Bull. Korean Math. Soc. **35** (1998), No. 1, pp. 173–187

APPLICATIONS OF ERGODIC THEORY TO PSEUDORANDOM NUMBERS

GEON HO CHOE, CHIHURN KIM AND DONG HAN KIM

ABSTRACT. Several aspects of pseudorandom number generators are investigated from the viewpoint of ergodic theory. An algorithm of generating pseudorandom numbers is proposed and shown to behave reasonably well.

1. Introduction to ergodic theory

The theory of dynamical systems studies the long term statistical behavior of the transformation T under iterations. The theory of measurable dynamical systems defined on a probability space is called the *ergodic theory*. Invariants are used to distinguish non-isomorphic dynamical systems in a given category of dynamical systems. For example, (measure theoretic) entropy completely classifies Bernoulli shifts. In other categories of dynamical systems we do not have enough knowledge of the complete invariants for the classification. In this paper we investigate the possible applications of ergodic theory to the problems related with pseudorandom number generators. Some of them will be regarded as discretized versions of measure preserving transformations on probability spaces.

Let (X, μ) be a probability measure space. A map $T : X \to X$ is said to be μ-preserving or μ-invariant if $\mu(T^{-1}(E)) = \mu(E)$ for every measurable subset $E \subset X$. The condition is equivalent to $\int_X f(x)d\mu = \int_X f(T(x))\, d\mu$ for every integrable function f. If μ is given by $d\mu = \rho\, dx$

Received May 31, 1997.

1991 Mathematics Subject Classification: Primary 65C10; Secondary 28D20, 11K45.

Key words and phrases: Ergodicity, random numbers, entropy.

Research supported by the Ministry of Information and Communication and the Center for Applied Mathematics.

for an integrable function $\rho(x) \geq 0$ with $\int_X \rho\, dx = 1$, then the measure is said to be absolutely continuous and ρ is called the density. Suppose T is μ-invariant. It is called *ergodic* if E satisfies $T^{-1}(E) = E$ modulo measure zero sets if and only if $\mu(E) = 0$ or 1. Equivalently, if $f(Tx) = f(x)$ for almost every x then f is constant almost everywhere. For general references on ergodic theory, see [Pe],[Wa],[CFS]. The following are examples of measure preserving transformations.

A trivial example of an ergodic invariant map is given by a translation $Tx = \{x + \alpha\}$ for some irrational α where $X = [0, 1)$. Another trivial example is $T(x) = \{2x\}$ on $X = [0, 1)$ where $\{t\}$ is the fractional part of a real number t. The Lebesgue measure is the invariant measure in both cases.

A nontrivial example is given by the logistic map $Tx = 4x(1 - x)$ on $X = [0, 1]$ with its invariant measure $\frac{1}{\pi\sqrt{x(1-x)}}dx$. It was first discovered by von Neumann and Ulam. It is known that the same measure is invariant under transformations obtained from Chebyshev polynomials.

An example with infinitely many discontinuities is given by the Gauss map $Tx = \{\frac{1}{x}\}$ with its invariant measure $\frac{1}{\ln 2}\frac{dx}{1+x}$. These are most of the known examples of absolutely continuous measures. It is not an easy task to find an explicit formula for invariant measure corresponding to a given map T.

The first fundamental fact in ergodic theory is the following.

Fact 1.1. (The Birkhoff Ergodic Theorem) If T is μ-invariant and f is integrable, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = f^*(x)$$

for some $f^* \in L^1(X, \mu)$ with $f^*(Tx) = f^*(x)$ for almost every x. Hence if T is ergodic, then f^* is constant and equal to $\int_X f d\mu$.

Let $1_E(x)$ denote the characteristic function of a measurable subset $E \subset X$. Choose $f(x) = 1_E(x)$ in the Birkhoff Ergodic Theorem. Then we see that if T is ergodic then the average number of times that the points $T^k x$ visit E is equal to the size of the subset E. In other words, ergodicity implies the uniform distribution.

