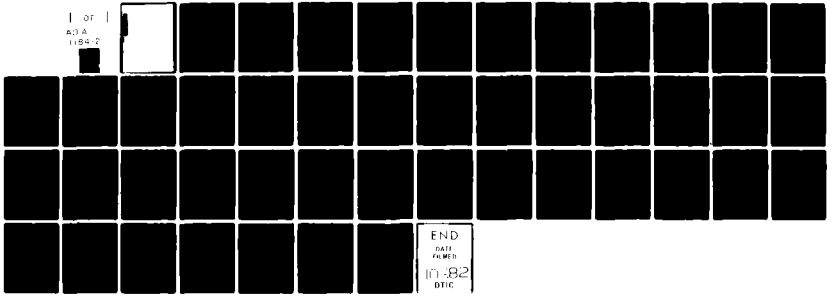
AD-A118 412 NAVAL SURFACE WEAPONS CENTER DAHLGREN VA F/G 12/1
RANDOM: A COMPUTER PROGRAM FOR EVALUATING PSEUDO-UNIFORM RANDOM--ETC(U)
AUG 82 J R CRIGLER, P A SHIELDS
UNCLASSIFIED NSWC-TR-82-93 NL



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NSWC TR 82-93	2. GOVT ACCESSION NO. AD-A118 418
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* RANDOM: A COMPUTER PROGRAM FOR EVALUATING PSEUDO-UNIFORM RANDOM NUMBER GENERATORS	5. TYPE OF REPORT & PERIOD COVERED Final
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) John R. Crigler Patricia A. Shields	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Surface Weapons Center (K106) Dahlgren, VA 22448	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Surface Weapons Center (K10) Dahlgren, VA 22448	12. REPORT DATE August 1982
	13. NUMBER OF PAGES 46
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*	15. SECURITY CLASS. *(of this report)* UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

RANDOM Computer Program
FORTRAN IV
CDC 6700 Computer
Pseudo-Uniform Random Number Generators

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The RANDOM computer program evaluates the usefulness of pseudo-uniform random number generators. RANDOM was written in FORTRAN IV for the CDC 6700 computer system at the Naval Surface Weapons Center, Dahlgren. This program can also be used to evaluate candidate pseudo-uniform random number generators designed for use on any other computer system. RANDOM subjects the candidate generator to 11 different statistical tests designed to detect departures from randomness. These tests serve as useful tools for determining the adequacy of a candidate generator.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

FOREWORD

The work documented in this technical report was performed at the Naval Surface Weapons Center (NSWC) by the Mathematical Statistics Staff (K106), Space and Surface Systems Division, Strategic Systems Department. The date of completion was March 1982.

This report was reviewed by Carlton W. Duke, Jr., Head, Space and Surface Systems Division.

Released by:

OLIVER F. BRAXTON, Head
Strategic Systems Department

CONTENTS

INTRODUCTION

The property of randomness, or random behavior, is an essential element in many areas of scientific research and application. For example, the presence of random behavior is a key requirement in digital computer simulation studies. Such a simulation requires a mechanism for generating sequences of events in which each sequence obeys a specific probability law. The probability law most frequently encountered in simulation work assumes that events in the sequence are independent and identically distributed; for example, each event in the sequence might be assumed to follow a normal distribution with the same mean and variance while each event occurs purely by chance and is not related to the occurrence of any other event in the sequence.

It is desirable to have a generating mechanism that can produce random variates from many different probability distributions. Probability theory establishes the fact that variates can be generated from a wide variety of distributions, provided that a sequence of independent *uniform* random variates on the interval (0,1) can be generated. In a uniform (0,1) distribution, each possible number in the range zero to one is equally likely to occur. Thus, the need for an efficient algorithm for generating uniform variates cannot be underestimated.

Computer algorithms for generating random numbers produce sequences that are deterministic; i.e., each number is completely determined by its predecessor and, therefore, all numbers in the sequence are determined by the starting number. While such sequences are not truly random, they appear to be so. Since the actual relationship between one number and its successor has no physical significance in most applications, this nonrandom character is not really undesirable. Sequences of numbers generated deterministically are referred to as *pseudo-random.*

This report is concerned with recommended statistical methodology for evaluating candidate pseudo-uniform random number generators; specifically, those that produce sequences of real numbers $U_0, U_1, U_2, \ldots$ that behave as though each number is independently selected at random from the uniform distribution with range zero to one. The symbol U(0,1) will be used to denote a continuous uniform random variable that takes on values between zero and one.

There are several ways to produce a sequence of numbers on a digital computer that looks like a sequence of U(0,1) random numbers. The most widely used method involves generating a sequence of integers $X_0, X_1, X_2, \ldots, X_n$ by means of a *linear congruential generator* of the form

$$X_{n+1} = (aX_n + c) \bmod m, \qquad n \geq 0$$

where $X_0 \geq 0$ represents the starting value, $a \geq 0$ is referred to as the multiplier, $c \geq 0$ is called the increment, and m is the modulus ($m > X_0$, $m > a$, $m > c$). A corresponding sequence of real numbers is then formed via the relationship $U_n = X_n/m$.

Ultimately, all congruential sequences produce a cycle of numbers that is repeated endlessly. This repeating cycle is referred to as the period of the sequence, and the length of the period can never exceed m. Since such sequences should have relatively long periods in order to be useful, the numbers X_0, a, c, and m must be properly chosen. The terms *multiplicative congruential generator* and *mixed congruential generator* are commonly used to refer to linear congruential generators with $c = 0$ and $c \neq 0$, respectively. For a detailed discussion of the construction of good linear congruential generators as well as a description of other methods for generating U(0,1) random numbers on digital computers see Knuth (1969).

Thus, any candidate uniform random number generator should be carefully examined to ensure that the numbers it produces are adequate for the desired experimental purposes. Of prime importance is that the candidate generator pass a collection of statistical tests designed to expose departures from independence and uniformity. The failure of a generator to possess these properties can produce severely misleading results in simulation studies. Fishman (1973) points out that it is also desirable for the candidate generator to be dense and efficient: a dense generator contains enough digits so that there are no wide gaps between assumable values on the unit interval; an efficient generator produces random numbers quickly and utilizes minimal storage in the computer. It should be emphasized that random number generators cannot be adequately evaluated in theory. Instead, one must generate a set of pseudo-random numbers from the candidate generator and perform statistical tests on them.

In the near future, NSWC will begin installing a new general-purpose computer system, which will include a machine-dependent pseudo-uniform random number generator. Since the use of random number generators is widespread among scientists and researchers in simulation and analysis studies at NSWC, the Mathematical Statistics Staff (K106) felt that it was important to develop a computer program that could be used to subject this generator to a battery of statistical "tests of randomness." The results of these tests could then be used to judge the adequacy of the generator for producing sequences of random variates that give the appearance of coming from the U(0,1) distribution. In addition, this program could also be used to test the adequacy of other candidate pseudo-uniform random number generators designed for use on this or any other computer system. The identification of a "bad" generator would result in its being rejected for use. One or more new generators would then be constructed and similarly tested until a "good" generator was found. Clearly, the early detection of "bad" pseudo-uniform random number generators is highly desirable.

The program written to meet the above requirements, RANDOM, is programmed in FORTRAN IV for the CDC 6700 computer system at NSWC. RANDOM performs 11 different statistical tests of randomness on a single sequence of 10,000 pseudo-uniform random numbers produced by the candidate generator (Appendices A and B present an input guide and sample output, respectively, for RANDOM). These tests are referred to as statistical *tests of hypothesis* and are designed to reveal departures from randomness. A statistical hypothesis

is a statement to be tested that is either accepted or rejected at a prescribed *level of significance*, which was chosen to be five percent for each of these tests.

For an elementary description of the theory of statistical hypothesis testing, the reader is referred to Walpole and Myers (1978). The number of empirical tests of randomness chosen for inclusion in RANDOM is by no means exhaustive. Rather, an attempt was made to include those tests that have proven most useful in characterizing randomness. These 11 tests are described in detail in the next section.

Much has appeared in the literature concerning random number generators. The interested reader is referred to Hull and Dobell (1962), Jansson (1966), and MacLaren and Marsaglia (1965), in addition to the previously mentioned sources.

TESTS PERFORMED IN PROGRAM RANDOM

This section is devoted to a detailed description of each of the 11 statistical tests of hypothesis employed in program RANDOM. The order of discussion here is identical to the order in which these tests are executed in RANDOM. Each test is applied to the same input sequence of 10,000 real numbers U_0, U_1, U_2, ... , U_{9999}, which purports to be uniformly distributed between zero and one. For additional details and historical remarks on most of these tests see Knuth (1969).

