GENERATION OF RANDOM NUMBERS BY DETERMINISTIC PROCESSES

ROLAND HUBERT EDDY

14503

GENERATION OF RANDOM NUMBERS BY DETERMINISTIC PROCESSES

BY

© ROLAND HUBERT EDDY

A THESIS

SUBMITTED TO THE COMMITTEE ON GRADUATE STUDIES

MEMORIAL UNIVERSITY OF NEWFOUNDLAND

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF MATHEMATICS

MEMORIAL UNIVERSITY OF NEWFOUNDLAND

ST. JOHN'S, NEWFOUNDLAND

AUGUST, 1971

TABLE OF CONTENTS

Abstract

Although the general principles of Monte Carlo and other simulation techniques have been known since the turn of the century, lack of efficient computational facilities has restricted their general application. The rapid advances in the field of electronic computing during the last three decades, however, have produced a new awareness of the potential of such techniques and as computing becomes even more sophisticated, such methods will no doubt play an increasingly important role in future scientific investigation.

Effective application of Monte Carlo methods requires access to long sequences of random numbers. Since perfectly random numbers can not, of course, be obtained by practical means, all sets produced to date must properly be termed "pseudo-random". It is generally accepted, in the published literature, that such sets will be more limited in their application than perfectly random sets would be. Even though considerable research has gone into producing sequences for general application, such sequences produced to date are not equally satisfactory for all purposes and must be considered in the light of the particular problem under investigation.

In this thesis, we consider the problem of finding sequences suitable for the Monte Carlo calculation of definite integrals. After a particular sequence is generated and tested for randomness, it is used in the evaluation of three definite integrals. The results of the statistical tests for each sequence are then compared with the values of the integrals produced by that sequence and an attempt is made to determine which properties a sequence should possess in order to produce good results in this application. Throughout the thesis, several small innovations are introduced which, we believe, have not been reported by other authors.

ACKNOWLEDGEMENTS

I wish to make the following acknowledgements:

(a) To my supervisor, Dr. M.K. Lewis, for guiding my research and giving invaluable assistance in the preparation of this thesis.

(b) To Dr. H.G. Benson of Computer Services Limited (Newfoundland and Labrador) for his assistance in obtaining the computed data.

(b) To Mrs. H. Tiller for typing the manuscript.

CHAPTER I

Introduction

Most generally, the principles of Monte Carlo techniques rely heavily upon properties of random variables and their associated sampling distributions. In order that the user may make effective application of such techniques, he must be able to observe values of a random variable, having a distribution relevant to the problem under investigation. This requirement has been facilitated by the construction of tables of "random numbers".

Specifically, we define a random number as an observed value of a random variable. Since, theoretically, all distributions may be derived from the uniform distribution by suitable functional relationships, we may, without loss of generality, redefine a random number as an observed value of a random variable uniformly distributed over the interval [0,1]. Of more practical value, however, is the following equivalent definition. A set of digits $S = \{s_1, s_2, s_3, \ldots, s_i, \ldots\}$ is called a "random sequence" if

(i) $s_i \in D$, where $D = \{0,1,2,\ldots,9\}$.

(ii) $P[s_i = \alpha] = \frac{1}{10}$, for all $\alpha \in D$; $i = 1,2,3,\ldots$

(iii) s_j is independent of s_i, for all $i < j$.

We now define a random number as a subset of the elements of S. Good [13] has given an equivalent definition in which he also includes the case of binary sequences, however, his definition is not truly accurate as he regards a random sequence as being finite.

Due to the limitations of the real world, it is not possible to produce even a single random number. However, it is possible by various classical

methods, such as throwing dice or picking numbers out of a hat, to generate small sets of numbers that resemble, in their properties, a sequence of random numbers.

In order to obtain reasonable accuracy using the Monte Carlo technique, one needs a large supply of random numbers. In an investigation involving the evaluation of definite integrals, Powell and Swann [36] used sets of size 10, 100, and 1,000. The best results were obtained when 1,000 numbers were used while a set of size 10 gave very poor results. Since classical methods of generation could not produce such large sequences efficiently, the Monte Carlo method did not enjoy widespread popularity until more sophisticated methods became available. In the last three decades, deterministic methods employing the facilities of high speed electronic computers have made possible the generation of long sequences of numbers.

Philosophical arguments have been raised against this procedure. Von Neumann [47], for example, has suggested that "anyone who uses arithmetical methods to produce random sequences is in a state of sin". We do not claim such sequences to be random in the sense defined above but, since many of them pass various statistical tests for randomness, they may, for practical purposes, be used in place of random sequences. Such sequences, possessing random-like properties and produced in a deterministic manner are more properly termed "pseudo-random" sequences. We shall, in this thesis, be concerned entirely with pseudo-random sequences but, for the sake of brevity, we shall use the term "random" throughout.

No sequence produced to date has possessed all the various properties of a random sequence to a high degree, in fact, it has been found that such sequences vary widely in the degree to which they possess random properties.

The suitability of any particular generated sequence, therefore, has to be considered in the light of the particular problem under investigation [17].

In this thesis, we attempt to determine which properties a generated sequence must necessarily possess in order to obtain reasonably good results in the Monte Carlo calculation of definite integrals. This we attempt to do by generating several sequences of numbers, testing them for randomness and then using them to evaluate three definite integrals. Conclusions are drawn from correlations worked out between the values of the integrals obtained and the results of the statistical tests applied.

In Chapter II, we survey some of the popular methods of generation, both past and present, and we describe how our own numbers were obtained. In addition, we present a mathematical treatment of a bias that exists in certain generated sequences.

In Chapter III, we discuss the six tests for randomness that were applied to each of the generated sequences.

In Chapter IV, we give the details of the Monte Carlo calculations of three definite integrals.

In Chapter V, all results obtained are presented and some conclusions are drawn.

CHAPTER II

The Generation of Random Variables

1. Tables and Physical Methods.

The idea of using random sampling methods in order to obtain a solution to a determinate mathematical problem must really be attributed to "student", who, in 1908, after theoretically predicting the "t" distribution and the distribution of the sample correlation coefficient, tested his results empirically using random sampling methods. Much later, Von Neumann and Ulam [31] presented what is now commonly known as the "Monte Carlo" method. They suggested that an approximate solution to certain mathematical problems could be obtained by random sampling from probabilistic analogues.

The only methods of obtaining random numbers in 1908, and for some time later, were by physical means such as drawing cards from a well-shuffled deck or by rolling dice. Not only were these methods extremely vulnerable to human bias, but it was practically impossible to get relatively large samples from them.

The first major break through in the field of large-scale sampling occured in 1927 when Tippett [46] who, after struggling with "ticket" sampling for some time, acted upon a suggestion from Karl Pearson that he replace his system of tickets by a single table of four digit numbers. Tippett formed his table of random numbers by taking 40,000 digits at random from census reports and combining them by fours to produce a set of 10,000 numbers. In order to completely fill 26 pages each containing 1600 digits, the published table actually contained 41,600 digits. Pearson [46] in his forward to the table, suggests further that even if certain applications demanded more numbers the table was still suitable as the numbers could be used over again. These digits were tested by Meyer, Gephart, and Rasmussen [32] who point out that the numbers should be used with caution especially

in small sampling surveys. Their results, however, were not completely accurate as the "serial test" used, which was based on the one proposed by Kendall and Babington Smith [21], was subsequently shown by Good [13] to be incorrect.

Kendall and Babington Smith [23] point out that repeated sampling from the same set as suggested earlier by Pearson [46] causes serious doubt to arise in regard to certain random properties. Consequently, they prepared a larger table consisting of 100,000 digits [23]. These digits were produced by a machine which consisted essentially of a disc divided into ten equal sections containing the digits 0 to 9 inclusive. The disc was rotated in a dark room by an electric motor and was illuminated from time to time by the flash of a neon lamp. At each flash, a fixed pointer selected a number from the apparently stationary disc. Four tests of randomness were applied to each block of 1,000 digits and of the hundred blocks tested, only five failed to pass all four tests.

As the Monte Carlo method grew in popularity, the demand for more random numbers increased. Consequently, in 1955, the RAND Corporation produced a table of one million digits [37]. The "randomising machine" used in this case consisted essentially of a random frequency pulse source and a five-place binary counter. Pulse standardization circuits passed the pulse through the counter and produced one number per second. Upon conversion to the decimal system the final digit only was retained and fed into an IBM punch. Standard tests of randomness were applied with satisfactory results.

Several other authors have reported successful experiments with other methods. In particular, we mention the use of radioactive noise by Isida and Ikeda [18], the arranging of roulette wheels in series by Horton [15],

and ERNIE (Electronic Random Number Generator Indication Equipment) by Thompson [45].

2. The Middle-Square and Related Methods.

The so-called "middle-square" method was proposed by John Von Neumann and N. Metropolis in 1946 and was the first attempt to produce randomlike numbers by a deterministic process. The following special case will serve to illustrate the method. Take a four-digit number x_o, e.g., $x_o = 5392$ and square it to obtain an eight-digit number, e.g., $x_o^2 = 29073664$. The middle four digits are then recorded as the first random number in the sequence, i.e., $x_1 = 0736$. Consequently $x_1^2 = 00541691$ and $x_2 = 5416$. Continuing the process we have $x_3 = 3330$, $x_4 = 0889$, etc.

This method is convenient for high speed calculations in a computer as it requires only one initial value and it has a very fast and short calculating procedure. A major weakness of the method is that all sequences ultimately degenerate into a cycle, usually small, or a sequence of zeros. In addition, it is very difficult to analyze important properties such as the length of the cycle.

An alternative method was proposed by Von Hoerner [19] in which a number is obtained by the middle-square method and recorded as the first random number in the sequence. A constant is then added to this number and the resulting number is squared. The middle of this number is extracted and recorded as the second random number in the sequence and so on. This method is an improvement over the previous one in that it never degenerates into a sequence of zeros, however, it will always reduce to a cycle, generally small, and is therefore unsuitable.