One might object to using the Birkhoff ergodic theorem in estimating an integral, but this may be regarded as the idea behind the Monte Carlo

simulation method since the error between $\frac{1}{n}\sum_{k=0}^{n-1} f(T^k x)$ and $\int_X f d\mu$ can be estimated by standard statistical methods. For a reference for Monte Carlo simulation see [Fi].

Now we consider the transformations defined on the unit interval $X = [0,1]$. If an invariant measure μ is of the form $d\mu = \rho\, dx$ for some differentiable map T, then

$$\rho(x) = \sum_{y \in T^{-1}(\{x\})} \frac{\rho(y)}{|T'(y)|}.$$

For piecewise differentiable maps on the unit interval the existence of absolutely continuous invariant measures is proved under various similar conditions. For the existence of absolutely continuous ergodic invariant measures on the unit interval, see [AF], [Bow], [MPV], [Re], [Ro], [Si].

Here is a brief introduction to the definition of randomness that was first formulated by C. Shannon[Sh] in 1948. Mathematical formulation of randomness in an experiment may be given as follows: An experiment ξ with possible outcomes a_i with probability p_i, $i = 1, 2, ..., N$, corresponds to a partition of a probability space consisting of finite elements a_i's with probability measure μ defined by $\mu(\{a_i\}) = p_i$. In general we can define entropy for a partition of a probability measure space as follows: Let $\xi = \{E_1, ..., E_N\}$ be a finite partition of a probability space $(X, \mathcal{A}, \mu)$ with $E_i \in \mathcal{A}$ for every i. Then we define the entropy of the partition by

$$H(\xi) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$$

where $p_i = \mu(E_i)$ where the base of the logarithm equals 2 throughout the article.

Given two partitions ξ_1 and ξ_2 we let $\xi_1 \vee \xi_2$ be the partition consisting of the subsets of the form $B \cap C$ where $B \in \xi_1$ and $C \in \xi_2$. Suppose $T : X \to X$ is a measure preserving transformation. For a finite partition ξ we define the entropy of the transformation with respect to ξ by

$$h(T, \xi) = \lim_{n \to \infty} \frac{1}{n} H(\xi \vee T^{-1}\xi \vee \cdots \vee T^{-(n-1)}\xi).$$

It is known that the limit exists. Finally we define the entropy $h(T)$ of the transformation T by

$$h(T) = \sup_{\xi} h(T, \xi)$$

where ξ ranges over all finite partitions ξ.

A set ξ of possible outcomes or source symbols is called an *alphabet*. Let $X = \prod_{i=1}^{\infty} \xi$ be the product of the alphabet $\xi = \{a_1, ..., a_N\}$. We consider the shift transformation $T : X \to X$ given by $(Tx)_n = x_{n+1}$ for every n. Suppose there is a T-invariant measure μ on X. For example, μ is the product measure of a measure ν on ξ defined by $\nu(\{a_n\}) = p_n$. In this case T is called a (one-sided) Bernoulli transformation. It is known to be ergodic and $h(T) = -\sum_i p_i \log p_i$.

Suppose that we are given a differentiable map T on the unit interval. For $|x - y| \approx 0$ we have

$$|T(x) - T(y)| \approx |T'(x)| \cdot |x - y|,$$

and

$$|T^n(x) - T^n(y)| \approx \prod_{i=0}^{n-1} |T'(T^i x)| \cdot |x - y|$$

hence

$$\frac{1}{n} \log |T^n(x) - T^n(y)| \approx \frac{1}{n} \sum_{i=0}^{n-1} \log |T'(T^i x)|.$$

If we let μ be the ergodic invariant measure for T then the right hand side converges to $\int_0^1 \log |T'(y)| d\mu(y)$ by the Birkhoff Ergodic Theorem. Therefore we see that $|T'(x)|$ measures the extent with which two neighboring points diverge and $\log |T'(x)|$ measures the exponent of the speed of the divergence, which is called the Lyapunov exponent. Oseledec's theorem states that the Lyapunov exponent is the entropy.