MEAN AND VARIANCE TESTS

This test (actually two separate tests that "belong" together) is more appropriately classified as a test on moments of a distribution vice a test on randomness. The U(0,1) distribution has a mean of 0.5 and a variance of 1/12 (equivalently, a standard deviation of 0.2887). The test on the mean is designed to determine whether or not the sample average of the 10,000 pseudo-uniform variates properly approximates the hypothesized mean of 0.5.

The sample mean is approximately distributed as a normal random variable with mean 0.5 and variance $(1/12)/10{,}000 = 8.3\overline{3} \times 10^{-6}$. Hence, for a test at the five-percent level of significance, the hypothesis that the true mean is 0.5 is rejected if the sample mean lies outside the interval $0.5 \pm 1.96 \times (8.3\overline{3} \times 10^{-6})^{1/2}$, or (0.4943, 0.5057).

Similarly, the test on the variance (actually on the standard deviation) is designed to determine whether or not the sample standard deviation properly approximates the hypothesized standard deviation of 0.2887. Hald (1952) shows that the sample standard deviation is approximately normally distributed with mean 0.2887 and variance $(1/12)/20{,}000 = 4.16\overline{6} \times 10^{-6}$ in this case. Thus, at the five-percent level of significance, the hypothesis

that the true standard deviation is 0.2887 is rejected if the sample standard deviation lies outside the interval $0.2887 \pm 1.96 \times (4.166\overline{6} \times 10^{-6})^{1/2}$, or (0.2847, 0.2927).

FREQUENCY TEST

To determine whether or not the input sequence of N = 10,000 real numbers consists of numbers that are, in fact, uniformly distributed between zero and one, divide the theoretical range of these numbers into 100 categories each of length 0.01. Let

p_i = probability that an observation falls into category i

E_i = expected number of observations in category i

O_i = number of observations that actually fall into category i

In this case, $p_i = 0.01$ and $E_i = Np_i = 100$ for all i. Then the statistic

$$X^2 = \sum_{i=1}^{100} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{100} \frac{(O_i - 100)^2}{100}$$

is approximately distributed as a chi-square random variable with 99 degrees of freedom. It is this statistic upon which the frequency (or equidistribution) test is based.

Note that the statistic X^2 is always either positive or zero. If the observed frequencies are close to the expected frequencies, the value of X^2 will be small, while observed frequencies that are not close to the expected frequencies will produce large values of X^2. Small values of X^2 support the hypothesis of uniformity, while large values of X^2 lead to rejection of the hypothesis. At the five-percent level of significance, the critical value for the test is found to be 123.2253. Thus, if the computed value of X^2 equals or exceeds this value, the hypothesis that the input sequence contains numbers that are uniformly distributed between zero and one is rejected.

In addition to performing the above test, program RANDOM displays a frequency table for the 100 categories and produces a graph of the cumulative frequency distribution, both based on the input sequence of 10,000 numbers.

KOLMOGOROV-SMIRNOV (K-S) TEST

Let X be a continuous random variable. Then, the cumulative distribution function (c.d.f.) of X, denoted by F(x), is defined by

$$F(x) = P(X \leq x) \qquad \text{for all } x.$$

Given a sample of N independent observations of X, the empirical c.d.f. $F_N(x)$ is defined as

$$F_N(x) = \frac{\text{number of observations} \leq x}{N} .$$

$F_N(x)$ will generally differ from F(x). However, if $F_N(x)$ differs from the assumed c.d.f. F(x) by too large a margin, one would have reasonable grounds for rejecting the hypothesis that F(x) is, in fact, the correct population c.d.f. This reasoning is the basis for the K-S test.

The test statistic is based upon the maximum absolute deviation between the ordinates of the empirical c.d.f. and the hypothesized c.d.f. at common abscissa values and is given by

$$D_N = \max_x \; | F_N(x) - F(x) | .$$

If the value of D_N is too large (i.e., if it exceeds a chosen critical value), the hypothesis that the assumed F(x) is the true F(x) is rejected.

Of concern is the comparison of a sample c.d.f. based on N = 10,000 observations with an assumed U(0,1) c.d.f. The comparison is to be made using 10,000 different abscissa values. Using the same symbols defined for the frequency test with $p_i = 0.0001$ and $E_i = Np_i = 1$ for all i, the appropriate K-S test statistic becomes

$$D_N = \max_i \left| \frac{\sum_{j=1}^{i} O_j - (i \times E_i)}{N} \right|, \qquad i = 1, 2, \ldots, N$$

where N = 10,000. According to Harter (1980), the critical value for this test at the five-percent level of significance is 0.01356. Hence, if the computed value of $D_{10,000}$ equals or exceeds this value, the hypothesis that the input sequence represents a sequence of random variates drawn from a U(0,1) distribution is rejected.

MAXIMUM OF t TEST

Consider the input sequence of N real numbers $U_0, U_1, \ldots, U_{N-1}$ assumed to have come from a U(0,1) distribution. Define

$$V_j = \max (U_{tj}, U_{tj+1}, \ldots, U_{tj+t-1})$$

for $0 \leq j < n$ where $N = nt$. It is easy to see that V_j has c.d.f. given by

$$F(x) = x^t, \quad 0 \leq x \leq 1 \tag{1}$$

since

$$P[\max(U_1, U_2, \ldots, U_t) \leq x]$$

$$= P[U_1 \leq x, U_2 \leq x, \ldots, U_t \leq x]$$

$$= x \cdot x \cdot \ldots \cdot x = x^t.$$

The K-S test can then be applied to the sequence $V_0, V_1, \ldots, V_{n-1}$ by evaluating

$$D_n = \max_x | F_n(x) - F(x) | .$$

If D_n exceeds the prechosen critical value, the hypothesis that the true c.d.f. is given by Equation 1 is rejected. This, in turn, implies that the input sequence $U_0, U_1, \ldots, U_{N-1}$ does not represent a sequence of random variates drawn from a U(0,1) distribution.

To apply the maximum of t test to the input sequence of N = 10,000 observations, n = 100 and t = 100 were selected so that Equation 1 becomes

$$F(x) = x^{100}, \quad 0 \leq x \leq 1 .$$

Next, the V's are ordered from smallest to largest; i.e.,

$$V'_0, V'_1, \ldots, V'_{99}$$

where $V'_0 \leq V'_1 \leq \ldots \leq V'_{99}$. The equation

$$x^{100} = c, \quad 0 \leq c \leq 1$$

is solved for x, yielding

$$x = c^{1/100}. \tag{2}$$

Equation 2 is then evaluated at n = 100 values of c; i.e., c = 0.01, 0.02, ... , 1.00. For example, for c = 0.01, $x \cong 0.9550$. Then, for each x the sequence $V'_0, V'_1, \ldots, V'_{99}$ is used to determine the value of the sample c.d.f.

$$F_{100}(x) = \frac{\text{number of } V' \text{ values} \leq x}{100}$$

These values are then used in computing the K-S test statistic

$$D_{100} = \max_{x} | F_{100}(x) - F(x) | .$$

The critical region for a test at the five-percent level of significance is $D_{100} \geq 0.13403$ (Owen, 1962). Hence, $F_{100}(x)$ represents the sample c.d.f. formed by the maximum values from consecutive blocks of t = 100 numbers and D_{100} is the maximum absolute deviation between this sample c.d.f. and the corresponding theoretical c.d.f. Therefore, the hypothesis to be tested is that this maximum deviation is not significantly different from the maximum deviation obtained had the input sequence come from a U(0,1) distribution. If D_{100} exceeds the above critical value, this hypothesis is rejected.

GAP TEST

Given the input sequence of N = 10,000 real numbers $U_0, U_1, \ldots, U_{9999}$, this test examines the lengths of "gaps" between occurrences of U_j in some prespecified range. A chi-square test statistic, akin to the one used in the frequency test, is then used to determine whether or not the observed numbers of gaps of each length are sufficiently close to their corresponding expected numbers, which would support the hypothesis that the input sequence represents a sequence of random variates coming from a U(0,1) distribution.

Choose two real numbers α and β such that $0 \leq \alpha < \beta \leq 1$. Consider the input sequence $U_0, U_1, \ldots, U_{N-1}$ to be a cyclic sequence with U_{N+j} identified with U_j. Next consider the lengths of consecutive subsequences $U_j, U_{j+1}, \ldots, U_{j+r}$ in which U_{j+r} lies between α and β but the other U values do not. Such a subsequence defines a *gap of length r*. If n of the N numbers $U_0, U_1, \ldots, U_{N-1}$ fall into the range $\alpha \leq U_j < \beta$, then there are n gaps in the cyclic sequence. Let

Z_r = number of gaps of length r, $0 \leq r < t$

Z_t = number of gaps of length t or greater

p = probability that $\alpha \leq U_j < \beta$.

