Technical Note

1969-49

C. M. Rader

A New Method of Generating Gaussian Random Variables by Computer

18 September 1969

Prepared under Electronic Systems Division Contract AF 19(628)-5167 by

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Lexington, Massachusetts

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

A NEW METHOD OF GENERATING GAUSSIAN RANDOM VARIABLES BY COMPUTER

C. M. RADER

Group 62

TECHNICAL NOTE 1969-49

18 SEPTEMBER 1969

LEXINGTON MASSACHUSETTS

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the Department of the Air Force under Contract AF 19(628)-5167.

ABSTRACT

It is relatively easy to generate, by digital computer, large numbers of seemingly independent random numbers with a uniform distribution over a fixed range, say $-\frac{1}{2} < X_n < \frac{1}{2}$. Methods of generating gaussian, or normal, random numbers generally are based on either non-linear transformations on random numbers from a uniform population, or the summing of enough independent numbers from a uniform population for the central limit theorem to be applicable. In the first case a time-consuming evaluation of a complicated function is involved. The second method is also slow because a large number of uniform random variables must be generated and summed for each normal random variable obtained. This note discloses a method based on the central limit theorem, except that the summing of N uniform random variables gives N normal random variables. The approach is to form an N dimensional vector whose components are uniform random variables, multiply the vector by a Hadamard matrix, and use the resulting components as normal random variables. It can be shown that the resulting N components have a uniform density inside a N-dimensional hypercube aligned with diagonals along the coordinate axes. However, the one dimensional marginal densities, the two dimensional marginal densities, indeed all the marginal densities tend toward the normal density as N gets large. Furthermore the components are uncorrelated and have equal variance, independent of N. However, some of the fourth moments, which should be zero for independent

normal random variables, are not zero for our derived set (although these moments do approach zero as N becomes large). These moments can be made zero, however, by randomly changing, or not changing, the sign of each component.

The method proposed is very fast because the principal step, Hadamard matrix multiplication, requires only $N \log_2 N$ additions to produce N components.

Accepted for the Air Force
Franklin C. Hudson
Chief, Lincoln Laboratory Office

A NEW METHOD OF GENERATING GAUSSIAN RANDOM VARIABLES BY COMPUTER

Many scientific computations require the generation by a computer of a large number of seemingly random numbers[1] with a multidimensional normal (gaussian) probability density function. However, although there are efficient algorithms available for generating uniformly distributed random numbers, the algorithms which generate normal random numbers are generally much slower. One typical algorithm begins by generating a moderate number of uniform random variables (say N) and forming the sum, which, for large N, is approximately normal, according to the central limit theorem. This method is slow because many uniformly distributed random variables must be developed for each resulting normal random variable, and because many additions must be performed for each normal random variable, as well. Another typical method, more attractive because it is less dependent on approximation, begins with a single uniform random variable and performs a non-linear transformation of the random variable to form a new random variable with a normal density. Unfortunately, the non-linear function which must be evaluated is complicated and cannot be computed quickly on most digital computers.

This note describes a variation of the central limit theorem approach, in which N uniform random variables (r.v.) are transformed into N approximately normal r.v. by a Hadamard transformation.

1. The term "pseudorandom" is often used to describe the deterministically generated, but seemingly random numbers.

This operation takes $\frac{1}{2}$ N $\log_2$ N additions and $\frac{1}{2}$ N $\log_2$ N subtractions, which is to say, $\log_2$ N additions or subtractions per normal r.v. obtained, and, of course, it is only necessary to generate one uniform r.v. for each normal r.v. obtained. The plan of the paper is to first describe the ideal uniform and normal multivariate densities, second to describe the Hadamard transformation used, third to derive the properties of the transformed variables, and fourth, to consider the properties of the transformed variables after they have been subjected to random sign changes.

I. Properties of the Ideal Uniform and Normal Densities

We define the ideal uniform r.v., X_n, as one whose probability density is

$$P_x(X_n) = \begin{cases} 1 & |X_n| < \frac{1}{2} \\ 0 & |X_n| > \frac{1}{2} \end{cases} \tag{1}$$

For several such r.v.s $X_0, X_1, \ldots, X_{N-1}$ to be independent implies that the joint density is

$$P_{xx\ldots x}(X_0, X_1, X_2, \ldots, X_{N-1}) = \begin{cases} 1 \text{ if all } |X_n| < \frac{1}{2} \\ 0 \text{ if any } |X_n| > \frac{1}{2} \end{cases} \tag{2}$$

On an N dimensional space, the joint density is unity on the inside and zero on the outside of a hypercube centered on the origin with coordinate axes passing from the center of each face to the center of the opposite face.

