Random Numbers Generated by Linear Recurrence Modulo Two

By Robert C. Tausworthe

1. Introduction. Many situations arise in various fields of interest for which the mathematical model utilizes a random sequence of numbers, events, or both. In many of these applications it is often extremely advantageous to generate, by some deterministic means, a sequence which appears to be random, even if, upon closer and longer observation, certain regularities become evident. For example, electronic computer programs for generating random numbers to be used in Monte Carlo experiments have proved extremely useful. This article describes a random number generator of this type with several outstanding properties. The numbers are generated by modulo 2 linear recurrence techniques long used to generate binary codes for communications.

2. Linear Recurrence Relations over GF(2). Let $a = \{a_k\}$ be the sequence of 0's and 1's generated by the linear recursion relation

$$a_k = c_1 a_{k-1} + c_2 a_{k-2} + \cdots + c_n a_{k-n} \qquad (\text{mod } 2)$$

for any given set of integers c_i $(i = 1, 2, \cdots, n)$, each having the value 0 or 1. We, of course, require $c_n = 1$, and say that the sequence has degree n.

From the recursion, a_k is determined solely (for fixed c_i) by the n-tuple $(a_{k-1}, a_{k-2}, \cdots, a_{k-n})$ of terms preceding it. Similarly, a_{k+1} is a function solely of $(a_k, a_{k-1}, \cdots, a_{k-n+1})$. Each such n-tuple thus has a unique successor governed by the recursion formula, and the period of a is clearly the same as the period with which an n-tuple repeats. The period p of a linear recurring sequence obviously cannot be greater than $2^n - 1$, for the n-tuple $(0, 0, \cdots, 0)$ is always followed by $(0, 0, \cdots, 0)$. The necessary and sufficient condition that $p = 2^n - 1$ is that the polynomial

$$f(x) = 1 + c_1 x + c_2 x^2 + \cdots + x^n$$

be primitive over GF(2) [1], [2].

We shall assume in the remainder of this article that $f(x)$ is a primitive nth degree polynomial over GF(2); the sequence a is then a *maximal-length linearly recurring sequence modulo* 2. These sequences have been studied and used as codes in communications and information-theoretic studies [3], [4]. The properties of interest to us at present are the following [1], [2]:

$$(1) \qquad \sum_{k=1}^{p} a_k = \frac{p+1}{2} = 2^{n-1}.$$

(2) For every distinct set of (0, 1) integers $s_1, s_2, \cdots, s_n$, not all zero, there

Received July 10, 1964. This paper presents the results of one phase of research carried at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

exists a unique integer v $(0 \leqq v \leqq p-1)$ such that for every k, $s_1a_{k-1} + s_2a_{k-2} + \cdots + s_na_{k-n} = a_{k+v} \pmod 2$. This is often referred to as the "cycle-and-add" property.

(3) Every nonzero (0, 1) binary n-vector $(e_1, e_2, \cdots, e_n)$ occurs exactly once per period as n consecutive binary digits in a.

Note that properties (1) and (3) follow directly from the fact that each possible nonzero binary n-tuple $(a_{k-1}, a_{k-2}, \cdots, a_{k-n})$ must occur exactly once per cycle if a has period $p = 2^n - 1$.

We shall, in what follows, find it convenient to use a slightly different version of the sequence a. Let us define

$$\alpha_k = (-1)^{a_k} = 1 - 2a_k .$$

Under this transformation, we see that, if a_k takes on the values 0 and 1, then α_k takes the values +1 and −1, respectively. The properties (1), (2), and (3) are then transformed into

(1′) $$\sum_{k=1}^{p} \alpha_k = -1.$$

(2′) For every distinct set of (0, 1) integers $s_1, \cdots s_n$, not all zero, there exists a unique integer v $(0 \leqq v \leqq p-1)$ such that $\alpha_{k-1}^{s_1}\alpha_{k-2}^{s_2} \cdots \alpha_{k-n}^{s_n} = \alpha_{k+v}$.

