

Orthogonal Array-Based Latin Hypercubes

Boxin TANG*

In this article, we use orthogonal arrays (OA's) to construct Latin hypercubes. Besides preserving the univariate stratification properties of Latin hypercubes, these strength r OA-based Latin hypercubes also stratify each r -dimensional margin. Therefore, such OA-based Latin hypercubes provide more suitable designs for computer experiments and numerical integration than do general Latin hypercubes. We prove that when used for integration, the sampling scheme with OA-based Latin hypercubes offers a substantial improvement over Latin hypercube sampling.

KEY WORDS: Computer experiments; Numerical integration; Orthogonal arrays.

1. INTRODUCTION

In many scientific and technological fields, we are often confronted with the problem of evaluating a complex integral over a high-dimensional domain. Among numerical integration techniques, Monte Carlo methods are especially useful and often competitive for high-dimensional integration (Davis and Rabinowitz 1984, chap. 5.10), and may be formulated as follows. Consider a deterministic function $Y = f(\mathbf{X})$ where $Y \in R$ and $\mathbf{X} \in R^m$ and f is known but expensive to compute. The random vector, $\mathbf{X} = (X^1, \dots, X^m)$, has a uniform distribution on the unit hypercube $[0, 1]^m$. We want to estimate the mean, $\mu = E(Y)$, of the random variable Y . This is equivalent to finding the integral of $y = f(\mathbf{x})$ with respect to the uniform measure on $[0, 1]^m$.

The simplest Monte Carlo way is to draw $\mathbf{X}_1, \dots, \mathbf{X}_n$ independently from $\text{Unif}[0, 1]^m$ and to use $\bar{Y} = n^{-1} \times \sum_{i=1}^n f(\mathbf{X}_i)$ as an estimate of μ . McKay, Conover, and Beckman (1979) introduced Latin hypercube sampling (LHS) as an alternative to iid sampling and showed that LHS can result in variance reduction for \bar{Y} when f is monotone in each variable. Stein (1987) obtained a more informative result. To describe Stein's result, we define the main effects as $f_j(X^j) = E[f(\mathbf{X}) | X^j] - \mu$. Then Stein showed that the variance of \bar{Y} under LHS is $n^{-1} \text{var}[f(\mathbf{X})] - n^{-1} \sum_{j=1}^m \text{var}[f_j(X^j)] + o(n^{-1})$, which is asymptotically smaller than the iid variance $n^{-1} \text{var}[f(\mathbf{X})]$.

The main feature of LHS is that it stratifies each univariate margin simultaneously. Stein's result simply states that stratifying each univariate margin filters out the main effects. One may expect that if stratification is also achieved on each bivariate margin, then one may filter out all the bivariate interactions as well as the main effects. In fact, if we let $f_{ij}(X^i, X^j) = E[f(\mathbf{X}) | X^i, X^j] - \mu - f_i(X^i) - f_j(X^j)$ be the interaction of X^i and X^j , then we will prove that the variance of \bar{Y} is $n^{-1} \text{var}[f(\mathbf{X})] - n^{-1} \sum_{j=1}^m \text{var}[f_j] - n^{-1} \sum_{i < j} \text{var}[f_{ij}] + o(n^{-1})$ under the sampling scheme using strength two OA-based Latin hypercubes.

In the theory and practice of experimental designs, it is

well known that when the assumed model is in doubt, we are led to concentrate on and minimize the bias part of the mean squared error (MSE). This usually can be done by distributing design points uniformly in the design region (Box and Draper 1959, Sacks and Ylvisaker 1984). OA designs are used extensively for planning experiments in industry, and their success is due at least in part to their uniformity properties. But when a large number of factors are to be studied in an experiment but only a few of them are virtually effective, OA designs projected onto the subspace spanned by the effective factors can result in replication of points. This is undesirable for physical experiments in which the bias of the proposed model is more serious than the variance, and it can be disastrous for deterministic computer experiments. In this case, Latin hypercube designs (LHD's) are the preferred alternatives (Welch et al. 1992). But the projections of such design points onto even bivariate margins cannot be guaranteed to be uniformly scattered. Current methods for improving an LHD control the correlations among its permutation columns (Iman and Conover 1982). In this article we present a method of constructing LHD's by exploiting the use of OA's. Such OA-based LHD's inherit the r -variate uniformity property of a strength r OA that is used for construction; and therefore, they are more appropriate designs for computer experiments than are general LHD's.

The article is organized as follows. Section 2 introduces OA's and Latin hypercubes and shows how OA's can be used to construct Latin hypercubes; a number of examples are presented for illustration. Section 3 presents some results on the sampling scheme using strength two OA-based Latin hypercubes. Section 4 gives a brief generalization of our method through the use of asymmetrical OA's.

2. CONSTRUCTION OF OA-BASED LATIN HYPERCUBES

An $n \times m$ matrix \mathbf{A} , with entries from a set of $s \geq 2$ symbols, is called an OA of strength r , size n , with m constraints and s levels if each $n \times r$ submatrix of \mathbf{A} contains all possible $1 \times r$ row vectors with the same frequency λ . The number λ is called the index of the array; clearly, $n = \lambda s^r$. The array is denoted by $\text{OA}(n, m, s, r)$. The s symbols are taken as $1, 2, \dots, s$ throughout this article.

