The Annals of Probability
1980, Vol. 8, No. 1, 170–175

PAIRWISE INDEPENDENT RANDOM VARIABLES[1]

BY G. L. O'BRIEN

York University, Ontario

Let $Y_1, \cdots, Y_r$ be independent random variables, each uniformly distributed on $\mathfrak{M} = \{1, 2, \cdots, M\}$. It is shown that at most $N = 1 + M + \cdots + M^{r-1}$ pairwise independent random variables, all uniform on $\mathfrak{M}$ and all functions of $(Y_1, \cdots, Y_r)$, can be defined. If $M = p^k$ for some prime p, the maximum can be attained by a strictly stationary sequence $X_1, \cdots, X_N$, for which any r successive random variables are independent.

1. Introduction. Lancaster (1965) showed that at most $n - 1$ pairwise independent nonconstant random variables can be defined on a probability space with n points, each with positive probability. He showed that the maximum can be attained for $n > 3$ but only if all the random variables take exactly two values. In this paper, we tackle a similar problem with the additional requirement that each random variable takes on M distinct values for some $M > 1$.

Our work was initiated as a result of a problem involving random number generation. Random numbers, that is independent random variables which are uniform on some set, are required in large quantities for computational problems such as simulation and Monte Carlo methods. The production of such quantities being difficult or impossible, many users of computers resort to the use of so-called pseudorandom numbers. One technique is to obtain a small set of r random numbers by some means such as rolling dice and then to generate a larger (but finite) sequence as functions of the first few. The generation procedure is often recursive in nature. It is clear that if one initially has r independent random variables all uniform on $\mathfrak{M} = \{1, 2, \cdots, M\}$, then any set of more than r of the generated random variables, all uniform on $\mathfrak{M}$, must be dependent. Thus this procedure does not generate true random numbers. Peskun (1977) recommends a particular generation procedure on the grounds that it gives independence between some pairs of random variables in the sequence and low correlations for the other pairs, provided the sequence is not too long. Further references to such procedures may be found in Sowey (1972). Here we make the requirement that the generated random variables should be pairwise independent and then investigate how long a sequence can be obtained. Our construction also provides two other useful properties, namely independence of any r successive terms in the sequence and strict stationarity.

Received July 27, 1977; revised November 22, 1978.

[1]Research supported in part by the National Research Council of Canada.

AMS 1970 *subject classifications.* Primary 60C05; secondary 65C10, 60B99, 62K10.

Key words and phrases. Pairwise independence, stationary sequences, pseudorandom numbers, block designs.

We now give some notation which will be fixed throughout. Let M and r be integers greater than 1 and let $\mathfrak{M}$ be a set with M elements. We assume $\mathfrak{M}$ is endowed with an algebraic ring structure. Let $Y_1, Y_2, \cdots, Y_r$ be independent uniform random variables on some probability space $(\Omega, \mathfrak{F}, P)$. By uniform, we mean that each Y_i takes values in $\mathfrak{M}$ and $P(Y_i = a) = M^{-1}$ for each $a \in \mathfrak{M}$. In what follows, there will be no loss of generality if we assume $(\Omega, \mathfrak{F}, P)$ is a discrete probability space with M^r points and the uniform measure. In fact, we may take Ω to be the module $\mathfrak{M}^r$ of r-tuples of elements in $\mathfrak{M}$. Then each point $\omega \in \Omega$ is determined by the r-tuple $(Y_1(\omega), Y_2(\omega), \cdots, Y_r(\omega))$. Let N_1 be the largest integer such that there exist pairwise independent random variables $X_1, X_2, \cdots, X_{N_1}$, each taking on at least M distinct values with positive probability and each a function of $Y_1, Y_2, \cdots, Y_r$. If $(\Omega, \mathfrak{F}, P)$ has the special form just indicated, then $X_1, X_2, \cdots, X_{N_1}$ may be any random variables on $(\Omega, \mathfrak{F}, P)$; they will automatically be functions of $Y_1, \cdots, Y_r$. Let N_2 be the largest integer such that there exists a stationary sequence $X_1, X_2, \cdots, ,X_{N_2}$ of pairwise independent uniform random variables, each a linear combination (in the ring $\mathfrak{M}$) of $Y_1, \cdots; Y_r$, such that any r successive terms of the sequence are independent.

