1980

Quality Evaluation of Some Combinations of Unit Uniform Random Number Generators and Unit Normal Transformation Algorithms.

Win-yeu Winnie Chen

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/3554

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
CHEN, WIN-YEU WINNIE

QUALITY EVALUATION OF SOME COMBINATIONS OF UNIT UNIFORM RANDOM NUMBER GENERATORS AND UNIT NORMAL TRANSFORMATION ALGORITHMS

The Louisiana State University and Agricultural and Mechanical Col. PH.D. 1980

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs ______
2. Colored illustrations ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Print shows through as there is text on both sides of page ______
6. Indistinct, broken or small print on several pages √
7. Tightly bound copy with print lost in spine ______
8. Computer printout pages with indistinct print √
9. Page(s) ______ lacking when material received, and not available from school or author
10. Page(s) ______ seem to be missing in numbering only as text follows
11. Poor carbon copy ______
12. Not original copy, several pages with blurred type √
13. Appendix pages are poor copy ______
14. Original copy with light type ______
15. Curling and wrinkled pages ______
16. Other ____________________________________________________________
Quality Evaluation of Some Combinations of Unit Uniform Random Number Generators and Unit Normal Transformation Algorithms

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

The Department of Quantitative Methods

by

Win-yeu Winnie Chen
B.A., National Taiwan University, 1964
M.B.A., University of Cincinnati, 1973
M.S., Louisiana State University, 1979
M.S., Louisiana State University, 1980
December, 1980
ACKNOWLEDGEMENTS

The author wishes to express her gratitude to Dr. Roger L. Burford and other professors of the Department of Quantitative Methods for their inspiration and advice which were generously given during this study. The author wishes to thank the director and staff of the Computer Research Center of Louisiana State University who provided the computer facilities used to generate the results of the study. The author wishes to express her particular thanks to Dr. Prentiss E. Schilling and Dr. Kenneth L. Koonce of the Department of Experimental Statistics who generously provided the CRT facilities used to complete this manuscript. Finally the author also wishes to express her special appreciation to Dr. John T. Marshall of the Department of Physics and Astronomy who gave her valuable help to solve some difficult problems in the spectral analysis section.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................. ii

ABSTRACT .......................................................... vii

Chapter

<table>
<thead>
<tr>
<th>Chapter</th>
<th>INTRODUCTION</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Objective and Justification of the Study</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Scope and Limitation of the Study</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Generators and Transformation Algorithms Investigated</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Criteria and Methodologies for Evaluation and Comparison</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Null Hypothesis, Significance Level, and Sample Size</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Preview of Conclusions</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Outline of the Study</td>
<td>12</td>
</tr>
</tbody>
</table>

| II.     | REVIEW AND SELECTION OF STATISTICAL TESTS | 14  |
|         | Goodness-of-Fit Tests                     | 15  |
|         | Introduction                              | 15  |
|         | Literature Review and Statistics Selection | 16 |
|         | Independence and Randomness Tests         | 36  |
|         | Introduction                              | 36  |
|         | Literature Review and Statistics Selection | 38 |

| III.    | THE CHI-SQUARE GOODNESS-FIT-TEST          | 47  |
|         | Introduction                              | 47  |
|         | Procedures of CS Test for Goodness of Fit | 49  |
|         | Empirical Test Results, Discussions and Evaluations | 56 |

<p>| IV.     | TESTS OF RUNS                             | 61  |
|         | Introduction                              | 61  |
|         | Runs Up and Down                          | 63  |
|         | Introduction                              | 63  |
|         | Procedures for Testing Runs Up and Down   | 66  |
|         | Runs Above and Below the Median (Mean)    | 78  |
|         | Introduction                              | 78  |</p>
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.</td>
<td>SERIAL CORRELATION TESTS BY AUTOCORRELATION ANALYSIS</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>The Autocovariance Function and the Autocorrelation Function</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Estimate of Autocovariance Function</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Sample Autocovariance Function</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Properties of COV(U)</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Procedures of Autocorrelation Analysis</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Number of Significant Autocorrelation Coefficients</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Portmanteau Chi-Square Statistic</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Empirical Test Results, Discussions and Evaluations</td>
<td>112</td>
</tr>
<tr>
<td>VI.</td>
<td>SERIAL CORRELATION TEST BY SPECTRAL ANALYSIS</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>Spectral Analysis</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>Power Spectrum</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>Smoothed Sample Spectral Estimators</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>Procedures of Spectral Analysis</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>Number of Significant Sample Spectra</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>Mean of Sample Spectra</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>Sample Variance of Sample Spectra</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>Normalized Sample Variance of Sample Spectra</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>Empirical Test Results, Discussions and Evaluations</td>
<td>157</td>
</tr>
<tr>
<td>VII.</td>
<td>SUMMARY AND CONCLUSION</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td>BIBLIOGRAPHY</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>Appendix</td>
<td>page</td>
</tr>
<tr>
<td>A.</td>
<td>UNIT UNIFORM RANDOM NUMBER GENERATORS</td>
<td>190</td>
</tr>
<tr>
<td>B.</td>
<td>NORMAL TRANSFORMATION ALGORITHMS</td>
<td>193</td>
</tr>
<tr>
<td>C.</td>
<td>DEFINITIONS OF STATISTICS INVESTIGATED FOR GOODNESS OF FIT</td>
<td>196</td>
</tr>
<tr>
<td>D.</td>
<td>MATHEMATICAL RELATIONS: POWER SPECTRUM AND AUTOCOVARIANCE</td>
<td>200</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>E. SAMPLE SPECTRUM AND SAMPLE AUTOCOVARIANCE</td>
<td>203</td>
<td></td>
</tr>
<tr>
<td>F. SMOOTHED SAMPLE SPECTRUM AND SAMPLE AUTOCOVARIANCE</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td>G. RELATIONS BETWEEN THE VARIANCE AND THE TRUNCATION POINT</td>
<td>207</td>
<td></td>
</tr>
<tr>
<td>H. PROPERTIES OF SAMPLE SPECTRAL ESTIMATORS: WHITE NOISE SERIES</td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>I. DISTRIBUTION OF THE SMOOTHED SAMPLE SPECTRAL ESTIMATORS</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>J. SYSTEM FLOW CHARTS, BLOCK DIAGRAMS AND PROGRAMS</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>VITA</td>
<td>270</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Goodness of Fit to Desired Distributions</td>
<td>59</td>
</tr>
<tr>
<td>2.</td>
<td>Runs up and Down</td>
<td>87</td>
</tr>
<tr>
<td>3.</td>
<td>Runs Above and Below the Median</td>
<td>90</td>
</tr>
<tr>
<td>4.</td>
<td>Portmanteau Test and Distribution of Lags</td>
<td>115</td>
</tr>
<tr>
<td>5.</td>
<td>Distribution of 50 Chi-squares from Dist. of Sig. Lags</td>
<td>119</td>
</tr>
<tr>
<td>6.</td>
<td>Distribution of 50 Chi-squares from Dist. of Sig. Lags</td>
<td>121</td>
</tr>
<tr>
<td>7.</td>
<td>Distribution of 50 Chi-squares from Dist. of Sig. Lags</td>
<td>120</td>
</tr>
<tr>
<td>8.</td>
<td>Means and Distribution of Sample Spectra</td>
<td>160</td>
</tr>
<tr>
<td>9.</td>
<td>Spectral Analysis by Variance and N. Variance</td>
<td>163</td>
</tr>
<tr>
<td>10.</td>
<td>Distribution of 50 Chi-squares from Dist. of Sig. Spectra</td>
<td>164</td>
</tr>
<tr>
<td>11.</td>
<td>Distribution of 50 Chi-squares from Dist. of Sig. Spectra</td>
<td>165</td>
</tr>
<tr>
<td>12.</td>
<td>Number of Rejections by Statistical tests by Sig. Levels</td>
<td>167</td>
</tr>
<tr>
<td>13.</td>
<td>Number of Rejections for All Levels by Statistical Tests</td>
<td>170</td>
</tr>
<tr>
<td>14.</td>
<td>Total Numbers of Rejections and Ranks</td>
<td>173</td>
</tr>
</tbody>
</table>
ABSTRACT

Due to the wide applications of normal random variates and the fact that applications are usually affected by their accuracy, an extensive study of the quality evaluation of unit uniform random number generators and unit normal transformation algorithms that are applications specific is recommended. In most simulation applications it is implicitly assumed that if any "good" unit uniform random number generator is used in combination with any "good" transformation algorithm. The resulting random variates will have good properties of randomness and will have the desired fit. The results of this research shows that this assumption is false. Use of a good, theoretically exact transformation in conjunction with a theoretically good unit uniform generator does not necessary guarantee that the resulting variates have good statistical properties. In addition, a generator is good for one type of transformations, but it is not necessarily good for the other type of transformations. Hence, the simulation practitioner must be cautious in the selection of both the unit uniform generator and the transformation algorithm to be used.
This study is concerned with the empirical study of quality evaluations of several different combinations of unit uniform random number generators and unit normal transformation algorithms for practical applications, and with procedures for determining if a given generator or a given combination is "good" or "bad". Major emphasis of this study is placed upon selection of the appropriate statistical tests for this evaluation, and development of the pertinent procedures as well as the corresponding computer programs for these tests. The statistical tests selected as appropriate for evaluating and comparing the quality of various generators are: chi-square goodness of fit tests for testing the desired distribution fits and other goodness of fits procedures for use in conjunction with other tests, runs up and down, runs above and below the median, autocorrelation tests, and spectral tests for randomness and independence. The corresponding procedures and the computer programs are set up for each test.

Seven widely used unit uniform random number generators (ADRAND, RANDU, L & L, M & M, R. Shore, URAND and GGUBS) and three well known normal transformation algorithms (Box-Muller partial inverse, Box-Muller rejection, and Hasting inverse routine) are investigated here. For each generator and each combination, a large sample size series is generated which contains 50 independent samples with 50 different
initial seed numbers. Each sample contains 50 sets, and each set contains 1200 generated numbers. Each selected generator and each selected combination is assessed by statistically testing the quality, both locally and globally, of the corresponding generated sequence. In total, we empirically evaluate (at significance levels 0.01, 0.05 and 0.2 respectively), the quality of seven generators, twenty-one combinations, and two IBM IMSL library unit normal generators: GGNML and GGNPM.

In addition, the statistics selected and the procedures developed as well as the corresponding computer programs established here are not restricted to use in this study. They can also be applied in assessing other generators and other combinations.
Chapter I
INTRODUCTION

1.1 OBJECTIVE AND JUSTIFICATION OF THE STUDY

Since the early forties, Monte Carlo methods\(^1\) have become more and more useful tools in numerical analysis and in simulation. This is largely due to the rapidly increasing use of electronic computers. Random numbers represent an indispensable ingredient in all Monte Carlo work, and in most cases, random numbers are needed in great quantities. In many applications, random variables with normal distribution are of interest, and the applications are usually affected by the quality of the random variates (Franklin, 1963). Consequently, there exists the problem of how to obtain reliable normally distributed random numbers for these applications.

People have searched for efficient ways to obtain random numbers in computers. Unit uniform random numbers generated deterministically in computers have worked quite well in nearly every application. So far, there are no algorithms

\[-----------------------------\]

\(^{1}\)This is a general term used to describe any algorithm that employs random numbers
to derive unit normal random numbers directly. If these are to be generated, it is first necessary to generate unit uniform random numbers then to use transformation algorithms to transform unit uniform random numbers to unit normal random numbers.

Many authors (Green, Smith and Klem, 1959; Maclaren and Marsaglia, 1965; Downham and Roberts, 1967; Whittlesey, 1967; Burford, 1973 and 1975; Golder, 1976; Fishman and Moore, 1978, etc.) have contributed to the study of various methods for generating unit uniform random sequences, and they have also developed a number of tests of generator effectiveness, randomness, and uniformity. Based on purely unit uniform random numbers, many authors (Box and Muller, 1958; Muller, 1959; Marsaglia and Bray, 1964; Golder and Settle, 1976, etc.) have developed a great number of theoretically good transformation algorithms for transforming unit uniform random deviates into unit normal random numbers. Hence, it appears that many accessible, and efficient algorithms for generating unit uniform random numbers as well as for transforming them into unit normal random numbers have been developed.

The transformation algorithms developed are, however, based on the idea that the unit uniform random numbers used are truly unit uniform random deviates; the numbers gener-
ated from computer generators are never truly random. Consequently, a study of the quality of combinations of unit uniform random number generators and unit normal transformation algorithms is necessary. Few studies of the combinations of generators and transformations, especially experimental applications, have been accomplished. Although some authors (Muller, 1959; Maclaren and Bray, 1964; Kinderman and Ramage, 1974 and 1976; Golden and Settle, 1976; Atkinson and Pearce, 1976; Burford and Willis, 1978, etc.) have recently been concerned with the comparison of several different unit uniform random number generators in conjunction with some transformation algorithms, they have either focused on the study of timing comparison, or they have concentrated on the study of a few generator combinations. Also their tests have been quite limited in scope. A very important consideration in evaluating random number generators is the quality of output. A more extensive study of the quality of combinations that are application specific was recommended by Kinderman and Ramage (1974).

This study is concerned with empirical quality evaluations of different combinations of unit uniform random number generators and unit normal transformation algorithms for practical applications, and with determining if a generated sequence is "good" or "bad". Major emphasis is placed upon selection of the appropriate statistical tests for this
evaluation, and development of pertinent procedures as well as of the corresponding computer programs for these tests. The statistics selected, the procedures developed and the corresponding computer programs set up here are not restricted to this study. They can also be applied to assessing other generators and other combinations.

1.2 **SCOPE AND LIMITATION OF THE STUDY**

1.2.1 **Generators and Transformation Algorithms Investigated**

Three basic congruential methods have been developed for generating pseudo-random numbers: additive congruential method, multiplicative congruential method, and mixed congruential method.\(^2\) There are numerous unit uniform generators that have been proposed. In this study, we investigate seven well-known and wide uses uniform generators which have already been proved good in terms of simplicity, accuracy, and speed. These are ADRAND, RANDU, L & L, M & M, R.Shore, URAND\(^3\), and one of the IBM IMSL library subroutines\(^4\) - GGUBS.

\(^2\)See Appendix-A.

\(^3\)ADRAND, an additive congruential generator, proposed by R. Burford; RANDU and L & L, multiplicative congruential generator, contained in the IBM SSF (Scientific Subroutine Package) library subroutine and developed by Lewis and Larmouth, respectively; M & M, R.Shore, and URAND, mixed congruential generator, proposed by Marsaglia and Marsaglia, R. Shore, and improved RANDU, respectively.

\(^4\)This generator is derived from the package LIRANDOM: L & L generator, of the Naval Postgraduate School in Monterey, California. Cosmetic changes and those necessary to cause
There are also three basic transformation methods to transform unit uniform random numbers into unit normal variates. They are the inverse transformation method, the rejection method, and the composition method. In this study, three well-known theoretically good unit normal transformation algorithms: Box-Muller partial inversion, Box-Muller rejection, and Hastings' inversion routines have been used in conjunction with the seven generators mentioned in the previous paragraph. Hence, in this study, we empirically evaluate the quality of seven generators, twenty-one combinations, and two IBM IMSI library unit normal generators: GGNML and GGPNM.  

1.2.2 Criteria and Methodologies for Evaluation and Comparison

Numbers generated by computers are based on mathematical relations (congruential methods) and are completely determined by the starting data. They are referred to as pseudo random numbers and are not truly random. However, so long as pseudo random numbers can pass a set of appropriate statistical tests implied by the randomness hypothesis, these pseudo random numbers can, for practical applications, be fitting into the IMSI library framework are the only changes made.

*See Appendix-B.

See Appendix-B.
treated as truly random even if they are not. Although the period of the sequence is one of the important criteria for evaluating the quality of random number generators, for practical purpose, the chances are that the sequence never will repeat. This by no means guarantees that the sequence will be useful in applications. Hence, for applications, the question is how to decide whether or not a sequence is satisfactory. There are two ways in which a sequence of pseudo unit uniform random numbers and unit random normal deviates can be deemed unsatisfactory for applications. The first is that the sequence may not constitute a sufficiently random sample from that distribution, and the second is that its distribution can fail to be an adequate representation of the desired distribution (Knuth, 1969). In the quality evaluation, criteria will be directed towards these two aspects in order to assess the suitability of various combinations for applications.

Since we can not judge by inspection whether or not a sequence of numbers is sufficiently random and has a desired distribution, some unbiased and objective tests must be applied. The theory of statistics provides several quantita-

7The desired distributions, in this study, are the unit uniform distribution and the standard normal distribution produced by a good unit uniform random number generator and a good combination of the unit uniform random number generator and the unit normal transformation, respectively.
tive tests for randomness and goodness of fit to a particular distribution (Knuth, 1969, p. 35).

Every series which is to be extensively used should be tested carefully. Two kinds of tests are distinguished: empirical tests and theoretical tests. Theoretical tests are computed over the entire cycle or period of the sequence. Full period of sequence is rarely used in practice, so while theoretical tests are desirable when possible, empirical tests are more convincing (Lewis, 1975, p. 35). This study emphasizes applications, thus empirical tests are used for evaluating the quality of the selected generators and the selected combinations by means of testing the generated sequence from them.

There is literally no end to the number of tests that can be conceived. We, therefore, must select and develop procedures for the tests which have proved to be powerful, instructive, and easily adapted to computer calculation. These will be used as objective methodologies for evaluating the

---

8Empirical test are conducted by means of the computer manipulating groups of numbers of sequences and evaluating certain statistics. Theoretical tests establish characteristics of the sequence by using number theoretical methods based on the recurrence value used to from the sequence.

9"Combination" is used throughout this study as an abbreviation of the combination of the unit uniform random number generator and the unit normal transformation algorithm.
quality of combinations. Also, for simplicity, we use the same statistics and the same procedures for evaluating the quality of the selected unit uniform random number generators. Hence we must select the statistics and develop the procedures as well as the corresponding computer programs which are suitable for both generators and combinations.

Based on their validity, simplicity, generality, wide usage, as well as the theoretical knowledge, the appropriate statistical methodologies and tests selected for evaluating and comparing the quality of the output in this study are as follows:

1. Tests of goodness of fit for the desired distributions.\(^1\)

   a) Pearson's Chi-square goodness of fit test.

2. Tests for randomness and independence.\(^1\)

\(^{10}\)Chi-square goodness of fit tests are also used extensively in the remaining test procedures.

\(^{11}\)Kendall and Stuart (1961) have noted that in some sense "good" independence tests are "good" randomness tests because the randomness hypothesis \(H_0: F_1=F_2=\ldots=F_n\) is equivalent to the hypothesis of independence of values and chronological order. More precisely, the univariate sample \(X_1, X_2, \ldots, X_n\) (where \(X_i\) has pdf \(f_i\)) can be represented by the bivariate sample \((1, X_1), (2, X_2), \ldots, (n, X_n)\) which, under \(H_0\), is from a bivariate population with \(X_i\) independence of \(i\).

Wolfowitz (1943) stated that to test whether \(x_1, x_2, \ldots, x_n\) is a "random"sequence means to test the hypothesis that \(X_1, X_2, \ldots, X_n\) are inde-
a) Runs up and down.
   i) Total number of runs.
   ii) The number of runs by length.

b) Runs above and below the median.
   i) Total number of runs.
   ii) The number of runs by length.

c) Autocorrelation analysis.
   i) Number of significant autocorrelation coefficients.
   ii) Portmanteau chi-square statistic.

d) Spectral analysis.
   i) Number of significant sample spectra.
   ii) Sample mean of sample spectra.
   iii) Sample variance of sample spectra.
   iv) Normalized sample variance of sample spectra.

---

pendently distributed and have identical distribution functions.
If a sequence behaves satisfactorily for applications with respect to tests T1, T2, T3, ..., TN, we can not be sure in general that it will not be an unpleasant failure when it is subjected to a further test TN+1. Yet each test gives us more and more confidence in the randomness of the sequence. In practice, we apply about half a dozen different kinds of statistical tests to a sequence, and if the sequence passes these tests satisfactorily we consider it to be random and to have the desired distribution for applications (Knuth, 1969).

1.2.3 Null Hypothesis, Significance Level, and Sample Size

Following the traditional approach, an objective test is made of an hypothesis about randomness, independence or some desired distributions. The testing process is roughly composed of five parts: null hypothesis (H0), test statistic, significance level (ALPHA) and sample size (N), sampling distribution, and conclusion. In this study the null hypothesis is a hypothesis of randomness, independence or some desired distribution. It states that the series is random, independent, and that the series is normally or unit uniformly distributed. Further for the goodness of fit test, the null hypothesis is a simple hypothesis with the known parameters.
Three levels of significance: 0.01, 0.05 and 0.2 will be used in this study. The 0.01 and 0.05 levels are a common value of ALPHA, and the 0.2 level is a stringent level in terms of the probability of accepting a false hypothesis, and for power computations. In order to obtain more powerful results from the statistical tests stated in subsection 1.2.2 and to reduce the probability of either type I errors, ALPHA, or type II errors, BETA, we use large samples for evaluating the generators and the combinations in this study. The proper choice of a sample size is somewhat obscure. Large values of $N$ will tend to smooth out locally nonrandom behavior, and such behavior is an undesirable characteristic that is of significant importance in most computer applications of random numbers. This makes a case for small values of $N$. A compromise may be taken by dividing a long series into $KL$ sets with $N$ numbers per set, where $N$ is relatively small; then testing both the local and the global characteristics of the series (Kunth, 1969). In addition, since the pseudo random numbers generated from computers are determined by the initial seed numbers, we may reduce this influence by generating ISET independent samples using ISET different initial seeds.

In this study, we investigate the quality of the generators and the combinations by generating 50 independent samples (ISET=50) as a whole series. Each sample contains 50
sets (KL=50), and each set contains 1200 numbers (N=1200). Next, we test the characteristics of each set, each sample, and the whole series. Finally, we use the test results to evaluate the quality of the generators and the combinations.

1.3 PREVIEW OF CONCLUSIONS

In most simulation applications it is implicitly assumed that if any "good" unit uniform random number generator is used in combination with any "good" transformation algorithm, the resulting random variates will have good properties of randomness and will have the desired fit. The results of this paper will likely show this assumption to be false more often than true. Thus if it is to be true the simulation practitioner must be cautious in the selection of both the unit uniform generator and the transformation algorithm to be used.

1.4 OUTLINE OF THE STUDY

The formal presentation of the study is divided into seven parts. Chapter one presents the objective and the justification of the study, the scope and the limitations of the study, a preview of conclusions, to the study and the outline of the study. Chapter two presents a review and selection of statistical tests for the goodness of fit, inde-
pendence, and randomness tests. Five statistical tests used as objective methodologies for quality evaluations and comparisons for generators/combinations are presented in Chapter three through Chapter six. Chapter three presents the selected goodness of fit tests (the chi-square goodness of fit test) for the distribution and for use in conjunction with other tests. Chapter 4 presents two selected randomness and independence tests—runs up and down, and runs above and below the median. The other two selected randomness and independence tests—autocorrelation analysis and spectral analysis, are described in chapters five and six, respectively. Each of these four chapters (Chapter 3-6) are composed of three parts. The theories of the tests are presented in Part I; the procedures for the tests and the test results are exhibited in Part II and Part III, respectively. Chapter seven gives a summary of the results obtained from each experimental statistical test and a statement of the evaluation and the conclusions drawn from this study.
Chapter II
REVIEW AND SELECTION OF STATISTICAL TESTS

Since numbers generated from the computer are based on mathematical congruential relations, they are not truly random. However, for practical applications, in general, the generated numbers should be sufficiently random and independent, and should fit well to their desired distribution. Thus the quality evaluations of the selected generators and the selected combinations in this empirical study must be based on objective statistical tests to assess the goodness of fit to the desired distribution, randomness and independence of their corresponding generated numbers.

There are numerous statistical goodness of fit, randomness, and independence tests in the literature. However, in most cases universally accepted no method of constructing a good test is known (Ievene, 1952). In addition it is not usually possible to say that one test is more powerful than another, unless the set of alternative hypotheses is sufficiently specified. In this study, however, the alternatives are vague and unclear. Thus the purpose of this chapter is to investigate and to compare the powers of the goodness of fit, randomness, and independence tests if they are availa-
ble, and to study the behavior and the characteristic of these tests. Then we select several tests which have been shown to be reliable and which seem to be the most suitable for the purpose of this study. Based on validity, simplicity, generality, and wide usage, the tests we selected are the Pearson chi-square goodness of fit statistic for the desired distribution and other goodness of fit tests in connection with other tests; runs up and down, runs above and below the median for random oscillatory tests; autocorrelation analysis and spectral analysis for sequence independence tests.

2.1 GOODNESS-OF-FIT TESTS

2.1.1 Introduction

A good unit uniform random number generator, or a good combination of a unit uniform random number generator and unit normal transformation algorithm, should generate numbers which have a unit uniform distribution, and a unit normal distribution, respectively. Thus one criterion for the evaluation of quality of the selected generators and the selected combinations is to assess the generated sequence for their fit to the desired distribution. Several methods of goodness of fit tests may be employed to determine how well the numbers generated fit their desired distribution. Over the years, many procedures for goodness of fit problems have
been proposed. In addition some procedures have been studied for their power properties. In the next section, we investigate and compare their powers and their characteristics as presented in the literature. Then we select Pearson's chi-square (CS) goodness of fit test as most appropriate for this study.

2.1.2 Literature Review and Statistics Selection

Shapiro, Wilk and Chen (1968) made an extensive empirical sampling study of the comparative sensitivities of nine statistical tests for evaluating the normality of a complete sample. They are the standard third moment ($\mu_1$), the standard fourth moment ($\mu_2$) for standard tests,\(^{12}\) Kolmogorov-Smirnov (KS) (Kolmogrov, 1933), Pearson chi-square goodness test (CS) (1900), Cramer-Von Mises (CVM) (Cramer, 1928 and Von Mises, 1931), modified KS (D) (Durbin, 1961), and a weighted Cramer-Von-Mises (WCVM) (Anderson and Darling, 1954) for distance tests, Range/standard deviation (u) (David, et al., 1954) for ratio tests, and Shapiro-Wilk test (W) (1965)\(^{13}\) for analysis of variance tests. Forty-five alternative distributions in twelve families with five samples each of sizes (N=10, 15, 20, 35, 50) were studied. For all

---

\(^{12}\)The earliest work on testing for departure from normality is based on $\mu_1$ and $\mu_2$.

\(^{13}\)The definitions of the tests are in the Appendix C.
statistics except W, the empirical null distribution was based on M=500 samples for each sample size. For the W statistic, it was based on M=5000 for N < 20, and on M=(100,000/N) for 20 < N < 50. Alternative distributions, in general, were classified into the discrete case and the continuous case. Furthermore, the continuous case was classified into five major groups, according to values of |NB1| and B2 as follows:

Group 1: |NB1| > 0.3, B2 > 3.0 (asymmetric, long-tailed).
Group 2: |NB1| > 0.3, B2 < 3.0 (asymmetric, short-tailed).
Group 3: |NB1| < 0.3, B2 > 4.5 (symmetric, long-tailed).
Group 4: |NB1| < 0.3, B2 > 2.5 (symmetric, short-tailed).
Group 5: |NB1| < 0.3, 2.5 < B2 < 4.5 (near normal).

In their study, the authors found that the power of the test procedures depends on the nature of the alternative distributions and the sample size (in most case the sensitivity increases markedly with the sample size). From the results of Shapiro and Wilk's power comparisons, for the sample of size N=50, the powers of the statistics considered are broadly in the following order of descending powers: W > b2 > CS > D > u > KS > WCM > NB1 > C4. Some general conclusions were gleaned by the authors from the empirical test results.

---

14 The average powers of the 44 alternatives studied at a 10% level for sample size 50 are W=69.98; b2=53.74; CS=52.88; D=51.79; u=49.29; KS=43.95; WCM=43.05; NB1=36.25; CM=35.62.
The distance tests (KS, CM, WCM, D-based on the comparison of the hypothesized cumulative distribution function with the empirical distribution function) are typically inferior in sensitivity against continuous distribution alternatives with some exceptions mainly in connection with discrete alternatives in the case of the D statistic. The D statistic has exceptional performance for discrete alternatives.

The u statistic has particularly good properties against symmetric, especially short tailed (e.g. the uniform) distributions, but it seems to have no power with respect to asymmetry (highly skewed and very long tailed distributions).

\(\text{Jb1}\) is quite sensitive to most forms of skewness, but it is in general insensitive in the case of symmetric alternatives such as the uniform distribution. It is very poor in comparison with the discrete distributions. The b2 statistic performs comparatively well with finite range distributions as well as with symmetric long tailed infinite range distributions. It is, however, not effective against skewed and discrete distributions. Generally \(\text{Jb1}\) responds to skewness and b2 to kurtosis. Combinations of both \(\text{Jb1}\) and b2 usually provide a sensitive judgement (D'Agostino and Pearson, 1973; Bowman and Shenton, 1975; Pearson, D'Agostino and Bowman, 1977)
CS (based on a numerical measure of the difference between the observed and the theoretical distribution) has good power against the highly skewed distributions and reasonable sensitivity to very long tailedness. Shapiro and Wilk indicated that for sample size 50 for many alternative distributions CS is the more powerful. However, it has the disadvantages that the number and the character of class intervals used is arbitrary. All information concerning sign and trend of discrepancies is ignored and that, for small samples, the number of cells must be very small.

The $W$ statistic (obtained by the ratio of the best or approximately best linear unbiased estimator of the population variance to the sample variance) exhibits sensitivity to non-normality over a wide range of alternative distributions. For continuous alternatives, it is the only test which never has very low power where another test has high power. For discrete alternatives it compares poorly only against the results for the D test and occasionally the CS statistics. It provides what might be termed a very good omnibus test.$^{15}$

$^{15}$An omnibus test is appropriate for detecting any form of deviation from normality due either to skewness or to kurtosis. So it is used if we do not have any prior knowledge of the expected direction of departure from normality.
However, from a practical viewpoint, this procedure has disadvantages. For each sample of size $N$, a different set of appropriate weights for the order observations for the numerator of $W$ is required. They are obtained by approximation to the covariance matrix of the normal order statistics. To date, the elements of the covariance matrix of the normal order statistics are available only for samples of sizes up to 20. These were presented by Sharhan and Greenberg (1956). For large sample sizes, it is awkward to approximate or to tabulate them. Furthermore, a set of significant points is needed for each sample of size, $N$. Because $W$'s moments beyond the first, are unknown, it is difficult to find the approximate null distribution of $W$. Simulation runs seem to be the only available way to obtain the null distribution. For large sample sizes, it may be difficult to determine the percentage points of its distribution. Thus Shapiro and Wilk did not extend their tests beyond samples of size 50. They supplied the weights and percentage points required for the test up to sample of size 50. Thus, although this statistic has surprisingly good power properties for a normality departure test, it is unavailable for large sample size, say $N > 100$ (D'Agostino, 1971).

Shapiro and Wilk did not extend the $W$ test beyond the sample of size 50, D'Agostino (1971) presented a new test, $DA$ (ratio of Downton's linear unbiased estimator of the po-
pulation standard deviation to the sample standard deviation), of normality applicable for sample of sizes of 50 or larger. To investigate the power of DA, he generated from 200 to 400 random samples of size 50 for each of several alternative distributions, 42 in all. He performed a two-sided 10 percent level significance test on them. Most of the alternatives were considered by Shapiro (1968). An empirical sampling study of the comparative power of the DA test with \( W, N_b 1, b_2 \) and \( u \) indicated that the DA test compares favorably with the \( W, N_b 1, b_2, \) and \( u \).

The statistic DA is an omnibus test. It can detect deviation from normality due either to skewness or to kurtosis. It is useful when the type of deviation from normality is unknown. It maintains a good power over a wide spectrum of alternatives. It is as powerful as or more powerful than \( N_b 1 \) for symmetric alternatives and for about half of the skewed distributions considered. For about three-quarters of both the symmetric and the skewed alternatives, DA is as powerful as or more powerful than \( b_2 \). Also DA is as powerful as or more powerful than \( u \) for about two-thirds of the symmetric alternatives while it is almost always so for skewed alternatives. The main use of DA comes when \( W \) is unavailable, for example when \( N > 100 \).

\[ \text{The average powers of 28 alternative studies at a 10% level for sample size } N=50 \text{ are } W=59.82; b_2=52.18; b_1=49.89; u=46.39; N_b 1=38.32. \]
Shapiro and Francia (1972) proposed a modification of the Shapiro- Wilk' $W$ statistic ($W'$), an approximate analysis of the variance test, for testing normality. It can be used for a large sample. The coefficients $W'$ needs depend only on the expected values of the normal order statistics and they are available (Harter, 1961; Pearson and Hartley, 1972, Table 9). An empirical sampling study was conducted to compare the sensitivities of the $W$ and the $W'$ tests in detecting non-normality. One hundred samples of size $N=35$ and 50 were generated randomly from 37 alternative distributions. The alternative distributions came from 12 different families and represented a wide variety of samples including skewed and symmetric as well as continuous and discrete distributions.

The results of this study indicate that, overall, the sensitivities of the $W$ test and the approximate $W'$ test are equivalent. Any difference depends on the characteristics of the alternative distributions. In general, the $W'$ test appears to be more sensitive than the $W$ test when the alternative distribution is continuous and symmetric with a high fourth moment, when it is near normal, and when it is discrete and skewed. The two tests appear to be equivalent for alternative distributions which are continuous and asymmetric with high fourth moment as well as being discrete and symmetric. The $W$ is superior to the $W'$ test for the other alternative distributions.
The use of the $W'$ procedure for sample sizes greater than 50 was recommended, and the extension to sample sizes over 100 can readily be accomplished. However, one must be careful in such extrapolations. Pearson and D'Agostino (1977) suggested that the $W'$ test must be used with caution if there are multiple ties existing in the ordered observations.

The power of Geary's $W_n$ test, the ratio of the mean deviation to the standard deviation (Geary, 1935), was investigated by D'Agostino and Rosman (1974). The power of Geary's $W_n$ test was compared with $\sqrt{b_1}$, $b_2$, $I_A$, $W$ (for sample sizes, $N = 20, 50$), and $W'$ (for sample size, $N = 100$). For samples of size $N = 20$, $50$, and $100$, they generated 200 random samples for each alternative and performed the tests under consideration at the 0.01, 0.05, 0.1, and 0.2 levels, both one and two-sided tests. The alternatives considered were the 35 given in the Table 3 of D'Agostino (1971) along with the 21 contaminated normals used in a power study by Chen (1971). The result of this test is that there was no specific situation where Geary's $W_n$ test clearly and for practical purposes dominates all other tests of normality. However, it still has good power properties and has the merit of simplicity.

Pearson (1935) investigated the power of the $W_n$ test. He found that for a skewness test, the $\sqrt{b_1}$ test is better than
For large sample size, the bias of $h_2$ is less than $W_n$. Thus for kurtosis test, $h_2$ is better than $W_n$.

The powers of KS, CM, WCM, DA, $W$, and $W'$ were also studied by Stephens (1974). For the normality test, he found that the power of KS is the poorest among the statistics considered. However, the CM and WCM tests are more powerful than those found by Shapiro, Wilk and Chen (1968). Stephens concluded that the CM and the WCM statistics are comparable to $W$. The reason is that he used the estimated mean and variance from sample data instead of using the true mean and the variance for the distribution function.

For the DA test, Stephens found relatively poorer power than for $W$ and $W'$. The reason is that the coefficients of $W$ and $W'$ need to have good properties and good approximation, respectively. In the DA test, these coefficients were replaced by values which can easily be calculated ($C_{in}=1/2*(N+1)-i$), but the power dropped considerably.

Dyer (1974) also suggested that even if one believes that the true mean and variance are known, one should still assume them unknown for the normality test and estimate them in making the test. The results he found indicate that the powers of $W$, CM, WCM, and KS are larger when using the estimated mean and variance than when using true values.
Pearson, D'Agostino and Bowman (1977) studied the power of four omnibus tests: $K$, $R$, $W$ and $DA$, and four directional Tests: $b_1$, $b_2$, right angle and the upper tail or lower tail $DA$ tests. They also considered the $u$ test, but in all of these cases it was less powerful than the appropriate single tail $b_2$ test. Under the assumption that the likely directions of departure from normality was known, They applied these tests at the 5% level for 200 samples of $N=20, 50, \text{ and } 100$. Each sample was drawn from among 58 nonnormal populations involving 12 distributions.

From the test results, they found that the power of the tests depends on the nature of the alternative distribution and the sample size. Among four omnibus tests, the $u$ test is the most sensitive test. In addition, if prior knowledge warrants the use of the correct directional test, the directional test seems to be more powerful than any one of the four omnibus tests.

17 $K$ and $R$ are omnibus tests involving the ratios of $b_1$ and $b_2$ jointly.

18 As an omnibus test, $DA$ should be two sided, i.e. both the upper and the lower percentage points of $DA$ are needed.

19 The directional test is especially sensitive to the expected type of departure. So it is used if it is possible to specify the way in which the data are likely to depart from normality.

20 Since $DA$ can be positive or negative simulation studies suggest that it can be used as a directional test in the same way as $b_2$. Extreme positive values indicate that the population will have kurtosis with $B2 < 3$, negative values imply $B2 > 3$ (Pearson, 1977).
For skewed populations and for sample sizes $N=20$ and 50, these authors estimated four omnibus tests and two directional tests, $b_2$ and $DA$. They found, that for populations 1-7 (i.e., $B_2 < 1.8$), the omnibus $W$ test is more powerful than the $K$ and $R$ tests, and the powers of $K$ and $R$ are equivalent. For leptokutic population ( $B_2 > 3.0$), i.e. populations 11-16 and 18-27, they found that the relative powers of the omnibus tests are in the following order: $K > R > DA > W$. However, for population 17, the $W$ test is the most powerful. For long tailed populations, 28-32 (i.e., $B_2 > 3.6$), the $DA$ test turns out to be the most sensitive of the omnibus tests.

When comparing two directional tests with the omnibus tests they found that for populations 1-7, the lower tail $b_2$ test is definitely more powerful than the upper tailed $DA$ test and the other four omnibus tests. For populations 11-32, the upper tailed $b_2$ and the lower tailed $DA$ tests are more powerful than any one of the omnibus tests.

For skewed populations and for sample sizes $N=20$ and 50, the authors estimated the power of four omnibus tests and two directional tests: the upper tailed $\chi^2$ distribution and the right angle tests. They found that the $W$ test is the most sensitive of the omnibus tests considered. However, if there is reason to believe that the population is positively
skewed, \( \text{jB1} > 0 \), the test based on the upper tail of the \( \text{jB1} \) distribution, (or if \( \text{jB1} < 0 \), the test based on the lower tailed of \( \text{jB1} \) distribution) is likely to be more sensitive than any one of the omnibus tests or the right angle test.

For sample size \( N=100 \), the powers of three omnibus tests (\( K \), \( R \), and \( DA \)) and one directional test (the upper and the lower tail of \( DA \)) were studied. The authors found that if the population is symmetric and leptokurtic, the directional test using the lower tail of the \( DA \) distribution is more powerful than the omnibus tests. Further if the population is skewed, the \( K \) and \( R \) tests are more sensitive than the \( DA \) test.

From the power comparison review for the goodness of fit to normality, two directional tests (\( \text{jB1} \) and \( b2 \)) have good power properties when we have prior knowledge of the likely departure from normality for the alternative distributions (Pearson, 1977). In this study, the alternative distributions are vague and unknown. Thus even if the \( \text{jB1} \) and the \( b2 \) tests have the merit of simplicity and good power properties, they are not suitable for use for the normality tests in this study. In addition, for normality tests, if we can not find a single test sensitive both as regards skewness and kurtosis, we can apply two separate tests to detect skewness and kurtosis (Pearson, 1935). \( b2 \) is good for a
kurtosis test; however, its distribution has proved difficult to handle (Pearson, D'Agostino and Bowman, 1977). Hence suitable tables of the critical values of the null distributions is not available for N>50. Thus this test is not available for large sample size applications. Neither is the less powerful U test considered in this study since the percentage points of its null distribution are not available for N>50.

Among the omnibus tests reviewed, the W test is the most powerful statistic for a goodness of fit test for the normal distribution. Due to the difficulties discussed before, however, this test is available only up to sample of sizes of 50. Thus it is not available for the sample size of 1200 per set considered in this study.

The next most powerful omnibus test is the W' statistic. Also, due to computational difficulty, it is not feasible to use it for large sample sizes, certainly not as large as N=1200. The null distribution theory of this test is difficult; even asymptotic theory is lacking. Simulation runs, therefore, seem to be the only practical method to estimate the null distribution. For large sample sizes, it is time consuming to obtain the critical values of the null distribution. In addition, Pearson, D'Agostino, and Bowman (1977) showed how difficulty it is to apply this test when in prac-
tice the sample data may contain "ties" resulting from rounding.

Although K and R omnibus tests are compatible with the W test and show good power in Bowman and Shenton's studies (1977), due to the limitations imposed by computation problems, we can not use these omnibus tests in this study. On these two tests, a quite good approximation to the probability integrals of \( J b_1 \) and \( b_2 \) is required. Although D'Agostino and Pearson (1973) presented a chart of an empirical probability integral, \( P \), of the distribution of \( b_2 \) from a normal population, the samples are only for sizes up to 200. Thus we can not use these omnibus tests for sample size \( N=1200 \) in this study.

The KS, WCM, CM, and D statistics are inferior in sensitivity in continuous distributions according to Shapiro and Wilk (1968). However, WCM appears to have good power, and it is comparable to the W test, if the estimated mean and variance are used in making the test (Stephens, 1974; Dyer, 1974).

---------------

21The average powers of 52 alternative studies at 5% level for sample size \( N=50 \) are \( W=65.23; \) \( K =62.19; \) \( R =62.12; \) \( DA=57.42 \).
Thus, in practice, for large sample applications, WCM, DA, Wn and CS may all be considered as good omnibus statistics for the normality test. Among these four statistics, CS will be used, in this study, for the normality test and other goodness of fit tests. This is due to its robustness, simplicity, flexibility and generality, and wide use. In addition, it is theoretically well-known.\textsuperscript{22}

In term of flexibility and generality as a goodness of fit test, the CS statistic is distribution free. No assumption is required concerning the underlying distribution. The application of a chi-square test is often justified in the following manner.

1. Given any type of underlying distribution, if mutually exclusive and exhaustive categories are established and random samples are taken, then the number falling into each category will follow a multinomial distribution.

\textsuperscript{22} Generally on theoretical grounds it is possible to specify which of these tests is the most efficient and powerful. It may happen, however, that owing to mathematical difficulties in putting the ideal test into working form or to practical difficulties arising from the extent of computation involved, the statistician will choose to employ a second but simpler test (Pearson, 1935).
2. If sufficiently large samples are taken, then the multinomial approaches the multivariate normal distribution.