Forsythe [12], after unsuccessful experiments with the middle-square method,and a similar method in which he extracted the middle from the product of two numbers, reported success with a method which generated a sequence $\{a_n\}$ of eight-digit random numbers. If two eight-digit numbers a_1 and a_2 are chosen arbitrarily, the sequence $\{a_n\}$ may be defined recursively as follows. From the eight-digit number a_n a five-digit number a^1_n is extracted. The eight-digit number a_{n+2} is then extracted from the product $a^1_n \cdot a_{n+1}$. The actual run gave the values, $a^1_1 = 34567$, $a_2 = 98765432$ and $a^1_2 = 76543$. Since $a^1_1 \cdot a_2 = 3414024687944$, $a_3 = 40246879$. Also, since $a^1_2 \cdot a_3 = 3080616859297$, $a_4 = 06168592$, etc. The resulting sequence contained 12,500 numbers which had no cycle and did not degenerate. Furthermore, it fared reasonably well on certain tests for randomness. Unfortunately, the method had one serious defect in that the distribution of pairs of digits was not uniform.

More sophisticated deterministic methods, free of the above disadvantages, have since replaced the middle-square type methods.

3. Congruential Methods.

These methods, based on congruence relations, lend themselves readily to use on high speed computers in that they are both easy to program and have a fast operating procedure.

Some of the most widely used generators of this type are derived from a congruence relation of the form

(2.1) ..... $x_n \equiv \lambda x_{n-1} + c \pmod{M}$,

where (mod M) with the congruence sign " $\equiv$ " means that "$\lambda x_{n-1} + c$ is

divided by M and x_n is the remainder or residue". When λ, x_o, c, and M are chosen, the sequence $\{x_n\}$ is determined recursively by (2.1). The possible values of x_n will be a subset of the set $\{x_n | 0 \leq x \leq M - 1\}$.

It has been shown, using number theoretic arguments [14,35], that all such sequences will repeat after a finite number of iterations. Furthermore, the number of terms obtainable before cycling begins is dependent upon the particular choices of λ, x_o, and c. A particular application involves choosing a fixed value for M, usually large, and suitable values for λ, x_o, and c that will produce a sequence of maximum length. The actual cycle of random numbers, distributed over the interval [0,1], is the sequence $\{\frac{x_n}{M}\}$.

If $c = 0$, the generator is called "multiplicative", in which case the sequence will always repeat before M numbers are obtained. If $c \neq 0$, the generator is referred to as a "mixed" or "mixed congruential" generator with maximum possible period equal to M.

The choice of $M = 2^k$ as a modulus is particularly convenient when the computer has a binary base. The calculation of a number $N(\text{mod } 2^k)$ involves simply retaining the k least significant digits of the binary representation of N.

E.g. To calculate $27(\text{mod } 2^4)$, 27(base 10) is represented in the binary base, i.e.

(2.2) ..... 27(base 10) = 11011 (base 2).

If the four least significant binary digits are retained, we have

(2.3) ..... 1011 (base 2) = 11 (base 10) $\equiv$ 27 (mod 16).

(a) The Multiplicative Generator.

This generator was introduced by Lehmer in 1949 [43]. He used the values $\lambda = 23$ and $M = 10^8 + 1$ and performed the calculations on ENIAC (a computer employing the decimal system). The resulting sequence produced eight-decimal digit numbers with period 5,882,352. He found that this value of λ was the best possible for this value of M as no larger multiplier produced a longer period and no smaller multiplier produced a period more than half as long. The multiplicative generator has been discussed by many authors including Barnett [1], Bofinger and Bofinger [3], Certaine [4], Dowhham and Roberts [7], Greenberger [14]. Hull and Dobell [17], Taussky and Todd [43], Thompson [44], and Stockmal [41].

Greenberger [14], using techniques of number theory, showed that the maximum period of such a generated sequence[1] is of length 2^{k-2} and can be obtained only when λ is of the form

(2.4) .... $\lambda \equiv \pm 3 \pmod 8$

and the starting value of x_o is odd. Hull and Dobell [17] showed that the generated sequences failed certain tests for randomness when particular values of λ satisfying (2.4) were used. Values that proved especially bad were those of the form

(2.5) .... $\lambda = 2^p \pm a$

where $a < 40$ and $p > 30$.

Our first set of random numbers was obtained by using the multiplicative generator

(2.6) .... $x_n \equiv 107x_{n-1} \pmod{2^{15}}$,

---

[1] One with modulus $M = 2^k$.

where the starting value x_o was chosen to be 15. This procedure is capable of producing 2^{13}, (8,192), distinct numbers which will always be a subset of the set $\{x | 0 \leq x \leq 2^{15} - 1,\ x \text{ is an integer}\}$. Hence the largest value obtainable by this method is $2^{15} - 1 = 32,767$. The first 8,000 numbers of this set were generated, normalized, and divided, for practical purposes, into 8 equal blocks.

(b) The Mixed Congruential Generator.

This generator has an advantage over the multiplicative one in that when $M = 2^k$, the maximum period is of length 2^k. To obtain this period it is necessary that the multiplier λ be of the form

(2.7) .... $\lambda \equiv 1 \pmod 4$

and c be any odd integer. This procedure will then generate all the integers in the set $\{x | 0 \leq x \leq 2^k - 1\}$ in random order. In addition, the maximum period is independent of the choice of x_o.

The proofs of these properties, using number theory arguments, are given by Greenberger [14] in the same paper in which the discussion of the multiplicative generator is presented. Peach [35], by using equations instead of congruence relations has presented an elegant discussion based on high school mathematics only and is particularly suited to those who in his own words, "are not at home in number theory".

As a result of extensive experimentation with various values of λ satisfying (2.7), Hull and Dobell [17] concluded that for practical applications there are various values that are unsuitable. Those that proved especially bad were $\lambda \equiv 1 \pmod{2^{13}}$, $\lambda < 30$, and $\lambda = 2^{12} + 1$. A

generally suitable value, $\lambda = 2^7 + 1$, was proposed by Rotenberg [38] and substantiated by Hull and Dobell [17].

Very little attention has been given in the literature concerning suitable values for c. Hull and Dobell [17] suggested that any odd integer was suitable. Later on, we prove this statement to be false.

Stockmal's paper [41] describes the problem of determining, algebraically, x_i from both the multiplicative and the mixed generators. The mixed generator has been discussed in detail by Coveyou [5], Jansson [19], MacLaren and Marsaglia [30], and Rotenberg [38].

Our first experiment with this procedure involved the generation of 12,000 numbers, which were recorded as 12 equal blocks, using the generator

(2.8) ..... $x_n \equiv (2^8 + 1)\, x_{n-1} + 517 \pmod{2^{15}}$.

This sequence failed very poorly on our battery of tests, and, in particular, all blocks failed a particular test[1]. Since the value of $\lambda = 2^8 + 1$ was in line with that suggested by the literature, it was evident that the value of $c = 517$ was not suitable. After carrying out considerable research, the best value that we found was $c = 21$. . We intend to do further study on this problem in order to classify c more generally.

Our second set of random numbers was obtained by using the generator

(2.9) .... $x_n \equiv (2^8 + 1)x_{n-1} + 21 \pmod{2^{15}}$.

This procedure is capable of producing a set of 2^{15} (32,768) integers in the range 0 to $2^{15} - 1$.

[1]This was the "runs test" which is dealt with in detail in Chapter III. The reported experiments in which good multipliers were sought did not involve the use of this test. The validity of this test as an indication of randomness, however, has been well established [7].

The first 10,000 elements of this set were generated, normalized, and recorded in 10 equal blocks. In order to observe elements in the upper half of the set, we used techniques of number theory and a bias (to be discussed in detail in Section 4 of this chapter) and computed the $26,384^{th}$ element from the $10,000^{th}$ element. Setting $x_0 = x_{26,384}$, we generated 2,000 additional numbers which were recorded in two equal blocks.

(c). Combination of Two Congruential Generators.

MacLaren and Marsaglia [30] reported that random numbers generated by mixed congruential methods gave poor results in a number of Monte Carlo calculations, notably, those involving order statistics. As an improvement they proposed using two different generators and having one shuffle the sequence produced by the other. The actual generators used were

$$(2.10) \quad \ldots\ldots \quad x_n \equiv (2^{17} + 3)x_{n-1} \pmod{2^{35}}$$

and

$$(2.11) \quad \ldots\ldots \quad y_n \equiv (2^7 + 1)y_{n-1} + 1 \pmod{2^{35}}.$$

Taking the initial values $x_0 = 1$ and $y_0 = 0$, a table of 128 locations in the core of the computer was filled with the numbers $x_1, x_2, x_3, \ldots, x_{128}$ The k^{th} random number to be used, i.e. Z_k, was generated by using the first seven bits of y_k as an index to select a value x_i, $i = 1,2,3,\ldots,128$, from the table. The location of x_i was refilled with the next number from the generator (2.10). Even though the time required to generate a sequence by this method was double that required by a single generator, the authors accepted this penalty as the resulting sequence exhibited better statistical properties.

Westlake [48] attempted to eliminate the storage problem in the previous method by using a combination of two multiplicative congruential generators.

The actual generators used were

$$(2.12) \quad \ldots\ldots \quad x_n \equiv (2^{31} + 3)x_{n-1} \pmod{2^{59}}$$

and

$$(2.13) \quad \ldots\ldots \quad y_n \equiv (2^{29} + 3)y_{n-1} \pmod{2^{59}}.$$

The generated sequence was obtained by first generating the numbers x_i and y_i. The bits of y_i were then permuted and added to x_i. The resulting number was converted to decimal form, normalized, and recorded as the i^{th} random number of the sequence. The results of tests on the individual generators were compared with those on the combined generator. Properties of non-randomness that were found to exist in the individual generators were not evident in the combined one.