The expected value, $E(X_n) = \overline{X}_n$, is zero and, by integration, all the other moments can be found. For the case of even order moments of X_n,

$$\overline{X_n^{2p}} = \frac{1}{2^{2p}(2p+1)} \tag{3}$$

and for odd order moments

$$\overline{X_n^{2p+1}} = 0 \tag{4}$$

Since the mean is zero, the concept of central moments is superfluous. For joint moments, we appeal to independence, which says that the moment of a product is the product of the moments. Therefore

$$\overline{(X_a^m X_b^n X_c^\ell)} = (\overline{X_a^m})(\overline{X_b^n})(\overline{X_c^\ell}) \tag{5}$$

and, of course, these joint moments are zero except when m, n, ℓ are all even.

It is worth taking note of the second moment, $\overline{X_n^2} = 1/12$.

The normal probability density of an r.v. Y_n, with zero mean, is given by

$$P_y(Y_n) = \frac{1}{\sqrt{2\pi}\,\sigma}\ e^{-Y_n^2/2\sigma^2} \tag{6}$$

where σ^2 is the variance. We shall be comparing several densities in this note and we shall force the variances to all be equal to 1/12. Therefore $\sigma^2 = 1/12$.

The multivariate density of several such independent variables, $Y_0, Y_1, \ldots, Y_{N-1}$, is

$$P_{yyy..y}(Y_0, Y_1, \ldots, Y_{N-1}) =$$

$$\frac{1}{\sqrt{2\pi}^N \sigma^N} e^{-(Y_0^2+Y_1^2+\ldots+Y_{N-1}^2)/2\sigma^2} \tag{7}$$

The odd order one dimensional moments of Y_n are

$$\overline{Y_n^{2p+1}} = 0 \tag{8}$$

while the even order moments may be shown to be

$$\overline{Y^{2p}} = \frac{1 \cdot 3 \cdot 5 \cdots (2p-1)}{(1/\sigma)^{2p}} = \frac{1 \cdot 3 \cdot 5 \cdots (2p-1)}{12^p} \tag{9}$$

Again, by independence, the joint moments are equal to the product of the individual moments, so that the moment

$$\overline{Y_a^m Y_b^n Y_c^\ell} = (\overline{Y_a^m})(\overline{Y_b^n})(\overline{Y_c^\ell}) \tag{10}$$

is zero unless m, n, ℓ are all even.

The normal density is never zero, and is, in fact, radially symmetrical, i.e., $P_{yyy..y}(Y_0,Y_1,\ldots,Y_{N-1})$ is only a function of the distance of $Y_0,Y_1,\ldots,Y_{N-1}$ from the origin in N-space. This is also a true statement for any marginal density. The same statement is not true for the multivariate uniform density, which takes on different probabilities in the directions of the vertices of the hypercube than along the direction of the faces of the cube.

II. The Hadamard Matrix

By definition, a Hadamard matrix is an orthogonal matrix whose elements are all either +1 or -1. Hadamard matrices do not exist for all dimensionalities; for example there is no 3×3 Hadamard matrix. It is relatively easy to construct an $N \times N$ Hadamard matrix, however, if N is a power of two. We shall content ourselves with a particular Hadamard matrix for each power of two, for our purpose is not to investigate Hadamard matrices but to apply them. Letting $N = 2^c$, and numbering our dimensions in the binary system

$$\begin{aligned} i &= i_0 + 2i_1 + 4i_2 + \ldots + 2^{c-1}i_{c-1} \qquad 0 \le i \le N-1 \\ j &= j_0 + 2j_1 + 4j_2 + \ldots + 2^{c-1}j_{c-1} \qquad 0 \le j \le N-1 \end{aligned} \tag{11}$$

the elements of our particular Hadamard matrix, H, are given by

$$h_{i,j} = (-1)^{i_0 j_0} (-1)^{i_1 j_1} (-1)^{i_2 j_2} \ldots (-1)^{i_{c-1} j_{c-1}} \quad (12)$$

We shall show in an appendix that multiplication of an N component vector by H can be accomplished with N $\log_2$ N additions and subtractions.

It is clear that $h_{i,j} = h_{j,i}$ so that H is a symmetric matrix. Also, the zeroth row and the zeroth column are composed of +1's only. The product of the elements in the i^{th} row with the k^{th} row give the elements of another row of the Hadamard matrix.

$$h_{i,j}\, h_{k,j} = h_{\ell,j} \quad (13)$$

where ℓ is the row whose number is

$$\ell = (i_0 \oplus k_o) + 2(i_1 \oplus k_1) + \ldots$$

$$+ \; 2^{c-1}(i_{c-1} \oplus k_{c-1}) \quad (14)$$

where $\oplus$ means "exclusive or". We shall say as a shorthand,

$$\ell = i \oplus k$$

in place of (14). By symmetry, the product of columns is also a column.