A Latin hypercube is an $n \times m$ matrix, each column of which is a permutation of $1, 2, \dots, n$. Therefore, by the

* Boxin Tang is a postdoctoral fellow in the Department of Preventive Medicine and Biostatistics, University of Toronto, Ontario, Canada, M5S 1A8. This work is part of his Ph.D. dissertation for the Department of Statistics and Actuarial Science, University of Waterloo. The research was supported by the Natural Sciences and Engineering Research Council of Canada, the Manufacturing Research Corporation of Ontario, and the Institute for Improvement in Quality and Productivity. The author is most indebted to Jeff Wu for the time and energy he spent on this work. The author also thanks Mary Thompson, William Welch, and the referees for their helpful comments and suggestions, and Joan Hu and Runchu Zhang for useful discussion.

definition of OA's, a Latin hypercube is an OA of strength one, $OA(n, m, n, 1)$.

Current users of Latin hypercubes randomly select permutation columns (Welch and Sacks 1990), usually with a supplemental control of the correlations among the permutation columns, in the hope of getting better uniformity properties for multivariate margins. The Latin hypercubes to be constructed here possess the r -variate uniformity properties inherent in a strength r OA used for construction.

Let \mathbf{A} be an $OA(n, m, s, r)$. For each column of \mathbf{A} , we replace the λs^{r-1} positions with entry k by a permutation of $(k-1)\lambda s^{r-1} + 1, (k-1)\lambda s^{r-1} + 2, \dots, (k-1)\lambda s^{r-1} + \lambda s^{r-1} = k\lambda s^{r-1}$, for all $k = 1, \dots, s$. After the replacement procedure is done for all m columns of \mathbf{A} , the newly obtained matrix, denoted by \mathbf{U} , is evidently a Latin hypercube. It enjoys the merits of \mathbf{A} in achieving uniformity in each r -variate margin. Note that \mathbf{U} becomes \mathbf{A} under the element-wise mapping

$$Z(i) = \lceil i/\lambda s^{r-1} \rceil, \quad i = 1, \dots, n, \quad (1)$$

where $\lceil x \rceil$ is the smallest integer $\geq x$.

The construction of OA-based Latin hypercubes depends on the existence of the corresponding OA. For related results on OA's, we refer to Plackett and Burman (1946), Rao (1947), Bose and Bush (1952), Dey (1985), and de Launey (1986). OA-based Latin hypercubes are referred to as U designs hereafter when they are used for designing experiments. U designs should be more appropriate designs than general LHD's for computer experiments. They may also be useful for physical experiments, in which some factors have many levels to be accommodated so that the number of experiments required by a suitable OA may be too large to be practical.

The nonuniqueness of U designs based on a single OA poses the problem of choosing a desirable U design. A paper in preparation by Tang and Wu discusses this problem, using both distance and correlation criteria. We do not give details here but merely point out this important issue.

We now consider how to use OA-based Latin hypercubes for numerical integration. For a given $OA(n, m, s, r) = \mathbf{A}$, we randomize its rows, columns, and symbols to obtain a randomized OA. Then for each column of this randomized array, we replace the λs^{r-1} positions with entry k by a random permutation (with each such permutation having an equal probability of being taken) of $(k-1)\lambda s^{r-1} + 1, (k-1)\lambda s^{r-1} + 2, \dots, (k-1)\lambda s^{r-1} + \lambda s^{r-1} = k\lambda s^{r-1}$, for all $k = 1, \dots, s$. This procedure generates a random OA-based Latin hypercube, denoted by $\mathbf{U} = (u_{ij})$. Now suppose that $X_i^j \sim \text{Unif}((i-1)/n, i/n)$, $i = 1, \dots, n$, $j = 1, \dots, m$, are all generated independently. Then the n points to be used for integration are formed by $\mathbf{X}_i = (X_{u_{i1}}^1, X_{u_{i2}}^2, \dots, X_{u_{im}}^m)$, $i = 1, \dots, n$. This sampling scheme is referred to as U sampling from now on. It should be noted that for the practical use of U sampling, we can in fact omit the step of randomizing the rows of \mathbf{A} . It is included simply to make the U sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ exchangeable and hence to ease the study of the theoretical properties of U sampling.

The description of U sampling can be made more compact using the mapping Z introduced in (1). We may simply draw

a Latin hypercube randomly from the collection of Latin hypercubes $\{\mathbf{U}\}$ (each such Latin hypercube having an equal chance to be drawn) such that $(Z(u_{ij}))$ forms an $OA(n, m, s, r)$ that is equivalent to the given \mathbf{A} . (Two OA's are said to be equivalent if one can be obtained by permuting the rows, columns, and symbols of the other.) This view is particularly useful for the derivations of various probabilities connected with U sampling. We postpone the study of the theoretical properties of U sampling to Section 3. The remainder of this section presents some examples of U designs.

Example 1. We use $OA(4, 2, 2, 2)$ to construct a four-point U design, \mathbf{U} . For comparisons, we also give a four-point LHD, \mathbf{L} , that is not a U design. The two designs are

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{bmatrix}^T \quad \mathbf{L} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{bmatrix}^T$$

and can be represented by Figure 1.

Example 2. We construct two U designs in this example, one based on $OA(9, 2, 3, 2)$ and the other on $OA(8, 2, 2, 2)$. We give their graphical representations in Figure 2.

Example 3. We consider a more realistic situation in this example. A random U design is generated using $OA(49, 8, 7, 2)$, and the pairwise plots of the columns of the design are given in Figure 3. To allow comparison, we also provide in Figure 4 the plots of this sort for a random LHD. It is seen that the design points in the former look more uniform than those in the latter.