Then, $p = \beta - \alpha$. Furthermore, let

p_r = probability of observing a gap of length r $(0 \leq r < t)$

p_t = probability of observing a gap of length t or greater.

Then,

$$p_r = \begin{cases} p(1-p)^r, & 0 \leq r \leq t-1 \\ (1-p)^t, & r = t \end{cases}$$

It can be shown that the test statistic

$$X^2 = \sum_{r=0}^{t} \frac{(Z_r - np_r)^2}{np_r}$$

is approximately distributed as a chi-square random variable with t degrees of freedom.

In program RANDOM, the values of α, β and t have been chosen to be

$$\alpha = 0.30$$

$$\beta = 0.60$$

$$t = 8$$

The critical value for the chi-square test at the five-percent level of significance is then found to be 15.5073. The hypothesis to be tested is that the observed number of gaps of each length is not significantly different from the number expected if the input sequence had come from a U(0,1) distribution. If the computed value of X^2 equals or exceeds the above critical value, this hypothesis is rejected.

POKER TEST

Consider subdividing the input sequence of N = 10,000 real numbers into k = 2000 groups of five successive numbers, $(U_{5j}, U_{5j+1}, \ldots, U_{5j+4})$, $0 \leq j < k$. Then, convert the U_i into the integers Y_i according to the following scheme:

$$Y_i = \begin{cases} 1 & \text{if } U_i \in [0, 0.2] \\ 2 & \text{if } U_i \in (0.2, 0.4] \\ 3 & \text{if } U_i \in (0.4, 0.6] \\ 4 & \text{if } U_i \in (0.6, 0.8] \\ 5 & \text{if } U_i \in (0.8, 1.0] \end{cases}.$$

Each quintuple is then classified into one of five disjoint categories based on the number of distinct values in the set of five integers. Each quintuple may be thought of as representing a "poker hand". The five categories are

five different	=	all different
four different	=	one pair
three different	=	two pairs, or three of a kind
two different	=	full house, or four of a kind
one different	=	five of a kind.

A chi-square test based on the observed and expected number of quintuples in each category can now be used, provided that an expression for the probability that a quintuple contains m different values can be formulated. Let p_m represent this probability. In general, consider k groups of n successive integers in which the integers may range from 1 to d, inclusive. The probability p_m can be formulated as the ratio

$$\frac{\text{number of n-tuples with exactly m different integers}}{\text{total number of n-tuples}}$$

from m = 1, 2, ... , d. The denominator of this ratio is d^n. The numerator is the product of the number of ways to partition a set of n elements into exactly m nonempty disjoint subsets, denoted by $S_n^{(m)}$, and the number of permutations of m things from a set of d objects, namely d (d - 1) ••• (d - m + 1). The required probability then becomes

$$p_m = \frac{d\,(d-1)\cdots(d-m+1)}{d_n}\, S_n^{(m)} .$$

The notation $S_n^{(m)}$ is used here to denote Stirling numbers of the second kind. (Tables of $S_n^{(m)}$ may be found in Abramowitz and Stegun, 1964). In this application, d = 5, n = 5, and m = 1, 2, 3, 4, 5. The required Stirling numbers are

$$S_5^{(1)} = 1$$

$$S_5^{(2)} = 15$$

$$S_5^{(3)} = 25$$

$$S_5^{(4)} = 10$$

$$S_5^{(5)} = 1.$$

Thus, the probabilities that a quintuple contains m = 1, 2, 3, 4, 5 different values are

$$p_1 = 0.0016$$
$$p_2 = 0.0960$$
$$p_3 = 0.4800$$
$$p_4 = 0.3840$$
$$p_5 = 0.0384$$

Note that $\sum_{m=1}^{5} p_m = 1$, as required.

Letting

E_i = expected number of quintuples in category i

O_i = observed number of quintuples in category i

for i = 1, 2, ... , t, the test statistic

$$X^2 = \sum_{i=1}^{t} \frac{(O_i - E_i)^2}{E_i}$$

is approximately distributed as a chi-square random variable with t - 1 degrees of freedom. Since there are k = 2000 quintuples in our application,

$$E_i = kp_i = 2000\, p_i, \quad i = 1, 2, 3, 4, 5 .$$

The above chi-square test statistic is employed with t = 5 categories and t - 1 = 4 degrees of freedom. The critical value for this chi-square test at the five-percent level of significance is 9.4877. The hypothesis being tested is that the number of poker hands of each type does not differ significantly from the expected number of hands of each type obtained when the input sequence does, in fact, come from a U(0,1) distribution. If X^2 equals or exceeds the above critical value, this hypothesis is rejected.

COUPON COLLECTOR'S TEST

This test is related to the poker test in much the same way as the gap test is related to the frequency test. The input sequence of N = 10,000 real numbers $U_0, U_1, \ldots, U_{9999}$ is converted into the sequence of integers $Y_0, Y_1, \ldots, Y_{9999}$ according to the following scheme:

$$Y_i = \begin{cases} 1 & \text{if } U_i \in [0, 0.2] \\ 2 & \text{if } U_i \in (0.2, 0.4] \\ 3 & \text{if } U_i \in (0.4, 0.6] \\ 4 & \text{if } U_i \in (0.6, 0.8] \\ 5 & \text{if } U_i \in (0.8, 1.0] \end{cases} .$$

These integers represent different "coupons" to be collected. A "coupon collector sequence" is a subsequence of $Y_0, Y_1, \ldots, Y_{9999}$ of the shortest length required to contain each of the integers one through five at least once. Begin by observing the length of the sequence $Y_0, Y_1, \ldots$ required to obtain a "complete set" of the integers one to five. If this length is denoted by r, then the first coupon collector sequence is $Y_0, Y_1, \ldots, Y_{r-1}$. This process is then repeated starting with $Y_r, Y_{r+1}, \ldots$ to obtain additional coupon collector sequences until the entire input sequence is exhausted.

In general, consider the lengths of coupon collector sequences that contain the integers one through d. Consider compiling the frequencies of occurrence of the lengths of these sequences. Choose an integer $t > d$ such that all sequences whose lengths are greater than or equal to t are considered to be in the same category. Then, each sequence length can be classified into one of $t - d + 1$ distinct categories.

A chi-square test based on the observed and expected number of coupon collector sequences in each category can now be developed. Expressions are needed for (1) p_r, the probability that a sequence containing at least one of each of the integers 1, 2, ... , d is of length r, and (2) p_t, the probability that a sequence containing at least one of each of the integers 1, 2, ... , d is of length t or greater. The derivations of p_r and p_t that follow reflect the fact that the shortest sequence that contains at least one of each of the integers one through d in each case is desired. Following the development in the poker test, the expression

$$\frac{d!}{d^r} S_r^{(d)}$$

represents the probability that an r-tuple contains exactly d different values, where $d! = d(d-1)(d-2)\cdots 1$ and $S_r^{(d)}$ denotes a Stirling number of the second kind, as before. Hence,

$$q_r = 1 - \frac{d!}{d^r} S_r^{(d)}$$

represents the probability that an r-tuple is "incomplete" (i.e., it does not contain all d different integers). It is clear, then, that

$$p_t = q_{t-1} = 1 - \frac{d!}{d^{t-1}} S_{t-1}^{(d)} .$$

Also, for $d \leq r < t$,

$$\begin{aligned} p_r &= q_{r-1} - q_r \\ &= \left[1 - \frac{d!}{d^{r-1}} S_{r-1}^{(d)}\right] - \left[1 - \frac{d!}{d^r} S_r^{(d)}\right] \\ &= \frac{d!}{d^r} S_r^{(d)} - \frac{d!}{d^{r-1}} S_{t-1}^{(d)} \\ &= \frac{d!}{d^r} [S_r^{(d)} - d\, S_{r-1}^{(d)}] \quad . \end{aligned}$$

From an addition identity involving Stirling numbers of the second kind,

$$S_{r-1}^{(d-1)} = S_r^{(d)} - d\, S_{r-1}^{(d)} \quad .$$

Hence, the required probability can be expressed as

$$p_r = \frac{d!}{d^r} S_{r-1}^{(d-1)}, \quad d \leq r < t \quad .$$

In programming this test for inclusion in RANDOM, d = 5 and t = 15 were chosen. These choices give rise to the following probabilities:

$$\begin{aligned} p_5 &= 0.038400000 \\ p_6 &= 0.076800000 \\ p_7 &= 0.099840000 \\ p_8 &= 0.107520000 \\ p_9 &= 0.104509440 \\ p_{10} &= 0.095477760 \\ p_{11} &= 0.083816448 \\ p_{12} &= 0.071639040 \\ p_{13} &= 0.060112994 \\ p_{14} &= 0.049791565 \\ p_{15} &= 0.212092753 \end{aligned}$$

Note that $\sum_{r=5}^{15} p_r = 1$ as required.