(3′) Every ±1 binary n-vector $(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)$, *except* the all-ones vector, occurs exactly once per period as n consecutive elements in α.

3. The Boolean Transform. Let $g(\mathbf{x})$ be a ±1-valued Boolean function of (0, 1) variables $x_1, x_2, \cdots, x_n$. For any $\mathbf{s} = (s_1, s_2, \cdots, s_n)$, $s_i = 0$ or 1, define

$$\phi(\mathbf{s}, \mathbf{x}) = 2^{-n/2}(-1)^{s_1x_1 + \cdots + s_nx_n}.$$

These 2^n functions of $\mathbf{x}$, the Rademacher-Walsh functions [5], form an orthonormal basis for 2^n-space. Relative to this basis, $g(\mathbf{x})$ has components $G(\mathbf{s})$ given by

$$G(\mathbf{s}) = 2^{-n/2} \sum_{\mathbf{x}} g(\mathbf{x})\phi(\mathbf{s}, \mathbf{x}).$$

That is, $G(\mathbf{s})$ is the projection of $g(\mathbf{x})$ on $\phi(\mathbf{s}, \mathbf{x})$, normalized so that

$$\sum_{\mathbf{s}} G^2(\mathbf{s}) = 1.$$

Similarly, we have

$$g(\mathbf{x}) = 2^{n/2} \sum_{\mathbf{s}} G(\mathbf{s})\phi(\mathbf{s}, \mathbf{x}).$$

Consider the effect of setting $x_i = a_{k-i}$ in $g(\mathbf{x})$. As a function of k, a binary ±1-sequence $\{\gamma_k\} = \gamma$ is generated:

$$\begin{aligned} \gamma_k &= \sum_{\mathbf{s}} G(\mathbf{s})(-1)^{s_1a_{k-1} + \cdots + s_na_{k-n}} \\ &= \sum_{\mathbf{s}} G(\mathbf{s})\alpha_{k-1}^{s_1}\alpha_{k-2}^{s_2} \cdots \alpha_{k-n}^{s_n} . \end{aligned}$$

By (2′), we now have the fourth property basic to our analysis:

(4) $$\gamma_k = G(0) + \sum_{s \neq 0} G(s)\alpha_{k+v(s)} ,$$

where the mapping $v(s)$ of all binary nonzero n-vectors onto $\{0, 1, 2, \cdots, p - 1\}$ is one-to-one.

4. Random Number Generation. Let $a = \{a_k\}$ be the (0, 1) sequence generated by an nth degree maximal-length linear recurrence modulo 2, as described previously, and define a set of numbers of the form

$$y_k = 0 \cdot a_{qk+r-1}a_{qk+r-2} \cdots a_{qk+r-L} \qquad \text{(base 2)},$$

where r is a randomly chosen integer, $0 \leqq r \leqq 2^n - 1$ and $L \leqq n$. That is, y_k is the binary expansion of a number whose binary representation is L consecutive digits in a; successive y_k are spaced q digits apart. For reasons essential to the analysis, we restrict $q \geqq L$, and $(q, 2^n - 1) = 1$.

We can also express y_k by

$$y_k = \sum_{t=1}^{L} 2^{-t} a_{qk+r-t} .$$

Such numbers always lie in the interval $0 < y_k < 1$. Because of condition (2), the randomness of the choice of r is equivalent to the statement that the initial value y_0 is a random choice.

5. Analysis of the Generator. We shall find it convenient to work with a transformed set of numbers w_k rather than the y_k. Specifically, let $\alpha = \{\alpha_k\}$ be the ± 1 sequence corresponding to $a = \{a_\kappa\}$, and define

$$w_k = \sum_{t=1}^{L} 2^{-t} \alpha_{qk+r-t} .$$

We see that y_k and w_k are related by

$$w_k = 1 - 2^{-L} - 2y_k .$$

There is thus an easy translation between w_k and y_k.

We generally may assume, merely from the applications to which we wish to suit the numbers, that n is moderately large, so that the numbers y_n and w_n are extremely numerous. For example, if $n = 35$, there are 3.43×10^{10} of them. We wish to consider only a portion of the total number of these, say N of them, and to discover, for moderately large N, how these are distributed.