3. SOME RESULTS ON U SAMPLING

In this section, we study U sampling with associated OA's being $OA(s^2, m, s, 2)$, as the use of such arrays leads to the most economical sample size, $n = s^2$, for a given s . (For a given s , the U sample is obtained by taking a fixed $OA(s^2, m, s, 2)$ and then applying the method of Section 2.) The basic ideas should carry through for general arrays, however, although derivations will become very complicated. An exception is Theorem 2, in which general OA's are considered.

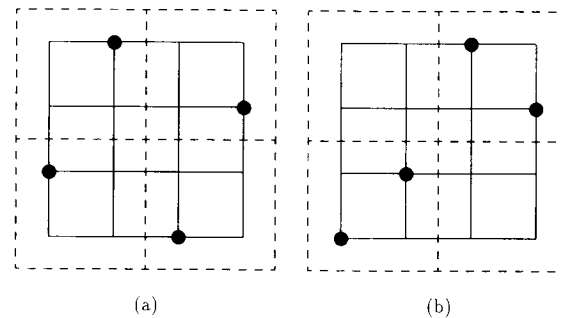


Figure 1. A Four-Point U Design and a Four-Point LHD. The points in (a) represent the U design; the points in (b), the LHD. On intuitive grounds, the former should be preferred to the latter. Indeed, suppose that the area within the dashed line box is the design region. The design points in (a) are more uniformly scattered than those in (b) in the two-dimensional region in the sense that each small dashed line box in (a) contains exactly one design point.

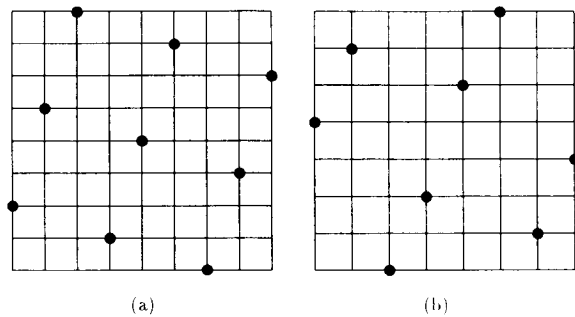


Figure 2. U Designs Based on OA(9, 2, 3, 2) and OA(8, 2, 2, 2). The points in (a) represent a nine-point U design; the points in (b), an eight-point U design.

We first introduce a generic two-stage sampling scheme that includes both LHS and U sampling as special cases. This general framework facilitates the theoretical development in the sequel, and it may also be used to suggest other useful sampling schemes. Let $\mathcal{P} = \{[0, 1/n], (1/n, 2/n], \dots, (1 - 1/n, 1]\}$ and $\mathcal{C} = \{P_1 \times \dots \times P_m | P_j \in \mathcal{P}, j = 1, \dots, m\}$. Members of \mathcal{C} are referred to as cells, and thus the n^m cells form a partition of $[0, 1]^m$. We use $c(P_j)$ to denote the cylinder set of P_j ; that is, $c(P_j)$ consists of points in $[0, 1]^m$ such that only the j th coordinate of such points are restricted and belong to P_j . To estimate the mean of Y , we would like to draw an \mathbf{X} from the uniform distribution for each cell in \mathcal{C} . Consideration of cost, however, leads us to use only a sample of cells and hence to consider the following two-stage sampling scheme.

Stage 1. Draw a sequence C_1, \dots, C_n of cells from \mathcal{C} using some sampling scheme satisfying the following two criteria:

1. The n cells are distinct, and the random vector (C_1, \dots, C_n) is exchangeable.

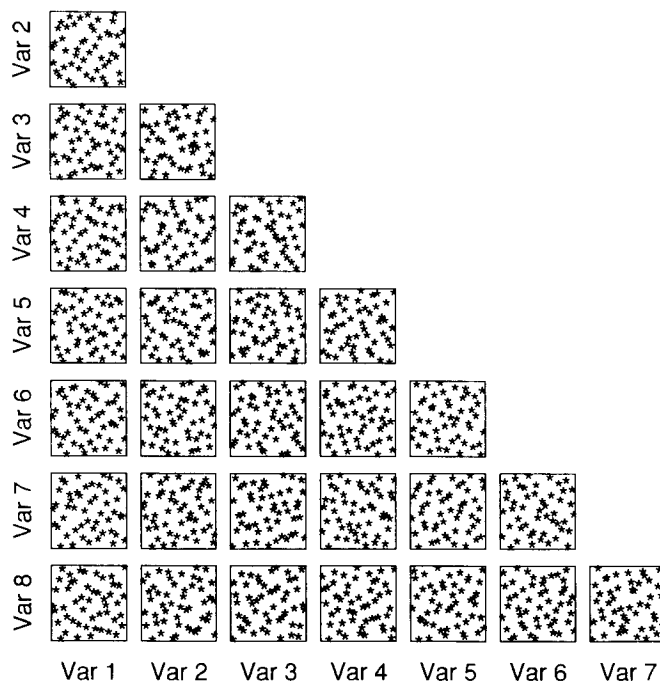


Figure 3. The Random Latin Hypercube Based on OA(49, 8, 7, 2).

