C. 2

BNWL-1743
Special
Distribution

AN EVALUATION OF FOUR PSEUDO-RANDOM NUMBER GENERATORS

Richard O. Gilbert

Battelle
Pacific Northwest Laboratories
Richland, Washington 99352

MAY 1973

Prepared for the U.S. Atomic Energy
Commission under Contract AT(45-1):1830

BNWL-1743

PACIFIC NORTHWEST LABORATORY
operated by
BATTELLE
for the
U.S. ATOMIC ENERGY COMMISSION
Under Contract AT(45-1)-1830

BNWL-1743
Special Distribution

AN EVALUATION OF FOUR PSEUDO-RANDOM NUMBER GENERATORS

By

Richard O. Gilbert

May 1973

Statistics Section
Systems Department
Pacific Northwest Laboratories
Battelle Memorial Institute
Richland, Washington

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

FIGURES		PAGE

1. INTRODUCTION

This paper is concerned with an evaluation of the performance of four pseudo-random number generators in generating numbers from the two parameter lognormal and gamma distributions. It is anticipated that one or two of these generators will be used in studies of sampling strategies for studying the movement of trace substances in food chains. The lognormal and gamma distributions have been suggested (Eberhardt (1972)) as suitable frequency distributions to represent observations of the concentration of substances such as radionuclides and DDT in samples of biota.

The four generators used here are denoted by RANDU1, RND, RN2 and RN4. Each of these is designed to generate independent uniformly distributed pseudo-random numbers between 0 and 1. These numbers are transformed to lognormal or gamma variates using equations (3) and (4) given below.

It is perhaps impossible to come to a definitive solution to the problem of choosing between RANDU1, RND, RN2 and RN4, since each can generate millions of different sequences of pseudo-random numbers. For a given generator, some of these sequences may have undesirable statistical properties and others may not. (See Marsaglia (1968) for a discussion of why certain random number generators may be unsuitable for some Monte Carlo problems.)

The approach used here to evaluate these generators was to generate just one sequence of pseudo-random U(0,1) numbers for each generator and to apply various distributional and randomness tests to the same numbers along this one sequence. This approach could be characterized as an intensive investigation of the numbers in one part of one sequence rather than an extensive evaluation of the generators by investigating many different sequences or different sections of the same sequence. This approach would appear to be adequate for our purposes since all that is required for the study of sampling strategies (and most any other Monte Carlo study) is a single sequence of pseudo-random U(0,1) numbers that is known as far as possible to have desirable statistical properties plus speed and long periodicity. The reader should not assume, however, that the results obtained here using just a part of one sequence necessarily apply to other sequences for that generator.

Each generator was evaluated for randomness by (i) computing the chi-square goodness-of-fit test on generated U(0,1) numbers falling in 100 equal intervals of size .01 between 0 and 1, and (ii) by performing the lag product test (Naylor, Balintfy, Burdick, and Chu (1968)) for lags k between 1 and 25 (a test for correlation between numbers in the sequence which are k units apart). In addition, each generator was used in conjunction with equations (1) through (4) below to generate log-normal and gamma variates. The goodness of fit of these generated distributions was evaluated by chi-square goodness-of-fit tests as well as other tests of significance based on normal theory described in sections 5, 6 and 7. In addition, some information is obtained on the speed and period of each generator. All computations were performed on the UNIVAC 1108 computer.

On the whole, the above tests point to the conclusion that RN4 and RN2 are about equal in performance followed by RND and RANDU1 in that order. RANDU1 seems to be particularly susceptible to serial correlation between members of the generated sequence. This in turn appears to result in relatively poor performance of RANDU1 in generating gamma variates which require summing several U(0,1) numbers. RANDU1 also appears to be the slowest of the 4 generators tested. RND is apparently only slightly less accurate than RN2 and RN4, and has the advantage that it is available on the teletype terminal. More detailed conclusions and discussions are given in section 8.

2. RANDOM NUMBER GENERATORS

RANDU1 is a slightly modified version of RANDU which is a MATH-PACK library subprogram on the UNIVAC 1108 computer described in the Computer Science Corporation (CSC) publication number UP-7542. RANDU1 and RANDU use the CSC subprogram NRAND (also described in UP-7542) which generates integers on the interval $(0, 2^{27})$ by a congruence method (Abramowitz and Stegun (1967)). The CALL statement for RANDU is

CALL RANDU (X,N)

where X is the starting value of the generator and N the number of U(0,1) pseudo-random numbers to be generated. RANDU was modified and renamed RANDU1 for two reasons:

(i) it was decided to always use N = 1 so that only one U(0,1) random number would be generated for each CALL statement.

(ii) It was anticipated that the evaluation of RANDU1 would require more than one computer execution (run) making it advantageous to know the arguments of NRAND at the end of each run so the sequence could be continued in the next computer run with no possibility of overlap. The CALL statement of RANDU1 is

CALL RANDU1 (X,1,J,K)

where X is the generated U(0,1) number and J and K are the arguments of NRAND which are printed out at the end of each computer run. The second argument 1 merely indicates that only one X is generated for each CALL RANDU1 statement. Program listings of RANDU, RANDU1 and NRAND are given in the appendix. The values of J and K used here are 4048130 and 2207286, respectively, which were chosen at random. These values identify the starting point of the sequence of U(0,1) numbers evaluated here.

In the CSC publication UP-7542, four sets of 500 uniform numbers were generated by RANDU (each set using a different starting number X) and the chi-square goodness-of-fit test applied to the data. None of the tests were significant, but it was concluded that the chi-square statistic obtained depends heavily on the initial number X supplied to RANDU.

RN4 was originally written by Kronmal (1964) for the IBM 7094. Two uniform random numbers U_1 and U_2 are generated with one CALL statement, each by a different mixed congruential generator. Box and Muller (1958) showed that if U_1 and U_2 are independent random variables from the U(0,1) distribution then

$$X_1 = (-2 \ln U_1)^{\frac{1}{2}} \cos (2\pi U_2) \qquad (1)$$

and

$$X_2 = (-2 \ln U_1)^{\frac{1}{2}} \sin (2\pi U_2) \qquad (2)$$

are normally distributed with mean equal to zero and variance equal to one. Kronmal tested several mixed congruential generators for generating independent U(0,1) numbers. These generators were evaluated by using equations (1) and (2) to produce a sequence of numbers which were statistically tested for normality and randomness. The mixed congruential generators ultimately chosen by Kronmal satisfied these tests on eight samples of one million numbers each. The tests for randomness were (i) serial correlations of lag 1 and lag 2, (ii) number of runs above and below zero, and (iii) the random dispersion of extreme values ($> \pm 3.891$). Normality was tested by computing means and variances of each sample, comparing empirical and cumulative distributions, and generating the distribution of the range and the order statistics for small samples. The results indicated close agreement with normal theory.

RN4 was modified for use on the UNIVAC 1108 and is called by

CALL RN4 (UU1, UU2, U1, U2)

where UU1 and UU2 are two 12 digit octal starting values (arguments) for the congruence generators and U1 and U2 are the generated uniform variates. The numbers UU1 = 233362477003 and UU2 = 212312312323 used by Kronmal were also used in the present paper.

RN2 is similar to RN4 except that it uses only one mixed congruential U(0,1) generator. It is called by

CALL RN2 (UU1, U1) ,

where UU1 and U1 are as defined above for RN4. The sequence of numbers obtained using RN2 was begun with UU1 = 011060471625 which was chosen at random. The last values of UU1 and UU2 for a given run of RN4 and RN2 were printed out and used to continue the sequence on the next occasion. Program listings of RN4 and RN2 are given in the appendix.

The generator RND is the function RND available on the Computer Sciences Conversational Executive (CSCX) BASIC programming language which is used when communication with the UNIVAC 1108 is by remote teletype terminal. No information is readily available on the method of generation used in RND or on its performance in generating U(0,1) numbers. A total of 200,000 U(0,1) numbers were generated using RND and stored on file for evaluation using the UNIVAC 1108 in Fortran V. Unfortunately, it is not possible to reconstruct this sequence of 200,000 numbers since the starting values are not supplied by the user and are consequently unknown. Hence, future Monte Carlo studies using RND will not in all probability use the sequence statistically evaluated in this paper.

3. METHODS

For the purposes of investigating sampling strategies using non-linear models, we wish to generate lognormal and gamma variables with expectation 1 and standard deviation C. Note that this implies the coefficient of variation defined to be C.V. = (standard deviation/mean) is also equal to C. Consider first the 2-parameter lognormal distribution whose density function is given by

$$f(y;\mu,\sigma) = \frac{1}{\sigma y\sqrt{2\Pi}} \exp\left[-\frac{1}{2\sigma^2}(\ln y - \mu)^2\right]$$

for $y > 0$, $-\infty < \mu < \infty$, $\sigma > 0$. We have that

$$E(y) = \exp(\mu + \tfrac{1}{2}\sigma^2)$$

and

$$\mathrm{Var}(y) = [\exp(2\mu + \sigma^2)][\exp(\sigma^2) - 1]$$

(Hahn and Shapiro (1967)). It follows that if $\mu = -\frac{\sigma^2}{2}$, then $E(y) = 1$ and $\mathrm{Var}(y) = \exp(\sigma^2) - 1 = C^2$. Furthermore, if $x \sim N(0,1)$ then

$$y = \exp(\mu + \sigma x) \tag{3}$$

is distributed lognormally with parameters μ and σ (Hahn and Shapiro (1967), p. 242). Thus, lognormal variates with expectation 1 and standard deviation C were generated by specifying a value for C, solving the equation $\exp(\sigma^2) - 1 = C^2$ for σ, setting $\mu = -\sigma^2/2$, generating x using one of the random number generators investigated here, and using equation (3) to obtain y. Clearly the distribution of y depends entirely on x for a given value of C. The lognormal distribution is studied in some detail in the monograph by Aitchison and Brown (1969).

The density function of the 2-parameter gamma distribution is given by

$$f(y;\eta,\lambda) = \frac{\lambda^{\eta}}{\Gamma(\eta)} \; x^{\eta-1} \; e^{-\lambda y}$$

for $y \geq 0$, $\lambda > 0$, $\eta > 0$, where $\Gamma(\eta)$ is the well known gamma function (Hahn and Shapiro (1967)). It follows that

$$E(y) = \eta/\lambda \quad \text{and} \quad Var(y) = \eta/\lambda^2$$

which implies

$$C^2 = 1/\eta \quad .$$

Our conditions that $E(y) = 1$ and $Var(y) = C$ are hence satisfied if we set $\eta = \lambda$. Furtherfore, if $U_i \sim U(0,1)$, then the transformation

$$y = -\frac{1}{\lambda} \sum_{i=1}^{\eta} \ln(1 - U_i) \tag{4}$$

yields a gamma variate y with parameters λ and η (for integer η) (Hahn and Shapiro (1967), p. 242). Consequently, gamma variates with expectation 1 and standard deviation C were generated by specifying a value for C, solving $C^2 = 1/\eta$ for η, setting $\eta = \lambda$, generating U_i using RANDU1, RND, RN2 or RN4 and solving equation (4) for y.

When using RN4 a question arises concerning which, or in what proportion the two mixed congruential number generators should be used to supply η uniform numbers U_i required in equation (4). An examination of this question is beyond the scope of this paper. Our results are based on using $\eta/2$ from each generator if η is even. When η is odd the extra uniform number required is taken from the first generator.

4. TESTS FOR PERIODICITY

By the period of a pseudo-random number generator we mean the length of the sequence of numbers which can be generated before the generator begins repeating itself, i.e., generating the same sequence over again. In general, generators are constructed so that their periods are as large as possible, usually 2^{27} or 2^{36}, while still retaining good statistical properties and speed of generation. The period of the four generators used here are believed to be at least 2^{27} (approximately 134 million). However, to insure that their periods are at least 2 or 3 million, the generators were allowed to generate a sequence of pseudo-random U(0,1) numbers for approximately 250 seconds of 1108 computer time. The starting values for each generator were as indicated in the previous section, i.e., the same sequences were used to evaluate the period of RANDU1, RN2 and RN4 as were used to test their statistical properties. The period of RND was examined using the teletype terminal and hence used different pseudo-random numbers than those stored on tape for statistical testing purposes using the UNIVAC 1108.