6. Correlation Properties. The mean value of w_k is easily found as

$$\begin{aligned} E(w_k) &= \frac{1}{p} \sum_{r=0}^{p-1} w_k = \frac{1}{p} \sum_{t=1}^{L} 2^{-t} \sum_{r=0}^{p-1} \alpha_{qk+r-t} \\ &= -2^{-n} \left(\frac{1 - 2^{-L}}{1 - 2^{-n}} \right), \end{aligned}$$

a number very nearly equal to zero for large n.

Define the sample autocorrelation function $\hat{R}(m)$ of w_k by

$$\hat{R}(m) = \frac{1}{N}\sum_{k=1}^{N} w_k w_{k+m}.$$

The expected value of $\hat{R}(m)$ is the true autocorrelation function $R(m)$ of the process,

$$R(m) = E[\hat{R}(m)],$$

and the value $R(0)$ is the mean-squared value of the process w_k.

$$R(0) = \frac{1}{p}\sum_{r=0}^{p-1} w_k^{\,2} = \frac{1}{p}\sum_{t=1}^{L}\sum_{u=1}^{L} 2^{-(t+u)} \sum_{r=0}^{p-1} \alpha_{qk+r-t}\,\alpha_{qk+r-u}.$$

The last sum is -1 if $t \neq u$, and p if $t = u$, by (2′). Hence

$$R(0) = \frac{1}{3} + 2^{-n}\left[\frac{1}{3}\left(\frac{1-2^{-2L}}{1-2^{-n}}\right) - \frac{(1-2^{-L})^2}{(1-2^{-n})}\right].$$

This shows that w_k has essentially the same variance as a uniformly distributed process.

Now consider $\hat{R}(m)$, $m \neq 0$. First, its mean value is

$$R(m) = E[\hat{R}(m)] = \frac{1}{pN}\sum_{k=1}^{N}\sum_{t=1}^{L}\sum_{u=1}^{L} 2^{-(t+u)} \sum_{r=0}^{p-1} \alpha_{qk+r-t}\,\alpha_{q(k+m)+r-u}$$

$$= \frac{1}{p}\sum_{t=1}^{L}\sum_{u=1}^{L} 2^{-(t+u)} \sum_{r=0}^{p-1} \alpha_r\,\alpha_{r+qm+t-u}.$$

The last sum is again -1 by (2′) unless $qm + t - u$ is a multiple of p. Obviously,

$$qm - L + 1 \leqq qm + t - u \leqq qm + L - 1.$$

Hence, if $q \geqq L$ and $m \leqq (p - L)/q$, we see that

$$0 < qm + t - u < p,$$

so $qm + t - u$ can never be a multiple of p. These conditions, mentioned earlier, shall now be assumed as one of our hypotheses. The mean value of $\hat{R}(m)$ is then

$$R(m) = -\frac{1}{p}(1 - 2^{-L})^2$$

$$= -2^{-n}\frac{(1-2^{-L})^2}{(1-2^{-n})}.$$

The mean behaviour of the process shows essentially no correlation between w_k and w_{k+m} for any nonzero integer m less in magnitude than $(p - L)/q$.

The sample autocorrelation function is a function of r, and is itself a random process; its mean-squared value for $m \neq 0$ is

$$E[\hat{R}^2(m)] = \sum_{t=1}^{L}\sum_{u=1}^{L}\sum_{i=1}^{L}\sum_{j=1}^{L} 2^{-(t+u+i+j)}\mu_{tuij},$$

where μ_{tuij} is defined by

$$\mu_{tuij} = \frac{1}{pN^2} \sum_{k=1}^{N} \sum_{l=1}^{N} \sum_{r=0}^{p-1} \alpha_{r+qk-t}\, \alpha_{r+(k+m)q-u}\, \alpha_{r+lq-i}\, \alpha_{r+(l+m)q-j} .$$