Let n represent the total number of coupon collector sequences observed in the sequence $Y_0, Y_1, \ldots, Y_{9999}$. Further, let

Z_r = observed number of coupon collector sequences of length r, $d \leq r < t$

and

Z_t = observed number of coupon collector sequences of length t or greater.

Then, the test statistic

$$X^2 = \sum_{r=d}^{t} \frac{(Z_r - np_r)^2}{np_r}$$

is approximately distributed as a chi-square random variable with t - d degrees of freedom. This test statistic is applied here with t - d + 1 = 11 categories and t - d = 10 degrees of freedom. The critical value for this chi-square test at the five-percent level of significance is 18.3070. The hypothesis to be tested is that the observed number of coupon collector sequences of each length does not differ significantly from the expected number obtained when the input sequence does, in fact, come from a U(0,1) distribution. Hence, if the computed value of X^2 equals or exceeds the above critical value, this hypothesis is rejected.

It is of interest to note that when it is assumed that the input sequence is random, the number of successive Y_i's that need to be examined, on the average, before a complete set of "coupons" has been found is $11.4166\overline{6}$. Hence, with N = 10,000 numbers, one would expect to observe 876 coupon collector sequences on the average if one applied the coupon collector's test repeatedly a large number of times, each time using a different input sequence of N = 10,000 real numbers $U_0, U_1, \ldots, U_{9999}$.

PERMUTATION TEST

In this test, the input sequence of real numbers is subdivided into k groups of n elements each, i.e., $(U_{jn}, U_{jn+1}, \ldots, U_{jn+n-1})$, $0 \leq j < k$. Assuming that equality between U's does not occur, the elements in each group can have n! possible relative orderings or permutations. Since the probability of occurrence of each of these orderings is $p_i = 1/n!$, i = 1, 2, ... , n!, a chi-square test with t = n! categories can be applied to the observed and expected number of n-tuples of each type.

In programming this test for inclusion in RANDOM, n = 3 was chosen. Thus, k = 3333 triples comprising the first 9999 numbers in the input sequence were observed. Let A, B, and C represent the elements in a triple where A is the smallest value, B the middle value, and C the largest value. Then, the 3! = 6 possible orderings are (A, B, C), (A, C, B), (B, A, C), (B, C, A),

(C, A, B), and (C, B, A). An algorithm was used to count the number of times each of these triple types was observed in the input sequence. These observed frequencies were then used in evaluating the chi-square test statistic with t = 6 categories.

If

E_i = expected number of triples in category i

O_i = observed number of triples in category i

for i = 1, 2, ... , t, then the test statistic

$$X^2 = \sum_{i=1}^{t} \frac{(O_i - E_i)^2}{E_i}$$

is approximately distributed as a chi-square random variable with t - 1 degrees of freedom. Note that in this application,

$$E_i = kp_i = (3333)(1/6) = 555.50$$

for all i. The critical value for the above chi-square test with t - 1 = 5 degrees of freedom at the five-percent level of significance is 11.0705. The hypothesis under consideration is that the observed number of permutations of each type is not significantly different from the expected number obtained when, in fact, the input sequence comes from a U(0,1) distribution. The hypothesis is rejected if the computed value of X^2 equals or exceeds the above critical value.

RUNS TEST

A method for testing the input sequence of real numbers U_0, U_1, ... , U_{9999} for "runs up" and "runs down" is presented. A *run* is defined as an unbroken sequence of observations in which all of the numbers are either increasing or decreasing. Consider the subsequence

$$U_i > U_{i+1} < U_{i+2} < \ldots < U_{i+p} > U_{i+p+1} \quad .$$

The numbers U_{i+1} through U_{i+p} define a "run up" of length p. Similarly, the numbers U_{i+1} through U_{i+p} would define a "run down" of length p if the direction of each of the inequality signs above were reversed. If the input sequence is representative of a sequence of random numbers drawn from a U(0,1) distribution, then neither the total number of runs up nor the total number of runs down should be excessively high or excessively low. Moreover, the frequency with which various lengths of runs up and runs down occur requires careful examination. The observed numbers of runs up and runs down of a particular length should not

differ greatly from their corresponding expected numbers. In the ensuing paragraphs, tests on both the number of runs and the lengths of runs will be discussed.

Define

$$R'_p = \text{number of runs of length} \geq p$$

and

$$R_p = \text{number of runs of length } p \text{ exactly}$$

$$= R'_p - R'_{p+1} \quad .$$

Expressions for the means of R_p and R'_p and the covariances between R_p and R_q, R'_p and R'_q, and R_p and R'_q are given in Knuth (1969). These expressions are functions of the length of the input sequence, N. The covariance between R_p and R_q measures the interdependence between these two variables. If $p = q$, the covariance between R_p and R_q is equivalent to the variance of R_p. Wolfowitz (1944) has shown that $R_1, R_2, \ldots, R_{t-1}, R'_t$ become normally distributed as $N \to \infty$, with means and covariances given by the aforementioned expressions. These results are sufficient to permit the development of tests on both the number and lengths of runs.

Let

$$R = \text{total number of runs}$$

$$= \sum_{i=1}^{t-1} R_i + R'_t \quad .$$

Then, the random variable

$$Z_R = \frac{R - E(R)}{[\text{Var}(R)]^{1/2}} \quad ,$$

has a standard normal distribution (i.e., mean zero and variance one). Here, E(R) and Var(R) denote the mean and variance of the random variable R, respectively.

The hypothesis to be tested is that the observed number of runs up is not significantly different from the expected number of runs up obtained when the input sequence comes from a U(0,1) distribution. Hence, to perform a test at the five-percent level of significance on the total number of runs up, compute Z_R and compare it to the interval (-1.960, 1.960). If Z_R falls outside this interval, the hypothesis is rejected. This same test is applied to the observed number of runs down.

An algorithm that counts both runs up and runs down is employed in RANDOM. The value of t was chosen to be six in this application. With N = 10,000, E(R) = 5000.50 and Var(R) = 833.4166 for both runs up and runs down.

In developing a test on the lengths of runs up (or runs down), we note that the usual chi-square test is not applicable, since adjacent runs are not independent. The following procedure circumvents this difficulty. Let

$$Q_i = \begin{cases} R_i - E(R_i), & i = 1,2, \ldots, t-1 \\ R_i' - E(R_i'), & i = t \end{cases}$$

Let $C = (c_{ij})$ denote the t x t matrix of covariances between the random variables $R_1, R_2, \ldots, R_{t-1}$, and R_t'; i.e., c_{14} is the covariance between R_1 and R_4, while c_{1t} is the covariance between R_1 and R_t'. Now, let the t x t matrix $A = (a_{ij})$ denote the inverse of C. Then, the test statistic

$$X^2 = \sum_{1 \le i,\, j \le t} Q_i\, Q_j\, a_{ij}$$

is distributed as a chi-square random variable with t degrees of freedom when N is large. Again, t = 6 was chosen in this application. Hence, the number of runs up and runs down of lengths one through five and six or longer are counted in program RANDOM and the same test is applied to both runs up and runs down. The critical value for this chi-square test with six degrees of freedom at the five-percent level of significance is 12.5916. The hypothesis to be tested is that the observed number of runs up (down) of each length is not significantly different from the expected number of runs up (down) observed when the input sequence comes from the U(0,1) distribution. If the computed value of X^2 equals or exceeds the above critical value, this hypothesis is rejected.

It is interesting to note that both the permutation and runs tests do not depend on the U's being uniformly distributed; they require only that the probability that $U_i = U_j$ is zero for $i \neq j$. Hence, these tests can be applied to pseudo-random sequences other than those generated from a U(0,1) distribution.

SERIAL TEST FOR SUCCESSIVE PAIRS

This test is designed to determine whether or not successive pairs of numbers are uniformly and independently distributed. Begin by converting the input sequence of N = 10,000 real numbers $U_0, U_1, \ldots, U_{9999}$ into a sequence of integers $Y_0, Y_1, \ldots, Y_{9999}$ by multiplying each U_i by 10 and truncating the decimal. Hence, $0 \le Y_i \le 9$ for all i. Then, subdivide the sequence $Y_0, Y_1, \ldots, Y_{9999}$ into 5000 pairs of two successive integers and count the frequency with which the pair $(Y_{2j}, Y_{2j+1}) = (q, r)$ occurs for $0 \le j < 5\,000$ where

$0 \leq q, r < 10$. These 5000 observed pairs of integers are then used to fill a 10 x 10 frequency table in which the first integer of the pair determines the row number and the second integer determines the column number of the cell to which that pair belongs. Since the probability that a randomly selected integer pair will belong to a particular cell is 0.01 for all cells, a chi-square test based on the observed and expected number of pairs in each of the 100 categories can be used.