Our third set of random numbers was obtained by using a combination of generators (2.6) and (2.8), i.e.

$$(2.14) \quad \ldots\ldots \begin{cases} x_n \equiv 107x_{n-1} \pmod{2^{15}} \\ \text{and} \\ y_n \equiv (2^8 + 1)y_{n-1} + 517 \pmod{2^{15}}. \end{cases}$$

The same starting value $x_o = 15$ was used as in (2.6) to generate the number x_1 which was then recorded. This number was then fed into the second generator to produce the element y_1 which was recorded and fed back into the first generator to produce a third number and so on.

This procedure, using the constant values specified, is capable of producing 2^{15} numbers in the range 0 to $2^{15} - 1$. However, all elements of the form $4n + 2$, $n = 0,1,2,\ldots,2^{13} - 1$, are repeated while no elements of the form $4n$ appear. Hence, to obtain all the integers in the range 0 to $2^{15} - 1$, it is necessary only to subtract 2 from all elements of the

form $4n + 2$ when they occur in the sequence for the second time. The first 10,000 numbers of the sequence were generated, normalized, and recorded in 10 equal blocks. We discovered also that the same type of bias exists as in sequences obtained by the mixed generator, hence we generated 2,000 additional numbers by a method analogus to that employed for (2.9).

3. Digits in Irrational Numbers.

Many authors have reported successful experiments with digits in the decimal expansion of irrational numbers. Reitweisner at the suggestion of Von Neumann computed the values of π and e to more than 2,000 decimal places. The frequency distribution of these digits was consequently studied by Metropolis, Reitweisner, and Von Neumann.

Pathria [34] conducted a study of randomness among the first 10,000 digits of π. After these digits were grouped into ten blocks of 1,000, the four tests of Kendall and Babington Smith [21] and a fifth one due to Yule [49] called the "five-digit sum" test, were applied to each block. The results, on the whole, were satisfactory.

At the same time, Dr. D.B. Gilles of the Digital Computer Laboratory, University of Illinois, had computed e to 60,000 places and was attempting to extend this to one million places.

Shanks and Wrench [39]have extended both π and e to 100,000 decimal places and have computed the frequency distributions of each with satisfactory results.

Several experiments have been carried out on the decimal expansion of $\sqrt{2}$. Takahashi and Sibuya [42] computed the value of $\sqrt{2}$ to 14,000

decimal places and this result was extended by Lal [15] who achieved the expansion to 19,600 decimal places. A further extendion of these results was obtained by Lal [27] who, using the "Newton-Raphson"[1] method, performed the expansion to 39,000 decimal places. The above results were considerably improved by Lal and Lunnen [28] who achieved the expansion to 100,000 places. For this investigation, the Newton-Raphson method was again used and the computations were performed on the Atlas computer of the University of Manchester, England. Due to limited core size, the results could not be extended further. The chi-square values for the frequency of digits in consecutive blocks of 1,000 digits were computed and, of the 100 blocks tested, only four fell outside the 1 to 99 percent range.

We formed our fourth set of random numbers by grouping the digits of $\sqrt{2}$ into a sequence of 20,000 five-digit numbers. After each number was normalized, the sequence was recorded and divided into twenty equal blocks.

---

[1]Briefly, the method is as follows. Let x_1 be an approximate value of $\sqrt{2}$. A value of $\sqrt{2}$, accurate to twice the number of digits as x_1, is then given by

$$x_2 = \frac{1}{2}x_1 + x_1^{-1} \quad .$$

In this case, x_1 was known to 19,600 places. It was shown that the computed value x_2 was actually accurate to 39,074 places although only the first 39,000 were recorded.

[2]Dr. M. Lal of the Mathematics Department, Memorial University of Newfoundland had stored these digits on IBM punch cards - 50 digits to a card. Consequently, they were readily available for reuse.

4. [1]Bias in Sequences produced by Congruential Generators.

In Chapter II, we noted that sequences generated by congruential methods are subject to the constraint that, in all cases, repetition occurs after a definite number (the period of the sequence) of elements are generated. This constraint however, does not seriously affect the variability of such sequences since different choices of constants will guarantee sufficiently long sequences before repetition.[2] We also note that proofs of these facts have been given by several authors using techniques of number theory.

Peach [35] has observed further bias in these sequences, however, before investigating it we shall require the following definition.

Definition 2.1.

If the period of a sequence $\{x_n\}$ generated by congruential methods is M, then we define a "cycle of length M" to be a subset of M numbers of the sequence recorded in order of generation.

---

[1]In this section, the term "bias" shall refer to certain regular patterns and periodicities that have been observed in sequences produced by congruential generators.

[2]In theory, the period of the generated sequence can be made equal to any positive integer M, in practice, however, the value of M is limited by the capacity of the available computing machinery.

Given any cycle of length M, Peach observed that the difference between corresponding elements of the half-cycles[1] is always $\frac{M}{2}$. He uses the following example to illustrate the bias.

Consider the following sequence generated by

(2.15) ..... $x_n \equiv 9x_{n-1} + 13 \pmod{32}$,

0	13	2	31	4	17	6	3	8	21	10	7	12	25	14	11
16	29	18	15	20	1	22	19	24	5	26	23	28	9	30	27
0	13	2	21	..	..	..									

The first two lines comprise the two halves of a cycle of length 32. It is readily seen that the difference between corresponding numbers in the two half-cycles is always 16. This relationship is an inherent property of the generator and can not be eliminated by different choices of the constants. Peach, however, did not attempt to justify the existence of this bias nor could we find any other author who had done so. Consequently, we turned to number theory and derived the following results.

Theorem 2.1.

If the numbers $x_1, x_2, x_3, \ldots, x_2k$ represents a cycle of length 2^k from the sequence generated by

(2.16) ..... $x_n \equiv \lambda x_{n-1} + c \pmod{2^k}$,

then

---

[1]He also observed that the difference between corresponding elements of quarter-cycles is always a multiple of $\frac{M}{4}$, however, we shall not investigate this particular bias at this time.

(2.17) ..... $x_{j+2^{k-1}} - x_j \equiv 2^{k-1} \pmod{2^k}$,

where $\lambda \equiv 1 \pmod 4$, c is odd, and $0 \leq j \leq 2^{k-1}$.

Proof. (2.16) may also be written in the form

(2.18) ..... $x_n \equiv \lambda^n x_o + \frac{(\lambda^n - 1)c}{\lambda - 1} \pmod{2^k}$.

If $i = 0,1,2, \ldots,$ then by repeated application, we have

(2.19) ..... $x_{i+n} \equiv \lambda^{i+n} x_o + \frac{(\lambda^{i+n} - 1)c}{\lambda - 1} \pmod{2^k}$,

which reduces to

(2.20) ..... $x_{i+n} \equiv \lambda^n x_i + \frac{(\lambda^n - 1)c}{\lambda - 1} \pmod{2^k}$.

Since the period of the sequence is 2^k ,

(2.21) ..... $x_{i+2^k} - x_i \equiv 0 \pmod{2^k}$.

If we put $n = 2^k$ in (2.20), then (2.21) may be written in the form

(2.22) ..... $\frac{\lambda^{2^k} - 1}{\lambda - 1} [(\lambda - 1)x_i + c] \equiv 0 \pmod{2^k}$.

However, since $(\lambda - 1)$ is even and c is odd, the expression $[(\lambda - 1)x_i + c]$ is odd, so that

(2.23) ..... $\frac{\lambda^{2^k} - 1}{\lambda - 1} \equiv 0 \pmod{2^k}$,

that is

(2.24) ..... $\lambda^{2^k} \equiv 1 \pmod{2^k}$.

Also, k is the smallest integer for which (2.24) is true.

Consider the difference $x_{j+2^{k-1}} - x_j$, where $0 \leq j \leq 2^{k-1}$. From (2.18) and (2.19), we may write

(2.25) ..... $x_{j+2^{k-1}} - x_j \equiv \frac{(\lambda^{2^{k-1}} - 1)}{\lambda - 1} [(\lambda - 1)x_o + c]\lambda^j$.

For convenience, set

$$(2.26) \quad \ldots\ldots \theta = \frac{\lambda^{2^{k-1}} - 1}{\lambda - 1} ,$$

hence

$$(2.27) \quad \ldots\ldots \frac{\lambda^{2^{k}} - 1}{\lambda - 1} = \theta(1 + \lambda^{2^{k-1}}) \equiv 0 \pmod{2^k}.$$

Since,

$$(2.28) \quad \ldots\ldots \lambda \equiv 1 \pmod 4,$$

$$(2.29) \quad \ldots\ldots \lambda^{i} \equiv 1 \pmod 4,$$

where i is a non-negative integer.

Therefore

$$(2.30) \quad \ldots\ldots \lambda^{2^{k-1}} \equiv 1 \pmod 4 ,$$

for $k \geq 1$. Consequently, we now put

$$(2.31) \quad \ldots\ldots \lambda^{2^{k-1}} = 1 + 2S ,$$

where S is even, then

$$(2.32) \quad \ldots\ldots \lambda^{2^{k}} = \left(\lambda^{2^{k-1}}\right)^2 = 1 + 4S(1 + S).$$

Now, from (2.24),

$$(2.33) \quad \ldots\ldots S(1 + S) \equiv 0 \pmod{2^{k-2}}$$

and since $(1 + S)$ is odd, 2^{k-2} divides S.

(2.31) may now be written in the form

$$(2.34) \quad \ldots\ldots \lambda^{2^{k-1}} = 1 + 2(2^{k-2} \cdot \rho) ,$$

for some $\rho \in \mathbb{Z}$. Hence

$$(2.35) \quad \ldots\ldots \lambda^{2^{k-1}} = 1 + 2^{k-1}\rho .$$

From (2.27), we have

(2.36) ..... $\theta(2 + 2^{k-1}\rho) \equiv 0 \pmod{2^k}$,

that is,

(2.37) ..... $\theta(1 + 2^{k-2}\rho) \equiv 0 \pmod{2^k - 1}$

Since $(1 + 2^{k-2}\rho)$ is odd, we have

(2.38) ..... $\theta = \frac{\lambda^{2^{k-1}} - 1}{\lambda - 1} \equiv 0 \pmod{2^{k-1}}$,

hence

(2.39) ..... $\frac{\lambda^{2^{k-1}} - 1}{\lambda - 1} = 2^{k-1}\beta$,

for some $\beta \in \mathbb{Z}$.