Now since we have restricted $q \geqq L$ and $1 \leqq m \leqq (p - L)/q$, there exist v_1 and v_2 such that

$$\alpha_{r+qk-t}\alpha_{r+kq+mq-u} = \alpha_{r+v_1} ,$$

$$\alpha_{r+lq-i}\alpha_{r+lq+mq-j} = \alpha_{r+v_2} .$$

For fixed values of t, u, i, and j, there is at *most* one value of l for each k such that $v_1 = v_2$, since $(q, p) = 1$. Hence

$$\mu_{tuij} \leqq \frac{1}{pN^2} [N(p + 1) - N^2]$$

produces the result, for $m \neq 0$,

$$E[\hat{R}^2(m)] \leqq (1 - 2^{-L})^4 \left(\frac{p + 1}{pN} - \frac{1}{p}\right),$$

and the value of the variance of $\hat{R}(m)$ is likewise bounded,

$$\text{var } [\hat{R}(m)] \leqq (1 - 2^{-L})^4 \left(\frac{p + 1}{pN} - \frac{1}{p} - \frac{1}{p^2}\right) < \frac{1}{N}\left(1 + \frac{1}{p}\right).$$

This indicates that the deviation of the sample autocorrelation function from its mean value is very small, and decreases inversely proportional to N.

7. The Distribution Properties. We have shown that w_k (and, consequently, y_k) has essentially the same mean and variance as a uniform distribution. Now consider actual distributions of N values of y_k on $(0, 1)$. To do this, we consider an arbitrary interval in $(0, 1)$ and observe what percentage of the N values of y_k lie in this range.

Since we are considering binary expansions of numbers, intervals of width 2^{-d} are most conveniently considered, and these will surely be sufficient to our needs. This is done efficiently by considering the first d positions of the vectors representing y_k for $k = 1, 2, \cdots , N$, and count the number of these having a specified pattern. This is equivalent to forming a Boolean function on the first d positions of y_k, whose value is, say, -1 if y_k has this initial pattern and $+1$ otherwise.

More specifically, let $(e_1 , e_2 , \cdots , e_d)$ be the initial pattern of *ones* and *zeros* we seek as a prefix to y_k . Then define the (± 1) Boolean function $g(\mathbf{x})$ by

$$g(\mathbf{x}) = \begin{cases} -1 & \text{if } x_1 = e_1 , x_2 = e_2 , \cdots , x_d = e_d , \\ 1 & \text{otherwise.} \end{cases}$$

The relative number of times $\hat{T}$ that a number y_k takes on the form $0 \cdot e_1 e_2 \cdots e_d x x \cdots x$, and thus falls in the specified interval, is

$$\hat{T} = \frac{1}{2N}\left[N - \sum_{k=1}^{N} \gamma_k\right],$$

where γ_k has the value

$$\gamma_k = G(0) + \sum_{s \neq 0} G(\mathrm{s})\alpha_{kq+r+v(s)}$$

by the Boolean transform. The expected value of $\hat{T}$ is

$$\begin{aligned} T = E[\hat{T}] &= \frac{1}{p} \sum_{r=0}^{p-1} \hat{T} \\ &= \frac{1}{2}\left[1 - G(\mathbf{0}) + \frac{1}{p} \sum_{\mathbf{s} \neq 0} G(\mathbf{s})\right] \\ &= \frac{1}{2}\left[1 - \left(\frac{p+1}{p}\right) G(\mathbf{0}) + \frac{1}{p} \sum_{\mathbf{s}} G(\mathbf{s})\right]. \end{aligned}$$

But it is easy to see from its definition that

$$g(\mathbf{0}) = \sum_{\mathbf{s}} G(\mathbf{s}),$$

and that

$$\begin{aligned} G(\mathbf{0}) &= 2^{-n} \sum_{\mathbf{x}} g(\mathbf{x}) = 2^{-n}(2^n - 2 \cdot 2^{n-d}) \\ &= 1 - 2^{-d+1}. \end{aligned}$$

Hence, we have

$$\begin{aligned} T &= \frac{1}{2}\left[1 - \left(1 + \frac{1}{p}\right)(1 - 2^{-d+1}) + \frac{g(\mathbf{0})}{p}\right] \\ &= \left(1 + \frac{1}{p}\right) 2^{-d} + \frac{1}{2p}[g(\mathbf{0}) - 1]. \end{aligned}$$

Thus, the y_k are equidistributed in the mean.