Let

E_{ij} = expected number of times the integer pair (i, j) occurred

O_{ij} = observed number of times the integer pair (i, j) occurred

for $0 \leq i, j \leq 9$. Then the test statistic

$$X^2 = \sum_{i=0}^{9} \sum_{j=0}^{9} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

is approximately distributed as a chi-square random variable with 99 degrees of freedom. Since 5000 integer pairs are to be assigned to 100 categories, $E_{ij} = 50$ for all i and j. The number of degrees of freedom used in this test is justified by the fact that the E_{ij} are not estimated from the observed frequencies (Freund, 1962). The critical value for this chi-square test at the five-percent level of significance is 123.2253. The hypothesis to be tested is that the observed distribution of successive pairs of numbers does not differ significantly from the expected distribution for an input sequence of random variates from a U(0,1) distribution. Hence, if the computed value of X^2 equals or exceeds the above critical values, this hypothesis is rejected.

SERIAL CORRELATION TEST

Consider the input sequence of real numbers $U_0, U_1, \ldots, U_{9999}$. To determine if the values in this sequence are related in any special way to the values h units apart for some h, the test for serial correlation, which computes a measure of the amount that U_{i+h} depends on U_i, is used. This measure of dependency is called the serial correlation of lag h and represents the correlation between pairs of equally spaced observations from a sample. Let N represent the number of elements in the sample to be considered. Serial correlation can be defined in one of two ways (Bennett and Franklin, 1954):

1. In the circular definition, $U_{i+h} = U_{i+h-N}$ is defined for $i + h \geq N$. Circular serial correlation is useful in detecting periodic effects in a sequence of observations.

2. In the noncircular definition, the pairs that depend upon use of the definition $U_{i+h} = U_{i+h-N}$ for $i + h \geq N$ are omitted and the serial correlation between the remaining pairs is computed. Noncircular serial correlation is useful in detecting trends in a series of observations.

The test for serial correlation employed in RANDOM is taken from Wald and Wolfowitz (1943) and is performed using both the circular and noncircular definitions. The theory behind this test requires that N be a prime number; hence, N was chosen to be 9973, the largest prime number less than or equal to 10,000. Serial correlations are usually computed only for small values of h in practice, since most of the important dependencies occur at those values. For this reason, only lags one through 10 were considered in RANDOM. Thus, the first 9973 numbers in the input sequence were used to compute the statistic

$$R'_h = \sum_i U_i U_{i+h}, \qquad h = 1, 2, \ldots, 10$$

where for $i + h \geq 9973$, $U_{i+h} = U_{i+h-9973}$ and $i = 0, 1, 2, \ldots, 9972$
(circular definition)

$i = 0, 1, 2, \ldots, 9972 - h$.
(noncircular definition)

The mean and variance of the random variable R'_h are given by

$$E(R'_h) = \frac{(S_1^2 - S_2)}{n - 1}$$

and

$$Var(R'_h) = \frac{S_2^2 - S_4}{n - 1} + \frac{S_1^4 - 4S_1^2S_2 + 4S_1S_3 + S_2^2 - 2S_4}{(n - 1)(n - 2)} - \frac{(S_1^2 - S_2)^2}{(n - 1)^2}$$

where $S_k = \sum_{i=0}^{9972} U_i^k$ is the kth power sum of the observations. It can be shown that R'_h approaches the normal distribution for large N. Hence, the random variable

$$Z_h = \frac{R'_h - E(R'_h)}{[Var(R'_h)]^{1/2}}$$

has a standard normal distribution (i.e., mean zero and variance one). To perform the tests for serial correlation, compute Z_h for all 10 values of h and for both the circular and noncircular forms. For tests at the five-percent level of significance, compare each computed value of Z_h to the interval (-1.960, 1.960). The hypothesis under consideration is that the serial correlation between observations seperated by lag h is not significantly different from the

correlation for an input sequence of random variates from the U(0,1) distribution. If any Z_h falls outside the above interval, the corresponding hypothesis is rejected.

It should be noted that the above test does not depend on the U's being uniformly distributed. The test is nonparametric in the sense that it assumes only that the U's represent a random sample from a distribution with continuous cumulative distribution function.

EVALUATING THE CANDIDATE RANDOM NUMBER GENERATOR

The previous section described 11 statistical tests of hypothesis designed to detect departures from randomness for the pseudo-uniform random number generator under consideration. This section will present some guidelines for interpreting the results of these tests; the case in which the candidate generator fails one or more tests is included.

Before proceeding with a discussion of the interpretation of test results, it would be beneficial to reconsider the testing process itself. In testing for randomness, one looks for behavior that is not actually present in one's sequence, since the process of pseudo-random number generation is deterministic; yet, we require that the numbers so produced exhibit the semblance of randomness (Overstreet, 1972). Thus, some degree of subjective judgment should accompany the evaluation of test results in order to reach a final decision to accept or reject the candidate generator.

A statistical test of an hypothesis that leads to its rejection is said to be "significant"; otherwise, it is "nonsignificant." Recall that each of the statistical tests of randomness is performed at the five-percent level of significance. For an individual test, this means that the probability is 0.05 that the test will incorrectly conclude that the candidate generator is not sufficiently random to be useful, when, in fact, the generator does produce sequences of numbers that exhibit the hypothesized characteristics of U(0,1) variates. Now, suppose that the "good" candidate generator was used to generate a large number of sequences of length N such that each sequence was obtained through the use of a different starting "seed" value. If each of these sequences were then subjected to the same statistical test of randomness, the test procedure would incorrectly conclude that approximately five-percent of these sequences were products of "bad" generators. In other words, even a "good" generator will fail any of these tests five percent of the time. It is in this sense that the rejection of the hypothesis of randomness is interpreted at the five-percent level of significance.

The collective interpretation of a set of tests for randomness is difficult. The need for making several tests is clear. Even "bad" generators will pass some of these tests, while failing others; hence, subjecting a "bad" generator to only a few tests may be insufficient to identify it as "bad." On the other hand, the use of numerous tests is not totally free from criticism,

since some of these tests are dependent, although the degree of dependence is either unknown or extremely difficult to assess. Hence, when interpreting the results of a series of tests, the significance levels should be viewed as a general indication rather than as a specific prediction.

Any statistical hypothesis testing procedure can result in one of two decision errors: rejection of a correct hypothesis, or acceptance of a false hypothesis. The size of each of these errors is measured in terms of its probability of occurrence. In this report, the size of the first error is controlled at 0.05, but no attempt has been made to assess the probability of accepting a false hypothesis. However, as N increases, this latter probability decreases. Hence, the large N value used in these tests ensures that the probability of accepting a false hypothesis will be reasonably small. Since a statistical test of hypothesis is not significant unless it results in a rejection, a test that does not detect lack of randomness does not imply that the sequence being tested is random. For this reason, the phrase "fail to reject the hypothesis" rather than "accept the hypothesis" is used when stating the conclusions. The advantage, then, of performing several tests of randomness on a sequence is that the feeling of confidence in using the proposed generator is increased if a relatively high number of nondetections is obtained. A reliable generator (i.e., one that produces numbers that are sufficiently random) is one that performs well when subjected to extensive testing.

One final comment regarding the testing of pseudo-uniform random number generators is in order. A candidate generator should not necessarily be discarded just because it fails one or two of the tests for randomness, since statistical testing permits failures for a *good* generator a small proportion of the time. If a failure is observed for a certain test, this test should be closely examined to see if the reasons for failure can be ascertained. Was the decision to reject the hypothesis a borderline one, or was the test highly significant? Was the failure merely a result of random variation, or was it the result of a serious deficiency of the generating scheme itself? It is recommended that program RANDOM be rerun one or more times using *different* input sequences, each generated by a different starting seed value, in the hopes of shedding some light on these questions.

Thus, no firm quantitative guidelines currently exist for deciding whether or not to accept a candidate generator. We note, however, that if we assume that the 11 tests in RANDOM are all independent, then the probability of obtaining one or more rejections in the 11 tests is about 0.43 if, in fact, all of the hypotheses are true! This observation can serve as a general guideline when deciding whether to accept or reject a candidate generator. The decision is still largely subjective and should be made only after careful examination of the test results.

BIBLIOGRAPHY

Abramowitz, M. and I. A. Stegun, *Handbook of Mathematical Functions*, Applied Mathematics Series 55, U.S. Department of Commerce (National Bureau of Standards, 1964).

Bennett, C. A. and N. L. Franklin, *Statistical Analysis in Chemistry and the Chemical Industry* (New York: John Wiley and Sons, Inc., 1954).

Fishman, G. S., *Concepts and Methods in Discrete Event Digital Simulation* (New York: John Wiley and Sons, Inc., 1973).