2. Shannon-MacMillan-Breiman Theorem

There are many tests for randomness and entropy test is one of them. For a recent extensive experimental result see [LCC]. In ergodic theory the entropy of a map is defined as the average rate of randomness in the complexity of the behavior of the map with respect to the given invariant probability measure. Consider the shift transformation $T : X \to X$, $(Tx)_i = x_{i+1}$ where $X = \prod_1^{\infty} \{0, 1\}$. A typical point in X is a randomly generated sequence of 0 and 1. The traditional entropy test for randomness is not sharp due to the fact that entropy measures the

randomness of the sequence on average. To overcome such a difficulty an equivalent pointwise formulation of entropy is used.

Suppose there is a shift invariant probability measure μ on X. A block b is a finite string of symbols and $|b|$ denotes its length, i.e., $|b| = n$ if b is of the form $b = b_1 b_2 \dots b_n$. For each n there are 2^n blocks of length n, some of which may have measure zero. Take a block b of length n. Put

$$P_b = \text{Prob}(x_s = b_1, \dots, x_{s+n-1} = b_n \text{ for some } s).$$

It does not depend on s since the probability is shift-invariant. It is the probability of observing the block b in a typical sequence generated by the given PRNG and is equal to the measure of the cylinder set $\{x \in X : x_1 = b_1, \dots, x_n = b_n\}$ since the measure is shift invariant. If we assume that the shift transformation gives an ergodic process, then the Birkhoff Ergodic Theorem implies that P_b is nothing but the relative frequency of observing b among all possible observations of blocks of length n in X. The entropy h_μ with respect to μ is defined by

$$h_\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{|b|=n} P_b \log \frac{1}{P_b}.$$

By definition the entropy is a global quantity to measure the randomness of the sequence from a random number generator. The local deviations from the full randomness might be invisible when averages are taken and the small discrepancy in the behavior of a generator cannot be detected. Therefore an equivalent pointwise formulation of entropy is needed. Let $P_n(x)$ be the relative frequency of the first n-block $x_1 x_2 \dots x_n$ in the sequence $x = x_1 x_2 x_3 \dots$, in other words,

$$P_n(x) = \lim_{K\to\infty} \frac{1}{K} \#\{0 \le t < K : x_{1+t} = x_1, \dots, x_{n+t} = x_n\}$$

Note that $P_n(x) = P_b$ where b is the first n-block in x, i.e., $b = x_1 \dots x_n$, hence $P_n(x)$ depends only on the first n digits of x. The Birkhoff Ergodic Theorem guarantees the existence of the limit. Note that if x and y satisfy the conditions $x_1 = y_1, \dots, x_n = y_n$, then $P_n(x) = P_n(y)$. Now we have

Fact 2.1. (Shannon-MacMillan-Breiman Theorem) For almost every x, i.e., with probability 1,

$$h_\mu = \lim_{n\to\infty} \frac{1}{n} \log \frac{1}{P_n(x)}.$$

For the proof, see [Pe].

In short, for sufficiently large n the random variable

$$Y_n(x) = \frac{1}{n} \log \frac{1}{P_n(x)}$$

measures the randomness of the sequence x, hence we choose Y_n as a test statistic. For example, the example in [KW] can be shown to have small entropy hence it is not a good generator.

For the fully random sequences the pointwise values $Y_n(x)$ is constant and equal to $\log 2$, therefore the variance is zero and it is hard to gauge the relative deviation from the theoretical predictions for $E(Y_n)$. To overcome such a difficulty we use a modified sequence of entropy less than $\log 2$, which is generated from a $(p, 1-p)$-Bernoulli process constructed out of the given uniform PRNG that is supposed to behave truly random.