The variance about this mean can also be bounded. First, however, we compute

$$\frac{1}{p} \sum_{r=0}^{p-1} \sum_{k=1}^{N} \sum_{l=1}^{N} \gamma_k \gamma_l = \sum_{k=1}^{N} \sum_{l=1}^{N} \sum_{\mathbf{s}} \sum_{\mathbf{u}} G(\mathbf{s}) G(\mathbf{u}) \frac{1}{p} \sum_{r=1}^{p-1} \alpha_{r-1}^{s_1} \cdots \alpha_{r-n}^{s_n} \alpha_{r+t-1}^{u_1} \cdots \alpha_{r+t-n}^{u_n}$$

using $t = q(l - k)$. If $\mathbf{s} \neq \mathbf{0}$, and $\mathbf{u} \neq \mathbf{0}$, then there exist integers v_1 and v_2 such that

$$\alpha_{r-1}^{s_1} \cdots \alpha_{r-n}^{s_n} = \alpha_{r+v_1},$$

$$\alpha_{r+t-1}^{u_1} \cdots \alpha_{r+t-n}^{u_n} = \alpha_{r+v_2},$$

and for each k there is at *most* one l such that $v_1 = v_2$. Using this fact and the Schwartz inequality, we see

$$\frac{1}{p} \sum_{r=0}^{p-1} \sum_{k=1}^{N} \sum_{l=1}^{N} \gamma_k \gamma_l \leqq N^2 \left\{G^2(\mathbf{0}) \left(1 + \frac{1}{p}\right) - \frac{1}{p}\right\} + N\left(1 + \frac{1}{p}\right).$$

This calculation then places a bound on the variance of $\hat{T}$,

$$\begin{aligned} \text{var}\,[\hat{T}] &= \frac{1}{4} E \left\{\frac{1}{N} \sum_{k=1}^{N} \gamma_k - \left(1 + \frac{1}{p}\right) G(\mathbf{0}) + \frac{1}{p} g(\mathbf{0})\right\}^2 \\ &\leqq \frac{1}{4}\left\{-[1 + G^2(\mathbf{0})] \frac{1}{p}\left(1 + \frac{1}{p}\right) + \frac{2}{p}\left(1 + \frac{1}{p}\right) g(\mathbf{0}) G(\mathbf{0}) + \frac{1}{N}\left(1 + \frac{1}{p}\right)\right\}. \end{aligned}$$

If the negative terms are omitted, the inequality is stronger,

$$\text{var}\,[\hat{T}] < \frac{1}{4}\left(1 + \frac{1}{p}\right)\left\{\frac{1}{N} + \frac{2g(\mathbf{0})(1 - 2^{-d+1})}{p}\right\} < \frac{1}{4}\left(1 + \frac{1}{p}\right)\left(\frac{1}{N} + \frac{2}{p}\right)$$

and, again, the deviation from expected behavior decreases as N grows larger.

8. Higher-Order Distributions. We have seen that the numbers w_k (or y_k) are "white" and uniformly distributed. We now consider the distribution of $(y_k, y_{k-l_2}, \cdots, y_{k-l_M})$ where $0 = l_1 < l_2 < \cdots < l_M$. It can be shown that this distribution is far from uniform if $q(l_M + 1) > n$. For $q(l_M + 1) \leqq n$, however, the distribution is uniform over the unit M-cube. To show this is the case, we shall count the relative number of times $(y_k, y_{k-l_2}, \cdots, y_{k-l_M})$ lies in an arbitrary given $2^{-d_1} \times \cdots \times 2^{-d_M}$ interval. Let the initial positions in the binary expansion of y_{k+l_i} be $0 \cdot e_1^i, e_2^i, \cdots, e_{d_i}^i$ for $i = 1, 2, \cdots, M$, and define $g(\mathbf{x})$ as follows:

$$g(\mathbf{x}) = \begin{cases} -1 & \text{if } x_{l_i q+j} = e_j^i \text{ for } i = 1, 2, \cdots, M \quad \text{and} \quad j = 1, 2, \cdots, d_i, \\ +1 & \text{otherwise.} \end{cases}$$