A consequence of (13) is the orthogonality of the Hadamard matrix. Strictly speaking the matrix

$$\frac{1}{\sqrt{N}} H$$

is not only its own transpose but its own inverse, and it is the matrix which we shall use in the transformation of uniform r.v. to normal r.v.

We give as an example the 4 x 4 Hadamard matrix

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

III. Hadamard Generation of Almost Normal r.v.

We begin with a set of uniformly distributed r.v. assumed independent. We form a new set of r.v. from the uniform set by the operation

$$W_m = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} h_{n,m} X_n \qquad m = 0,1,\ldots,N-1 \qquad (15)$$

and we shall be interested in the properties of the set of r.v. W_m. For any single random variable, it is easy to see that $P_w(W_m)$ is the density of a sum of independent r.v. with zero mean and

equal variance. It is known from the central limit theorem that such a sum approaches a gaussian r.v. in the limit of large N. The approximation is good in the region of W_m near zero, bad in the tails. It is, of course, clear that

$$|W_m| \leq \frac{1}{2}\sqrt{N} \tag{16}$$

so that $P_w(W_m)$ is zero outside this range. However, for large N, $\frac{1}{2}\sqrt{N}$ is many standard deviations from the mean, so that the difference can be tolerated. Indeed, the approximation is the same as for the common method of generating normal r.v. except that we hope to use a larger value of N and so obtain a good approximation.

We have generated, however, N different r.v. from the same N uniform r.v. and we should not expect the W_m's to be independent. This leads us to investigate the multivariate density $P_{ww\ldots w}(W_0,W_1,\ldots,W_{N-1})$. But

$$P_{ww\ldots w}(w_0,W_1,\ldots,W_{N-1}) = P_{xx\ldots x}(X_0,X_1,\ldots,X_{N-1}) \Big/ |D|$$

where

$$D = \begin{vmatrix} \frac{\partial W_o}{\partial X_o}, & \cdots, & \frac{\partial W_o}{\partial X_{N-1}} \\ \vdots & & \vdots \\ \frac{\partial W_{N-1}}{\partial X_0}, & \cdots, & \frac{\partial W_{N-1}}{\partial X_{N-1}} \end{vmatrix}$$

$$P_{ww\ldots w}(W_0,W_1,\ldots,W_{N-1}) = P_{xx\ldots x}(X_0,X_1,\ldots,X_{N-1}) \tag{17}$$

since the magnitude of the determinant of $\frac{1}{\sqrt{N}}$ H is unity. Of course, in (17), $P_{xx\ldots x}(X_0,X_1,\ldots,X_{N-1})$ is to be evaluated at

$$\begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix} = \frac{1}{\sqrt{N}} H \begin{bmatrix} W_0 \\ W_1 \\ \vdots \\ W_{N-1} \end{bmatrix}$$

but this is only a change of coordinate system. In fact, therefore, the N-dimensional joint density $P_{ww\ldots w}$ is not normal but uniform. The density is one inside and zero outside a hypercube. This hypercube, however, is rotated. Whereas for the variables X_n, the hypercube was situated with the coordinate axes passing through its faces, for the W_n's we shall see that the coordinate axes pass through vertices of the hypercube. In fact, the coordinates of some of the vertices of the hypercube surrounding $P_{xx\ldots x}(X_0,X_1,\ldots,X_{N-1})$ are

$$X_j = \frac{1}{2} h_{i,j}$$

for any i, and such a vertex becomes rotated into the position with coordinates

$$W_j = \begin{cases} \frac{1}{2}\sqrt{N} & j = i \\ 0 & j \neq i \end{cases}$$

which lies on the coordinate axis of the i^{th} coordinate in W-space. The vertex at $-X_j$ rotates to the opposite position, thus proving that the coordinate axes of W-space pass through diagonals of the hypercube.

A hypercube in N-dimensions has 2^N vertices. There are, however, only N coordinate axes, so only 2N vertices can lie on the coordinate axes. The remaining 2^N - 2N vertices of the hypercube do not map onto the coordinate axes. At least some of the vertices are found in the "interior" of hyperquadrants, but not every hyperquadrant of the N-space can be lucky enough to have a vertex within it, for there are 2^N hyperquadrants and at most 2^N - 2N vertices to go around. Therefore, we learn that $P_{ww\ldots w}(W_0,W_1,\ldots,W_{N-1})$ is not only radially asymmetrical but is not symmetrical in each hyperquadrant of W-space. This shows up when we compute the moments of $P_{ww\ldots w}$. However, the most important moment of $P_{ww\ldots w}$, namely $\overline{W_m W_n}$, is zero, showing that the variables W_m are at least uncorrelated. We shall now compute all the first and second moments of $P_{ww\ldots w}(W_0,W_1,\ldots,W_{N-1})$.