Suppose we have x obtained from the $(p, 1-p)$-Bernoulli process, $0 < p < 1$. For notational simplicity, put $p_0 = p$, $p_1 = 1 - p$. Then for $x = x_1 \dots x_n \dots$ we have

$$P_n(x) = p_{x_1} \cdots p_{x_n}.$$

Then

$$Y_n(x) = -\frac{1}{n} \log(p_{x_1} \cdots p_{x_n}) = -\frac{1}{n} \sum_{i=1}^{n} \log p_{x_i}$$

hence

$$\begin{aligned} E(Y_n) &= -\frac{1}{n} \sum_{i=1}^{n} E(\log p_{x_i}) \\ &= -E(\log p_{x_1}) \\ &= -p \log p - (1-p) \log(1-p) \\ &= -p \log p - q \log q \end{aligned} \tag{2.1}$$

where $q = 1 - p$, and

(2.2)

$$
\begin{aligned}
n^2 \cdot E(Y_n^2) &= E\left(\left(-\sum_{i=1}^{n} \log p_{x_i}\right)^2\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} E(\log p_{x_i} \log p_{x_j}) \\
&= \sum_{i \neq j} E(\log p_{x_i}) E(\log p_{x_j}) + \sum_{i=1}^{n} E\left((\log p_{x_i})^2\right) \\
&= \left(\sum_{i} E(\log p_{x_i})\right)^2 - \sum_{i} \left(E(\log p_{x_i})\right)^2 + \sum_{i} E\left((\log p_{x_i})^2\right) \\
&= n^2 (E(\log p_{x_1}))^2 - n(E(\log p_{x_1}))^2 + nE\left((\log p_{x_1})^2\right).
\end{aligned}
$$

Therefore

(2.3)

$$
\begin{aligned}
Var(Y_n) &= E(Y_n^2) - E(Y_n)^2 \\
&= \frac{1}{n}\left[-(E(\log p_{x_1}))^2 + E((\log p_{x_1})^2)\right] \\
&= \frac{1}{n}\left[-(p \log p + q \log q)^2 + p(\log p)^2 + q(\log q)^2\right]
\end{aligned}
$$

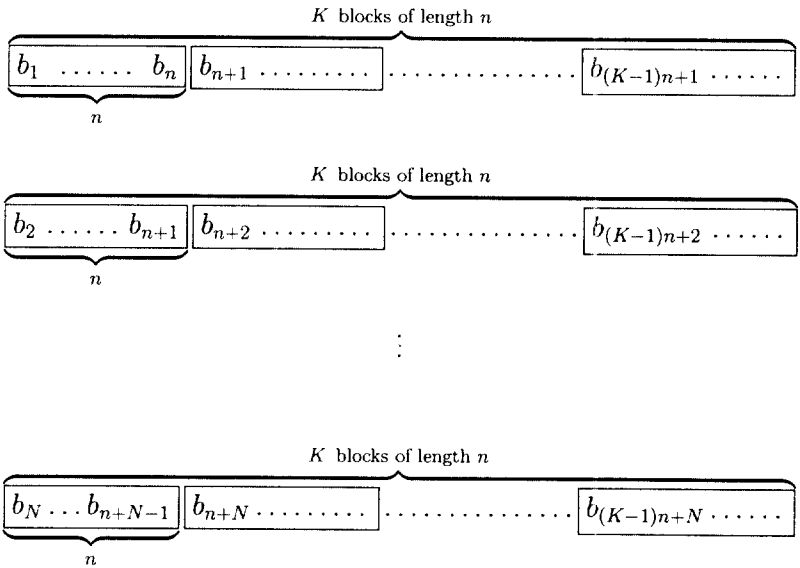
To use the result as a test for PRNG's, first we compute the expectation $E(Y_n)$ and the variance $Var(Y_n)$ theoretically. Next, we estimate $E(Y_n)$ experimentally. Let T be the shift transformation defined by $(Tx)_j = x_{j+1}$. According to the Birkhoff Ergodic Theorem the average

$$\frac{1}{N}\sum_{i=0}^{N-1} Y_n(T^i x),$$

converges to $E(Y_n)$ as $N \to \infty$, hence it is expected to approximate $E(Y_n)$ for sufficiently large N. Their theoretical distribution should be close to the normal distribution with its mean and variance given as above. Therefore, to test a PRNG, we first use it to construct a stochastic process and have it generate a typical sequence x. Then consider the random variable Y_n for some reasonably large n, for example, $n = 10$,

and estimate experimentally $Y_n(x), Y_n(Tx), \ldots, Y_n(T^{N-1}x)$ as accurately as possible for very large N. And finally compare it with the theoretical prediction for the model under consideration.