RN2 and RN4 each generated between 6 and 6.5 million U(0,1) pseudo-random numbers in the allotted time period, while RANDU1 generated between 3 and 3.5 million. RND generated 2,400,000 numbers in 137 seconds indicating approximately 4,400,000 numbers would be generated in 250 seconds. None of the 4 generators began repeating during this time.

This was determined for RANDU1, RN4 and RN2 by testing whether the arguments of each newly generated U(0,1) number in the sequence were equal to the arguments used for the first number in the sequence. (These arguments are J and K for RANDU1, UU1 and UU2 for RN4, and UU1 for RN2.) In the case of RND, since the argument starting the generator is not supplied by the used and is unknown, the periodicity was evaluated by testing whether each newly generated U(0,1) number was equal to the first such number generated in the sequence. The results of the above tests indicate the periods of RANDU1, RND and both generators of RN4 are at least 3,000,000 and that of RN2 at least 6,000,000.

The above generation times give only a rough guide to the speed of the generators since the length of the computer programs varied slightly between generators, and RND was used over a remote terminal using the BASIC language. However, the above results suggest that RN2, RN4 and RND may be somewhat faster than RANDU1.

5. TESTS FOR RANDOMNESS

(a) Goodness-of-Fit Tests

Eight samples of pseudo-random U(0,1) numbers from each generator were constructed and the chi-square test computed on each. Two samples each of sizes 500,1000,5000 and 10000 were generated. In addition, the chi-square test was computed on the pooled frequencies of each of the 100 intervals over the eight samples (a total of 33,000 observations). The sum of the resulting eight chi-square values was also obtained since under the hypothesis of randomness it is distributed as a chi-square with (in this instance) 792 degrees of freedom. These results are given in Table 1. We see that RANDU1 and RN4 had no significant chi-square values while RND and RN2 had one and two significant results, respectively. Furthermore, the pooled and sum chi-squares for RND were significant. Figure 1 is a plot of the pooled frequencies obtained using RND. There are no apparent patterns to the deviations from the expected frequency of 330 in each interval.

The computer time required to generate the eight samples of pseudo-random numbers and to perform the above indicated chi-square tests are also given in Table 1 for each generator except RND. The computer programs written to calculate these chi-square tests are nearly identical except for the specific CALL, READ and WRITE statements which of necessity must be unique for each generator. Hence the different times reported in Table 1 are believed to be due primarily to differences in the speeds of the number generators. RN2 and RN4 appear to be considerably faster than RANDU1, the same conclusion reached above in our discussion of periodicity. The time required for RND is not given since it is a function of opening and closing files containing the U(0,1) numbers and thus cannot be compared with the times shown for the other generators.

b. Lag Product Test

The lag product test for a given lag k is a test of the hypothesis of no correlation between numbers k positions apart in a sequence of numbers of length N. The statistic is defined to be

$$C_k = \frac{1}{N-k} \sum_{i=1}^{N-k} U_i \, U_{i+k} \tag{5}$$

where k is the length of the lag. If there is no correlation between U_i and U_{i+k} for i = 1,2,. . ., N-k then C_k is approximately normally distributed with μ = .25 and standard deviation

$$\sigma_k = (13N - 19k)/12(N-k) \tag{6}$$

for large N (Naylor, et.al., (1968)).

The procedure used here to compare the 4 random number generators was to generate 100 sequences of 200 U(0,1) pseudo-random numbers on each generator and compute C_k for k = 1,2, . . ., 25 on each sequence. For each value of k between 1 and 25, the mean $\overline{C}_k$ and standard deviation s_k of the 100 computed values of C_k were obtained and compared with their expected values of .25 and σ_k (equation (6)), respectively. A frequency distribution of 100 values of the transformed variables

$$Z_{ki} = \frac{C_{ki} - .25}{\sigma_k} \quad , \quad i = 1,2, \ldots, 100$$

was constructed for each k from 1 to 25 and compared with the expected frequencies using a chi-square goodness-of-fit test assuming Z_{ki} is N(0,1). The results are given in Table 2, where each chi-square value has 7 degrees of freedom. The numbers in columns 1 and 2 under each generator in Table 2 are deviations defined as

$$C_{kd} = (\bar{C}_k - .25) \times 10^4$$

and

$$s_{kd} = (s_k - \sigma_k) \times 10^4$$

respectively (d stands for deviation). C_{kd} and s_{kd} should be approximately zero if the generators are operating properly, i.e., generating U(0,1) numbers.

No significant chi-square values at the $\alpha = .05$ level were obtained for RN2 and RN4 for any value of the lag k between 1 and 25. RND gave one significant chi-square for k = 19, but RANDU1 resulted in 9 significant chi-squares. These were for k equal to 5, 11, 13, 16, 19, 22, 23, 24 and 25. The sum of the 25 chi-squares obtained for RANDU1 was also significant with 175 degrees of freedom. Thus, it appears that the numbers generated here by RANDU1 are correlated. We also note from Table 2 that the mean $\overline{C}_k$ of the computed C_k values for most values of k between 1 and 25 overestimates slightly the expected mean .25 (none were statistically significant using the t test with 99 d.f.). This is especially true of RND. Similarly, the standard deviation s_k of the 100 values of C_k tends to underestimate the expected value σ_k, particularly for RANDU1 and RND although it also occurs for RN2 and RN4. It is not known whether these deviations from expectation are a manifestation of the significant chi-squares in the case of RANDU1. It seems unlikely since these deviations are for the most part quite small. Another possibility is that the length of the sequence (200-k for lag k) is not sufficiently large for the normal approximation of the distribution of C_k to be accurate. For purposes of comparing the four generators, however, the adequacy of the approximation is not of crucial importance since it is the same for each generator.

6. EVALUATION OF GENERATED LOGNORMAL VARIATES

The performance of RANDU1, RND, RN2 and RN4 in generating lognormal variates was evaluated for values of the standard deviation C equal to .2, .5 and .7. These distributions are plotted in Figure 2. We remind the reader that since $\sigma^2 = \ln [C^2 + 1]$, C completely specifies the parameters μ and σ of the lognormal distribution under our specification that $\mu = -\sigma^2/2$. The restriction that $\mu = -\sigma^2/2$ insures that $E(y) = 1$ and $Var(y) = C^2$.

These generators were evaluated for each value of C by generating a total of 50 samples of lognormal variates of size 500 each using equation (3). The normal variate x required in equation (3) was obtained using equations (1) and (2) (Box and Muller (1958)). Two chi-square goodness-of-fit tests to the lognormal distribution were obtained for each C by splitting the 50 samples into two equal groups of size 25 samples each, which gave a total of 12,500 numbers for each chi-square test. The number of intervals for these tests were 44, 70 and 90 for C = .2, .5 and .7, respectively. More intervals were used for larger values of C since the lognormal distribution becomes highly skewed to the right as C increases (Figure 2). The length of the intervals in the tails of the distribution is generally greater than in the middle of the distribution to insure that the expected frequency in each cell in the tail is at least 5 for sample sizes of 12,500.

The chi-square results are given in Table 3. The only significant departures from lognormality were for RANDU1 and RN2 for C = .2. The observed and expected frequencies for these cases are given in Figures 3 and 4. We note that the sum of the two chi-square values is significant for RANDU1 but not for RN2.

In addition to these chi-square tests, each of the 50 samples of size n = 500 were investigated for fit to the lognormal distribution by transforming each lognormal variate y to a normal variate x using the transformation x = ln y. If y is lognormal with parameters μ and σ, then x is normal with these same parameters (Aitchison and Brown (1969)). Thus, lognormality of the y's is implied if tests of significance on the transformed variate x indicates no departure from normality.

Four tests of normality on the x's were computed for each of the 50 samples of size n = 500: (i) that the variate

$$Z = \left(\frac{\bar{x} - \mu}{\sigma}\right)\sqrt{n}$$

is distributed N(0,1), where μ and σ are determined from C as described above, (ii) that $\frac{s^2}{\sigma^2}(n-1)$ is distributed as a chi-square with n-1 d.f., where

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad ,$$

(iii) that the coefficient of skewness γ_1 is equal to zero, and (iv) that the coefficient of kurtosis γ_2 is equal to 3. (γ_1 and γ_2 are defined in

Hahn and Shapiro (1967), p. 45, 46). If the y's are lognormally distributed with parameters μ and σ so that $x_i \sim N(\mu, \sigma^2)$ then the probability is .05 that each of these tests on each sample of size 500 will be statistically significant if the significance level of each test is α = .05. The estimate of γ_1 is computed as

$$\hat{\gamma}_1 = m_3/(m_2)^{3/2} \tag{7}$$

and that of γ_2 by

$$\hat{\gamma}_2 = m_4/m_2^2 \quad , \tag{8}$$

where

$$m_j = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^j \quad .$$

The proportion of statistically significant results (upper and lower α = .05 level tests) for tests (i) through (iv) are given in Tables 4 and 5. Table 4 also gives the proportion of the 50 computed values of Z and $s^2(n-1)/\sigma^2$ (tests (i) and (ii) above) that fall in the upper 50% of their respective null distributions. Of course, the expected proportion in each tail is .05 and in the upper half of the null distribution is .5 if the generators have indeed generated sequences of independent U(0,1) pseudo-random numbers. The sum of the two observed proportions (each based on 50 tests) for the two tails in Tables 4 and 5 were tested for departure from expectation .10 using the two-tailed binomial test at the exact significance level α = .05833.

Considering first the test (i) that $\overline{x}$ is significantly different from μ, we see from Table 4 that of the 4 generators only RND gave a proportion of significant results different from the expected .10. However, each generator gave at least one significantly different proportion when test (ii) was applied to each of the 50 samples. RND produced too many significant results for C = .5 and .7 as did RN2 for C = .2. Both RANDU1 and RN4 produced too few significant results for C = .5. We have also tested (Table 4) that the proportion in the upper half of the null distributions of tests (i) and (ii) are significantly different from .5 using the 2-tailed binomial test at significance level α = .03284. RND and RN4 each gave one significant proportion (.34); the former for C = .2 and the latter for C = .7. In Table 5 the only significant proportions obtained for tests (iii) and (iv) on γ_1 and γ_2 were for RANDU1 and RN4 each of which gave too few significant results for the test on γ_2 when C = .5.

It is difficult indeed to choose among the four generators on the basis of the information in Tables 3, 4 and 5. RN2 had the least number of significant results in Tables 4 and 5, but one of its chi-square goodness-of-fit tests in Table 3 was significant for C = .2. RND had the poorest performance under tests (i) and (ii) (Table 4) but it had no significant proportions in Table 5 or chi-square tests in Table 3. RANDU1 had two significant goodness-of-fit tests in Table 3, but its performance in Tables 4 and 5 was good (one significant result in each table). Finally, RN4 had no significant chi-square tests in Table 3) but had three significant results in Tables 4 and 5. Considering Tables 3, 4 and 5 as a whole, more significant results (a total of six) were obtained for C = .2 than for any other value of C. Only two were obtained for C = .7. Further interpretations of the data are deferred until we examine the results of the next section concerning the generation of gamma variates.