The first moments are all equal and are all zero,

$$\overline{W}_m = \overline{\frac{1}{\sqrt{N}} \sum_n h_{n,m} X_n} = \frac{1}{\sqrt{N}} \sum_n h_{n,m} \overline{X}_n = 0 \qquad (18)$$

because $\overline{X}_n = 0$. The second moments are

$$\overline{W_a W_b} = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} h_{m,a}\, h_{n,b}\, \overline{X_m X_n} \tag{19}$$

but $\overline{X_m X_n}$ is zero unless m = n, in which case it is σ^2. Thus (19) becomes

$$\overline{W_a W_b} = \frac{1}{N} \sum_{m=0}^{N-1} h_{m,(a \oplus b)}\, \sigma^2 \tag{20}$$

By definition of a $\oplus$ b, it is zero if a = b, and the zeroth column of a Hadamard matrix consists of N ones. Therefore

$$\overline{W_a^2} = \sigma^2 \tag{21}$$

as expected. But if a $\neq$ b, a $\oplus$ b is not zero and since all the other columns of H contain as many -1's as 1's, the sum

$$\sum_m h_{m,(a \oplus b)}$$

is zero, giving

$$\overline{W_a W_b} = 0 \tag{22}$$

as stated above.

It is not difficult to argue that all the odd order moments of $P_{ww\ldots w}$ are zero. We focus now on fourth moments. By definition

$$\overline{W_a W_b W_c W_d} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} h_{i,a} h_{j,b} h_{k,c} h_{\ell,d} \overline{X_i X_j X_k X_\ell} \tag{23}$$

To evaluate the fourth moments we note that $\overline{X_i X_j X_k X_\ell} = 0$ for most combinations of i,j,k,ℓ with the exceptions listed below:

1. $i=j$, $k=\ell$, $i\neq k$ for which $\overline{X_i^2 X_k^2} = \sigma^4 = \frac{1}{144}$
2. $i=k$, $j=\ell$, $i\neq j$ for which $\overline{X_i^2 X_j^2} = \sigma^4 = \frac{1}{144}$
3. $i=\ell$, $j=k$, $i\neq j$ for which $\overline{X_i^2 X_j^2} = \sigma^4 = \frac{1}{144}$
4. $i=j=k=\ell$ for which $\overline{X_i^4} = \frac{1}{80}$

These are all disjoint cases. Therefore

$$\begin{aligned}
\overline{W_a W_b W_c W_d} = {} & \frac{1}{N^2} \left(\frac{1}{12}\right)^2 \left(\sum_{i=0}^{N-1} h_{i,a} \theta b \left(\sum_{\substack{k=0 \\ k\neq i}}^{N-1} h_{k,c} \theta d \right) \right) \\
& + \frac{1}{N^2} \left(\frac{1}{12}\right)^2 \left(\sum_{i=0}^{N-1} h_{i,a} \theta c \left(\sum_{\substack{j=0 \\ j\neq i}}^{N-1} h_{k,b} \theta d \right) \right) \\
& + \frac{1}{N^2} \left(\frac{1}{12}\right)^2 \left(\sum_{i=0}^{N-1} h_{i,a} \theta d \left(\sum_{\substack{j=0 \\ j\neq i}}^{N-1} h_{j,b} \theta c \right) \right) \\
& + \frac{1}{N^2} \left(\frac{1}{80}\right) \sum_{i=0}^{N-1} h_{i,a} \theta b \theta c \theta d
\end{aligned} \tag{24}$$

Equation (24) looks very complicated but it may be straightforwardly evaluated for any special case. Thus, for $\overline{W_a^4}$ we have a=b=c=d and we find

$$\overline{W_a^4} = \frac{1}{N^2}\left(\frac{1}{12}\right)^2 (N^2-N) + \frac{1}{N^2}\left(\frac{1}{12}\right)^2 (N^2-N) +$$

$$\frac{1}{N^2}\left(\frac{1}{12}\right)^2 (N^2-N) + \frac{1}{N^2}\left(\frac{1}{80}\right)N$$

$$\overline{W_a^4} = \frac{3}{144} + \frac{1}{N}\left(\frac{1}{80} - \frac{3}{144}\right) = 3\sigma^4 - \frac{\sigma^2}{10N} \qquad (25)$$

Similarly, if we let a=c, b=d, $a \neq b$ we get

$$\overline{W_a^2 W_b^2} = \frac{1}{N^2}\left(\frac{1}{12}\right)^2 (N^2-N) + \frac{1}{N^2}\left(\frac{1}{80}\right)N$$

$$\overline{W_a^2 W_b^2} = \frac{1}{144} + \frac{1}{180N} = \sigma^4 + \frac{\sigma^2}{15N} \qquad (26)$$