3. By squaring and summing standardized normal variates a chi-square distribution is obtained.

Thus the CS statistic will be used for normality tests, uniformity tests and other goodness of fit tests in this study.

In terms of simplicity, CS is the simplest of the tests considered. WCM and DA require order observations for their computation, so more computing time is required. CS does not require order observations; this will save a large amount of sorting time that would be required for the large sample used in this study. Because of its simplicity, flexibility and generality, it has become a widely used statistic for goodness of fit tests. Cochran (1952) stated this test is most commonly use when there is a lack of clear knowledge of alternative hypotheses. Thus it is suitable for this study.

Due to its popularity, more theoretical and empirical studies have been done of this statistic than have been done of WCM, Wn or DA. Under the null hypothesis, the CS statistic for large sample size has approximately a chi-square
distribution (Fisher, 1942). Thus to test the null hypothesis the upper tail of the chi-square distribution is used as the critical region. Tables of percentage points of the chi-square distribution up to degrees of freedom 100 are available. Even if degrees of freedom are greater than 100, the approximation formula \( C_S(p) = \left( Z_p + (2v - 1) * 0.5 \right) * 2 / 2 \), can be used for estimating the percentage points.\(^2\) Where \( Z_p \) is the 100th percentile of the standard normal distribution, (Pearson and Hartley, 1972). In addition, the limiting power function of the chi-square test may be approximated by a non-central chi-square distribution. It depends on two parameters: the degrees of freedom and a parameter of non-centrality.

Many authors have shown that the CS statistic is valid and robust even if the expected frequencies are small but equal in each class. The larger the sample size used the more power the CS test will provide. The size of sample determines whether the test really is a severe test of the null hypothesis. Cochran (1952) reported that the power of the CS to detect an underlying disagreement between theory and data is controlled by the size of the sample. He showed

\(^2\)D'Agostino (1971) offered the table of percentile points (.5, 1, 2.5, 5, 10, 90, 95, 97.5, 99, 99.5) of DA for samples of size, \( N = 50 (10) 100, 150 (50) 1000 \). Geary (1935) supplied the table of 10% and 5% (upper and lower) probability points of \( Wu \) at samples of size \( N = 5 (5) 50, 75 (25) 100, 500, 1000 \).
that the limiting power distribution of the CS tends to one as the sample size \( N \) becomes large. Hence the CS statistic is consistent.

Since for large sample sizes, the CS statistic is approximated by the chi-square distribution, the validity and the robustness of CS is beyond doubt for large sample sizes. However, not much is known for certain concerning the degree of approximation for a given sample size, and there is considerable difference of opinion as to the sample size required. Kendall (1952) reported that the chi-square approximation may confidently be applied when the theoretical cell frequencies are 20 or more. Cramer (1946) suggested that the expectations in the cells should be at least 10. Fisher (1941) states that they should be at least 5. Cochran (1952) recommended a compromise in which 80\% of the expectations should exceed five and the remainder should exceed one. Vesseran (1958), Good (1961), Wise (1963) and Slakter (1965) provided evidence that the chi-square distribution may give a good approximation to the distribution of the CS statistic, even when the expected frequencies are as low as 1 or 2 if they are equal. Slakter (1966) offered some further evidence on the robustness of the CS goodness of fit with respect to validity when the expected frequencies are as small as 1/2 or 1/4 but equal. Roscoe and Byars (1971) stated that excellent approximations are obtained with aver-
age expected frequencies of one or two in tests of goodness of fit to uniform, and two or three with the non-uniform tests.

Although different authors have different opinions as to the sample size required for the CS test. We can conclude that the CS statistic is valid and robust even if the expected frequencies are small but equal. In this study, for normality tests, we use \( N=1200 \) for each set. The expected values in each cell are about 30 and equal. Thus the validity of the CS test is assured for each set in this study. As for the normality tests and other goodness of fit tests for each sample and for the whole series, the sample size is 50 for each (i.e. each sample contains 50 sets, and \( ISET=50 \) samples comprise one series). The expected values of each cell are 5 and equal. According to Slakter, Wise, etc. the validity of the CS test is assured here.

Although difficulties (arbitrary choice of the number of classes and the class interval points) accompany the use of the CS test with continuous distributions, these difficulties can be solved. If we use equal probabilities for each cell, the arbitrary choice of the class interval points will be avoided. With regard to this matter, Mann, Wald (1942) and Gumbel (1943) recommended that, for computational simplicity, the expected value be the same in all classes. In
addition many authors (as stated above) have demonstrated the robustness of the CS test if the expected values are the same in each cell. Thus equal probabilities will be used in this study. As for the choice of the number of classes, Cochran (1952) suggested that a moderate number of classes, say between 10 and 25, is applied in the common practice. The paper by Mann and Wald (1942) deals with the choice of the number of classes. Williams (1950) showed that the optimum number of classes obtained by Mann and Wald is too large, and he suggested that half the optimum number of classes obtained by Mann and Wald is more suitable.

In this study, for the normality test of each set, the number of sets, $N=1200$. According to Mann and Wald's optimum number of classes and Williams' modification, 32 classes is the optimum number of classes for sample size 1200. Thus we use 32 classes for the normality test of each set. As for the normality test or other goodness of fit tests for each sample and for the whole series, the sample size $KL$ and the series size $ISST$ are 50 for each. According to Cochran's moderate number of classes, the suggestions of equal probability per cell by Slakter, Wise, etc and the conventionally used minimum of expected value 5, 10 classes with 5 expected frequencies per class are used for the normality tests and other goodness of fit tests for each sample and for the whole series.
2.2 INDEPENDENCE AND RANDOMNESS TESTS

2.2.1 Introduction

The other more important criteria to evaluate the quality of the generators and the combinations are randomness and independence. Since the numbers generated on computers are based on mathematically congruential relations, and the numbers are determined by the starting value, each number of the generated series is calculated solely from its predecessors and some fixed parameters in the congruential equation. Coveyou (1960) and Greenberger (1961) found that the serial correlations of a generated sequence are functions of the parameters of the sequence. Thus the randomness of a generated series depends on the choice of parameters. The major type of non-randomness of a generated sequence is serial correlation. Thus a vital requirement for the quality of generators and combinations is that the elements of the generated sequence be independent of each other or at least uncorrelated.

There has been some number-theoretic work, such as that of Coveyou (1960) and Greenberger (1961), relating the va-

---

24 Serial correlation \( NR(X_n, X_{n+1}) \), between a number and its immediate successor in a full sequence \( X_n \) generated by \( X_{n+1} = A \times X_n + C \mod P \) \( (0 < X_n < P, \ \text{and} \ P = 2^m \) used in the formula are in Appendix-A) is: a. \( NR(X_n, X_{n+1}) = \frac{1 - (6\times C/p) \times (1 - (C/p))}{A} \) according to Coveyou. b. \( NR(X_n, X_{n+1}) = \frac{(6\times C**2/p**2) - (6\times C/p) + 1}{A} \) according to Greenberger.
values of parameters to certain statistical properties of the resultant generator. This work has not been sufficiently comprehensive to indicate a choice of parameters that can insure against the unlimited possibilities for non-randomness (Greenberger, 1965). The parameters are selected primarily on the basis of a limited set of objectives for generators, such as high speed, long period, etc.; these appear favorable only on intuitive grounds, and the user need to examine a generated sequence carefully by means of statistical tests for serial correlation, runs, etc. (Greenberger, 1961). In addition, simulation is usually affected by a non-random deterministically generated sequence which is supposed to resemble an independence-like or randomness-like sequence (Franklin, 1963). Hence, in practical applications, randomness and independence tests for a generated sequence are advised, and a good generator and a good combination should pass these tests.

To devise a randomness or an independence test is, in general, a difficult problem. Chiefly, this may be due partly to the large classes of alternatives to the null hypothesis (Wolffowitz, 1943). Next, the principles sometimes used to derive procedures (mirimax, likelihood, Bayes) appear to be especially difficult to apply in the problem of randomness even when the class of alternatives is restricted (Lindgren, 1962).
Although numerous tests for randomness and independence have been offered in the literature, many tests are proposed on an intuitive basis, such as rank tests, sign tests, etc., and the alternative hypotheses are restricted to such classes as trend, cycle, regression, etc. It is difficult to find uniformly most powerful test for detecting all of the possibilities for non-randomness. And there is no altogether satisfactory way of testing the hypothesis that a sample is random (Tate, 1957). Hence we select four statistics for randomness and independence tests in this study.

In addition, due to the difficulty of doing so, few power comparisons among randomness and independence tests have been accomplished. The tests we select here are based on their wide use and wide suggestions by respected authors. As will be shown in the next subsection, we have selected runs up and down and runs above and below the median for random oscillation tests and autocorrelation functions (serial correlation coefficients) and spectral analysis for sequence correlation tests.

2.2.2 Literature Review and Statistics Selection

Mosteller (1941) suggested using the length of the longest run above (or below) the median as a test for trend. Wallis (1941) recommended using a goodness of fit test on the number of runs of length 1, 2 or more than 2 as a test for randomness.
Anderson (1941) defined the serial correlation coefficient and introduced it as a measure of the relationship between successive values of a variable ordered in time or space. Wald and Wolfowitz (1943) reported that the test based on the serial correlation seems to be suitable if the alternative to randomness is existence of a trend or of some regular cyclical movement in the data. Kendall (1945) remarked that there are three methods of determining oscillatory elements in time series in current use: a) direct counting of intervals between peaks in the primary series; b) analysis of the data by the correlogram method (i.e., serial correlation); c) analysis by the periodogram method (i.e., spectral analysis).

Leven (1952) noted that runs up and runs down statistics have been independently discovered by a number of different authors and have been widely advocated for testing randomness. Cox and Stuart (1955) pointed out that many distribution free tests have been devised to test the hypothesis of randomness of a series of $N$ observations. Of these, the rank correlation tests are the most efficient test against normal trend alternatives. However, tests based on ranks throw away information and lose power for the tests eventually.
Tate (1957) stated that when the observations have a meaningful time order or sequence, it may be possible to detect certain types of deviation from randomness by serial correlation (good but complicated) or by number of runs (simpler and more practical). Green (1959) reported that serial dependency can be checked by finding the distribution of runs above and below some constant. He used serial correlations and distributions of run lengths in connection with the CS goodness of fit tests to check sequence dependence. From his research, he found the run test to be very sensitive, and he stated that the failure to pass the runs test does not necessarily mean that the numbers are badly awry.

Lindgreen (1962) suggested that randomness tests for a single sample may be based on runs tests, standardized mean square successive difference tests and ranks of the observation tests. Naylor (1966) and Knuth (1969) stated that random oscillation of the sequence of pseudo-random numbers can be tested by "tests of runs". They also stated that serial tests can be used to check the degrees of randomness between successive numbers in a sequence. The serial correlation coefficient is the other measure of independence of the pseudo random numbers. It can detect dependence between two numbers which are k numbers (lag k) apart.
Downham and Roberts (1967) applied tests for runs above and below the median and runs up and down in conjunction with the CS goodness of fit tests and serial correlations up to lag six in conjunction with the CS goodness of fit tests on some multiplicative congruential pseudo-random number generators for the study of their statistical properties. Coweycu and Macpherson (1967) used spectral test (Fourier analysis) as a guideline for choosing multipliers for any multiplicative congruential generator. Knuth (1969) formulated the spectral test as a sequence of steps, but did not give any computer coding. He stated that this test is especially sensitive. Golder (1976) developed "AS 98 algorithm" for this test in computer code form. He suggested that the use of this test for assessing multiplicative congruential generators is an effective prelude to the generated sequence. This is because a good generator will pass this test but not a bad generator.

Whittlesey (1968) mentioned that the autocorrelation function is a measure widely used in studying stochastic processes, and he investigated certain empirical and theoretical properties of autocorrelation functions calculated from sequence of numbers generated by several different pseudo random generators.
Naylor, Wertzy, and Wonnacott (1969) recommended that it is essential to quantify and evaluate the autocorrelation at which spectral analysis is aimed. In addition many authors (e.g. Fishman and Kiviat, 1967; Jenkins, 1961; Sovereign, Nolan, and Mandel, 1971) have applied the spectral analysis to time series data generated by stochastic models. Box and Pierce (1970) used the individual autocorrelation function at each lag K and the portmanteau statistic as an overall test to check the residual of an autoregressive integrated moving average time series for the correlation test.

Emshoff (1970) commented that the two most frequently used tests of randomness are the frequency test and the serial test. The frequency test examines whether or not all numbers are equal likely to occur. so it is improper to use it for testing randomness on normally distributed variates. The serial test is used to determine whether pairs of digits occur randomly. This is the primary test for independence of sequences of numbers. In addition to these, he said one of the most frequently used is the runs test.

Totitle, Robinson and Adams (1971) reported that if the generator has good statistical properties, its runs up and down performance will be an adequate approximation to the formula derived by Leven and Wolfowitz (1941) from the assumption of a continuous rectangular universe. Totitle used
runs up and runs down of different length explicitly to exhibit the distinct behavior patterns for certain generators.

Siegel (1959) and Swanson (1974) suggested the run tests for testing the randomness of a single sample. They commented that the analysis of randomness must be based on the order of occurrence rather than upon the frequency. The runs test focuses on the order of the observations.

The Runs chart is a simple but powerful statistical tool. A run of six or seven consecutive points all lying above or below the median indicates the existence of a special case of variation, usually a trend. A run of six or seven successively progressing downward or successively progressing upward has the same significance. Indeed, it is an important statistical point that some of the most powerful statistical techniques are the most simple (Tanur, Mosteller, Link, Pieters, and Rising, 1972).

Alan (1974) found that various statistics have been considered in the statistical literature for testing randomness against trend alternatives in time series data. For instance, they are Spearman's rank correlation, Kendall's rank correlation, rank serial correlation, number and lengths of runs of elements above and below the median, the runs up and down, etc. Lewis (1975) suggested that fractional period
sequences and other random generators must be tested empirically by "correlograms" (i.e. autocovariance functions) calculated from \( N \) length sequences, and normalized autocorrelation functions. Anderson (1977) used the portmanteau chi-square statistic for checking on the serial correlations taken as a whole. Fisher and Moore (1978) stated that successive short (long) runs imply more mixing (clustering) than one should expect to find in a purely random sequence.

From the review of literature on randomness and independence tests, we have found that runs tests and serial correlation coefficients (i.e., autocorrelation functions for a single time series.)\(^{25}\) are the two most popular statistics; many authors use them for randomness and independence tests. In addition some other tests, such as rank tests, rank serial correlation tests, standardized mean square successive difference tests, and serial tests, are recommended by some authors. They are, however, not suitable for the randomness and independence tests for the generated sequence. The explanation will be discussed in the following paragraphs.

With regard to rank tests or rank serial correlation tests, if the data are nonnumeric, the methods of ranks are

\(^{25}\)The generated numbers can be regarded as a single stochastic time series with discrete lag unit equal one. A stochastic time series is a set of random observations generated sequentially in time.
often the most powerful methods available. If the data are numeric, however, a test based on ranks accounts to throwing away information (Mann, 1945). Such a loss of information usually results in a corresponding loss of power (Conover, 1971).

As for standardized mean square successive differences, this test is based on the statistic \( r = \frac{d}{s} \), where \( d \) denotes the mean square successive difference, \( s \) measures variance independent of the order of the observations; consequently it includes the effect of the trend. If trend is present, \( d \) is not altered nearly so much as the variance estimate \( s \) which increases greatly. The critical region \( r < \) some constant should be employed in testing against the alternative of a trend. Thus we need to know the distribution of \( r \). However, the distribution of \( d \) in samples of \( N \) has not been derived. Tables of critical values are only available up to \( N=50 \). In addition, the \( d \) statistic is not consistent (Newmann and Hart, 1941). Hence it is improper for this study.

Some authors have suggested that serial tests can be used for detecting serial associations or the degrees of randomness between the successive numbers. However, these may be regarded as special cases of autocorrelation functions. For lag one, the autocorrelation function is the same as the serial test.
Autocorrelation analysis not only can discover serial correlations between successive numbers but also can examine serial correlations between numbers lagged \( k \) time units apart. Thus it is more powerful and more general than the serial test. In addition it must be used for randomness and independence tests for generated sequences since the correlation between successive numbers is not the only one of importance, the correlation between the current random number and at least its next 10 or 20 successors should be controlled (Coveyou, 1960).

In all, for the quality evaluation of randomness and independence of the generators and the combinations, we have selected runs tests and autocorrelation tests. For runs tests, we will use two kinds: runs up and down, and runs above the median and below the median. In addition we use spectral analysis as a complementary test for autocorrelation analysis since both of them are Fourier transform of each other, and spectral analysis can offer more information. Also, many previous authors have applied spectral analysis as a test of congruential generators, and they have found it to be especially sensitive: good generators pass the test but not bad ones. Hence, in this study, we will use this analysis as an empirical test on the selected generators and the selected combinations.
Chapter III
THE CHI-SQUARE GOODNESS-FIT-TEST

3.1 INTRODUCTION

Among the numerous goodness of fit tests for distributions, we have selected the CS test for this study. This is due to its simplicity, flexibility, robustness and because it is well-known. In addition CS is most commonly used when there is not a clear alternative hypothesis. In the standard applications of this test, the N observations in a random sample from a population are classified into K mutually exclusive and exhaustive classes. It gives a numerical measure of the difference between the observed and the theoretical distribution. And the theoretical chi-square distribution can be used to define a critical region for the hypothesis test since for large sample sizes, the CS statistic is approximated by the chi-square distribution with K-1 degrees of freedom.

CS can be used to test whether or not a significant difference exists between an observed number of objects falling in each category and an expected number based on the null hypothesis. The number of observations falling into a given category is called the observed frequencies of that cate-
The null hypothesis states the expected frequencies for each exclusive and exhaustive category. For each category, there is a probability that an observation randomly selected from the hypothesized population will fall in that category. Under the null hypothesis, we may designate the probability as \( P_1, P_2, P_3, \ldots, P_k \), for categories 1, 2, \ldots, \( k \), respectively, and we can obtain the expected frequencies for each category by computing the product of \( N \) and the corresponding category probability.

The test consists of establishing mutually exclusive and exhaustive categories. It compares the observed and expected number of occurrences in each category and calculates corresponding chi-square statistic as an appropriate measure of agreement (or disagreement) between observed and expected frequencies. The usual Pearson (1900) chi-square statistic is computed by dividing the squared difference between observed and expected frequencies by the corresponding expected frequencies in each category and summing over all categories. Using the notation, the Pearson chi-square statistic is defined as

\[
CS = \sum \left( \frac{(O_i - N \times P_i)^2}{(N \times P_i)} \right)
\]

where \( O_i \) is the number of observations for the \( i \)th category and \( N \times P_i \) is the expected numbers of frequencies for the \( i \)th category under the null hypothesis, \( H_0 \). \( N \) is the sample size, and \( P_i \) is the probability for the \( i \)th category. It is assumed in this study that \( P_1 = P_2 = \ldots = P_k \).
If the agreement between the observed and expected frequencies is close, the differences \((O_i - N*P_i)^2\) will be small and consequently \(CS\) will be small. If the divergence is large, however, the value of \(CS\) as computed by the formula stated above will also be large. Roughly speaking, the larger the \(CS\) is, the more likely it is that the observed frequencies do not come from the population on which the null hypothesis is based. Since for large samples, \(CS\) is distributed approximately as chi-square with \(K-1\) degrees of freedom. The upper tail (one-tailed test) of the chi-square distribution is used as the critical region. If the computed \(CS\) value is larger than the critical value obtained from the chi-square table at significance level \(ALPHA\), the null hypothesis will be rejected at that significance level. Otherwise, the null hypothesis will be accepted.

3.2 **PROCEDURES OF CS TEST FOR GOODNESS OF FIT**

Three procedures are used here for testing the series for fit to the desired distribution.

1. Test each set for fit to the desired distribution.

   For large sample sizes, the \(CS\) statistic approximates the chi-square distribution. Thus if a series has a desired distribution under the null hypothesis at the significance level \(ALPHA\), the \(CS\)
value computed from that series is less than or equal to the critical value obtained from the chi-square distribution table.

The procedure for accomplishing the test is expected as follows:

a) Determine the sample size $N$ of each set, say $N=1200$.

b) Generate the series $X$ from the computer.

c) By means of Mann and Wald's optimum number of classes with Williams' modification, determine the number of classes and establish $K$ mutually exclusive and exhaustive categories, $K=32$ for $N=1200$.

d) Use the equal probability $1/32$ per class to compute expected frequencies $E_i$ for each category. $E_i=N/K$. $E_i=37.5$ for $N=1200$ and $K=32$.

e) Under the null hypothesis and equal probability $1/32$ per class, compute the upper and the lower bounds for each class.

f) Count the observed frequencies $O_i$ for the $i$th category.

g) Compute the CS statistic

$$CS = \sum ((O_i - E_i) ** 2 / E_i), \text{ for } i=1,2,...,K=32.$$
h) Set up significance level $\text{ALPHA}$. Let $\text{ALPHA}=0.01$, 0.05 and 0.2.

i) Find the critical value, $\text{CRT}$, for the hypothesis test. At significance levels $\text{ALPHA}=0.01$, 0.05, and 0.2 with $K-1=31$ degrees of freedom, $\text{CRT}=52.2$, 45.0 and 37.4, respectively.

The hypothesis and its alternative are:

$H_0$: At significance level $\text{ALPHA}$, if the series $X$ generated is from a population under the null hypothesis, the computed CS value is less than or equal to $\text{CRT}$.

$H_A$: At significance level $\text{ALPHA}$ if the series $X$ is not from a population consistent with the null hypothesis, the computed CS value is greater than $\text{CRT}$.

2. Test each sample of $KL=50$ sets for the desired distribution fit by means of the CS goodness of fit test.

If the sample of $KL=50$ sets has a desired distribution fit, then the $KL=50$ CS statistics computed from each set have a chi-square with $K-1=31$ degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the sample has a desired distribution fit.
The procedure for accomplishing the test is as follows:

a) Use Procedure mentioned above to compute the CS statistic for each set. There are KL=50 CS statistics which have chi-square distribution with K-1=31 degrees of freedom if the sample has a desired distribution fit.

b) By the conventional method using minimum expected value 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, K1, for KL samples and let K1=10 for KL=50.

c) Use equal probabilities 1/10 to compute the expected frequency Ei for the ith category. Use Ei=5 for KL=50 and Pi=0.1.

d) Under the null hypothesis and equal probability 1/10 per class, compute the upper and lower bounds for each class.

e) Count the observed frequencies Oi for the ith category in K1 categories.

f) Compute the CS statistic

\[ CS = \text{SUM} \left( \frac{(E_i - G_i)^2}{E_i} \right), \text{ for } i = 1, 2, \ldots, K = 10. \]
g) Set up significance levels ALPHAl=0.01, 0.05 and 0.2 for the hypothesis tests.

h) Find the critical value, CRT, for the hypothesis tests. At significance levels ALPHAl=0.01, 0.05 and 0.2 with K1-1=9 degrees of freedom, CRT=21.7, 16.9 and 12.2, respectively.

The hypothesis and its alternative are:

HO: The sample of KL=50 sets is from the desired distribution under the null hypothesis at significance level ALPHAl if the computed CS value is less than or equal to CRT.

HA: The sample of KL=50 sets is not from the desired distribution under the null hypothesis at significance level ALPHAl if the computed CS value is greater than CRT.

3. Test the whole series of ISET=50 samples for the desired distribution by number of significant samples counted above and by means of the CS goodness of fit test.

At significance level ALPHAl, we use two ways to judge whether or not the whole generated series has a desired distribution fit. First, If the whole series of ISET=50 samples has a desired dis-
tribution fit, then the number of significant samples counted over ISET=50 samples is less than or equal to the upper bound for number of significant samples \((ISET*ALPHA+Z*(ISET*ALPHA*(1-ALPHA)))^{**.5}\), where \(Z=2.328, 1.645\) and \(0.845\) for \(ALPHA=0.01, 0.05\) and \(0.2\), respectively). Second, if the whole generated series has a desired distribution fit, the ISET=50 CS statistics computed from all samples have a chi-square distribution with \(K1-1=9\) degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the whole generated series has a desired distribution fit. The expectation of this procedure is as follows:

a) Use previous Procedure to compute the CS statistic for each sample. There are ISET=50 CS statistics which have a chi-square distribution with \(K1-1=9\) degrees of freedom if the whole generated series has a desired distribution fit.

b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, \(K2\), and let \(K2=10\) for ISET=50 in this study.
c) Use equal probabilities $1/10$ to compute the expected frequencies $E_i$ for the $i$th category. Thus, $E_i=5$ for ISET=50 and $P_i=0.1$.

d) Under the null hypothesis and equal probabilities $1/10$ per class, compute the upper and lower bounds for each class.

e) Count the observed frequencies, $O_i$ for the $i$th category in $K_2$ categories.

f) Compute the CS statistic

$$CS=\sum ((E_i-C_i)^2/E_i), \text{ for } i=1,2,\ldots,K_2=10.$$ 

$g)$ Set up significance levels $ALPHA=0.01$, 0.05 and 0.2 for the hypothesis tests.

$h)$ Find the critical value, CRT, for the hypothesis tests. At significance levels $ALPHA=0.01$, 0.05, and 0.2 with $K_2-1=9$ degrees of freedom, $CRT=21.7$, 16.9 and 12.2, respectively.

The hypothesis and its alternative are:

$H_0$: The whole series of ISET=50 samples is from the desired distribution under the null hypothesis at the significance level $ALPHA$ if the number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to CRT.
HA: The whole series of ISET=50 samples is not from the desired distribution under the null hypothesis at significance level ALPHA if the number of significant samples is greater than its upper bound, or if the computed CS value is greater than CRT.

3.3 EMPIRICAL TEST RESULTS, DISCUSSIONS AND EVALUATIONS

The goodness of fit test results for 7 generators and 23 combinations are shown in Table 1. To test the fit to a desired distribution, we generated 50 independent samples. Each sample contained 50 sets of 1200 numbers per set.

For each set, we established 32 mutually exclusive and exhaustive categories, then computed chi-square values for the distribution of the 1200 numbers in the set. We thus obtained one chi-square value with 31 degrees of freedom for each set and 50 chi-square values for each sample.

Next, we set up 10 mutually exclusive and exhaustive classes for the 50 independent chi-square values for each sample. Then we derived one chi-square value with 9 degrees of freedom from these and computed 50 chi-square values for all independent samples. Finally, at the significance level ALPHA, we counted the number of significant (rejected) sam-
amples by comparing the observed sample chi-square value with the appropriate critical chi-square value.

Columns 3, 4, and 5 list the numbers of significant samples at levels 0.2, 0.05, and 0.01, respectively. At significance levels 0.2, 0.05, and 0.01, the upper bounds for the numbers of significant samples are 12, 5, and 2, respectively. For example, at level 0.2, there are 6 significant samples among the 50 independent samples generated from AD-RAND. The upper critical bound of the number of significant samples is 12 at level 0.2. So the numbers generated from AD-RAND are judged to be unit uniform distribution by investigating the number of significant samples generated.

Again, we casted 50 samples' chi-square values into 10 mutually exclusive and exhaustive groups and computed one grand chi-square value with 9 degrees of freedom for each whole generated series. Thirty grand chi-square values are exhibited in Column 6, and their critical chi-square values for the whole series are 12.2, 16.9, and 21.7 for significance levels 0.2, 0.05, and 0.01, respectively. For instance, the observed chi-square value for the whole series generated from AD-RAND is 5.6 at significance level 0.2. Since the critical chi-square value is 12.2, the numbers generated from AD-RAND are judged to have a good fit to the unit uniform distribution at the stringent 0.2 level.
Column 2 summarizes the actual total number of rejections from the goodness of fit tests. RANDU, for example, has 2 rejections in the test for goodness of fit to the distribution. The distributions tested for are unit uniform and unit normal.

Table 1 informs us of 6 types of possible rejections. Three of them are indicated by the number of significant samples (Columns 3, 4, and 5) and the other three (Column 6) are pointed out by overall chi-square values at levels 0.2, 0.05, and 0.01, respectively. In addition, the actual total numbers of rejections that occurred in the test for goodness of fit to the unit uniform or unit normal distribution are listed in column 2.

From Table 1 it is clear that most generators or combinations seem to fit to their desired (i.e., unit uniform or unit normal) distributions well. At significance level 0.2, the numbers of significant samples generated both by R. Shore and by M & M with the Box-Muller rejection transformation are 14 (Column 3). This is slightly over 12, the upper bound of the significant samples. Except for these two, all selected generators and combinations pass the chi-square goodness of fit test based on the number of significant samples.
### Table 1

Goodness of Fit to Desired Distributions

<table>
<thead>
<tr>
<th>Generators/</th>
<th>[No.0] # of Sig. Samples at Level</th>
<th>Chi-square Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>ADRAND</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>RANDU</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>URAND</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>R. Shore</td>
<td>1</td>
<td>14 s</td>
</tr>
<tr>
<td>GGUSS</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>GGUSS-BMI</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>GGUSS-HSI</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>1</td>
<td>14 s</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>GGUSS-BMJ</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>GGNPM</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>GGNML</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

*: Significant at level 0.2
**: Significant at levels 0.2 and 0.05.
****: Significant at levels 0.2, 0.05 and 0.01.
s: Significant at level ALPHA.
#: Total number of possible rejections is 6.
BMI: Stands for Box-Muller Inverse.
BMJ: Stands for Box-Muller Reject.
HSI: Stands for Hasting Inverse.
Upper bound for the number of significant samples at levels 0.2, 0.05 and 0.01 are 12, 5, and 2, respectively. Critical chi-square values for the whole generated series are 12.2, 16.9 and 21.7 for significance levels 0.2, 0.05 and 0.01, respectively.
From the investigation of the observed grand chi-square values (Column 6), it is clear that unit uniform generators, RANDU and M & M, and unit normal combinations, GGUBS with Box-Muller inverse, L & L with Box-Muller rejection, R. Shore with Box-Muller rejection, and GGUBS with Box-Muller rejection do not pass the goodness of fit tests for their desired distributions at significance level 0.2 or 0.05. Their corresponding observed grand chi-square values are 17.2, 17.2, 15.2, 15.6, 16.8, and 12.8. However, the critical chi-square values are 12.2, 16.9, and 21.7 for levels 0.2, 0.05, and 0.01, respectively. Hence, the uniform generators, RANDU and M & M are rejected at levels 0.2 and 0.05, and the other three unit normal combinations are rejected at level 0.2. The rest of the generators and combinations pass the goodness of fit tests to the desired distributions using the grand chi-square values derived from 50 independent samples.

Overall, there are 2 rejections cut of 6 computed for the RANDU and M & M uniform generators and 1 rejection each for R. Shore alone and GGUBS combined with the Box-Muller inverse routine, L & L with Box-Muller rejection, M & M with Box-Muller rejection, R. Shore with Box-Muller rejection, and GGUBS with Box-Muller rejection routines. There are no rejections in rest of the uniform generators or normal combinations (Column 2).
Chapter IV

TESTS OF RUNS

4.1 INTRODUCTION

According to Siegel (1959) and Swanson (1974), the analysis of randomness must be based on order of occurrence rather than upon frequency. Runs tests, focusing on the order of the events, give us information which is not indicated by the frequency of the events. In addition, a review of the literature indicates that the runs test is a simple but powerful statistical tool (Tanur, 1972; Green, 1959, et al.). Thus the random oscillation of the sequence of pseudo random numbers (i.e. randomness in order of appearance, or permutational arrangement) will be tested by tests for runs in this study.

A run is defined as a succession of similar events preceded and succeeded by different events or by no event at all. The number of elements in a run will be referred to as the length of the run. Among runs tests, some authors use runs above and below the median; some suggest runs up and down. The test for runs up and down is sensitive for detecting cyclical or seasonal variations in time series but is not sensitive to trend (Wallis and Moore, 1941). How-
ever, runs above and below the median can be used as a test for trend (Mosteller, 1941). Thus in this study both of these are used for randomness tests.

The total number of runs and the distribution of the number of runs by length will be used as test criteria both in tests for runs up and down and in tests for runs above and below the median. The total number of runs in a sample gives an indication of whether or not the sample is random. It is applicable to cases where the alternative hypothesis is either abnormally long or abnormally short runs. Too few runs in total might indicate a tendency for like observations to follow like observations; a time trend or some bunching due to lack of independence may be suggested. Too many runs may indicate that each observation tends to follow and be followed by an observation of the other type; a systematic short-period, cyclical fluctuations seem to be influencing the events (Swanson, 1974). However, an "abnormally" long run in one or both variates which would be regarded as an indication of nonrandomness, might be accompanied by a large number of runs of length one which might make the total number of runs not critically low. Hence another possible test is advised which compares the number of runs of various lengths with their corresponding expected numbers by the proper chi-square (Wolfowitz, 1943). This test is normally more sensitive than tests based on the to-
tal number of runs (Wallis and Moore, 1941). It is based on more information, and it can exhibit distinct behavioral patterns in the generated sequence.

Furthermore, Mood (1940) showed that the total number of runs (regardless of the length of the run) and the number of runs of each length are asymptotically normally distributed. Thus the test statistic based on the total number of runs, for large sample size, can be treated as a normal variate and the standard normal table can be used to determine the critical values for the hypothesis test. As for the test based on the distribution of the number of runs by length, a proper chi-square goodness of fit test of significance can be used to test the "randomness of elements with respect to order" when dealing with large samples.

4.2 RUNS UP AND DOWN

4.2.1 Introduction

A sequence may be tested for "runs up" and "runs down". This means we examine the length of monotone subsequences of the original sequence, i.e., segments which are increasing or decreasing. Thus this test is based on sequences in direction of movement, that is, upon the sequence of like sign in the differences between successive observations. And the test is designed to result in a rejection of the hypothesis
of randomness if the numbers in the sequence are repeatedly larger (or smaller) than the previous numbers.

Under the runs up and down hypothesis, the run is defined as an unbroken sequence of increasing or decreasing observations. A monotonically ascending subsequence of observations will be defined as a run "up" denoted as a sequence of "+" signs. A steadily decreasing subsequence of observations is called a run "down" and is denoted by a sequence of "-" signs. Each point at which the series under analysis ceases to decline and starts to rise, or ceases to rise and starts to decline, is called a turning point. A turning point is a "peak" if it is a (relative) maximum or a "trough" if a (relative) minimum. The interval between consecutive turning points is called a run. The length of a run is the number of elements (i.e., number of "+" signs or number of "-" signs in it) between its initial and terminal points. Thus a series of N observations may contain as few as zero or as many as N-2 turning points. And the length of a run may be as short as one event or as long as N events.

Under the randomness hypothesis, the mean and variance of the total number of runs are \( \frac{2N-1}{3} \) and \( \frac{16N-29}{90} \).

---

1We subtract each element from its predecessor and replace the result by "+" or "-" according as the difference is positive or negative. We get runs of "+" signs and "-" signs, called respectively runs up and down.
respectively, and the total number of runs is asymptotically normally distributed. Thus for large samples, say \( N > 25 \), for a test based on the total number of runs, an approximation to the probability of the total number of runs may be obtained by treating the total number of runs as a normal deviate and obtaining the critical ratio from the standard normal table (Bradley, 1968; Lewis, 1975). A two-tailed test is used here since too many runs or too few runs will indicate nonrandomness.

The expected value of the number of runs by length \( k \) assuming a "truly" random sample is \( 2*\frac{(k**2+3*k+1)*N-(k**3+3*k**2-4)}{(k+3)!} \) (Levene and Wolfowitz, 1944; Naylor, 1966; Tootitle, 1971; Lewis, 1975). The test based on the distribution of number of runs by length involves counting the actual number of occurrences of runs of different lengths and comparing these actual counts with their corresponding expected theoretical values. Then the chi-square goodness of fit test may be used to check whether or not a generated sequence is acceptable at a given level of significance (Naylor, Balintfy, Furdick, and Chu, 1966; Downham and Roberts, 1967).
4.2.2 Procedures for Testing Runs Up and Down

There are two approaches using runs up and down for testing the generated sequence for random oscillation. They are total number of runs and number of runs by length, respectively. The following two sub-subsections will outline each approach.

4.2.2.1 Total number of runs

Three procedures are used here for testing the generated series for randomness by means of this approach.

1. First, test each set for randomness.

Under the randomness null hypothesis, for large sample sizes, the total number of runs is approximated by the normal distribution with mean \((2N-1)/3\) and variance \((16N-29)/90\). Thus, the procedure for testing whether or not the series is random is as follows:

a) Determine the sample size \(N\) of each set, say \(N=1200\).

b) Generate the series \(X\).

c) List in order of occurrence the signs (+/-) of the differences between successive items.
d) Count the total number of runs, $U$, in the sequence. That is, count the number of unbroken subsequences of "+" signs and "−" signs.

e) Compute the statistic:

$$ZU = \frac{U - ((2N - 1)/3)}{\sqrt{((16N - 29)/90)^{1/2}}}$$

By hypothesis, $ZU$ is approximately a standard normal distribution.

f) Choose a significance level ALPHA. Let ALPHA=0.01, 0.05 and 0.2.

g) Set up the theoretical confidence interval for the total number of observed runs. That is,

$$(\frac{(2N - 1)}{3} - Z\sqrt{\frac{(16N - 29)}{90}}) < U < \left(\frac{(2N - 1)}{3} + Z\sqrt{\frac{(16N - 29)}{90}}\right)^{0.5}$$. From the standard normal table, we find the $Z=2.325$, 1.96 and 1.28, for two-tailed hypothesis test, at the significance levels ALPHA=0.01, 0.05 and 0.2 respectively.

The hypothesis and its alternative are:

$H_0$: At significance level ALPHA, if the series $X$ generated is random, the total sample number of runs is within the theoretical confidence interval computed above.

$H_A$: At significance level ALPHA, if the series $X$ is not random, the total sample number of
runs is not within the theoretical confidence interval.

2. Second, test each of the samples of KL=50 sets for randomness.

If the sample of KL sets is random, then the KL=50 ZU values have a standard normal distribution. Thus we can use a chi-square goodness of fit test to test the normality of KL=50 ZU statistics.

a) Use the procedure above to compute the ZU statistic for each set. There are KL=50 ZU statistics which have an approximately standard normal distribution if the sample is random.

b) Use the conventional expected value of at least 5 and Cochran's moderate number of classes between 10 to 25, establish K1=10 mutually exclusive and exhaustive categories for KL samples.

c) Use equal probabilities 1/10 to compute the expected frequencies Ei for the ith category. Thus, Ei=5 for KL=50 and Pi=0.1.

d) Under the null hypothesis and equal probability 1/10 per class, compute the upper and the lower bounds for each class.
e) Count the observed frequencies, $O_i$, for the ith category in $K_1$ categories.

f) Compute the CS statistic

$$CS = \sum ((E_i - C_i)^2 / E_i), \text{ for } i=1,2,\ldots,K_1=10.$$ 

g) Set up significance levels $\text{ALPHA}=0.01, 0.05$ and $0.2$ for the hypothesis tests.

h) Find the critical value, CRT, for the hypothesis test. At significance levels $\text{ALPHA}=0.01, 0.05$ and $0.2$ with $K_1-1=9$ degrees of freedom, CRT=21.7, 16.9 and 12.2, respectively.

The hypothesis to be tested and its alternative are:

$H_0$: The sample of KL=50 sets has random oscillatory characteristics at the significance level $\text{ALPHA}$ if the computed CS value is less than or equal to CRT.

$H_1$: The sample of KL=50 sets does not have random oscillatory characteristics at the significance level $\text{ALPHA}$ if the computed CS value is greater than CRT.

3. Third, test the combined series of ISET=50 samples for randomness by number of significant samples counted above and by means of the CS goodness of fit test.
At significance level ALPHA, we use two ways to judge whether or not the whole generated series is random. First, if the whole series of ISET=50 samples is random, then the number of significant samples counted over ISET=50 samples is less than or equal to the upper bound for number of significant samples (ISET*ALPHA+Z*(ISET*ALPHA*(1-ALPHA))**.5). Second, if the whole generated series of ISET=50 samples is random, the ISET=50 CS values have a chi-square distribution with K1-1=9 degrees of freedom. Thus we can use a chi-square goodness of fit test to check whether or not the ISET=50 CS values have a chi-square distribution with K1-1=9 degrees of freedom. The procedure for accomplishing this test is as follows:

a) Use the second procedure above to compute the CS statistic for each sample. There are ISET=50 CS statistics which have a chi-square distribution with K1-1=9 degrees of freedom if the whole series is random.

b) Using the conventional expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish K2=10 mutually exclusive and exhaustive categories.
c) Use equal probabilities \(1/10\) to compute the expected frequencies \(E_i\) for the \(i\)th category. Thus, \(E_i=5\) for \(ISET=50\) and \(P_i=0.1\).

d) Under the null hypothesis and with equal probability \(1/10\) per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, \(O_i\), for the \(i\)th category among the \(K_2\) categories.

f) Compute the CS statistic

\[
CS = \sum \left( \frac{(E_i-G_i)^2}{E_i} \right), \text{ for } i=1,2,\ldots,K_2=10.
\]

g) Set up significance levels \(ALPHA=0.01, 0.05\) and \(0.2\) for the hypothesis tests.

h) Find the critical value, \(CRT\), for the hypothesis test. At significance levels \(ALPHA=0.01, 0.05\) and \(0.2\) with \(k_2-1=9\) degrees of freedom, \(CRT=21.7, 16.9\) and \(12.2\), respectively.

The hypothesis and its alternative are:

\(H_0\): The whole series of \(ISET=50\) samples is random at significance level \(ALPHA\) if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to \(CRT\).
HA: The whole series of ISPI=50 samples is not random at significance level ALPHA if number of significant samples is greater than its upper bound, or if the computed CS value is greater than CRT.

4.2.2.2 The Number of Runs by Length

Three procedures are used in this approach for testing the generated series for randomness.

1. First, test each set for randomness.

   If the series is random, the number of runs of various lengths is asymptotically normally distribution. Thus for large sample sizes, the chi-square goodness of fit test can be used to compare the number of runs of different lengths with the corresponding expected number of runs to check the generated sequence for randomness.

   This procedure is expected as follows:

   a) Determine the sample size N of each set, say N=1200.

   b) Generate the series X from the computer.
c) List in order the signs (+/-) of the difference between successive items.

d) Count the frequencies of runs, $O_i$, by length $i$.

e) Compute the expected theoretical frequencies for length $i$ from the formula:

$$E_i = \frac{2(i^2+3i+1)N-(i^3+3i^2-2i-4)}{(i+3)!}$$

for $i=1,2,...,(K-1)$ and

$$E_k = \frac{(2N-1)}{3} - \text{SUM}(E_i), \text{ for } i=1,2,...,(K-1).$$

Let $K=5$ in this study.

f) Adjust the expected frequency, $E_i$

$$AE_i = E_i \times \frac{\text{SUM}(O_j)}{((2N-1)/3)}$$

for $i=1,2,...,K$, $j=1,...,K$.

g) Compute the CS statistic

$$CS = \text{SUM}((O_i-AE_i)^2/AE_i), \text{ for } i=1,2,...,K=5.$$  

h) Set up the significance levels $\alpha = 0.01, 0.05$ and 0.2.