Freund, J. E., *Mathematical Statistics* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962).

Hald, A., *Statistical Theory with Engineering Applications* (New York: John Wiley and Sons, Inc., 1952).

Harter, H. L., "Modified Asymptotic Formulas for Critical Values of the Kolmogorov Test Statistic," *The American Statistician*, 34 (1980), No. 2, pp. 110 - 111.

Hull, T. E. and A. R. Dobell, "Random Number Generators," *SIAM Review*, 4 (1962), pp. 230 - 254.

Jansson, B., *Random Number Generators* (Stockholm: Almqvist and Wiksell, 1966).

Knuth, D. E., *The Art of Computer Programming: Volume 2/Seminumerical Algorithms* (Reading, Mass.: Addison-Wesley, 1969).

MacLaren, M. D. and G. Marsaglia, "Uniform Random Generators," *Journal of the Association of Computing Machinery*, 12 (1965), pp. 83 - 89.

Overstreet, C., Jr., *A FORTRAN V Package for Testing and Analysis of Pseudorandom Number Generators*, Technical Report CP - 72009, Computer Science/Operations Research Center, Institute of Technology, Southern Methodist University (Dallas, Texas, 1972).

Owen, D. B., *Handbook of Statistical Tables* (Reading, Mass.: Addison-Wesley, 1962).

Wald, A. and J. Wolfowitz, "An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation," *Annals of Mathematical Statistics*, 14 (1943), pp. 378 - 388.

Walpole, R. E. and R. H. Myers, *Probability and Statistics for Engineers and Scientists*, 2nd Edition (New York: Macmillan, 1978).

Wolfowitz, J., "Asymptotic Distribution of Runs Up and Down," *Annals of Mathematical Statistics*, 15 (1944), pp. 163 - 172.

APPENDIX A

INPUT GUIDE

The 10,000 data points used as input may be read in from punched cards or a permanent file. The following data card is required for either option:

CARD TYPE 1

Columns	Variable	Description	Format
1-5	IØP	IØP = 1: data follows on punched cards IØP ≠ 1: data attached via TAPE 5 with input format (E22.14)	I5
11-80	FØRM	Format of input data cards (used only if IØP = 1)	7A10

DECK SET UP

ATTACH, LGØ, RANDØM, ID = N1W.
ATTACH, SYSLIB.
LIBRARY, SYSLIB.
ATTACH, TAPE5, datafile. [used if IØP ≠ 1]
LGØ.
7/8/9
Card Type 1
.
.
.
.
data cards if IØP = 1
.
.
.
.
6/7/8/9

APPENDIX B

SAMPLE OUTPUT

The U(0,1) random number generator currently in use on the CDC 6700 computer system at NSWC is called RANF. This generator has been widely used in simulation and analysis studies at NSWC. Ten thousand U(0,1) variates were generated from RANF using the starting floating point seed value of 3571.0 and stored on a permanent file. These numbers were then processed by program RANDOM to illustrate the program's output. The sample printout for this case is shown in this appendix.

* * * TEST ON THE MEAN AND THE VARIANCE OF A UNIFORM (0,1) DISTRIBUTION * * *

MEAN = .4980082

VARIANCE = .0825229

STANDARD DEVIATION = .2873

RECIPROCAL OF VARIANCE = 12.1178415

* TEST ON THE MEAN *

LOWER BOUND = .4943
UPPER BOUND = .5057

FAIL TO REJECT THE HYPOTHESIS THAT THE TRUE MEAN = .5 AT THE 5 PERCENT LEVEL OF SIGNIFICANCE

* TEST ON THE VARIANCE *
(COMPARE BOUNDS TO STANDARD DEVIATION)

LOWER BOUND = .2847
UPPER BOUND = .2927

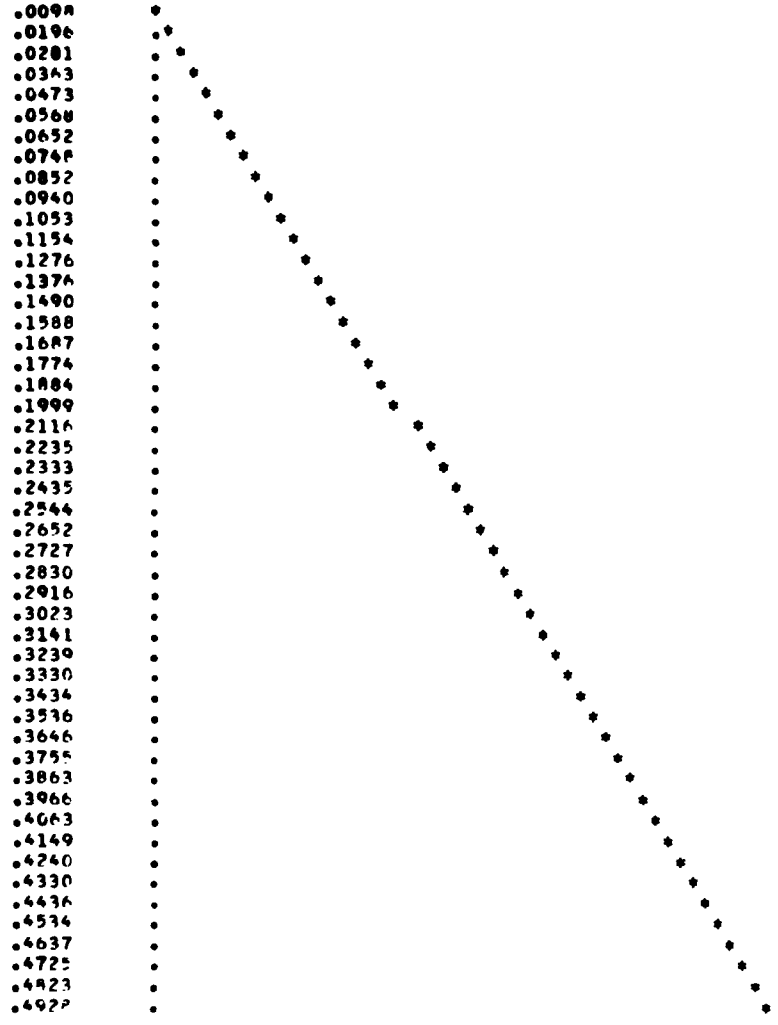
FAIL TO REJECT THE HYPOTHESIS THAT THE TRUE VARIANCE = 1/12 AT THE 5 PERCENT LEVEL OF SIGNIFICANCE

FREQUENCY DISTRIBUTION FOR
10000 UNIFORM (0,1) RANDOM NUMBERS

CATEGORY	RANGE	OBSERVED FREQUENCY	EXPECTED CUMULATIVE FREQUENCY	OBSERVED CUMULATIVE FREQUENCY
1	(0.00 - .01)	98	100	98
2	(.01 - .02)	98	200	196
3	(.02 - .03)	85	300	291
4	(.03 - .04)	82	400	363
5	(.04 - .05)	110	500	473
6	(.05 - .06)	95	600	568
7	(.06 - .07)	84	700	652
8	(.07 - .08)	96	800	748
9	(.08 - .09)	104	900	852
10	(.09 - .10)	88	1000	940
11	(.10 - .11)	113	1100	1053
12	(.11 - .12)	101	1200	1154
13	(.12 - .13)	122	1300	1276
14	(.13 - .14)	100	1400	1376
15	(.14 - .15)	114	1500	1490
16	(.15 - .16)	98	1600	1588
17	(.16 - .17)	99	1700	1687
18	(.17 - .18)	87	1800	1774
19	(.18 - .19)	110	1900	1884
20	(.19 - .20)	115	2000	1999
21	(.20 - .21)	117	2100	2116
22	(.21 - .22)	119	2200	2235
23	(.22 - .23)	98	2300	2333
24	(.23 - .24)	102	2400	2435
25	(.24 - .25)	109	2500	2544
26	(.25 - .26)	108	2600	2652
27	(.26 - .27)	75	2700	2727
28	(.27 - .28)	103	2800	2830
29	(.28 - .29)	86	2900	2916
30	(.29 - .30)	107	3000	3023
31	(.30 - .31)	118	3100	3141
32	(.31 - .32)	98	3200	3239
33	(.32 - .33)	91	3300	3330
34	(.33 - .34)	104	3400	3434
35	(.34 - .35)	102	3500	3536
36	(.35 - .36)	110	3600	3646
37	(.36 - .37)	109	3700	3755
38	(.37 - .38)	108	3800	3863
39	(.38 - .39)	103	3900	3966
40	(.39 - .40)	97	4000	4063
41	(.40 - .41)	86	4100	4149
42	(.41 - .42)	91	4200	4240
43	(.42 - .43)	90	4300	4330
44	(.43 - .44)	106	4400	4436
45	(.44 - .45)	98	4500	4534
46	(.45 - .46)	103	4600	4637
47	(.46 - .47)	88	4700	4725