We obtain lower bounds for N_2 (Theorems 2 and 3) and an upper bound for N_1 (Theorem 1). Since the two bounds are often equal and since

$$N_2 \leqslant N_1, \tag{1}$$

we obtain the actual value of N_1 and N_2 in those cases. The values of N_1 and N_2 depend on r and M, while the particular ring structure of $\mathfrak{M}$ only affects N_2. The rings we principally consider are fields or direct products of fields. When M is a power of a prime number, taking $\mathfrak{M}$ to be a field gives the maximum possible value of N_2.

The reader may note a connection between what follows and the theory of block designs and related combinatorial concepts, such as latin squares, as described for example in Chapters 10 and 13 of Hall (1967). In fact, the existence of a balanced incomplete block design, with M^2 objects and $M^2 + M$ blocks of M objects each, is equivalent to having $N_1 \geqslant M + 1$ in the case $r = 2$. In Theorem 2 we show $N_2 \geqslant M + 1$ if $r = 2$ and $\mathfrak{M}$ is a field, thereby effectively producing a block design with additional properties of interest here.

2. An upper bound for N_1.

THEOREM 1. *The following bound holds*:

$$N_1 \leqslant 1 + M + \cdots + M^{r-1} = (M^r - 1)/(M - 1). \tag{2}$$

PROOF. We may assume without loss of generality that $(\Omega, \mathfrak{F}, P)$ is a discrete probability space with M^r points and we may then consider random variables $X_1, \cdots, X_k$, each defined directly on Ω and each taking M distinct values with positive probability. Let V be the set of real-valued random variables on Ω with expectation 0. Then V is an $(M^r - 1)$-dimensional vector space. For each X_i, let V_i

be the set of real-valued random variables on Ω which may be written in the form $f = \varphi(X_i)$ for some real-valued function φ on M. Then V_i has dimension M and $V \cap V_i$ has dimension $M - 1$. Define the inner product (f, g) on V as the covariance:

$$(f, g) = Efg = \Sigma_{\omega \in \Omega} f(\omega) g(\omega) P(\{\omega\}).$$

By the independence of X_i and X_j for $i \neq j$, the subspaces $(V \cap V_i)$ and $(V \cap V_j)$ are orthogonal. The direct sum of all such subspaces has dimension $k(M - 1)$, which cannot exceed $M^r - 1$. This proves the theorem.

REMARK. As is clear from the proof, the random variables $Y_1, \cdots, Y_r$ have no essential role. The essential assumptions *are* that the X_i's each take M distinct values with positive probability and that they be defined on a set with at most M^r points.

3. A lower bound for N_2. The principal result of this section is a lower bound for N_2, in the sense that for a given M the ring $\mathfrak{M}$ can be chosen so as to make N_2 at least as large as the lower bound. We first study the case when M is the power of a prime and $\mathfrak{M}$ is a field with M elements.

LEMMA. *Let $\mathfrak{M}$ be a finite field and let $\mathfrak{M}^r$ be the r-dimensional vector space over $\mathfrak{M}$. Suppose $\mathfrak{M}^r$ is endowed with the uniform probability measure. If $X_1, X_2, \cdots, X_n : \mathfrak{M}^r \to \mathfrak{M}$ are nonzero and linear (so that they are in the dual space of $\mathfrak{M}^r$), then they are stochastically independent* iff *they are linearly independent. Also, if X_1 is nonzero and linear, then it is uniform.*

PROOF. If $X_1, \cdots, X_n$ are not linearly independent, then some X_j is a function of the others, which implies stochastic dependence. Now suppose they are linearly independent. Let $a_1, a_2, \cdots, a_n \in \mathfrak{M}$. There is a solution $x \in \mathfrak{M}^r$ to the set of equations $X_i(x) = a_i$, $i = 1, \cdots, n$. The set A of all solutions of this system is obtained by adding x to each solution of the homogeneous system $X_i(y) = 0$, $i = 1, \cdots, n$. The latter system has as its set of solutions an $(r - n)$-dimensional subspace of $\mathfrak{M}^r$, which has M^{r-n} points. Thus A has M^{r-n} points and $P(A) = (M^{r-n})(M^{-r}) = M^{-n}$. Applying the foregoing argument to the equations $X_i = a_i$ one at a time yields $P(X_1 = a_1) \cdots P(X_n = a_n) = (M^{-1})^n = P(A)$, which proves the lemma.

REMARK. There are exactly $(M^r - 1)/(M - 1)$ nonzero pairwise linearly independent linear functions from $\mathfrak{M}^r \to \mathfrak{M}$. Thus $N_1 \geqslant (M^r - 1)/(M - 1)$ when M is a power of a prime. Equality follows from (2).

In the following theorem a few basic facts are used from the theory of finite fields. The reader is referred to Chapter 5 of van der Waerden (1953) for background material.