Now, β must be odd, since, otherwise

(2.40) ..... $\lambda\frac{2^{k-1}}{\lambda - 1} - 1 = 2^k\gamma \equiv \pmod{2^k}$,

for some $\gamma \in \mathbb{Z}$, which contradicts (2.24). (2.25) must now be written in the form

(2.41) ..... $x_{j+2^{k-1}} - x_j \equiv 2^{k-1}\{\beta[(\lambda - 1)x_o + c]\lambda^j\} \pmod{2^k}$.

Since the coefficient of 2^{k-1} is odd,

(2.42) ..... $x_{j+2^{k-1}} - x_j = 2^{k-1} + 2^k\omega$,

for some $\omega \in \mathbb{Z}$, hence

(2.43) ..... $x_{j+2^{k-1}} - x_j \equiv 2^{k-1} \pmod{2^k}$.

Theorem 2.2

If $x_1, x_2, x_3, \ldots, x_{2^{k-2}}$ represents a cycle of length 2^{k-2} from the sequence generated by

(2.44) ..... $x_n \equiv \lambda x_{n-1} \pmod{2^k}$,

then

(2.45) ..... $x_{i+2^{(k-3)}} - x_i \equiv 2^{k-1} \pmod{2^k}$,

where $\lambda \equiv \pm 3 \pmod 8$, x_o is odd, and $0 \leq i \leq 2^{k-3}$.

Proof. Since $\lambda \equiv \pm 3 \pmod 8$, it is easily established that

(2.46) ..... $\lambda^{2t} \equiv 1 \pmod 4$,

for $t \in \mathbb{Z}$, hence

(2.47) ..... $\lambda^{2^{k-3}} \equiv 1 \pmod 4$,

for $k \geq 3$.

Now, the period of the sequence is 2^{k-2} ,

hence

(2.48) ..... $\left(\lambda^{2^{k-3}}\right)^2 = \lambda^{2^{k-2}} \equiv 1 \pmod{2^k}$

and no smaller value of k will satisfy this relation. Using (2.47) and (2.48), the derivation of

(2.49) ..... $\lambda^{2^{k-3}} = 1 + 2^{k-1} \cdot b$,

is analogus to that of (2.35). Also b is odd, otherwise

(2.50) ..... $\lambda^{2^{k-3}} = 1 + 2^k h \equiv 1 \bmod (2^k)$,

for some $h \in \mathbb{Z}$, which contradicts (2.48).

(2.44) may be written in the form

(2.51) ..... $x_n \equiv \lambda^n x_o \pmod{2^k}$,

therefore

(2.52) ..... $x_{i+2^{(k-3)}} - x_i \equiv \lambda^i x_o (\lambda^{2^{k-3}} - 1)$.

From (2.49), we have

(2.53) ..... $x_{i+2^{(k-3)}} - x_i \equiv (\lambda^i x_o b) 2^{k-1} \pmod{2^k}$.

Since λ^i, x_o and b are odd, the coefficient of 2^{k-1} is odd, hence

(2.54) ..... $x_{i+2^{(k-3)}} - x_i \equiv 2^{k-1} \pmod{2^k}$.

CHAPTER III

Tests for Randomness

We have defined a truly random sequence as one possessing two major qualities;

(i) The numbers in the sequence are uniformly distributed over the interval [0,1].

(ii) Each element in the sequence is independent of any other.

If an artifically generated sequence possesses these qualities to a reasonably high degree then, for practical purposes, it may be regarded as a random sequence. The degree to which a given sequence possesses these qualities is determined by specially designed tests that are referred to as "tests for randomness".

Many authors including Edmonds [9] and Taussky and Todd [43][1] have pointed out that both the numbers and the digits must be tested separately. Consequently, our own battery of tests consists of three designed to test the randomness of the numbers and three to test the randomness of digits.

(1) The Chi-Square Test for Goodness of Fit.

This test, the most widely used in the literature, is used to test the hypothesis that a given generated sequence has uniform distribution over the interval [0,1].

---

[1]Taussky and Todd report that experiments carried out by M.L. Juncosa show that sequences which were very good random numbers gave rise to sequences of random digits which could, at best, be classed as fair.

Suppose, in general, we wish to test the hypothesis that a given population is distributed according to the distribution function $F(x)$. A random sample of n observations is drawn from the population and divided into k mutually exclusive categories. Karl Pearson, in 1900, suggested computing the statistic

$$(3.1) \quad \chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} ,$$

where n_i denotes the number of observations in the i^{th} category and p_i is the probability that a given observation falls into the i^{th} category. The exact probability distribution of the random variable χ^2 is quite complicated but, as $n \to \infty$, its distribution is approximately chi-square with $(k - 1)$ degrees of freedom. A value r is then chosen and the hypothesis is rejected if the computed value of χ^2 exceeds r.

In the case of random sequences, $F(x)$ is the uniform distribution over the interval $[0,1]$ and the categories consist of k equal subintervals of $[0,1]$. Since the measure of all categories is the same, $p_i = 1/k$ and $np_i = n/k$, $i = 1,2,3,\ldots,k$. Hence, for this application, the χ^2 statistic (3.1) reduces to the simpler form

$$(3.2) \quad \chi^2 = \left(\frac{k}{n} \sum_{i=1}^{k} n_i^2\right) - n ,$$

also with $(k - 1)$ degrees of freedom[1]. This statistic is more efficient for computation purposes than (3.1). The value r was chosen from the tables such that

$$(3.3) \quad P[\chi^2 > r] = 0.05.$$

---

[1]After we had derived (3.2), we received the work of Newman and Odell [33]. They had also recognized the existence of a simpler statistic for this application. However, they gave the simplified statistic as $\chi^2 = k\left(\sum_{i=1}^{k} n_i^2\right) - n$ which is incorrect.

(2) The Kolmogorov-Smirnov Test for Goodness of Fit.

This is the most important of the general tests of fit alternative to chi-square [24] and is also used to test the hypothesis that a given sequence is uniformly distributed over the interval [0,1]. The chi-square test tends to be rather insensitive and also depends upon the arbitrary division of [0,1] into subintervals. The Kolmogorov-Smirnov test requires no such division as the test statistic is computed over the whole of the interval [0,1]. However, it is more difficult to apply than the chi-square as all observations must be ranked.

The mathematical basis for the test is as follows. Suppose that an ordered sample $x_1, x_2, x_3, \ldots, x_n$, where $x_i < x_{i+1}$ has been drawn from a sequence with distribution function $F(x)$ - uniform on the interval [0,1]. The observed cumulative relative frequency $S_n(x)$ is defined as the step function

(3.4) ..... $S_n(x) = \frac{k}{n}$,

where $x_k \leq x < x_{k+1}$ and $k = 1,2,3,\ldots,n$.

From the following theorem we will expect that $S_n(x)$ will be a reasonably good approximation to $F(x)$. In addition, this approximation should improve as n increases.

Theorem 3.1.

If $F(x)$ is the uniform distribution on the interval [0,1] and if $S_n(x)$ is defined as in (3.4), then

(3.5) ..... $\lim_{n \to \infty} P[|S_n(x) - F(x)| < \epsilon] = 1,$

for all $\epsilon > 0$.

Proof:

Suppose that $y_1, y_2, y_3, \ldots, y_n$ are the unranked values of the generated sequence. Define

$$(3.6) \quad \ldots\ldots \quad Z_i = \begin{cases} 0, & \text{if } y_i > x \\ 1, & \text{if } y_i \leq x \end{cases},$$

where $i = 1,2,3,\ldots,n$ and Z_i are all independent.

Define

$$(3.7) \quad \ldots\ldots \quad G_n(Z) = Z_1 + Z_2 + Z_3 + \ldots + Z_n \ ,$$

then

$$(3.8) \quad \ldots\ldots \quad S_n(x) = \frac{G_n(Z)}{n} = \frac{k}{n} .$$

From (3.6) it is clear that

$$(3.9) \quad \ldots\ldots \quad E(Z_i) = F(x).$$

Now, the strong law of large numbers states that

$$(3.10) \quad \ldots\ldots \quad \lim_{n \to \infty} P[|n^{-1}G_n(z) - E(Z_i)| < \epsilon] = 1,$$

for all $\epsilon > 0$.

Hence (3.5) is verified.

The test statistic, D_n, is the least upper bound of the absolute deviation of $S_n(x)$ from $F(x)$, i.e.

$$(3.11) \quad \ldots\ldots \quad D_n = \underset{(x)}{\ell.u.b} |S_n(x) - F(x)| .$$

The distribution of D_n is completely distribution-free when the null hypothesis holds provided that F is continuous [24]. Kendall and

Stuart [24] have given a very terse justification of this fact in the following way. If $S_n(x)$ and $F(x)$ are plotted as ordinates against x as abscissa, D_n is simply the largest vertical difference between them. Clearly, if we make any one-to-one transformation of x, this will not affect the vertical difference at any point and, in particular, the value of D_n will be unaffected. Hence we can take

(3.12) ..... $F(x) = x \quad , \quad 0 \leq x \leq 1.$

Kolgomorov [25] proved that, for any $\epsilon > 0$,

(3.13) ..... $\lim_{n \to \infty} P[n^{1/2} D_n \geq \epsilon] = L(\epsilon),$

where

(3.14) ..... $L(\epsilon) = 2 \sum_{n=1}^{\infty} (-1)^{n+1} e^{-2n^2\epsilon^2} .$

We shall choose the value $\epsilon = 1.36$ so that, $L(\epsilon) = 0.05$. Therefore,

(3.15) ..... $P[D_n > \frac{1.36}{n^{1/2}}] = 0.05.$

Since we shall take $n = 1000$ for each sequence, (3.15) becomes

(3.16) ..... $P[D_n > 0.043007] = 0.05.$

The statistic D_n is computed from the observed data and the hypothesis is rejected if the computed value exceeds 0.043007.