---

The expected value for the 5th category for $N=1200$ is 4.84. It is slightly less than the conventional value 5. The probabilities for the number of runs by length are not equal. However, according to Cochran's (1952) and Roscoe and Byars' (1971) suggestions mentioned in Chapter 2, we are still confident using the CS goodness of fit test in conjunction with this test.
i) Find the critical value, CRT, for the hypothesis test. At the significance levels \( \text{ALPHA}=0.01, 0.05 \) and 0.2 with \( K-1=4 \) degrees of freedom, CRT=13.3, 9.49 and 5.99, respectively.

The hypothesis and its alternative are:

\( H_0: \) At significance level \( \text{ALPHA} \), if the generated series \( X \) is random, the computed CS value is less than or equal to CRT.

\( H_A: \) At significance level \( \text{ALPHA} \), if the generated series \( X \) is not random, the computed CS value is greater than CRT.

2. Second, test each sample of \( KL=50 \) sets for randomness.

If the sample of \( KL=50 \) sets is random, then the \( KL=50 \) CS statistics computed from each set have a chi-square distribution with \( K-1=4 \) degrees of freedom. Thus we can use CS goodness of fit test to check whether or not the \( KL=50 \) values of CS statistics have a chi-square distribution with \( K-1=4 \) degrees of freedom.

a) Bed the procedure described above to compute the CS statistic for each set. There are \( KL=50 \) CS statistics which have the chi-square distri-
bution with $K-1=4$ degrees of freedom if the sample is random.

b) Using the conventional expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish $K1=10$ mutually exclusive and exhaustive categories for the KL samples.

c) Use equal probabilities $1/10$ to compute expected frequencies $E_i$ for the $i$th category. Thus, $E_i=5$ for $KL=50$ and $P_i=0.1$.

d) Under the null hypothesis and equal probability of $1/10$ per class, compute upper and lower bounds for each class.

c) Count the observed frequencies $O_i$ for $K1$ categories.

f) Compute the CS statistic

$$CS = \sum \{(E_i-C_i)^2/E_i, \text{for } i=1,2,\ldots,K1=10.$$  

\(g\) Set up significance levels for $\text{ALPHA}=0.01, 0.05$ and 0.2 for the hypothesis tests.

h) Find the critical value, CRT, for the hypothesis tests. At significance levels $\text{ALPHA}=0.01, 0.05$ and 0.2 and $K1-1=9$ degrees of freedom, CRT=21.7, 16.9 and 12.2, respectively.

The hypothesis and its alternative are:
HO: The sample of KL=50 sets is random at significance level ALPHA if the computed CS value is less than or equal to CRT.

HA: The sample of KL=50 sets is not random at the significance level ALPHA if the computed CS value is greater than CRT.

3. Third, test the whole series for randomness by number of significant samples counted above and by means of CS goodness of fit test.

Again, at significance level ALPHA, we use two methods to judge whether or not the whole generated series is random. First, if the whole series of ISET=50 samples is random, then the number of significant samples counted over ISET=50 samples is less than or equal to the upper bound for number of significant samples
\[(ISET\times ALPHA + Z \times (ISET\times ALPHA \times (1-ALPHA))^{**.5}).\] Second, if the whole series of ISET=50 sets is random, then the ISET=50 CS statistics computed from all samples have a chi-square distribution with \(K1-1=9\) degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the ISET=50 values of CS statistics have a chi-square distribution with \(K1-1\) degrees of freedom. This procedure is as follows:
a) Use the second procedure above to compute the CS statistic for each sample. There are ISET=50 CS statistics which have a chi-square distribution with K1-1=9 degrees of freedom if the whole series is random.

b) By using the conventional minimum expected value of 5 and Cochran's moderate number of classes (between 10 to 25) establish K2=10 mutually exclusive and exhaustive categories.

c) Use equal probability 1/10 to compute the expected frequencies Ei for the ith category. Thus, Ei=5 for ISET=50 and Pi=0.1.

d) Under the null hypothesis and equal probability 1/10 per class, compute the upper and lower bounds for each class.

e) Count the observed frequencies, Oi for the ith category in the K2 categories.

f) Compute the CS statistic

\[ CS = \sum \frac{(E_i - O_i)^2}{E_i} \text{ for } i=1,2,\ldots,K2=10. \]

g) Set up significance levels \( \text{ALPHA}=0.01, 0.05 \) and 0.2 for the hypothesis tests.

h) Find the critical value, \( \text{CRT} \), for the hypothesis tests. At significance levels \( \text{ALPHA}=0.01, \)
0.05 and 0.2 with \( K2-1=9 \) degrees of freedom, CRT=21.7, 16.9 and 12.2, respectively.

The hypothesis and its alternative are:

**H0:** The whole series of ISET=50 samples is random at significance level \( \text{ALPHA} \) if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to CRT.

**HA:** The whole series of ISET=50 samples is not random at significance level \( \text{ALPHA} \) if number of significant samples is greater than its upper bound, or if the computed CS value is greater than CRT.

4.3 **RUNS ABOVE AND BELOW THE MEDIAN (MEAN)**

4.3.1 **Introduction**

The other tests for runs used in this study are tests for runs above and below the median (or mean).\(^3\) Runs above and

\(^3\)For the unit uniform distribution or the standard normal distribution, the median of the sequence should equal the mean of the sequence, and the mean=median=0.5 or 0, respectively. Thus for these two distributions, runs above and below the mean is the same as the runs above and below the median.
below the median may be considered as runs from a binomial population. The observations are classified into two classes (i.e., above the median and below the median). Under the randomness null hypothesis each observation has a constant probability, \( P_i = 0.5 \), of falling into the \( i \)th class \((i=1,2)\). Then a sequence of \( K \) observations belonging to the same class—above the median (or below the median)—and which is preceded and followed by observations belonging to another class—below the median (or above the median)—is called a run of length \( K \).

A run is an unbroken sequence of similar event (above the median or below the median). All observations below the median may be designated by "-" signs; all above the median may be designated by "+" signs. Thus the runs of "+" signs are called runs above the median, and the runs of "-" signs are called runs below the median.

The test for runs above and below the median is designed to result in a rejection of the hypothesis of randomness if the sequence of numbers is such that any number which is greater (or less) than the median of the numbers in the sequence, is repeatedly followed by a number which is greater (or less) than the median of the numbers in the sequence. If the generator or the combination is producing a random sequence, then the conditional probability that some number,
\( R_{i+1} \) is greater than the median, given that \( R_i \) is greater than the median, is equal to the probability that \( R_{i+1} \) is less than the median, given that \( R_i \) is greater than the median. A similar equality holds for the conditional probability in the case when \( R_i \) is less than the median (Buffa, 1971).

For large samples, a good approximation to the sampling distribution of the total number of runs, \( U \), is the normal distribution with mean = \((N/2) + 1\) and variance = \((N \times (N-2))/(4 \times (N-1))\). Therefore when the sample size is large, the null hypothesis may be tested by \( Z_U = (U - (N/2) + 1)/(N \times (N-2))/(4 \times (N-1))\)*0.5 (Bradley, 1963; Lewis, 1975).

Assuming random samples, the expected value of the number of runs by length \( K \) is \((N-K+3)/(2 \times (K-1))\). Again the chi-square goodness of fit test may be used to check whether or not a generated sequence is acceptable at a given significance level (Naylor, 1966).

4.3.2 Procedures for Testing Runs Above and Below the Median

Two approaches are used with runs above and below the median for testing the generated series for random oscillation. They are the total number of runs and the number of runs by length. The following two sub-subsections will discuss each, respectively.
4.3.2.1 Total Number of Runs

Three procedures are used for testing the generated series for randomness using of this approach.

1. First, test each set for randomness.

Under the randomness null hypothesis, for large sample sizes, the total number of runs is approximated by the normal distribution with mean \(\frac{(N/2) + 1}{4(N-1)}\) and variance \(\frac{(N^2(N-2))}{(4(N-1))}\).

The procedure is expected as follows:

a) Determine the sample size \(N\) of each set, say \(N=1200\).

b) Generate the series \(X\).

c) List the ordered sequence of signs (+/-) for each observation. "+" if the observation is above the median, and "-" if the observation is below the median.

d) Count the total number of runs, \(U\), in the sequence. That is, count number of "+"'s in an unbroken subsequence of "+" signs and do the same for "-" signs.
e) Compute the statistic:
\[ Z_U = \frac{(U - (N/2) + 1))}{((N^2(N-2))/(4(N-1)))^{0.5}}. \]

\( Z_U \) is distributed approximately as a standard normal distribution.

f) Set up the significance level \( \text{ALPHA} \) with \( \text{ALPHA}=0.01, 0.05 \) and 0.2.

g) Set up theoretical confidence interval for total number of observed runs. That is,
\[(N/2) + 1) - Z \times (N^2(N-2))/(4(N-1)))^{0.5} < U < (N/2) + 1) + Z \times (N^2(N-2))/(4(N-1)))^{0.5}. \]

From the standard normal table, the value of \( Z \) is 2.325, 1.96 and 1.28 for a two-tailed hypothesis test, at significance levels \( \text{ALPHA}=0.01, 0.05 \) and 0.2, respectively.

The hypothesis and its alternative are:

\( H_0 \): At significance level \( \text{ALPHA} \), if the series X is random, the observed number of total runs is within the theoretical interval computed above.

\( H_A \): At significance level \( \text{ALPHA} \), if the series X is not random, the observed number of total runs is not within the theoretical confidence interval.
2. Second, test each sample and the whole series for randomness.

The criteria and the procedures are the same as those in the approach used for the runs up and down test for each sample and the whole generated series.

4.3.2.2 The Number of Runs by Length

All the criteria and the procedures used here are the same as those used in the test for "runs up and down" with the exception of the theoretically expected numbers of runs by length. That is, the expected number of runs of length \( K \) is given by \( (N-K+3)/(2^{*}(K+1)) \) (Naylor, 1966).* 

4.4 EMPIRICAL TEST RESULTS, DISCUSSIONS AND EVALUATIONS

The results of randomness tests using two non-parametric tests, runs up and down and runs above and below the median, are shown in Tables 2 and 3, respectively. In order to evaluate the characteristic of random oscillation of the se-

*The number of lengths in runs above and below the median is \( K=7 \). At significance levels \( \text{ALPHA}=0.01, 0.05 \) and \( 0.2 \) with \( df=6 \), the critical chi-square values are 16.8, 12.6 and 8.56, respectively.
lected generators/combinations, we generated 50 independent samples which contain 50 sets per sample and 1200 numbers per set.

**Test by total number of runs:** First, we counted the number of runs for each set. Then, for each set we normalized the observed number of runs by its expected number of runs and standard deviation. Next, for each sample we computed one chi-square value with 9 degrees of freedom from 50 normalized total numbers of observed runs. We tested whether or not these 50 normalized total numbers of observed runs have a standard normal distribution by comparing the computed chi-square value with its critical chi-square value. If the sample is not random, we counted it as a significant sample. Column 7, 8, and 9 (in Tables 2 and 3) list the number of significant samples resulting from this approach. The corresponding upper bounds for the number of significant samples are 12, 10 and 5.

Column 10 shows the grand chi-square values with 9 degrees of freedom for the whole generated series. They are derived from the 50 independent chi-square values of the samples. Their critical values are 12.2, 16.9, and 21.7, respectively, for levels 0.2, 0.05 and 0.01.
Test by using of runs by length: For each set, we counted the number of runs by length 1, 2, 3, 4, 5, for runs up and down, and by the number of runs by length 1, 2, 3, 4, 5, 6, 7, for runs above and below the median. Based on the observed number of runs by length and the expected number of runs by length, we calculated the chi-square values with 4 and 6 degrees of freedom for runs up and down and runs above and below the median, respectively. For each sample, we computed 50 chi-square values with 4 degrees of freedom for runs up and down, and 50 chi-square values with 6 degrees of freedom for runs above and below the median. Then we fitted these 50 independent chi-square values with 4 or 6 degrees of freedom into the other chi-square distribution; we obtained one chi-square value with 9 degrees of freedom.

Next, we tested for significant samples by comparing the observed chi-square value with its critical value at significant level ALPHA. This information is presented in Columns 3, 4 and 5 in Tables 2 and 3 for significance levels 0.2, 0.05 and 0.01, respectively. Again the grand chi-square values with 9 degrees of freedom computed from all samples' chi-square values are demonstrated in Column 6 of Tables 2 and 3.

Both from Table 2 and from Table 3, there are 12 possible types of significant results. Six of them are obtained from tests by length (Columns 3, 4, 5, and 6); the other six are
from tests by total number of runs (Columns 7, 8, 9, and 10). Among each of the six kinds of rejections, three are indicated by numbers of significant samples, and the other three by overall chi-square values at levels 0.2, 0.05, and 0.01. The actual total number of rejections from runs up and down and from runs above and below the median are summarized in Column 2 of Tables 2 and 3, respectively.

From Table 2 (runs up and down tests) it is clear that most generators/combinations do not pass the tests by length of run at significance level 0.2. This is especially true for ADRAND and ADRAND combined with Hasting inverse transformation; not one sample of the 50 that are generated passes this test at level 0.2 or 0.05 or 0.01. Their grand chi-square values are 450 against critical chi-square values 12.2, 16.9, and 21.7 for levels 0.2, 0.05, and 0.01, respectively. So they are rejected at all levels investigated here either from the examination of the numbers of significant samples or from the investigation of overall chi-square values. Also, ADRAND combined with the Box-Muller rejection transformation has large numbers of significant samples, 38, 30, and 17 (the upper bounds are 12, 10 and 5) for levels 0.2, 0.05, and 0.01, respectively and a very large grand chi-square value (173) in the tests by length. In addition, URAND combined with Box-Muller inverse and R. Shore with Hasting inverse are also rejected by the tests by length at all significance levels considered here.
### Table 2

**Runs up and Down**

<table>
<thead>
<tr>
<th>Generators/</th>
<th>No.</th>
<th>Comb</th>
<th>By Length</th>
<th>By Total No. of Runs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **ACRAND**: 750s 50s 50s 450*** 14s 3 0 10.4
- **RANDU**: 20s 5 1 20.8** 11 6s 3s 11.2
- **L & L**: 19s 7s 3s 18.0** 14s 2 1 7.2
- **M & M**: 12 3 1 11.2 15s 3 1 12.8*
- **URAND**: 22s 4 0 27.2*** 17s 4 2 16.0*
- **R. Shore**: 22s 13s 1 42.8*** 9 1 0 6.8
- **GGUBS**: 20s 6s 2 20.0** 14s 3 1 8.8
- **ADRAND-BMI**: 0 11 4 1 7.6 6 3 0 11.6
- **RANDU-BMI**: 18s 2 1 24.0** 10 0 0 14.8*
- **L & L-BMI**: 18 8s 5 1 12.4* 14s 3 0 17.6**
- **M & M-BMI**: 20s 9s 1 25.2*** 14s 5 1 9.6
- **URAND-BMI**: 21s 8s 5s 32.4*** 11 4 1 9.2
- **R. Shore-BMI**: 21s 5 2 19.6** 13s 5 0 15.6*
- **GGUBS-BMI**: 20s 6 1 14.0* 12 4 1 12.0
- **ADRAND-HSI**: 50s 50s 50s 450*** 14s 3 0 11.2
- **RANDU-HSI**: 24s 3 0 27.6*** 12 5 2 8.4
- **L & L-HSI**: 19s 8s 3s 17.2** 14s 3 0 6.0
- **M & M-HSI**: 11 2 0 4.8 11 3 1 2.8
- **URAND-HSI**: 20s 4 1 23.2*** 16s 4 1 9.6
- **R. Shore-HSI**: 19s 12s 3s 39.2*** 11 2 0 4.8
- **GGUBS-HSI**: 20s 7s 2 22.4*** 13s 4 0 4.4
- **ADRAND-BMJ**: 38s 30s 17s 173*** 9 2 0 10.0
- **RANDU-BMJ**: 12 3 0 9.2 15s 3 0 11.2
- **L & L-BMJ**: 19s 7s 0 18.4** 13s 6s 1 7.6
- **M & M-BMJ**: 16s 5 2 12.4* 14s 6s 2 16.4*
- **URAND-BMJ**: 0 12 5 0 12.0 11 2 0 4.8
- **R. Shore-BMJ**: 15s 9s 1 20.8** 13s 6s 1 18. **
- **GGUBS-BMJ**: 18s 7s 0 16.4* 12 6s 1 7.2
- **GGNPM**: 12 5 0 14.8* 10 3 1 12.0
- **GGNML**: 5 20s 6s 2 20.0** 14s 3 1 8.8

*: Significant at level 0.2
**: Significant at levels 0.2 and 0.05.
**: Significant at levels 0.2, 0.05 and 0.01.
*: Significant at level ALPHA.
@: Total number of possible rejections is 12.
BMI: Stands for Box-Muller Inverse.
BMJ: Stands for Box-Muller Rejection.
HSI: Stands for Hasting Inverse.
Upper bound for the number of significant samples at levels 0.2, 0.05 and 0.01 are 12, 5, and 2, respectively.
Critical chi-square values for the whole generated series are 12.2, 16.9 and 21.7 for significance levels 0.2, 0.05 and 0.01, respectively.
The five generators and combinations mentioned above can be regarded as "bad" generators or "bad" combinations according to tests by length of runs up and down. They, however, perform well in tests by total number of runs. They have 1, 1, 0, 0 and 0 rejections respectively. Hence, the corresponding total numbers of rejections of these five generators and combinations in tests for runs up and down are 7, 7, 6, 6, and 6 out of 12 (Column 2). They can not be considered as "good" generators or "good" combinations. One other bad combination is that of R. Shore combined with BMJ; it has 8 significant values, 4 in tests by length, the other 4 in tests by total number of runs.

There are some good generators/combinations from the point of view of the runs up and down tests. ADRAND combined with the Box-Muller inverse routine, M & M combined with Hastings inverse and GRAND combined with Box-Muller rejection transformation all perform very well in the runs up and down tests. They pass both tests by length and tests by total number of runs at all levels of ALPHA. Next, RANDU combined with Box-Muller rejection and GGFM pass 11 types of tests out of 12 in runs up and down. Next to these two, M & M, and GGUBS with Box-Muller inverse are rejected only by 2 types out of 12.
RANDU combined with the Box-Muller inversion routine, RANDU with the Hastings inversion routine, and GGUBS with the Box-Muller rejection routine perform moderately well in the runs up and down tests. They pass 8 out of 12 of the tests. The Rest of the generators/combinations are rejected on more than 4 but less than 7 tests. Most rejections occur in tests by length of run. Thus, it appears that, tests by length of runs are more powerful than tests of the total number of runs.

With respect to runs above and below the median (in Table 3) most generators/combinations perform well. GGUBS alone, GGUBS with HSI, GGUBS with BMJ, L & L with BMJ and GGNML pass all 12 types of tests. In addition, there 14 of the generators/combinations pass 11 out of 12 tests. These are GGUBS with BMI, L & L, L & L with HSI, RANDU, RANDU with HSI, RANDU with BMJ, URAND, URAND with BMJ, URAND with HSI, R. Shore, R. Shore with BMI, R. Shore with HSI, M & M with BMJ and ADRAND with EMI.

However, ADRAND with BMJ, M & M with HSI, R. Shore with BMJ and M & M can be considered as "bad" combinations or generators based on the results of the runs above and below the median tests. They have 8, 7, 7 and 6 rejections respectively out of 12. This is especially true with respect to the tests by length of runs; ADRAND with BMJ has 50 sig-
<p>| Generators/ | No.of Sig Samp | By Length | By Total No. of Runs |</p>
<table>
<thead>
<tr>
<th>Combinations</th>
<th># of Sig Samp</th>
<th>Chi-S</th>
<th># of Sig Samp</th>
<th>Chi-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRAND</td>
<td>3</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RANDU</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>6</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>URAND</td>
<td>1</td>
<td>13</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>R. Shore</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3GUBS</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>2</td>
<td>13s</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>2</td>
<td>12</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>1</td>
<td>11</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3GUBS-BMI</td>
<td>1</td>
<td>14s</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>3</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>7</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>1</td>
<td>13s</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3GUBS-HSI</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>8</td>
<td>50s</td>
<td>50s</td>
<td>50s</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>1</td>
<td>16s</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>7</td>
<td>13s</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3GUBS-BMJ</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GGNPM</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GGNML</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*: Significant at level 0.2
**: Significant at levels 0.2 and 0.05.
***: Significant at levels 0.2, 0.05 and 0.01.
s: Significant at level ALPHA.
: Total number of possible rejections is 12.
BMI: Stands for Box-Muller Inverse.
BMJ: Stands for Box-Muller Rejection.
HSI: Stands for Hasting Inverse.
Upper bound for the number of significant samples at levels 0.2, 0.05 and 0.01 are 12, 5, and 2, respectively.
Critical chi-square values for the whole generated series are 12.2, 16.9 and 21.7 for significance levels 0.2, 0.05 and 0.01, respectively.
significant samples at all levels and has a large grand chi-square value 450. The next worse generators/combinations are URAND with BMJ, ADRAND alone and ADRAND with HSI at that order; their number of rejections are 4, 3, and 3 (Column 2).
Chapter V

SERIAL CORRELATION TESTS BY AUTOCORRELATION ANALYSIS

5.1 INTRODUCTION

Numbers generated on computers are not truly random. They may be serially correlated with each other if the parameters in the congruential generating equations are not selected properly or depending on the form of the equation. Hence the evaluation of the quality of the generator and the generator-transform combination requires the serial correlation tests.

Serial correlation tests can be used for evaluating the randomness of a sequence. If a sequence is correlated, then it is not random. On the other hand, if a sequence is not correlated, it may be but is not necessarily random. If the variate is normally distributed, its sequence is random if it is not serially uncorrelated. So the serial correlation test is one good randomness test, especially when the variables in the sequence are normal.

The autocovariance function attempts to answer the question "how much does the value of a series at time T depend on the value at a previous time, T-K". So serial correla-
tion or statistical dependence may be detected by the autocovariance function which gives the correlation between the observations separated by lag U units.\textsuperscript{5} Since periodicities in the autocovariance function do reveal similar periodicities in the original series itself, fluctuation in the autocovariance function can give an indication of the fluctuation in the original sequence itself (Chow, 1975). Hence autocovariance analysis can be used to test the sequence generated from computers for white noise.\textsuperscript{6} That is particularly true, when we want to detect "local correlation", that is, whether neighboring points of the sequence are correlated.

In the autocorrelation analysis used here, the autocorrelation coefficient will be used for each individual lag K and the Portmanteau chi-square statistic will be used to test for serial correlation for all lags collectively. If a series is uncorrelated, or if a process is purely random, the theoretical autocorrelation function at discrete lag K has the property: \( NR(K) = 1 \) (0) for \( K=0 \) (\( K\neq 0 \)) (Jenkins and Watts, p. 149, 1969). In addition, when the number of terms,

\textsuperscript{5}In this study, U denotes a continuous lag; K denotes a discrete lag.

\textsuperscript{6}The white noise series is a statistically independent series (or uncorrelated series) with expected value, \( E(X) = 0 \), and autocovariance function at lag K, \( R(K) = \text{variance of the series} \) for \( K=0 \)

\[ = 0 \quad \text{for} \quad K \neq 0 \]
in the series is sufficiently large, the distribution of sample autocorrelation coefficients can be taken to be normally distributed with mean zero and variance $1/N$ (Anderson, 1942). Thus the standard normal table can be used for establishing the theoretical confidence interval for the sample autocorrelation function for the white noise hypothesis tests.

Rather than consider only the autocorrelation function at each lag $K$ individually, a test is often needed of the first $M$ autocorrelations of the series, taken as a whole, for checking independence of the whole generated series. If the series is random, then the sample autocorrelation function $\text{COR}(K)$ is uncorrelated with variance $V(\text{COR}(K)) = (N-K)/(N^2(N+2)) = 1/N$. From which it follows that the statistic computed from the first $(M-1)$ autocorrelation functions $\text{PORT} = N \times (N-2) \times \sum (\text{COR}(K) \times 2/(N-K))$ (for Lag $K=1,2,\ldots,(M-1)$) would, for large sample size $N$, be distributed as chi-square with $M-1$ degrees of freedom (Box and Pierce, 1970). Thus the chi-square distribution table can be used for the hypothesis tests.

5.2 The Autocovariance Function and the Autocorrelation Function

In general, a stationary stochastic process $X(T)$ with mean zero, variance $\text{VAR}$, and size $N$ has an autocovariance
function (ACVF) and an autocorrelation function (ACF). The ACVF at lag $U$ may be defined as

$$P(U) = \lim_{N \to \infty} \frac{1}{N} \int_{-N/2}^{N/2} X(T)X(T+U) \, dT = \mathbb{E}(X(T)X(T+U)),$$

for $-N/2 < T < N/2$, and $-T < U < +T$ (5-1)

and shows how the dependence between adjacent values in the stochastic process changes with lag $U$.

The autocorrelation function can be obtained by normalizing the autocovariance function

$$NR(U) = R(U)/R(0) = R(U)/\text{VAR}.$$  \hspace{1cm} (5-2)

From Equation 5-2, we can obtain

$$R(U) = \text{VAR} \cdot NR(U).$$

The plot of the function $NR(U)$ against lag $U$ is called the autocorrelogram. For a continuous process, the lag $U$ can take on any value from positive infinity to negative infinity, but for a discrete process, it will be defined for only integer value of $U$. $K$ will denote an integer value of $U$ throughout this paper.

If a series is uncorrelated, or if a process is a purely random process, the autocorrelation function at discrete lag $K$ has the following property

$$NR(K) = 1 \text{ for } K=0,$$

$$0 \text{ for } K \neq 0.$$
This will be used as one of the criteria for testing the series for autocorrelation.

5.3 **ESTIMATE OF AUTOCOVARIANCE FUNCTION**

In practice, it is impossible to know these covariance functions exactly, and hence it is necessary to estimate them from records of finite length. Suppose \( x(t) \) is one sample record of the stochastic process with sample mean zero, sample variance \( S \), and sample size \( N \). According to Bendat and Piersol (1966), an estimate for the autocovariance between the values of \( x(t) \) at times \( T \) to \( T+U \) may be obtained by taking the product of two values and averaging over the observations \( N \). The resulting average product will approach an exact autocovariance function as \( N \) approaches infinity. In equation form

\[
R(U) = \lim_{N \to \infty} \frac{1}{N} \sum_{T} x(T) x(T+U) dT,
\]

for \(-N/2 < T < N/2\), and \(-T < U < +T\). (5-3)

5.3.1 **Sample Autocovariance Function**

The sample ACVF was seen to arise quite naturally as an estimate of the theoretical autocovariance function. From Equation 5-1, the estimator may be written

\[
\text{COV}(U) = \frac{1}{N} \int_{0}^{N-U} (x(T) x(T+U)) dT, \quad (5-4)
\]

---

7 The autocovariance function is an even function.
for $0 < |U| < n$, and $0 < T < n-|U|$.

for the continuous case. For the discrete case,

\[ COV(K) = \frac{1}{N} \cdot \text{SUM}(x(T) \cdot x(T+|K|)) \]

for $K=0,1,2,\ldots,+ (N-1)$, and $T=1,2,\ldots,(N-|K|)$ or

\[ COV(K) = 0, \text{ otherwise} \]

An alternative estimator which is also widely used may be written

\[ COV1(U) = \left( \frac{1}{(N-|U|)} \right) \int_{0}^{N-|U|} x(T) \cdot x(T+|U|) \, dT, \]

for $0 < |U| < n$, and $0 < T < n-|U|$. 

= 0, otherwise \hspace{1cm} (5-6)

for the continuous case, and

\[ COV1(K) = \frac{\text{SUM}(x(T) \cdot x(T+|K|))}{(N-|K|)}, \]

for $T=1,2,\ldots,(N-|K|)$, $K=0,1,2,\ldots,+ (N-1)$

=0, otherwise \hspace{1cm} (5-7)

for the discrete case.

The estimators $COV(U)$ and $COV1(U)$ have been used in statistical work mainly because they have intuitive appeal, not because they are necessary best in any known sense. However, these estimators may be compared according to some criteria, such as minimum mean square error, and are generally the best of the available estimators (Jenkins and Watts, 1968).

$COV1(U)$ is an unbiased estimator of $R(U)$ because
\[ E(\text{COV}(U)) = R(U), \quad |U| < N \quad \text{(5-8)} \]

\[ 0, \quad |U| \geq N \]

However, the estimator \( \text{CCV}(U) \) is only asymptotically unbiased as the record length \( N \) tends to infinity because

\[ E(\text{COV}(U)) = R(U) \times (1 - (U/N)), \quad |U| < N \quad \text{(5-9)} \]

\[ 0, \quad |U| \geq N \]

Since \( \text{COV}(U) \) has a smaller mean square error than \( \text{COV}1(U) \), the biased estimator will be used in this paper to compute the autocovariance at each discrete lag \( K \) (Jenkins and Watts, 1968).

5.3.2 Properties of \( \text{COV}(U) \)

1. \( \text{COV}(U) \) is a biased and a symmetric estimator, but it has a smaller mean square error than \( \text{COV}1(U) \).

2. Except for purely random numbers or white noise series, correlation estimators are highly autocorrelated. This can produce misleading correlation function estimates in the sense that they do not reduce as rapidly as they should. Thus spectral analysis will be used (next Chapter) as a complement to autocorrelation analysis (Fishman and Kiviat, 1971).

3. The estimator is consistent so that, given a very large record, the estimator tends to the true covariance function.
4. When the number of terms, $N$, in the series is sufficiently large, the distribution of sample autocorrelation coefficients can be taken to be normal with mean zero and variance $1/N$. Thus the standard normal table and the theoretical confidence interval can be used for the hypothesis tests to check whether or not the generated sequence is autocorrelated.

5. It is a descriptive and nonparametric statistic.

5.4 PROcedURES OF AUTOCORRELATION ANALYSIS

Two approaches are used here for testing the generated series for white noise. They are tests of the number of significant autocorrelation coefficients and the Portmanteau chi-square statistic of the autocorrelation functions. The following two subsections will discuss each, respectively.

5.4.1 Number of Significant Autocorrelation Coefficients

Three procedures are used for testing the series for white noise by means of this approach.

1. Test each set for independence.
If the series is independent, or is a purely random process, then its autocorrelation function at discrete time lag \( K \) is

\[
\begin{align*}
NR(K) &= 1, \quad \text{for } K=0, \\
&= 0, \quad \text{for } K \neq 0.
\end{align*}
\]

In addition when the sample size is large, the distribution of sample autocorrelations approaches a normal distribution with mean zero and variance \( 1/N \). This test is conducted as follows.

a) Determine the sample size, \( N \), of each set.

b) Generate the series \( x \); treat the series as a discrete series with sampling time one. Subtract the sample mean from each term such that the mean of the series \( Mx=0 \).

c) Determine a truncation point, \( M \). To allow use of the central limit theorem and spectral analysis, \( M=64 \) for lags \( K=0,1,2,\ldots,63 \) is used in this paper.

d) Compute sample autocovariance functions

\[
\text{COV}(K) = \left( \frac{\sum (x(T) \cdot x(T+K))}{N} \right),
\]

for \( K=0,1,2,\ldots,(M-1) \), and \( T=1,2,\ldots,(N-K) \).

e) Compute sample autocorrelation functions

\[
\text{COR}(K) = \frac{\text{COV}(K)}{\text{COV}(0)}, \quad \text{for } K=1,2,\ldots,(M-1).
\]
f) Set up the significance level ALPH A; let ALPH A=0.01, 0.05 and 0.2.

g) Set up the theoretical confidence interval for the sample autocorrelation function, COR(K). When the number of terms in the series is sufficiently large, the distribution of sample autocorrelations approaches a normal distribution with mean zero and variance 1/N, i.e. COR(K) \( \sim N(0, 1/N) \). Then the hypothetical confidence interval is

\[
N_R(K) - Z^* (1/N)^{**.5} < COR(K) < N_R(K) + Z^* (1/N)^{**.5}
\]

If a series is white noise, then \( N_R(K)=0 \) for all \( K=0 \), and the confidence interval for white noise is

\[
- Z^*(1/N)^{**0.5} < COR(K) < + Z^*(1/N)^{**0.5}
\]

At significance levels ALPH A=0.01, 0.05 and 0.2, the value of \( Z=2.325, 1.96, \) and 1.28, respectively for the two tailed test.

h) Count the total number of significant correlations observed (i.e., where COR(K) does not lie within the theoretical confidence interval).

i) At significance level ALPH A, compute the expected number of significant COR(K), EXP, and upper bound UCINOL, over M-1=63 lags, respectively:

\[
EXP = ALPH A \times (M-1)
\]
UCINOL = EXP + Z*((M-1)*ALPHA*(1-ALPHA))**0.5

where \( Z \) is the number of standard deviations corresponding to the confidence interval.

The hypothesis to be tested and its alternative are as follows:

\( H_0: \) The series \( x \) is white noise at significance level \( \text{ALPHA} \) if the number of significant correlations \( \text{CCF}(K) \) is less than or equal to the upper bound, \( \text{UCINCL} \), for significant \( \text{COR}(K) \).

\( H_A: \) The series \( x \) is not white noise at significance level \( \text{ALPHA} \) if the number of significant \( \text{COR}(K) \) is greater than upper bound, \( \text{UCINCL} \), of significant correlations.

2. Test each sample of \( KL=50 \) sets for independence using the chi-square goodness of fit test.

Since the distribution of significant \( \text{COR}(K) \) values, \( K=1,2,\ldots,(M-1) \), is uniform over the \( KL \) sets generated (Burford and willis, 1978), a chi-square test can be used to test whether or not the distribution of significant \( \text{COR}(K) \) is uniform over \( KL=50 \) sets. If it is, then the process of \( KL=50 \) sets is white noise; i.e., it is uncorrelated.

The procedure is accomplished as follows:
a) For each lag \( K=1,2,\ldots,(M-1)=63 \), count the number of significant \( \text{COR}(K) \) values, \( \text{CTLAG1}(K) \), over the \( KL \) sets as done in the first procedure in this section.

b) For each lag \( K \) compute the expected number of significant \( \text{COR}(K) \), \( \text{EXP} \), over the \( KL \) sets at significance level, \( \text{ALPHA} \).

\[
\text{EXP} = KL \times \text{ALPHA}
\]

c) Compute the chi-squared value, \( \text{CHICOR} \)

\[
\text{CHICOR} = \text{SUM} \left( \frac{\text{EXP} - \text{CTLAG1}(K)}{\text{EXP}} \right)^2 / \text{EXP},
\]

for \( K=1,2,\ldots,(M-1) \).

d) Find the chi-square critical value for a hypothesis test at \( (M-1)-1=62 \) degrees of freedom and significance level \( \text{ALPHA} \). In this study, we used \( \text{CRT}=90.8,\ 81.4,\ \text{and}\ 71.1 \) for \( \text{d.f.}=(M-1)-1=62 \) and \( \text{ALPHA}=0.01,\ 0.05 \) and \( 0.2 \), respectively.

The hypothesis to be tested and its alternative are as follows:

\( \text{HO} \): The process of all \( KL=50 \) sets is white noise at significance level \( \text{ALPHA} \) if \( \text{CHICOR} \) is less than or equal to its critical value, \( \text{CRT} \).

\( \text{HA} \): The process of all \( KL=50 \) sets is correlated at significance level \( \text{ALPHA} \) if \( \text{CHICOR} \) is greater than its critical value, \( \text{CRT} \).
3. Test the whole generated series of ISET=50 samples for independence by number of significant samples counted above and by means of the CS goodness of fit test.

At significance level ALPHA, we use two ways to judge whether or not the whole generated series is random. First, if the whole series of ISET=50 samples is random, then the number of significant samples counted over ISET=50 samples is less than or equal to the upper bound for number of significant samples

\[(ISET*ALPHA + Z*(ISET*ALPHA*(1-ALPHA))**.5)\].

Second, if the whole generated series of ISET samples is a random sequence, the ISET=50 CHICOR values computed from all samples are from a chi-square distribution with \((M-1)-1=62\) degrees of freedom. Thus we can use a chi-square goodness of fit test to check whether or not the ISET=50 CHICOR values have a chi-square distribution with \((M-1)-1=62\) degrees of freedom. The expectation of this procedure is as follows:

a) Use the second procedure in this section to compute the CHICOR statistic for each sample. There are ISET=50 CHICOR statistics which have chi-square distributions with \((M-1)-1=62\) degrees of freedom if the whole series is random.
b) Using the conventional minimum expected value 5 and Cochran's moderate number of classes between 10 to 25, establish $K_2$ mutually exclusive and exhaustive categories and let $K_2=10$ for $ISET=50$ in this study.

c) Use equal probability $1/10$ to compute the expected frequencies $E_i$ for the $i$th category. Thus, $E_i=5$ for $ISET=50$ and $P_i=0.1$.

d) Under the null hypothesis and equal probability $1/10$ per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, $O_i$, for the $i$th category in $K_2$ categories.

f) Compute the CS statistic

$$CS = \sum \left( \frac{(E_i - O_i)^2}{E_i} \right), \text{ for } i = 1, 2, \ldots, K_2=10.$$ 

g) Set up significance levels $\alpha=0.01, 0.05$ and 0.2 for the hypothesis tests.

h) Find the critical values, $CRT$, for the hypothesis tests. At significance levels $\alpha=0.01$, 0.05 and 0.2 with $K_2-1=9$ degrees of freedom, $CRT=21.7, 16.9$ and 12.2, respectively.
The hypothesis to be tested and its alternative are as follows:

**H0:** The whole series of $ISET=50$ samples is random at the significance level $\text{ALPHA}$ if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to $\text{CRT}$.

**HA:** The whole series of $ISET=50$ samples is not random at the significance level $\text{ALPHA}$ if number of significant samples is greater than its upper bound, or if the computed CS value is greater than $\text{CRT}$.

5.4.2 **Portmanteau Chi-Square Statistic**

Three procedures are used in this approach to test the series as to whether or not it is independent.

1. Test each set for independence.

   If the series is white noise, then the Portmanteau chi-square statistic of its autocorrelation functions at lag $K$, $K=1,2,\ldots,(M-1)=63$, is less than or equal to the chi-square critical value with $M-1=63$ degrees of freedom.

This procedure is expected as follows:
a) Determine the sample size, $N$, of each set.

b) Generate the series $x$; treat the series as a discrete time series with sampling time one. Subtract the sample mean from each term such that mean of the series, $M_x=0$.

c) Determine a truncation point, $M$. To allow use of the central limit theorem and the spectral analysis in this paper, set $M=64$ for lag $K=0,1,2,\ldots,63$.

d) Compute sample autocovariance functions

$$COV(K)=(1/N)*(\sum x(T)x(T+K)),$$

for $K=0,1,2,\ldots,(M-1)$ and $T=1,2,\ldots,(N-K)$.

e) Compute sample autocorrelation functions

$$COR(K)=COV(K)/COV(0), \text{ for } K=1,2,\ldots,(M-1).$$

f) Compute the Portmanteau chi-square statistic,

$$PTMANT=M*(N+2)*\sum (COR(K)*2/(N-K)),\text{ for } K=1,2,\ldots,(M-1)=63.$$

g) Set up significance levels $\alpha=0.01, 0.05$ and $0.2$ for the hypothesis tests.

h) Find the chi-square critical value, $CRT$. For $df=M-1=63$ and significance levels $\alpha=0.01, 0.05$ and $0.2$, $CRT=92.0, 82.5$ and $72.2$, respectively.
The hypothesis to be tested and its alternative are as follows:

H0: The series $x$ is uncorrelated at significance level $\alpha$ if the Portmanteau chi-square statistic, $PT\text{MANT}$, is less than or equal to $CRT$, the critical value.

HA: The series $x$ is correlated at significance level $\alpha$ if the chi-square Portmanteau statistic, $PT\text{MANT}$, is greater than $CRT$.

2. Test each sample of $KL=50$ sets for independence by means of the CS goodness of fit test.

If the sample is random, then the $KL=50$ $PT\text{MANT}$ statistics computed from each set have a chi-square distribution with $M-1=63$ degrees of freedom.

This procedure is expected as follows:

a) Use the first procedure in this section to compute chi-square statistics for each set. If the sample is random, there are $KL=50$ sample statistics which have chi-square distributions with $M-1=63$ degrees of freedom.

b) Using the conventional minimum expected value 5 and Cochran's moderate number of classes bet-
ween 10 to 25, establish $K_1=10$ mutually exclusive and exhaustive categories for $K_L=50$ sets.

c) Use equal probability $1/10$ to compute the expected frequencies $E_i$ for the $i$th category. Thus, $E_i=5$ for $K_L=50$ and $P_i=0.1$.

d) Under the null hypothesis and equal probability $1/10$ per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, $O_i$, in the $i$th category.

f) Compute the CS statistic

$$CS = \sum \left( \frac{(F_i - C_i)^2}{E_i} \right), \text{ for } i=1,2,\ldots,K_1=10.$$ 

g) Set up significance levels $\alpha=0.01$, 0.05 and 0.2 for the hypothesis test.

h) Find the critical value, $C_R$, for the hypothesis test. At significance levels $\alpha=0.01$, 0.05 and 0.2 with $K_1-1=9$ degrees of freedom, $C_R=21.7$, 16.9 and 12.2.

The hypothesis to be tested and its alternative are as follows:
H0: The sample of KL = 50 sets is independent at significance level ALpha if the computed CS value is less than or equal to CRT.

HA: The sample is not independent at significance level ALpha if the computed CS value is less than or equal to CRT.

3. Test the whole series of ISET=50 samples for independence by number of significant samples counted above and by means of the other CS goodness of fit test.