7. EVALUATION OF GENERATED GAMMA VARIATES

The methods used to evaluate the effectiveness of RANDU1, RND, RN2 and RN4 in generating gamma variates were similar to those used for the generation of lognormal variates. Equation (4) was used for values of $\lambda = \eta = 2$, 4 and 25, i.e., for values of $C = (1/\eta)^{1/2} = .7071$, .5 and .2, respectively. These distributions are plotted in Figure 5. Fifty samples of size 500 were generated using RANDU1, RN2 and RN4. Only 100,000 RND U(0,1) numbers were stored on file for purposes of generating gamma variates. Thus, it was necessary in the case of RND to reduce the number and size of the samples. Consequently, for C = .2, .5 and .7071 we generated 20 samples of size 100, 25 samples of size 250, and 25 samples of size 500, respectively, which in light of equation (4) makes use of all 100,000 RND numbers stored on file.

For each generator, the statistics $\overline{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ were computed on each sample and tests of significance obtained as described below. As was done with the lognormal distribution the first 25 samples of size 500 were pooled to provide a chi-square goodness-of-fit test based on 12,500 gamma variates for RANDU1, RN2 and RN4. A second goodness-of-fit test was similarly obtained by pooling the second 25 samples of size 500. For RND, one goodness-of-fit test was computed using sample sizes of 20 x 100 = 2000, 25 x 250 = 6250 and 25 x 500 = 12,000 for C = .2, .5 and .7071, respectively. Thus, the tests of significance using the statistics $\overline{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are based on the same generated numbers as the goodness-of-fit tests.

From Table 6 we see that none of the goodness-of-fit tests were significant for RND, RN2 and RN4. For RANDU1, however, the chi-square tests were highly significant for C = .2 and .5. The deviations from expectation and the contribution to total chi-square are given for each interval of the two samples in Tables 7 and 8 for C = .2 and .5, respectively. In Figures 6 and 7 the observed and expected frequencies of sample 1 from Tables 7 and 8 are plotted.

From Table 7 we see that for both samples 1 and 2, the greatest contribution to chi-square occurs in the upper tail (large values of y) where there are more observations than expected, whereas the lower tail is characterized by too few observations. Figure 6 illustrates, however, that the deviations are also large in the central portion of the distribution. Similar results were observed for RANDU1 when C = .5 (Table 8, Figure 7) except that the major contribution to chi-square occurred in the lower tail (small values of y), where the expected frequency substantially exceeded the observed. There were no large deviations from expectation in the upper tail.

The cause of these large chi-square values obtained using RANDU1 is unknown, but it might be explained by the large number of significant results obtained for RANDU1 using the lag-product test for lags 1 through 25 (Table 2). We recall from equation (4) that a gamma variate is obtained by finding the sum of a function of η independently distributed uniform variates with mean 0 and variance 1, where η equals

2, 4 and 25 for C = .7071, .5 and .2, respectively. If the sequence of uniform numbers generated by RANDU1 are not independent (as suggested by the lag-product tests (Table 2)) then it seems reasonable to suppose that as η becomes large, the pseudo-gamma variates generated may tend to deviate from that expected. Our results suggest this hypothesis since the chi-square values obtained for RANDU1 (Table 6) increase rapidly as η increases from 2 to 25. Whether or not the correlations detected by the lag-product tests are related to the significant chi-squares in Table 6 is unknown, but clearly this question should be investigated further before the particular sequence of pseudo-random numbers generated here by RANDU1 is used in monte carlo studies where such correlations could seriously bias the results.

As indicated above, tests using $\bar{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ were also made on the individual samples of pseudo-gamma variates. Unlike the lognormal distribution, there is apparently no transformation available which when applied to a sample of gamma variates will allow the use of normal theory to obtain tests of significance on the transformed sample. Linhart (1965) gives a procedure for finding approximate confidence limits for the coefficient of variation of gamma distributions. Unfortunately, the data obtained here were not in a form suitable for the use of his method. We have, however, obtained tests of significance by appealing to the central limit theorem. Considering first the statistic $\bar{x}$, i.e., the mean of a sample of size n of pseudo-gamma variates. Applying the central limit theorem we have that if $E(\bar{x}) = \mu$ and $Var(\bar{x}) = \sigma^2/n$, then

$$Z = \frac{\bar{x}-\mu}{\sigma}\sqrt{n}$$

is approximately normally distributed with mean zero and variance 1 for large n. Z could be used to test the composite hypothesis that $\mu = 1$ and $\sigma = C$, so that Z could be significantly large if either μ or σ were substantially different from 1 or C, respectively. To avoid this ambiguity of interpretation we have instead used the statistic

$$Z' = \frac{\overline{x}-1}{s}\sqrt{n} \tag{9}$$

to test the hypothesis that $E(\overline{x}) = 1$, its significance being determined by reference to the t distribution with n-1 d.f. at the $\alpha = .10$ level. Table 9 gives the proportion of such significant results for the four generators when C = .2, .5 and .7071. Each proportion is based on $n' = 50$ samples of size n = 500 each, i.e., each mean $\overline{x}$ is based on 500 observations, except for RND (see footnote to Table 9). Table 9 also gives the proportion of positive Z' values, the expected proportion being .5 if Z' is distributed as a t_{n-1}. Each proportion was tested for significant deviation from expectation using the binomial test described above for the lognormal distribution. Significantly low proportions were obtained for RN2 when C = .2 and RN4 for C = .5 and .7071. We note in particular that in contrast to the significant goodness-of-fit tests in Table 6, no significant results were obtained for RANDU1 in Table 9.

Table 10 contains the results of tests made on $\overline{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ by computing each of these statistics on each of the n' samples of size n, finding the mean and standard deviation of each over the n' samples and using the central limit theorem. If the generators are in fact generating independent gamma variates then we expect that the means of $\overline{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ should be approximately equal to 1, C, γ_1 and γ_2, respectively, where for the gamma distribution it is easily shown that $\gamma_1 = 2C$ and $\gamma_2 = 6C^2$ when $E(y) = 1$ and $Var(y) = C^2$ (Hahn and Shapiro, p. 124). Then illustrating with $\overline{x}$, we have that

$$Z_{\bar{x}} = \left(\frac{\bar{\bar{x}} - 1}{s_{\bar{x}}}\right)\sqrt{n'}$$

is approximately distributed as a t with n' - 1 d.f. if n' is large, where

$$s_{\bar{x}} = \frac{1}{n' - 1}\sum_{i=1}^{n'} (\bar{x}_i - \bar{\bar{x}})^2$$

and

$$\bar{\bar{x}} = \frac{1}{n'}\sum_{i=1}^{n'} \bar{x}_i \quad .$$

Similarly the statistic used for testing that $E(\hat{\gamma}_1) = \gamma_1$ is

$$Z_{\gamma_1} = \left(\frac{\bar{\hat{\gamma}}_1 - 2C}{s_{\hat{\gamma}_1}}\right)\sqrt{n'} \quad .$$

We see from Table 10 that there were no significant results obtained for RN2 or RN4, but that the differences $\bar{\hat{\gamma}}_1 - 2C$ and $\bar{\hat{\gamma}}_2 - 6C^2$ were highly significant under RANDU1 for C = .2 and .5. The difference $\bar{\hat{\gamma}}_2 - 6C^2$ was also statistically significant under RND for C = .2.

The results for RANDU1 in Table 10 are not completely unexpected in light of the significant chi-square results in Table 6. The deviation $\bar{\hat{\gamma}}_1 - 2C$ in Table 10 for C = .2 and .5 is positive which agrees with the observed shift in distribution of the generated gamma values from the expected distribution as illustrated in Figures 6 and 7. The interpretation of the significance of $\bar{\hat{\gamma}}_2 - 6C^2$ is more difficult (for a discussion on the interpretation of $\hat{\gamma}_2$, see Darlington (1970)) but its significance here indicates in general that the shape of the distributions of the generated observations is not that expected.

8. CONCLUSIONS AND DISCUSSION

While making no claims that the present study is definitive, or that no more tests need to be considered in order to be exhaustive in our effort, our results do suggest certain conclusions concerning the choice of a random number generator for generating lognormal and gamma variates.

A rather crude indication of the relative performance of RANDU1, RND, RN2 and RN4 can be obtained by simply adding up the number of significant tests obtained for each in Tables 1 through 10. These results are tabulated in Table 11. We find that 20.9% of the tests performed on RANDU1 generated numbers (24 out of 115) were statistically significant. These data for RND are 8.1% (9 out of 111), and for both RN2 and RN4 are 4.3% (5 out of 115). (These numbers are used here only as a general guide since the tests are not in all cases independent. For example, the sum and pooled chi-square tests in Table 1 are not independent of the eight individual chi-square tests.) RANDU1 did poorest on the lag-product tests for lack of serial correlation (Table 2). This may be related to its poor performance in generating gamma distributions (Tables 6, 7, 8, 10; Figures 6 and 7) for C = .2 and .5 since equation (4), which is used to obtain gamma variates, involves summing uniform numbers from 1 to $\lambda = 1/C^2$. It seems reasonable to suppose that as λ increases (C decreases) the variates produced using equation (4) would tend to deviate from the expected gamma distribution. We note from

Table 1, however, that RANDU1 evidently generates numbers which are uniformly distributed over the interval (0,1) at least for sample sizes of 500 or greater. From Table 1 it also appears that RANDU1 is the slowest of the generators considered.

RND had 9 significant results, most of which (7) occurred in Tables 1 and 4. From Table 1 we see that RND had the largest pooled and "sum" chi-squares of the four generators considered although the plot of the 33,000 pooled frequencies of RND in Figure 1 show no discernable pattern in the deviations of the frequencies from expectation. From Table 4 we see that the distribution of the 50 values of $\overline{x}$ from the generated lognormal variates tends to have too few values in the tails for C = .2. For C = .5 and .7, however, the tails of the distribution of s^2 tend to be too large. We note also that the number of significant values of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ obtained for RND (Table 5) do not deviate from expectation. It should also be recalled that the same generated numbers used to compute the statistics in Tables 4 and 5 were pooled to provide the goodness of fit tests to the lognormal distribution in Table 3 for which no significant results were obtained for RND. Of course the tests in Tables 4 and 5 are evaluating the generated lognormal variates in relatively small batches of size 500, whereas the chi-square tests pool these data into much larger sample sizes. Thus the local tests may be picking out deviations from expectation which are masked by the chi-square test.

Of the total of 5 significant tests obtained for RN2, three were goodness-of-fit tests (2 from Table 1 and 1 from Table 3). None of the lag-product tests (Table 2) or the goodness of fit tests on the pseudo-gamma variates were statistically significant. While RN4 had the same total number of significant results as RN2, all are found in Table 4, 5 and 9, i.e. no goodness-of-fit tests were significant.

On the basis of the above results the sequences generated using RND, RN2 and RN4 appear to have somewhat better statistical properties than that generated by RANDU1. Of these remaining three, RN4 is our first choice since Kronmal (1964) has shown the particular sequence generated here to have good statistical properties. However, our test results indicate that the RN2 sequence has properties about equal to that of RN4, and RND has the feature that it is available on the time-share terminal to the UNIVAC 1108 which could be an asset in some Monte Carlo work. We also remind the reader that we have not investigated the most appropriate way of using or combining the two sequences of uniform numbers generated by RN4 (using 2 mixed congruence generators) to obtain gamma variates.

As indicated previously, the present study should not be considered definitive since (i) only a few of the many statistical tests available were applied to the data, and (ii) only a relatively small segment of a single sequence of pseudo-random numbers was examined for each generator. At the very least, however, our results suggest some hypotheses concerning the relative performance of the generators which should, perhaps, be investigated more thoroughly.

ACKNOWLEDGMENTS

The help of Dr. Lee Eberhardt and Dr. John Thomas in the design, analysis and interpretation of the data is gratefully acknowledged. Special thanks also go to Clarence Oster who modified RN2 and RN4 to operate on the UNIVAC 1108. We also thank Jeff Buchanan who was of great assistance on the computer.

REFERENCES

1. Abramowitz, M. and Stegun, I. A. (1967). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. U.S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series, 55.