3. The Runs Up and Down Test.

This test is used to test the hypothesis that the elements of a generated random sequence are independent.

Consider a sequence $\{x_i | i = 1,2,3,\ldots,n\}$ of random numbers. A subsequence

(3.16) ..... $x_{i-1}, x_i, \ldots, x_{i+r}, x_{i+r+1}$ $(2 \leq i \leq n - r - 1)$

of $(r + 3)$ consecutive numbers is said to form a "run up" of length r if

(3.17) ..... $x_{i-1} > x_i < x_{i+1} < \ldots < x_{i+r} > x_{i+r+1}$.

If all the signs in (3.17) are reversed, then a "run down" of length r is defined.

The "runs up and down" test is based on a comparison of the expected and actual numbers of runs of various lengths under the hypothesis of independence. All the relevant results that are necessary to derive the test statistic have been given in detail by Downham [7]. A summary of these results now follow.

The expected number of runs of length r is given by

$$(3.18) \quad \ldots\ldots \quad E(r) = \begin{cases} \dfrac{2[(r^2 + 3r + 1)n - (r^3 + 3r^2 - r - 4)]}{(r + 3)!}, & r < n - 1 \\ \dfrac{2}{n!}, & r = n - 1 \end{cases}$$

where n is the number of elements in the sequence.

The expected number of runs of length r or greater is given by

$$(3.19) \quad \ldots\ldots \quad E^1(r) = \sum_{\alpha=r}^{n-1} E(\alpha).$$

The well known Pearson chi-square statistic now takes the form

$$(3.20) \quad \ldots\ldots \quad \sum_{i=1}^{r-1} \frac{(N_\alpha - E(\alpha))^2}{E(\alpha)} + \frac{(\sum_{i=1}^{n-1} N_\alpha - E^1(r))^2}{E^1(r)}$$

which is assymptotically distributed as chi-square with $(r - 1)$ degrees of freedom. N_α is the observed number of runs of length α and $E(\alpha)$ and $E^1(r)$ are computed from (3.18) and (3.19) respectively. The value

of r is usually taken to be 5, in which case, the critical value at the 0.05 percent level is 9.490.

4. Frequency Test.

If a sequence is truly random, each digit should occur an equal number of times. The frequency test is used to determine whether or not a generated sequence possesses this property to a reasonably high degree. If a particular sequence is composed of m digits[1], the expected number of occurances of each digit is $m/10$. The observed frequencies m_i of the digits are obtained from the row or column totals in a table of digital pairs. An example of such a table is shown in fig. (3.1).

Now, the statistic

$$(3.21) \quad \ldots\ldots \quad \chi^2 = \sum_{i=1}^{10} \frac{(m_i - m/10)^2}{m/10}$$

which, for computational convenience, reduces to

$$(3.22) \quad \ldots\ldots \quad \chi^2 = \frac{1}{500} \left(\sum_{i=1}^{10} m_i^2 \right) - 5{,}000$$

is asymptotically distributed as chi-square with 9 degrees of freedom. Also, if $r = 16.92$,

$$(3.23) \quad \ldots\ldots \quad P[\chi^2 > r] = 0.05.$$

The test is applied by computing the value of χ^2 in (3.22) and comparing it with the value $r = 16.92$.

---

[1]The generated random numbers x_i, distributed over the unit interval are worked out to five significant digits, hence $m = 5n$, where n is the number of random numbers in the sequence. Since we are testing all numbers in blocks of 1,000, $m = 5000$.

5. Serial Test.

In a sequence of truly random numbers no digit in any given number will show a tendency to be followed by another particular digit. The serial test is used to determine whether or not a generated sequence possesses this property to a reasonably high degree.

In the sequence of digits

(3.24) ..... $s_1, s_2, s_3, \ldots, s_m$,

we consider the $(m - 1)$ pairs of the form (s_n, s_{n+1}), $(n = 1,2,3,\ldots,m - 1)$ and the pair (s_m, s_1). Under the hypothesis of independence, each pair has an expected frequency of $\frac{m}{100}$. A table, such as shown in fig. (3.1) is constructed by observing the number $m_{ij} (i,j = 1,2,3,\ldots,10)$, of the digit pairs $(i - 1, j - 1)$ and entering this number in the cell determined by the intersection of the i^{th} row and the j^{th} column.

Kendall and Babington Smith [21] proposed the statistic

$$(3.25) \quad \ldots\ldots \quad \chi_1^2 = \sum_{i=1}^{10} \sum_{j=1}^{10} \frac{(m_{ij} - \frac{m}{100})^2}{m/100}$$

which, they claimed, is approximately chi-square with 90 degrees of freedom. However, it was pointed out by Good [13] that

$$(3.26) \quad \ldots\ldots \quad E(\chi_1^2) = 99,$$

hence the random variable χ_1^2 cannot have an approximate chi-square distribution with 90 degrees of freedom. He then proceeds to show that if

$$(3.27) \quad \ldots\ldots \quad m_i = \sum_{i=1}^{10} m_{ij} \quad ,$$

then the statistic

$$(3.28) \quad \ldots\ldots \quad \chi_2^2 = \left(\frac{100}{m}\right) \sum_{i=1}^{10} \sum_{j=1}^{10} \left(m_{ij} - \frac{m}{100}\right)^2 - \left(\frac{10}{m}\right) \sum_{i=1}^{10} \left(m_i - \frac{m}{10}\right)^2$$

is distributed approximately as chi-square with 90 degrees of freedom. For computational convenience, (3.28) reduces to

$$(3.29) \quad \ldots\ldots \quad \chi_2^2 = \frac{1}{50} \left(\sum_{i=1}^{10} \sum_{j=1}^{10} m_{ij}^2 \right) - \left(\frac{1}{500} \sum_{i=1}^{10} m_i^2 \right)$$

Also, if $r = 113.14$,

$$(3.30) \quad \ldots\ldots \quad P[\chi_2^2 > r] = 0.05.$$

The serial test is applied by computing the statistic χ_2^2 in (3.29) and comparing the result with the value of r in (3.30).

Table of Digital Pairs

(First block of generated numbers)

First digit \ Second digit	0	1	2	3	4	5	6	7	8	9	TOTAL
0	55	44	75	57	47	52	50	38	51	54	523
1	48	50	49	42	40	51	41	59	47	58	485
2	54	45	40	49	51	44	64	44	52	54	497
3	54	52	45	44	49	58	48	49	51	52	502
4	44	48	50	55	47	51	52	47	53	57	504
5	47	47	35	55	55	42	47	49	50	56	483
6	59	49	56	55	54	47	48	46	45	55	514
7	52	50	56	39	52	46	39	44	48	47	473
8	52	48	48	49	56	47	65	41	46	43	495
9	58	52	43	57	53	45	60	56	52	48	524
Total	523	485	497	502	504	483	514	473	495	524	5,000

fig. (3.1)

6. The Poker Test.

The combination of digits in each generated number corresponds to a combination similar to that which occurs in poker. Under the hypothesis of independence, the probabilities of all possible combinations are readily computed using the multinomial theorem. The table, fig. (3.2), showing the various combinations and their corresponding probabilities is taken from Jansson [19].

Combination	Probability
Bust (abcde)	.3024
Pair (aabcd)	.5040
2 pairs (aabbc)	.1080
3 of a kind (aaabc)	.0720
Full house (aaabb)	.0090
4 of a kind (aaaab)	.0045
5 of a kind (aaaaa)	.0001

fig. (3.2)

The statistic

$$(3.31) \quad \ldots\ldots \quad \chi^2 = \sum_{i=1}^{k} \frac{(w_i - wp_i)^2}{wp_i} \quad ,$$

where w is the total number of combinations, w_i is the observed number of the i^{th} combination, and p_i is the probability that the i^{th} combination occurs, is distributed approximately as chi-square with $(k - 1)$ degrees of freedom.

Since, in this case , $k = 7$,

(3.32) ..... $P[\chi^2 > r] = 0.05$

for $r = 12.59$.

The poker test is applied by computing the statistic χ^2 in (3.31) and comparing the result with the value of r in (3.32).

CHAPTER IV

Methods of Monte Carlo Integration

In this chapter, we examine how simulation techniques may be used to evaluate definite integrals. The basic technique is straightforward but, unfortunately, requires an excessively large supply of random numbers to produce a reasonable degree of accuracy. However, this difficulty has been partially aleviated by the introduction of various refinements of the basic technique. We examine the basic method and two popular refinements of it.

1. The Crude Estimator (The Basic Method).

The Monte Carlo calculation of a definite integral may be considered as a statistical problem of estimating the parameter θ, where

$$(4.1) \quad \theta = \int_0^1 f(x)dx.$$

We may estimate θ in the following way.

Consider a sample $X_1, X_2, X_3, \ldots, X_n$ of n independent random variables uniformly distributed over the interval $[0,1]$. Then the random variable $\hat{\theta}$ defined by

$$(4.2) \quad \hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} f(X_i)$$

is an unbiased estimator of θ. Furthermore, the variance of $\hat{\theta}$ is given by

$$(4.3) \quad \mathrm{Var}(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{Var}(f(X_i)).$$

From the Chebychev inequality, we have, for any estimator $\hat{\theta}$ with mean θ and variance σ^2,

(4.4) ..... $P[|\hat{\theta} - \theta| \geq c\,\sigma] \leq \frac{1}{c^2}$,

for $c > 0$. Hence if we put $\sigma = \sqrt{Var(\hat{\theta})}$ and $c = \frac{1}{\sqrt{\epsilon}}$, $\epsilon > 0$, then (4.4) takes the form

(4.5) ..... $P\left[|\hat{\theta} - \theta| \geq \frac{\sqrt{\sum_{i=1}^{n} Var(f(X_i))}}{\sqrt{\epsilon\, n}}\right] \leq \epsilon$.