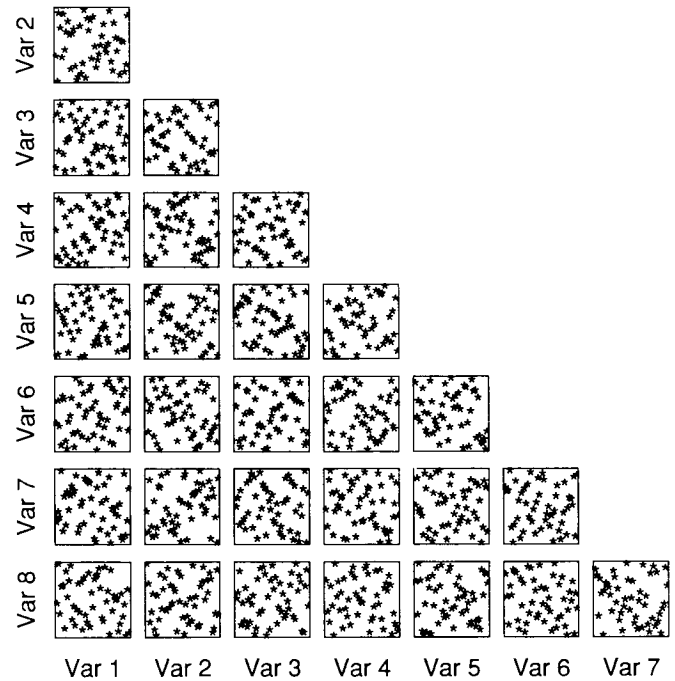


Figure 4. The Random Latin Hypercube With 49 Columns and 8 Rows.

2. The marginal distribution of C_i is uniform on \mathcal{C} , $i = 1, \dots, n$.

Stage 2. For each $i = 1, \dots, n$, an \mathbf{X}_i is drawn from the uniform distribution on C_i , the drawing being carried out independently for each $i = 1, \dots, n$.

Denote the variance of Y by σ^2 . Then we have the following obvious assertion.

Assertion. Under the two-stage sampling scheme, \bar{Y} is an unbiased estimator of μ , and the variance of \bar{Y} is given by $\text{var}(\bar{Y}) = n^{-1}\sigma^2 + n^{-1}(n-1)\text{cov}(f(\mathbf{X}_1), f(\mathbf{X}_2))$.

It is easy to see that both LHS and U sampling are special cases of the two-stage sampling scheme, where the i th cell C_i corresponds to the i th row of a random Latin hypercube for LHS and a random OA-based Latin hypercube for U sampling. Therefore, the assertion holds for both.

Let $\mathcal{Q} = \{[0, 1/s], (1/s, 2/s], \dots, (1 - 1/s, 1]\}$ and $\mathcal{L} = \{Q_1 \times \dots \times Q_m | Q_j \in \mathcal{Q}\}$. We refer to members of \mathcal{L} as large cells and members of \mathcal{C} introduced previously as small cells from now on. We similarly define the cylinder sets of Q_j and $Q_i \times Q_j$, denoted by $c(Q_j)$ and $c(Q_i \times Q_j)$. We are now in a position to give the following theorem.

Theorem 1. If f is bounded, then under U sampling we have

$$\begin{aligned} \text{cov}[f(\mathbf{X}_1), f(\mathbf{X}_2)] &= -n^{-1} \sum_{j=1}^m \text{var}[E(f(\mathbf{X}_1) | c(P_j))] \\ &\quad - n^{-1} \sum_{i < j}^m \text{var}\{E[f(\mathbf{X}_1) | c(Q_i \times Q_j)] - E[f(\mathbf{X}_1) | c(Q_i)] \\ &\quad - E[f(\mathbf{X}_1) | c(Q_j)]\} + o(n^{-1}), \end{aligned}$$

where $P_1 \times \cdots \times P_m$ and $Q_1 \times \cdots \times Q_m$ are the small and large cells to which \mathbf{X}_1 belongs and, for example, conditioning on $c(P_j)$ means conditioning on $\mathbf{X}_1 \in c(P_j)$.

The proof of Theorem 1 is given in Appendix A. If f is continuous on $[0, 1]^m$, consequently also bounded, then we can easily show the following result using Theorem 1.

Corollary 1. If f is continuous on $[0, 1]^m$, then under U sampling we have

$$\begin{aligned} \text{cov}[f(\mathbf{X}_1), f(\mathbf{X}_2)] &= -n^{-1} \sum_{j=1}^m \text{var}[f_j(X^j)] \\ &\quad - n^{-1} \sum_{i < j}^m \text{var}[f_{ij}(X^i, X^j)] + o(n^{-1}), \end{aligned}$$

where f_j and f_{ij} are main effects and interactions as defined in Section 1.

Under LHS, we know that $\text{cov}[f(\mathbf{X}_1), f(\mathbf{X}_2)] = -n^{-1} \times \sum_{j=1}^m \text{var}[f_j(X^j)] + o(n^{-1})$ (Stein 1987). Therefore, Corollary 1 indicates a stronger negative correlation between $f(\mathbf{X}_1)$ and $f(\mathbf{X}_2)$ under U sampling than under LHS in the asymptotic sense.

Corollary 2. If f is continuous on $[0, 1]^m$, then under U sampling we have

$$\begin{aligned} \text{var}(\bar{Y}) &= n^{-1} \sigma^2 - n^{-1} \sum_{j=1}^m \text{var}[f_j(X^j)] \\ &\quad - n^{-1} \sum_{i < j}^m \text{var}[f_{ij}(X^i, X^j)] + o(n^{-1}). \end{aligned}$$

This follows from Corollary 1 and our Assertion.