Since the random variable Y_n centers around the entropy value, we take the sample of a large number (say, $K = 10^5$) of nonoverlapping blocks of equal length n (the blocks in each row of the diagram given below). More precisely, to estimate $P_n(b_1 \cdots b_n)$ we use the first row, i.e., the n-blocks given by $[b_1 \cdots b_n]$, $[b_{n+1} \cdots b_{2n}]$, $\ldots$, $[b_{(K-1)n+1} \cdots b_{Kn}]$ and count the number of times that the block $[b_1 \cdots b_n]$ appears and compute the relative frequency. Similarly, we start with the n-block $[b_2 \cdots b_{n+1}]$ on the second row and obtain $P_n(b_2 \cdots b_{n+1})$ using $[b_2 \cdots b_{n+1}]$, $[b_{n+2} \cdots b_{2n+1}]$, $\ldots$, $[b_{(K-1)n-2} \cdots b_{Kn+1}]$, and so on. Thus N values of Y_n from N rows are obtained. These values are correlated, so to reduce the correlation we may use only some of the rows, for example, use nk-th rows for $k = 1, 2, ..., [\frac{N}{n}]$. From the computer experiment $[\frac{N}{n}]$ values of P_n's obtained from those rows seem to have negligible correlation coefficient.



The pointwise entropy test introduced in the above has a drawback. There is an unavoidable error in estimating $P_n(x)$ for any finite sample size K. Hence we need to take sufficiently large N to minimize the error.

In practice, most pseudorandom number generators are very good so we need exact estimates. To achieve that goal we may modify the definition of P_n. Let $P_n(b)$ denote the relative frequency of the arrearances of the n-block b in the finite string $b_1 \dots b_{n+N-1}$. Then the expectation of Y_n is given by

$$E(Y_n) = \frac{1}{n} \sum_{j=0}^{K-1} \log \left(\frac{1+j}{K} \right) \binom{K-1}{j} (\frac{1}{2^n})^j (1 - \frac{1}{2^n})^{K-1-j},$$

and

$$E(Y_n^2) = \frac{1}{n} \sum_{j=0}^{K-1} \left(\log \left(\frac{1+j}{K} \right) \right)^2 \binom{K-1}{j} (\frac{1}{2^n})^j (1 - \frac{1}{2^n})^{K-1-j},$$

3. PRNG's from ergodic theory

Let $T : X \to X$ be a transformation on a probability space X. By partitioning X into

$$X = \bigcup_{i=0}^{k-1} E_i$$

we can associate to every point $x \in X$ an infinite sequence of symbols from the alphabet $\{0, 1, \dots, k-1\}$ by the relation

$$x \mapsto (a_0, a_1, a_2, \dots, a_n, \dots)$$

where $a_n \in \{0, 1, \dots, k-1\}$ and $T^n x \in E_{a_n}$. If the given partition $\{E_0, \dots, E_{k-1}\}$ generates the σ-algebra, then the randomness of T is inherited by the shift map on the sequence space and this is how to obtain a pseudorandom number generator from a dynamical system. For example, if $X = [0, 1)$ and $Tx = 2x \pmod 2$ and if we choose the partition $E_1 = [0, \frac{1}{2})$, $E_2 = [\frac{1}{2}, 1)$, then the above method gives a binary expansion of a real number $0 \leq x < 1$. The transformation T is measure preserving and ergodic with its entropy $\log 2$ and the resulting shift map, which is nothing but the fair coin tossing, displays random behavior. If we had chosen other partitions, the extent of the randomness is less than optimal since by definition the entropy of T is the optimal upper bound for any entropy with respect to a given partition. In the above example, the partition is optimal.

EXAMPLE 3.1. (Linear congruential generators) The linear congruential generators defined by $x_n \equiv ax_{n-1} + b \pmod{m}$ are modeled after the ergodic transformations $x \mapsto ax \pmod 1$ on the unit interval, which is invariant with respect to the Lebesgue measure. Here the partition of the unit interval is given by m subintervals $[0, \frac{1}{m})$, $[\frac{1}{m}, \frac{2}{m})$, $\dots$, $[\frac{m-1}{m}, 1)$. In many practical implementations, $m = 2^{31}$ or $m = 2^{32}$.