2. Aitchison, J. and Brown, J. A. C. (1969). The Lognormal Distribution. Cambridge University Press.

3. Box, G. P., and Muller, M. E. (1958). A Note on the Generation of Random Normal Deviates. Annals of Mathematical Statistics, 29, 610-611.

4. Darlington, R. B. (1970). Is Kurtosis Really "Peakedness"?, The American Statistician, 24, pp. 19-21, April.

5. Eberhardt, L. L. (1973). Gamma and Log-Normal Distributions as Models in Studying Food-Chain Kinetics. Battelle Memorial Institute, Pacific Northwest Laboratories, Richland, Washington. (Report in preparation)

6. Hahn, G. J. and Shapiro, S. S. (1967). Statistical Models in Engineering, John Wiley and Sons, Inc., New York.

7. Kronmal, R. (1964). Evaluation of a Pseudorandom Normal Number Generator, Journal of the Association for Computing Machinery, 11, pp. 357-363.

8. Linhart, H. (1965). Approximate Confidence Intervals for the Coefficient of Variation of Gamma Distributions, Biometrics 21, 733-738.

9. Marsaglia, G. (1968). Random Numbers Fall Mainly in the Planes. Proceedings of the National Academy of Sciences, 61, September-December, pp. 25-28.

10. Naylor, T. H., Balintfy, J. L., Burdick, B. S. and Kong Chu (1968). Computer Simulation Techniques, John Wiley and Sons, Inc., New York.

11. Univac 1108 MATH PACK No. UP-7542, Computer Science Corporation, Los Angeles, California.

12. Anonymous (1969). CSCX BASIC Reference E00007-00-02, November 3, Computer Sciences Corporation Information Network Division, Los Angeles, California.

Table 1. Chi-Square Goodness-of-Fit Tests† For Randomness

SAMPLE SIZE	RANDU1	RND	RN2	RN4
500	95.60	116.80	90.40	88.00
500	88.00	118.40	123.60*	74.40
1000	114.80	123.20	94.60	110.00
1000	94.20	95.60	113.40	90.80
5000	117.04	92.08	89.48	85.84
5000	85.80	100.44	126.00*	73.04
10000	110.52	155.24**	98.92	114.78
10000	91.20	97.68	102.38	112.96
33000	86.15	132.78*	86.98	89.23
SUM (792 d.f.)	797.16	899.44**	838.78	749.82
Time	7.09 sec.	——	4.94 sec.	5.22 sec.

† Each test has 99 degrees of freedom (d.f.)

* Significant at α = .05

** Significant at α = .01

Table 2. Summary of Lag Product Test Results

LAG	RANDU1			RND			RN2			RN4		
	C_{kd}	S_{kd}	χ^2	C_{kd}	S_{kd}	χ^2	C_{kd}	S_{kd}	χ^2	C_{kd}	S_{kd}	χ^2
1	3	-20	9.37	7	-16	2.67	1	-10	3.76	5	- 7	9.87
2	- 2	-15	12.72	5	-18	4.09	14	-15	6.28	12	-10	4.59
3	- 9	-31	13.27	16	-13	3.95	10	2	3.64	0	-18	9.89
4	- 2	- 9	6.53	13	- 5	6.99	-16	-16	2.57	-29	- 1	7.61
5	-11	-11	27.70**	1	-11	2.33	12	-11	7.35	11	- 9	8.11
6	0	- 4	5.53	23	-11	4.60	- 7	-12	2.14	14	- 6	2.86
7	1	-12	5.73	18	-19	8.33	3	- 3	7.93	14	1	5.51
8	- 1	6	0.68	11	-11	1.96	33	- 5	3.59	35	30	12.59
9	- 2	-32	15.24	6	- 6	6.95	5	- 7	1.77	9	6	8.90
10	3	6	7.05	- 2	-22	7.34	7	-27	8.33	8	- 9	11.00
11	5	-31	18.16*	8	-20	13.78	12	- 1	3.39	6	-11	7.64
12	6	- 4	3.42	13	- 8	9.46	0	-15	9.34	- 3	24	6.30
13	1	-40	16.52*	3	- 8	7.82	8	-18	4.19	3	3	5.85
14	- 1	-16	3.09	13	-12	8.78	8	-12	6.11	3	-11	8.50
15	4	-12	8.62	10	-19	10.53	9	- 7	3.75	3	- 2	10.37
16	2	-36	14.16*	9	+ 1	5.72	11	3	6.04	1	-10	7.84
17	8	-14	5.42	8	- 8	7.89	12	- 8	2.21	2	-10	9.26
18	1	-15	7.49	4	- 6	7.22	- 2	-25	6.80	16	-12	4.18
19	- 4	-28	15.86*	18	-23	14.13*	9	- 2	3.48	11	- 1	4.38
20	8	6	4.11	13	-14	10.93	4	- 7	4.23	- 4	31	2.54

Table 2. (Continued)

	RANDU1			RND			RN2			RN4		
	C_{kd}	s_{kd}	χ^2	C_{kd}	s_{kd}	χ^2	C_{kd}	s_{kd}	χ^2	C_{kd}	s_{kd}	χ^2
21	10	-30	7.76	23	-13	3.93	7	-12	3.72	6	0	6.03
22	8	-38	16.56*	25	-11	4.82	2	-21	7.30	- 3	-11	7.12
23	- 4	-14	25.56**	14	-18	7.91	10	-23	5.20	1	10	2.15
24	4	-37	19.93**	9	-10	7.79	- 6	- 5	12.05	4	27	11.24
25	- 1	- 2	20.30**	7	-12	9.91	9	- 3	0.92	7	- 5	5.12
SUM			290.79**			179.83			126.08			179.43
AVE.	1.08	-17.32	11.63	11	-12.52	7.19	6.20	-10.4	5.04	5.28	-.04	7.18

† $C_{kd} = (\overline{C}_k - .25) \times 10^4$; $s_{kd} = (s_k - \sigma_k) \times 10^4$; see text for definition of C_k and s_k.

Table 3. Chi-Square Goodness-of-Fit Tests of Pseudo-Lognormal Numbers

	RANDU1	RND	RN2	RN4
C = .2				
Sample 1 (43 d.f.)	63.66*	57.46	31.33	32.85
Sample 2 (43 d.f.)	45.74	42.94	67.78**	43.55
SUM (86 d.f.)	109.40*	100.40	99.11	76.40
C = .5				
Sample 1 (69 d.f.)	45.44	61.77†	59.72	53.14
Sample 2 (69 d.f.)	57.15	69.58	81.98	64.36
SUM (138 d.f.)	102.59	131.35	141.70	117.50
C = .7				
Sample 1 (89 d.f.)	111.51	71.85	88.89	95.64
Sample 2 (89 d.f.)	92.98	96.44	81.47	85.48
SUM (178 d.f.)	204.49	168.29	170.36	181.12

†Two additional chi-square tests were obtained for the RND generator when C = .5; the chi-square values were 64.87 and 70.32 with 69 d.f., both non-significant at the $\alpha = .05$ level.

Table 4. Proportion of Significant Tests† for $\overline{x}$ and s^2 on Pseudo-Lognormal Variates

	RANDU1		RND		RN2		RN4	
C = .2	$\overline{x}$	s^2	$\overline{x}$	s^2	$\overline{x}$	s^2	$\overline{x}$	s^2
UT	.02	.04	.02	.06	.14	.06	.06	.04
LT	.08	.02	.00	.06	.00	.12	.04	.04
UT + LT	.10	.06	.02*	.12	.14	.18*	.10	.08
$P_{.5}$	.50	.38	.34*	.52	.62	.36	.60	.50
C = .5								
UT	.02	.02	.04	.08	.10	.08	.08	.02
LT	.08	.00	.02	.12	.06	.08	.00	.00
UT + LT	.10	.02*	.06	.20*	.16	.16	.08	.02*
$P_{.5}$	.52	.38	.40	.38	.40	.56	.40	.36
C = .7								
UT	.08	.06	.06	.08	.06	.04	.10	.04
LT	.04	.06	.08	.10	.08	.06	.04	.02
UT + LT	.12	.12	.14	.18*	.14	.10	.14	.06
$P_{.5}$	.56	.38	.52	.56	.54	.58	.34*	.58

†UT = upper tail; LT = lower tail; based on α = .05 in each tail

$P_{.5}$ = proportion of the 50 sample values of $\overline{x}$ and s^2 that were greater than μ and $(499)^{-1}\sigma^2\chi^2_{499,\ \alpha\ =\ .50}$, respectively, where $x \sim N(\mu,\sigma^2)$ if y is lognormal with parameters μ and σ^2.

*Indicates the proportion so marked is significantly different from that expected using the binomial test; see text for explanation.

Table 5. Proportion of Significant Tests† for Skewness (γ_1) and Kurtosis (γ_2) on Pseudo-Lognormal Variates

	RANDU1		RND		RN2		RN4	
C = .2	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2
UT	.04	.04	.02	.04	.04	.08	.00	.04
LT	.02	.06	.02	.00	.04	.00	.06	.00
UT + LT	.06	.10	.04	.04	.08	.08	.06	.04
C = .5								
UT	.04	.02	.06	.06	.04	.06	.00	.00
LT	.06	.00	.04	.04	.00	.06	.04	.02
UT + LT	.10	.02*	.10	.10	.04	.12	.04	.02*
C = .7								
UT	.02	.06	.06	.06	.06	.08	.06	.06
LT	.08	.00	.04	.02	.02	.02	.10	.04
UT + LT	.10	.06	.10	.08	.08	.10	.16	.10

† UT = upper tail; LT = lower tail; α = .05 in each tail

* Indicates the proportion of significant test results at the α = .10 (two-tailed) level is significantly different (α = .05833) from that expected from binomial theory.

Table 6. Chi-Square Goodness-of-Fit Tests of Pseudo-Gamma Variates[†]

	RANDU1	RND	RN2	RN4
C = .2				
Sample 1 (54 d.f.)	650.24**	59.18	66.85	54.14
Sample 2 (54 d.f.)	587.34**	—	55.10	35.86
SUM (108 d.f.)	1237.58**	—	121.95	90.00
C = .5				
Sample 1 (49 d.f.)	348.19**	57.47	54.08	31.18
Sample 2 (49 d.f.)	344.42**	—	36.02	53.51
SUM (98 d.f.)	692.61**	—	90.10	84.69
C = .7071				
Sample 1 (48 d.f.)	29.14	34.53	42.46	42.20
Sample 2 (48 d.f.)	48.68	—	49.62	50.13
SUM (96 d.f.)	77.82	—	92.08	92.33

† Each chi-square computed on n = 12,500 generated gamma variates except for RND under C = .2 and .5 when n = 2000 and 6250, respectively.