These results approach the gaussian independent results $3\sigma^4$, σ^4 for large N. But consider the curious case of a,b,c,d all unequal but with $a \ominus b \ominus c \ominus d = 0$. For this

$$\overline{W_a W_b W_c W_d} = \frac{1}{80N} \qquad (27)$$

An example of such a fourth moment is $\overline{W_0 W_1 W_2 W_3}$. This moment also approaches the expected result for independent normal r.v. as $N \to \infty$. However, it is somewhat disturbing that it is non-zero. We can, after all, convince ourselves that if a certain moment is within, say, 2% of the expected result for normal independent r.v. that this is somehow good enough. But what can we say if the expected moment should be zero? It is not clear how important this is, and indeed it probably depends on the application for which the r.v. are desired.

Before considering the effect of a random sign change, it is well to summarize what we have shown so far. We have proposed a method by which uniformly distributed zero-mean random noise can be converted to approximately normal zero mean random noise. The one dimensional marginal densities are sums of N independent uniform r.v. and thus tend nicely toward normal for large N. The r.v. are also uncorrelated (in the sense that the expected value of a product is zero). Furthermore, based on the moments we have computed, the difference between a given joint moment and the same joint moment for independent gaussian noise seems to go to zero as 1/N. N can reasonably be made quite large since the amount of computation is proportional to $\log_2 N$ per normal r.v. generated.

IV. HADAMARD GENERATION OF MORE NEARLY INDEPENDENT RANDOM VARIABLES

Let us now suppose that we have generated N random variables W_m. We have seen that some of the moments $\overline{W_a W_b W_c W_d}$, which would be zero for an independent gaussian set, are non-zero. We can form a set of N new random variables from the W's by changing the sign of each W, or not changing it, at random. That is, consider an easily obtained set of r.v., r_m, independent of each other and of W_m, with equal likelihood of being +1 or -1. For these r.v. the moments are all either zero or one. Even order moments are equal to 1,

$$\overline{r_m^{2p}} = 1 \quad ,$$

odd order moments are equal to zero,

$$\overline{r_m^{2p+1}} = 0 \quad ,$$

and, by independence, the moment of the product of different r_m's is the product of the moments. Now consider the set of random variables

$$V_m = r_m W_m \quad . \tag{28}$$

We shall compute some of the moments of the joint density $P_{vv\ldots v}(V_0, V_1, \ldots . V_{N-1})$ to show that this density is nearly gaussian, for large N. First consider the moments which violated the gaussian assumption for the W's. The $\overline{V_a V_b V_c V_d}$ is equal to

$$\overline{r_a r_b r_c r_d W_a W_b W_c W_d} = (\overline{r_a r_b r_c r_d})\ (\overline{W_a W_b W_c W_d})$$

and unless a, b, c, d are equal at least in pairs, $\overline{r_a r_b r_c r_d}$ will be zero. A similar argument will show that all joint moments of $P_{vv\ldots v}$ which would be zero for independent, zero mean gaussian noise are indeed zero. However, all other moments will be unaffected by the conversion of W_m to V_m. Thus $\overline{V_a^2} = \sigma^2$, $\overline{V_a V_b} = 0$, and

$$\begin{aligned} \overline{V_a^2 V_b^2} &= \overline{W_a^2 W_b^2} = \sigma^4 + \frac{\sigma^2}{15N} \\ \overline{V_a^4} &= \overline{W_a^4} = 3\sigma^4 - \frac{\sigma^2}{10N} \end{aligned} \tag{29}$$

We next consider the 6^{th} order moments

$$\overline{V_a V_b V_c V_d V_e V_f} = \overline{r_a r_b r_c r_d r_e r_f}\ \ \overline{W_a W_b W_c W_d W_e W_f}$$

Discarding the trivially zero moments, we have for the other

$$\overline{V_a V_b V_c V_d V_e V_f} = \frac{1}{N^2} \sum_{i\,j\,k\,\ell\,m\,n} h_{i,a}\, h_{j,b}\, h_{k,c}\, h_{\ell,d}$$

$$h_{m,e}\, h_{n,f}\, \overline{X_i X_j X_k X_\ell X_m X_n}$$

We must again consider moments $\overline{X_i X_j X_k X_\ell X_m X_n}$, most of which are zero. The exceptions are listed below with the equal subscripts grouped and the ungrouped subscripts unequal. Thus

Disjoint Cases	Moment
i j k ℓ m n	$\frac{1}{448}$
i j k ℓ m n	
i j k m ℓ n	
i j k n ℓ m	
i j ℓ m k n	
i j ℓ n k m	$\frac{1}{960}$
i k ℓ n j n	
i k ℓ n j m	
j k ℓ m i n	
j k ℓ n i m	
k ℓ m n i j	
j ℓ m n i k	
j k m n i ℓ	
i k m n j ℓ	
i ℓ m n j k	
i j m n i ℓ	