EXAMPLE 3.2. (Toral automorphisms and lagged Fibonacci generators) Consider the transformations defined on n-dimensional torus given by an (invertible) integral matrix A with $\det A = \pm 1$. In practice most of the elements of A are 0 so that the iteration of multiplication of A is fast. Let $X = \mathbb{R}^n/\mathbb{Z}^n$ be the n-torus, which is identified with the unit cube in $\mathbb{R}^n$. For every integral matrix A we define the transformation T_A on X by $T_A x = Ax$. Since $|\det A| = 1$, the transformation T_A preserves the Lebesgue measure on the torus. It is known that T_A is ergodic if and only if any eigenvalue of A is not a root of the unity. It is known that the entropy of T_A equals the sum of $\log \lambda$ with $|\lambda| > 1$, where λ is an eigenvalue of A.

Suppose a matrix $A \in GL(2, \mathbb{Z})$ has two eigenvalues λ_1 and λ_2 with $|\lambda_1 \lambda_2| = 1$. This transformation is called *hyperbolic* if $|\lambda_1| > 1$ and $|\lambda_2| < 1$. The hyperbolic toral automorphism has two distinct eigenvectors. Under the action of A distances in the plane expand to one direction by a factor of $|\lambda_1|$ and contract to the other direction by $|\lambda_2|$.

For the matrix $A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ we can find a partition of the torus that gives the maximal value for the entropy. Let $\mathcal{P} = \{P_1, P_2, \dots, P_k\}$ be a partition of the torus for A, for example, k vertical strips of equal width for some large k, say $k = 2^{32}$. Each $x \in \mathbb{T}^2$ defines a sequence as explained previously. As long as x never lands in the boundary of $\mathcal{P}$ this produces a single well-defined sequence. No two points in $\mathbb{T}^2$ correspond to the same sequence.

In general, we use integral matrices with $|\det A| = 1$ so to use the invariance of the Lebesgue measure on the torus, which in turn implies the uniform distribution, and we choose A with large $|\lambda_1|$. The coded/discretized version is nothing but a Fibonacci generator $x_{n+2} = x_{n+1} + x_n \pmod m$. We can generalize the idea to the n-dimensional case. For the related results see [MZ, Ma].

EXAMPLE 3.3. (Inversive congruential generators) From the viewpoint of the ergodic theory we try to explain why purely algebraic algorithms such as inversive congruential generators $ICG(p, a, b, x_0)$ defined by $x_n \equiv a{x_{n-1}}^{-1}+b \pmod p$ display random behaviors. In the finite commutative group $\mathbb{Z}_p \setminus \{0\}$ we have the identity $x^{p-1} = 1$, hence $x^{-1} = x^{p-2}$. Therefore a typical ICG is given by $x_n \equiv a x_{n-1}^{p-2} + b \pmod p$. This algorithm corresponds the interval map $Tx = ap^{p-2}x^{p-2} + b \pmod 1$. Since p is large, we expect that the distribution is approximately equal to the Lebesgue measure and that the entropy is large from the Lyapunov exponent formula so that the coded sequences display randomness with respect to any reasonable partition of the unit interval. Almost the same argument can be applied for the PRNG's obtained from polynomials with steep slopes.

4. A piecewise linear map and a new PRNG

Now we propose a new class of generators based on observations in ergodic theory. Consider the following piecewise linear map on the unit interval defined by

$$f(x) = -2^n(x - \frac{1}{2^{n-1}}) = -2^n x + 2 \quad \text{for} \quad \frac{1}{2^n} < x < \frac{1}{2^{n-1}}.$$

It is obtained by modifying the map $L(x) = \log \frac{1}{x} \pmod 1$, $0 \le x < 1$, hence we call it the *linearized logarithmic map*.