Table 7. Contributions of Each Interval to the Chi-Square Goodness-of-Fit Tests to the Gamma Distribution Obtained Using RANDU1 for C = .2

Interval	O-E†		Contribution To Chi-Square		Interval	O-E		Contribution To Chi-Square	
	Sample 1	Sample 2	Sample 1	Sample 2		Sample 1	Sample 2	Sample 1	Sample 2
0- .54	-39.2	-37.2	38.22	34.42	1.10	- 5.3	-53.3	0.07	6.86
.56	-14.5	-12.5	9.34	6.94	1.12	-66.8	-54.8	11.48	7.72
.58	-23.2	-15.2	16.72	7.18	1.14	3.0	-29.0	0.02	2.32
.60	-21.6	-19.6	10.46	8.61	1.16	-48.6	-26.6	7.06	2.11
.62	-18.2	-27.2	5.50	12.29	1.18	-66.0	-42.0	14.19	5.75
.64	-15.2	-19.2	2.92	4.65	1.20	-52.7	-63.7	9.93	14.51
.66	-44.6	-45.6	19.58	20.47	1.22	-47.1	-41.1	8.76	6.67
.68	-65.5	-43.5	33.65	14.84	1.24	-36.5	-60.5	5.86	16.09
.70	-23.6	-26.6	3.56	4.52	1.26	-47.2	-55.2	10.96	15.00
.72	-18.5	-37.5	1.82	7.46	1.28	-36.4	-34.4	7.34	6.56
.74	-30.7	- 3.7	4.23	0.06	1.30	-40.3	-46.3	10.20	13.46
.76	- 0.4	-16.4	0.001	1.04	1.32	-38.7	-44.7	10.72	14.30
.78	46.3	29.3	7.27	2.91	1.34	-29.9	-16.9	7.33	2.34
.80	75.3	3.3	17.15	0.03	1.36	-16.8	-22.8	2.67	4.91
.82	45.6	48.6	5.69	6.46	1.38	-20.4	-26.4	4.55	7.63
.84	63.0	64.0	9.97	10.29	1.40	-16.5	- 7.5	3.47	0.72
.86	55.6	89.6	7.23	18.78	1.42	3.9	-11.1	.23	1.84
.88	105.0	64.0	24.33	9.04	1.44	- 2.0	-18.0	.07	5.68
.90	114.9	79.9	27.85	13.47	1.46	- 4.3	- 0.3	.38	0.002
.92	28.8	68.8	1.69	9.66	1.48	5.3	11.3	.69	3.14
.94	71.1	58.1	10.09	6.74	1.50	- 1.1	13.9	.04	5.67
.96	75.7	77.7	11.32	11.92	1.52	6.5	12.5	1.48	5.48
.98	47.7	58.7	4.49	6.81	1.54	15.3	4.3	9.88	0.78
1.00	23.8	69.8	1.13	9.72	1.58	11.1	4.1	3.43	0.47
1.02	6.7	37.7	0.09	2.89	1.62	10.8	18.8	4.82	14.60
1.04	3.9	9.9	0.03	0.21	1.68	25.0	29.0	28.4	38.23
1.06	- 6.1	36.9	0.08	2.97	>1.68	69.9	58.9	211.52	150.18
1.08	-16.0	10.0	0.58	0.23			Total χ^2_{54}	650.2	587.3

† O-E = Observed-Expected frequencies based on 12,500 generated gamma variates in each sample.

Table 8. Contributions of Each Interval to the Chi-Square Goodness-of-Fit Tests to the Gamma Distribution Obtained Using RANDU1 for C = .5

Interval	O-E†		Contribution To Chi-Square		Interval	O-E		Contribution To Chi-Square	
	Sample 1	Sample 2	Sample 1	Sample 2		Sample 1	Sample 2	Sample 1	Sample 2
0-.15	- 42.0	- 42.0	42.00	42.00	1.4	-19.4	- 3.4	1.33	0.04
.2	- 57.5	- 57.5	46.24	46.24	1.45	- 5.2	11.8	0.10	0.54
.25	- 77.9	- 83.9	48.98	56.81	1.5	10.5	-12.5	0.47	0.67
.3	- 84.8	- 96.8	38.91	50.70	1.55	-14.2	-16.2	0.95	1.24
.35	- 3.5	30.5	0.05	3.73	1.6	- 4.4	15.6	0.10	1.27
.4	73.4	94.4	17.18	28.42	1.65	24.9	0.9	3.60	0.00
.45	112.3	125.3	33.75	42.01	1.7	1.7	-13.3	0.02	1.15
.5	127.9	85.9	38.30	17.28	1.75	8.0	- 7.0	0.46	0.36
.55	91.9	32.9	17.89	2.29	1.8	2.9	- 4.1	0.07	0.14
.6	56.1	67.1	6.20	8.86	1.85	1.4	8.4	0.02	0.64
.65	39.0	12.0	2.85	0.27	1.9	13.7	- 7.3	1.93	0.55
.7	29.2	- 7.8	1.55	0.11	1.95	13.8	- 1.2	2.21	0.02
.75	16.2	37.2	0.47	2.48	2.0	2.8	4.8	0.10	0.30
.8	- 16.9	- 8.9	0.51	0.14	2.05	- 4.3	- 6.3	0.27	0.59
.85	- 48.0	- 7.0	4.17	0.09	2.1	- 0.2	- 4.2	0.00	0.30
.9	- 23.3	- 35.3	1.01	2.31	2.15	9.9	1.9	1.88	0.07
.95	- 25.6	- 13.6	1.26	0.35	2.2	- 2.7	- 4.7	0.16	0.48
1.0	- 39.2	- 34.2	3.07	2.34	2.25	6.9	0.9	1.19	0.02
1.05	- 53.8	10.2	6.08	0.22	2.3	- 0.1	17.9	0.00	9.13
1.1	- 40.4	- 24.4	3.63	1.32	2.35	- 9.7	4.3	3.06	0.60
1.15	- 41.8	- 38.8	4.14	3.57	2.4	1.2	- 1.8	0.05	0.12
1.2	- 26.5	- 51.5	1.78	6.74	2.45	- 0.3	0.7	0.00	0.02
1.25	- 34.0	- 4.0	3.17	0.04	2.5	8.7	7.7	3.73	2.92
1.3	19.0	14.0	1.07	0.58	3.0	8.4	16.4	0.70	2.67
1.35	2.3	- 6.7	0.02	0.14	>3.0	- 6.6	- 6.6	1.52	1.52
							Total χ^2_{49}	348.2	344.4

† O-E = Observed-Expected frequencies based on 12,500 generated gamma variates in each sample.

Table 9. Proportion of Statistically Significant Means $\overline{x}$ of Generated Gamma Variates, and Proportion of Means $\overline{x}$ Greater Than Zero (see text)

	RANDU1	RND†	RN2	RN4
C = .2				
α = .10	.06	0.00	.02*	.12
P = .50	.42	.50	.58	.42
C = .5				
α = .10	.06	.12	.04	.18*
P = .50	.46	.60	.52	.50
C = .7071				
α = .10	.16	.04	.16	.02*
P = .50	.48	.44	.54	.54

† Proportions based on 20 samples of size 100, 25 samples of size 250 and 25 samples of size 500 for C = .2, .5 and .7071, respectively. All other proportions are based on 50 samples each of size 500.

* The proportion so marked is significantly different from expected using the binomial test at $\alpha = .05833$.

Table 10. Significance Tests on the Mean Values of $\bar{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ Computed on Pseudo-Gamma Variates††

	RANDU1			RND†			RN2			RN4		
C = .2	O-E††	S.E.††	t	O-E††	S.E.††	t	O-E††	S.E.††	t	O-E††	S.E.††	t
$\bar{x}$	-8.65	11.0	- 0.79	8.8	28.0	0.31	3.35	10.1	0.33	-13.4	14.9	-0.90
	1.26	11.4	1.11	12.0	28.0	0.43	7.35	18.0	0.41	2.5	17.2	0.14
$\hat{\gamma}_1$	5329.5	220.0	24.2**	- 681.2	483.6	-1.41	305.2	248.3	1.23	95.85	175.2	0.55
γ_2	12776.4	883.6	14.5**	-2393.0	910.5	-2.63**	1882.7	980.7	1.92	103.85	517.7	0.20
C = .5												
$\bar{x}$	-9.5	29.1	-.33	31.8	60.5	0.53	12.45	26.3	0.47	8.1	35.6	0.23
s	3.45	31.1	.11	42.95	54.1	0.79	9.91	27.4	0.36	4.76	28.6	0.17
$\hat{\gamma}_1$	458.45	179.6	2.55**	175.9	447.8	0.39	423.3	306.1	1.38	18.9	347.1	0.05
γ_2	-2068.2	903.4	2.29**	-201.4	2360.4	-0.09	2994.2	1791.7	1.67	1042.0	2269.1	0.46
C = .707												
$\bar{x}$	-10.15	46.8	- 0.22	16.1	58.4	0.28	-11.3	50.8	-0.22	-26.35	33.1	-0.80
s	71.04	52.1	1.36	9.1	75.0	0.12	0.478	42.6	0.01	-15.34	46.2	-0.33
$\hat{\gamma}_1$	568.25	431.6	1.32	82.8	496.5	0.17	- 283.0	329.0	-0.86	- 201.9	338.2	-0.60
$\hat{\gamma}_2$	3335.25	3348.1	1.00	762.7	3853.8	0.20	-2084.8	2272.4	-0.92	-1031.25	2462.4	-0.42

† Results for RND are based on 20 samples of size 100, 25 samples of size 250, and 25 samples of size 500 for C = .2, .5 and .7, respectively. All results for RANDU1, RN2 and RN4 are based on 50 samples of size 500.

†† The tabled values for O-E and S.E. must be multiplied by 10^{-4} to obtain the actual values.

O-E equals $\bar{\bar{x}} - 1$, $\bar{s} - C$, $\bar{\hat{\gamma}}_1 - 2C$ and $\bar{\hat{\gamma}}_2 - 6C^2$ for $\bar{x}$, s, $\hat{\gamma}_1$ and $\hat{\gamma}_2$, respectively.

Table 11. Number of Significant Tests In Tables 1 Through 10 For Each U(0,1) Generator

Table	RANDU1	RND	RN2	RN4	No. of Tests For Each Generator	
1	0	3	2	0	10	
2	10	1	0	0	51	
3	2	0	1	0	9	(11 for RND)
4	1	4	1	2	12	
5	1	0	0	1	6	
6	6	0	0	0	9	(3 for RND)
9	0	0	1	2	6	
10	4	1	0	0	12	
Sum:	24	9	5	5	115	(111 for RND)
Percent Significant Tests :	20.9%	8.1%	4.3%	4.3%		

Figure 1. Frequency Distribution of 33000 U(0,1) Numbers Generated by RND

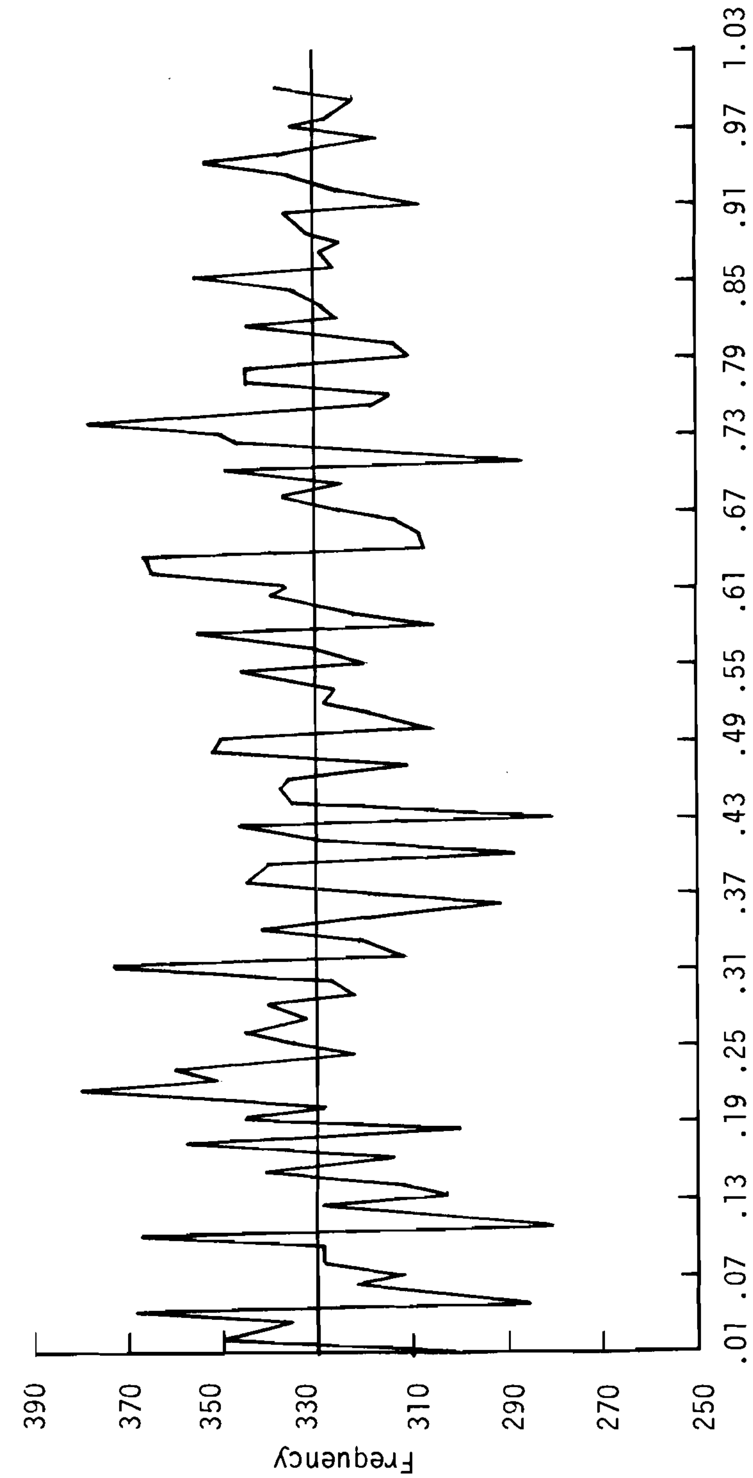


Figure 2. Density Functions of the Lognormal Distribution for Standard Deviations C = .2, .5 and .7 When $\mu = -\sigma^2/2$