Hence the expression, $|\hat{\theta} - \theta|$, which is the error of the method, is proportional to $\frac{1}{\sqrt{n}}$. Since this error can be relatively large[1] , $\hat{\theta}$ is often referred to as the "crude" estimator of θ .

Consider the integral

(4.6) ..... $\theta_1 = \int_0^1 [x^3 + 1]dx$.

We may evaluate this integral by the estimator $\hat{\theta}_1$, where

(4.7) ..... $\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} [X_i^3 + 1]$.

The variance of this estimator is given by

(4.8) ..... $Var(\hat{\theta}_1) = \frac{1}{n}[Var\ X_i^3 + 1]$

$= \frac{1}{n}[Var\ X_i^3]$

$= \frac{1}{n}[\int_0^1 x^6 dx - [\int_0^1 x^3 dx]^2]$

$= \frac{9}{112n}$

---

[1] A reduction of this error to even $\frac{1}{10}$ of its former size would require a sample 100 times as large.

2. The Importance Sampling Estimator.

The main weakness of the previous method is that the variance of $\hat{\theta}$ can not be reduced without a significant increase in the sample size. To overcome this difficulty, other estimators have been developed which, for a given n, have smaller variances than $\hat{\theta}$. One such estimator may be developed as follows.

Let X be a random variable with probability density function $g(x)$ defined on the interval $[0,1]$. The integral (4.1) may now be rewritten in the form

$$(4.9) \quad \ldots\ldots \quad \theta_2 = \int_0^1 \left(\frac{f(x)}{g(x)}\right) g(x)dx.$$

Since the mathematical expectation of the function $\frac{f(X)}{g(X)}$ is equal to θ_2, it follows that (4.9) may be estimated from

$$(4.10) \quad \ldots\ldots \quad \hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} \frac{f(X_i)}{g(X_i)}$$

where $X_1, X_2, X_3, \ldots, X_n$ are values of the random variable X with density function $g(x)$.

The variance of $\frac{f(X)}{g(X)}$ is given by

$$(4.11) \quad \ldots\ldots \quad \mathrm{Var}\left(\frac{f(X)}{g(X)}\right) = \int_0^1 \frac{f^2(x)}{g(x)}\,dx - \theta_2^2 .$$

The minimum variance is obtained [1] when X has the distribution given by

---

[1] Using elementary techniques of the calculus of variations [1].

$$(4.12) \quad \ldots\ldots \quad g(x) = \frac{|f(x)|}{\int_0^1 |f(x)|dx} \quad ,$$

Hence, (4.11) may be rewritten in the form

$$(4.13) \quad \ldots\ldots \quad \mathrm{Var}\left(\frac{f(X)}{g(X)}\right) = \left[\int_0^1 |f(x)|dx\right]^2 - \theta_2^2.$$

If the function $f(X)$ does not change sign then

$$(4.14) \quad \ldots\ldots \quad \mathrm{Var}\left(\frac{f(X)}{g(X)}\right) = 0.$$

Since the evaluation of the integral

$$(4.15) \quad \ldots\ldots \quad \int_0^1 |f(x)|dx$$

is almost equivalent to evaluating θ_2, this is of theoretical interest only. The practical alternative is to choose the distribution of X such that

$$(4.16) \quad \ldots\ldots \quad \frac{g(x)}{|f(x)|} \approx \text{constant}.$$

If $f(x)$ has a power series expansion, then $g(x)$ can take the form

$$(4.17) \quad \ldots\ldots \quad g(x) = \frac{|h(x)|}{\int_0^1 |h(x)|dx} \quad ,$$

where $h(x)$ is a function, preferably linear, composed from the first few terms of the expansion.

Consider the integral

$$(4.18) \quad \ldots\ldots \quad \theta_2 = \int_0^1 e^x dx \quad .$$

Since

(4.19) ..... $e^x = 1 + x + x^2 + \ldots$,

we may consider the random variable X with probability density

(4.20) ..... $g(x) = \frac{2}{3}(1 + x)$, $0 \leq x \leq 1$.

Hence, by (4.10), θ_2 may be estimated by

(4.21) ..... $\hat{\theta}_2 = \frac{3}{2n} \sum_{i=1}^{n} \frac{e^{X_i}}{1 + X_i}$.

The variance of $\hat{\theta}_2$ is given by

(4.22) ..... $\mathrm{Var}(\hat{\theta}_2) = \frac{\frac{3}{2}\int_0^1 \frac{e^{2x}}{1 + x} dx - \theta_2^2}{n}$

$= \frac{0.0269}{n}$.

If X^1 is uniformly distributed on the interval $[0,1]$, the values of X may be obtained from the formula

(4.23) ..... $X = \sqrt{1 + 3X^1} - 1$.

Hence (4.21) becomes

(4.24) ..... $\hat{\theta}_2 = \frac{3}{2n} \sum_{i=1}^{n} \frac{e^{\sqrt{1 + 3X_i^1}} - 1}{\sqrt{1 + 3X_i^1}}$,

where X_i^1 is uniformly distributed on $[0,1]$.

When $g(x)$ is chosen according to (4.16), sampling from X will insure that the larger values of $f(x)$ will have a higher probability of occurance. Since these values will make a greater contribution to the integral, this method is referred to as "importance sampling".

3. The "Weighted Uniform Sampling" Estimator.

Since the "importance sampling" technique usually requires the user to sample from a random variable X with non-uniform distribution, it may be difficult to devise a practical method of generating these values. If, as in (4.24), a transformation is used to derive X from X^1, uniformly distributed on $[0,1]$, then the evaluation of the estimator could be quite complicated. The "weighted uniform sampling" method, introduced by Handscomb [50] and analysed by Powell and Swann[36], allows the user to apply samples from the uniform distribution directly to the estimator.

The estimator employed is

$$\hat{\theta}_3 = \frac{\sum_{i=1}^{n} f(X_i)}{\sum_{i=1}^{n} g(X_i)} \quad , \tag{4.25}$$

where X_i is uniform over the interval $[0,1]$. The function $g(X)$ has the property that

$$\int_0^1 g(x)dx = 1, \tag{4.26}$$

although its values are not necessarily all non-negative.

Furthermore, if $g(X)$ is chosen so that

$$\text{Var}\left(\frac{f(X)}{g(X)}\right) \approx \text{constant}, \tag{4.27}$$

then the variance of $\hat{\theta}_3$ will be relatively small.

The estimator $\hat{\theta}_3$ is, of course, biased however it was shown by Powell and Swann [36] that

$$\left[\frac{\beta^2}{\delta}\right] = O\left(\frac{1}{n}\right) \quad , \tag{4.28}$$

where β is the bias and δ is the mean square error of $\hat{\theta}_3$. They also note that the magnitude of β depends upon the variance of $g(X)$.

Consider the integral

$$(4.29) \quad \ldots\ldots \theta_3 = \int_0^{10} [x + \cos^2 x]dx$$

and the corresponding function $g(x)$, where

$$(4.30) \quad \ldots\ldots g(x) = \frac{57.7(x + \cos^2 x)}{x + 0.77} \quad ,$$

then the estimator $\hat{\theta}_3$ is given by

$$(4.31) \quad \ldots\ldots \hat{\theta}_3 = \frac{57.7\left[\sum_{i=1}^{n} X_i + \cos^2 X_i\right]}{\sum_{i=1}^{n} X_i + 0.77} \quad ,$$

where X_i is uniformly distributed over the unit interval. Powell and Swann [36] showed that, in general, the accuracy of this method is comparable to importance sampling and the bias β is not damaging.

CHAPTER V

Results and Conclusions

A total of fifty-two blocks, each consisting of one thousand random numbers, were generated using the methods described in Chapter II. Each block was tested for randomness and applied to the evaluation of the three definite integrals (4.6), (4.18), and (4.31) described in the previous chapter.

Since we were interested mainly in investigating randomlike properties of the generated sequences, no attempt was made to evaluate the various methods of integration used. The rationale of choosing three entirely different integrals, each computed by a different method, was to obtain results that were independent of a particular type of integral or a particular method of computation.

The generation and testing were programmed in Fortran IV language and executed using an IBM System 360/40.

1. Results of Tests.

The results of the six tests for randomness for each block of generated random numbers are given in Tables 5.1 to 5.4 inclusive. All tests, except the Kolmogorov-Smirnov test for goodness of fit, involve either the direct application of the Pearson chi-square statistic or some variation of it.

After each table, we note the cases where the 95% level was exceeded and we refer to these as instances where particular blocks "fail" particular tests.

TABLE 5.1

The Multiplicative Generator

Block	Chi-square Test	Kolmogorov-Smirnov Test	Runs Test	Frequency Test	Serial Test	Poker Test
1	9.920	0.022	2.311	5.196	71.044	2.416
2	17.952	0.033	4.247	3.072	60.808	5.334
3	12.480	0.022	0.448	3.464	63.536	20.690
4	10.080	0.028	3.352	3.572	57.868	6.272
5	9.568	0.017	1.552	3.036	63.444	5.919
6	14.880	0.036	3.569	2.204	51.716	2.091
7	13.472	0.022	8.918	2.220	50.740	4.547
8	13.664	0.294	3.907	8.416	68.104	3.977
Critical value at 95% level	25.000	0.043	9.490	16.920	124.34	12.590

Of the eight blocks tested, block 3 failed the poker test and block 8 failed the Kolmogorov-Smirnov test.