Disjoint Cases				Moment
i j	k ℓ	m n		
i j	k m	n ℓ		
i j	k n	m ℓ		$\frac{1}{1728}$
i k	j ℓ	m n		
i k	j m	ℓ n		
i k	j n	ℓ m		
i ℓ	j k	m n		
i ℓ	j m	n k		
i ℓ	j n	m k		
i m	j k	ℓ n		
i m	j ℓ	k n		
i m	j n	k ℓ		
i n	j k	ℓ m		
i n	j ℓ	k m		
i n	j m	k ℓ		

This table allows us to compute all the 6^{th} order moments for special cases. For example, if we were to compute $\overline{V_a^6}$, we would have as part of the computation a sum

$$\frac{1}{N^3} \sum_{i=0}^{N-1} h_i, \; (a \; \theta \; a \; \theta \; a \; \theta \; a \; \theta \; a \; \theta \; a) \; \overline{X_i^6} = \frac{1}{448} \frac{1}{N^2}$$

and fifteen sums similar to

$$\frac{1}{N^3} \sum_{i=0}^{N-1} h_i (a \ominus a \ominus a \ominus a) \sum_{\substack{m=0 \\ m\neq i}}^{N-1} h_m, \ a \ominus a \ \overline{X_i^4 X_m^2}$$

$$= \frac{1}{960} \left(\frac{N^2-N}{N^3}\right)$$

and fifteen sums similar to

$$\frac{1}{N^3} \sum_{i=0}^{N-1} h_i, \ (a \ominus a) \sum_{\substack{k=0 \\ k\neq i}}^{N-1} h_k, \ (a \ominus a) \sum_{\substack{m=0 \\ m\neq i \\ m\neq k}}^{N-1} h_m, \ (a \ominus a) \overline{X_i^2 X_k^2 X_m^2}$$

$$= \frac{1}{1728} \ \frac{N(N-1)\ (N-2)}{N^3}$$

Gathering all the terms together, we get, for the 6^{th} moment

$$\overline{V_a^6} = \frac{15}{1728} \left(\frac{N(N-1)(N-2)}{N^3}\right) + \frac{15}{960} \left(\frac{N(N-1)}{N^3}\right) + \frac{1}{448} \left(\frac{N}{N^3}\right)$$

$$\overline{V_a^6} = 15\sigma^6 + \text{terms in } \frac{1}{N} \text{ and } \frac{1}{N^2} \qquad (30)$$

Evaluating $\overline{V_a^4 V_b^2}$ we set, for example, a = c = d = e, b = f, $a \neq b$. Therefore, the sum

$$\frac{1}{N^3}\sum_{i=0}^{N-1} h_i,\ (a\ \theta\ b\ \theta\ a\ \theta\ a\ \theta\ a\ \theta\ b)\ \overline{X_i^6} = \frac{1}{448}\ (\frac{N}{N^3})$$

is the same as before, but many of the other thirty sums are zero, leaving only the cases

i	j	ℓ	n	k	m
i	k	ℓ	m	j	n
j	k	ℓ	n	i	m
j	ℓ	m	n	i	k
j	k	m	n	i	ℓ
i	j	m	n	k	ℓ

and

i	k	j	n	ℓ	m
i	ℓ	j	n	m	k
i	m	j	n	k	ℓ

We quickly count

$$\overline{V_a^4\ V_b^2} = \frac{3}{1728}\ (\frac{N(N-1)(N-2)}{N^3}) + \frac{6}{960}\ (\frac{N(N-1)}{N^3}) + \frac{1}{448}\ (\frac{N}{N^3})$$

(31)

$$\overline{V_a^4\ V_b^2} = 3\sigma^6 + \text{terms in } \frac{1}{N} \text{ and } \frac{1}{N^2}$$

Similarly for $\overline{V_a^2 V_b^2 V_c^2}$ we obtain

$$\overline{V_a^2 V_b^2 V_c^2} = \frac{1}{1728}\ (\frac{N(N-1)(N-2)}{N^3}) + \frac{3}{960}\ (\frac{N(N-1)}{N^3}) + \frac{1}{448}\ (\frac{N}{N^3})$$

(32)

$$\overline{V_a^2 V_b^2 V_c^2} = \sigma^6 + \text{terms in } \frac{1}{N} \text{ and } \frac{1}{N^2}$$

These results also approach the expected results for normal r.v. with an error like 1/N, i.e., $15\sigma^6$, $3\sigma^6$, and σ^6.