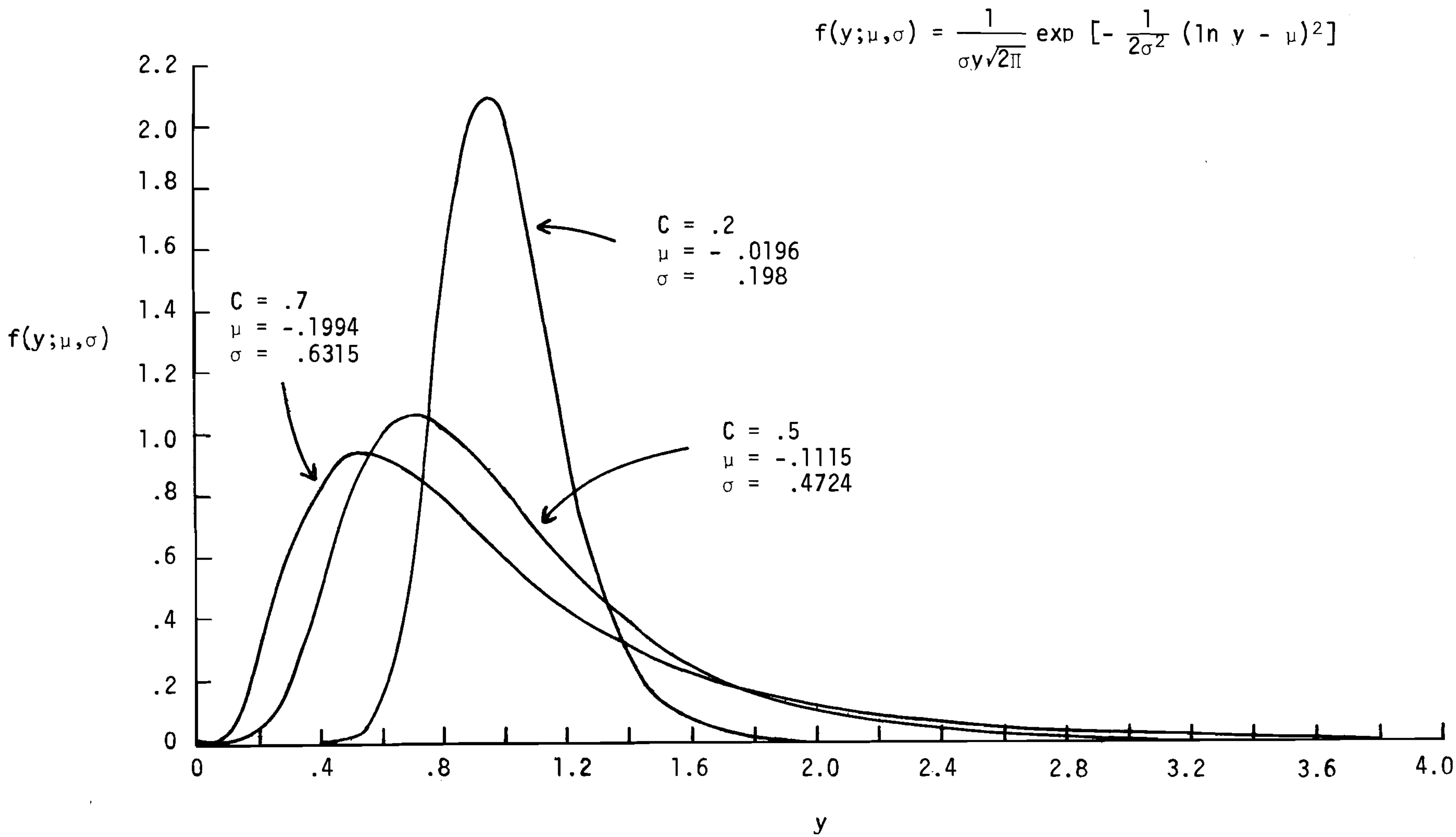


Figure 3. Observed and Expected Frequencies of Generated Lognormal Variates for C = .2 Using RANDU1 (Sample 1, Table 3)

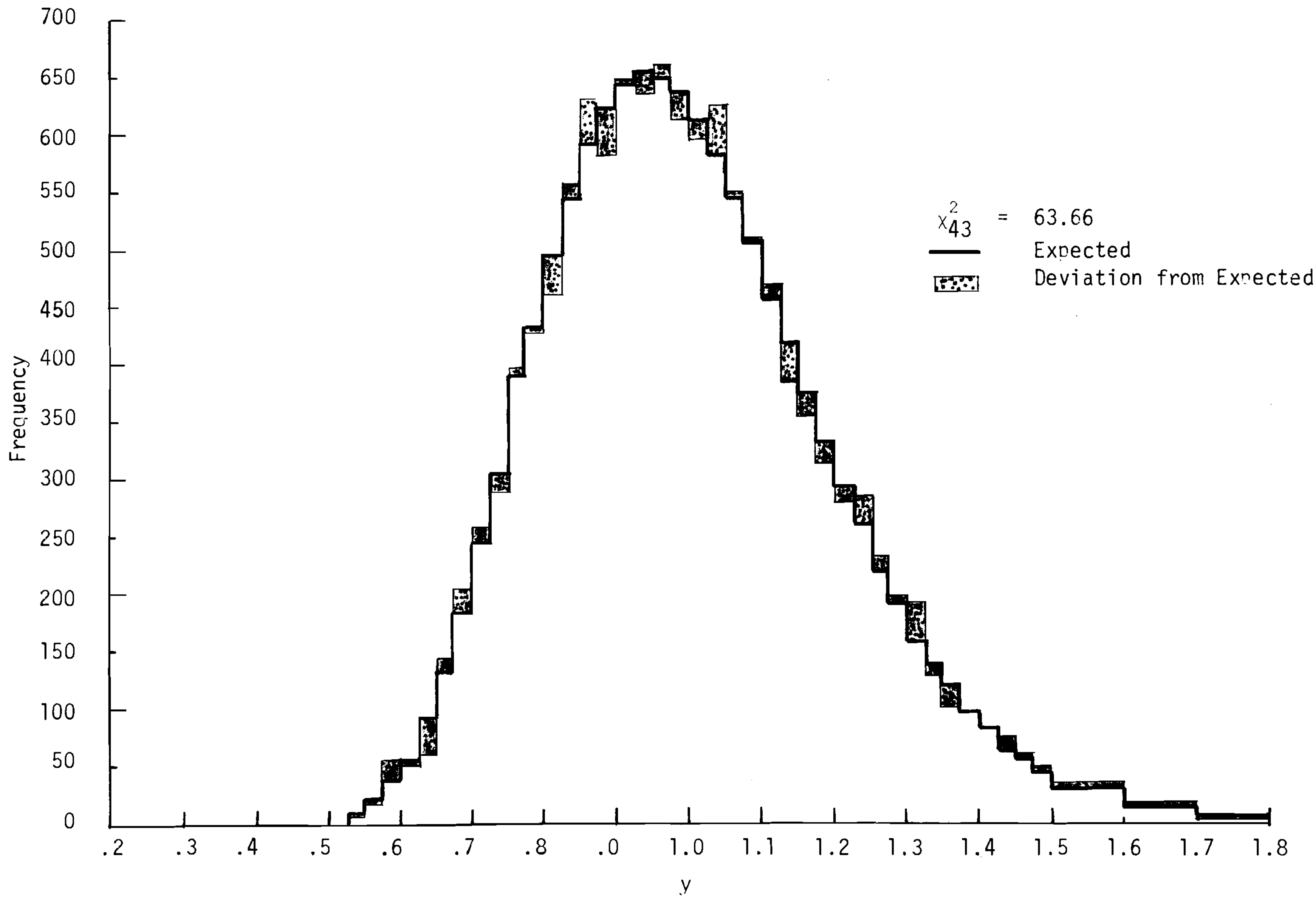


Figure 4. Observed and Expected Frequencies of Generated Lognormal Variates for C = .2 Using RN2 (Sample 2, Table 3)

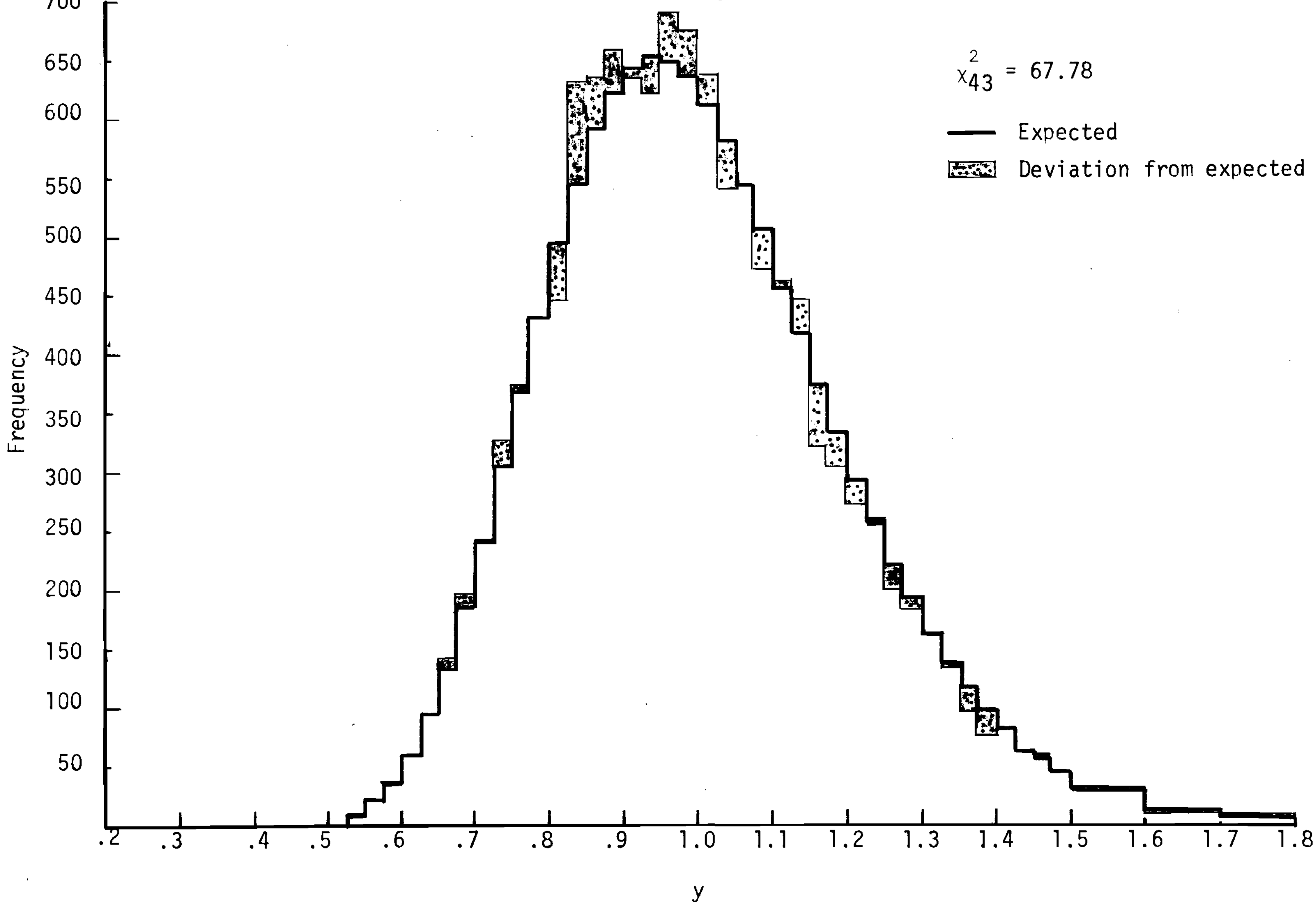


Figure 5. Density Function of the Gamma Distribution for Standard Deviations $C = .2$, $.5$ and $.7071$ When $C^2 = 1/\eta$ and $\eta = \lambda$