THEOREM 2. *Let $\mathfrak{M}$ be a finite field with M elements. Then*

$$N_2 = 1 + M + \cdots + M^{r-1}. \tag{3}$$

PROOF. Let $\Omega = \mathfrak{M}^r$ be an r-dimensional vector space over $\mathfrak{M}$ and let $Y_i : \Omega \to \mathfrak{M}$ be the ith component map. Let $\mathfrak{K}$ be the dual vector space of Ω. Let $\mathfrak{M}_1$ denote the field with M^r elements. We may regard $\mathfrak{M}$ as a subfield of $\mathfrak{M}_1$, in which case $\mathfrak{M}_1$ may be treated as an r-dimensional vector space over $\mathfrak{M}$. As vector spaces, $\mathfrak{M}_1$ and $\mathfrak{K}$ have the same dimension so that there exists a one-to-one linear map T from $\mathfrak{M}_1$ onto $\mathfrak{K}$. The multiplicative group of nonzero elements of $\mathfrak{M}_1$ is cyclic; let g be a generator of this group. Finally, let X be a nonzero element of $\mathfrak{K}$ and define X_n in $\mathfrak{K}$ by

$$X_n = T(g^n T^{-1}(X)) \tag{4}$$

for $n > 0$. (In (4), g^n and $T^{-1}(X)$ are multiplied as elements of $\mathfrak{M}_1$ and the product is then mapped to $\mathfrak{K}$ by the action of T. An alternative approach would be to induce a multiplication in $\mathfrak{K}$ by means of T, to let h be a generator of the induced multiplicative group of nonzero elements of $\mathfrak{K}$, and then let $X_n = h^n X$).

Assume P is the uniform probability measure on Ω. Since each X_n is a nonzero element of $\mathfrak{K}$, each is uniform by the lemma. We will show that the sequence $\{X_n\}$ has the properties needed to prove the theorem.

Suppose that X_n and X_m are linearly dependent for some n and m with $1 \leqslant m < n$. Then $X_n = cX_m$ for some $c \in \mathfrak{M}$. Since T is linear, we obtain from (4) that

$$g^n T^{-1}(X) = cg^m T^{-1}(X).$$

Since $T^{-1}(X)$ is a nonzero element of the field $\mathfrak{M}_1$, we deduce that $g^n = cg^m$. Since the nonzero elements of $\mathfrak{M}$ form a group of order $M - 1$, $g^{(n-m)(M-1)} = c^{M-1} = 1$. Since g is a generator of a cyclic group with $M^r - 1$ elements, this implies that $(n - m)(M - 1) \geqslant M^r - 1$ and hence that $n > (M^r - 1)(M - 1)^{-1}$. Thus, the first $(M^r - 1)(M - 1)^{-1}$ terms of the sequence $\{X_n\}$ are pairwise linearly independent and, by the lemma, pairwise stochastically independent.

Suppose X_{m+i} is a linear combination of $X_{m+1}, X_{m+2}, \cdots X_{m+i-1}$ for some $m \geqslant 0$ and $i > 1$, say

$$X_{m+i} = \Sigma_{j=1}^{i-1} b_j X_{m+j},$$

where $b_1, b_2, \cdots, b_{i-1} \in \mathfrak{M}$. By the definition of T,

$$g^{m+i} T^{-1}(X) = \Sigma_{j=1}^{i-1} b_j g^{m+j} T^{-1}(X).$$

Multiplication by g^k for $k \geqslant -m$ and application of T yields

$$X_{m+i+k} = \Sigma_{j=1}^{i-1} b_j X_{m+j+k}. \tag{5}$$

By induction we see that for $k \geqslant 0$, X_{m+i+k} is also a linear combination of $X_{m+1}, \cdots, X_{m+i-1}$. Now it is clear from the definition of $\{X_n\}$ that this sequence is periodic with period $M^r - 1$ and that any $M^r - 1$ successive terms are distinct elements of $\mathfrak{K}$. It follows that $X_{m+1}, \cdots, X_{m+i-1}$ spans $\mathfrak{K}$ so that $i > r$. This implies that any r successive terms in the sequence are linearly (and hence stochastically) independent.

If $n \leqslant r$, the last paragraph shows that $X_1, \cdots, X_n$ and $X_{t+1}, \cdots, X_{t+n}$ have the same joint distributions for any $t > 0$. For $n > r$, X_n is the same linear combination of $X_1, \cdots, X_r$ as X_{t+n} is of $X_{t+1}, \cdots, X_{t+r}$, by an argument like that used to prove (5). Thus $\{X_n\}$ is strictly stationary.