TABLE 5.2

The Mixed Congruential Generator

Block	Chi-square Test	Kolmogorov-Smirnov Test	Runs Test	Frequency Test	Serial Test	Poker Test
1	7.232	0.010	41.251	11.120	60.728	6.983
2	4.768	0.012	38.612	2.528	63.992	17.290
3	8.672	0.014	33.831	2.428	54.612	3.785
4	5.504	0.014	58.214	2.252	63.108	2.203
5	7.136	0.013	44.020	2.012	45.068	7.257
6	5.728	0.012	41.793	1.692	53.308	4.164
7	5.952	0.015	48.949	1.708	59.812	4.164
8	6.656	0.019	40.654	2.004	47.596	8.009
9	7.456	0.014	41.056	1.080	62.400	4.419
10	6.592	0.010	38.768	2.660	51.580	3.063
11	6.880	0.011	40.213	2.048	67.112	2.362
12	5.440	0.015	45.161	2.260	64.500	10.430
Critical Value at 95% level.	25.000	0.043	9.490	16.920	124.34	12.590

Of the twelve blocks tested, block 2 failed the poker test and all 12 failed the runs test. As noted in Chapter II, section 3(b), this generator generally performs badly on this test and this performance seems to depend upon the value of c used. We tried the values $c = 21$, 517, 17,005 and 29,001 and of these $c = 21$ and $c = 29,001$ gave the best results. The

results in Table 5.2 are for $c = 21$.

We feel that the best values for c may be those near 0 or 2^k and we intend to further investigate this matter at a later date.

TABLE 5.3

The Combination of Two Congruential Generators

Block	Chi-square Test	Kolmogorov-Smirnov Test	Runs Test	Frequency Test	Serial Test	Poker Test
1	10.944	0.021	8.457	3.404	69.196	2.059
2	12.864	0.030	1.329	4.700	55.220	5.285
3	12.288	0.037	4.478	4.620	63.020	15.860
4	16.544	0.019	4.094	1.224	69.816	12.620
5	12.832	0.026	3.035	7.840	67.680	4.992
6	8.768	0.023	0.601	6.324	75.656	17.210
7	13.920	0.024	3.826	5.292	69.468	3.644
8	17.408	0.040	4.664	2.312	76.648	3.481
9	13.760	0.021	1.226	3.368	72.592	8.793
10	15.072	0.034	2.865	7.900	69.740	4.480
11	21.056	0.034	0.739	4.404	67.676	5.294
12	15.072	0.024	2.648	7.772	78.868	15.860
Critical value at 95% level	25.000	0.043	9.490	16.920	124.34	12.590

Of the twelve blocks tested, blocks 3, 4, 6, and 12 failed the poker test.

TABLE 5.4

Expansion of $\sqrt{2}$ to 100,000 places

Block	Chi-square Test	Kolmogorov-Smirnov Test	Runs Test	Frequency Test	Serial Test	Poker Test
1	2.192	0.051	2.954	9.732	101.670	4.566
2	7.300	0.038	1.548	4.292	91.348	10.430
3	7.500	0.037	11.695	4.340	108.900	1.205
4	13.240	0.032	5.488	6.720	83.960	2.037
5	14.040	0.050	4.119	11.732	106.150	6.505
6	4.720	0.020	5.875	7.192	76.008	4.268
7	15.940	0.032	7.145	14.100	80.500	18.630
8	6.340	0.034	5.402	8.136	91.824	3.140
9	7.840	0.034	13.779	8.856	84.984	0.649
10	11.900	0.040	2.631	8.120	77.800	5.410
11	8.400	0.020	1.247	4.312	94.288	2.952
12	7.740	0.021	2.593	11.044	102.400	19.510
13	6.700	0.019	9.764	5.336	83.584	1.364
14	11.563	0.032	5.906	8.092	104.030	1.996
15	10.600	0.034	7.683	11.856	79.024	3.064
16	4.320	0.032	7.268	14.876	110.600	3.760
17	8.140	0.024	3.439	15.832	100.810	14.000
18	9.760	0.025	6.088	2.444	78.516	7.517
19	3.820	0.015	1.941	3.340	73.340	4.398
20	9.260	0.022	1.541	10.824	102.940	13.940
Critical value at 95% level	25.000	0.043	9.490	16.920	124.34	12.590

In general, the performance of these numbers on the tests for randomness was comparable to those generated by congruential methods. Table 5.5 indicates the blocks that failed certain of the tests.

TABLE 5.5

TESTS FAILED	BLOCKS
Kolmogorov-Smirnov	1, 5
Runs	3, 9, 13
Poker	7, 12, 17, 20

Even though no block actually failed the frequency or the serial tests, the chi-Square values were generally higher than for the congruential generators.

2. Results of the Monte Carlo Integration

In Tables 5.6 to 5.9 inclusive, the results of the Monte Carlo integration are presented. Each block of generated numbers is regarded as a random sample from the distribution uniform on [0,1] and the integrals are evaluated in accordance with the methods described in Chapter IV. The true value of each integral is given in each of the tables.

TABLE 5.6

The Multiplicative Generator

Block	$\int_0^1 (x^3 + 1)dx$	$\int_0^1 e^x dx$	$\int_0^{10} (x + \cos^2 x)dx$
1	1.2614	1.7236	55.4403
2	1.2574	1.7254	55.3463
3	1.2444	1.7170	55.8706
4	1.2380	1.7108	56.1970
5	1.2456	1.7165	55.8766
6	1.2421	1.7130	56.0851
7	1.2561	1.7191	55.6904
8	1.2560	1.7222	55.5221
True Value	1.2500	1.7183	55.2282

TABLE 5.7

The Mixed Congruential Generator

Block	$\int_0^1 (x^3 + 1)dx$	$\int_0^1 e^x dx$	$\int_0^{10} (x + \cos^2 x)dx$
1	1.2518	1.7195	55.6906
2	1.2495	1.7182	55.7778
3	1.2488	1.7169	55.8316
4	1.2536	1.7206	55.6276
5	1.2542	1.7207	55.6214
6	1.2468	1.7166	55.8614
7	1.2470	1.7160	55.8883
8	1.2565	1.7226	55.5140
9	1.2483	1.7171	55.8362
10	1.2514	1.7190	55.7124
11	1.2495	1.7162	55.7736
12	1.2441	1.7201	55.9525
True Value	1.2500	1.7183	55.2282

TABLE 5.8

The Combination of Two Congruential Generators

Block	$\int_0^1 (x^3 + 1)dx$	$\int_0^1 e^x dx$	$\int_0^{10} (x + \cos^2 x)dx$
1	1.2494	1.7192	55.7160
2	1.2560	1.7230	55.5113
3	1.2383	1.7111	56.2038
4	1.2564	1.7226	55.5103
5	1.2522	1.7180	55.7611
6	1.2394	1.7137	56.0486
7	1.2508	1.7203	55.6495
8	1.2377	1.7083	56.3510
9	1.2467	1.7157	55.9100
10	1.2495	1.7159	55.8586
11	1.2553	1.7089	55.4851
12	1.2407	1.7244	56.0950
True Value	1.2500	1.7183	55.2282

TABLE 5.9

Expansion of $\sqrt{2}$ to 10,000 places

Block	$\int_0^1 (x^3 + 1)dx$	$\int_0^1 e^x dx$	$\int_0^{10} (x + \cos^2 x)dx$
1	1.2662	1.7289	55.1283
2	1.2342	1.7093	56.3216
3	1.2430	1.7122	56.1043
4	1.2470	1.7170	55.8033
5	1.2258	1.7033	56.6743
6	1.2430	1.7150	55.9713
7	1.2596	1.7245	55.4146
8	1.2638	1.7254	55.3264
9	1.2637	1.7258	55.2972
10	1.2416	1.7130	56.1010
11	1.2544	1.7191	55.6929
12	1.2493	1.7187	55.7585
13	1.2517	1.7184	55.7439
14	1.2394	1.7151	55.9908
15	1.2604	1.7257	55.3548
16	1.2574	1.7236	55.4538
17	1.2423	1.7137	56.0388
18	1.2509	1.7167	55.8332
19	1.2516	1.7195	55.6938
20	1.2455	1.7154	55.9140
True Value	1.2500	1.7183	55.2282

3. Correlations and Conclusions.

Several authors, including Hull and Dobell [17] and Jansson [19], have stated that no finite class oftests can guarantee the general suitability of a finite sequence of numbers. Therefore, to determine the suitability for a particular application, the generated sequences have to be studied in the light of that application.

To the computed data, we applied the "Descriptive Statistics Package with Data Transformations" [11] provided by the Division of Educational Research Service, University of Alberta. The correlations between the test scores and the absolute error in the computed values of the integrals are summarized in Tables 5.10 and 5.11. The actual values of the correlation coefficient are given in Appendix 1, page 54.

The following legend is used in Tables 5.10 and 5.11 and Appendix 1.

$I_1 - \int_0^1 (x^3 + 1)dx$

$I_2 - \int_0^1 e^x dx$

$I_3 - \int_0^{10} (x + \cos^2 x)dx$

G_1 - The multiplicative generator

G_2 - The mixed generator

G_3 - The combination of two generators

G_4 - Digits in $\sqrt{2}$.

T_1 - The chi-square goodness of fit

T_2 - The Kolmogorov-Smirnov test

T_3 - The runs test

T_4 - The frequency test

T_5 - The serial test

T_6 - The poker test

TABLE 5.10

Most Significant[1] Tests for each Integral

	G_1	G_2	G_3	G_4	Highest Frequency of Occurance
I_1	T_3, T_6, T_1	T_2, T_3, T_5	T_2, T_6, T_5	T_2, T_6, T_4	T_6 and T_2
I_2	T_6, T_1, T_4	T_2, T_5, T_3	T_3, T_1, T_4	T_2, T_3, T_4	T_3 and T_4
I_3	T_5, T_4, T_1	T_5, T_4, T_2	T_5, T_2, T_6	T_1, T_3, T_5	T_4 and T_5

From these results, we make the following conclusions:

(i) When employing the crude estimator, the Kolmogorov-Smirnov and poker tests appear to be the most useful.

(ii) When employing variance reducing techniques the runs, frequency, and serial tests appear to be the most useful.