Now since $q(l_M + 1) \leqq n$, if we let the Boolean function variables be

$$x_t = a_{qk+r-t},$$

then we can use the transformed equation

$$\gamma_k = G(\mathbf{0}) + \sum_{\mathbf{s} \neq 0} G(\mathbf{s}) \alpha_{kq+r+v(\mathbf{s})}$$

to reveal the desired properties. The previous analysis is valid, with $d = d_1 + d_2 + \cdots + d_M$. Therefore, the relative number of times $\hat{T}$ that $(y_k, y_{k-l}, \cdots, y_{k-l_M})$ lies in the specified interval has mean value

$$T = E(\hat{T}) = \left(1 + \frac{1}{p}\right) 2^{-(d_1 + \cdots + d_M)} + \frac{1}{2p}[g(\mathbf{0}) - 1]$$

and the variance about this mean is bounded by

$$\text{var}\,(\hat{T}) < \frac{1}{4}\left(1 + \frac{1}{p}\right)\left(\frac{1}{N} + \frac{2}{p}\right).$$

8. Summary. The conclusions reached by this analysis are stated in the following

THEOREM. *If $\{a_k\}$ is a $(0, 1)$ binary sequence generated by an nth degree maximal-length linear recursion relation modulo 2, if for $(q, 2^n - 1) = 1$ and $q \geqq L$, $y_k = 0 \cdot a_{kq-1} a_{kq-2} \cdots a_{qk-L}$ is the binary expansion of a real positive number in the interval $(0, 1)$, and if w_k is a real number in the interval $(-1, +1)$ related to y_k by $w_k = 1 - 2y_k - 2^{-L}$, then, averaged over all possible (assumed equally likely) initial values y_0 (or w_0):*

1. *The mean value μ of the sequence w_k*

$$\mu = -2^{-n}\left(\frac{1 - 2^{-L}}{1 - 2^{-n}}\right) \approx 0$$

and variance σ^2

$$\sigma^2 = \frac{1}{3} + 2^{-n}\left[\frac{1}{3}\left(\frac{1-2^{-2L}}{1-2^{-n}}\right) - \frac{(1-2^{-L})^2}{1-2^{-n}} - 2^{-n}\left(\frac{1-2^{-L}}{1-2^{-n}}\right)^2\right]$$

$$\approx \frac{1}{3}.$$

2. *The sample autocorrelation function, defined by*

$$\hat{R}(m) = \frac{1}{N}\sum_{k=1}^{N} w_k w_{k+m},$$

has as its mean value $R(m)$, *given by*

$$R(m) = -2^{-n}\left(\frac{1-2^{-L}}{1-2^{-n}}\right)$$
$$\approx 0$$

for nonzero integral values of $|m|$ *less than* $(p-L)/q$. *The variance of* $\hat{R}(m)$ *about* $R(m)$ *is bounded by*

$$\operatorname{var}[\hat{R}(m)] < \frac{1}{N}\left[1 + \frac{1}{(2^n-1)}\right] \approx \frac{1}{N}.$$

3. *The relative number of times* $\hat{T}$ *that* y_k *falls in the interval for which the first* d *positions of the binary expansion are fixed, i.e., a neighborhood of length* 2^{-d} *in the interval* $(0, 1)$, *has mean*

$$T = E[\hat{T}] = 2^{-d}\left[1 + \frac{1}{(2^n-1)}\right] + \frac{1}{2}[g(0)-1]\left(\frac{1}{2^n-1}\right)$$
$$\approx 2^{-d}$$

for any number N *of points* y_k. *The variance of* $\hat{T}$ *is bounded by*

$$\operatorname{var}[\hat{T}] < \frac{1}{4}\left[1 + \frac{1}{(2^n-1)}\right]\left[\frac{1}{N} + \frac{2}{(2^n-1)}\right] \approx \frac{1}{4N}.$$