At this point, we comment that the behavior, for large N, of any moment is a matter of combinatorial analysis. For example, the moment

$$\overline{V_a^{2p}} = 1 \cdot 3 \cdot 5 \cdots (2p-1) \cdot \sigma^{2p} + \text{terms in } \frac{1}{N} \text{ , etc.} \qquad (33)$$

can be derived by a counting process. For more complicated moments, the analysis is so complicated that we have not undertaken it.

We previously described the N dimensional distribution of the W's in terms of a hypercube which we argued was asymmetically distributed among the 2^N hyperquadrants in N space. We cannot make a simple statement about $P_{vv\ldots v}(V_0, V_1, \ldots V_{N-1})$ because it is not uniformly distributed, but rather has some density which is a continuous function of spacial coordinates. However, it is clearly symmetrically distributed in the hyperquadrants, by construction.

Programming Considerations

The method proposed requires at least the following two or three steps:

1. Compute N uniform r.v. $X_0, X_1, \ldots, X_{N-1}$. Suppose that each r.v. requires a time T_u to generate. Then part 1 will require a time NT_u.

2. Compute the "Hadamard transform" of the array of N numbers from part 1. This can be done in-place (i.e., without any extra memory) with $N \log_2 N$ additions or subtractions. Assuming that an addition and a subtraction take the same amount of time T_a, part 2 will require a time $N(\log_2 N)T_a$. If it is desired to adjust the variance of the resulting W_n, we should add to this time an additional time $N T_m$, where T_m is the time for a multiply (or a scale if that is acceptable).

3. If the improved approximation to independent gaussian noise is desired, generate N additional random variables r_j which can be used to randomly change or not change the sign of the r.v. W_j generated in part 2. This should take at most a time $N(T_u + T_a)$ assuming the r_j can be generated as quickly as a uniform r.v. and sign changing can be accomplished as quickly as addition or subtraction.

The total time, therefore, is like

$$T_{total} = N(2T_u + \log_2(2N)\ T_a)$$

but this is the time for generating N r.v. whereas the relevant

time would be the time per random variable. This latter time is

$$T_g = 2T_u + (1\log_2 N + 1)\ T_a \quad .$$

If many r.v. are needed, as will usually be the case, they should be generated N at a time by a subroutine. This subroutine could, of course, supply them one at a time to a user program, recomputing a block of N when its supply was exhausted. Thus, unless fewer than N r.v. were required, the limitation on computing the r.v. N at a time should not be important.

Conclusions

Random variables with an almost normal distribution can be computed by first obtaining uniform random variables, using standard computer algorithms, and then performing a linear transformation on a set of $N = 2^c$ of the uniform random variables to obtain N random variables whose joint distribution is not gaussian but uniform in a hypercubic region. Nevertheless, many of the marginal densities are nearly gaussian, in accord with the central limit theorem, and the r.v. are uncorrelated. In the limit of large N, all the moments which have been determined seem to approach the moments expected for gaussian independent r.v. However, some of the joint moments which would be zero for

independent gaussian r.v. are finite for the derived set. By randomly changing or not changing the sign of the linearly transformed r.v., a new set of r.v. is obtained for which all the moments, including joint moments, show considerable similarity to the moments expected from the gaussian independent case, for large N. The procedure has the advantage of being faster than other decent methods of generating gaussian r.v. on a digital computer, with relatively good accuracy.

APPENDIX

NUMBER OF OPERATIONS IN HADAMARD MATRIX MULTIPLICATION

If we consider the multiplication of an arbitrary vector by the matrix H, with elements $h_{i,j}$ given by (12), an abvious proceedure is to partition the matrix into four quadrants,

$$H = \begin{bmatrix} a_{i,j} & \vdots & b_{i,j} \\ \cdots\cdots & \vdots & \cdots\cdots \\ c_{i,j} & \vdots & d_{i,j} \end{bmatrix}$$

$$\left.\begin{aligned} a_{i,j} &= h_{i,j} \\ b_{i,j} &= h_{i,j} + \tfrac{N}{2} \\ c_{i,j} &= h_{i} + \tfrac{N}{2},j \\ d_{i,j} &= h_{i} + \tfrac{N}{2},j + \tfrac{N}{2} \end{aligned}\right\} \qquad \begin{aligned} 0 \le i \le \tfrac{N}{2} - 1 \\ 0 \le j \le \tfrac{N}{2} - 1 \end{aligned}$$

It is seen that for

$$a_{i,j} = (-1)^{i_0 j_0} (-1)^{i_1 j_1} \cdots (-1)^{i_{c-2} j_{c-2}} (-1)^{i_{c-1} j_{c-1}}$$