If μ is the Lebesgue measure, then for an open interval $I = (a, b)$,

$$\begin{aligned} \mu(f^{-1}(I)) &= \mu(f^{-1}(a, b)) \\ &= \mu(\bigcup_{n=1}^{\infty}(-\frac{b}{2^n} + \frac{1}{2^{n-1}}, -\frac{a}{2^n} + \frac{1}{2^{n-1}})) \\ &= \sum_{n=1}^{\infty} \mu((-\frac{b}{2^n} + \frac{1}{2^{n-1}}, -\frac{a}{2^n} + \frac{1}{2^{n-1}})) \end{aligned}$$

Since the Lebesgue measure is translation invariant,

$$\mu(f^{-1}(I)) = \sum_{n=1}^{\infty} \mu((-\frac{b}{2^n} - \frac{a}{2^n})) = \sum_{n=1}^{\infty} -\frac{1}{2^n}(a - b) = \mu(I)$$

Generators	Seed x_0	Increment b	Period
L1	3838758767	1965377237	2103452906
L2	3552705551	1281215333	1625984253
L3	2631754467	904400631	1209855518
L4	443889273	1710812443	1483585494
L5	4044416047	411920269	1497237346

TABLE 1. Periods of LinLog's for different choices of x_0 and b

Hence the Lebesgue measure is its invariant measure. From the Lyapunov exponent formula its entropy equals $\sum_{n=1}^{\infty} \frac{n}{2^n} \log 2 = \frac{3}{2} \log 2$. To use f as a model of a pseudorandom number generator, we discretize it as follows: we will scale up its domain and simplify the map where the slope of the map is too steep. Define the piecewise linear map as follows:

$$f(x) = \begin{cases} -2^{31}x + s \pmod{2^{32}}, & 0 \le x < 2 \\ -2^{32-n}x + s \pmod{2^{32}}, & 2^n \le x < 2^{n+1}, \quad 1 \le n \le 31, \end{cases}$$

for some integer $0 \le s \le 2^{32} - 1$. We call it the linearized logarithmic generator(LinLog).

Its algorithm is faster than most of other generators due to the nature of computer hardware structure, i.e., LinLog shifts a binary number $x = (x_1, x_2, \ldots, x_{32})$, $x_i \in \{0, 1\}$, consisting of 32 bits to the left until the first nonzero bit is placed in the first slot.

To find generators with sufficiently long periods suitable choices of the increment b are needed. Some increments having long periods are given in Table ??. They were obtained by extensive computer experiments.

To test the uniform distribution of the pseudorandom numbers generated by PRNG's we apply the chi-square test from statistics. The following tables (Table 2 – Table 4) show the results of applying four different types of chi-square tests on each of sequences generated by LinLog's and by some standard LCG's. Each generator has been tested with block lengths $n = 4, 5, 6, 7, 8$. The number of observations in the tests A, B, C and D are $2^{4+n}, 2^{9-n}, 2^{14+n}$ and 2^{19+n} respectively, where n is the block size. In the chi-square test only the first bit among 32 bits was tested. The symbol $\times$ indicates that the statistic belongs to 0 – 1 % or 99 – 100 % area, the symbol $\triangle$ indicates that the statistic belongs to 1

– 5 % or 95 – 99 % area, and the blank spaces imply that the generator passes test.

Now we apply the pointwise entropy test that was introduced in Section 2. We choose $N = 10^5$, $K = 10^5$ for each n-block. In L3 only the first bit was tested and the whole bit was tested in the others. Theoretical values of expectations and standard deviations for some n are given in Table 4.