4. *The relative number of times* $\hat{T}$ *that* $(y_k, y_{k-l_2}, \cdots, y_{k-l_M})$ *falls in the interval of the unit* M*-cube for which the first* d_i *positions of the binary expansion of* y_{k+l_i} *are fixed, i.e., in a* $2^{-d_1} \times 2^{-d_2} \times \cdots \times 2^{-d_M}$ *interval in the unit* M*-cube, has mean value*

$$T = E(\hat{T}) = 2^{-(d_1+\cdots+d_M)}\left(1 + \frac{1}{2^n-1}\right) + 2^{-n-1}\left(\frac{g(0)-1}{1-2^{-n}}\right)$$
$$\approx 2^{-(d_1+d_2+\cdots+d_M)}$$

for any number N *of points* $(y_k, y_{k-l_2}, \cdots, y_{k-l_M})$, *provided* $0 < l_2 < \cdots < l_M < n/q - 1$. *The variance of* $\hat{T}$ *is then bounded by*

$$\operatorname{var}[\hat{T}] < \frac{1}{4}\left[\frac{1}{N} + \frac{2}{2^n-1}\right]\left[1 + \frac{1}{2^n+1}\right] \approx \frac{1}{4N}.$$

9. Primitive Polynomials. In order to implement the generator, it is necessary to find a primitive polynomial $f(x)$ over GF(2). A complete tabulation up through degree 34 appears in Peterson [6]. The form easiest to implement is usually one in which the recursion relation has fewest terms. Golomb et al. [7] have found primitive trinomials for most degrees through degree 36.

Watson [8] has published a table giving one primitive polynomial for each degree up to 100. A degree 35 polynomial $f(x) = x^{35} + x^2 + 1$ is very useful for generating numbers on an IBM-7094, whose numerical register contains 35 digits. In this case the period $p = 2^{35} - 1$ is relatively prime to 35, so q may be set equal to 35 for maximal precision ($L = n$) numbers. Preliminary experimental results indicate that the bounds given here are indeed valid for arbitrary sample sequences y_k.

Additional tests have shown that with $L = q = 17$, the pair (y_k, y_{k+1}) is uniform on the unit square.

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

1. S. W. GOLOMB, *Sequences with Randomness Properties*, Martin Co., Baltimore, Md., 1955.

2. NEAL ZIERLER, "Linear recurring sequences," *J. Soc. Indust. Appl. Math.*, v. 7, 1959, pp. 31–48. MR **21** #781.

3. L. BAUMERT, ET AL., *Coding theory and its Applications to Communications Systems*, Report 32-167, Jet Propulsion Laboratory, Pasadena, Calif., 1961.

4. R. C. TITSWORTH & L. R. WELCH, *Modulation by Random and Pseudo-Random Sequences*, Report 20-387, Jet Propulsion Laboratory, Pasadena, Calif., 1959.

5. A. ZYGMUND, *Trigonometrical Series*, Monogr. Mat., Bd. 5, Warsaw, 1935; reprint, Dover, New York, 1955; 2nd ed., Chelsea, New York, 1952; Russian transl., Moscow, 1939. MR **17**, 361; MR **17**, 844.

6. W. W. PETERSON, *Error-Correcting Codes*, M.I.T. Press, Cambridge and Wiley, New York, 1961, pp. 251–270. MR **22** #12003.

7. S. W. GOLOMB, L. R. WELCH & A. HALES, *On the Factorization of Trinomials Over* GF(2)," Report 20-189, Jet Propulsion Laboratory, Pasadena, Calif., 1959.

8. E. J. WATSON, "Primitive polynomials (mod 2)," *Math. Comp.* v. 16, 1962, pp. 368–369. MR **26** #5764.