48	(.47 - .48)	98	4800	4823
49	(.48 - .49)	105	4900	4928
50	(.49 - .50)	121	5000	5049
51	(.50 - .51)	99	5100	5148
52	(.51 - .52)	91	5200	5239
53	(.52 - .53)	98	5300	5337
54	(.53 - .54)	102	5400	5439
55	(.54 - .55)	97	5500	5536
56	(.55 - .56)	113	5600	5649
57	(.56 - .57)	90	5700	5739
58	(.57 - .58)	101	5800	5840
59	(.58 - .59)	100	5900	5940
60	(.59 - .60)	85	6000	6025
61	(.60 - .61)	100	6100	6125
62	(.61 - .62)	101	6200	6226
63	(.62 - .63)	93	6300	6319
64	(.63 - .64)	86	6400	6405
65	(.64 - .65)	94	6500	6499
66	(.65 - .66)	91	6600	6590
67	(.66 - .67)	116	6700	6706
68	(.67 - .68)	114	6800	6820
69	(.68 - .69)	110	6900	6930
70	(.69 - .70)	103	7000	7033
71	(.70 - .71)	102	7100	7135
72	(.71 - .72)	109	7200	7243
73	(.72 - .73)	93	7300	7336
74	(.73 - .74)	96	7400	7432
75	(.74 - .75)	100	7500	7532
76	(.75 - .76)	96	7600	7628
77	(.76 - .77)	95	7700	7723
78	(.77 - .78)	89	7800	7812
79	(.78 - .79)	112	7900	7924
80	(.79 - .80)	109	8000	8033
81	(.80 - .81)	120	8100	8153
82	(.81 - .82)	89	8200	8242
83	(.82 - .83)	96	8300	8338
84	(.83 - .84)	94	8400	8432
85	(.84 - .85)	109	8500	8541
86	(.85 - .86)	86	8600	8627
87	(.86 - .87)	116	8700	8743
88	(.87 - .88)	92	8800	8835
89	(.88 - .89)	98	8900	8933
90	(.89 - .90)	94	9000	9027
91	(.90 - .91)	80	9100	9107
92	(.91 - .92)	111	9200	9218
93	(.92 - .93)	106	9300	9324
94	(.93 - .94)	106	9400	9430
95	(.94 - .95)	99	9500	9529
96	(.95 - .96)	101	9600	9630
97	(.96 - .97)	88	9700	9718
98	(.97 - .98)	100	9800	9818
99	(.98 - .99)	93	9900	9911
100	(.99 - 1.00)	89	10000	10000

GRAPH OF THE CUMULATIVE DISTRIBUTION FUNCTION
FOR 10000 UNIFORM (0,1) RANDOM NUMBERS IN 100 INTERVALS



.5049
.5148
.5239
.5337
.5435
.5536
.5646
.5739
.5840
.5940
.6025
.612[illegible]
.6226
.6315
.6405
.6495
.6590
.6706
.6820
.6930
.7033
.7135
.7243
.7336
.7432
.7532
.7628
.7723
.7813
.7924
.8033
.8153
.8242
.8338
.8432
.8541
.8627
.8743
.8835
.8933
.9027
.9107
.9218
.9324
.9430
.9530
.9630
.9718
.9818
.9911
1.0000

* * * FREQUENCY TEST (EQUIDISTRIBUTION TEST) * * *

COMPUTED CHI-SQUARE TEST STATISTIC = 100.3400

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 123.2253
(DEGREES OF FREEDOM = 99)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF RANDOM NUMBERS IN EACH CATEGORY IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * KOLMOGOROV-SMIRNOV TEST FOR RANDOMNESS * * *

COMPUTED K-S STATISTIC = .0074

CRITICAL VALUE FOR K-S TEST = .0136

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE MAXIMUM DEVIATION OF THE SAMPLE CUMULATIVE DISTRIBUTION FUNCTION FROM THE THEORETICAL CUMULATIVE DISTRIBUTION FUNCTION IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * MAXIMUM OF T TEST * * *

(T = 100)

COMPUTED VALUE OF MAXIMUM OF T TEST STATISTIC = .0400

CRITICAL VALUE OF MAXIMUM OF T TEST STATISTIC (ALPHA = .05) = .1340

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE MAXIMUM DEVIATION OF THE SAMPLE CUMULATIVE DISTRIBUTION FUNCTION OF THE MAXIMA OF CONSECUTIVE BLOCKS OF 100 NUMBERS FROM THE CORRESPONDING THEORETICAL CUMULATIVE DISTRIBUTION FUNCTION IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * GAP TEST * * *

GAP LENGTH	OBSERVED FREQUENCY OF GAPS OF THIS LENGTH	EXPECTED FREQUENCY OF GAPS OF THIS LENGTH	CONTRIBUTION TO THE CHI-SQUARE TEST STATISTIC
0	882.0000	900.6000	.3841
1	639.0000	630.4200	.1168
2	429.0000	441.2940	.3425
3	313.0000	308.9058	.0543
4	232.0000	216.2341	1.1495
5	146.0000	151.3638	.1901
6	116.0000	105.9547	.9524
7	86.0000	74.1683	1.8875
8	159.0000	173.0593	1.1422

(LAST CATEGORY INCLUDES GAPS OF LENGTH 8 OR MORE)

NUMBER OF OBSERVED GAPS IN THE SEQUENCE OF 10000 PSEUDO UNIFORM (0,1) RANDOM NUMBERS = 3002

COMPUTED CHI-SQUARE TEST STATISTIC = 6.2193

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 15.5073
(DEGREES OF FREEDOM = 8)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF GAPS OF EACH LENGTH IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * POKER TEST * * *

K	OBSERVED NUMBER OF POKER HANDS WITH K DISTINCT CARDS	EXPECTED NUMBER OF POKER HANDS WITH K DISTINCT CARDS	CONTRIBUTION TO THE CHI-SQUARE TEST STATISTIC
1	6.0000	3.2000	2.4500
2	195.0000	192.0000	.0469
3	953.0000	960.0000	.0510
4	773.0000	768.0000	.0326
5	73.0000	76.8000	.1880

COMPUTED CHI-SQUARE TEST STATISTIC = 2.7685

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 9.4877
(DEGREES OF FREEDOM = 4)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF POKER HANDS OF EACH TYPE IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * COUPON COLLECTOR'S TEST * * *

COUPON SEQUENCE LENGTH	OBSERVED FREQUENCY OF COUPON SEQUENCES OF THIS LENGTH	EXPECTED FREQUENCY OF COUPON SEQUENCES OF THIS LENGTH	CONTRIBUTION TO THE CHI-SQUARE TEST STATISTIC
5	29.0000	33.6350	.6387
6	55.0000	67.2701	2.2381
7	90.0000	87.4511	.0743
8	90.0000	94.1781	.1854
9	85.0000	91.5416	.4675
10	79.0000	83.6302	.2563
11	63.0000	73.4159	1.4777
12	71.0000	62.7495	1.0848
13	47.0000	52.6537	.6071
14	50.0000	43.6130	.9353
15	196.0000	185.7747	.5628

(LAST CATEGORY INCLUDES COUPON SEQUENCES OF LENGTH 15 OR MORE)

NUMBER OF OBSERVED COUPON SEQUENCES IN THE SEQUENCE OF 10000 PSEUDO UNIFORM (0,1) RANDOM NUMBERS = 855

COMPUTED CHI-SQUARE TEST STATISTIC = 8.5280

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 18.3070
(DEGREES OF FREEDOM = 10)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF COUPON COLLECTOR SEQUENCES OF EACH LENGTH IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * PERMUTATION TEST * * *

PERMUTATION TYPE	OBSERVED FREQUENCY OF PERMUTATION TYPE	EXPECTED FREQUENCY OF PERMUTATION TYPE	CONTRIBUTION TO THE CHI-SQUARE TEST STATISTIC
(C,A,B)	569.0000	555.5000	.3281
(B,C,A)	570.0000	555.5000	.3785
(B,A,C)	507.0000	555.5000	4.2345
(C,B,A)	554.0000	555.5000	.0041
(A,C,B)	592.0000	555.5000	2.3983
(A,B,C)	541.0000	555.5000	.3785

(A = SMALLEST VALUE , B = MIDDLE VALUE , C = LARGEST VALUE)

COMPUTED CHI-SQUARE TEST STATISTIC = 7.7219

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 11.0705
(DEGREES OF FREEDOM = 5)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF PERMUTATIONS OF EACH TYPE IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * TEST ON RUNS * * *