If we write

$$f(\mathbf{X}) = \mu + \sum_{j=1}^m f_j(X^j) + \sum_{i < j}^m f_{ij}(X^i, X^j) + r(\mathbf{X}),$$

then it is easy to check that the terms on the right side of the equality are uncorrelated with each other. Therefore, we obtain the corresponding variance decomposition

$$\sigma^2 = \sum_{j=1}^m \text{var}(f_j) + \sum_{i < j}^m \text{var}(f_{ij}) + \text{var}[r(\mathbf{X})],$$

and consequently the variance of \bar{Y} under U sampling is $n^{-1} \text{var}[r(\mathbf{X})] + o(n^{-1})$. This generalizes Stein's result. Owen

(1992) independently obtained a result similar to Corollary 2 by using randomized OA's, with the aid of the results of Patterson (1954). Owen's sampling procedure may be briefly described as follows. Suppose that $\mathbf{A} = (a_{ij})$ is an OA(n, m, s, r) with its symbols randomized and that $X_i^j \sim \text{Unif}(0, 1)$, $i = 1, \dots, n$, $j = 1, \dots, m$, are generated independently. Then the n points Owen used for integration are the rows of the matrix $((a_{ij} - X_i^j)/s)$. To compare U sampling with Owen's procedure, let V_1 and V_2 be the variances of \bar{Y} using U sampling and Owen's procedure. Then we can show the following theorem.

Theorem 2. When f is additive—that is, when $f(\mathbf{X}) = \mu + \sum_{j=1}^m f_j(X^j)$ —we have

$$\begin{aligned} V_2 - V_1 &= n^{-1} \sum_{j=1}^m E\{\text{var}[E(f_j(X^j) | X^j \in P \subset Q) | X^j \in Q]\} > 0, \end{aligned}$$

where $X^j \sim \text{Unif}[0, 1]$ and $P \in \mathcal{P}$, $Q \in \mathcal{Q}$.

Thus Theorem 2 shows that U sampling gives a smaller variance of \bar{Y} than does Owen's procedure whenever f is additive. Note that Theorem 2 is true for any OA. We prove Theorem 2 in Appendix B.

Before concluding this section, we note the close connections between our results and the work in the numerical analysis literature on quasi-Monte Carlo. Hua and Wang (1981) gave a systematic account of quasi-Monte Carlo methods. For historical developments, we refer to the review papers by Niederreiter (1978, 1988). Our method has two features not shared by quasi-Monte Carlo methods. First, OA-based Latin hypercubes stress low-dimensional margins and specifically achieve uniformity in each t -dimensional margin ($t \leq r$). Second the uniformity properties of OA-based Latin hypercubes are in fact finite sample properties in the sense that, for example, each $c(P_j)$ contains exactly one point. I am currently exploring further connections between OA-based Latin hypercubes and quasi-Monte Carlo methods, both theoretically and numerically, and hope to report them in a future communication.

4. GENERALIZATION

Generalizations of OA-based Latin hypercubes can be made easily using asymmetrical OA's. Relevant references on asymmetrical OA's include Rao (1973), Dey (1985), Wu (1989), Wang and Wu (1991), and Wu, Zhang, and Wang (1992). The notation OA($n, r; s_1, \dots, s_m$) is used to denote an asymmetrical OA of size n and strength r , with s_j levels in the j th column, $j = 1, \dots, m$, where the s_j 's may not be distinct. Let \mathbf{A} be a given OA($n, r; s_1, \dots, s_m$). Then an LHD, $U = (u_{ij})$, is said to be an asymmetrical U design based on \mathbf{A} , if $(Z_j(u_{ij})) = \mathbf{A}$, where $Z_j(i)$ is defined to be $\lceil is_j/n \rceil$. The generalization of U sampling can be made accordingly. A simple example illustrates the underlying idea.

Example 4. An asymmetrical U design is constructed using the OA(6, 2; 2, 3), and is represented graphically in Figure 5. An LHD that is not an asymmetrical U design is also given in Figure 5.

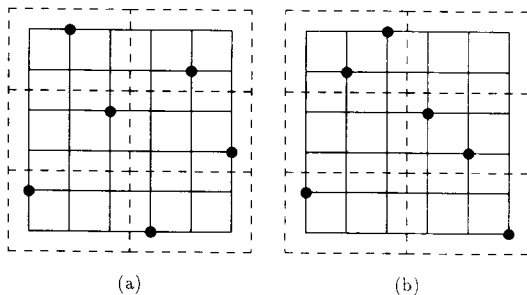


Figure 5. A Six-Point U Design Together With a Six-Point LHD. The points in (a) represent the U design; the points in (b), the LHD. Note that each small dashed line box in (a) contains exactly one design point.

APPENDIX A: PROOF OF THEOREM 1

Several lemmas are needed for proving Theorem 1.

Let $c_1 = P_1 \times \cdots \times P_m$ be a small cell in \mathcal{C} and $L = Q_1 \times \cdots \times Q_m$ be the large cell containing c_1 . Further, let $G = (\cup_{j=1}^m c(P_j)) \cup (\cup_{j=1}^m c(Q_j \times Q_j))$. For U sampling using $OA(s^2, m, s, 2)$, we have $P(C_2 = c_2 | C_1 = c_1) = 0$ for all $c_2 \subset G$. We derive below an expression for $P(C_2 = c_2 | C_1 = c_1)$, for $c_2 \subset G^c$, which plays a key role in the proof of Theorem 1. The following lemma is obvious by the definition of U sampling.