At significance level ALpha, we use two ways to judge whether or not the whole generated series is random. First, If the whole series of ISET=50 samples is random, then the number of significant samples counted over ISET=50 samples is less than or equal to the upper bound for number of significant samples (ISET*ALPHA+Z*(ISET*ALPHA*(1-ALPHA))**.5). Second, If the whole generated series of ISET samples is a random sequence, the ISET=50 CS values have a chi-square distribution with K1-1=9 degrees of freedom. Thus we can use a chi-square goodness of fit test to check whether or not the ISET=50 CS values have a chi-square distribution with K1-1=9 degrees of freedom. This procedure is expected as follows:
a) Use the second procedure in this section to compute the CS statistic for each sample. There are ISET=50 CS statistics which have a chi-square distribution with K1-1=9 degrees of freedom if the whole series is random.

b) Using the conventional minimum expected value 5 and Cochran's moderate number of classes between 10 to 25, establish K2=10 mutually exclusive and exhaustive categories.

c) Use equal probability 1/10 to compute the expected frequencies Ei for the ith category. Thus, Ei=5 for ISET=50 and Pi=0.1.

d) Under the null hypothesis and equal probability 1/10 per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, Oi, for the ith category in K2 categories.

f) Compute the CS statistic

\[ CS = \sum \left( \frac{(E_i - O_i)^2}{E_i} \right), \text{ for } i=1,2, \ldots, K2=10. \]

g) Set up significance levels ALPHA=0.01, 0.05 and 0.2 for the hypothesis tests.

h) Find the critical values, CRT, for the hypothesis tests. At significance levels ALPHA=0.01,
0.05 and 0.2 with $K_2-1=9$ degrees of freedom, \( \text{CRT}=21.7, 16.9 \) and 12.2, respectively.

The hypothesis to be tested and its alternative are as follows:

\( H_0 \): The whole series of ISET=50 samples is random at the significance level ALPHA if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to CRT.

\( H_A \): The whole series of ISET=50 samples is not random at the significance level ALPHA if number of significant samples is greater than its upper bound, or if the computed CS value is greater than CRT.

5.5 **EMPIRICAL TEST RESULTS, DISCUSSIONS AND EVALUATIONS**

Results of the autocorrelation tests are listed in Table 4. The information in Table 4 was derived from 50 independent samples with 50 sets per sample and 1200 numbers per set.

For each set, we computed 63 sample autocorrelation functions for lag 1 to lag 63. At significance level ALPHA, we counted the number of significant lags by comparing the sam-
ple autocorrelation functions with its corresponding theoretical confidence interval.

For each sample, we derived a distribution of significant lags for lags 1, 2, ..., 63 over 50 sets. In order to check whether or not the distribution of significant lags has a uniform distribution with $50 \times \text{ALPHA}$ expected as the number of significant lags, a chi-square calculated value was computed with 62 degrees of freedom. Then we tested for significant samples by comparing the observed chi-square value with its critical value at significance level $\text{ALPHA}$. The critical chi-square values with 62 degrees of freedom for levels 0.2, 0.05, and 0.01 are 71.1, 81.4, and 90.8, respectively. Columns 7, 8, and 9 in Table 4 outline the numbers of significant samples at levels 0.2, 0.05, and 0.01. Their corresponding upper bounds are 12, 5, and 2. Columns 10, 11, and 12 exhibit the grand chi-square values with 9 degrees of freedom. They are derived by fitting 50 independent chi-square values with 62 degrees of freedom into a chi-square distribution with 9 degrees of freedom. These grand chi-square values are used to evaluate whether or not the whole generated series is correlated.

The information in columns 3, 4, 5, and 6 is the results of Portmanteau chi-square tests. Based on the 63 autocorrelation functions taken as a whole, we derived one chi-square
value with 63 degrees of freedom for each set. Then from 50 chi-square values with 63 degrees of freedom for all sets, we computed one chi-square value with 9 degrees of freedom for each sample. Then we compared the observed chi-square value with its critical value for each sample, the result is the number of significant samples for the whole series. The numbers of significant samples are arranged in Columns 3, 4, and 5 for significance levels 0.2, 0.05, and 0.01, respectively. Finally, we computed one grand chi-square value with 9 degrees of freedom from all samples for each whole generated series. The grand chi-square values with 9 degrees of freedom for the whole generated series are presented in Column 6.

Like Tables 2 or 3, Table 4 shows 12 types of possible rejections. Six of them are derived from Portmanteau chi-square tests; the other six are from the distribution of significant lags. Three of each six are indicated by number of significant samples; the other three are shown by overall chi-square values. The observed total numbers of rejections from the autocorrelation analysis are listed in Column 2.

The results of the Portmanteau chi-square tests show that ADRAND combined with BMJ is the worst combination. It is rejected at all significance levels either from the point of view of the numbers of significant samples or the grand
Table 4

Portmanteau Test and Distribution of Lags

| Generators/ Combinations | | No. | Portmanteau Tests | | Distribution of Lags | | Chi-square-Values |
|---------------------------|-----------------|-----|-------------------|-------------------|------------------|------------------|
|                           |                 | No. | Sig. | Chi-S. | Sig. | Chi-S. | Sig. | Chi-S. |
|                           |                 |     | S    | Values | S    | Values | S    | Values |
|                           |                 |     | 0.2  | 0.05  | 0.01 | 0.2  | 0.05  | 0.01  |
|                           |                 |     | (1)  | (2)  | (3)  | (4)  | (5)  | (6)  | (7)  | (8)  | (9)  | (10) | (11) | (12) |
| AD RAND                  | 3               | 9   | 2    | 0     | 4.0  | 1    | 0    | 2    | 49.2s | 39.0s | 69.2s |
| RANDU                   | 3               | 6   | 2    | 0     | 8.0  | 1    | 0    | 1    | 100.8s | 20.8s | 66.8s |
| L & L                   | 3               | 16s | 3    | 1     | 5.2  | 1    | 1    | 1    | 102.8s | 8.8   | 23.2s |
| M & M                   | 4               | 12  | 4    | 0     | 12.8*| 0    | 0    | 1    | 92.4s  | 18.0s | 60.0s |
| URAND                   | 3               | 11  | 5    | 1     | 4.4  | 0    | 0    | 0    | 130.0s | 13.6  | 43.2s |
| R. Shore                | 4               | 15s | 6s   | 1     | 8.8  | 0    | 1    | 2    | 70.0s  | 14.4  | 22.0s |
| GGBUS                   | 3               | 15s | 4    | 2     | 7.2  | 1    | 3    | 2    | 69.6s  | 6.4   | 22.8s |
| AD RAND-BMI             | 3               | 8   | 2    | 0     | 2.8  | 1    | 1    | 2    | 107.2s | 26.4s | 38.0s |
| RANDU-BMI               | 2               | 8   | 2    | 0     | 6.8  | 1    | 1    | 0    | 50.3s  | 14.8  | 42.0s |
| L & L-BMI               | 6               | 11  | 6s   | 2     | 17.2*| 0    | 0    | 1    | 119.6s | 17.2s | 69.2s |
| M & M-BMI               | 4               | 11  | 4    | 1     | 12.8*| 0    | 1    | 2    | 103.6s | 35.6s | 60.4s |
| URAND-BMI               | 3               | 9   | 1    | 0     | 6.8  | 0    | 0    | 1    | 99.4s  | 32.8s | 84.0s |
| R. Shore-BMI            | 3               | 10  | 3    | 1     | 6.8  | 0    | 2    | 1    | 112.8s | 17.2s | 50.0s |
| GGBUS-BMI               | 3               | 10  | 5    | 1     | 11.6 | 0    | 0    | 1    | 121.2s | 20.0s | 79.2s |
| AD RAND-HSI             | 3               | 10  | 4    | 0     | 9.6  | 0    | 0    | 1    | 71.2s  | 35.6s | 56.4s |
| RANDU-ksi               | 3               | 6   | 2    | 0     | 12.8*| 0    | 0    | 0    | 109.6s | 16.0  | 59.2s |
| L & L-ksi               | 4               | 14s | 1    | 1     | 5.2  | 0    | 1    | 1    | 92.8s  | 22.8s | 37.6s |
| M & M-ksi               | 3               | 10  | 1    | 1     | 9.2  | 0    | 1    | 5s   | 125.6s | 16.4  | 41.2s |
| URAND-ksi               | 4               | 11  | 7s   | 1     | 14.8*| 1    | 0    | 2    | 133.6s | 14.4  | 53.2s |
| R. Shore-ksi            | 5               | 5   | 1    | 0     | 21.2*| 0    | 0    | 2    | 67.6s  | 16.8s | 30.3s |
| GGBUS-ksi               | 4               | 13s | 1    | 1     | 3.3  | 1    | 1    | 1    | 76.8s  | 19.2s | 30.4s |
| AD RAND-BMJ             | 12              | 42s | 3    | 1     | 23s  | 2     | 4*   | 3    | 71.6s  | 14.0  | 43.2s |
| RANDU-BMJ               | 3               | 8   | 2    | 1     | 12.4*| 3    | 1    | 1    | 71.6s  | 14.0  | 43.2s |
| L & L-BMJ               | 5               | 16s | 2    | 1     | 16.0*| 0    | 0    | 1    | 182.4s | 35.2s | 40.4s |
| M & M-BMJ               | 3               | 7   | 1    | 0     | 8.8  | 1    | 3    | 0    | 98.0s  | 34.4s | 45.2s |
| URAND-BMJ               | 4               | 14s | 2    | 0     | 5.2  | 0    | 1    | 0    | 96.4s  | 28.4s | 39.4s |
| R. Shore-BMJ            | 3               | 12  | 3    | 0     | 11.2 | 0    | 2    | 0    | 96.6s  | 20.0s | 25.6s |
| GGBUS-BMJ               | 6               | 16s | 2    | 1     | 17.2*| 0    | 0    | 2    | 181.2s | 32.0s | 39.4s |
| GGSMF                   | 3               | 7   | 3    | 1     | 5.2  | 0    | 0    | 0    | 92.6s  | 42.0s | 43.2s |
| GGHAJ                   | 4               | 13s | 1    | 1     | 8.8  | 0    | 1    | 2    | 72.6s  | 16.4s | 32.8s |

*: Significant at level 0.2
**: Significant at levels 0.2 and 0.05.
***: Significant at levels 0.2, 0.05 and 0.01.
s: Significant at level ALPHA.
3: Total number of possible rejections is 12.
BMI: Stands for Box-Muller Inverse.
BMJ: Stands for Box-Muller Rejection.
HSI: Stands for Hastings Inverse.
Upper bound for the number of significant samples at levels 0.2, 0.05 and 0.01 are 12, 5, and 2, respectively.
Critical chi-square values for the whole generated series are 12.2, 16.9 and 21.7 for significance levels 0.2, 0.05 and 0.01, respectively.
chi-square values. In addition, it generated the greatest numbers of significant samples, 42, 31, and 23 for corresponding levels 0.2, 0.05, and 0.01 and the largest grand chi-square value, 245. The worse generators or combinations are L & L with BMI, GGUBS with BMJ, URAND with HSI, R. Shore alone, R. Shore with HSI, and L & L with BMJ. These are rejected by 3, 3, 2, 2, 2, out of 6 possible significant tests, respectively.

Thirteen generators or combinations pass the tests at all levels from the point of view of numbers of significant samples as well as the grand chi-square values. These are ADRAND alone, RANDU alone, URAND alone, ADRAND with BMI, RANDU with BMI, URAND with BMI, R. Shore with BMI, GGUBS with BMI, ADRAND with HSI, M & M with HSI, M & M with BMJ, R. Shore with BMJ, and GGNEM alone. The rest of the generators/combinations pass 5 out of 6 tests.

Considering the distribution of significant lags, most generators or combinations pass the tests from the viewpoint of numbers of significant samples. Again ADRAND with BMJ is the worst one. It neither passes any sample nor passes any level. Each sample generated from it has a very large autocorrelation in lag 4. So the chi-square values for each sample and for each whole generated series are big significant. In lags other than 4 its chi-square values are
not so large. URAND and M & M with HSI have, respectively, 3 and 5 significant samples over the upper bound, 2, at level 0.01. The rest of the generators or combinations pass the number of significant lags tests at all levels.

Looking at the grand chi-square values, however, no generator nor combination passes at significance level 0.2 or level 0.01. Especially at level 0.2, each generator or combination has a very large grand chi-square value compared to its critical value, 12.2. One of the primary reasons is, with the exception of ADRAND with BMJ, that they generate fewer autocorrelated samples than expected (Column 7, 8, 9 in Table 4 explain this). The other reason is that the chi-square values computed from the distribution of significant lags for each sample are too small for all cases except for ADRAND with BMJ. Most sample chi-square values fall with the first two intervals in their chi-square distributions with 9 degrees of freedom. That is most chi-square values are less than 48.23 or between 48.23 and 52.49. These are shown in Tables 5, 6, and 7 for levels 0.2, 0.05, and 0.01, respectively. For instance, there are 50 chi-square values generated from ADRAND at level 0.2; 15 of them fall in interval 1 and 11 fall in interval 2 while the expected number of chi-square values in each category is 5. Consequently, ADRAND generates a large grand chi-square value, 49.2 (critical value, 12.2) for the whole series, and
it is rejected on the randomness hypothesis test. Hence, the generators or the combinations are rejected because they generate too few autocorrelated samples to be considered random.

Above all, ADRAND with BMJ is the worst combination in terms of autocorrelation analysis with 12 rejections. Next worse ones are L & L with BMI and GGUBS with BMJ, with 6 rejections each. The best combination is RANDU with BMI; it has 2 rejections. The other combinations do not pass the tests having 3 to 4 kinds cut of 12 (Column 2).
### Table 5

**Distribution of 50 Chi-squares from Dist. of Sig. Lags**

At Significance Level 0.2

<table>
<thead>
<tr>
<th>Generators/Combinations</th>
<th># of Chi-Square Values in ith Category</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRAND</td>
<td>15 11 8 8 2 4 1 0 1 0</td>
<td>49.2s</td>
</tr>
<tr>
<td>RANDU</td>
<td>25 7 4 7 3 2 1 0 1 0</td>
<td>100.8s</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>24 12 4 4 1 1 3 0 0 1</td>
<td>102.8s</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>22 11 9 1 3 0 4 0 0 0</td>
<td>92.4s</td>
</tr>
<tr>
<td>URAND</td>
<td>28 8 4 4 4 2 0 0 0 0</td>
<td>130.0s</td>
</tr>
<tr>
<td>R. Shore</td>
<td>20 10 5 4 3 7 1 0 0 0</td>
<td>70.0s</td>
</tr>
<tr>
<td>GGUUBS</td>
<td>23 11 4 3 3 2 3 0 0 1</td>
<td>89.6s</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>24 13 5 3 2 1 1 0 1 0</td>
<td>107.2s</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>16 9 11 5 3 3 1 1 0 1</td>
<td>50.8s</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>26 11 5 0 4 3 0 1 0 0</td>
<td>119.6s</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>24 7 11 4 2 0 1 1 0 0</td>
<td>103.6s</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>24 5 10 4 2 3 1 1 0 0</td>
<td>96.4s</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>26 6 8 3 5 0 2 0 0 0</td>
<td>112.8s</td>
</tr>
<tr>
<td>GGUUBS-BMI</td>
<td>25 14 5 1 2 2 0 1 0 0</td>
<td>121.2s</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>21 8 7 3 5 3 3 5 0 0</td>
<td>71.2s</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>26 8 6 3 2 1 2 2 0 0</td>
<td>109.6s</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>23 11 4 6 1 3 1 1 0 0</td>
<td>92.8s</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>27 10 5 2 2 4 0 0 0 0</td>
<td>125.6s</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>29 6 3 3 2 4 1 1 1 0</td>
<td>133.6s</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>19 11 7 4 6 1 2 0 0 0</td>
<td>67.6s</td>
</tr>
<tr>
<td>GGUUBS-HSI</td>
<td>21 11 4 6 3 2 1 1 1 0</td>
<td>76.0s</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>0 0 0 0 0 0 0 0 0 0</td>
<td>50 450.0s</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>20 12 4 3 5 1 2 0 3 0</td>
<td>71.6s</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>33 6 4 4 3 0 2 0 1 0</td>
<td>182.4s</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>22 15 2 3 3 2 2 0 1 0</td>
<td>98.0s</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>23 8 11 2 1 3 2 0 0 0</td>
<td>96.4s</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>22 14 4 3 5 0 2 0 0 0</td>
<td>96.8s</td>
</tr>
<tr>
<td>GGUUBS-BMJ</td>
<td>33 6 4 3 0 2 1 1 0 0</td>
<td>181.2s</td>
</tr>
<tr>
<td>GGNPM</td>
<td>20 16 6 4 1 0 2 1 0 0</td>
<td>92.8s</td>
</tr>
<tr>
<td>GGNML</td>
<td>20 12 4 6 3 2 1 1 1 0</td>
<td>72.4s</td>
</tr>
</tbody>
</table>

s: Stands for significant series.

Critical chi-square value for grand chi-square value is 12.2.

Total number of chi-square values in each sample is 50.

The expected number of chi-square values in each interval is 5.

Bounding Chi-Square Values with equal probability for 10 intervals are:
1: under 46.23, 2: under 52.49, 3: under 55.72
4: under 58.58, 5: under 61.33, 6: under 64.18, 7: under 67.32,
8: under 71.11, 9: under 76.62, 10 above 76.62.
### Table 6

Distribution of 50 Chi-squares from Dist. of Sig. Lags.

At Significance Level 0.05

<table>
<thead>
<tr>
<th>Generators/Combinations</th>
<th># of Chi-Square Values in ith Category</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRAND</td>
<td>7 5 5 6 4 2 5 4 0 2 30.0s</td>
<td></td>
</tr>
<tr>
<td>RANDU</td>
<td>9 7 6 8 8 7 0 3 1 1 20.8</td>
<td></td>
</tr>
<tr>
<td>L &amp; L</td>
<td>6 5 8 3 6 8 5 5 3 1 8.8</td>
<td></td>
</tr>
<tr>
<td>M &amp; M</td>
<td>8 4 6 11 2 5 6 6 1 1 18.0s</td>
<td></td>
</tr>
<tr>
<td>URAND</td>
<td>7 5 6 3 9 8 2 5 5 0 13.6</td>
<td></td>
</tr>
<tr>
<td>R. Shore</td>
<td>11 7 4 3 5 3 7 6 2 2 14.4</td>
<td></td>
</tr>
<tr>
<td>GGUBS</td>
<td>6 6 7 2 5 8 5 5 3 3 6.4</td>
<td></td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>9 10 8 7 8 0 4 0 2 2 26.4s</td>
<td></td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>11 9 6 4 2 3 4 3 4 4 14.8</td>
<td></td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>12 7 2 6 2 4 3 7 4 3 17.2s</td>
<td></td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>10 10 8 12 2 3 1 2 1 1 35.6s</td>
<td></td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>12 4 12 7 6 4 2 1 2 2 32.8s</td>
<td></td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>13 7 4 3 4 5 5 3 3 3 17.2s</td>
<td></td>
</tr>
<tr>
<td>GGUBS-BM</td>
<td>12 7 3 5 2 3 4 9 3 2 20.0s</td>
<td></td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>14 3 9 8 6 2 6 1 1 0 35.6s</td>
<td></td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>10 6 6 5 9 5 3 3 3 0 16.0</td>
<td></td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>13 9 6 3 3 4 3 5 3 1 22.8s</td>
<td></td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>8 10 6 7 5 6 4 1 2 1 16.4</td>
<td></td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>11 3 4 3 7 8 5 3 4 2 14.4</td>
<td></td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>10 6 5 11 3 5 4 2 2 2 18.8s</td>
<td></td>
</tr>
<tr>
<td>GGUBS-HSI</td>
<td>11 11 4 2 5 5 3 6 3 1 19.2s</td>
<td></td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>0 0 0 0 0 0 0 0 0 0 50 450.0s</td>
<td></td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>9 10 5 6 4 6 3 3 2 2 14.0</td>
<td></td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>15 9 2 6 5 1 4 1 6 1 35.2s</td>
<td></td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>15 6 10 5 2 3 2 1 3 3 34.4s</td>
<td></td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>13 8 8 6 0 1 6 2 3 3 28.4s</td>
<td></td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>13 7 6 2 5 6 3 2 3 3 20.0s</td>
<td></td>
</tr>
<tr>
<td>GGUBS-BMJ</td>
<td>15 8 1 6 5 1 5 2 5 2 32.0s</td>
<td></td>
</tr>
<tr>
<td>GGNPM</td>
<td>15 8 9 7 3 4 4 0 0 0 42.0s</td>
<td></td>
</tr>
<tr>
<td>GGNML</td>
<td>10 11 4 2 4 6 4 5 2 2 18.4s</td>
<td></td>
</tr>
</tbody>
</table>

s: Stands for significant series.

Critical chi-square value for grand chi-square value is 16.9.

Total numbers of chi-square values in each sample is 50.

The expected numbers of chi-square values in each interval is 5.

Bounding Chi-Square Values with equal probability for 10 intervals are 1: under 48.23, 2: under 52.49, 3: under 55.72
4: under 58.58, 5: under 61.33, 6: under 64.18, 7: under 67.32, 8: under 71.11, 9: under 76.62, 10 above 76.62.
Table 7

Distribution of 50 Chi-squares from Dist. of Sig. Lags

At Significance Level 0.01

<table>
<thead>
<tr>
<th>Generators/Combinations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRAND</td>
<td>21</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>69.2s</td>
</tr>
<tr>
<td>RANDU</td>
<td>21</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>66.8s</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>12</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>23.2s</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>19</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>60.0s</td>
</tr>
<tr>
<td>URAND</td>
<td>13</td>
<td>7</td>
<td>12</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>43.2s</td>
</tr>
<tr>
<td>R. Shore</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>22.0s</td>
</tr>
<tr>
<td>SGUBS</td>
<td>12</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>22.8s</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>38.0s</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>16</td>
<td>7</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>42.0s</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>21</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>69.2s</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>20</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>60.4s</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>23</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>84.0s</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>18</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>50.0s</td>
</tr>
<tr>
<td>GGUBS-BMI</td>
<td>22</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>79.2s</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>19</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>56.4s</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>19</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>59.2s</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>37.6s</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>13</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>41.2s</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>18</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>53.2s</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>11</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>30.8s</td>
</tr>
<tr>
<td>GGUBS-HSI</td>
<td>14</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>30.4s</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>40</td>
<td>450.0s</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>16</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>43.2s</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>16</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>40.4s</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>14</td>
<td>14</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>49.2s</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>13</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>38.4s</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>11</td>
<td>5</td>
<td>11</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>25.6s</td>
</tr>
<tr>
<td>GGUBS-BMJ</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>38.4s</td>
</tr>
<tr>
<td>GGNEM</td>
<td>16</td>
<td>8</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>43.2s</td>
</tr>
<tr>
<td>GGNML</td>
<td>14</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>32.8s</td>
</tr>
</tbody>
</table>

s: Stands for significant series.

Critical chi-square value for grand chi-square value is 21.7.

Total numbers of chi-square values in each sample is 50.
The expected numbers of chi-square values in each interval is 5.
Bounding Chi-Square Values with equal probability for 10 intervals are 1: under 48.23, 2: under 52.49, 3: under 55.72
4: under 58.58, 5: under 61.33, 6: under 64.18, 7: under 67.32,
8: under 71.11, 9: under 76.62, 10 above 76.62.
Chapter VI
SERIAL CORRELIATION TEST BY SPECTRAL ANALYSIS

6.1 INTRODUCTION

Although autocovariance or autocorrelation analysis can be used to test a sequence generated from a computer for white noise, in general, the autocovariance or autocorrelation analysis is used to estimate the inter relationship of the elements of a time series over discrete lags. In order to measure the behavior of a time series using autocovariance analysis, it could be necessary to examine a very large number of autocovariance lags. In addition, except for purely random sequences, or white noise series, the sample autocovariance estimators are themselves highly autocorrelated. Therefore, theoretically, autocovariance invalidates the use of standard statistical tests which assume that observations are independent of one another (Sovereign, Nolan, and Mandel, 1971).

Alternatively, the techniques of Fourier analysis can be used to transform the autocovariance function of a time ser-

---

8If a series has sample size N, then the number of possible autocorrelation lags is (N-1).

- 122 -
ies from the time domain into the frequency domain which is a continuous function termed a spectrum. The spectrum has a one to one correspondence to the autocovariance for a sequence, and it has the advantage of representing all possible autocorrelations over discrete lags without computing a very large number of autocovariance lags. In addition, the adjacent sample spectral estimators are essentially uncorrelated; consequently, the theoretical study of sampling variability and bias is much simpler in the case of the estimates of the spectrum than in the case of the estimates of the autocovariance function. Thus, the spectrum over the frequencies can be used as an alternative approach to determine the sequence generated from computers for white noise. It is more appropriate for detecting departures from white noise caused by periodic effect. In this chapter we will analyze the spectrum of a series to evaluate its randomness.

6.2 **SPECTRAL ANALYSIS**

Spectral analysis (Fourier transform of the autocovariance function?) is a technique for analyzing autocorrelated variables more fully than the usual statistical techniques using such measures as sample means and variances. It is

---

9See Appendix-D.

10A stationary time series is a time series with time invariance of the mean value, the variance, and the autocovari-
the decomposition of a stationary time series\textsuperscript{10} into periodical\textsuperscript{11} (cyclical) components, and it determines which of the component cycles exert most influence on the original series. In other words, it decomposes the variance of a stationary time series into the variances over the harmonic frequencies\textsuperscript{12} termed the power spectrum, the variance spectrum, or the spectrum. It is like the ANOVA which is regarded as the decomposition of the variance into factor components; so the variance of a time series is equal to the sum of the variances of each of its frequencies. The analysis can tell which harmonic frequency variance contributes the most to the total variance of the original stationary time series. Therefore, spectral analysis focuses on the power spectrum which provides a measure of the relative contribution of cycles in a band of frequencies to the total variance of the data. Also, the power spectrum can be plotted against frequencies as a measure of the contribution of periods with particular frequencies to the total variance of the series.

\textsuperscript{11}Period is a measurement of the time required for a full cycle.

\textsuperscript{12}Frequency is the reciprocal of period, i.e., a measurement of number of cycles per unit of time. The fundamental frequency has its corresponding period equal to the length of the whole period. Harmonic frequency is a multiplier of the fundamental frequency, such Kth harmonic frequency as \( F_k = K/(D*\text{N}) \), where D is sampling time, and N is number of sampling data points.
6.2.1 Power Spectrum

The power spectrum for a discrete white noise time series \( X(T) \) is constant. It is equal to a multiple of the variance of a time series. If the variance of a time series is VAR, and the variance of the power spectrum at each harmonic frequency equal to \( P(F) \), then \( P(F) = D \times VAR \) (where \( D \) is the sampling time). Hence, we can analyze the power spectrum or variance over each harmonic frequency and determine whether or not the power spectrum is constant. If the power spectrum is constant, then we can conclude that the time series from which the power spectrum was computed is a white noise series.

In practice, we can obtain only a limited number of sample of time series of finite length. The reduction of data will, therefore, yield no more than an estimate of the autocovariance function and of the power spectrum. Thus, we need to compute the sample spectrum first and then use it to estimate the power spectrum. Like the theoretical power spectrum and the autocovariance function, the sample spectrum and the sample autocovariance function are Fourier transforms of each subseries.\(^{13}\) Hence, we can obtain the sample spectrum by a Fourier transform of the sample autocovariance function.

\(^{13}\)See Appendix-E.
However, the sample spectrum estimator at each harmonic frequency is an unbiased but inconsistent\textsuperscript{14} estimator. Its behavior is so erratic as to render it useless as an estimator when it is needed to analyze a stochastic random process. Hence, in order to obtain a useful spectrum estimate, one must first smooth the sample spectrum estimator. One device which can be used to produce a spectrum estimator which has smaller and smoother variance than the sample spectral estimator is Bartlett's smoothed spectral estimator.\textsuperscript{15} It involves the averaging or smoothing of the sample spectrum over sub-series at each harmonic frequency (Blackman and Tukey, 1958; Jenkins, 1961; Bendat and Piersol, 1966).

6.2.2 **Smoothed Sample Spectral Estimators**

Bartlett's smoothed sample spectral estimators are the Fourier transform of the averaged or weighted sample autocovariance functions over sub-series. Since the weights called Bartlett's window lag in Bartlett's original estimates vanish for $K > M$ (where $K$ is a discrete lag, and $M$ is the length of sub-series and is called the truncation point). The number of sample autocovariance functions we need to

\textsuperscript{14}See Appendix-E.

\textsuperscript{15}See Appendix-F.
compute at different lags is only up to the length of the subseries. Furthermore, the truncation point, \( M \), is proportional to the variance\(^{16}\). In other words, the variance of smoothed sample spectral estimators at each harmonic frequency can be reduced by making the truncation point of the window lag small. Therefore, the smaller the truncation point (or the wider the bandwidth\(^{17}\)) we use in the sample spectrum, the smaller the variance of the smoothed sample spectral estimators that will be obtained.

In general, the more we want to smooth the estimators, the larger the bias we will introduce. Consequently, a compromise must be effected between achieving small variability (termed high stability) and small bias (termed high fidelity). However, for a white noise series, the smoothed sample spectral estimators are unbiased and consistent. So the variance could be reduced without introducing bias by using a small truncation point, \( M \), or a wide bandwidth of spectral window.

The sample spectral estimator at each harmonic frequency, \( C(F) \), is approximately distributed in proportion to a chi-square distribution with two degrees of freedom. That is

\[^{16}\text{See Appendix-G.}\]
\[^{17}\text{See Appendix-G.}\]
2C(F)/P(F) is approximately distributed as chi-square with two degrees of freedom (d.f.) \(^{18}\) (where C(F) is the sample spectral estimator at frequency F, and P(F) is the power spectrum at frequency F). The corresponding result for the smoothed spectral estimator at each harmonic frequency, SC(F), is approximately distributed as chi-square with v = d.f. That is the variable \(v \cdot SC(F)/P(F)\) is approximately distributed as chi-square with v d.f. \(^{19}\) and \(v = (A \cdot N / M)^2\) (where A is a multiplier and N is the total length of the original series).

The covariance between smoothed spectral estimators is proportional to the amount of overlap of the spectral windows centered at frequencies F1 and F2. Hence, if the spectral windows overlap only slightly, the covariance will be very small. And the covariance between two estimators at a sufficiently wide frequency spacing is approximately zero. Hence, independent confidence intervals may be constructed at this frequency separation (Jenkins, 1961).

Since the random variable \(v \cdot SC(F)/P(F)\) is distributed according to a chi-square distribution with v = d.f., it follows that the probability that the variable \(v \cdot SC(F)/P(F)\)
will be within the interval \( CS(\text{ALPHA}/2) \) and \( CS(1-(\text{ALPHA}/2)) \) is \((1-\text{ALPHA})\). That is,

\[
P(\text{CS}(\text{ALPHA}/2) < v*SC(F)/P(\text{F}) < \text{CS}(1-(\text{ALPHA}/2))) = 1-\text{ALPHA}
\]

where \( P(\text{CS}<\text{CS}(\text{ALPHA}/2)) = \text{ALPHA}/2 \), and \( \text{ALPHA} \) is the significance level. So, the interval between \( P(\text{F})*\text{CS}(\text{ALPHA}/2)/v \) and \( P(\text{F})*\text{CS}(1-(\text{ALPHA}/2))/v \) is a 100 \((1-\text{ALPHA})\) percent theoretical confidence interval for \( \text{SC}(F) \). This confidence interval will be used as one of the criteria to drive a hypothesis test for detecting whether or not the series generated on the computer is a white noise series. That is, if the sequence is white noise, then the theoretical white noise confidence interval at each frequency \( F \) will contain the sample spectrum computed from the generated series.

The other three approaches in the spectral analysis are sample mean of sample spectra, sample variance of sample spectra, and normalized sample variance of sample spectra. For a set of \( L \) statistically independent smoothed sample spectral estimators: \( SC(\text{Fi}) \), for \( i=1,2,3,...,L \), the sample mean of \( SC(\text{Fi}) \) approaches to the normal distribution with mean equal to the true mean, \( \mu \), and variance equal to the true variance, \( \text{VARsc} \), divided by the sample size, \( L \), if \( L \) is large.\(^{20}\) That is, if \( L \) is large, \( \text{MSC} = N(\mu, \text{VARsc}/L) \). Thus the theoretical white noise confidence intervals can be es-

\(^{20}\)An extension of the central limit theorem.
tablished for the sample mean of a generated series for the hypothesis test. That is, if the generated sequence is a white noise, then the theoretical white noise confidence interval will contain the sample mean of generated series. Hence we can use it as one of the criteria for the randomness test in spectral analysis.

The sample variance for SC(Fi), SVARsc approaches proportionality to the chi-square distribution with \((L-1)\) = d.f., if \(L\) of the smoothed sample spectral estimators is large.\(^{21}\) That is, for large \(L\), \((L-1)\) * SVARsc/VARsc approaches the chi-square with \((L-1)\) = d.f. In addition, the normalized sample variance also becomes proportional to chi-square with \((L-1)\) = d.f.. That is

\[(L-1) \times \frac{NSVARsc}{NVARsc} - CS\text{ with } (L-1) \text{ degrees of freedom.}\]

where \(NSVARsc=SVARsc/Msc\)**2, and \(NVARsc=VARsc/P**2. Again, theoretical white noise confidence intervals can be set up for the sample variance and normalized variance for the test of the white noise hypothesis.

\(^{21}\)This is true because a chi-square distribution approaches the normal distribution if the number of degrees of freedom becomes large.
6.3 **PROCEDURES OF SPECTRAL ANALYSIS**

There are four approaches for testing the series generated from computers for white noise by using spectral analysis. They are the number of significant sample spectra, sample mean of sample spectra, sample variance of sample spectra, and normalized sample variance of sample spectra. These four approaches will be discussed in the following four subsections.

6.3.1 **Number of Significant Sample Spectra**

Three procedures are used here for testing the series for white noise.

1. Test each set for white noise.

The power spectrum, $P(F)$, for a discrete white noise series $X(T)$ is constant; it is equal to a multiplier of the variance of a stationary time series. And, the smoothed sample spectrum estimator approaches proportionality to the chi-square distribution with $v=\text{d.f.}$. The expectation of this procedure is as follows:

a) Determine the sample size, $N$. 

b) Generate series $x$ from the computer; treat the series as a discrete time series with sampling time, $D = 1$. Subtract the sample mean from each term such that the mean of the series, $M_x = 0$.

c) Select window—Bartlett's window—for smoothing.

$$W(K) = \begin{cases} 
1 - (|K|/M), & |K| < M \\
0 & |K| \geq M 
\end{cases}$$

d) Determine the truncation point, $M$. In order to use the central limit theorem, we let $M = 64$ in this study.

e) Compute sample autocovariance functions

$$COV(K) = \frac{1}{M} \sum (x(T) \times x(T+K)),$$

$K = 0, 1, 2, \ldots, (M-1) = 63$ and $T = 1, 2, \ldots, K$.

f) Compute smoothed autocovariance functions.

$$SCOV(K) = W(K) \times COV(K), \text{ for } K = 0, 1, 2, \ldots, (M-1)$$

g) Take Fourier transforms of the smoothed autocovariance functions to obtain a smoothed sample spectrum.

$$SC(F_i) = (SCOV(0) + 2 \times \sum (SCOV(K) \times \text{COS}2\pi F_i(K))),$$

where $K = 1, 2, \ldots, (M-1)$, and $F_i = 0, 1/M, 2/M, \ldots, 1/2, \ldots, (M-1)/M$ for $i = 1, 2, \ldots, M$ respectively. After the Fourier transform of the autocorrelations of $M$ lags, there are $M$
frequencies. For the hypothesis test, we only need to analyze \((M/2)-1\)=31 frequencies; that is
\[f=1/M,2/M,...,31/M.\]
In this paper, we will analyze \(L=31\) frequencies.

h) Define the significance level \(\text{ALPHA}\), and let \(\text{ALPHA}=0.01, 0.05\) and 0.2.

i) Find theoretical white noise power spectrum at each frequency: \(P(F)=B*VAR=VAR\). In this study the \(VAR\) is \(1/12\) and 1 for the uniform distribution and the unit normal distribution, respectively.

j) Set up theoretical confidence interval for sample spectra \(SC(Fi)\) at each harmonic frequency
\[
P(F)\cdot CS(\text{ALPHA}/2)/v, \quad P(F)\cdot CS(1-\text{ALPHA}/2)/v.
\]
For Bartlett's window \(v=3*N/M=56\) if \(N=1200\) and \(M=64.\)

\[22\text{Since the highest frequency is such that its period corresponds to two sample points. In this case, the highest frequency is 1/2. Moreover, SC(Fi) is an even function, so the variance at the frequencies (1/2+1/M)...(M-1)/M is the same as the variance at 1/M...(1/2 -1/M). In addition, the degrees of freedom of frequencies 0 and 1/2 are one half of degrees of freedom of frequencies 1/M,2/M,...31/M. For simplicity, we do not include frequencies 0 and 1/2 for analyses in this study.}\]
k) Count the total number of observed significant intervals, SIGSPF, (i.e., the hypothetical confidence interval not include the sample spectra SC(Fi) over L=31 lags.

l) At significance level ALPHA, compute the total number of expected significant intervals, EXP, and then compute upper bound, UCINO, over L=31 harmonic frequencies, respectively.

\[ \text{EXP} = \text{ALPHA} \times L \]

\[ \text{UCINO} = \text{EXP} + Z \times (L \times \text{ALPHA} \times (1 - \text{ALPHA}))^{0.5} \]

Where \( Z \) is the number of standard deviations corresponding to the confidence interval.

The hypothesis to be tested and its alternative are as follows:

\( H_0: \) The series \( x \) is a white noise series at the significance level ALPHA if the observed number of significant intervals, SIGSPF, is less than or equal to upper bound of the number of significance intervals, UCINO.

\( H_A: \) The series \( x \) is not a white noise series at the significance level ALPHA if the observed number of significant intervals, SIGSPF, is greater than the upper bound of the number of significant confidence intervals, UCINO.
2. Test each sample of KL=50 sets for white noise by means of the chi-square test.

As with the distribution of significant COR(K), the distribution of significant SC(Fi) values, Fi=1/M,2/M,...,31/M for i=1,2,...,L=31, respectively, should be uniform over the KL sets generated. The chi-square goodness of fit test can be used to test whether or not the distribution of significant SC(Fi) is uniform over the KL=50 sets. If it is, then the process of KL=50 sets is white noise; if not it is correlated.

The procedure is expected as follows:

a) For each harmonic frequency, count the number of significant SC(Fi) values, SCFREQ(Fi), (Fi=1/M,2/M,...,31/M for i=1,2,...,L, respectively) over the KL sets by means of the procedure stated above.

b) For each frequency compute the expected value of significant SC(F), EXP, over KL sets at the significant level ALPHA.

\[ \text{EXP} = KL \times \text{ALPHA} \]

c) Compute the chi-square value, CHISCF

\[ \text{CHISCF} = \text{SUM} ((\text{EXP} - \text{SCFREQ(Fi)})^2/\text{EXP}) \]
d) Find the chi-square critical value for the hypothesis test at \((L-1)\) degrees of freedom and the significance level \(\alpha\). In this study, we found \(C_R = 50.9, 43.8, 36.3\) with \(d.f. = (L-1) = 30\) for \(\alpha = 0.01, 0.05, 0.2\), respectively.

The hypothesis to be tested and its alternative are as follows:

\(H_0\): The process of all \(kL = 50\) sets is white noise at significance level \(\alpha\) if \(\text{CHISCF} \leq \text{CRT}\).

\(H_A\): The process of all \(kL = 50\) sets is not white noise at significance level \(\alpha\) if \(\text{CHISCF} > \text{CRT}\).

3. Test the whole series generated of \(ISET = 50\) samples for whit noise by number of significant samples counted above and by means of the CS goodness of fit test.

At significance level \(\alpha\), we use two ways to judge whether or not the whole generated series is whit noise. First, If the whole series of \(ISET = 50\) samples is random, then the number of significant samples counted over \(ISET = 50\) samples is less than or equal to the upper bound for number of signifi-
can samples

\((\text{ISET} \cdot \text{ALPHA} \cdot Z \cdot (\text{ISET} \cdot \text{ALPHA} \cdot (1 - \text{ALPHA}))^{**.5})\). Second, if the whole generated series of ISET samples is a random sequence, the ISET=50 CHISCF values computed from each sample have a chi-square distribution with \((L-1)=30\) degrees of freedom. Thus we can use a chi-square goodness of fit test to check whether or not the ISET=50 CHISCF values have a chi-square distribution with \(L-1=30\) degrees of freedom. The expectation of this procedure is as follows:

a) Use the second procedure derived above to compute the CHISCF statistic for each sample. There are ISET=50 CHISCF statistics which have a chi-square distribution with \((L-1)=30\) degrees of freedom if the whole series is random.

b) Using the conventional expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, \(K_2\), and let \(K_2=10\).

c) Use equal probabilities \(1/10\) to compute the expected frequencies \(E_i\) for the \(i\)th category. \(E_i=5\) for \(\text{ISET}=50\) and \(P_i=0.1\).
d) Under the null hypothesis and with equal probabilities 1/10 per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, $O_i$, for the $i$th category in all $K_2$ categories.

f) Compute the CS statistic

$$CS = \sum \left( \frac{(E_i - O_i)^2}{E_i} \right), \text{ for } i=1,2,\ldots,10.$$ 

g) Set up significance levels $\alpha=0.01$, 0.05 and 0.2 for the hypothesis tests.

h) Find the critical values, CRT, for the hypothesis tests. At the significance levels $\alpha=0.01$, 0.05 and 0.2 with $K_2-1=9$ degrees of freedom, CRT=21.7, 16.9 and 12.2, respectively.

The hypothesis to be tested and its alternative are as follows:

$H_0$: The whole series of $ISET=50$ samples is random at the significance level $\alpha$ if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to CRT.

$H_A$: The series of $ISET=50$ samples is not random at the significance level $\alpha$ if number of significant samples is greater than its up-
per round, or if the computed CS value is greater than CR1.

6.3.2 **Mean of Sample Spectra**

There are three procedures for testing the series for white noise with this approach.

1. Test each set for white noise.

   By the central limit theorem, if the sample size is large, and if the population has finite variance and finite mean, then the sample mean approaches the normal distribution with its mean equal to the population mean and its variance equal to the population variance divided by the sample size. In this study, the number of independent sample spectra to be analyzed is \( L = 31 \). Using the central limit theorem, we assume that the sample mean of spectra approaches the normal distribution with the mean \( \mu = D \cdot \text{VAR} \), and with the variance \( \text{VAR} = \frac{1}{L} \). In addition, for white noise series, the true mean \( \mu = D \cdot \text{VAR} \) and the true variance \( \sigma^2 = \left( \mu^2 \right) \frac{1}{L} \), where \( L = \frac{(2*M)}{3} \) for Bartlett's window.