(iii) Although most authors regard the chi-square goodness of fit test as basic when testing for randomness, it does not appear to be useful for this application.

TABLE 5.11

Most Significant Tests for each Generator

G_1	T_1, T_4, T_6
G_2	T_2, T_5, T_3
G_3	T_2, T_5, T_6
G_4	T_2, T_3, T_4

[1]The significance of a test for a particular integral evaluation is conveniently measured by the absolute value of the correlation coefficient.

We conclude from these results that

(i) When testing the randomness of numbers, for this application, the Kolmogorov-Smirnov test is more useful than the chi-square goodness of fit test or the runs test.

(ii) When testing the randomness of digits it appears tnat, generally, no one test is more useful than the others. We suggest that, for this aspect of testing, each generator has to be considered individually. From our investigation it appears that

(a) for congruential generators, the serial and poker tests are the most useful.

(b) for the digits in $\sqrt{2}$, the frequency test is the most useful.

APPENDIX 1

Values of the Correlation Coefficient

		T_1	T_2	T_3	T_4	T_5	T_6
G_1	I_1	-0.2650	-0.2135	-0.3576	-0.0523	0.0918	-0.3223
	I_2	0.2128	0.0115	-0.0988	0.1007	0.0716	-0.4217
	I_3	-0.3862	-0.2980	-0.1861	-0.4254	0.5123	0.1944
G_2	I_1	-0.2257	0.7357	0.3973	-0.1745	-0.3938	0.0984
	I_2	0.0924	0.7458	0.3178	-0.2148	-0.3989	-0.2505
	I_3	-0.1320	-0.1707	-0.1309	-0.1936	0.5072	0.0233
G_3	I_1	0.0259	0.3964	-0.1812	-0.1919	0.2234	0.6425
	I_2	0.5733	0.6568	-0.2320	-0.3063	0.1277	0.2377
	I_3	-0.2166	0.4036	-.1188	0.1274	0.5316	0.3752
G_4	I_1	0.1392	0.8250	0.4073	0.3030	0.2912	-0.1186
	I_2	0.1154	0.8706	0.7883	0.3424	0.2900	-0.0545
	I_3	0.3051	0.1396	-0.3108	-0.1698	0.2332	0.1274

BIBLIOGRAPHY

[1] Barnett, V.D., The behaviour of pseudo-random sequences generated on computers by the multiplicative congruential method, Math. Comp. 16 (1962), pp. 63-69.

[2] Bliss, G.A., Lecture on the Calculus of Variations, Phoenix Science Series, The University of Chicago Press (Chicago and London, 1963).

[3] Bofinger, E. and V.J. Bofinger, On a periodic property of pseudo-random sequences, J. Assoc. Comp. Math., 5 (1958), pp. 261-265.

[4] Certaine, J., On sequences of pseudo-random numbers of maximal length, J.Assoc. Comp. Math., 5 (1958), pp. 353-356,

[5] Coveyou, R., Serial correlation in the generation of pseudo-random numbers, J. Assoc. Comp. Math., 7 (1960), pp. 72-74.

[6] Downham, D.Y., The runs up and down test, Comp. J., Vol. 12, No. 4 (1969), pp. 373-376.

[7] Downham, D.Y. and F.D.K. Roberts, Multiplicative congruential random number generators, Comp. J., 10 (1967), pp. 74-77.

[8] Dudley, V., Elementary Number Theory, Freeman and Company, (San Francisco, 1969).

[9] Edmunds, A.R., The generation of pseudo-random numbers on electronic digital computers, Comp. J., 2 (1960), pp. 181-185.

[10] Feller, W.E., An Introduction to Probability Theory and its Applications, Vol. 1, Third Edition, John Wiley and Sons Inc., (New York, 1950).

[11] Ferguson, G.A., Statistical Analysis in Psychology and Education, McGraw-Hill, (New York, 1966).

[12] Forsythe, G.E., Generation and testing of random digits at the National Bureau of Standards, Los Angeles, Monte Carlo Method, Nat. Bur. Stand., Appl. Math. Series 12 (1951), pp. 34-35.

[13] Good, I.J., The serial test for sampling numbers and other tests for randomness, Proc. Camb. Phil. Soc. 49 (1953), pp. 276-284.

[14] Greenberger, M., Notes on a new pseudo-random number generator, J. Assoc. Comp. Assoc., 8 (1961), pp. 163-167.

[15] Horton, H.B., A method for obtaining random numbers, Annals. of Math. Stat., 19 (1948), pp. 81-85.

[16] Hull, T.E. and A.R. Dobell, Mixed congruential random number generators for binary machines, J. Assoc. Comp. Math., 11 (1969), pp. 31-40.

[17] Hull, T.E. and A.R. Dobell, Random number generators, SIAM Review, 4 (1962), pp. 230-254.

[18] Isida, M. and H. Ikeda, Random number generators, Ann. Inst. Statist. Math., 8 (Tokyo, 1956), pp. 119-126.

[19] Jansson, B., Random Number Generators, Victor Pattersons Bokindustri Aktiekolog, (Stockholm, 1966).

[20] Keeping, E.S., Introduction to Statistical Inference, D. Van Nostrand Company, Inc., (Princeton, 1962).

[21] Kendall, M.G. and B. Babington Smith, Randomness and random sampling numbers, J. Roy. Stat. Soc., 101 (1938), pp. 147-166.

[22] Kendall, M.G. and B. Babington Smith, Second paper on random sampling numbers, J. Roy. Stat. Soc., (Supplement) 6 (1939), pp. 51-61.

[23] Kendall, M.G. and B. Babington Smith, Tables of random sampling numbers, Tracts for Computers, No. XXIV (Cambridge, 1960).

[24] Kendall, M.G. and A. Stuart, The Advanced Theory of Statistics, Charles Griffin & Co. Ltd., (London, 1967).

[25] Kolmogorov, A., Sulla determinazione empirica di una legge di distribuzione, G. Ist. Ital. Attuari, 4 (1933). p. 83.

[26] Lal, M. Expansion of $\sqrt{2}$ to 19,600 decimals, reviewed in Math. Comp., 21 (1967), pp. 258-259.

[27] Lal, M., First 39,000 decimal digits of $\sqrt{2}$, reviewed in Math. Comp., 22 (1968), p. 226.

[28] Lal, M. and W.F. Lunnen, Expansion of $\sqrt{2}$ to 100,000 decimals, reviewed in Math. Comp., 22 (1968), pp. 899-900.

[29] Lehmer, D.H., Mathematical methods in large-scale computing units, Annals. Comp. Laboratory Harvard Univ., 26 (1951), pp. 141-146.

[30] MacLaren, M.D. and G. Marsaglia, Uniform random number generators, J. Assoc. Comp. Math., 12 (1965), pp. 83-89.

[31] Marshall, A.W., An introductory note, Symposium on Monte Carlo Methods, ed. Herbert A. Mayer, John Wiley and Sons Inc., (New York, 1956), pp. 1-14.

[32] Meyer, H.A., L.S. Gephart, and N.L. Rasmussen, On the generation and testing of random digits, Air Res. Div. Command, WADC, Tech. Rep. 54-55, Wright-Patterson Air Force Base, (Ohio, 1954).

[33] Newman, T.G. and P.L. Odell, The Generation of Random Variates. Charles Griffin & Company Ltd., (London, 1971).

[34] Pathria, R.K., A statistical study of randomness among the first 10,000 digits of π, Math. Comp., 16 (1962), pp. 188-197.

[35] Peach P., Bias in pseudo-random numbers, J. Amer. Stat. Assoc., 56 (1961), pp. 610-618.

[36] Powell, M.J.D. and J. Swann, Weighted Uniform Sampling - a Monte Carlo technique for reducing variance, J. Inst. Maths. Applics., 2 (1966), pp. 228-236.

[37] Rand Corporation, A Million Random Digits with 100,000 Normal Deviates, Free Press, (Glencoe, 1955).

[38] Rotenberg, A., A new pseudo-random number generator, J. Assoc. Comp. Math., 7 (1960), pp. 75-77.

[39] Shanks, D. and J.W. Wrench Jr., Calculation of π to 100,000 decimals, Math. Comp., 16 (1962), pp. 76-79.

[40] Shreider, Yu. A. (ed.), Mathods of Statistical Testing (Monte Carlo Method), Elsevier Publishing Co., (Amsterdam, London, and New York, 1964).

[41] Stockmal, F., Calculations with pseudo-random numbers, J. Assoc. Comp. Math., 11 (1964), pp. 41-52.

[42] Takahashi, K. and M. Sibuya, Stastics of the digits of $\sqrt{n}$, Joho Shori (Information Processing) 6, (1965), pp. 221-223 (Japanese).

[43] Taussky, O. and J. Todd, Generation and testing of pseudo-random numbers, Symposium on Monte Carlo Methods, W.A. Meyer (ed.), John Wiley and Sons Inc,, (New York, 1956), pp. 15-28.

[44] Thompson, W.E., A modified congruence method of generating pseudo-random numbers. Comp. J., 1 (1958), pp. 83,86.

[45] Thompson, W.E., ERNIE - A mathematical and statistical analysis, J. Roy. Stat. Soc., A122 (1959), pp. 301-333.

[46] Tippett, L.H.C., Random sampling n-mbers, Tracts for Computers, No. XV, Cambridge University Press, (1969).

[47] Von Neumann, T., Various techniques used in connection with random digits, Monte Carlo Method, Nat. Bur. Stand., Appl. Math. Series, 12 (1951), pp. 36-38.

[48] Westlake, W.J., A uniform random number generator based on the combination of two congruential generators, J. Assoc. Comp. Math., 14 (1967), pp. 337-340.

[49] Yule, G.U., A test of Tippett's random sampling numbers, J. Roy. Statist. Soc., 101 (1938), pp. 167-172.

[50] Handscomb, D.C., Numer. Math., 6 (1964), pp. 261-268.