Lemma 1. We have

$$P(C_2 = c_2 | C_1 = c_1)$$

$$= \begin{cases} a, & \forall c_2 \subset \left(\bigcup_{j=1}^m c(Q_j) \right)^c = \bigcap_{j=1}^m c(Q_j)^c, \\ b, & \forall c_2 \subset \left(\bigcup_{j=1}^m c(Q_j) \right) \setminus G = \bigcup_{j=1}^m (c(Q_j) \setminus G), \end{cases}$$

where a and b are constants. Note that $G^c = (\cap_{j=1}^m c(Q_j)^c) \cup (\cup_{j=1}^m (c(Q_j) \setminus G))$.

In Lemmas 2 and 3 we derive the expressions for a and b .

Lemma 2. We have

$$\frac{a}{b} = \frac{s - m + 1}{s}.$$

Proof. Let $c_1 = (u_{11}, u_{12}, \dots, u_{1m})$ and $c_2 = (u_{21}, u_{22}, \dots, u_{2m})$. We want to count the number of Latin hypercubes $U = (u_{ij})$, with the first and second rows fixed to be c_1 and c_2 , such that $(Z(u_{ij}))$ is an $OA(s^2, m, s, 2)$ that is equivalent to the given array, A , where Z is defined in (1).

Step 1. We first obtain the elements u_{i1} , $i = 3, \dots, n$, in the first column. The number of ways for doing this is $(n - 2)!$.

Step 2. Now consider an $OA(s^2, m, s, 2)$, $B = (b_{ij})$ that is equivalent to A , such that the symbols in the first and second rows are given by $b_{ij} = Z(u_{ij})$, $i = 1, 2, j = 1, \dots, m$, and the symbols in the first column are given by $b_{i1} = Z(u_{i1})$, $i = 3, \dots, n$.

In understanding Step 2, it is helpful to think of there being ms symbols τ_j , $j = 1, \dots, m$ and $\tau = 1, \dots, s$; then $b_{ij} = \tau_j$ if column j has level τ in the i th row. Thus the symbols look like $2_1, 3_2, 1_3$, and so on.

For the m pairs (b_{1j}, b_{2j}) of symbols, $j = 1, \dots, m$, two cases arise:

Case 1. When $c_2 \subset \cup_{j=1}^m (c(Q_j) \setminus G)$, there must be one pair having the same symbol in it. Without loss of generality, suppose that this is the first pair, which implies $b_{11} = b_{21}$, and $b_{1j} \neq b_{2j}$, $j = 2, \dots, m$.

Case 2. When $c_2 \subset \cap_{j=1}^m c(Q_j)^c$, no pair has the same symbol in it; that is, $b_{1j} \neq b_{2j}$, $j = 1, \dots, m$.

We now divide the array B into s subarrays of s rows each such that within each subarray, the symbol in the first column is the same for all rows. Therefore, within each subarray, each symbol for every other column appears once. Let us call the subarray to which the first row $(b_{11}, b_{12}, \dots, b_{1m})$ belongs the first block and the subarray to which the next row not in the first block belongs the second block.

Case 1. The two rows $(b_{11}, b_{12}, \dots, b_{1m})$ and $(b_{21}, b_{22}, \dots, b_{2m})$ are both in the first block, because $b_{11} = b_{21}$. Because the symbols b_{12}, \dots, b_{1m} already appear in the top row of the first block, in the second block they must all be in different rows. Thus in the second block there are $s^{(m-1)} = s(s-1) \cdots (s-m+2)$ ways to arrange the symbols.

Case 2. The row $(b_{11}, b_{12}, \dots, b_{1m})$ is in the first block, and the row $(b_{21}, b_{22}, \dots, b_{2m})$ is in the second block, as $b_{11} \neq b_{21}$. Here again, the symbols b_{12}, \dots, b_{1m} must all be in different rows in the second block. Because the row $(b_{21}, b_{22}, \dots, b_{2m})$ already occupies one row in the second block, there are $(s-1)^{(m-1)}$ ways to arrange the symbols b_{12}, \dots, b_{1m} in the second block.

It is well known that for the existence of an $OA(s^2, m, s, 2)$, it is necessary that $m \leq s+1$. If $m = s+1$, then Lemma 2 is obviously true, because it is impossible to arrange b_{12}, \dots, b_{1m} in the second block for Case 2. Now suppose that $m \leq s$.

Case 1. We now have b_{12}, \dots, b_{1m} placed in the first and second blocks, and a subrow (b_{22}, \dots, b_{2m}) of different symbols in the second row of the first block. Because $m \leq s$, at least one of the rows of the second block (without loss of generality, say, the first row) has none of $b_{12}, b_{13}, \dots, b_{1m}$ in it. Let (b_2, b_3, \dots, b_m) be an assignment of symbols to that row.

Case 2. Here we have b_{12}, \dots, b_{1m} placed in the first and second blocks and a subrow (b_{22}, \dots, b_{2m}) of different symbols in the first row of the second block. Let (b_2, \dots, b_m) be an assignment of symbols to the second row of the first block.