The procedure is expected as follows:
a) Steps a to i are the same as those in the procedure of "Test each set for white noise" in section 6.3.1 in this chapter

b) Compute the sample mean for L independent smoothed sample spectra

\[ M_{sc} = \frac{\text{SUM} (SC(i))}{L}, \text{ for } i=1,2,\ldots,L. \]
where \( L = (M/2)-1 \) and \( L=31 \) in this study.

c) Compute the normalized sample mean, \( Z_{Msc} \)

\[ Z_{Msc} = \frac{(M_{sc}-P)}{(\text{VAR}_{sc}/L)^{\times0.05}}. \]

d) Compute the theoretical confidence interval for the sample mean \( M_{sc} \) at significance level \( \alpha \).

\[ P - Z*\left(\frac{2*\text{M}*(P**2)}{3*N*L}\right)^{\times0.5}, \]
\[ P + Z*\left(\frac{2*\text{M}*(P**2)}{3*N*L}\right)^{\times0.5} \]
Since \( L > 30 \), \( M_{sc} \sim N(P, \text{VAR}_{sc}/L) \)

The hypothesis and its alternative are:

\( H_0: \) The series \( x \) is white noise at the significance level \( \alpha \) if \( M_{sc} \) is in the hypothetical confidence intervals.

---

\(^{23}\)From the Appendix-G, we know that the variance of a smoothed power spectrum of a white noise series is \( \text{VAR}_{sc} = (P**2)*(I/N) \), and for Bartlett's window \( \text{VAR}_{sc} = (P**2)*2M/(3*N) \). So, \( \text{VAR}_{sc}/L = P**2*2M/(3*N*L) \).
HA: The series $x$ is not white noise at the significance level $\text{ALPHA}$ if $\text{Msc}$ is not in the hypothetical confidence intervals.

2. Test each sample of $\text{KL}=50$ sets for randomness.

If the sample of $\text{KL}$ sets is random, then the $\text{KL}=50$ $\text{Msc}$ values have a normal distribution. Thus we can use a chi-square goodness of fit test to test the normality of the $\text{KL}=50$ $\text{Msc}$ statistics.

The Procedure is expected as follows:

a) Use procedure stated above to compute the $\text{Msc}$ statistic for each set. There are $\text{KL}=50$ $\text{Msc}$ statistics which have an approximately normal distribution with mean=$P$ and variance=$\text{VARsc}$ if the sample is random. So the normalized $\text{Msc}$, $Z\text{Msc}$, has a unit normal distribution.

b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, $K1$, for $\text{KL}$ samples and let $K1=10$ for $\text{KL}=50$.

c) Use equal probabilities $1/10$ to compute the expected frequencies $Ei$ for the $i$th category. That is $Ei=5$ for $\text{KL}=50$ and $pi=0.1$. 
d) Under the null hypothesis and with equal probability $1/10$ per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, $G_i$, for the $i$th category in $K_1$ categories.

f) Compute the CS statistic

$$CS = \sum (E_i - O_i)^2 / E_i$$

for $i = 1, 2, \ldots, 10$

g) Set up significance levels $\alpha = 0.01, 0.05$ and $0.2$ for the hypothesis tests.

h) Find the critical value, $CRT$, for the hypothesis tests. At significance levels $\alpha = 0.01, 0.05$ and $0.2$ with $K_1 - 1 = 9$ degrees of freedom, $CRT = 21.7, 16.9$ and $12.2$, respectively.

The hypothesis to be tested and its alternative are as follows:

$H_0$: The sample of $KL=50$ sets is a white noise sequence at significance level $\alpha$ if the computed CS value is less than or equal to $CRT$.

$H_A$: The sample of $KL=50$ sets is not a white noise sequence at significance level $\alpha$ if the computed CS value is greater than $CRT$.

3. Test the whole generated series with $ISET=50$ for randomness by number of significant samples.
counted above and by means of the CS goodness of fit test.

At significance level $\text{ALPHA}$, we use two ways to judge whether or not the whole generated series is whit noise. First, if the whole series of $\text{ISET}=50$ samples is random, then the number of significant samples counted over $\text{ISET}=50$ samples is less than or equal to the upper bound for number of significant samples

$$(\text{ISET} \ast \text{ALPHA} + z \ast (\text{ISET} \ast \text{ALPHA} \ast (1 - \text{ALPHA})) \ast 0.5).$$

Second, if the whole generated series of $\text{ISET}=50$ samples is random sequence, the $\text{ISET}=50$ CS values have a chi-square distribution with $K1-1=9$ degrees of freedom. Thus we can use a chi-square goodness of fit test to check whether or not the $\text{ISET}=50$ CS values have a chi-square distribution with $K1-1=9$ degrees of freedom. The procedure for accomplishing this test is as follows:

a) Use the second procedure mentioned above to compute the CS statistic for each sample. There are $\text{ISET}=50$ CS statistics which have a chi-square distribution with $K1-1=9$ degrees of freedom if the whole series is random.
b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, $K_2$, and let $K_2=10$ for ISET=50 in this study.

c) Use equal probabilities $1/10$ to compute the expected frequencies $E_i$ for the $i$th category. $E_i=5$ for ISET=50 and $P_i=0.1$.

d) Under the null hypothesis with equal probability $1/10$ per class, compute the upper and the lower bounds for each class.

e) Count the observed frequencies, $O_i$, for the $i$th category in $K_2$ categories.

f) Compute the CS statistic

$$CS = \text{SUM} ((E_i - O_i)^2/E_i), \text{ for all } i=1,2,\ldots,10.$$  

g) Set up significance levels $\alpha = 0.01$, 0.05 and 0.2 for the hypothesis tests.

h) Find the critical value, $CRT$, for the hypothesis test. At significance levels $\alpha = 0.01$, 0.05 and 0.2 with $k_2-1=9$ degrees of freedom, $CRT=21.7$, 16.9 and 12.2, respectively.
The hypothesis to be tested and its alternative are as follows:

\( H_0: \) The whole series of ISET=50 samples is random at the significance level \( \alpha \) if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to CRT.

\( H_A: \) The whole series of ISET=50 samples is not random at the significance level \( \alpha \) if number of significant samples is greater than its upper bound, or if the computed CS value is greater than CRT.

6.3.3 Sample Variance of Sample Spectra

Three procedures are used here for testing the series for white noise using this approach.

1. Test each set for white noise.

The smoothed sample spectrum estimator has an approximately chi-square distribution with \( v=d.f. \). A chi-square distribution approaches a normal distribution if the number of degrees of freedom becomes large. In this study, the number of degrees of freedom of the smoothed sample spectral estimator is 56; thus the sample spectral estimator ap-
proaches the normal distribution (Bendat and Pier-sol, 1966, p.324). Hence the sample variance is approximately proportional to the chi-square distribution.

The Procedure is as follows:

a) Steps a through i are the same as those in the procedure of "Test each set for white noise" in section 6.3.1 in this chapter.

b) Compute the sample variance: \( SVARsc \)

\[
SVARsc = \frac{\text{SUM}(SC(Fi)-Msc)**2}{(L-1)}, \quad \text{for } i=1,2,...,L.
\]

And \( SVARsc \) is approximately chi-square distribution with \( L-1=30 \) degrees of freedom.

c) Set up theoretical confidence intervals for the sample variance: \( SVARsc \)

\[
VARsc*CS(\alpha/2)/(L-1), \quad VARsc*CS(1-\alpha/2)/(L-1)
\]

The hypothesis and its alternative are:

\( H0: \) The series \( x \) is white noise at the significance level \( \alpha \) if \( SVARsc \) is in the hypothetical confidence intervals.

\( HA: \) The series \( x \) is not white noise at the significance level \( \alpha \) if \( SVARsc \) is outside the theoretical confidence intervals.
2. Test each sample of KL=50 sets for white noise by means of a CS goodness of fit test.

If the sample of KL=50 sets is random, then the KL has fifty statistics \(((L-1) \times SVAR_{sc}/VAR_{sc})\) computed from each set have a chi-square with \(L-1=30\) degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the sample is random.

The procedure is expected as follows:

a) Use procedure derived above to compute the \(SVAR_{sc}\) statistic for each set. There are KL=50 statistics \(((L-1) \times SVAR_{sc}/VAR_{sc})\) which have the chi-square distribution with \(L-1=30\) degrees of freedom if the sample is random.

b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, \(K1\), for KL samples and let \(K1=10\) for KL=50.

c) Use equal probabilities \(1/10\) to compute the expected frequencies \(E_i\) for the \(i\)th category. Find \(E_i=5\) for KL=50 and \(P_i=0.1\).
d) Under the null hypothesis and with equal probabilities 1/10 per class, compute the upper and lower bounds for each class.

e) Count the observed frequencies $O_i$ for the $i$th category in $K_1$ categories.

f) Compute the CS statistic

$$CS = \sum \frac{(E_i - O_i)^2}{E_i}, \text{ for } i=1,2,\ldots,K_1=10.$$ 

g) Set up the significance levels $\alpha = 0.01, 0.05$ and 0.2 for the hypothesis tests.

h) Find the critical values, CRT, for the hypothesis tests. At significance levels $\alpha = 0.01, 0.05$ and 0.2 with $K_1-1=9$ degrees of freedom,

$$CRT = 21.7, 16.9, 12.2,$$ 

respectively.

The hypothesis to be tested and its alternative are as follows:

$H_0$: The sample of $K_1=50$ sets is a white noise sequence at the significance level $\alpha$ if the computed CS value is less than or equal to CRT.

$H_A$: The sample of $K_1=50$ sets is not a white noise sequence at the significance level $\alpha$ if the computed CS value is greater than CRT.

3. Test the whole series of $ISET=50$ samples for white noise by number of significant samples counted above and by means of the CS goodness of fit test.
At significance level $\alpha$, we use two ways to judge whether or not the whole generated series is white noise. First, if the whole series of $ISET=50$ samples is random, then the number of significant samples counted over $ISET=50$ samples is less than or equal to the upper bound for number of significant samples ($ISET*\alpha + z(ISET*\alpha*(1-\alpha))^{*0.5}$). Second, if the whole series of $ISET=50$ samples is random, then the $ISET=50$ values of CS statistic computed from each sample have a chi-square with $K1-1=9$ degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the series is random. The procedure of this test is as follows:

a) Use the second procedure mentioned above to compute the CS statistic for each sample. There are $ISET=50$ CS statistics which have a chi-square distribution with $K1-1=9$ degrees of freedom if the whole generated series is random.

b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, $K2$, and let $K2=10$ for $ISET=50$ in this study.
c) Use equal probabilities 1/10 to compute the expected frequencies $E_i$ for the $i$th category. $E_i=5$ for ISET=50 and $P_i=0.1$.

d) Under the null hypothesis and with equal probabilities 1/10 per class, compute the upper and lower bounds for each class.

e) Count the observed frequencies, $O_i$ for the $i$th category in $K_2$ categories.

f) Compute the CS statistic

$$CS=\text{SUM}((E_i-C_i)^2/E_i), \text{ for } i=1,2,...,K_2=10.$$  

g) Set up the significance levels $\text{ALPHA}=0.01, 0.05$ and 0.2 for the hypothesis tests.

h) Find the critical value, CRT, for hypothesis tests. At significance levels $\text{ALPHA}=0.01, 0.05$ and 0.2 with $K_2-1=9$ degrees of freedom, CRT=21.7, 16.5 and 12.2, respectively.

The hypothesis to be tested and its alternative are as follows:

$H_0$: The whole series of ISET=50 samples is random at the significance level ALPHA if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to CRT.
HA: The whole series of ISET=50 samples is not white noise at the significance level ALPNA if number of significant samples is greater than its upper bound, or if the computed CS value is greater than CBT.

6.3.4 Normalized Sample Variance of Sample Spectra

Three procedures are used for testing the series for white noise using this approach.

1. Test each set for white noise.

If the number of degrees of freedom of the smoothed sample spectrum is large, the normalized sample variance of sample spectra, like the sample variance, is approximately proportional to the chi-square distribution (Bendat and Piersol, 1966, p. 324).

The Procedure for accomplishing the test is as follows:

a) Steps a through i are the same as those in the procedure of "Test each set for white noise" in section 6.3.1 in this chapter.
b) Compute the normalized sample variance: \( nsvar_{sc} \)

\( nsvar_{sc} = \frac{svar_{sc}}{m_{sc}^2} \).

c) Compute the normalized theoretical variance: \( nvar_{sc} \)

\( nvar_{sc} = \frac{var_{sc}}{p^2} \).

d) Set up the confidence interval for the sample normalized variance

\[ \begin{align*}
nsvar_{sc} \cdot cs \left( \frac{\alpha}{2} \right) / (L-1), \\
nsvar_{sc} \cdot cs \left( 1 - \frac{\alpha}{2} \right) / (L-1).
\end{align*} \]

The hypothesis and its alternative are:

\( H_0 \): The series \( x \) is white noise at the significance level \( \alpha \) if the \( nsvar_{sc} \) is in the hypothetical confidence intervals.

\( H_A \): The series \( x \) is not white noise at the significance \( \alpha \) if the \( nsvar_{sc} \) is outside the confidence intervals.

2. Test each sample of \( KL=50 \) sets for white noise by means of a CS goodness of fit test.

If the sample of \( KL=50 \) sets is random, then the \( KL \) has fifty statistics \( \left( (L-1) \cdot nsvar_{sc} / nvar_{sc} \right) \) computed from each set have a chi-square with \( L-1=30 \) degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the sample is random.
The procedure is as follows:

a) Use procedure derived above to compute the \( \text{NSVARsc} \) statistic for each set. If the sample is random, there are \( \text{KL}=50 \) statistics 
\[ (L-1) \times \text{NSVARsc}/\text{NVARsc} \]
which have the chi-square distribution with \( K-1=30 \) degrees of freedom.

b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, \( K1 \), for \( KL \) samples and let \( K1=10 \) for \( KL=50 \).

c) Use equal probabilities \( 1/10 \) to compute the expected frequencies \( Ei \) for the ith category. Find \( Ei=5 \) for \( KL=50 \) and \( Pi=0.1 \).

d) Under the null hypothesis and equal probabilities \( 1/10 \) per class, compute the upper and lower bounds for each class.

e) Count the observed frequencies \( Oi \) for the ith category in \( K1 \) categories.

f) Compute the CS statistic
\[
\text{CS} = \sum ( (Ei-Oi)^2/Ei ) \text{, for } i=1,2,...,K1=10.
\]
g) Set up the significance levels $\text{ALPHA}=0.01$, 0.05 and 0.2 for the hypothesis tests.

h) Find the critical value, CRT, for hypothesis tests. At significance levels $\text{ALPHA}=0.01$, 0.05 and 0.2 with $K1-1=9$ degrees of freedom, CRT=21.7, 16.9 and 12.2, respectively.

The hypothesis to be tested and its alternative are as follows:

$H0$: The sample of $KL=50$ sets is a white noise series at the significance level $\text{ALPHA}$ if the computed CS value is less than or equal to CRT.

$HA$: The sample of $KL=50$ sets is not a white noise series at the significance level $\text{ALPHA}$ if the computed CS value is greater than CRT.

3. Test the whole series of $\text{ISET}=50$ samples for white noise by number of significant samples counted above and by means of the CS goodness of fit test.

At significance level $\text{ALPHA}$, we use two ways to judge whether or not the whole generated series is whit noise. First, If the whole series of $\text{ISET}=50$ samples is random, then the number of significant samples counted over $\text{ISET}=50$ samples is less than or equal to the upper bound for number of signifi-
cant samples

(ISET*ALPHA+Z*(ISET*ALPHA*(1-ALPHA)))**.5). Second, if the whole series of ISET=50 samples is random, then the ISET=50 values of the CS statistic computed from each sample have a chi-square distribution with $K1-1=9$ degrees of freedom. Thus we can use the CS goodness of fit test to check whether or not the series is random. The expectation of this procedure is as follows:

a) Use the second procedure derived above to compute the CS statistic for each sample. If the whole generated series is random, there are ISET=50 CS statistics which have a chi-square distribution with $K1-1=9$ degrees of freedom.

b) Using the conventional minimum expected value of 5 and Cochran's moderate number of classes between 10 to 25, establish mutually exclusive and exhaustive categories, $K2$, and let $K2=10$ for ISET=50 in this study.

c) Use equal probabilities 1/10 to compute the expected frequencies $Ei$ for the $ith$ category. $Ei=5$ for ISET=50 and $Pi=0.1$.

d) Under the null hypothesis and equal probabilities 1/10 per class, compute the upper and lower bounds for each class.
e) Count the observed frequencies, $C_i$ for the $i$th category in $K_2$ categories.

f) Compute the CS statistic

$$CS = \sum((E_i - O_i)^2/E_i), \text{ for } i=1,2,\ldots,K_2=10.$$  

g) Set up significance levels $\text{ALPHA}=0.01, 0.05$ and $0.2$ for the hypothesis tests.

h) Find the critical values, $\text{CRT}$, for the hypothesis tests. At the significance levels $\text{ALPHA}=0.01$, $0.05$ and $0.2$ with $K_2-1=9$ degrees of freedom, $\text{CRT}=21.7$, $16.9$ and $12.2$, respectively.

The hypothesis to be tested and its alternative are as follows:

$H_0$: The whole series of $\text{ISET}=50$ samples is a white noise at the significance level $\text{ALPHA}$ if number of significant samples is less than or equal to its upper bound, or if the computed CS value is less than or equal to $\text{CRT}$.

$H_A$: The whole series of $\text{ISET}=50$ samples is not a white noise at the significance level $\text{ALPHA}$ if number of significant samples is greater than its upper bound, or if the computed CS value is greater than $\text{CRT}$. 
6.4 **EMPIRICAL TEST RESULTS, DISCUSSIONS AND EVALUATIONS**

Results of tests of the series for white noise by spectral analysis are outlined in Tables 8 and 9. Distributions of sample spectra and sample means of spectra are listed in Table 8. Sample variances and normalized sample variances of spectra are exhibited in Table 9. All information in the tables were derived from 50 independent samples with 50 sets per sample and 1200 numbers per set.

Like the autocorrelation tests, we computed 31 sample spectra over 50 sets for each sample, and we obtained a distribution of sample spectra for that sample. If the series is white noise, the distribution of significant spectra has a uniform distribution with 50*ALPHA expected numbers of significant spectra. Columns 7, 8, and 9 in Table 8 show the numbers of significant samples for all 50 independent samples at levels 0.2, 0.05, and 0.01, respectively. They were obtained by comparing the observed chi-square values with 30 degrees of freedom derived from the distribution of significant spectra with their corresponding critical values. Columns 10, 11, and 12, however, summarize the grand chi-square values with 9 degrees of freedom for the whole series at significance levels 0.2, 0.05 and 0.01, respectively. They were derived from 50 chi-square values with 30 degrees of freedom.
In addition, we computed the sample means, variances, and normalized variances of the 31 sample spectra. For each sample, we have 50 sample means, 50 sample variances and 50 normalized sample variances. We then used chi-square goodness of fit test to test whether or not the distribution of sample means have a normal distribution and the distribution of sample variances and normalized sample variances have a proportional chi-square distribution with 30 degrees of freedom.

Similarly the other types of statistical tests in the previous chapter, we counted the numbers of significant samples and computed the grand chi-square values for the whole generated series. The numbers of significant samples and the grand chi-square values of sample means are arranged in columns 3, 4, 5, and 6 in Table 8. However, the numbers of significant samples and the grand chi-square values of sample variances and normalized sample variances are listed in Table 9.

Tables 8 and 9, each shows 12 possible rejections. In Table 8, six of them result from testing means of sample spectra (Columns 3, 4, 5, and 6); the other six result from testing distributions of sample spectra (Columns 7, 8, 9, 10, 11, and 12). In Table 9, 6 of 12 possible significant results come from investigating variances of sample spectra.
(Columns 3, 4, 5, and 6) and the other 6 are from normalized variances of sample spectra (Columns 7, 8, 9 and 10).

For each set of six possible rejections, three are indicated by the numbers of significant samples, and the other three are pointed out by the grand chi-square values at levels 0.2, 0.05 and 0.01. Column 2 exhibits actual total numbers of rejections from the distributions of sample spectra and the means of sample spectra and from variances and normalized variances of sample spectra in Table 8 and 9, respectively.

In the test using sample spectra, (in Table 8), at levels 0.2 and 0.05, no generator or combination passes the test either from studying numbers of significant samples or grand chi-square values (Columns 3, 4, and 6). Even at level 0.01, only a few generators or combinations pass the test. RANDU, R. Shore, and M & M with BMI pass the test both from the views of numbers of significant samples and grand chi-square values. ADRAND, URAND, R. Shore with BMI, M & M with HSI, and L & L with BMJ pass the test from the view of the numbers of significant samples (Column 5 and 6 in Table 8).

The results of the tests by distribution of sample spectra are very similar as the ones from distribution of lags. It is to be expected since spectral analysis and autocorre-
Table 3

Means and Distribution of Sample Spectra

| Generators/ | Means | Distribution of Spectra |
| Combinations | [\# of Sig. S Chi-Square-Values] | [\# of Sig. S Chi-Square-Values] |
|              |       | Ref] 0.2 | 0.05 | 0.01 | Ref] 0.2 | 0.05 | 0.01 |
| (1)          |       | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| ADRAND       | 7     | 20s  | 7s  | 1   | 30.0** | 4  | 0  | 0  | 29.2s | 13.2 | 75.6s |
| RANDU        | 7     | 15s  | 7s  | 0   | 17.6** | 1  | 0  | 0  | 33.6s | 20.4s | 90.9s |
| L & L        | 8     | 19s  | 9s  | 3s  | 32.8***| 5  | 0  | 0  | 14.8s | 15.6 | 100.4s|
| M & H        | 8     | 24s  | 7s  | 3s  | 32.8***| 2  | 0  | 0  | 24.8s | 12.4 | 57.2s |
| URAND        | 7     | 20s  | 8s  | 2   | 22.0***| 2  | 2  | 2  | 23.6s | 12.8 | 53.2s |
| R. Shore     | 7     | 19s  | 8s  | 2   | 21.2** | 3  | 0  | 1  | 31.6s | 17.6s | 66.4s |
| GGUBS        | 9     | 19s  | 10s | 4s  | 32.4***| 5  | 0  | 0  | 15.6s | 20.4s | 66.8s |
| ADRAND-BMI   | 8     | 25s  | 12s | 4s  | 42.8***| 2  | 2  | 2  | 36.8s | 13.6 | 56.8s |
| RANDU-BMI    | 9     | 25s  | 13s | 4s  | 45.6***| 1  | 5  | 3s | 46.0s | 10.4 | 46.4s |
| L & L-BMI    | 9     | 25s  | 13s | 4s  | 45.6***| 2  | 0  | 0  | 39.2s | 20.0s | 67.2s |
| M & M-BMI    | 6     | 20s  | 7s  | 2   | 17.6** | 2  | 0  | 1  | 40.4s | 14.0 | 28.8s |
| URAND-BMI    | 9     | 22s  | 12s | 6s  | 27.6***| 1  | 2  | 1  | 61.2s | 23.2s | 37.2s |
| R. Shore-BMI | 9     | 19s  | 9s  | 1   | 34.0***| 1  | 2  | 3s | 53.2s | 16.3s | 52.8s |
| GGUBS-BMI    | 9     | 23s  | 14s | 5s  | 40.0***| 3  | 0  | 0  | 31.6s | 17.2s | 63.6s |
| ADRAND-HSI   | 9     | 26s  | 15s | 8s  | 52.8***| 3  | 1  | 5s | 14.4s | 14.0 | 51.2s |
| RANDU-HSI    | 8     | 23s  | 7s  | 3s  | 39.6***| 0  | 1  | 0  | 39.6s | 8.4  | 53.2s |
| L & L-HSI    | 8     | 30s  | 10s | 7s  | 68.0***| 5  | 2  | 1  | 27.2s | 9.2  | 41.2s |
| M & M-HSI    | 7     | 22s  | 11s | 2   | 36.4***| 2  | 3  | 2  | 23.6s | 6.8  | 35.6s |
| URAND-HSI    | 8     | 25s  | 15s | 3s  | 43.6***| 1  | 3  | 2  | 56.4s | 13.6 | 40.0s |
| R. Shore-HSI | 8     | 25s  | 10s | 3s  | 35.6***| 1  | 2  | 2  | 37.2s | 13.6 | 54.0s |
| GGUBS-HSI    | 8     | 29s  | 17s | 7s  | 66.0***| 6  | 2  | 0  | 28.8s | 8.4  | 46.0s |
| ADRAND-BMJ   | 12    | 32s  | 17s | 9s  | 63.9***| 6s  | 4s  | 50s | 23.2s | 39.2s | 374.8s|
| RANDU-BMJ    | 8     | 21s  | 11s | 3s  | 27.2**  | 1  | 0  | 1  | 33.4s | 15.6 | 50.4s |
| L & L-BMJ    | 8     | 25s  | 9s  | 1   | 41.6***| 1  | 4  | 1  | 28.9s | 18.4s | 54.9s |
| M & M-BMJ    | 3     | 20s  | 11s | 4s  | 35.6***| 2  | 0  | 1  | 45.2s | 11.2 | 41.6s |
| URAND-BMJ    | 8     | 24s  | 13s | 5s  | 52.4***| 2  | 1  | 2  | 30.8s | 14.0 | 45.2s |
| R. Shore-BMJ | 9     | 23s  | 12s | 7s  | 28.0***| 2  | 3  | 3s | 48.4s | 10.8 | 32.8s |
| GGUBS-BMJ    | 7     | 26s  | 10s | 2   | 42.0***| 1  | 2  | 1  | 32.8s | 16.0 | 54.8s |
| GGUBS        | 9     | 27s  | 14s | 4s  | 71.6***| 1  | 1  | 1  | 42.8s | 26.0s | 65.2s |
| SNFL         | 8     | 30s  | 14s | 6s  | 87.6***| 4  | 2  | 0  | 31.2s | 9.2  | 46.0s |

*: Significant at level 0.2
**: Significant at levels 0.2 and 0.05.
**:**: Significant at levels 0.2, 0.05 and 0.01.
$: Significant at level ALPHA.
@: Total number of rejections is 12.
BMI: Stands for Box-Muller Inverse.
BMJ: Stands for Box-Muller Rejection.
HSI: Stands for Hastings Inverse.
Upper bound for the number of significant samples at levels 0.2, 0.05 and 0.01 are 12, 5, and 2, respectively.
Critical chi-square values for the whole generated series are 12.2, 16.9 and 21.7 for significance levels 0.2, 0.05 and 0.01, respectively.
lation analysis are simply Fourier transforms each other. Like the results of the distribution of lags, ADRAND with BMJ yields very poor results as indicated by the numbers of significant samples and the grand chi-square values. Judged by the number of significant samples, with the exception of ADRAND with BMJ, all of generators and combinations pass the test at levels 0.2 and 0.05; a few, RANDE with BMI, R. Shore with BMI, and ADRAND with SL, do not pass the test at level 0.01. Judged by the grand chi-square values, no generators or combinations pass the test at level 0.2 or 0.01. The reasons are the same as discussed relative to the distribution of autocorrelation lags. They are rejected because they generated too many white noise samples (Column 7, 8, and 9 in Table 8) and too evenly distributed numbers (Table 10 and 11) in the series.

Briefly, no generator or combination performs well in the tests of means and distributions of sample spectra; especially ADRAND with BMJ which has 12 rejections out of the 12 possible. Even the best combination, M & M with BMI, in this test, has 6 significant values out of 12 possible. The reason probably is that spectral analysis is a difficult test to pass. Hence few generators or combinations can pass the test.
Concerning the tests of variances and normalized variances (in Table 9) again, ADRAND with BMJ is the worst performer. It does not pass any of the tests. The next worse ones in order, are M & M with BMI, URAND with HSI, URAND with BMJ, ADRAND with BMI, L & L with BMI, GGUES, GGUBS with BMI, GGUBS with HSI, and GGNML, they have 7, 7, 7, 6, 6, 6, 6, 6, 6 rejections, respectively.

The best generators or combinations by these tests are L & L and R. Shore with BMJ. They pass all of the tests at all significance levels investigated here. Next to these two, RANDU with HSI, RANDU with BMJ, have 1 significant test out of 12 possible ones. These are followed by ADRAND, RANDU, L & L with BMJ, GGUBS with BMJ and M & M with BMJ, with 2 rejections each. The remaining generators or combinations have 3 to 5 rejections each (Column 2).
<table>
<thead>
<tr>
<th>Generators/</th>
<th>Variance</th>
<th>Normalized Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ofSig Samp</td>
<td>Chi-S</td>
</tr>
<tr>
<td></td>
<td>Rej</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADRAND</td>
<td>2</td>
<td>14s</td>
</tr>
<tr>
<td>RANDU</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>3</td>
<td>17s</td>
</tr>
<tr>
<td>URAND</td>
<td>5</td>
<td>14s</td>
</tr>
<tr>
<td>R. Shore</td>
<td>3</td>
<td>17s</td>
</tr>
<tr>
<td>GGUBS</td>
<td>6</td>
<td>14s</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>6</td>
<td>14s</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>6</td>
<td>19s</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>7</td>
<td>17s</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>3</td>
<td>17s</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>5</td>
<td>19s</td>
</tr>
<tr>
<td>GGUBS-BMI</td>
<td>6</td>
<td>18s</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>5</td>
<td>13s</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>5</td>
<td>13s</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>4</td>
<td>13s</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>7</td>
<td>19s</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>5</td>
<td>17s</td>
</tr>
<tr>
<td>GGUBS-HSI</td>
<td>6</td>
<td>15s</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>12</td>
<td>50s</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>1</td>
<td>15s</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>7</td>
<td>24s</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>GGUBS-BMJ</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>GGNPM</td>
<td>4</td>
<td>19s</td>
</tr>
<tr>
<td>GGNML</td>
<td>6</td>
<td>17s</td>
</tr>
</tbody>
</table>

*: Significant at level 0.2
**: Significant at levels 0.2 and 0.05.
**:**: Significant at levels 0.2, 0.05 and 0.01.
s: Significant at level ALPHA.
@: Total number of rejections is 12.
BM: Stands for Box-Müller Inverse.
BM: Stands for Box-Muller Rejection.
HSI: Stands for Hastings Inverse.
Upper bound for the number of significant samples at levels 0.2, 0.05 and 0.01 are 12, 5, and 2, respectively. Critical chi-square values for the whole generated series are 12.2, 16.9 and 21.7 for significance levels 0.2, 0.05 and 0.01, respectively.
Table 10

Distribution of 50 Chi-squares from Dist. of Sig. Spectra

At Level 0.2

<table>
<thead>
<tr>
<th>Generators/Combinations</th>
<th># of Chi-Square Values in ith Category</th>
<th>Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRAND</td>
<td>10 10 4 4 4 4 5 3 1 2 2</td>
<td>29.2s</td>
</tr>
<tr>
<td>RANDU</td>
<td>7 7 10 9 4 4 7 1 0 1 0</td>
<td>33.6s</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>6 9 4 4 4 4 3 3 2 3 2</td>
<td>14.8s</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>4 8 5 6 10 2 2 2 0</td>
<td>24.8s</td>
</tr>
<tr>
<td>URAND</td>
<td>9 6 4 5 7 3 2 2 0</td>
<td>23.6s</td>
</tr>
<tr>
<td>R. Shore</td>
<td>10 4 7 6 5 0 2 3 0</td>
<td>31.6s</td>
</tr>
<tr>
<td>GGBES</td>
<td>6 9 4 4 3 4 3 4 4 1</td>
<td>15.6s</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>10 14 4 7 4 4 5 3 1 1</td>
<td>36.8s</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>12 4 12 7 2 0 1 1 0</td>
<td>46.0s</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>11 7 11 3 4 3 2 3 2 0</td>
<td>39.2s</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>15 5 12 4 4 3 3 2 2 0</td>
<td>40.4s</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>19 7 8 3 3 2 2 0 1 0</td>
<td>61.2s</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>19 8 4 5 2 4 5 4 2 0 1</td>
<td>53.2s</td>
</tr>
<tr>
<td>GGUBS-BMI</td>
<td>13 9 11 3 3 3 3 3 3 0</td>
<td>31.6s</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>6 9 8 5 5 3 5 3 3 0</td>
<td>14.4s</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>10 12 6 1 6 1 3 0 0</td>
<td>39.6s</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>13 11 4 2 3 5 5 2 2 3</td>
<td>27.2s</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>11 8 6 9 3 3 2 1 4 0</td>
<td>23.6s</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>18 9 7 2 1 2 8 2 1 0</td>
<td>56.4s</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>14 9 10 2 5 2 3 4 3 1 0</td>
<td>37.2s</td>
</tr>
<tr>
<td>GGUBS-HSI</td>
<td>14 10 4 2 3 5 5 1 3 3</td>
<td>28.8s</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>0 0 1 0 1 1 2 2 6 37</td>
<td>233.2s</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>10 14 9 6 3 3 1 1 0</td>
<td>38.4s</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>11 9 9 7 6 4 1 2 0 1</td>
<td>28.0s</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>17 9 6 6 3 2 1 2 0</td>
<td>45.2s</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>13 7 11 4 3 2 5 3 1 1</td>
<td>30.8s</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>16 13 3 5 3 3 2 3 1 1</td>
<td>48.4s</td>
</tr>
<tr>
<td>GGUBS-BMJ</td>
<td>11 11 8 8 5 4 1 0 1</td>
<td>32.8s</td>
</tr>
<tr>
<td>GGNPM</td>
<td>14 13 7 4 4 2 3 2 1 0</td>
<td>42.8s</td>
</tr>
<tr>
<td>GGNML</td>
<td>15 8 5 2 4 7 2 1 3 3</td>
<td>31.2s</td>
</tr>
</tbody>
</table>

Total numbers of chi-square values in each sample is 50.
The expected numbers of chi-square values in each category is 5.Bounding Chi-Square Values with equal probability for 10 categories are 1: under 48.23, 2: under 52.49, 3: under 55.72 4: under 58.58, 5: under 61.33, 6: under 64.18, 7: under 67.32, 8: under 71.11, 9: under 76.62, 10 above 76.62.
Table 11

Distribution of 50 Chi-squares from Dist. of Sig. Spectra

At Level 0.01

<table>
<thead>
<tr>
<th>Generators/Combinations</th>
<th># of Chi-Square Values in ith Category</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRAND</td>
<td>21 0 9 0 6 0 7 4 2 1</td>
<td>75.6s</td>
</tr>
<tr>
<td>RANDU</td>
<td>24 0 7 0 3 0 9 4 0 3</td>
<td>93.0s</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>25 0 5 0 7 0 6 3 2 2</td>
<td>100.4s</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>15 0 14 0 7 0 6 5 1 2</td>
<td>57.2s</td>
</tr>
<tr>
<td>URAND</td>
<td>18 0 9 0 6 0 4 5 3 5</td>
<td>53.2s</td>
</tr>
<tr>
<td>R. Shore</td>
<td>19 0 11 0 7 0 4 3 1 5</td>
<td>66.4s</td>
</tr>
<tr>
<td>GGUUS</td>
<td>23 0 7 0 7 0 6 4 2 1</td>
<td>86.8s</td>
</tr>
<tr>
<td>ADRAND-BMI</td>
<td>18 0 6 0 6 0 10 5 2 3</td>
<td>56.8s</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>14 0 10 0 10 0 3 3 2 8</td>
<td>46.4s</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>10 0 12 0 16 0 9 2 0 1</td>
<td>67.2s</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>11 0 7 0 8 0 8 8 4 4</td>
<td>28.8s</td>
</tr>
<tr>
<td>URAND-BMI</td>
<td>13 0 9 0 8 0 9 4 4 3</td>
<td>37.2s</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>14 0 15 0 3 0 4 4 4 6</td>
<td>52.8s</td>
</tr>
<tr>
<td>GGUUS-BMI</td>
<td>11 0 13 0 15 0 6 4 0 1</td>
<td>63.6s</td>
</tr>
<tr>
<td>ADRAND-HSI</td>
<td>14 0 10 0 9 0 8 0 1 8</td>
<td>51.2s</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>15 0 13 0 8 0 5 5 1 2</td>
<td>53.2s</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>12 0 11 0 9 0 3 8 1 6</td>
<td>41.2s</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>14 0 5 0 9 0 7 5 4 6</td>
<td>35.6s</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>13 0 9 0 12 0 4 4 4 5</td>
<td>40.0s</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>15 0 12 0 10 0 5 4 1 3</td>
<td>54.0s</td>
</tr>
<tr>
<td>GGUUS-HSI</td>
<td>11 0 13 0 10 0 3 8 1 4</td>
<td>46.0s</td>
</tr>
<tr>
<td>ADRAND-BMJ</td>
<td>0 0 1 0 0 0 1 1 4 6</td>
<td>374.0s</td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>16 0 10 0 10 0 8 2 2 2</td>
<td>56.4s</td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>10 0 11 0 16 0 5 3 2 3</td>
<td>54.8s</td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>14 0 8 0 6 0 11 4 3 4</td>
<td>41.6s</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>16 0 9 0 5 0 8 3 4 5</td>
<td>45.2s</td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>12 0 9 0 7 0 8 6 2 6</td>
<td>32.8s</td>
</tr>
<tr>
<td>GGUUS-BMJ</td>
<td>10 0 11 0 16 0 5 3 2 3</td>
<td>54.8s</td>
</tr>
<tr>
<td>GGNUM</td>
<td>16 0 11 0 12 0 7 1 2 1</td>
<td>65.2s</td>
</tr>
<tr>
<td>GGML</td>
<td>11 0 13 0 10 0 4 8 1 3</td>
<td>46.0s</td>
</tr>
</tbody>
</table>

Total numbers of chi-square values in each sample is 50.
The expected numbers of chi-square values in each category is 5.
Bounding Chi-Square Values with equal probability for 10
categories are: 1: under 48.23, 2: under 52.49, 3: under 55.72
4: under 58.53, 5: under 61.33, 6: under 64.18, 7: under 67.32,
8: under 71.11, 9: under 76.62, 10 above 76.62.
Chapter VII
SUMMARY AND CONCLUSION

In order to evaluate the quality of unit uniform random number generators and unit normal random number transformations, we have selected five sets of statistical tests to test the generators and combinations. A summary of the results of all of these tests are presented in Tables 12, 13 and 14.

Table 12 lists the total numbers of rejections for generators and combinations by statistical tests and by significance levels. That is, for each generator and combination, under each significance level, we list numbers of rejections on tests of goodness of fit, runs up and down, runs above and below the median, autocorrelation analysis, means and distribution of sample spectra, and variance and normalized variance of sample spectra. Under each significance level each statistical test except for goodness of fit has 4 possible rejections; the goodness of fit has two. Hence, at each level of ALPHA there are 22 possible rejections. Also for each generator or combination, there are 66 possible numbers of rejections considering all levels.

- 166 -
<table>
<thead>
<tr>
<th>Generators</th>
<th>Significance Levels</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>F</td>
</tr>
<tr>
<td>AERAND</td>
<td>0 2 0 1 2 0 5</td>
<td>22 14</td>
</tr>
<tr>
<td>RANDU</td>
<td>0 1 0 1 1 0 3</td>
<td>20 7.5</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>0 1 0 1 3 0 5</td>
<td>18 3.5</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>0 1 0 1 3 0 5</td>
<td>18 3.5</td>
</tr>
<tr>
<td>SAND</td>
<td>0 1 0 1 2 0 5</td>
<td>18 3.5</td>
</tr>
<tr>
<td>R. Shore</td>
<td>0 1 0 1 1 0 3</td>
<td>21 10.5</td>
</tr>
<tr>
<td>GUES</td>
<td>0 1 0 1 1 0 3</td>
<td>21 10.5</td>
</tr>
<tr>
<td>AERAND-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>RANDU-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>L &amp; L-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>M &amp; M-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>SAND-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>GUES-BMI</td>
<td>0 1 0 1 3 0 5</td>
<td>23 18.5</td>
</tr>
<tr>
<td>AERAND-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>RANDU-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>L &amp; L-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>M &amp; M-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>SAND-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>R. Shore-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>GUES-NS1</td>
<td>0 1 0 1 3 0 5</td>
<td>27 27</td>
</tr>
<tr>
<td>AERAND-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>RANDU-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>L &amp; L-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>M &amp; M-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>SAND-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>R. Shore-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>GUES-MS1</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>AERAND-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>RANDU-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>L &amp; L-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>M &amp; M-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>SAND-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>R. Shore-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>GUES-VV</td>
<td>0 1 0 1 3 0 5</td>
<td>24 21.5</td>
</tr>
<tr>
<td>DD: Stands for goodness of fit tests.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RU: Stands for runs up and down tests.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT: Stands for autocorrelation tests.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM: Stands for distribution of spectra and mean of sample spectra in spectral analysis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VW: Stands for variance and normalized variance of sample spectra in spectral analysis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RU, RA, AT, MS and VW have 4 possible rejections each at level ALPHA. PD, however, has 2 at each level ALPHA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TI: Stands for total number of rejections from DD, RU, RA, AT, MS and VW. The number of possible rejections at each level of ALPHA is 22.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From level 0.01, except for ADRAND with BMJ, each generator or combination passes two-thirds of the tests (Column 8). The best generators or combinations are RANDU, and R. Shore with BMJ. Each has 3 significant tests out of the 22 that are possible. ADRAND with BMJ is the worst one, it has 16 rejections. This can clearly be considered as a bad combination. The rest of the generators or combinations have 4 to 7 significant tests out of 22. Hence, at level 0.01, with the exception of ADRAND with BMJ, the quality of generators or combinations is not substantially different.

At significance level 0.05, just a few generators or combinations do not pass two-thirds of the tests. ADRAND with HS1, L & L with BMI, R. Shore with BMJ have 8, 9 and 9 rejections out of 22, respectively. Again, ADRAND with BMJ is the worst one, it has 17 significant tests out of 22. The best combination is RANDU with BMJ, 2 rejections. This is followed by RANDU with BMI and RANDU with HS1, each has 3 rejections. M & M with BMJ has 4 significant tests. The rest of them have numbers of rejections between 5 and 7 (Column 15).

At level 0.2, however, no generators or combinations pass two-thirds of the tests. The best ones are L & L, ADRAND with BMI, and RANDU with BMJ; each has 8 significant tests out of 22. These are followed by RANDU with HS1 and GGNPM,
9 rejections each. RANDU, GGUBS, URAND with BMI, M & M with HSI, GGUBS with BMJ have 10 significant tests each. Again, ADRAND with BMI, the worst one, only passes 5 out of 22. L & L with BMI, M & M and M & M with BMJ pass 7, 8 and 8, respectively. The rest of the generators or combinations have total numbers of rejections between 11 to 13 (Column 24).

From the total number of rejections from all tests and all levels, it appears that RANDU with BMI is the best one; it is rejected by 14 out of 66 tests. Next is RANDU with HSI, 17 rejections. Two, 3.5 rank, are L & L and ADRAND with BMI, 18 significant tests each. GGNEM has 19 rejections. RANDU, RANDU with BMI, M & M with BMJ, and GGUBS with BMJ, are rank 7.5 with 20 rejections each. With 50 rejections out of 66, ADRAND with BMI is the worst combination. Next to this are L & L with BMI and R. Shore with BMJ with 28 significant tests each. This is followed by ADRAND with HSI which has 27 rejections. The rest of them have 21 to 25 significant tests (Column 26).