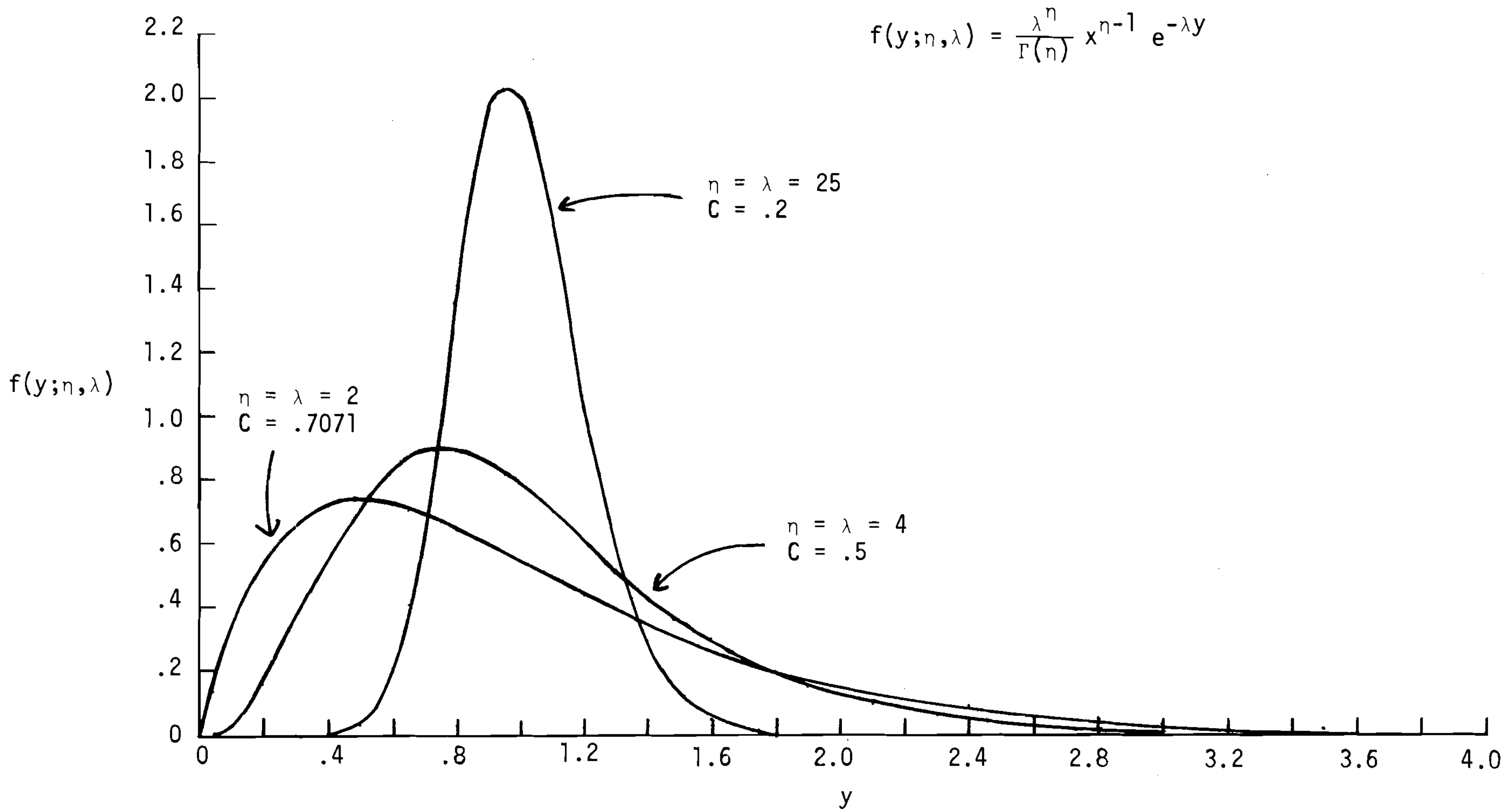


Figure 6. Observed and Expected Frequencies of Generated Gamma Variates for C = .2 Using RANDU1 (Sample 1 from Table 7)

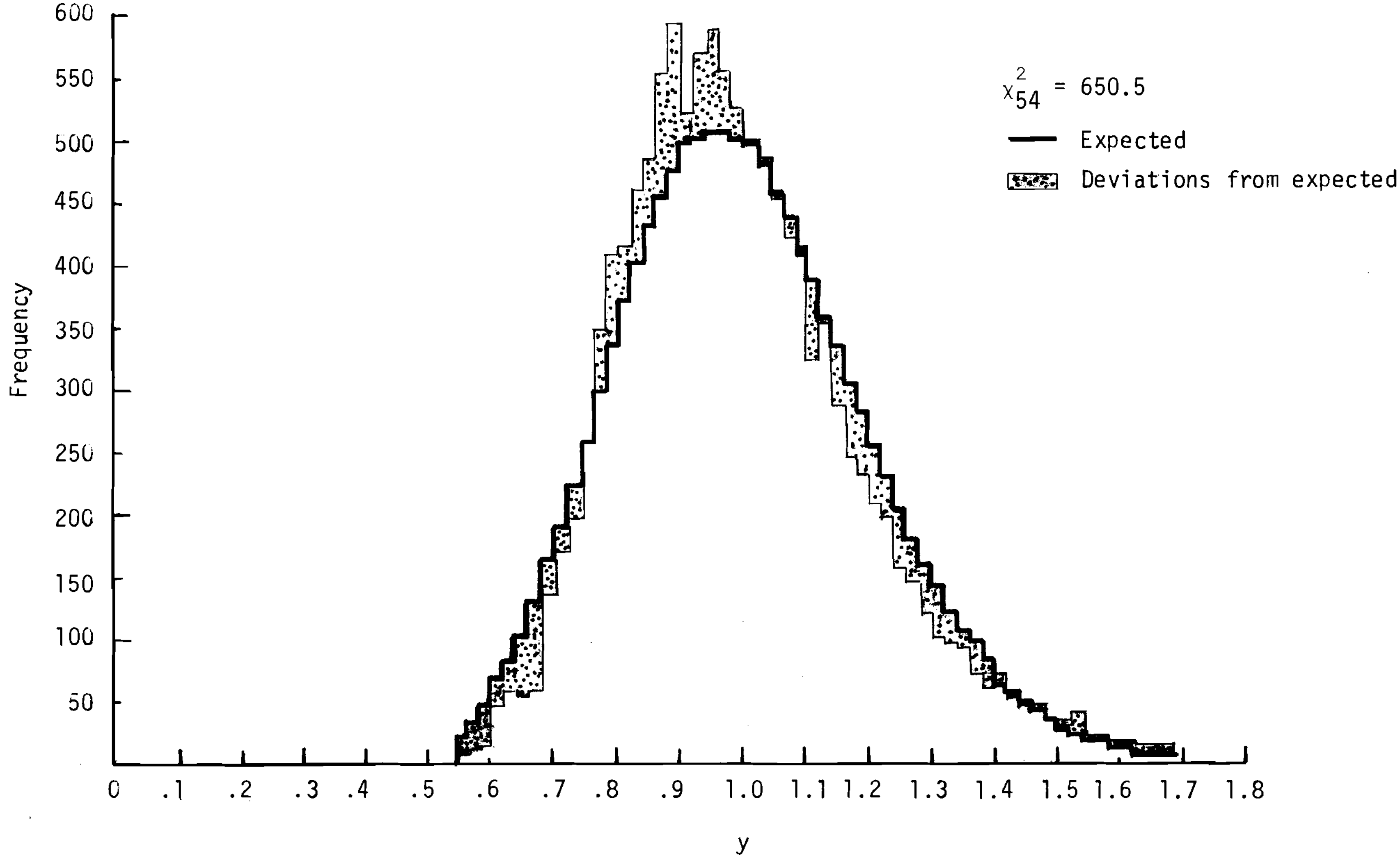
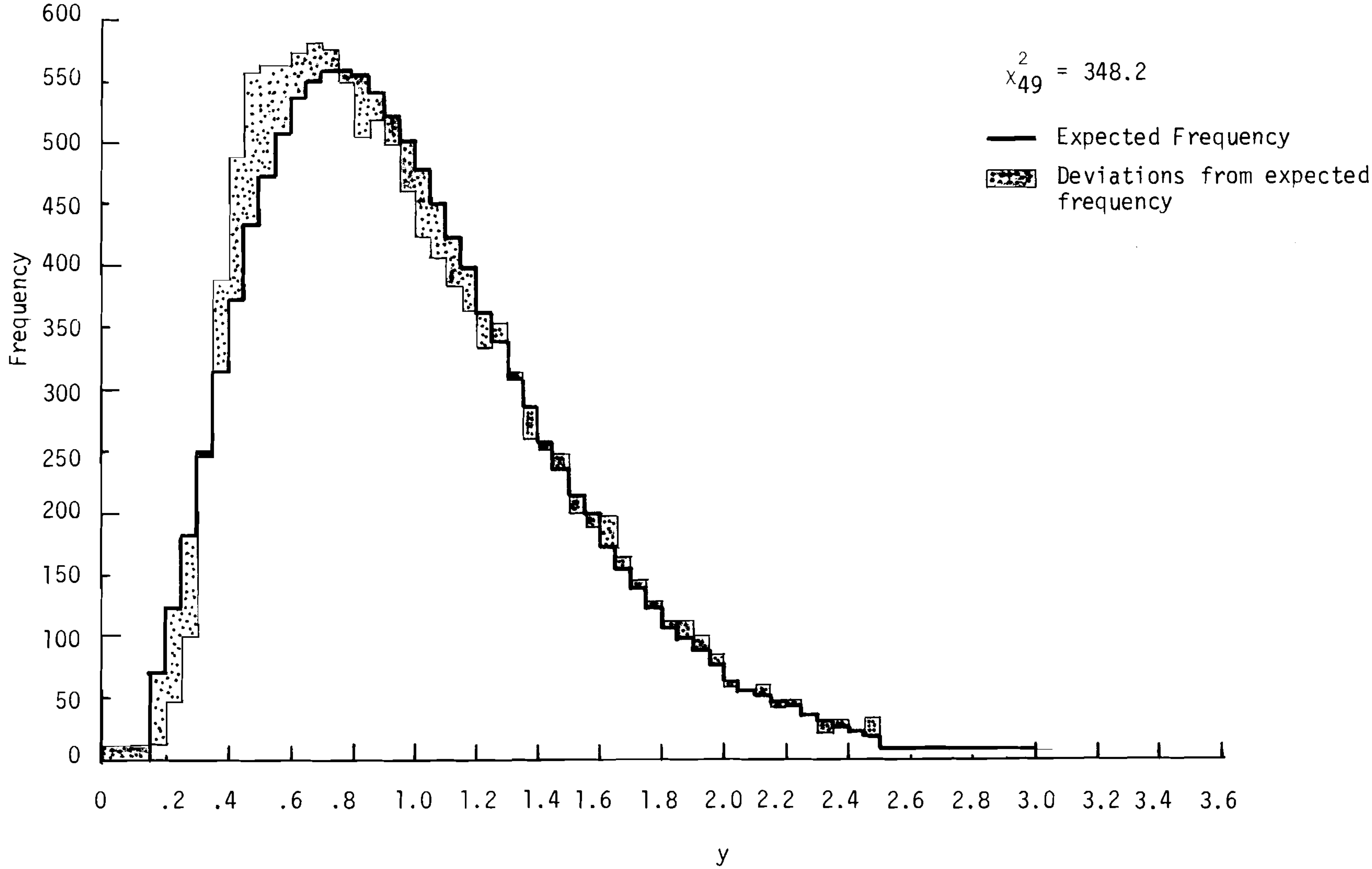