At this point, each Case 1 assignment compatible with the given A could have been arrived at as a Case 2 assignment, and vice versa. To be specific, with any particular arrangement for the symbols b_{12}, \dots, b_{1m} , let \mathcal{B}_1 for Case 1 and \mathcal{B}_2 for Case 2 be the collections of OA 's that can be obtained by filling the remaining positions with appropriate symbols. For any $B \in \mathcal{B}_1$, if we switch the two symbols b_j and b_{2j} in column j , $j = 2, \dots, m$ and then supplement the symbol-switching procedure with a suitably chosen permutation of the rows, we will obtain an OA in \mathcal{B}_2 . This defines a mapping from \mathcal{B}_1 to \mathcal{B}_2 . It is easy to check that this mapping actually establishes a one-to-one correspondence between \mathcal{B}_1 and \mathcal{B}_2 . Thus the numbers of OA 's in \mathcal{B}_1 and in \mathcal{B}_2 are equal. Note that for any $B_1 \in \mathcal{B}_1$, the array in \mathcal{B}_2 (say, B_2) given by the mapping is equivalent to B_1 . So either both B_1 and B_2 are equivalent to the given A or neither is equivalent to A . Therefore, \mathcal{B}_1 and \mathcal{B}_2 contain the same number of OA 's that are equivalent to A .

Step 3. Obtain u_{ij} , $i \geq 3, j \geq 2$, such that $Z(u_{ij}) = b_{ij}$. It is easily seen that the number of ways of getting such u_{ij} 's is the same for both cases.

In summary, we conclude that the number of Latin hypercubes based on A , with the first two rows equal to c_1 and c_2 , is $ks \cdots (s-m+2)$ for Case 1 and $k(s-1) \cdots (s-m+1)$ for Case 2. Therefore,

$$\frac{a}{b} = \frac{k(s-1) \cdots (s-m+1)}{ks \cdots (s-m+2)} = \frac{s-m+1}{s}.$$

It is easy to see that the numbers of small cells in $\cap_{j=1}^m c(Q_j)^c$ and in $\cup_{j=1}^m (c(Q_j) \setminus G)$ are $(n-s)^m$ and $m(s-1)(n-s)^{m-1}$. Therefore, we have

$$a(n-s)^m + bm(s-1)(n-s)^{m-1} = 1.$$

Combining this equation with Lemma 2, we can easily derive the exact expressions for a and b . The results are provided in Lemma 3.

Lemma 3. We have

$$a = \frac{s-m+1}{s^m(s-1)^m(s+1)}, \quad b = \frac{1}{s^{m-1}(s-1)^m(s+1)}.$$

Let ν be the uniform measure with total mass 1 on G^c , and let $g(x | C_1 = c_1)$ be the probability density function of X_2 conditionally on $C_1 = c_1$ with respect to ν . Then Lemma 4 is a direct consequence of Lemma 3.

Lemma 4. We have

$$g(\mathbf{x}|C_1 = c_1) = \begin{cases} \frac{s^2 - (m-1)^2}{s^2 - 1}, & \text{for } \mathbf{x} \in \bigcap_{j=1}^m c(Q_j)^c, \\ \frac{s^2 + (m-1)s}{s^2 - 1}, & \text{for } \mathbf{x} \in \bigcup_{j=1}^m (c(Q_j) \setminus G). \end{cases}$$

We are ready to prove Theorem 1.

Proof of Theorem 1. Without loss of generality, we assume $\mu = 0$ in the proof. We now have

$$\text{cov}[f(\mathbf{X}_1), f(\mathbf{X}_2)] = E[f(\mathbf{X}_1)f(\mathbf{X}_2)] = E[E(f(\mathbf{X}_1)f(\mathbf{X}_2)|C)],$$

where $C = P_1 \times \cdots \times P_m$ and $L = Q_1 \times \cdots \times Q_m$ are the small and large cells to which \mathbf{X}_1 belongs. Let $G = (\bigcup_{j=1}^m c(P_j)) \cup (\bigcup_{i < j} c(Q_i \times Q_j))$ as before. Then, conditionally on C , \mathbf{X}_1 and \mathbf{X}_2 are independent and \mathbf{X}_2 has the probability density function $g(\mathbf{x}|C)$ given in Lemma 4. Therefore,

$$E[f(\mathbf{X}_1)f(\mathbf{X}_2)] = E[E(f(\mathbf{X}_1)|C)E(f(\mathbf{X}_2)|C)],$$

so that

$$\text{cov}[f(\mathbf{X}_1), f(\mathbf{X}_2)] = E[E(f(\mathbf{X}_1)|C)E(f(\mathbf{X}_2)|C)]. \quad (\text{A.1})$$

Further, we have

$$E(f(\mathbf{X}_2)|C) = \frac{s^2 + (m-1)s}{s^2 - 1} \int_{\bigcup_{j=1}^m (c(Q_j) \setminus G)} f(\mathbf{x}) d\nu + \frac{s^2 - (m-1)^2}{s^2 - 1} \int_{\bigcap_{j=1}^m c(Q_j)^c} f(\mathbf{x}) d\nu.$$

