(\mu_1 \leq x, \mu_2 \leq y) \leq P(\mu_1 \leq x)P(\mu_2 \leq y).$$

Using Hoeffding's equation

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [P(X \leq x, Y \leq y) - P(X \leq x)P(Y \leq y)]\, dx\, dy,$$

(see Lehmann [10] for a proof), we have $\mathrm{Cov}(\mu_1, \mu_2) \leq 0$. Since $\mathrm{Var}(T_L) = \mathrm{Var}(T_R) + (N-1)/N\, \mathrm{Cov}(\mu_1, \mu_2)$, the theorem is proved.

Since $g(t)$ as used in both Sections 3 and 5 is an increasing function of t, we can say that if $Y = h(\mathbf{X})$ is a monotonic function of each of its arguments, Latin hypercube sampling is better than random sampling for estimating the mean and the population distribution function.

REFERENCES

[1] ANSCOMBE, F. J. (1959). Quick analysis methods for random balance screening experiments. *Technometrics, 1,* 195–209.

[2] BUDNE, T. A. (1959). The application of random balance designs. *Technometrics, 1,* 139–155.

[3] DEMPSTER, A. P. (1960). Random allocation designs I: On general classes of estimation methods. *Ann. Math. Statist., 31,* 885–905.

[4] DEMPSTER, A. P. (1961). Random allocation designs II: Approximate theory for simple random allocation. *Ann. Math. Statist., 32,* 387–405.

[5] EHRENFELD, S., and ZACKS, S. (1951). Randomization and factorial experiments. *Ann. Math. Statist., 32,* 270–297.

[6] EHRENFELD, S., and ZACKS, S. (1967). Testing hypotheses in randomized factorial experiments. *Ann. Math. Statist., 38,* 1494–1507.

[7] HIRT, C. W., NICHOLS, B. D., and ROMERO, N. C. (1975). SOLA—a numerical solution algorithm for transient fluid flows. *Los Alamos Scientific Laboratory Report LA-5852,* Los Alamos.

[8] HIRT, C. W., and ROMERO, N. C. (1975). Application of a drift-flux model to flashing in straight pipes. *Los Alamos Scientific Laboratory Report LA-6005-MS,* Los Alamos.

[9] JESSEN, RAYMOND J. (1975). Square and cubic lattice sampling. *Biometrics, 31,* 449–471.

[10] LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist., 35,* 1137–1153.

[11] RAJ, DES. (1968). *Sampling Theory,* pp. 206–209. New York: McGraw-Hill.

[12] SATTERTHWAITE, F. E. (1959). Random balance experimentation. *Technometrics, 1,* 111–137.

[13] STEINBERG, H. A. (1963). Generalized quota sampling. *Nuc. Sci. and Engr., 15,* 142–145.

[14] TOCHER, K. D. (1963). *The Art of Simulation.* pp. 106–107. Princeton, N.J.: D. Van Nostrand.

[15] YOUDEN, W. J., KEMPTHORNE, O., TUKEY, J. W., BOX, G. E. P., and HUNTER, J. S. (1959). Discussion of the papers of Messrs. Satterthwaite and Budne. *Technometrics, 1,* 157–193.

[16] ZACKS, S. (1963). On a complete class of linear unbiased estimators for randomized factorial experiments. *Ann. Math. Statist., 34,* 769–779.

[17] ZACKS, S. (1964). Generalized least squares estimators for randomized fractional replication designs. *Ann. Math. Statist., 35,* 696–704.