Generators	ANSI C					MS C					RANDU				
Block size	4	5	6	7	8	4	5	6	7	8	4	5	6	7	8
A	△		△	△		△	△	△	△			△			
B					△			×					×		
C	×		△				△	×	×	△			×		
D								×	×			×	×	×	×

TABLE 2. Chi-square test of LCG's

Generators	L1					L2					L3				
Block size	4	5	6	7	8	4	5	6	7	8	4	5	6	7	8
A									×	×					×
B	△	△	△			×	△				×				
C			×	×			×	×							
D		×		×				△	×	×	×				

TABEL 3. Chi-square test of LinLog

Generators	L4					L5				
Block size	4	5	6	7	8	4	5	6	7	8
A	△		△	×	×				△	△
B	×	×				△	△		×	
C			×	△	△		×	△		
D			△				×	×	×	×

TABLE 4. Chi-square test of LinLog

Block length n	Expectation of Y_n	Standard deviation
8	0.99976997	0.00910075
9	0.99959008	0.01144418
10	0.99926079	0.01455424
11	0.99865296	0.01866667
12	0.99752084	0.02406902
13	0.99538672	0.03104823

TABLE 5. Theoretical predictions of Y_n for $K = 10^5$

Block length n	Estimation of $E(Y_n)$			
8	0.99978845	0.99983396	0.99977429	0.99983305
9	0.99963740	0.99965458	0.99963268	0.99964900
10	0.99916076	0.99939345	0.99960046	0.99918736
11	0.99859457	0.99877539	0.99868825	0.99875253
12	0.99741595	0.99758465	0.99759730	0.99747080
13	0.99533120	0.99548402	0.99523713	0.99537788
Generators	ANSI C	MS C	RANDU	L3

TABLE 6. Experimental estimation of $E(Y_n)$, $N = 10^5$, $K = 10^5$

To find a very good generator needed in today's serious computer simulations more extensive experimental work for the coice of seed and increment will be required.

References

[AF] R. L. Adler and L. Flatto *Geodesic flows, interval maps, and symbolic dynamics*, Bull. Amer. Math. Soc., **25** (1991), 229–334.

[Bow] R. Bowen, *Invariant measures for Markov maps of the interval*, Comm. Math. Phys., **69** (1979), 1–17.

[CT] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, 1991.

[Fi] George S. Fishman, *Monte Carlo: concepts, algorithm, and applications*, Springer Verlag, New York 1996.

[CFS] I. P. Cornfeld, S. V. Fomin, Ya. G. Sinai, *Ergodic Theory* SpringerVerlag, New York, 1982.

[KW] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods, vol 1, 2nd ed.*, Addison-Wesley, 1980.

[Kn] D. Knuth, *The Art of Computer Programming, Vol 2, 2nd ed.*, Addison-Wesley, 1980.

[LY] A. Lasota and J. Yorke, *On the existence of invariant measures for piecewise monotone transformations*, Trans. Amer. Math. Soc., **186** (1973), 481–488.

[LCC] P. L'Ecuyer, A. Compagner and J-F. Cordeau, *Entropy tests for random number generators*, preprint (1996).

[LM] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics*, Cambridge Univ. Press, 1995.

[Ma] G. Marsaglia, *The mathematics of random number generators*, Proc. Symp. App. Math., **46** (1992), 73–90.

[MZ] G. Marsaglia and A. Zaman, *A new class of random number generators*, Ann. App. Prob., **1** (1991), 462–480.

[MPV] J. Moser, E. Phillips and S. Varadhan, *Ergodic Theory, a seminar*, New York Univ., New York, 1975, 111–120.

[Ni] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Vermont, 1992.

[Pe] K. Petersen, *Ergodic Theory* , Cambridge Univ. Press, New York, 1983.

[Re] A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Akad. Sci. Hungary **8** 1957, 477–493.

[Ro] V. A. Rohklin, *Exact endomorphisms of a Lebesgue space*, Amer. Math. Soc. Transl., Series 1, **39** (1964), 1–37.

[Sh] C. Shannon, *The mathematical theory of communication*, Bell Sys. Tech. **J.**, **27** (1948) 379–423 and 623–656.

[Si] Ya. Sinai, *Topics in Ergodic Theory*, Princeton Univ. Press, Princeton, New Jersey, 1994

[Wa] P. Walters, *An Introduction to Ergodic Theory, 2nd ed.*, Springer-Verlag, New York, 1981.

Department of Mathematics, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea
E-mail: choe@euclid.kaist.ac.kr
E-mail: chihurn@euclid.kaist.ac.kr
E-mail: kdh@euclid.kaist.ac.kr