As s becomes large, the set G^c approaches the unit cube $[0, 1]^m$. Thus ν has a density $1 + o(1)$ with respect to the Lebesgue measure on G^c . Noting that $\mu = \int_{[0,1]^m} f(\mathbf{x}) d\mathbf{x} = 0$, we can easily show that

$$\begin{aligned} E(f(\mathbf{X}_2)|C) &= -s^{-2} \sum_{j=1}^m E[f(\mathbf{X}_1)|c(P_j)] \\ &\quad - (m-1)s^{-2} \sum_{j=1}^m E[f(\mathbf{X}_1)|c(Q_j)] \\ &\quad - s^{-2} \sum_{i < j} E[f(\mathbf{X}_1)|c(Q_i \times Q_j)] + o(s^{-2}), \end{aligned}$$

where, for example, the first \mathbf{X}_1 in the expression can be any random variable uniformly distributed on $c(P_j)$. With this expression for $E(f(\mathbf{X}_2)|C)$ substituted into (A.1), it is easy to obtain

$$\begin{aligned} \text{cov}[f(\mathbf{X}_1), f(\mathbf{X}_2)] &= -s^{-2} \sum_{j=1}^m \text{var}[E(f(\mathbf{X}_1)|c(P_j))] \\ &\quad - s^{-2} \sum_{i < j} \text{var}[E(f(\mathbf{X}_1)|c(Q_i \times Q_j)) - E(f(\mathbf{X}_1)|c(Q_i)) \\ &\quad \quad - E(f(\mathbf{X}_1)|c(Q_j))] + o(s^{-2}). \end{aligned}$$

This proves the result claimed in Theorem 1, as $n = s^2$.

APPENDIX B: PROOF OF THEOREM 2

Because $\bar{Y} = \mu + n^{-1} \sum_{j=1}^m \sum_{i=1}^n f_j(X_i^j)$, we have

$$V_1 = n^{-2} \sum_{j=1}^m \sum_{i=1}^n \text{var}[f_j(X_i^j)],$$

where $X_i^j \sim \text{Unif}((i-1)/n, i/n)$, and

$$V_2 = n^{-2} \sum_{j=1}^m \sum_{k=1}^s \sum_{l=1}^{s^{r-1}} \text{var}[f_j(X_{kl}^j)],$$

where $X_{kl}^j \sim \text{Unif}((k-1)/s, k/s)$. With this observation, the proof of Theorem 2 is straightforward.

[Received July 1991. Revised January 1993.]

REFERENCES

- Bose, R. C., and Bush, K. A. (1952), "Orthogonal Arrays of Strength Two and Three," *The Annals of Mathematical Statistics*, 23, 508–524.
- Box, G. E. P., and Draper, N. R. (1959), "A Basis for the Selection of a Response Surface Design," *Journal of the American Statistical Association*, 54, 622–654.
- Davis, P. J., and Rabinowitz, P. (1984), *Methods of Numerical Integration* (2nd ed.), Orlando, FL: Academic Press.
- Dey, A. (1985), *Orthogonal Fractional Factorial Designs*, New York: Halsted.
- Hua, L. K., and Wang, Y. (1981), *Applications of Number Theory to Numerical Analysis*, Beijing: Springer-Verlag, Science Press.
- Iman, R. L., and Conover, W. J. (1982), "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics, Part B—Simulation and Computation*, 11, 311–334.
- de Launey, W. (1986), "A Survey of Generalized Hadamard Matrices and Difference Matrices $D(k, \lambda; G)$ With Large k ," *Utilitas Mathematica*, 30, 5–29.
- McKay, M. D., Conover, W. J., and Beckman, R. J. (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code," *Technometrics*, 21, 239–245.
- Niederreiter, H. (1978), "Quasi-Monte Carlo Methods and Pseudo-Random Numbers," *Bulletin of the American Mathematical Society*, 84, 957–1041.
- (1988), "Quasi-Monte Carlo Methods for Multidimensional Integration," in *Numerical Integration III*, eds. H. Braß and G. Hämmerlin, Basel: Birkhäuser Verlag, pp. 157–171.
- Owen, A. (1992), "Orthogonal Arrays for Computer Experiments, Integration, and Visualization," *Statistica Sinica*, 2, 439–452.
- Patterson, H. D. (1954), "The Errors of Lattice Sampling," *Journal of the Royal Statistical Society, Ser. B*, 16, 140–149.
- Plackett, R. L., and Burman, J. P. (1946), "The Designs of Optimum Multifactorial Experiments," *Biometrika*, 33, 305–325.
- Rao, C. R. (1947), "Factorial Experiments Derivable from Combinatorial Arrangements of Arrays," *Journal of the Royal Statistical Society, Suppl.* 9, 128–139.
- (1973), "Some Combinatorial Problems of Arrays and Applications to Design of Experiments," in *A Survey of Combinatorial Theory*, eds. J. N. Srivastava et al., Amsterdam: North-Holland, pp. 349–359.
- Stein, M. (1987), "Large-Sample Properties of Simulations Using Latin Hypercube Sampling," *Technometrics*, 29, 143–151.
- Sacks, J., and Ylvisaker, D. (1984), "Some Model Robust Designs in Regression," *The Annals of Statistics*, 12, 1324–1348.
- Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., and Morris, M. D. (1992), "Screening, Predicting, and Computer Experiments," *Technometrics*, 34, 15–25.
- Welch, W. J., and Sacks, J. (1990), "A System for Quality Improvement Via Computer Experiments," Research Report RR-90-09, University of Waterloo, Institute for Improvement in Quality and Productivity.
- Wang, J. C., and Wu, C. F. J. (1991), "An Approach to the Construction of Asymmetrical Orthogonal Arrays," *Journal of the American Statistical Association*, 86, 450–456.
- Wu, C. F. J. (1989), "Construction of $2^{m/4}$ Designs Via a Grouping Scheme," *The Annals of Statistics*, 17, 1880–1885.
- Wu, C. F. J., Zhang, R., and Wang, R. (1992), "Construction of Asymmetrical Orthogonal Arrays of the Type $OA(s^k, s^{m_1}(s^{r_1})^{n_1} \cdots (s^{r_n})^{n_n})$," *Statistica Sinica*, 2, 203–219.