$i_{c-1} = j_{c-1} = 0$ for the entire quadrant. Therefore $a_{i,j}$ is actually an $\frac{N}{2} \times \frac{N}{2}$ Hadamard matrix. Similarly

$$a_{i,j} = b_{i,j} = c_{i,j} = -d_{i,j}$$

It follows that if the vector to be multiplied by H in a column vector $[X_i]$ we may form two shorter vectors with components

$$X_i' = X_i + X_{i+\frac{N}{2}} \qquad 0 \le i \le \frac{N}{2} - 1$$

$$X_i'' = X_i - X_{i+\frac{N}{2}} \qquad 0 \le i \le \frac{N}{2} - 1$$

and we shall find that the first $\frac{N}{2}$ components of the product are given by

$$\sum_{i=0}^{\frac{N}{2}-1} a_{i,j} \; X_i'$$

and the second $\frac{N}{2}$ components are given by

$$\sum_{i=0}^{\frac{N}{2}-1} a_{i,j} \; X_i''$$

each of which are $\frac{N}{2} \times \frac{N}{2}$ Hadamard matrix multiplications. This reduction cost us N/2 additions and N/2 subtractions. Of course $a_{i,j}$ may be similarly partitioned and this may be done repeatedly. Each reduction costs and additional $\frac{N}{2}$ additions and $\frac{N}{2}$ subtractions so that a total of $Nc = N \log_2 N$ operations is required in all.

REFERENCES

1. S. W. Golomb and L. D. Baumert, "The Search for Hadamard Matrices," Am. Math. Monthly, Vol. 70, p. 12, January 1963.

2. W. R. Crowther and C. M. Rader, "Efficient Coding of Vocoder Channel Signals Using Linear Transformation," Proc. IEEE, Vol. 54, No. 11, pp. 1594-1595, November 1966.

3. J. L. Shanks, "Computation of the Fast-Walsh-Fourier Transform," IEEE Trans. on Computers, Vol. , pp. 457-459, May 1969.

4. W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard Transform Image Coding," Proc. IEEE, Vol. 57, pp. 58-68, January 1969.

5. J. E. Whelchel and D. F. Guinn, "The Fast Fourier-Hadamard Transform and its Use in Signal Representation and Classification," Melpar, Inc., Falls Church, Va., Tech. Report PRC68-11, 1968.

6. T. E. Hull and A. R. Dobell, "Random Number Generators," Soc. Ind. App. Math, Vol. 4, pp. 230-254, 1962.

7. A. Papoulis, Probability, Random Variables, and Stochastic Processes, pp. 234-235, McGraw-Hill, N.Y., 1965.

ACKNOWLEDGMENT

My grasp of the field of probability and statistics is far from complete and I would have been unable to discuss the properties of the r.v. generated by the Hadamard method without the guidance of many other Laboratory staff. I am especially indebted to Drs. Robert Lerner and Irvin Stiglitz for many suggestions.

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Lincoln Laboratory, M.I.T.

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP
None

3. REPORT TITLE
A New Method of Generating Gaussian Random Variables by Computer

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Technical Note

5. AUTHOR(S) *(Last name, first name, initial)*
Rader, Charles M.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
18 September 1969	34	7

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF 19(628)-5167	Technical Note 1969-49
b. PROJECT NO. 649L	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* ESD-TR-69-266
d.	

10. AVAILABILITY/LIMITATION NOTICES
This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Air Force Systems Command, USAF

13. ABSTRACT

It is relatively easy to generate, by digital computer, large numbers of seemingly independent random numbers with a uniform distribution over a fixed range, say $-1/2 < X_n < 1/2$. Methods of generating gaussian, or normal, random numbers generally are based on either non-linear transformations on random numbers from a uniform population, or the summing of enough independent numbers from a uniform population for the central limit theorem to be applicable. In the first case a time-consuming evaluation of a complicated function is involved. The second method is also slow because a large number of uniform random variables must be generated and summed for each normal random variable obtained. This note discloses a method based on the central limit theorem, except that the summing of N uniform random variables gives N normal random variables. The approach is to form an N dimensional vector whose components are uniform random variables, multiply the vector by a Hadamard matrix, and use the resulting components as normal random variables. It can be shown that the resulting N components have a uniform density inside a N-dimensional hypercube aligned with diagonals along the coordinate axes. However, the one dimensional marginal densities, the two dimensional marginal densities, indeed all the marginal densities tend toward the normal density as N gets large. Furthermore the components are uncorrelated and have equal variance, independent of N. However, some of the fourth moments, which should be zero for independent normal random variables, are not zero for our derived set (although these moments do approach zero as N becomes large). These moments can be made zero, however, by randomly changing, or not changing, the sign of each component.

The method proposed is very fast because the principal step, Hadamard matrix multiplication, requires only $N \log_2 N$ additions to produce N components.

14. KEY WORDS

Gaussian processes
digital computers
Hadamard matrix
random variables