The information in Table 13 is the same as that in Table 12 but for different arrangement of data.

The numbers of observed rejections for generators and combinations are listed by significance levels and by statistical tests. From this table, it is clear that, among
Table 13
Number of Rejections for All Levels by Statistical Tests

<table>
<thead>
<tr>
<th>Generators/</th>
<th>Distri.</th>
<th>Runs Up</th>
<th>Runs All</th>
<th>Autocor</th>
<th>Spectrum</th>
<th>Variance</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combinations</td>
<td>Fit &amp; Down</td>
<td>E NL.</td>
<td>Analysis</td>
<td>E Mean</td>
<td>E Var</td>
<td>Tl</td>
<td>E</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>--------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AERAND</td>
<td>0 0 0 0</td>
<td>2 2 3 7</td>
<td>0 1 2 3</td>
<td>1 1 1 3</td>
<td>2 2 3 7</td>
<td>0 0 2</td>
<td>2 2 14</td>
</tr>
<tr>
<td>RANDU</td>
<td>0 1 1 2</td>
<td>1 2 2 5</td>
<td>0 0 1 1</td>
<td>1 1 1 3</td>
<td>3 3 3 7</td>
<td>0 0 2</td>
<td>2 2 10</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>0 0 0 1</td>
<td>1 2 3 6</td>
<td>0 1 0 1</td>
<td>1 1 1 3</td>
<td>3 2 3 8</td>
<td>0 0 0</td>
<td>0 1 19</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>0 1 1 2</td>
<td>0 2 2 1</td>
<td>2 1 3 6</td>
<td>1 1 1 2</td>
<td>4 3 2 8</td>
<td>0 0 0</td>
<td>0 1 24</td>
</tr>
<tr>
<td>URAND</td>
<td>0 0 0 1</td>
<td>1 4 6 0</td>
<td>0 1 1 2</td>
<td>0 2 0 1</td>
<td>3 2 3 7</td>
<td>1 2 2</td>
<td>5 2 2</td>
</tr>
<tr>
<td>R. Shore</td>
<td>0 0 0 1</td>
<td>1 2 2 5</td>
<td>0 1 1 1</td>
<td>1 1 1 2</td>
<td>4 3 2 8</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>GGSUS</td>
<td>0 0 0 0</td>
<td>2 3 5 0</td>
<td>0 0 0 0</td>
<td>1 1 1 2</td>
<td>3 2 3 8</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>AERAND-BMI</td>
<td>0 0 0 0</td>
<td>0 3 6</td>
<td>0 1 0 1</td>
<td>1 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>RANDU-MI</td>
<td>0 0 0 0</td>
<td>1 3 4</td>
<td>0 0 0 0</td>
<td>2 2 1 0</td>
<td>1 1 2</td>
<td>4 2 3</td>
<td>9 1 0</td>
</tr>
<tr>
<td>L &amp; L-MI</td>
<td>0 0 0 0</td>
<td>1 4 5</td>
<td>0 0 0 0</td>
<td>2 2 1 3</td>
<td>2 2 3</td>
<td>8 0 1</td>
<td>1 1</td>
</tr>
<tr>
<td>M &amp; M-MI</td>
<td>0 0 0 0</td>
<td>1 2 3</td>
<td>6 0 0 0</td>
<td>2 2 1</td>
<td>2 1 4</td>
<td>2 1 2</td>
<td>5 2 6</td>
</tr>
<tr>
<td>URAND-MI</td>
<td>0 0 0 0</td>
<td>2 2 6</td>
<td>0 0 1 1</td>
<td>1 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>R. Shore-BMI</td>
<td>0 0 0 0</td>
<td>1 4 5</td>
<td>0 0 1 1</td>
<td>1 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>GGSUS-BMI</td>
<td>0 0 0 0</td>
<td>1 2 2</td>
<td>2 2 0</td>
<td>0 1 1 1</td>
<td>1 1 1 3</td>
<td>3 3 3 9</td>
<td>0 0 0</td>
</tr>
<tr>
<td>AERAND-HSI</td>
<td>0 0 0 0</td>
<td>2 2 7</td>
<td>7 0 1 2</td>
<td>3 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>RANDU-HSI</td>
<td>0 0 0 0</td>
<td>1 1 2</td>
<td>4 0 0 0</td>
<td>1 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>L &amp; L-HSI</td>
<td>0 0 0 0</td>
<td>1 2 3</td>
<td>1 0 1 1</td>
<td>1 1 1 3</td>
<td>3 2 3 8</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>M &amp; M-HSI</td>
<td>0 0 0 0</td>
<td>1 2 3</td>
<td>7 0 1 2</td>
<td>3 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>URAND-HSI</td>
<td>0 0 0 0</td>
<td>1 1 3</td>
<td>5 0 1 1</td>
<td>1 1 1 3</td>
<td>3 2 3 7</td>
<td>0 0 0</td>
<td>0 1 12</td>
</tr>
<tr>
<td>R. Shore-HSI</td>
<td>0 0 0 0</td>
<td>2 2 6</td>
<td>0 0 1 1</td>
<td>1 1 2</td>
<td>5 3 2 8</td>
<td>1 1 3</td>
<td>5 2 5</td>
</tr>
<tr>
<td>GGSUS-HSI</td>
<td>0 0 0 0</td>
<td>1 2 3</td>
<td>6 0 0 0</td>
<td>1 1 2</td>
<td>2 3 8</td>
<td>1 1 3</td>
<td>5 2 5</td>
</tr>
<tr>
<td>AERAND-BMJ</td>
<td>0 0 0 0</td>
<td>2 2 6</td>
<td>2 3</td>
<td>3 8</td>
<td>4 4</td>
<td>4 2 4 4 4 4 7 4 4 12 5 0 30</td>
<td></td>
</tr>
<tr>
<td>RANDU-BMJ</td>
<td>0 0 0 0</td>
<td>0 1 1</td>
<td>0 1 1 0</td>
<td>2 3</td>
<td>3 2 3 8</td>
<td>0 0 1 1 14 1</td>
<td></td>
</tr>
<tr>
<td>L &amp; L-BMJ</td>
<td>0 0 0 0</td>
<td>1 2 3</td>
<td>6 0 0 0</td>
<td>1 1 1 3</td>
<td>3 2 3 8</td>
<td>0 0 1 1 14 1</td>
<td></td>
</tr>
<tr>
<td>M &amp; M-BMJ</td>
<td>0 0 0 0</td>
<td>1 1 0</td>
<td>1 4 5</td>
<td>1 1 1</td>
<td>1 1 3</td>
<td>3 2 3 8</td>
<td>0 0 2 2 20 7</td>
</tr>
<tr>
<td>URAND-BMJ</td>
<td>0 0 0 0</td>
<td>0 0 0</td>
<td>1 3 4 1</td>
<td>1 1 2</td>
<td>4 3 2 8</td>
<td>1 2 4 7 23 10.5</td>
<td></td>
</tr>
<tr>
<td>R. Shore-BMJ</td>
<td>0 0 0 0</td>
<td>1 1 0</td>
<td>4 8</td>
<td>7 2 4</td>
<td>1 1 1</td>
<td>3 2 3 8</td>
<td>0 0 2 2 28 23.5</td>
</tr>
<tr>
<td>GGSUS-BMJ</td>
<td>0 0 0 0</td>
<td>1 1 0</td>
<td>2 2 4</td>
<td>0 0 0</td>
<td>0 1 2</td>
<td>3 6</td>
<td>2 2 3 7</td>
</tr>
<tr>
<td>GGSUS-HMI</td>
<td>0 0 0 0</td>
<td>0 0 1 1 0 1</td>
<td>1 1 1 3</td>
<td>3 2 3 8</td>
<td>0 0 1 1 19 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGHSML</td>
<td>0 0 0 0</td>
<td>2 3 5 0</td>
<td>0 0 0 0</td>
<td>1 1 2 4</td>
<td>3 2 3 8</td>
<td>1 1 4 6 23 18.5</td>
<td></td>
</tr>
</tbody>
</table>

There are 4 possible numbers of rejections for each statistical test but for desired distribution fit (2 ones) at each significance level. TT: Stands for total number of rejections from 3 levels, 0.01, 0.05 and 0.2. Number of possible rejections at each statistical test is 12 but for desired distribution fit test (6 ones).
four randomness and independence tests, spectral analysis is
the most powerful test, especially the mean and distribution
of sample spectra. The total number of rejections of all
generators and combinations at all levels for this test is
245 (sum of 30 rows in column 21) out of 360 (30 * 4 possi-
ble rejections * 3 significance levels). The next most pow-
erful test is the runs up and down tests with 134 (Column 9)
out of 360 possible significant tests. This is followed by
the variance and normalized variance tests of the sample
spectra with 126 rejections (Column 2). Next to this is au-
tocorrelation tests with 116 rejections. The least powerful
test is runs above and below the median test, 60 rejections
(Column 13).

It is as expected; the spectral analysis is more powerful
than autocorrelation analysis since the former uses more in-
formation than the latter. In this study, we analyzed 64
autocorrelation lags among the 1200 numbers in the autocor-
relation analysis. In other words, the autocorrelation ana-
lysis uses the first 64 lags to judge whether or not the
1200 numbers correlated. The spectral analysis, after the
Fourier transformation on all possible 1199 lags to analyze
whether or not the 1200 numbers respects white noise. The
runs up and down test uses more information than does the
runs above and below the median test, also.
The other summary table is Table 14. This table lists the total number of rejections for each of the 7 generators itself and in combination with each of the 3 transformation algorithms and 2 IBM IMSL unit normal variate generators. From this table it is clear that, using the same transformation algorithm to transform 7 generators, good combinations are not necessarily obtained from good unit uniform generators. We also see that some generators are good in combination with one type of transformation but not necessarily good with other types of transformations.

From Column 2, we see that the quality of the 7 unit uniform generators is not substantially different. With respect to the total number of rejections, their range is 7, their mean is 21.57 and their standard deviation is 2.147. With 18 rejections, L & L is the best one. Next to this are RANDBU and R. Shore with 20 and 21 rejections, respectively. Next are followed by ADRAND and URAND; each has 22 rejections. The worst one is & M with 25 rejections. GGUBS has 23 rejections.

Using the BMI transformation, the quality of combinations is substantially more variable than that of the unit uniform generators. The range of the total number of rejections is 10. The mean and standard deviation are 22.57 and 3.259, respectively. The best combination does not come from the
<table>
<thead>
<tr>
<th>Combinations</th>
<th>BMI</th>
<th>HSI</th>
<th>BMJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADFAND</td>
<td>22</td>
<td>4.5</td>
<td>18</td>
</tr>
<tr>
<td>RANDU</td>
<td>20</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>L &amp; L</td>
<td>18</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>M &amp; M</td>
<td>25</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>URAND</td>
<td>22</td>
<td>4.5</td>
<td>22</td>
</tr>
<tr>
<td>R. Shore</td>
<td>21</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>GGUBS</td>
<td>23</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>GGNPM *</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGNML *</td>
<td>23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range 7 ** 10 10 36.0/14.0***
Mean 21.57 ** 22.57 23.286 25.29/21.17***
Standard De 2.147 ** 3.259 3.047 10.81/4.18***

* IMSL normal random variate generator.
** Does not include GGNPM and GGNML.
*** Does not include ADFAND combined with BMJ.
Total number of possible rejections is 66.
Range, mean and standard deviation of total numbers of rejections.
best generator, L & L, but from ADRAND, rank 4.5. The rank 1 generator, L & L, when combined with BMI yields the worst combination with 28 rejections. The ranks 3, 4.5, and 6 generators, R. Shore, URAND and GGUBS, respectively, when combined with BMI yield ranks 5, 3.5, 3.5 with their corresponding rejections 23, 22 and 22 each. The quality of combination with the worst generator, M & M, has 25 rejections.

When using the HSI transformation, ADRAND (the best combination in BMI) becomes the worst combination, 27 rejections. L & L and M & M, the worst and second worse combinations with BMI, move to ranks 3.5 and 2 in conjunction with HSI; they have 24 and 21 rejections, respectively. RANDU improves in quality in conjunction with HSI; with 17 rejections, it is the best generator in combination with the HSI transformation. URAND, R. Shore, and GGUBS yield worse combinations with HSI than with BMI; their corresponding numbers of significant tests are 25, 25 and 24, respectively.

Using the BMJ algorithm, the ranks of the various combinations are nearly the same as those with HSI. With 14 rejections, RANDU has a better quality than in BMJ and retains the first rank. ADRAND keeps the last rank, but with 50 rejections, it is substantial margin the poorest quality combination. The ranks of M & M and L & L with BMJ are close to those in HSI, but each has a slightly better quality.
Their corresponding numbers of rejections are 20 and 22. GGUPS and URAND also are better in combinations with BMJ than with HSI. Their numbers of rejections are reduced from 24 to 20 and from 25 to 23, respectively and their corresponding ranks move from 3.5 to 2.5 and 5.5 to 5. With 28 rejections R. Shore combined with BMJ is worse than that combined with BMI and HSI.

Among these three unit normal transformation algorithms, in general, it appears the BMJ is the best one. Excluding combined with ADRAND, the average number of rejections is 21.17. However, the quality of combination with BMI and HSI is not substantially different. Their means number of rejections are 22.57 and 23.286, respectively.

Finally, the findings of this study can be summarized as follows. First, among 21 combinations between 7 good generators and 3 good unit normal transformation algorithms and 2 IBM IMSL unit normal random variate generators, we have found that some combinations seem to be good but some are bad. RANDU with BMJ, RANDU with HSI, ADRAND with BMI, GGNEM, RANDU with BMI, M & M with BMJ, GGUBS with BMJ are good combinations. However, ADRAND with EMJ is the worst one. In addition, L & L with BMI, R. Shore with BMJ and ADRAND with HSI are bad. The rest of combinations fall somewhere between these extremes. An extensive study to improve
the quality of available combinations is necessary. Second, we found that spectral analysis is a tough test, few generators or combinations can pass the test. Third, in general, we found BMJ transformation algorithm appears better than BMI and HSI. Fourth, we found that the quality of combination of a "good" generator and a "good" transformation algorithm is not necessarily "good". Fifth, we found some "good" generators have "good" combination with one type of transformations but a "bad" quality with the other type of transformation. However, in most simulation algorithms, it is implicitly assumed that in any "good" generator can be used in conjunction with any "good" transformation; the resulting random variates will have good properties of randomness and desired distribution fit. The results of this research have shown that this assumption is false. Hence, the simulation practitioner must be cautious in the selection of both the generator and the transformation algorithm to be used and their interaction; especially he needs to select a stringent significance level. Further, a research would be valuable in identifying which unit uniform generators (in terms of additive, multiplicative or mixed congruential) is good for which kind of transformations (in terms of inverse, rejection and other type of distribution transformation) in the simulation area.
BIBLIOGRAPHY


Haldane, J. B. S., (1937), "The Exact Value of the Moments of the Distribution of X, Used as A Test of Goodness of Fit When Expectations are Small." Biometrika, Vol. 29, pp. 133-143.


Appendix A

UNIT UNIFORM RANDOM NUMBER GENERATORS

Additive Congruential Method

The additive congruential method tested here (ADRAND) (Burford, 1973, 1975) assumes K starting values, and it computes a sequence of numbers by the following congruence relation:

\[ R_i = (\text{SUM}(R_{i-j})) \pmod{b^{**m}}, \text{ for } j=1,2,\ldots,K. \]

Where \( b \) is a base of the number system (i.e. base=10, or base=2).

\( m \) is the number of digits desired for the pseudo random numbers generated.

\( R_i \) is the pseudo random numbers determined from \( R_{i-1} \) up to \( R_{i-k} \).

\( K \) is the number of previous numbers used to generate one pseudo random number.

The congruence relation of ADRAND is

\[ R_i = (\text{SUM}(R_{i-j})) \pmod{10^{**8}}, \text{ for } j=1,2,3,4. \]

Multiplicative Congruential Method
The multiplicative congruential method computes a sequence $R_i$ of nonnegative integers by means of the congruence relation:

$$R_i = A \cdot R_{i-1} \pmod{b^m}$$

where $A$ is a multiplier.

Two multiplicative congruential generators are tested here. They are the generators of Lewis and Learmonth (L & L) (Lewis and Miller, 1969, Kleijnen, 1974) and IBM's RANDU (IBM, SSP, 1968). The congruence relations of L & L and RANDU are as follows:

$$R_i = (7 \cdot 5) \cdot R_{i-1} \pmod{(2\cdot31) -1}$$ for L & L, and

$$R_i = 65539 \cdot R_{i-1} \pmod{2\cdot31}$$ for RANDU.

**Mixed Congruential Method**

The basic congruence relation for mixed congruential generators is

$$R_i = A \cdot R_{i-1} + C \pmod{b^m},$$ where $C$ is a constant.

Three mixed congruential generators investigated in this study are URAND (Forsythe, Maclcolm, and McIer, 1977), M & M (Marsaglia and Maclaren, 1970), and R. Shore. (Kinderman and Ramage, 1974). 'A' and 'C' in URAND are computed according to the theory and suggestions of D. E. Knuth (1969, p. 78 and p. 155).
The IMSL generator GGUBS (IBM IMSL, 1979) is also tested here. GGUBS is derived from the package LLRANDOM (i.e. L & L) with some cosmetic changes and other minor changes necessary to allow it to fit into the IMSL Library.
Appendix B

NORMAL TRANSFORMATION ALGORITHMS

Box-Muller Inversion Transformation

The well known Box - Muller transformation algorithm (Box and Muller, 1958) is a partial inversion method based on properties of pairs of normal random variables. Let Z1 and Z2 be independent random variables with the standard normal distribution which define the Cartesian co-ordinates of a random point. Then the polar co-ordinates are also independently distributed. The angle is uniformly distributed on (0,2PI), and the square of the radius has a chi-square distribution with two degrees of freedom (Atkison and Pearce, 1976). Both of these distributions may be inverted to give the pair

\[ Z1 = (-2 \times \ln R1) \times 0.5 \times \cos 2\pi R2, \text{ and} \]
\[ Z2 = (-2 \times \ln R1) \times 0.5 \times \sin 2\pi R2 \]

where \( R1 \) and \( R2 \) are two unit uniform random numbers.

Although many alternatives for transforming to normality do exist, no other method has the simplicity of this transformation. Also, this transformation is a theoretically ex-
act one if the variates to which it is applied are from a truly random unit uniform sequence. This method developed from the desire to have a way of generating normal deviates which would be reliable in the tails of the distribution. Mathematically this approach has the attractive advantage that the transformation for changing from uniform deviates to normal deviates is exact (Muller, 1959).

**Box-Muller Rejection Transformation Algorithm**

Marsaglia and Bray (1964) improved the Box-Muller inversion stated above by replacing the trigonometric functions with

\[ X = V_1 * ((-2*\ln W) / W) ^ * 0.5, \text{ and} \]
\[ Y = V_2 * ((-2*\ln W) / W) ^ * 0.5 \]

where \( W = V_1**0.5 + V_2**0.5 \leq 1 \); \( V_1 \) and \( V_2 \) are uniformly distributed on \((-1, 1)\). And \( W \) is uniformly distributed on \((0, 1)\) independently of \( V_1 \) and \( V_2 \) (Atkison and Pearce, 1976).

**Hastings Rational Approximation**

Rational approximations to transform a uniform deviate to a normal deviate have been suggested by several people. The best known version is due to Hastings (1955) (Muller, 1959, Schmidt and Taylor, 1970). Using this approach one obtains a reliable normal deviate \( X \) with mean \( MU \) and standard deviation \( SIGMA \) from a uniform deviate \( U = q \) as follows:
\[ X = MU + \left(\frac{U - 0.5}{|U - 0.5|}\right) \cdot \text{SIGMA} \cdot AA, \text{ where} \]

\[ AA = V - \left(\frac{A0 + A1 \cdot V + A2 \cdot V^2}{1 + D1 \cdot V + B2 \cdot V^2 + E3 \cdot V^3}\right) \]

and \[ V = (-2 \ln 0.5 \cdot (1 - 1 - 2U \cdot 0.5)^0.5) \]

A0 = 2.515517, A1 = 0.802853, A2 = 0.010328, B1 = 1.432786, B2 = 0.189269, B3 = 0.001308.

**IBM IMSL Library Subroutines**

GGNML and GGNPM (IBM IMSL, 1979) are two normal or Gaussian random deviate generators included in the IBM IMSL Library Subroutines. GGNML generates pseudo-random normal deviates by transforming unit uniform variates produced from the GGUBS routine to normal variates by means of the inverse approach. The problem is to find the inverse relationship

\[ U = F(X) \text{ given that} \]

\[ U = F(X) = \left(\frac{1}{2 \pi}\right)^{0.5} \int_{-\infty}^{X} \exp \left(-\frac{1}{2}t^2\right) dt \]

That is, the uniform random deviates, \( U \), generated by GGUBS are transformed to normal deviates \( X \) using IMSL inverse normal probability distribution function MDNRIS (IBM IMSL, 1979).

GGNPM generates pseudo-random normal deviates by the polar method. It generates pseudo-random normal deviates by transforming unit uniform variates generated from GGUBS to normal variates using Box-Muller rejection transformation algorithm.
Appendix C

DEFINITIONS OF STATISTICS INVESTIGATED FOR GOODNESS OF FIT

1. Standard Third Moment—\( b_1 \)

\( b_1 \) is the ratio of the third moment to the standard deviation.

\[ b_1 = N^{0.5} \frac{\sum (X_i - M_x)^3}{\left( \sum (X_i - M_x)^2 \right)^{1.5}}, \]

for \( i = 1, 2, \ldots, N \).

\( M_x \) is the sample mean of the observations \( X_i \).

2. Standard Fourth Moment—\( b_2 \)

\( b_2 \) is the ratio of the fourth moment to the square of the variance.

\[ b_2 = N \frac{\sum (X_i - M_x)^4}{\left( \sum (X_i - M_x)^2 \right)^2}, \]

for \( i = 1, 2, \ldots, N \).


KS is a function of the largest difference between the empirical distribution function of data and the distribution function hypothesized

\[ KS(+) = \max \left( \frac{i}{N} - F(Y_i) \right), \quad i = 1, 2, \ldots, N. \]

\[ KS(-) = \max \left( F(Y_i) - \frac{(i-1)}{N} \right), \quad i = 1, 2, \ldots, N. \]

\[ KS = \max (KS(+) , KS(-)) \]

\( Y_1 < Y_2 \ldots < Y_n \) denotes the ordered observations from a complete sample of size \( N \).
F is the hypothesized cumulative normal distribution function.

4. Cramer-Von Mises (1928, 1931)—CM

CM is the sum of the squared difference between the empirical distribution function of the data and the hypothesized distribution function.

$$CM = \text{SUM} (F(Y_i) - \left( \frac{(2i-1)}{(2*N)} \right)) \cdot \frac{2}{(2*N)} + \frac{1}{(12*N)}$$

for i=1,2,...,N.

5. Weighted Cramer-Von Mises (Anderson and Darling, 1954)—WCM

WCM is a modification of the CM in which the tails are given greater weight.

$$WCM = \left( \text{SUM} \left( \frac{2i-1}{(1/N)} \cdot \left( \ln F(Y_i) + \ln (1-F(Y_{n+1-i})) \right) \right) / N \right) - N$$

for i=1,2,...,N.

6. Modified KS (Durbin, 1961)—D

$$D = \text{Max} \left( \frac{i}{N} - \text{SUM} (G_j) \right), \ i=1,2,...,N.$$  

$$G_j = \left( \frac{N+2-j}{N} \right) \ast \left( C_j - C_{j-1} \right), \ j=1,2,...,N.$$  

$$0 < C_0 < C_1 < ... < C_n$$ obtained by ordering. 

$$C_1 = n, \ C_2 = n-2, \ C_3 = n-3, ..., \ C_n+1 = 1-n$$

$$u_i = F(Y_i), \ i=1,2,...,N.$$  

7. Chi-squared (Pearson, 1900)—CS

With K mutually exclusive and exhaustive classes, the CS statistic is the sum of the ratios of the square of the difference between the num-
bers of data points observed in the classes and the numbers expected in the classes under the null hypothesis to the numbers expected.

\[ CS = \text{SUM}(O_i - E_i)^2 / E_i \], for \( i = 1, 2, \ldots, K \).

\( E_i \) = Expected number per cell.

\( O_i \) = Number of observations per cell.

8. Shapiro and Wilk (1965) -- W test

\( W \) is the ratio of the square of the best, or approximately best linear unbiased estimator of the population standard deviation to the sample variance.

\[ W = \left( \text{SUM}(a_i Y_i) \right)^2 / \left( \text{SUM}(Y_i - My)^2 \right), \text{ where} \]

\( a_i = (a_1, a_2, \ldots, a_n) = \left( \text{m'V} / \left( \text{m'V * V * m'} \right)^{0.5} \right) \)

\( \text{m'} = (m_1, m_2, \ldots, m_n) \) denotes the vector of expected values of standard normal ordered statistics.

\( V = (v_{ij}) \) is the corresponding N*N inverse covariance matrix.

\( My \) is the sample mean of the ordered observations.


Up to a constant, DA is the ratio of Downton's linear unbiased estimator\(^1\) of the population standard deviation to the sample standard deviation.

\[ DA = T / \left( 2^{(N-2)} \text{SVAR} \right), \]

\(^1\text{Downton's (1966) original unbiased estimator of the normal distribution standard deviation is } 2 \ast T / N(N-1). \]
where $T_\text{VAR} = (\text{SUM} (i - ((N+1)/2)) * Y_i$ and

$T = (\text{SUM} (i - ((N+1)/2)) \times Y_i$ and

$\text{VAR} = (\text{SUM} (X_i - M_x) \times 2) / N$

10. Modified $W$ (Shapiro and Francia, 1972)--$W'$

$W' = (\text{SUM} (b_i Y_i) \times 2) / (\text{SUM} (Y_i - M_y) \times 2)$,

Where coefficients $b_i = (b_1, b_2, \ldots, b_n) = m'/(m' \times m)$ depend only on the expected values of the normal order statistics (Harter, 1961; Pearson and Hartley, 1972, Table 9).

11. Geary test (Geary, 1935)--$W_n$

$W_n$ is the ratio of the mean deviation to the standard deviation.

$W_n = \text{SUM} |X_i - M_x| / S$

where $S$ denotes the sample standard deviation.

12. Standardized range test (David, et al., 1954)--$u$

$u$ is the ratio of the range to the standard deviation.

$u = (N-1) \times 0.5 \times (Y_n - Y_1) / (((\text{SUM} (Y_i - M_y) \times 2) \times 0.5))$

13. Ratio of $b_1$ and $b_2$ jointly (Bowman and Shenton, 1975)--$K$

$K = X(b_1) + X(b_2)$

where $X(b_1)$ and $X(b_2)$ are standardized normal equivalents to the sample skewness, $b_1$, and kurtosis, $b_2$, statistics. $K$ is distributed as chi-squared with 2 degrees of freedom.
Appendix D

MATHEMATICAL RELATIONS: POWER SPECTRUM AND AUTOCOVARIANCE

Suppose \( X(T) \) is a stationary stochastic process with mean zero, variance \( \text{VAR} \), and length \( N \). The mathematical relations between the power spectrum and the autocovariance function of \( X(T) \) are as follows:

1. Continuous time domain case.

The autocovariance function with lag \( U \) may be defined as

\[
R(U) = \lim_{N \to \infty} \frac{1}{N} \int_{-N/2}^{N/2} X(T) \ast X(T+U) \, dT,
\]

\(-N/2 < T < N/2, -N < U < N\) \hspace{1cm} (D-1)

This may be reduced to the form:

\[
R(U) = \int_{-\infty}^{\infty} P(F) \ast \exp\left(j2\pi F U\right) dF, \quad \text{where} \hspace{1cm} (D-2)
\]

\[
P(F) = \lim_{N \to \infty} \frac{1}{N} \left( \int_{-N/2}^{N/2} X(T) \ast \exp\left(-j2\pi F T\right) dT\right)^2 \hspace{1cm} (D-3)
\]

D-3 may be reduced to

\[
P(F) = \int_{-\infty}^{\infty} R(U) \ast \exp\left(-j2\pi F U\right) dU, \hspace{1cm} (D-4)
\]

where \( \exp\left(j2\pi F U\right) = \cos2\pi F U + j\sin2\pi F U \). \hspace{1cm} (D-5)

\( j = (-1)^{0.5} \),
F denotes frequencies from positive infinity to negative infinity.

\( P(F) \) denotes a function of frequencies, and it is called the periodogram power spectrum.

D-2 shows that the autocovariance function \( R(U) \) is a Fourier transform of the power spectrum, \( P(F) \). And A-4 shows that the power spectrum, \( P(F) \), is an inverse Fourier transform of the autocovariance function, \( R(U) \).

In addition, the autocovariance function \( R(U) \) and the power spectrum \( P(F) \) are even functions of their respective arguments. Hence, the relation between the two may be expressed more simply as two-sided cosine transformations. That is

\[
R(U) = \int_{-\infty}^{\infty} P(F) \cos 2\pi F U \, dF, \tag{D-7}
\]

and

\[
P(F) = \int_{-\infty}^{\infty} R(U) \cos 2\pi F U \, dU. \tag{D-8}
\]

For lag \( U = 0 \)

\[
R(0) = \int_{-\infty}^{\infty} P(F) \, dF = \text{VAR}, \tag{D-9}
\]

the sum of \( P(F) \, dF = \text{VAR} \).

\( P(F) \, dF \) is a measurement of the variance in frequency band \( F \) to \( F + dF \). Thus, \( P(F) \) shows how the variance of the \( X(T) \) process is distributed over frequencies.

2. Discrete time domain case.
The autocovariance function with lag $K$, and sampling interval $D$ may be defined as follows

$$R(K) = \lim_{N \to \infty} \frac{1}{N} \sum_{T=1}^{N-K} X(T)X(T+K)$$  \hspace{1cm} (D-10)

for $T=1, 2, \ldots, (N-K)$.

The power spectrum may be defined as

$$P(F) = D \sum R(K) \exp(-j2\pi F K D),$$
$$P(F) = D \sum R(K) \cos(2\pi F K D),$$

$-\infty < K < +\infty, -1/2D < F < 1/2D$, (D-11)

where $P(F)$ is a Fourier transform of the autocovariance function $R(K)$ over discrete lags, and it is a periodical function. Also $R(K)$ is the inverse Fourier transform of the power spectrum $P(F)$. That is

$$R(K) = \int_{-1/2D}^{1/2D} P(F) \exp(j2\pi F K D) \, dF,$$

$-1/2D < F < 1/2D, K=0, \pm 1, \pm 2, \ldots$  \hspace{1cm} (D-12)
Appendix E
SAMPLE SPECTRUM AND SAMPLE AUTOCOVARIANCE

Suppose \( x(T) \) is one sample record of the stochastic process with sample mean zero, sample variance \( \text{SVAR} \), sampling interval \( D \), and sample size \( N \). The mathematical relations between the sample spectrum and the sample autocovariance function are as follows:

1. Continuous time domain case.

Using equation (D-3), the sample spectrum may be defined as follows:

\[
C(F) = \frac{1}{N} \left( \frac{1}{N} \int_{-N/2}^{N/2} x(T) \cdot \text{EXP}(j2\pi F T) \, dT \right) ** 2,
\]

where \( N \) is the finite length of the sample data.

Equation E-1 may be reduced to the form

\[
C(F) = \int_{-N}^{N} \text{CCV}(U) \cdot \text{EXP}(-j2\pi F U) \, dU,
\]

where \( -N < U < N \).

Equation E-2 shows that the sample spectrum is the Fourier transform of the sample autocovariance function. This may be inverted to express the sample autocovariance function as the Fourier transform of the sample spectrum. Thus we have
\[ \text{COV}(U) = \int_{-N}^{N} C(F) \exp(j2\pi F U) \, dF, \quad -N < U < N \] (E-3)

and, for \( U=0 \), equation E-3 becomes
\[ \text{COV}(0) = \text{SVAR} = \int_{-\frac{1}{2D}}^{\frac{1}{2D}} C(F) \, dF = \left( \frac{1}{N} \right) \sum C(F). \] (E-4)

Thus the sample spectrum shows how the sample variance of \( x(T) \) is distributed over frequencies.

2. Discrete time domain case.

The sample spectrum is a Fourier transform of the sample autocovariance function. Thus,
\[ C(F) = D \sum \text{COV}(K) \exp(-j2\pi F K D), \quad -\frac{1}{2D} < F < \frac{1}{2D}, \]
\[ K = -(N-1), -(N-2), \ldots, 0, \ldots, (N-1). \] (E-5)

In addition, the sample autocovariance function is a Fourier transform of the sample spectrum
\[ \text{COV}(K) = \int_{-\frac{1}{2D}}^{\frac{1}{2D}} C(F) \exp(j2\pi F K D) \, dF, \]
\[ -\frac{1}{2D} < F < \frac{1}{2D}, \quad K = 0, \pm 1, \pm 2, \ldots, \pm (N-1). \] (E-6)
Appendix F

SMOOTHED SAMPLE SPECTRUM AND SAMPLE AUTOCOVARIANCE

Bartlett suggested in 1948 how spectral analysis could be modified so as to lead to improved spectral estimators. His was the simple device of splitting the series into $Q$ sets of $M$ terms so that $N=M*Q$. For each set of subseries, we conduct a spectral analysis and estimate the sample spectrum, $C_h(F_i)$, for $h=1,2,3,...,Q$. Then the smoothed sample spectrum, averaging the sample spectrum over sub-series at each frequency, may be written

$$SC(F_i) = \frac{1}{Q} \sum_{h=1}^{Q} C_h(F_i),$$  \hspace{1cm} (F-1)

The sample spectrum for the $h$th sub-series with length $M$ is

$$C_h(F) = \int_{-\frac{1}{2}D}^{\frac{1}{2}D} \text{COV}(U) \exp(-j2\pi F U) \, dU, \quad -M < U < M. \hspace{1cm} (F-2)$$

Define the average sample autocovariance at each lag as

$$SCOV(U) = \frac{1}{Q} \sum_{h=1}^{Q} \int_{(h-1)M}^{hM-U} X(T)X(T+U) \, dT,$$

$$(h-1)M < T < hM-U, \quad U > C. \hspace{1cm} (F-3)$$

From C-2 and C-3, Equation C-1 becomes

$$SC(F) = \int_{-\infty}^{\infty} SCOV(U) \exp(-j2\pi F U) \, dU, \quad M < U < M. \hspace{1cm} (F-4)$$

Since $E(SCOV(U))=R(U)*(1-(|U|/M))$, \hspace{1cm} (F-5)
and since Bartlett's lag window is

\[ w(U) = 1 - \frac{|U|}{M}, \quad |U| < M \]

\[ = 0, \quad |U| \geq M \]

Equation F-4 becomes

\[ S_C(F) = \int_{-\infty}^{\infty} \text{COV}(U) * w(U) * \exp(-j2\pi F U) \, dU. \]

Equation F-7 shows that the smoothed sample spectrum is the Fourier transform of the averaged (or weighted) sample autocovariance function. Furthermore, the number of sample autocovariance functions, that must be computed, is only up to \( M \): a truncation point.
Appendix G

RELATIONS BETWEEN THE VARIANCE AND THE TRUNCATION POINT

The variance of the smoothed spectral estimators is theoretically given by

\[ \text{VAR}( \hat{S}(F) ) = \left( \frac{\hat{P}(F)^2}{N} \right) \int_{-\infty}^{\infty} \hat{W}(U)^2 dU \]

\[ = \left( \frac{\hat{P}(F)^2}{N} \right) \frac{I}{B} = \left( \frac{\hat{P}(F)^2}{N \cdot B} \right) \]

where \( I = \int_{-\infty}^{\infty} \hat{W}(U)^2 dU \), \( B = 1/I = 1/\int_{-\infty}^{\infty} \hat{W}(U)^2 dU \)

\[ = 1/\int_{-\infty}^{\infty} \hat{W}(U)^2 dU \int_{-\infty}^{\infty} \hat{P}(F)^2 dF. \]  \( (G-1) \)

For Bartlett's lag window, \( I = (2/3)^{\text{th}} \), where 2/3 is a multiplier, and \( B = 3/(2^{\text{th}}) \).

So the variance is proportional to \( M \) and is inversely proportional to \( F \).

\( B \) is called the frequency resolution bandwidth in cycles per unit of time. \( \hat{W}(U) \) is called the lag window; \( \hat{P}(F) \) is called the spectral window. The lag window and the spectral window are Fourier transforms of each other.
Appendix H

PROPERTIES OF SAMPLE SPECTRAL ESTIMATORS: WHITE NOISE SERIES

Suppose \( X \) is a normal stochastic process with its expected value equal to zero, variance equal to \( \text{VAR} \), sampling interval equal to \( D \), and sample size equal to \( N \). From Appendix-A and Appendix-B, the sample spectrum estimators for the discrete case may be defined as

\[
C(F) = \left( \frac{D}{N} \right) \left( \sum (X \cdot \cos 2\pi F \cdot T \cdot D) \right)^2 + \left( \sum (X \cdot \sin 2\pi F \cdot T \cdot D) \right)^2
\]

where \( T = \left\{ \frac{-N}{2}, \frac{-N}{2} + 1, \ldots, 0, \ldots, \frac{N}{2} - 1 \right\} \),

\[-\frac{1}{2}D < F < \frac{1}{2}D . \tag{H-1} \]

This may be reduced to the form

\[
C(F) = \left( \frac{D}{N} \right) \left( A(F)^2 + B(F)^2 \right), \quad -\frac{1}{2}D < F < \frac{1}{2}D , \tag{H-2} \]

where \( A(F) = \sum (X \cdot \cos 2\pi F \cdot T \cdot D) \), \( \tag{H-3} \)

\[
B(F) = \sum (X \cdot \sin 2\pi F \cdot T \cdot D) . \tag{H-4} \]

Since \( E(X) = 0 \), \( E(A(F)) = 0 = E(B(F)) \). \( \tag{H-5} \)

Hence, at the harmonic frequencies \( F_i = i/(N \cdot D) \),

\[
\text{VAR}(A(F_i)) = E(A(F_i)^2) = \text{VAR} \cdot \cos^2 2\pi i \cdot F \cdot T \cdot D = \text{VAR} \cdot N/2, \quad i = \pm 1, \pm 2, \ldots, \pm (N-1) \]

\[
= \text{VAR} \cdot N, \quad i = 0, -N \tag{H-6} \]

- 208 -
Similarly,
\[ \text{VAR}(B(F_i)) = \text{VAR} \times \frac{N}{2}, \quad i=1, \pm 2, \ldots, \pm (N-1) \]
\[ = 0 \quad i=0, -N \]  
\[ \text{(H-7)} \]

Furthermore, when \( i=1 \)
\[ \text{COV}(A(F_i), A(F_l)) = \text{VAR} \times \text{SUM}(\cos 2\pi F_i \times l \times D) \cos 2\pi F_l \times l \times D) \]
\[ \text{COV}(A(F_i), A(F_l)) = 0 \]  
\[ \text{CCV}(B(F_i), B(F_l)) = 0. \]

In addition, for all \( K \) and \( L \)
\[ \text{COV}(A(F_i), B(F_l)) = 0. \]  
\[ \text{(H-9)} \]

Equation H-7 and H-8 show that the random variables are uncorrelated and thus independent. Now, since \( A(F_i) \) and \( B(F_i) \) are linear functions of normal random variables, \( A(F_i) \) and \( B(F_i) \) are also distributed normally. Hence, the random variables
\[ A(F_i)^2 / \text{VAR}(A(F_i)) = 2 \times A(F_i)^2 / (N \times \text{VAR}), \text{ and} \]
\[ B(F_i)^2 / \text{VAR}(B(F_i)) = 2 \times B(F_i)^2 / (N \times \text{VAR}) \]

are each distributed as CS. Since \( A(F_i) \) and \( B(F_i) \) are normal random variables. Hence their sum
\[ (2/(N \times \text{VAR})) \times (A(F_i)^2 + B(F_i)^2) \]
\[ = 2 \times C(F_i) / (D \times \text{VAR}) = Y(F_i) \text{ is distributed as CS.} \]  
\[ \text{(H-10)} \]

When \( k=0, -N, B(F_i) \) is identically zero. Hence,
\[ Y(F_i) = A(F_i)^2 / \text{VAR}(A(F_i)) = C(F_i) / (D \times \text{VAR}), \quad K=0, -N, \]  
\[ \text{(H-11)} \]

is distributed as CS.
Equation H-7 and H-8 imply that the random variables \( Y(F_i) \) at distinct frequencies are uncorrelated and hence independent because of the normality assumption.

Since \( 2\cdot C(F_i)/(D\cdot VAR) \) is distributed as CS with 2 d.f.,
\[
E\left(2\cdot C(F_i)/(D\cdot VAR)\right) = \text{number of d.f.} = 2.
\]
Therefore, \( E(C(F_i)) = D\cdot VAR \).

For white noise series, this is an unbiased estimator since the theoretical power spectrum \( P(F_i) = D\cdot VAR \).

and \( \text{VAR}(2\cdot C(F_i)/(D\cdot VAR)) = 2\cdot \text{d.f.} \).

Thus, \( \text{VAR}(C(F_i)) = (D^*2)\cdot \text{VAR}^*2 \).

This shows that at the harmonic frequencies, at least, the variance of the estimator is a constant independent of the sample size. So, it is inconsistent.

Even if \( X \) is not a normal process, if the sample size \( N \) is large, by the central limit theorem, the sample spectral estimator: \( 2\cdot C(F_i)/(D\cdot VAR) \) is approximately distribution as CS with 2 d.f. for all \( F \).
Appendix I

DISTRIBUTION OF THE SMOOTHED SAMPLE SPECTRAL ESTIMATORS

The smoothed sample spectrum $S_C(F)$ is a weighted sum of chi-square variables. Hence it is reasonable to approximate its distribution by a chi-square distribution. That is,

$$SC(F) \sim A \ast CS\nu$$

$$E(SC(F)) = E(A \ast CS\nu) = A \ast \nu,$$

and

$$\text{VAR}(SC(F)) = \text{VAR}(A \ast CS\nu) = 2(A^2) \ast \nu$$

where $\nu = 2(E(SC(F)))^2 / \text{VAR}(SC(F))$, and $A = E(SC(F)) / \nu$.

Since $\text{VAR}(SC(F)) = P(F) * 2 / (N \ast \int_{-\infty}^{\infty} (W(U))^2 dU)$,

and $E(SC(F)) = P(F)$,

$\nu = 2N / (\int_{-\infty}^{\infty} (W(U))^2 dU) = 2N / I$, and $A = P(F) / \nu$,

where $I$ is the lag window, $N$ is the sample size, and $M$ is the truncation point.
For generating U.U.R.N. and U.N.D.