Figure 7. Observed and Expected Frequencies of Generated Gamma Variates for C = .5 Using RANDU1 (Sample 1 from Table 8)



APPENDIX

LISTING OF COMPUTER DECKS OF PSEUDO-RANDOM NUMBER GENERATORS RANDU, RANDU1, NRAND, RN2 AND RN4

RANDU U(0,1) PSEUDO-RANDOM NUMBER GENERATOR

```
'IT  FR5  RANDU,RANDU
      SUBROUTINE RANDU(X,N)                                             RANDU
C-------- ---                                                           RANDU
C     GENERATION OF UNIFORMLY DISTRIBUTED PSEUDO-RANDOM NUMBERS.        RANDU
C------------                                                           RANDU
C     METHOD                                                            RANDU
C        CONGRUENCE SCHEME WITH PERIOD 2**27                            RANDU
C------------                                                           RANDU
C     ARGUMENT DEFINITION                                               RANDU
C        X IS THE ARRAY OF RANDOM NUMBERS (OUTPUT)                      RANDU
C        N IS THE NUMBER OF RANDOM NUMBERS DESIRED (INPUT)              RANDU
C        ALSO, UPON ENTRY, X(1) IS THE INITIAL NUMBER FROM WHICH        RANDU
C        ALL OTHERS ARE GENERATED.  THIS VALUE IS REPLACED BY THE SECONDRANDU
C        RANDOM NUMBER.  IN GENERAL, X(I) IS THE I+1-TH RANDOM NUMBER   RANDU
C        GENERATED BY RANDU.                                            RANDU
C-------- ---                                                           RANDU
C     OTHER SUBPROGRAMS REFERENCED                                      RANDU
C        NRAND SLEUTH FUNCTION SUBPROGRAM--GENERATION OF RANDOM NUMBERS RANDU
C              ON THE INTERVAL (0,2**27)                                RANDU
C-------- ---                                                           RANDU
C     REFERENCE                                                         RANDU
C        M. ABRAMOWITZ AND I. STEGUN, HANDBOOK OF MATHEMATICAL FUNCTIONSRANDU
C           WITH FORMULAS, GRAPHS, AND MATHEMATICAL TABLES.             RANDU
C           U.S. DEPT. OF COMMERCE, NATIONAL BUREAU OF STANDARDS,       RANDU
C           APPLIED MATHEMATICS SERIES NO. 55, 1965.                    RANDU
C------------                                                           RANDU
      DIMENSION X(N)                                                    RANDU
      DATA T1,T2/8192.,16384./                                          RANDU
      J=X(1)                                                            RANDU
      K=J                                                               RANDU
      I=NRAND(J,K)                                                      RANDU
      DO 1 I=1,N                                                        RANDU
    1 X(I)=NRAND(J,K)/T1/T2                                             RANDU
      RETURN                                                            RANDU
      END                                                               RANDU
```

RANDU1 U(0,1) PSEUDO-RANDOM NUMBER GENERATOR

```
'IT  FR5  RANDU1,RANDU1
      SUBROUTINE RANDU1(X,N,J,K)
      DIMENSION X(N)
      DATA T1,T2/8192.,16384./
    2 X(1)=NRAND(J,K)/T1/T2
      RETURN
      END
'IT  ASB  NRAND,NRAND
. RANDOM NUMBER GENERATOR                                          NRAND
. CALLING SEQUENCE         NR = NRAND(J,I)                         NRAND
.                                                                  NRAND
        NBPR    .              REGISTER NAMES AND J FACTORS        NRAND
NRAND*  L       A1,*0,B11      LOAD FIRST ARGUMENT J               NRAND
        L       A2,*1,B11      LOAD SECOND ARGUMENT I              NRAND
        A       A1,A2          ADD ARGUMENTS J+I                   NRAND
        A       A2,(0576167244) ADD CONSTANT B                     NRAND
        L       A3,A2          LOAD A3 WITH I+B                    NRAND
        TOP     A2,(01000,0)   TEST IF CARRY IN BIT 27 OF A2       NRAND
        J       NOCARY         NO CARRY                            NRAND
        A,M     A1,1           CARRY IN A2 - ADD ONE TO A1         NRAND
        XOR     A2,(01000,0)   ELIMINATE CARRY OF A2 IN A3         NRAND
NOCARY  L       A2,A1          LOAD A2 WITH J+I                    NRAND
        TEP     A1,(01000,0)   TEST IF CARRY IN BIT 27 OF A1       NRAND
        XOR     A1,(01000,0)   YES - ELIMINATE CARRY OF A1 IN A2   NRAND
        DSC     A1,36          A1=J+I WITH ZEROES IN BITS 27-35    NRAND
        DSC     A2,36          A2=I+B WITH ZEROES IN BITS 27-35    NRAND
        L       A0,A1          LOAD FUNCTION ANSWER                NRAND
        S       A1,*0,B11      STORE A1 INTO ARGUMENT J            NRAND
        S       A2,*1,B11      STORE A2 INTO ARGUMENT I            NRAND
        J       3,B11          RETURN                              NRAND
        END     .                                                  NRAND
```

RN2 U(0,1) PSEUDO-RANDOM NUMBER GENERATOR

```
'IT   ASM  RN2,RN2
B11    EQU    11                     .  SUBROUTINE RN2(U,FU)
A0     EQU    12                     .     U IS A RANDOM 35 DIGIT INTEGER
A1     EQU    13                     .     FU FLOATING POINT NUMBER RANDOM ON (0,1)
A2     EQU    14                     .         INITIALLY U MUST BE FURNISHED AFTER
XM     EQU    15                     .         THAT IT IS KEPT CURRENT BY RN2
.                                    . SUGGESTED INITIAL VALUE IS
.                                    .      U=011060471625    (OCTAL)
RN2*   LUF    A0,*0,B11              .  LOAD U
       SSC    A1,27                  .  MULTIPLY BY 2**9
       JP     A1,$+3                 . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                  .
       AA     A1,*0,B11              .  ADD U
       JP     A1,$+3                 . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                  .
       A,XM   A1,1                   .  ADD 1
       JP     A1,$+3                 . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                  .
       SA     A1,*0,B11              .  SAVE NEW U
       AND    A1,(0777777777)        .  FLOAT U
       AA     A2,TWO                 .
       FA     A2,TWO                 .
       SA     A2,*1,B11              .  RETURN FU
       J      3,B11                  .
TWO    +      0200000000000          .
       END                           .                                       END-
```

RN4 U(0,1) PSEUDO-RANDOM NUMBER GENERATOR

```
'IT   ASM   RN4,RN4
. REFERENCE      RICHARD KRONMAL, EVALUATION OF A PSEUDORANDOM NORMAL NUMBER
.                GENERATOR  JACM VOL 11,NO 3 (JULY 1964) PP 357-363
B11    EQU    11                    .  SUBROUTINE RN4(U1,U2,FU1,FU2)
A0     EQU    12                    .     U-S ARE  RANDOM 35 DIGIT INTEGERS
A1     EQU    13                    .     FU-S FLOATING POINT NUMBERS RANDOM ON (0,1)
A2     EQU    14                    .         INITIALLY U-S MUST BE FURNISHED AFTER
XM     EQU    15                    .         THAT THEY ARE KEPT CURRENT BY RN4
.                                   . SUGGESTED INITIAL VALUES IF NONE ARE AVAILABLE
.                                   . U1=233362477003 AND U2=212312312323  (OCTAL)
RN4*   LA     A1,*0,B11             .  LOAD U1
       LSSL   A1,12                 .  MULTIPLY BY 2**12
       JP     A1,$+3                . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                 .
       AA     A1,*0,B11             .  ADD U1
       JP     A1,$+3                . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                 .
       A,XM   A1,1                  .  ADD 1
       JP     A1,$+3                . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                 .
       SA     A1,*0,B11             .  SAVE NEW U1
       JP     A1,$+3                . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                 .
       AND    A1,(0777777777)       .  FLOAT U1
       AA     A2,TWO                .
       FA     A2,TWO                .
       SA     A2,*2,B11             .  RETURN FU1
       LA     A1,*1,B11             .  LOAD U2
       LSSL   A1,7                  .  MULTIPLY BY 2**7
       JP     A1,$+3                . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
       AND    A1,(0377777777777).
       LA     A1,A2                 .
       AA     A1,*1,B11             .  ADD U2
```

```
        JP     A1,$+3               . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
        AND    A1,(0377777777777).
        LA     A1,A2                .
        AA     A1,C                 .  ADD C
        JP     A1,$+3               . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
        AND    A1,(0377777777777).
        LA     A1,A2                .
        SA     A1,*1,B11            .  SAVE NEW U2 AND REPEAT AGAIN
        LSSL   A1,7                 .  MULTIPLY BY 2**7
        JP     A1,$+3               . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
        AND    A1,(0377777777777).
        LA     A1,A2                .
        AA     A1,*1,B11            .  ADD U2
        JP     A1,$+3               . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
        AND    A1,(0377777777777).
        LA     A1,A2                .
        AA     A1,C                 .  ADD C
        JP     A1,$+3               . TEST FOR OVERFLOW AND RETAIN A1(MOD(2**35))
        AND    A1,(0377777777777).
        LA     A1,A2                .
        SA     A1,*1,B11            .  SAVE NEW U2
        AND    A1,(0777777777)      .  FLOAT U2
        AA     A2,TWO               .
        FA     A2,TWO               .
        SA     A2,*3,B11            .  RETURN FU2
        J      5,B11                .
TWO     +      0200000000000        .
C       +      0311564570652        .  =.788*2**35 (DECIMAL)
        END                         .
```

END-

BNWL 1743
Special Distribution
in Category UC-32

DISTRIBUTION

OFFSITE

1 AEC Chicago Patent Group

G. H. Lee

3 AEC Division of Biomedical & Environmental Research

Washington, D.C.

H. Hollister
C. L. Osterbery
J. Swinebroad

3 AEC Technical Information Center

1 Oak Ridge National Laboratory

R. O'Neil

1 University of Washington

R. J. Kronmal

ONSITE

1 AEC/RL Patent Attorney

R. M. Poteat

2 AEC Richland Operations Office

W. E. Lotz

60 Battelle-Northwest

T. M. Beetle
G. E. Driver
L. L. Eberhardt (5)
J. C. Fox
R. O. Gilbert (40)
W. T. Hinds
R. L. Hooper
R. J. Olson
C. A. Oster
R. Sauer
J. M. Thomas
B. E. Vaughan
B. B. Vinson
Technical Information (3)
Technical Publications