We have shown that $N_2 \geqslant (M^r - 1)(M - 1)^{-1}$. Equality follows from (1) and (2).

Theorem 2 can easily be modified to give a result for some rings with M elements for M not a power of a prime. If $r = 2$ and $M = 6$, it was shown by Tarry (1900) that $N_1 = 3$. Thus (3) does not hold for every M no matter what ring of M elements is used. The following theorem shows that with an appropriate ring we can obtain a lower bound for N_2 that in the case $r = 2$ was once conjectured to be the value of N_1 (c.f. Hall (1967), page 192):

THEOREM 3. *Let $M = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ where $p_1, \cdots, p_m$ are distinct primes and $p_1^{k_1} < p_2^{k_2} < \cdots < p_m^{k_m}$. Let $\mathfrak{M}$ be the direct product of the fields $\mathfrak{M}_1, \cdots, \mathfrak{M}_m$ where $\mathfrak{M}_i$ has $M_i = p_i^{k_i}$ elements. Then*

$$N_2 \geqslant 1 + M_1 + \cdots + M_1^{r-1}.$$

PROOF. The random variables $Y_1, \cdots, Y_r$ can each be considered to be ordered m-tuples with independent uniform components in $\mathfrak{M}_1, \cdots, \mathfrak{M}_m$. Define $X_1, X_2, \cdots$ separately for each component using Theorem 2. The linear form of $X_1, X_2, \cdots$ is assured by the component-wise nature of the ring operations in $\mathfrak{M}$. It is easily checked that $X_1, X_2, \cdots$ is a stationary sequence of uniform random variables, that any r successive ones are independent and that the first $(M_1^r - 1)(M_1 - 1)^{-1}$ are pairwise independent.

FURTHER REMARKS. Let M be a prime and let $r \in \{2, 3, \cdots, M - 1\}$. Joffe (1974) constructs a sequence $X_1, X_2, \cdots, X_{M+1}$ of uniform random variables as functions of $Y_1, \cdots, Y_r$ such that any r of X_i's are independent. For $r = 2$, his result coincides with our result that $N_1 = M + 1$.

One might suppose that the uniformity requirements are not essential in the above construction. That this supposition is false is shown by looking at the case $M = r = 2$. It is easily seen that three pairwise independent random variables which each take two distinct values with positive probability can be defined on a probability space consisting of four points if and only if the space has the uniform probability measure on the four points. If it does, then, of course, the three random variables must be uniform.

Although finding the element g used in the proof of Theorem 2 involves a finite procedure, it is difficult for large M and r. Consequently, using the proof for the generation of pseudorandom numbers is difficult. One fact worth noting is that it is only necessary to go through the procedure once for any M and r; several sequences can be drawn for the same g.

Acknowledgments. This paper could not have been written without considerable assistance from Professors A. Karrass and A. Pietrowski, who assisted me on

the original proof of Theorem 2. I am grateful to the referee for indicating a proof of that theorem which involves much less field theory than the original. I also benefitted from discussions with Professors D. Bean, A. Brunner, R. Burns, D. A. Dawson, H. O. Lancaster, T. MacHenry, M. Muldoon, R. Pyke and D. Solitar.

REFERENCES

[1] HALL, M. JR. (1967). *Combinatorial Theory*. Blaisdell, Waltham, Massachusetts.

[2] JOFFE, A. (1974). On a set of almost deterministic k-independent random variables. *Ann. Probability* **2** 161–162.

[3] LANCASTER, H. O. (1965). Pairwise statistical independence. *Ann. Math. Statist.* **36** 1313–1317.

[4] PESKUN, P. H. (1977). Improving the apparent randomness of pseudo-random numbers generated by the mixed congruential method. In *Proceedings of Computer Science and Statistics: Tenth Annual Symposium on the Interface*. (D. Hogben and D. Fife, eds.) 323–328. National Bureau of Standards Special Publication 503, Gaithersburg, Maryland.

[5] SOWEY, E. R. (1972). A chronological and classified bibliography on random number generation and testing. *Internat. Statist. Rev.* **40** 355–371.

[6] TARRY, G. (1900). Le problème des 36 officiers. *C. R. Assoc. Fr. Av. Sci.* **1** 122–123.

[7] VAN DER WAERDEN, B. L. (1953). *Modern Algebra, I*. Ungar. New York.

DEPARTMENT OF MATHEMATICS
YORK UNIVERSITY
DOWNSVIEW, ONTARIO
M3J 1P3
CANADA