For translating U.U.R.N. into U.N.D. by Box-Muller inverse and Hastings inverse
System Flow Charts

For translating U.U.R.N. into U.N.D. by Box-Muller rejection

For statistical tests
START

Read, Write, Print Heading Parameters and Seeds From Cards

Generate 50 Sets of Seeds for 50 Samples

No More Samples to be Generated?

Yes

Write and Print No. of Samples, Seeds onto Tapes

Call "GETRAN" and Corresponding Generator RTNS

End of One Set?

Yes

Write Generated Numbers for Set onto Tapes

End of Sets?

Yes

No

STOP

A

B

A

B

B

B
List of Programs

UURNGT, GETRAN, ADRAND, RANDU, RANNOS, RANDOM, RANC, URAND.
*** UURNGT ROUTINE ***

** This program generates seven types of unit uniform random numbers and two kinds of unit normal random variates (GGNML and GGNPM).

Each generated sequence contains 50 samples, and each sample contains 50 sets with 1200 no. per set. In this study, the generated sequence will be stored in the tape. The created tape will be used as the input file for normal translating routine: MAINNOR and statistical test routines.

** Required subroutines—GETTRAN, ADRAND, RANDU, RANNOS, RANDOM, RANC, URAND, and three IMSL RNS—GGUBS, GGNML, GGNPM.

** Input file:

1) Free format:
   ISET—N3. of independent samples. ISET=50.
   KL—No. of sets per sample. KL=50.
   N—NC. of random numbers per set. N=1200.

2) Heading and initial seeds for whole series:
   Heading: Format 4A10.
   Initial seeds: IA, IB, IC, ID——4110.

** Output file:

1) Magnetic tape:
   1) Heading and initial seeds for whole series (4A10. 4110).
   2) Number of samples and seeds for each sample (30X, 5110).
   3) 1200 generated numbers per set. 50 sets per sample.

<table>
<thead>
<tr>
<th>Heading, Seeds</th>
<th>No. of Sample, Seeds</th>
<th>Set 1</th>
<th>Set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>Set 50</td>
<td>...</td>
<td>No. of Sample, Seeds</td>
</tr>
<tr>
<td>...</td>
<td>Set 50</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

** Other key variables in the program

U—Array for unit uniform random number per set.
IA, IB, IC, ID—Seeds per sample.
IAA, IBE, ICC, IDD—Array for seeds of 30 samples.
REAL*8 RX
DIMENSION HEAD(10),U(5000),R(4),IAA(100),IBB(100),
1 ICC(100),IDD(100)
DATA IAA/100*0/,IBB/100*0/,ICC/100*0/,IDD/100*0/
CC * INPUT AND PRINT PARAMETERS--ISET, KL, N, AND IGG.
READ(5,*) ISET,KL,N,IGG
WRITE (6,155) ISET,KL,N,IGG
155 FORMAT (10X,'ISET=','I5.5X,'KL=','I5.5X,'N=','I5.5X,'1' IGG=','I5)
CC * INPUT HEADING AND SEEDS FOR WHOLE GENERATED SERIES
CC AND WRITE INPUT DATA ON THE TAPE
READ (5,5) (HEAD(I),I=1,10),IA,IB,IC,ID
WRITE (6,156) (HEAD(I),I=1,10),IA,IB,IC,ID
156 FORMAT (10X,10A4,'INITIAL SEEDS=','4I10/')
WRITE (1,5) (HEAD(I),I=1,10)
5 FORMAT (10A4,'4I10')
CC * GENERATE 50 INITIAL INTEGER SEEDS FOR 50
CC INDEPENDENT SAMPLES
RX=DFLOAT(IA)
DO 40 I=1,ISET
CALL GGL3S (RX,4,R)
IAA(I)=R(1)*100000000
IBB(I)=R(2)*100000000
ICC(I)=R(3)*100000000
IDD(I)=R(4)*100000000
40 CONTINUE
CC * GENERATE U,U,R,N FOR ISET=50 INDEPENDENT SAMPLES
CC WITH KL=50 SETS PER SAMPLE AND N=1200 NOS. PER SET.
DO 50 III=1,ISET
CC * INITIAL SEEDS FOR EACH SAMPLE
IA=IAA(III)
IB=IBB(III)
IC=ICC(III)
ID=IDD(III)
RX=DFLOAT(IA)
CC * WRITE INITIAL SEEDS AND NUMBER OF SAMPLES
CC ONTO THE OUTPUT TAPE FOR EACH SAMPLE
WRITE (1,25) III,IA,IB,IC,ID
25 FORMAT (30X,'E10.0')
WRITE (6,200) III,IA,IB,IC,ID
200 FORMAT (10X,'SAMPLE=','I5','INITIAL VALLE=','4I10')
DC 7000 II=1,KL
DO 30 I=1,II
CC * CALL CORRESPONDING GENERATOR RTNS AND GENERATE
CC U,U,R,N.
CALL GETRAN (IA,IB,IC,ID,IGG,YFL,R,RX)
IF (IGG .GT. 6 ) GOTO 31
CC * STORE U,U,R,N FOR GENERATORS--
CC ADRAND, RANDU, RANNCS, RANDOM, RANC, AND URAND
U(I)=YFL
GOTO 30
CC * STORE U,U,R,N FOR IMSL GENERATORS--GGU3S,
CC GGNML AND GGAPM.
31 U(I)=R(1)
30 CONTINUE
CC * WRITE THE GENERATED U,U,R,N, AND U,N,D. CNTQ
CC THE CUTFUT TAPE
CC WRITE (1,2C) (U(I),I=1,N)
20 FORMAT (20A4)
CC * GENERATE NEXT SET
7000 CONTINUE
CC * GENERATE NEXT SAMPLE
50 CONTINUE
STOP
END
*** GETRAN ROUTINE ***

** THIS ROUTINE WILL CALL EACH CORRESPONDING UNIT UNIFORM RANDOM NUMBER GENERATORS (U.R.N.G.) OR IBM IMSL SUBROUTINES (GGNML, GGNPM) BY MEANS OF PARAMETER IGG.

** SUBROUTINES REQUIRED ARE ADRAND, RANDU, RANNOS, URAND, RANDOM, RAN, AND THREE IBM IMSL ROUTINES--GGUBS, GGNML, GGNPM.

SUBROUTINE GETRAN (IA, IB, IC, ID, IGG, KLKL, YFL, R, RX)

** KEY VARIABLES IN THIS ROUTINE
IGG--IT IS THE SAME AS "UURNGT" ROUTINE.
YFL--GENERATED NUMBER.
RX--INITIAL SEED FOR IMSL ROUTINES.
R--ONE DIMENSIONAL ARRAY FOR GENERATED NUMBER FOR IMSL ROUTINES.
KLKL--NO. OF DIGITS PER GENERATED NUMBER.
IA, IB, IC, ID--INITIAL SEEDS FOR THE ROUTINES OTHER THAN IMSL.

DIMENSION R(1)
REAL*8 RX

1       GOTO (10, 20, 30, 40, 50, 60, 70, 90, 90), IGG
10      CALL ADRAND (IA, IB, IC, ID, KLKL, YFL)
        IF (YFL .EQ. 0. OR. YFL .EQ. 1.) GOTO 11
        RETURN
20      CALL RANDU (IA, YFL)
        IF (YFL .EQ. 0. OR. YFL .EQ. 1.) GOTO 11
        RETURN
30      CALL RANNOS (IA, YFL)
        IF (YFL .EQ. 0. OR. YFL .EQ. 1.) GOTO 11
        RETURN
40      YFL = URAND(IA)
        IF (YFL .EQ. 0. OR. YFL .EQ. 1.) GOTO 11
        RETURN
50      CALL RANDOM (IA, YFL, 1)
        IF (YFL .EQ. 0. OR. YFL .EQ. 1.) GOTO 11
        RETURN
60      YFL = RAN(1)
        IF (YFL .EQ. 0. OR. YFL .EQ. 1.) GOTO 11
        RETURN
70      CONTINUE
        CALL GGUBS (RX, 1, R)
IF ( R(1) .EQ. 0 . .OR. R(1) .EQ. 1. ) GOTO 11
RETURN
80 CONTINUE
CALL GGNML(RX,1,R)
RETURN
90 CONTINUE
CALL GGNPM (RX,1,R)
RETURN
END
*** ADRAND ROUTINE ***

** THIS IS R. BURFORD'S ADDITIVE GENERATOR ROUTINE. 

SUBROUTINE ADRAND (NA, NB, NC, ND, KLKL, YFL)
LOGICAL*1 ISW/.FALSE./

IF (ISW) GOTO 10
ITEMP=10. ** KLKL
TEMP=DFLOAT(ITEMP)
ISW=.TRUE.

10 MR=NA+NB+NC+ND
MR=MR-(ME/ITEMP) * ITEMP
YFL=MR
YFL=YFL/ITEMP
NA=NB
NB=NC
NC=ND
ND=MR
RETURN
END
*** RANDU ROUTINE ***

** THIS ROUTINE IS A CONGRUENTIAL GENERATOR FURNISHED BY IBM SSP PACKAGE.

SUBROUTINE RANDU (IX,YFL)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:
IX--FOR THE FIRST ENTRY THIS MUST CONTAIN ANY ODD INTEGER NUMBER WITH NINE OR LESS DIGITS.
AFTER THE FIRST ENTRY IX SHOULD BE THE PREVIOUS VALUE OF IY COMPUTED BY THIS ROUTINE.
IY--A RESULTANT INTEGER RANDOM NUMBER REQUIRED FOR THE NEXT ENTRY TO THIS SUBROUTINE.
The RANGE OF THIS NUMBER IS BETWEEN ZERO AND 2**31.
YFL--THE RESULTANT UNIFORMLY DISTRIBUTED, FLOATING POINT, RANDOM NUMBER IN THE RANGE 0 TO 1.

1 IY=IX * 65539
IF (IY) 5,6,6
5 IY=IY+2147483647 +1
6 YFL=IY
YFL=YFL * 0.4656613 E-9
IX=IY
RETURN
END
*** RANNO ROUTINE ***

THIS IS MACLAREN AND Marsaglia's MIXED MULTIPLICATIVE
CONGRUENTIAL GENERATOR ROUTINE (EMSHOFF AND
Sisson, 1970).

SUBROUTINE RANNO (MSED, YFL)
REAL RNEW
INTEGER PLUS/Z1000000/, IBIT/Z4000000/, HEX/Z10003/
INTEGER OLD, NEW, SEED
LOGICAL*1 NORMAL/.FALSE./
EQUIVALENCE (NEW, RNEW)
IF (NORMAL) GOTO 10
OLD = IABS (MSED)
NORMAL = .TRUE.
10 OLD = OLD * HEX
NEW = OLD / 256
IF (NEW .LT. 0) NEW = NEW + PLUS
NEW = NEW + IBIT
YFL = RNEW + 0.0
RETURN
END
*** RANDOM ROUTINE ***

THIS IS LEWIS-LEAHMONT'S UNIFORM RANDOM NUMBER GENERATOR ROUTINE. NAVAL POSTGRADUATE SCHOOL MONTEREY CALIFORNIA (LEWIS AND MILLER, 1969, KLEIJNEN, 1974).

0VFLOW CSFECT
ENTRY RANDOM
USING OVFLOW, R12
SAVE (14,12), *
B 12 (0,15)
DC AL1(6)
DC CL6'0VFLOW'

STM 14, 12, 12(13)
LR R12, R15
ST R13, SA+4
LR R2, R13
LA R13, SA
ST R13, 8(, E2)

ISSUE SPIE MACRO TO FIELD INTERRUPTS

SPIE XIT, (3, 9, 11, 12, 13, 15)
CNOP 0, 4
PAL 1, **10
DC BL1'000001010'
DC AL3(XIT)

SVC 14
ST R1', PICA
L R13, SA+4
RETURN (14, 12)
LM 14, 12, 12(13)
BR 14

HERE ON INTERRUPT

USING XIT, R15
TM 7(R1), X'F7'
BC 5, FORT
CLC 17(3, R1), ARANDOM+1
BNE 0(, R14)

MAKE THE CORRECTION, I.E. ADD 2**31-3 TO R4

A R4', PH2
AR R4, R2
BR R14
53 * LET US HANDLE THE INTERRUPT
54 *
55 PORT L R15, PICA
56 L R15, 0 (, R15)
57 BR R15
59 *
60 * GENERATE SOME RANDOM NUMBERS
61 *
62 * USING *, R15
63 RANDOM SAVE (14, 12), *
64 RANDOM B 12 (0, 15)
65* DC AL1 (6)
66* DC CL6* RANDOM *
67* STM 14, 12, 12 (13)
68 ST R13, SAI4
69 LR R2, R13
70 LA R13, SA
71 ST R13, 0 (, R2)
72 SUR FR0, FR0
73 LM R9, R11, A75
74 *
75 *
76* LA R2, 4
77 LM R5, R7, 0 (R1)
78* L R5, 0 (, R5)
79* L R3, 0 (, R7)
80* S LA R3, 2
81* SR R6, R2
82* LR R7, R2
83* LA R12, NONORM
84* CNOP 2, 8
85 *
86 * LOOP FOR N RANDOM NUMBERS
87 *
88 LOOP MR R4, R9
89* SLDA R4, 1
90* SRL R5, 1
91* AR R4, R5
92* LR R5, R4
93* SRL R4, 7
94* CR R4, 10
95* ST R4, 0 (R7, R6)
96* CR R4, R11
97* BNL 0 (, R12)
98* LE R32, 0 (R7, R6)
99* AER FR2, FR0
100* STE FR2, 0 (R7, R6)
101* NONORM BXLE R7, R2, LOOP
102 *
103* L R4, 0 (, R1)
104* ST R5, 0 (, R4)
105* L R13, SAI4
106* RETURN (14, 12)
107* LM 14, 12, 12 (13)
108* BR 14
* CONSTANTS AND STORAGE

PICA DC '0'
ARAND0M DC '0'(RANDOM)
PM2 DC '2147483645'
A75 DC '16807'
DC X'40000001'
DC X'40100000'
SA DS '18F'
R0 EQU 0
R1 EQU 1
R2 EQU 2
R3 EQU 3
R4 EQU 4
R5 EQU 5
R6 EQU 6
R7 EQU 7
R8 EQU 8
R9 EQU 9
R10 EQU 10
R11 EQU 11
R12 EQU 12
R13 EQU 13
R14 EQU 14
R15 EQU 15
FR0 EQU 0
FR2 EQU 2
END
*** RANC ROUTINE ***

This is R. Shore's assembler uniform generator routine. Each entry has 3 points. A single-precision uniform deviates on (0,1) is stored in U by

setting U=RAN(1) (Kinderman & Ramage, 1974).

RANC

CSECT
ENTRY RAN, SET, IGET

M1 EQU 0
M2 EQU 1
BASE EQU 15

USING RAN, BASE

This program is a FORTRAN-callable routine.

To generate unit RANDCM

L M2, CURNT GET THE LAST INTEGER
M M1, MULTR TIMES THE MULTIPLIER
AL M2, ADER PLUS THE CONSTANT
ST M2, CURNT ADD STORED IN THE RIGHT PLACE

SLL M2, 1 GET RLD OF THE EXTRAMEOUS BIT.
ST M2, CONV AND SET UP FOR FLOAT CONVERSION
LD M1, FLOTR NOW TO CLEARLY CONVERT TO FLOATING
AD M1, ZERO ALL DONE

BR 14 BACK TO THE MAIN

USING SET, BASE

SET L M2, 0 (M2)
L M2, 0 (M2) GET THE NUME
ST M2, CURNT PUT IT IN THE RIGHT PLACE
BR 14 BACK WE GO

USING IGET, BASE

IGET L 0, CURNT
BR 14

NOW THE CONTENTS

DS 0D GET THE ALIGNMENT RIGHT
ZERO DC X'4600000000000000'
MULTR DC F'1611715109'
ADER DC F'453815693'
FLOTR DC X'46000000'
CONV DC F'0'
CURNT DC F'0'
END
** Urand Routine **

Urand is a uniform random number generator based on theory and suggestions given in D. E. Knuth (1969) Vol. 2. The integer IY should be initialized to an arbitrary integer prior to the first call to Urand. The calling program should not alter the value of IY between subsequent calls to "Urand". Value of "Urand" will be returned in the interval (0,1) (Forsythe, Malcolm, 1977).

REAL FUNCTION Urand(IY)
INTEGER IY
INTEGER IA, IC, ITWO, M2, M, MIC
DOUBLE PRECISION HALFM
REAL S
DOUBLE PRECISION DATAN, DSORT
DATA M2/0/, ITWO/2/
IF (M2 .NE. 0) GOTO 20
10 M2=M
M=ITWO * M2
IF (M .GT. M2) GOTO 10
* IF first entry, compute machine integer word length
M=1
M2=M
M=ITWO * M2
IF (M .GT. M2) GOTO 10
* Compute multiplier and increment for linear congruential method
HALFM=M2
IA=8 * INT(HALFM * DATAN(1.0D0/8.0D0) + 5)
IC=2 * INT(HALFM * (0.5D0 - DSORT(3.0D0/6.0D0))) + 1
MIC=(M2 - IC) * M2
* S is the scale factor for converting to floating point
S =0.5/HALFM
* Compute next random number
20 IY=IY*IA
* The following statement is for computers which do not allow integer overflow no addition
IF (IY .GT. MIC) IY=(IY - M2) - M2
IY=IY+IC
* The following statement is for computers where the word length for addition is greater than for multiplication
IF (IY/2 .GT. M2) IY=(IY-M2) -M2

* THE FOLLOWING STATEMENT IS FOR COMPUTERS WHERE INTEGER OVERFLOW AFFECTS THE SIGN BIT

IF (IY .LT. 0) IY=(IY+M2) +M2

URAND =FLOAT(IY) *S
RETURN
END
START

Read, Write &
Print Heading
Parameters &
Seeds From
Cards

Generate 50
Sets of Seeds
for 50 Samples

Call "BW" & Corresponding RMS for
Generating & Translating

Call "GR" & Corresponding RMS
for Generating & Translating

Write & Print
UO onto Tapes

Write & Print
No. of Samples & Seeds onto Tapes

Yes

STOP

No

Yes

More Samples to be generated & translated?

No

Yes

No

A

B

STOP

Yes

No

No

B

A

No

End

No

End

No

End

No
List of Programs

URGBMJ, BM, GETRAN, ADRAND, RANDU, RANNOS, RANDOM, RANC, URAND.
** UGVRMBJ ROUTINE **

** This program generates seven types of unit uniform random numbers and translate these numbers into unit normal random numbers by box-muller rejection algorithms. Each generated sequence contains 50 samples. Each sample contains 50 sets with 1200 numbers per set. The generated series will be stored in the tape. The created tape will be used as the input file for statistical test routines. **

** Required subroutines—GETRAN, ADRAND, RANDU, RANNOS, RANDOM, RANC, URAND, IBM IMSL-GGUBS, and BM. **

** Input file:**

1) Free format:
   - ISET—No. of independent samples. ISET=50.
   - KL—No. of sets per sample. KL=50.
   - N—No. of random numbers per set. N=1200.
   - IGG—Index for unit uniform random number generators (U.U.R.N.G.).
     1: ADRAND; 2: RANDU; 3: RANNOS; 4: RANDOM;
     5: RANC; 6: URAND; 7: GGUBS.

2) Heading and initial seeds for whole series:

** Output file:**

1) Magnetic tape:
   1) Heading and initial seeds for whole series (4A10, 4I10).
   2) Number of samples and seeds for each sample (30X, 5I10).
   3) Generated numbers per set, 50 sets per sample.

** Other key variables in this routine are as follows:**

- U—Array for unit uniform random number per set.
- IA, IB, IC, ID—Seeds per sample.
- IAA, IBB, ICC, IID—Array for seeds of 30 samples.
- Head—Array for heading.

REAL*8 RX
DIMENSION HEAD(10), U(5000), R(4), IAA(100), IBB(100),
ICC (100) . IDD (100)
DATA IAA/100*0/ , IBB/100*0/ , ICC/100*0/ , IDD/100*0/
CC * INPUT AND PRINT PARAMETERS -- ISET, KL, N, AND IGG.
READ (5, *) ISET, KL, N, IGG
WRITE (6, 155) ISET, KL, N, IGG
CC * INPUT HEADING AND SEEDS FOR WHOLE GENERATED
CC SERIES AND WRITE INPUT DATA ONTO THE TAPE
READ (5, 5) (HEAD(I), I = 1, 10), IA, IB, IC, ID
WRITE (6, 156) (HEAD(I), I = 1, 10), IA, IB, IC, ID
156 FORMAT (10X, 10A4, 'INITIAL SEEDS=', 4110)
WRITE (1, 5) (HEAD(I), I = 1, 10), IA, IB, IC, ID
5 FORMAT (10A4, 4I10)
CC * GENERATE 50 INITIAL INTEGER SEEDS FOR 50
CC INDEPENDENT SAMPLES
RX=DFLOAT (IA)
DO 40 I = 1, ISET
CALL GGUBS (RX, 4, R)
IAA(I) = R(1) * 100000000
IBB(I) = R(2) * 100000000
ICC(I) = R(3) * 100000000
IDD(I) = R(4) * 100000000
40 CONTINUE
CC * GENERATE U.U.R.N. FOR ISET=50 INDEP. SAMPLES WITH
CC KL=50 SETS PER SAMPLE AND N=1200 NOS. PER SET.
DO 50 III = 1, ISET
CC * INITIAL SEEDS FOR EACH SAMPLE
IA = IAA(III)
IB = IBB(III)
IC = ICC(III)
ID = IDD(III)
RX=DFLOAT(IA)
CC * WRITE NUMBER OF SAMPLES AND INITIAL SEEDS ONTO
CC THE OUTPUT TAPE FOR EACH SAMPLE
WRITE (1, 25) III, IA, IB, IC, ID
25 FORMAT (30X, SI10)
WRITE (6, 200) III, IA, IB, IC, ID
200 FORMAT (10X, 'SAMPLE=', IS, ', INITIAL VALUE=', 4I10)
DO 7000 II = 1, KL
DO 30 I = 1, N
CC * CALL CORRESPONDING GENERATOR RTNS
CC AND GENERATE NORMAL RANDOM NUMBERS
U(I) = BM(IA, IB, IC, ID, IGG, KL, YFL, E, RX)
30 CONTINUE
CC * WRITE THE GENERATED UNIT NORMAL VARIATES ONTO
CC THE TAPE
WRITE (1, 20) (U(I), I = 1, N)
20 FORMAT (20A4)
CC * GENERATE NEXT SET
7000 CONTINUE
CC * GENERATE NEXT SAMPLE
50 CONTINUE
STOP
END
** THIS ROUTINE WILL BE CALLED FOR TRANSLATING U.N.D. GENERATED FROM GETRAN AND CORRESPONDING GENERATOR ROUTINES INTO U.N.D. BY BOX-MULLER REJECTION ALGORITHM. THE U.N.D. GENERATED HERE WILL BE RETURNED TO THE CALLING ROUTINE.

** REQUIRED SUBROUTINE: GETRAN

REAL FUNCTION BM(IA,IB,IC,ID,IGG,KLKL,YFL,R,RX)

* KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS: 
  IGG--INDEX FOR GENERATOR ROUTINES.
  IA, IB, IC, ID--INITIAL SEEDS FOR GENERATOR OTHER THAN IMSL'S GENERATORS.
  IX--INITIAL SEED FOR IMSL'S GENERATORS.
  YFL--GENERATED NUMBERS FROM GENERATORS OTHER THAN GGUBS.
  R--GENERATED NUMBERS FROM GGUBS.

DIMENSION R(1)
REAL*8 RX
LOGICAL*1 CK/.TRUE./
CK=.NOT.CK
IF (CK) GOTO 20
10 CALL GETRAN (IA,IB,IC,ID,IGG,KLKL,YFL,R,RX)
   IF (IGG .EQ. 7) GOTO 15
   V=2.*{YFL -0.5)
   GOTO 16
15 V=2.*{R(1) -0.5}
16 CALL GETRAN (IA,IB,IC,ID,IGG,KLKL,YFL,R,RX)
   IF (IGG .EQ. 7) GOTO 17
   W=2.*{YFL -0.5)
   GOTO 18
17 W=2.*{R(1) -0.5)
18 BM=V**2 + W**2
   IF ( BM .GT. 1 ) GOTO 10
   X=SQRT((-2.*ALOG (BM))/BM)
   BM=V*X
   RETURN
20 BM=W*X
RETURN
END
START

Read, Write, Print Heading, Parameters, Seeds from Cards

More Samples to be Translated?

No

STOP

Yes

Read & Write No. of Samples and Seeds for Each Sample From Tapes

Read Generated Nos. to be Translated From Tapes

A

A

Call "UNTRAN" for Translating

End of One Set?

No

Yes

Write UND onto Tapes

End of All Sets?

No

Yes

B
Lists of Programs

MAINNOR, URTRAN.
** MAINNOR ROUTINE **

** THIS PROGRAM TRANSLATES U.U.R.N. INTO U.N.D. **
A. READ U.U.R.N. FROM TAPE CREATED IN "UURNGT" ROUTINE.
B. TRANSLATE U.U.R.N. INTO U.N.D.
C. WRITE U.N.D. ONTO TAPE.

** THE REQUIRED SUBROUTINE IS UNTRAN. **

** INPUT FILE: **
1) FREEFORMAT: N, KL, ISET, IGG.
2) HEADING FROM CARD FILE, FORMAT: 20A4.
3) HEADING & STARTING SEEDS FOR WHOLE SERIES FROM TAPE FILE CREATED IN "UURNGT" RTN, FORMAT: 10A4, 4I10.
4) NUMBER OF SAMPLES AND SEEDS PER SAMPLE FROM TAPE FILE CREATED IN "UURNGT" RTN, FORMAT: 30X, 5I10.
5) U.U.R.N. FROM INPUT TAPE FILE CREATED IN "UURNGT" RTN, FORMAT: 20A4. N=1200 PER SET, 50 SETS PER SAMPLE, 50 SAMPLES FOR WHOLE SERIES.

** OUTPUT: TAPE FILE **
1) HEADING & STARTING SEEDS FOR WHOLE SERIES. FORMAT: 10A4, 4I10.
2) NUMBER OF SAMPLES AND SEEDS PER SAMPLE, FORMAT: 30X, 5I10.

** KEY VARIABLES USED IN THIS PROGRAM ARE AS FOLLOWS: **
ISET—NO. OF INDEPENDENT SAMPLES, LET ISET=50.
KL—NO. OF SETS PER SAMPLE, LET KL=50.
N—NO. OF RANDOM NUMBERS PER SET, LET N=1200.
IGG—TYPE OF TRANSFORMATION ALGORITHMS, IGG=1 FOR BOX-MULLER, IGG=2 FOR HASTING'S ROUTINE.
HEAD—AREA FOR HEADING OF U.U.R.N. INPUT FROM TAPE FILE.
HEADN—AREA FOR HEADING OF U.N.D. INPUT FROM CARD FILE.
U—AREA FOR U.U.R.N. INPUT FROM THE TAPE.
RMORM—AREA FOR U.N.D. PRODUCED IN THIS ROUTINE.

DIMENSION HEAD (10), HEADN (20), U (6000), RNORM (6000)
* INPUT PARAMETERS—ISET=50, KL=50, N=1200, IGG=1
* OR 2
READ (5,*) ISET, KL, N, IGG
WRITE (6, 155) ISET, KL, N, IGG
155 FORMAT (10X, 'ISET=', I5, 5X, 'KL=', I5, 5X, 'N=', I5, 5X, 'IGG=', I2)

** INPUT HEADING FROM CARD FILE
CC * INPUT HEADING & INITIAL SEEDS FROM INPUT TAPE FILE
READ (1,10) (HEAD(I),I=1,10),IA,IB,IC,ID
10 FORMAT (10A4,4I10)

CC * PRINT HEADING
WRITE (6,20) (HEAD(I),I=1,10),IA,IB,IC,ID,(HEADN(I)
1 ,I=1,20)
20 FORMAT (5X,'U,U,R,N.--',10A4,5X,'STARTING SEEDS=','
1 4I10 /5X,'NORMAL TRANSFORMATION--',5X,20A4/)

CC * WRITE HEADING & STARTING SEEDS ONTO THE OUTPUT TAPE
WRITE (2,10) (HEADN(I),I=1,10),IA,IB,IC,ID

CC * TRANSFORM U,U,R,N. INTO U,N,N. AND WRITE U,N,N
CC ONTO TAPE
DO 200 III=1,ISET

CC * READ NUMBER OF SAMPLES AND SEEDS FOR EACH SAMPLE
CC FROM THE INPUT TAPE FILE
READ (1,15) ISET1,IA,IB,IC,ID

CC * WRITE NUMBER OF SAMPLES AND SEEDS OF EACH SAMPLE
CC ONTO THE TAPE FILE
WRITE (2,15) ISET1,IA,IB,IC,ID
WRITE (6,16) ISET1,IA,IB,IC,ID
16 FORMAT (30X,5I10)
15 FORMAT (30X,5I10)
30 FORMAT (20A4)

TWOPA= FLOAT(2)*3.141592654
DO 201 II=1,KI

CC * READ U,U,R,N. FROM INPUT TAPE, AND TRANSLATE
CC THESE NUMBERS INTO U,N,D.
READ (1,30) (U(I),I=1,N)

CC * NORMAL TRANSFORMATION BY BOX-MULLER AND HASTING
DO 40 I=1,N,2
CALL UNTTRAN (U,RNORM,I,IGG,TWOPA)
40 CONTINUE

CC * WRITE U,N,D. ONTO TAPE
WRITE (2,30) (RNORM(I),I=1,N)

CC * PROCESS NEXT SET
201 CONTINUE

CC * PROCESS NEXT SAMPLE
200 CONTINUE

STOP
END
START

Read & Print Parameters

Compute Some Parameters

No

Yes

No More Samples to be Tested?

STOP

Read One Set of Nos. for Testing

Deviating Nos. from Mean

Call "DISFIT" for Fit Test

Call "RUNUD" & "RUNAB" for Runs Test

Call "AUTO" for Computing Autocovar.

Call "TSTAUT" for Autocor. Test

Call "TEST" for Spectral Test

End of All Sets

Call "DISFIT" for Fit Test

Call "RUNUD" & "RUNAB" for Runs Test

Call "AUTO" for Computing Autocovar.

Call "TSTAUT" for Autocor. Test

Call "TEST" for Spectral Test

End of All Sets

Yes
List of Programs

STSMAIN, DISFIT, RUNUD, RUNAB, TSTAUT, TEST, EQPDIS, SRUNLT, AUTO, FFT, RMEAN.
** THIS IS THE MAIN ROUTINE OF STATISTICAL TEST PROGRAMS. IT WILL CALL FIVE SUBROUTINES FOR TESTING DESIRED DISTRIBUTION FIT, RUNS UP AND DOWN, RUNS ABOVE AND BELOW THE MEDIAN, AUTOCORRELATION TESTS AND SPECTRAL TESTS.

** REQUIRED SUBROUTINES IN THIS ROUTINE ARE DISFIT, RUNUP, RUNAB, AUTO, TSTAOT, AND TEST.

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:

N—SAMPLE SIZE PER SET, LET N=1200.
KL—NUMBER OF SETS PER SAMPLE, LET KL=50.
ISFT—NUMBER OF SAMPLES FOR WHOLE SERIES, LET ISET=50.
U—ARRAY FOR RANDOM NUMBERS.
X—ARRAY FOR RANDOM NUMBERS DEVIATED FROM THE MEAN.
USFT—UPPER BOUND OF EXPECTED NO. OF SIGN. SETS PER SAMPLE.
HEAD--HEADING OF RANDOM NUMBERS FROM INPUT.
IA, IB, IC, ID—INITIAL SEEDS OF RANDOM NUMBERS.
ER1—SIGNIFICANCE LEVEL, TYPE I ERROR, LET ER1=0.2/0.05/0.01.
ZVAL—Z VALUE AT SIGNIFICANCE LEVEL ER1 (TWO TAILED TEST).
ZVALU—Z VALUE AT SIGNIFICANCE LEVEL ER1 (ONE TAILED TEST).
CH63—CRITICAL CHI-SQUARE VALUE WITH D.F.=63 FOR PORTMANTEAU STATISTIC TEST AT ER1 LEVEL.
CH62—CRITICAL CHI-SQUARE VALUE WITH D.F.=62 FOR DISTRIBUTION OF AUTOCORRELATION LAG TESTS AT ER1 LEVEL.
COV—AUTOCOVARIANCE OF TIME SERIES FOR LAGS 0-63.
COV0—VARIANCE OF TIME SERIES.
TPS—POWER SPECTRUM FOR WHITE NOISE SERIES.
L—TRUNCATION POINT, M, LET L=M=64.
K—K POWER OF 2 EQUALS TO L FOR FAST FOURIER TRANSFORM ROUTINE.
L2—NUMBER OF SAMPLE SPECTRA NEEDED FOR ANALYSIS.
W—ARRAY FOR BARTLETT'S LAG WINDOW.
DF—DEGREES OF FREEDOM OF SMOOTHED SAMPLE SPECTRA, DF=2M/1 FOR BARTLETT'S WINDOW DF=3M/M.
CHI1—CRITICAL CHI-SQUARE VALUE AT ER1/2 SIG. LEVELS WITH DF=L2-1. IT IS USED IN HYPOTHESIS TEST FOR SAMPLE VARIANCE.
CHI2—SAME AS CHI1 AT (1-ER1/2) SIGNIFICANCE LEVEL.
CH32—CRITICAL CHI-SQUARE VALUE WITH D.F.=30 FOR VAR. AND NORMALIZED VARIANCE OF SAMPLE SPECTRA TESTS.
CHIH—CHI-SQUARE VALUE FOR D.F.=DF=56 AT ER1/2 LEVELS.
CHIL—CHI-SQUARE VALUE FOR D.F.=DF=56 AT 1-ER1/2 LEVELS.
CC CH31—CRITICAL CHI-SQUARE VALUE WITH D.F. = 31
CC FOR DESIRED DISTRIBUTION.
CC RMEN—MEDIAN OF GENERATED SERIES FOR RUNS TEST.
CC CH4—CRITICAL CHI-SQUARE VALUE WITH D.F. = 4 FOR RUNS UP
CC AND DOWN TESTS.
CC CH6—CRITICAL CHI-SQUARE VALUE WITH D.F. = 6 FOR
CC RUNS ABOVE AND BELOW THE MEDIAN TESTS.

IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION X(5000),U(5000),COV(100),HEAD(10).
REAL*4 R,U

CC * INPUT PARAMETERS
CC * FOR GENERAL CASE
READ (5,*) N,KL,ISET,ER1,ZVAL,ZVALU

CC * FOR SPECTRAL ANALYSIS
READ (5,*) TPS,DF,L,K,CH32

CC * FOR DISTRIBUTION FIT
READ (5,*) CH31,CH9,IDX

CC * FOR AUTOCORRELATION ANALYSIS
READ (5,*) CH63,CH62

CC * FOR RUNS TESTS
READ (5,*) CH4,CH6,RMEN
L1=L/2-1
IF (DF.GT.100.D0) GOTO 2

CC * DF LE 100 Input CRITICAL VALUES FROM THE
CC CHI-SQUARE TABLE
READ (5,*) CHIL,CHIH
R D  G FR=CHIL*TPS/DF
HDGFR=CHIH*TPS/DF
GOTO 3

2 CONTINUE
CC * DF GT 100 COMPUTE CHI-SQUARE VALUE BY FORMULA
HDGFR=((2.0*DF-1.0)**0.5+ZVAL)**2/2.0*TPS/DF
RDGFR=((2.0*DF-1.0)**0.5-ZVAL)**2/2.0*TPS/DF

3 CONTINUE
IF (L1.GT.100) GOTO 4

CC * SAMPLE SIZE L1 LE 100, INPUT CRITICAL CHI-SQUARE
CC FROM THE TABLE
READ (5,*) CHI1,CHI2
GOTO 551

4 CONTINUE
CC * SAMPLE SIZES L1 GT 100, ESTIMATE CRITICAL
CC CHI-SQUARE VALUE
CHI1=((2.0*DFLOAT(L1)-1.0)**0.5-ZVAL)**2/2.0
CHI2=((2.0*DFLOAT(L1)-1.0)**0.5+ZVAL)**2/2.0

551 CONTINUE
CC * COMPUT UPPER BOUND FOR NO. OF SIGNIFICANT SETS
CC AND SAMPLES
USET =DFLOAT(KL)*ER1+DSQRT(DFLOAT(KL)*ER1*(1.0-ER1))
1 *ZVALU
USAMP=DFLOAT(ISET)*ER1+DSQRT(DFLOAT(ISET)*ER1*
1 *(1.0-ER1))*ZVALU

CC * PRINT INPUT DATA
WRITE (6, 11) N, KL, ISET, ER1, ZVAL, ZVALU, TPS, DF, L, K, CH32, CHIH
1 CHI1, CHI2, CH63, CH62, CH31, CH9, IDX, CH4, CH6, RMEN, USET, USAMP
11 FORMAT ('1', 1X, 'N', 15, 'KL', 'I3', 'ISET', 'I3', 'ER1='
1 , F6.4, 'ZVAL=', F6.4, 'ZVALU=', F6.4,
3 / 1X, '** SPECTRAL ANALYSIS:**', 'TPS=', 'F10.8', 'DF=', 'F4.1,
3 ' L=', 'I3,
4 ' K=', 'I3, ' CH32=', 'F6.2,' L. CHI-DF=', 'F6.2,' H. CHI-DF'
5 '** AUTOCORRELATION ANALYSIS **, ' CH63=', 'F6.2
5 ' ' CH62=', 'F6.2/
5 / 1X, '** DIST. FIT **', ' CH31=', 'F6.2, ' CH9=', 'F6.2
5 ' IDX=', 'I3,
6 / 1X, '** RUN TEST **', ' CH4=', 'F6.2, ' CH6=', 'F6.2
5 ' RMEN=', 'F6.2 / ** UPPER BOUND **
6 ' FOR NO. OF SIG. SETS=', 'F6.2, ' FOR NO. OF SIG.,
7 ' SAMPLES=', 'F6.2/)
CC
* INPUT HEADING, SEEDS AND GENERATED RANDOM NO S. AND
CALL STATISTICAL TEST ROUTINES
1 READ (1, 12, END=1000) (HEAD(I), I=1, 10), IA, IB, IC, ID
12 FORMAT (10 A4, 4110)
WRITE (6, 25) (HEAD (I) ,1=1 ,10) ,IA  , I B , I C , I D
25 FORMAT ( '1 10X,10A4 , 5 X,* INIT IA L SEEDS=»,411 0)
CC
* ISET=50 SAMPLES FOR THE WHOLE SERIES
DO 7000 I I I = 1 ,ISET
READ (1, 21) ISET1, IA, IB, IC, ID
21 FORMAT (30X, 5110)
WRITE (6, 35) ISET1, IA, IB, IC, ID
35 FORMAT ('//10X,'SET=», 'I5,5X, STARTING NUMBER*,
1 ' FOR THAT SET=», '411 0)
CC
* KL=50 SETS PER SAMPLE
DO 7001 III = 1, KL
CC
* INPUT RANDOM NUMBERS
READ (1, 14) (U(1), I=1, N)
14 FORMAT (20A4)
SU=0.00
DO 5 I=1, N
SU=SU+U(I)
5 CONTINUE
CC
* SUBTRACT FROM MEAN
SU=SU/DFLOAT(N)
DO 6 I=1, N
X(I)=U(I)-SU
6 CONTINUE
CC
* DISTRIBUTION FIT TESTS
CALL DISFIT (CH31, CH9, U, N, III, IIII, KL, ISET, USET
1 ,USAMP, IDX)
NN=N
CC
* RUNS UP AND DOWN TESTS
CALL RNUD (U, NN, III, IIII, KL, ISET, USET, USAMP, CH4,
1 CH9, ZVAL)
NN=N
CC
* RUNS ABOVE AND BELOW THE MEDIAN TESTS
CALL RUNAB (U, NN, III, IIII, KL, ISFT, USET, USAMP, CH6, CH9, 1  
         RMEN, ZVAL)
CALL AUTO (X, N, L, COV, COV0)
CC * AUTOCORRELATION TESTS
CALL TSTAUT (N, L, TPS, COV, COVO, IIII, IIII, KL, ISET, USET, 1  
           USAMP, ZVAL, ZVALU, CH9, CH62, CH63, ER1)
CC * SPECTRAL TESTS
CALL TEST (N, L, K, TPS, CHI1, CHI2, COV, RDGFR, HDGFR, IIII, IIII, 1  
          KL, ISET, USET, USAMP, ER1, ZVAL, ZVALU, CH9, CH32)
CC * PROCESS NEXT SET
7001 CONTINUE
CC * PROCESS NEXT SAMPLE
7000 CONTINUE
CC * END OF JOB
1000 WRITE (6, 2000)
2000 FORMAT (/55X, '*** END OF JOB ***')
STOP
END
**DISFIT ROUTINE**

**THIS ROUTINE WILL BE CALLED BY ROUTINE-STSMAIN FOR DESIRED DISTRIBUTION TESTS. IT COMPUTES NUMBER OF SIGNIFICANT SAMPLES AND GRAND CHI-SQUARE VALUES USING TESTS OF PEARSON'S CHI-SQUARE FOR WHOLE SERIES.**

**THE REQUIRED SUBROUTINE IS EQPDIS.**

SUBROUTINE DISFIT (CH31, CH9, U, N, III, IIII, KL, ISET, USAMP, IDX)

**KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:**
- **U**—ARRAY FOR GENERATED NUMBERS IN SINGLE PRECISION.
- **DU**—ARRAY FOR GENERATED NUMBERS IN DOUBLE PRECISION.
- **N**—NUMBER OF OBSERVATIONS.
- **IDX**—INDEX.
  1: UNIT UNIFORM DISTRIBUTION.
  2: UNIT NORMAL DISTRIBUTION.
  3: CHI-SQUARE DISTRIBUTION.
- **CHI**—CHI-SQUARE VALUE COMPUTED.
- **CTFT**—NUMBER OF SIGNIFICANT SETS.
- **TCTFT**—NUMBER OF SIGNIFICANT SAMPLES.
- **STCHI**—ARRAY FOR CHI-SQUARE VALUES OF EACH SET.
- **SMCHI**—ARRAY FOR CHI-SQUARE VALUES OF EACH SAMPLE.