I. TEST ON NUMBER OF RUNS

EXPECTED NUMBER OF RUNS OF EITHER TYPE = 5000.50
ESTIMATED STANDARD DEVIATION FOR THE NUMBER OF RUNS OF EITHER TYPE = 28.86896

LOWER CRITICAL VALUE OF STANDARD NORMAL TEST STATISTIC (ALPHA = .05) = -1.9600
UPPER CRITICAL VALUE OF STANDARD NORMAL TEST STATISTIC (ALPHA = .05) = 1.9600

A. VALUE OF STANDARD NORMAL TEST STATISTIC FOR RUNS UP = 1.2643
(NUMBER OF RUNS UP = 5037.)

B. VALUE OF STANDARD NORMAL TEST STATISTIC FOR RUNS DOWN = -1.2643
(NUMBER OF RUNS DOWN = 4964.)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF RUNS OF EITHER TYPE IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

II. TEST ON LENGTH OF RUNS

RUN LENGTH	EXPECTED NUMBER OF RUNS OF THIS LENGTH	OBSERVED NUMBER OF RUNS UP OF THIS LENGTH	CONTRIBUTION TO THE CHI-SQUARE TEST STATISTIC	OBSERVED NUMBER OF RUNS DOWN OF THIS LENGTH	CONTRIBUTION TO THE CHI-SQUARE TEST STATISTIC
1	1667.33	1702.0000	-86.9808	1629.0000	-331.7590
2	2083.38	2093.0000	-48.3033	2062.0000	-370.0137
3	916.55	921.0000	-33.4992	942.0000	660.8418
4	263.82	269.0000	-51.9568	250.0000	-478.5999
5	57.52	41.0000	207.2582	66.0000	367.0346
6	11.90	11.0000	13.9793	15.0000	166.3252

(LAST CATEGORY INCLUDES RUNS OF LENGTH 6 OR MORE)

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 12.5916
(DEGREES OF FREEDOM = 6)

A. COMPUTED CHI-SQUARE TEST STATISTIC FOR RUNS UP = 5.7271

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF RUNS UP OF EACH LENGTH IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

B. COMPUTED CHI-SQUARE TEST STATISTIC FOR RUNS DOWN = 5.7204

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE NUMBER OF RUNS DOWN OF EACH LENGTH IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * SERIAL TEST FOR SUCCESSIVE PAIRS * * *

FREQUENCY TABLE FOR OBSERVED PAIRS OF RANDOM INTEGERS

	0	1	2	3	4	5	6	7	8	9	ROW TOTALS
0	46.0	53.0	47.0	40.0	47.0	41.0	42.0	45.0	38.0	52.0	451.0
1	48.0	57.0	66.0	54.0	46.0	42.0	37.0	58.0	56.0	54.0	518.0
2	56.0	52.0	44.0	55.0	48.0	44.0	67.0	55.0	53.0	50.0	524.0
3	59.0	66.0	47.0	53.0	45.0	53.0	58.0	50.0	49.0	48.0	527.0
4	47.0	47.0	48.0	53.0	44.0	43.0	48.0	47.0	64.0	51.0	492.0
5	55.0	39.0	57.0	55.0	50.0	58.0	50.0	42.0	56.0	47.0	509.0
6	56.0	66.0	50.0	53.0	45.0	44.0	45.0	51.0	43.0	44.0	517.0
7	49.0	55.0	46.0	53.0	49.0	55.0	48.0	44.0	44.0	52.0	495.0
8	39.0	50.0	47.0	43.0	53.0	54.0	58.0	60.0	42.0	58.0	504.0
9	34.0	56.0	48.0	54.0	47.0	33.0	38.0	53.0	46.0	54.0	463.0
COLUMN TOTALS	489.0	541.0	500.0	513.0	494.0	467.0	491.0	505.0	490.0	510.0	5000.0

EXPECTED NUMBER OF PAIRS OF RANDOM INTEGERS IN EACH CELL = 50.0

COMPUTED CHI-SQUARE TEST STATISTIC = 98.3200

CRITICAL VALUE OF CHI-SQUARE TEST STATISTIC (ALPHA = .05) = 123.2253
(DEGREES OF FREEDOM = 99)

FAIL TO REJECT THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE DISTRIBUTION OF SUCCESSIVE PAIRS OF RANDOM NUMBERS IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

* * * SERIAL CORRELATION TEST * * *

LOWER CRITICAL VALUE OF STANDARD NORMAL TEST STATISTIC (ALPHA = .05) = -1.9600
UPPER CRITICAL VALUE OF STANDARD NORMAL TEST STATISTIC (ALPHA = .05) = 1.9600

CALCULATED VALUES OF STANDARD NORMAL TEST STATISTIC

LAG (L)	CIRCULAR (PERIODIC EFFECT)	NONCIRCULAR (TRENDS)
1	1.6672 FAIL TO REJECT	1.6601 FAIL TO REJECT
2	-.5283 FAIL TO REJECT	-.5365 FAIL TO REJECT
3	.7203 FAIL TO REJECT	.7041 FAIL TO REJECT
4	-.1720 FAIL TO REJECT	-.2139 FAIL TO REJECT
5	.8717 FAIL TO REJECT	.8299 FAIL TO REJECT
6	-.1700 FAIL TO REJECT	-.2110 FAIL TO REJECT
7	.5299 FAIL TO REJECT	.4445 FAIL TO REJECT
8	-.1623 FAIL TO REJECT	-.2685 FAIL TO REJECT
9	-.9451 FAIL TO REJECT	-1.0484 FAIL TO REJECT
10	-.3901 FAIL TO REJECT	-.5609 FAIL TO REJECT

(FAIL TO REJECT / REJECT) THE HYPOTHESIS, AT THE 5 PERCENT LEVEL OF SIGNIFICANCE, THAT THE SERIAL CORRELATION BETWEEN RANDOM NUMBERS SEPARATED BY LAG (L) IS AS WOULD BE EXPECTED IF THE RANDOM NUMBERS CAME FROM A UNIFORM (0,1) DISTRIBUTION

DISTRIBUTION

Commander
Naval Air Development Center
ATTN: Technical Library
Warminster, PA 18974

Commander
U. S. Army Missile Research and
Development Command
ATTN: Technical Library
Redstone Arsenal, AL 35809

Director
Naval Research Laboratory
ATTN: Technical Information Division
Washington, DC 20375

Office of Naval Research
Department of the Navy
ATTN: ONR-436
Washington, DC 20360

Commanding Officer
U. S. Air Force Armament Laboratory
ATTN: Technical Library
Eglin Air Force Base, FL 32542

Commanding Officer
Naval Weapons Center
ATTN: Code 324C (Kurotori)
Technical Library
China Lake, CA 93555

Superintendent
U. S. Naval Postgraduate School
ATTN: Technical Library
Monterey, CA 93940

Superintendent
U. S. Naval Academy
ATTN: Library
Annapolis, MD 21402

Commanding Officer
U. S. Army Research Office
Durham, NC 27706

Library of Congress
ATTN: Gift and Exchange Division (4)
Washington, DC 20540

Department of Statistics
VPI&SU
ATTN: Statistical Laboratory
Blacksburg, VA 24060

Department of Statistics
Purdue University
ATTN: Dr. George McCabe, Jr.
Lafayette, IN 47905

Commander
U. S. Army Armament Research and
Development Command
ATTN: DRDAR-LCS-E (Dr. Einbinder)
Technical Library
Dover, NJ 07801

Commander
U. S. Army Armament Material
Readiness Command
ATTN: DRSAR-PES (Mr. Michels)
Technical Library
Rock Island, IL 61299

Director
U. S. Army Ballistic Research Laboratory
ARRADCOM
ATTN: Technical Library
Aberdeen Proving Ground, MD 21005

DISTRIBUTION (Continued)

Oklahoma State University
Field Office
P. O. Box 1925
ATTN: Mrs. Peebles
Eglin Air Force Base, FL 32542

Director
Naval Ship Research and Development Center
ATTN: Technical Library
Washington, DC 20034

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314 (2)

Local:

E23 (McCleskey)
E31 (GIDEP)
E33 (Holbrook)
E411 (Hall)
F28 (Petranka)
G12 (Ahlquist)
G22 (Clawson)
G32 (Seidl)
G51 (Meyerhoff)
K
K10
K105 (Harvey)
K106 (35)
K13 (Meyerhoff)
K30
K33 (Gass)
K33 (Krein)
K33 (Morris)
K34
K35 (McLaughlin)
K40
K42 (Nordai)
K50
K51 (Bond)
K51 (Mendenhall)
K51 (Owens)
K52 (Farr)
K52 (Lipscomb)
K54 (Schmoyer)
N21 (Case)
N52 (Weddell)
R11
R13
R14
R16
R32
R34
R44
E431 (10)


DATE
ILME