IMPLICIT REAL*8 (A-H, O-Z)
REAL*4 U
DIMENSION STCHI(100), SMCHI(100), U(1), DU(5000)
DATA CTFT/0.0/, TCTFT/0.0/

**TEST DISTRIBUTION FIT PER SET**
DO 500 I = 1, N
DU(I) = U(I)
500 CONTINUE

**COMPUTE CHI-SQUARE VALUES FOR SETS**
CALL EQPDIS (0., 32, DU, N, IDX, CHI, III, ISET)

**STORE CHI-SQUARE VALUE FOR EACH SET**
STCHI(III) = CHI

**TEST SIGNIFICANT SETS**
IF (CHI GT. CH31) CTFT = CTFT + 1.D0

**END OF SAMPLE**
IF (III EQ. KL) GOTO 100

**PROCESS NEXT SET OF DATA**
RETURN

100 CONTINUE

**COMPUTE CHI-SQUARE VALUES FOR SAMPLES**
CALL EQPDIS (31., 10, STCHI, KL, 3, CHI, III, ISET)

**PRINT STATISTICS FOR EACH SAMPLE**
WRITE (6, 200) III, CTFT, CHI
200 FORMAT (/50X, '*** DISTRIBUTION FIT TESTS FOR SAMPLE', 1, 'I4', '***', '/1X', 'NO. OF SIG. SETS', 'F5.0', '5X',
`CHI-SQUARE VALUE FOR SAMPLE', F7.2/

CC  * CLEAR COUNTERS
CTFT=0.D0
CC  * DIST. FIT TEST PER SAMPLE
SMCHI(III)=CHI
IF (SMCHI(III) .GT. CH9) TCTFT=TCTFT+1.D0
CC  * TEST END OF SAMPLE
IF (III .EQ. ISET) GOTO 300
CC  * PROCESS NEXT SAMPLE
RETURN
CC  * DISTRIBUTION FIT TEST FOR WHOLE SERIES
300 CONTINUE
CC  * COMPUTE CHI-SQUARE VALUES FOR WHOLE SERIES
CALL EQPDIS (9.,10,SMCHI,ISET,3,CHI,III,ISET)
WRITE (6,400) (SMCHI(I),I=1,ISET)
400 FORMAT (1X,'DISTFIT=',10F10.5)
CC  * PRINT STATISTICS FOR THE WHOLE SERIES
WRITE (6,201) TCTFT,CHI
201 FORMAT (/55X,'*** DISTRIBUTION FIT TESTS FOR ',
1 'NO. OF SIG. SAMPLES FOR THE WHOLE SERIES IS',F5.0
1 ,5X,'CHI-SQUARE VALUE FOR THE SERIES IS',F6.2/
TCTFT=0.D0
RETURN
END
*** RUNUD ROUTINE ***

** This routine will be called by routine--STSMAIN for runs up and down tests. It computes no. of significant samples and grand chi-square values using tests of total number of runs and number of runs by length for the whole generated series.

** The required subroutines are SRUNLT and EOPDIS.

SUBROUTINE RUNUD (U, N, III, IIII, KL, ISET, USET, USAMP 1, CH46, CH9, ZVAL)

** Key variables in this routine are as follows:

U--Array for generated numbers to be tested.
RMEN--Median of generated sequences.
ICOT--Array for signs of observations.
N1--Number of signs.
DP--Degrees of freedom for number of runs by length.
NOLTH--No. of run by length=5.
CTRUN--No. of sig. sets by testing no. of runs by len.
CMRUN--No. of sig. samples by testing no. of runs by length.
CTTOL--No. of sig. sets by testing total no. of runs.
CMTOL--No. of sig. samples by testing total no. of runs.
EXPTL--Expected total no. of runs.
STD--Standard deviation of total no. of runs.
UPER--Upper bound of C.I. for sample total no. of runs.
FLOW--Lower bound of C.I. for sample total no. of runs.
EXPL--Expected no. of runs by length.
SMCHI--Array for chi-squares for each sample by length.
SMTOL--Array for chi-square values for each sample by total no. of runs.
STCHI--Array for chi-squares for each set by length.
STTOL--Array for chi-squares for each set by total no. of runs.

IMPLICIT REAL*8 (A-H, O-Z)
DIMENSION ICOT(3000),EXPL(20),STCHI(100),SMCHI(100)
1 ,STTOL(100),SMTOL(100),U(1)
REAL*4 DF,U
DATA EXPL/498.75D0,219.27D0,63.07D0,13.74D0,4.84D0,
1 15*0.D0/,CMTOL/0.D0/,CMRUN/0.D0/,ICOT/3000*9/
LOGICAL*1 KEY./. FALSE./

* IEQU=0
IF (KEY) GOTO 3
* First time enter runs up and down RTN.,
* Compute expect value, standard dev.
* and upper and lower bound
EXPFL=DFLOAT (2*N-1)/DFLOAT (3)
STD = DSQRT (DFLOAT(16*N-29) / DFLOAT(90))
UPER = EXPTL + (STD*ZVAL)
LOW = EXPTL - (STD*ZVAL)
NOLTH = 5
KEY = .TRUE.

CC * NOT FIRST ENTER THIS ROUTINE
3 CONTINUE
N1 = N - 1

CC * LIST IN ORDER OF OCCURRENCE THE SIGN (0/1) OF
CC THE DIFFERENCE BETWEEN SUCCESSIVE ITEMS
DO 100 I = 1, N1
IF (U(I) .EQ. U(I+1)) GOTO 101
IF (U(I) .LT. U(I+1)) ICOT(I) = 0
IF (U(I) .GT. U(I+1)) ICOT(I) = 1
GOTO 100
101 IEQU = IEQU + 1
WRITE (6,10) U(I)
10 CONTINUE

CC * CREATE INDICATOR FOR END OF RUN
IF (ICOT(N1) .EQ. 0) GOTO 203
ICOT(N1+1) = 0
GOTO 204
203 ICOT(N1+1) = 1
204 CONTINUE

CC * CLEAR COUNTERS
N1 = N1 - IEQU
CTRUN = 0. DO
CTTOL = 0. DO
IEQU = 0

CC * BUILD TOTAL NO. OF RUNS, NO. OF RUNS BY LEN.
CC AND CHI-SQ. VALUE OF NO. OF RUNS BY LENGTH
CALL SRUNLT (ICOT, N1, EXPT1, EXPTL, NOLTH, TOL, CS
1 , III, ISET)

CC * STORE CHI-SQ. VALUES OF NUMBER OF RUNS BY LENGTH
STCHI(III) = CS

CC * COUNT NUMBER OF SIGNIFICANT SETS
IF (CS .GT. CH46) CTRUN = CTRUN + 1. DO
CC * STORE NO. OF TOTAL NORMALIZED RUNS
ZTOL = (TOL - EXPTL) / STD
STTOL(III) = ZTOL
IF (TOL .LE. UPER .AND. TOL .GE. LOW) GOTO 105
CTTOL = CTTOL + 1. DO

CC * CHECK END OF SETS
105 CONTINUE
IF (III .EQ. KL) GOTO 290
RETURN

290 CONTINUE
DF = NOLTH - 1.

CC * COMPUTE CHI-SQUARE VALUES FOR SAMPLES
CALL EQPDIS (DF, 10, STCHI, KL, 3, CHI, III, ISET)
SMCHI(III) = CHI
CALL EQPDIS (1., 10, STTOL, KL, 2, CHI, III, ISET)
SMTOL(IIII) = CHI
CC
* COUNT NO. OF SIGNIFICANT SAMPLES FOR WHOLE SERIES
IF (SMCHI(IIII) .LT. CH9) CMRUN = CMRUN + 1. DO
IF (SMTOL(IIII) .LT. CH9) CMTOL = CMTOL + 1. DO
CC
* PRINT STATISTICS FOR SAMPLE
WRITE (6, 210) IIII, CMRUN, CMTOL, SMCHI(IIII), SMTOL(IIII)
210 FORMAT (/50X, '*** RUNS UP AND DOWN FOR SAMPLE ' , I4,
1 ' ***/1X, ' NO. OF SIG. SETS: BY LENGTH=' , F5.0,
1 ' BY TOTAL NO. OF RUNS'
1 , F5.0, 3X , ' CHI OF SAMPLE BY LENGTH=' , F6.2,
1 ' BY TOTAL NO. OF RUNS=' , F7.2/)
CC
* CHECK END OF SAMPLE
IF (I III .EQ. ISET) GOTO 300
RETURN
CC
* SUMMARY STATISTICS FOR THE WHOLE SERIES
300 CONTINUE
CC
* COMPUTE CHI-SQUARE VALUES FOR WHOLE SERIES
CALL EQPDIS (9., 10, SMCHI, ISET, 3, CHI, I III, ISET)
CALL EQPDIS (9., 10, SMTOL, ISET, 3, CHI1, I III, ISET)
WRITE (6, 311) (SMCHI(I), I = 1, ISET)
311 FORMAT (1X, ' BY LEN.=', 10F10.5)
WRITE (6, 312) (SMTOL(I), I = 1, ISET)
312 FORMAT (1X, 'TOTAL RUN=', 10F10.5)
CC
* PRINT STATISTICS FOR THE WHOLE SERIES
WRITE (6, 310) CMRUN, CMTOL, CHI, CHI1
310 FORMAT (/50X, '*** RUNS UP AND DOWN FOR WHOLE SERIES' ,
1 ' ***/1X, ' NO. OF SIG. SAMPLES BY LENGTH=' , F5.0, 1X,
1 ' BY TOTAL NO. OF RUNS=' , F5.0, 5X,
3 ' CHI BY LENGTH=' , F7.2, ' BY TOTAL NO. OF RUNS=' , F7.2/)
** RUNAB ROUTINE **

** THIS ROUTINE WILL BE CALLED BY ROUTINE--STSMAIN
IT COMPUTES NUMBER OF SIGNIFICANT SAMPLES AND
GRAND CHI-SQUARE VALUES USING TESTS OF TOTAL
NUMBER OF RUNS AND NUMBER OF RUNS BY
LENGTH FOR THE WHOLE GENERATED SERIES.

** REQUIRED SUBROUTINES ARE SPUNLT AND EQPDIS.

SUBROUTINE RUNAB (U, N, III, IIII, KL, ISET, USET, USAMP,
1 CH46, CH9, RMEN, ZVAL)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:
U--ARRAY FOR GENERATED NUMBERS TO BE TESTED.
RMEN--MEDIAN OF GENERATED SEQUENCES.
ICOT--ARRAY FOR SIGNS OF OBSERVATIONS.
N1--NUMBER OF SIGNS.
DF--DEGREES OF FREEDOM FOR NUMBER OF RUNS BY LENGTH.
NOLTH--NO. OF RUNS BY LENGTH = 7
CTRUN--NUMBER OF SIG. SETS BY TESTING NO. OF RUNS
BY LENGTH
CMRUN--NUMBER OF SIG. SAMPLES BY TESTING NO. OF RUNS
BY LENGTH
CTTOL--NO. OF SIG. SETS BY TESTING TOTAL NO. OF RUNS
CMTOL--NUMBER OF SIG. SAMPLES BY TESTING TOTAL NO. OF
RUNS
EXPTL--EXPECTED TOTAL NO. OF RUNS
STD--STANDARD DEVIATION OF TOTAL NO. OF RUNS
UPER--UPPER BOUND OF C.I. FOR SAMPLE TOTAL NO. OF RUNS
RLOW--LOWER BOUND OF C.I. FOR SAMPLE TOTAL NO. OF RUNS
EXPL--EXPECTED NO. OF RUNS BY LENGTH
SMCHI--ARRAY FOR CHI-SQUARE VALUES FOR EACH SAMPLE
BY LENGTH
SMTOL--ARRAY FOR CHI-SQUARE VALUES FOR EACH SAMPLE BY
TOTAL NO. OF RUNS
STCHI--ARRAY FOR CHI-SQUARE VALUES FOR EACH SET BY LENGTH
STTOL--ARRAY FOR CHI-SQUARE VALUES FOR EACH SET BY
TOTAL NO. OF RUNS

IMPLICIT REAL*8 (A-H, O-Z)
DIMENSION ICOT(3000), EXPL(20), STCHI(100), SMCHI(100),
1 STTOL(100), SMTOL(100), U(1)
REAL*4 DF, U
DATA EXPL/300.50D0, 150.13D0, 75.00, 37.47D0, 18.72D0, 9.35D0,
1 9.86D0, 13*0.D0/, CMRUN/0.D0/, CMTOL/0.D0/
LOGICAL*1 KEY/, FALSE./

* IEQU=0
IF (KEY) GOTO 4

* FIRST TIME ENTER THIS ROUTINE, COMPUTE PARAMETERS
SXPTI. = DFLOAT(A) /DFLOAT(2) + DFLOAT(1)
STD= DSQR D FLOAT(A*(N-2)) /DFLOAT(4*(N-1))
UPER= EXPTL+ (STD*ZVAL)
RLOW= EXPTL- (STD*ZVAL)
NOLTH= 7
KEY= .TRUE.
CONTINUE
N1=N
CC  * LIST IN ORDERED SEQUENCE OF SIGN (0/1) FOR EACH OBS.
    DO 200 I=1,N1
        IF (U(I) .EQ. PMEN) GOTO 201
        IF (U(I) .LT. RMEN) ICOT(I)=0
        IF (U(I) .GT. RMEN) ICOT(I)=1
    GOTO 200
201  IEQU=IEQU+1
    WRITE (6,10) U(I)
10  FORMAT (1X,'EQUA L=',F10.5)
    CONTINUE
CC  * CREATE INDICATOR FOR END OF RUN
    IF (ICOT(N1) .EQ. 0) GOTO 203
    ICOT(N1+1)=0
    GOTO 204
203  ICOT(N1+1)=1
204  CONTINUE
CC  * CLEAR COUNTERS
    N1=N1-IEQU
    CTRUN=0
    CTTOL=0
    IEQU=0
CC  * COUNT TOTAL NO. OF RUNS AND NO. OF RUNS BY LENGTH
    CALL SRUNIT (ICOT,N1,EXPTL,EXPL,NOLTH,TOL,CS
1 ,III, ISET)
CC  * STORE CHI-SQ. VALUES OF NUMBER OF RUNS BY LENGTH
    STCHI(III)=CS
CC  * COUNT NUMBER OF SIGNIFICANT SETS
    IF (CS .GT. CH46) CTRUN=CTRUN+1.
    DO
CC  * NORMALIZED TOTAL NUMBER OF RUNS AND STORE IT
    ZTOL=(TOL-EXPTL)/STD
    STTOL(III)=ZTOL
CC  * COUNT NO. OF SIG. SETS OF TOTAL NUMBER OF RUNS
    IF (TOL .LE. UPER .AND. TOL .GE. RLOW) GOTO 105
    CTTOL=CTTOL+1.
    DO
CC  * CHECK END OF SETS
105 CONTINUE
    IF (III .EQ. KL) GOTO 290
    RETURN
CC  * SUMMARY STATISTIC FOR SAMPLES
290 CONTINUE
    DF=NOLTH-1
CC  * COMPUTE CHI-SQUARE VALUE FOR SAMPLES
    CALL EQPDIS (DF,10,STCHI,KL,3,CHI,III, ISET)
    CHI(III)=CHI
    CALL EQPDIS (1.,10,STTOL,KL,2,CHI,III, ISET)
SMTOL(I) = CHI

CC  * COUNT NO. OF SIG. SAMPLES FOR WHOLE SERIES
    IF (SMCHI(I) .GT. CHI) CMRUN=CMRUN+1.D0
    IF (SMTOL(I) .GT. CHI) CMTOL=CMTOL+1.D0

CC  * PRINT STATISTICS FOR SAMPLE
    WRITE (6,210) I, CMRUN, CMTOL, SMCHI(I), SMTOL(I)

210 FORMAT (/50X,'*** RUNS ABOVE AND BELOW THE MEDIAN:
    1 ' FOR SAMPLE I, 13/
    1 'X,' NO. OF SIG. SETS: BY LENGTH=', F5.0,
    2 ' BY TOTAL NO. OF RUNS', F5.0,
    3 'X,' CHI OF SAMPLE BY LENGTH', F7.2,
    4 ' BY TOTAL NO. OF RUNS', F7.2/

CC  * CHECK END OF SAMPLE
    IF (I .EQ. ISET) GO TO 300
    RETURN

CC  * SUMMARY STATISTICS FOR THE WHOLE SERIES
300 CONTINUE

CC  * COMPUTE CHI-SQUARE VALUES FOR WHOLE SERIES
    CALL EQPDIS (9., 10, SMCHI, ISET, 3, CHI, I, ISET)
    CALL EQPDIS (9., 10, SMTOL, ISET, 3, CHI1, I, ISET)
    WRITE (6,311) (SMCHI(I), I = 1, ISET)
311 FORMAT (1X, 'BY LEN.=', 10F10.5)
    WRITE (6,312) (SMTOL(I), I = 1, ISET)
312 FORMAT (1X,'TOTAL RUN=', 10F10.5)

CC  * PRINT STATISTICS FOR THE WHOLE SERIES
    WRITE (6,310) CMRUN, CMTOL, CHI, CHI1
310 FORMAT (/50X,'*** RUNS ABOVE AND BELOW MEDIAN:
    1 ' FOR WHOLE SERIES ***/', 1X,
    1 'NO. OF SIG. SAMPLES BY LENGTH', F5.0,
    3 ' BY TOTAL NO. OF RUNS', F5.0,
    3 'CHI BY LENGTH', F7.2,
    4 ' BY TOTAL NO. OF RUNS', F7.2/

CC  * CLEAR COUNTERS
    CMRUN=0.D0
    CMTOL=0. D0
    RETURN
END
*** TEST ROUTINE ***

** THIS ROUTINE WILL BE CALLED BY ROUTINE_STSMAIN FOR SPECTRAL ANALYSIS TEST. IT COMPUTES NUMBER OF SIG. SAMPLES AND GRAND CHI-SQUARE VALUES USING TESTS OF DISTRIBUTION OF SAMPLE SPECTRA, SAMPLE MEAN OF SPECTRA, SAMPLE VARIANCE OF SPECTRA AND NORMALIZED SAMPLE VARIANCE OF SPECTRA FOR WHOLE GENERATED SERIES.

** REQUIRED SUBROUTINE IS FFT AND RMEN.

SUBROUTINE TEST(N,L,K,TPS,CHI1,CHI2,COV,RCGPR,HDGFR,  
1 III,III,KL,ISET,USET,USAMP,ER1,ZVAL,ZVALU,CH9,CH32)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:
L2—NO. OF SAMPLE SPECTRA IN THIS TEST, LET L2=31
W—ARRAY FOR LAG WINDOWS.
COV—ARRAY FOR AUTOCOVARIANCE FUNCTIONS.
COW—ARRAY FOR SMOOTHED AUTOCOVARIANCE FUNCTIONS.
SS—SAMPLE VARIANCE OF SAMPLE SPECTRA.
SSN—NORMALIZED SAMPLE VARIANCE OF SAMPLE SPECTRA.
PVAR—THEORETICAL VARIANCE OF POWER SPECTRA.
UCIPVR—U. BOUND FOR THEORETICAL VAR. OF P. SPECTRA.
RCIPVR—L. BOUND FOR THEORETICAL VAR. OF P. SPECTRA.
COTVAR—NO. OF SIGNIFICANT SETS BY TESTING VARIANCE.
PNVAR—NORMALIZED THEORETICAL VAR. OF P. SPECTRA.
UCIPNV—U. BOUND FOR NORMALIZED VAR. OF P. SPECTRA.
RCIPNV—L. BOUND FOR NORMALIZED VAR. OF P. SPECTRA.
COTNVR—NO. OF SIG. SETS BY TESTING NORMALIZED VAR.
TPS—THEORETICAL MEAN OF POWER SPECTRUM.
SUM—SAMPLE MEAN OF SAMPLE SPECTRA.
RMEN—ARRAY FOR NORMALIZED SAMPLE MEAN OF SPECTRA.
VAR—ARRAY FOR RATIO OF SAMPLE VAR. AND THEORETICAL VARIANCE OF SPECTRA.
VARN—ARRAY FOR RATIO OF NORMALIZED SAMPLE VAR. AND NORMALIZED THEORETICAL VARIANCE OF SPECTRA.
CHISPS—ARRAY FOR CHI-SQ. VALUES OF DISTRIBUTION OF SPECTRA.
UCIPMN—U. BOUND FOR THEORETICAL MEAN OF POWER SPECTRA.
RCIPMN—L. BOUND FOR THEORETICAL MEAN OF POWER SPECTRA.
COTAVG—NO. OF SIG. SETS BY TESTING MEAN.
TCTAVG—NO. OF SIG. SAMPLES BY TESTING MEAN.
COTSPS—TOTAL NO. OF SIGN. SETS BY TESTING S. SPECTRA.
COTPS—NO. OF SIG. SETS BY TOTAL NO. OF SIG. SPECTRA.
TCTPS—NO. OF SIG. SAMPLES BY TOTAL NO. OF SIG. SPECTRA.
DCTSPS—NO. OF SIG. SETS BY KTH SIG. SPECTRUM.
TDSPS—NO. OF SIG. SAMPLES BY KTH SIG. SPECTRUM.

IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION COV(1), DCTSPS(100), TDSPS(100), CHISPS(100),  
1 W(100), RMEN(100), VAR(100), VARN(100), SCHI1(100),
CC IF (KEY) GOTO 1
CC * FIRST TIME ENTER THIS ROUTINE, COMPUTE PARAMETERS
L2=L/2-1
CC * COMPUTE U. BOUND FOR NO. OF SIGNIFICANT P. SPECTRA
CC — PER SET, LET L2=31
EXPS=DFLOAT (KL) *ER1
UCINO=DFLOAT(L2) *ER1+DSQRT (DFLOAT (L2) *ER1 * (1. D0-ER1) )
1  *ZVAL
CC * COMPUTE BARTLETT'S LAG WINDOW
DO 100 I=1,L
W(I)=DFLOAT (1) -(DFLOAT (I-1)/DFLOAT (L) )
100 CONTINUE
CC * THEORETICAL VAR., STANDARD DEV. AND NORMALIZED VAR.
PNVAR=DFLOAT(2*(L2-1) *2 )/DFLOAT (N*3)
PVAR=PNVAR *TPS**2
PSTDSC=D SQRT (PVAR/DFLOAT (L2) )
CC * CONFIDENCE INTERVAL FOR SAMPLE MEAN OF SPECTRA
UCIPMN=TPS+ZVAL * PSTDSC
RCIPMN = TPS-ZVAL, * PSTDSC
CC * CONFIDENCE INTERVAL FOR VAR. OF SAMPLE SPECTRA
UCIPVR=CHI2* PVAR/DFLOAT (L2-1)
RCIPVR = CHI1* PVAR/DFLOAT (L2-1)
CC * CONFIDENCE INTERVAL FOR NORM. VAR. OF S. SPECTRA
UCIPNV = CHI2* PNVAR/DFLOAT (L2-1)
RCIPNV = CHI1* PNVAR/DFLOAT (L2-1)
CC * NOT THE FIRST TIME ENTER THIS SUBROUTINE
KEY=.TRUE.
1 CONTINUE
CC * WEIGHT AUTOCOVARIANCE BY BARTLETT'S WINDON
CC AND FOURIER TRANSFORM
DO 15 I=1,L
COW(I)=COV (I) *W (I)*DPLOAT (2)
15 FI(I)=0.D0
COW(1)=COV (1) *W (1)
CC * F. TRANSFORM OF WEIGHTED AUTOCOV. FUNCTIONS
CALL FFT(COW,FI,K)
CC * ANALYST SPECTRA: 34-64
DO 29 I=1,L2
COW(I)=COW(I+33)
29 CONTINUE
CC * COMPUTE SAMPLE MEAN AND VAR. OF SPECTRA
CALL RMEAN (COW,L2,SMN,SMN2,SS,SSE,TPS)
CC * NORMALIZED SAMPLE VARIANCE
SSN=SS/SMN**2
CC * STORE SAMPLE MEANS, S. VAR. AND NORMALIZED VAR.
RMEN (III) = (SUM-TPS) /PSTDSC
VAR (III) = DFLOAT (L2-1) * SS/PVAR
VARN (III) = DFLOAT (L2-1) * SSN/PNVAR

CC * SIG. SETS BY MEAN, VARIANCE, AND N. VAR.
IF (SUM.GT.UCIPMN.OR.SUM.LT.RCIPMN) COTAVG = COTAVG + 1.D0
IF (SSN.GT.UCIPVNR.OR.SSN.LT.RCIPVNR) COTNVR = COTNVR + 1.D0
IF (SS.GT.UCIPVR.OR.SS.LT.RCIPVR) COTVAR = COTVAR + 1.D0

CC * COUNTS NO. OF SPS ABOVE & BELOW EXP. C. I.
COTSPS = 0.D0
DO 10 I = 1, L2
IF (COW(I).LE.HDGFR.AND.COW(I).GE.RDGFR) GOTO 10
DCTSPS(I) = DCTSPS(I) + 1.D0
COTSPS = COTSPS + 1.D0
10 CONTINUE
IF (COTSPS.GT.UCINO) COTPS = COTPS + 1.D0
IF (III.EQ.KL) GOTO 50

CC * PROCESS NEXT SET OF DATA
RETURN

CC * SUMMARY STATISTICS FOR EACH SAMPLE
50 CONTINUE

CC * CALCULATE CHI-SQ. VALUE OF DIST. OF S. SPECTRA
DO 101 I = 1, L2
CHISPS(III) = CHISPS(III) + ((EXPS-DCTSPS(I)) ** 2) / EXPS
101 CONTINUE

CC * COMPUTE CHI-SQUARE VALUES OF MEAN, VAR. AND N.
CC VARIANCE FOR SAMPLES
CALL EQPDIS (1., 10, EMEN, KL, 2, CHI, III, ISET)
SCH1(III) = CHI
CALL EQPDIS (30., 10, VAP, KL, 3, CHI, III, ISET)
SCH2(III) = CHI
CALL EQPDIS (30., 10, VARN, KL, 3, CHI, III, ISET)
SCH3(III) = CHI

CC * PRINT STATISTICS FOR SAMPLES
WRITE (6, 51) III, COTAVG, COTVAR, COTNVR, COTPS, CHISPS(III),
1 SCH1(III), SCH2(III), SCH3(III), CH32, CH9
51 FORMAT (/1X, '** SUMMARY STATISTICS OF EACH SAMPLE*',
1 ' FOR POWER SPECTRUM **',
1 ' NO. OF SIGNIFICANCE OVER TOTAL SETS OF EACH SAMPLE=',
1 ' MEAN=', F5.0, ' VARIANCE=', F5.0, ' N. VARIANCE=',
1 ' TOTAL NO. OF SIG. SPECTRA=', F5.0, /
5 ' CHI-SQUARE VALUES FOR', ' SAMPLE MEANS', ' SAMPLE VARIANCE',
6 ' SAMPLE VARIANCES', ' CRITICAL CHI-VALUE FOR', ' DIST. OF S.',

CC * WRITE DIST. OF SAMPLE SPECTRA FOR EACH FREQ.
WRITE (6, 52) (DCTSPS(I), I = 1, L2)
52 FORMAT (1X, 'NO. OF SPS AT FREQ. K OVER TOTAL',
1 ' NO. OF SETS', F10.2, 1X)

CC * TEST NO. OF SIGNIFICANT SAMPLES
DO 53 I = 1, L2
IF (DCTSPS(I).GT.USET) TDSPS(I) = TDSPS(I) + 1.D0
53 CONTINUE
IF (CHISPS(III).GT.CH32) TCTPS = TCTPS + 1.D0
IF (SCH1 (III) .GT. CH9) TCTAVG = TCTAVG + 1.D0
IF (SCH2 (III) .GT. CH9) TCTVAR = TCTVAR + 1.D0
IF (SCH3 (III) .GT. CH9) TCTNVR = TCTNVR + 1.D0

CC * CLEAR COUNTERS FOR NEXT SET OF DATA
DO 54 I = 1, L2
DCTSPS (I) = 0.D0
CONTINUE

CCTPS = 0.D0
COTAVG = 0.D0
COTVAR = 0.D0
COTNVR = 0.D0

CC TEST END OF JOB
IF (III .GE. ISET) GOTO 60
RETURN

CC * SUMMARY STATISTICS FOR WHOLE SERIES

60 CONTINUE

CC * COMPUTE CHI-SQUARE VALUES FOR WHOLE SERIES
CALL EQPDIS (30., 10, CHISPS, ISET, 3, CHI1, IIII, ISET)
CALL EQPDIS (9., 10, SCHI1, ISET, 3, CHI2, IIII, ISET)
CALL EQPDIS (9., 10, SCHI2, ISET, 3, CHI3, IIII, ISET)
CALL EQPDIS (9., 10, SCHI3, ISET, 3, CHI4, IIII, ISET)
WRITE (6, 73) (CHISPS (I), I = 1, ISET)
    FORMAT (1X, 'CHISPS=', 10F10.5)
WRITE (6, 74) (SCHI1 (I), I = 1, ISET)
    FORMAT (1X, 'SCHI1=', 10F10.5)
WRITE (6, 75) (SCHI2 (I), I = 1, ISET)
    FORMAT (1X, 'SCHI2=', 10F10.5)
WRITE (6, 76) (SCHI3 (I), I = 1, ISET)
    FORMAT (1X, 'SCHI3=', 10F10.5)

CC * PRINT STATISTICS FOR THE WHOLE SERIES
WRITE (6, 70) IIII, TCTAVG, TCTVAR, TCTNVR, TCTPS, USAMP
    1, CHI1, CHI2, CHI3, CHI4, CH9
70 FORMAT (/ / / 1X, '*** SUMMARY STATISTICS OF ALL SAMPLES' 
        1, ' FOR POWER SPECTRUM ***' / 
        2, '1X, 'NO. OF SIGNIFICANT SAMPLES OVER NO. OF TOTAL' 
        3, 'SAMPLES=', I5/
        4, '3X, 'MEAN=', F5.0, 2X, 'VARIANCE=', F5.0, 2X, 'N. VARIANCE=', 
        5, 'F5.0, 2X, 'TOTAL NO. OF SIG. SPECTRUM=', F5.0, 
        6, '2X, 'NO. OF TOTAL SIG. SAMPLES--UPPER BOUND', F5.0/ 
        8, '1X, 'CHI VALUE FOR DIST. OF S. SPECTRA', F7.2, 
        9, 'FOR SAMPLE MEAN', F7.2, 
        3, 'SAMPLE VARIANCES', F7.2, 'N. SAMPLE VARIANCES', F7.2/1X, 
        9, 'CRITICAL CHI-SQUARE VALUE IS', F7.2/)

CC * PRINT DISTRIBUTION OF SPECTRA OVER 50 SAMPLES
WRITE (6, 72) (TDSPS(I), I = 1, L2)
72 FORMAT (1X, 'NO. OF SIG. SAMPLES AT FREQ. K OVER TOTAL' 
        1, 'NO. OF SAMPLES=', I1, 10(F10.2, 1X))
** TSTAUT ROUTINE **

** THIS ROUTINE WILL BE CALLED BY MAIN ROUTINE-STSMAIN FOR AUTOCORRELATION TESTS. IT WILL COMPUTE NUMBER OF SIGNIFICANT SAMPLES AND GRAND CHI-SQUARE VALUES USING TESTS OF DISTRIBUTION OF AUTOCORRELATION FUNCTIONS AND PORTMANTEAU CHI-SQUARE STATISTIC FOR WHOLE GENERATED SERIES. **

** THE REQUIRED SUBROUTINE IS EQPDIS. **

AUTOCOVARIANCE SUBROUTINE

SUBROUTINE TSTAUT (N,L,TPS,COV,COVO,III,III,KL,
1 ISET,USET,USAMP, ZVAL,ZVALU,CH9,CH62,CH63,ER1)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS: **

L—NO. OF LAGS COMPUTED (INCLUDING LAG 0), LET L=64.
N—NO. OF RANDOM NUMBERS PER SET, LET N=1200.
COV—ARRAY FOR AUTOCOVARIANCE.
COR—ARRAY FOR AUTOCORRELATION.
UCOV—UPPER BOUND OF THEORETICAL CONFIDENCE INTERVAL FOR SAMPLE AUTOCORRELATION FUNCTIONS.
RCOV—LOWER BOUND OF THEORETICAL CONFIDENCE INTERVAL FOR SAMPLE AUTOCORRELATION FUNCTIONS.
UCINOL—UPPER BOUND FOR NUMBER OF SIG. AUTOCOVARIANCE FUNCTIONS PER SET.
EXP—EXPECTED NUMBER OF SIG. AUTOCOR. PER SAMPLE.
USET—UPPER BOUND FOR EXPECTED NO. OF SIG. SETS.
USAMP—UPPER BOUND FOR EXPECTED NO. OF SIG. SAMPLES.
SCLAG1—TOTAL NO. OF SIG. LAGS PER SET.
SIGNO—NO. OF SIG. SETS BY TOTAL NO. OF SIG. LAGS.
CHICOR—COMPUTED CHI-SQUARE VALUE OF THE DIST. OF AUTOCORRELATIONS.
TCTCOR—NO. OF SIG. SAMPLES USING TEST BY UNIFORM DISTRIBUTION OF SIGNIFICANT LAGS.
CTLAG1—NO. OF SIG. SETS PER SAMPLE AT LAG K USING TEST BY DIST. OF AUTOCORRELATION COEFFICIENTS.
K=1,2,...,(L-1).
TCTLAG—NO. OF SIG. SAMPLES AT LAG K USING TEST BY DISTRIBUTION OF AUTOCORRELATION COEFFICIENTS.
K=1,2,...,(L-1).
CH62—CRITICAL CHI-SQUARE VALUE WITH (L-1) D.F.=62 AT SIG. LEVEL ALPHA FOR DIST. OF SIG. LAGS TEST.
CH63—CRITICAL CHI-SQUARE VALUE WITH (L-1) D.F.=63 AT SIG. LEVEL ALPHA FOR PORTMANTEAU CHI-SQUARE TEST.
PTMANT—COMPUTED CHI-SQ. PORTMANTEAU STATISTIC PER SET.
COTPTM—NUMBER OF SIGNIFICANT SETS PER SAMPLE USING TEST BY PORTMANTEAU CHI-SQUARE STATISTIC.
TCTPM—NO. OF SIG. SAMPLES USING TEST BY PORTMANTEAU CHI-SQUARE STATISTIC.
IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION COV(1),CTLAG1(100),COR(100),TCTLAG(100)
1,CHICOR(100),PTCHI(100),PTMANT(100)
LOGICAL*1 KEY/.FALSE./
DATA C'TLAG1/100*0.D0/,COTPTM/0.D0/,TCTLAG/100*0.D0/
1,TCTCOR/0.D0/,SIGNO/0.D0/
2,CHICOR/100*0.D0/

CC
IF (KEY) GOTO 1
CC * FIRST TIME ENTER THIS ROUTINE, CALCULATE PARAMETERS
CC * COMPUTE CONFIDENCE INTERVALS FOR SAMPLE AUTOCOR.
RCOV=-(((1.D0/DFLOAT(N))**0.5))*ZVAL
UCOV=((1.D0/DFLOAT(N))**0.5)*ZVAL
CC * COMPUTE U.BOUND FOR NO. OF SIG. AUTO LAGS PER SET
UCINOL=DFLOAT(L-1)*ER1
1+DSQRT (DFLOAT (L -1 ) *ER1 *(1. DO-ER1))*ZVALU
CC * COMPUTE EXPECTED NO. OF SIG. AUTOCOR. PER SAMPLE
EXP=DFLOAT(KL)*ER1
CC * NOT THE FIRST TIME ENTER THIS SUBROUTINE
KEY=.TRUE.
1 CONTINUE
SUMCOV=0.DO
SCLAG1=0.DO
CC * COMPUTE AUTOCORRELATION, LAGS=1 TO L-1=63
DO 100 I=2,L
COR(I)=COV(I)/COVO
CC * COUNT NO. OF SAMPLE AUTOCOR. OUT OF CONF. INTERVAL
IF (COR(I) .GE. RCOV .AND. COR(I) .LE. UCOV) GOTO 100
CC * COUNT TOTAL NO. OF SIG. LAGS AND PORTMANT E U PER SET
SCLAG1=SCLAG1+1.DO
CC * NUMBER OF SIGNIFICANT LAGS FOR ITH LAG PER SAMPLE
CTLAG1(I-1)=CTLAG1(I-1)+1.DO
100 SUMCOV=((COR(I)**2)/(DFLOAT(N-I+1))) +SUMCOV
CC * CALCULATE PORTMAENTEUA STATISTIC
PTMANT(III)=SUMCOV*DFLOAT(N*(N+2))
CC * COUNT NO. OF SIGNIFICANT SETS
IF (PTMANT(III) .GT. CH63) COTPTM=COTPTM+1.DO
IF (SCLAG1 .GT. UCINOL) SIGNO=SIGNO+1.DO
IF (III .EQ. KL) GOTO 50
CC * PROCESS NEXT SET OF DATA
RETURN
CC * SUMMARY STATISTICS FOR EACH SAMPLE
50 CONTINUE
CC * COMPUTE CHI-SQUARE VALUES FOR SAMPLES
CALL EQPRDIS (63.,10,PTMANT,KL,3,CHI,III,ISET)
PTCHI(III)=CHI
CC * CALCULATE CHI-SQUARE VALUES OF DIST. OF AUTO. LAGS
L1=L-1
DO 60 I=1,L1
60 CHICOR(III)=CHICOR(III)+((EXP-CTLAG1(I))**2)/EXP
CC * PRINT STATISTICS FOR SAMPLES
WRITE (6,51) IIII,SIGNO,L,CHICOR(III),COTPTM,CHI
**AUTO CORRELATION ANALYSIS**

**FORMAT**

261

```
51 FORMAT (///55X,'*** AUTOCORRELATION ANALYSIS FOR ',
4 ,' SAMPLE ***',I4/' TOTAL NO. OF SIG.' ,
5 ' SETS BY TOTAL NO. OF SIG. LAGS',F5.0,2X ,
6 ' CHI-SQUARE VALUE OF DIST.',I4,' AUTO. LAGS',F7.2/1X ,
7 ' NO. OF SIG. SETS BY COUNTING PORTMANTEAU TEST',F7.2 ,
8 2X, 'PORTMANTEAU CHI-SQUARE VALUE',F7.2/)
WRITE (6,52) SIGNO, (CTLA3(I),I=1,L1)
52 FORMAT (1X,'TOTAL NO. OF SIG. SETS',F5.0/1X ,
1 ' NO. OF SIG. SETS AT LAG K OVER TOTAL NO. OF SETS PER ',
1 ' SAMPLE'/,1X,10(F10.5))

CC

* TEST SIGNIFICANT SAMPLES
  IF (PTCHI(I)) .GT. CH9 ) TCTPTM=TCTPTM+1.D0
  IF (CHICOR(I)) .GT. CH62 ) TCTCOR=TCTCOR+1.D0
DO 65 I=1,L1
  IF (CTLAG(I) .GT. USET) TCTLAG(I)=TCTLAG(I)+1.D0
65 CONTINUE
CC

* CLEAR COUNTERS FOR NEXT SET OF SAMPLE
SIGNO=0.D0
COTPTM=0.D0
DO 20 I=1,11
  COTLAG1(I)=0.D0
20 CONTINUE
CC

* TEST END OF JOB
IF (III .GE. ISET) GOTO 70
CC

* PROCESS NEXT SAMPLE OF DATA
RETURN
CC

* SUMMARY STATISTICS FOR THE WHOLE SERIES
70 CONTINUE
CC

* COMPUTE CHI-SQUARE VALUES FOR WHOLE SERIES
CALL EQPDIS (62.,10,CHICOR,ISET,3.,CHI,III,ISET)
CALL EQPDIS (9.,10,PTCHI,ISET,3.,CHI,III,ISET)
WRITE (6,92) (CHICOR(I),I=1,1,ISET)
92 FORMAT (1X,'CHICOR=',10F10.5)
WRITE (6,93) (PTCHI(I),I=1,1,ISET)
93 FORMAT (1X,'PTCHI=',10F10.5)
CC

* PRINT STATISTICS FOR THE WHOLE SERIES
WRITE (6,80) III,TCTPTM,TCTCOR,USAMP,CHI,CH1
80 FORMAT (///1X,'*** SUMMARY STATISTICS OF ALL SAMPLES FOR ' ,
1 ' AUTOCORRELATION ANALYSIS--NO. OF SIGNIFICANT SAMPLES',
2 ' OVER', ' NO. OF TOTAL SAMPLES=',I5//' ,
3 3X,'CHI-SQUARE PORTMANTEAU TEST=',
4 ,F5.0,2X,'UNIFORM DISTRIBUTION OF NO. OF SIG. AUTO.' ,
5 ' COR. COEF',F5.0
6 /1X,'NO. OF EXPECT SAMPLES --UPPER BOUND',F7.2/
7 3X,'CHI-SQUARE VALUES BY DIST. OF SIG. LAGS',
8 F7.2,' BY PORTMANTEAU STATISTIC',F7.2/
WRITE (6,90) (TCTLAS(I),I=1,L1)
90 FORMAT (1X,'NO. OF SIGNIFICANT SAMPLES AT LAG K OVER ',
1 ' TOTAL NO.', ' OF SAMPLES/(1X,10(F10.5))
CC

* CLEAR COUNTERS
TCTPTM=0.D0
```
TCTCOR=0. DO
DO 91 I=1,L1
TCTLAG(I)=0. D0
91 CONTINUE
RETURN
END
** AUTO ROUTINE **

THIS ROUTINE WILL COMPUTE AUTO. FUNCTIONS FOR LAGS 0 TO L-1 AND RETURN THESE VALUES TO THE CALLING ROUTINE.

SUBROUTINE AUTO(A, N, L, R, COVO)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS: **
- FIND AUTOCOVARIANCES OF SERIES A FOR LAGS 0 TO L-1.
- A -- INPUT VECTOR OF LENGTH N CONTAINING THE TIME SERIES OF AUTOCOVARIANCE FUNCTION IS DESIRED.
- N -- LENGTH OF THE VECTOR A.
- L -- AUTOCOVAR. IS CALCULATED FOR LAGS OF 0, 1, 2, ..., L-1.
- R -- OUTPUT VECTOR OF LENGTH L CONTAINING AUTOCOVARIANCES OF SERIES A.
- R(K) = 1/N(SUM(X(T)*X(T+K))) FOR ALL K=0, 1, 2, ..., (L-1).
- COVO -- VARIANCE (COVARIANCE AT LAG 0).
- AVER -- MEAN OF SERIES.

IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION A(1), R(1)

* CALCULATE AVERAGE OF TIME SERIES A
COVO=0.DO
AVER=0.DO
IF (N-L) .LE. L RETURN TO MAIN PROGRAM
50 R(1)=0.DO
RETURN
100 DO 110 I=1,N
110 AVER=AVER+A(I)
FN=DFLOAT(N)
AVER=AVER/FN

* CALCULATE AUTOCOVARIANCES
DO 130 J=1,L
NJ=N-J+1
SUM=0.DO
DO 120 I=1,NJ
IJ=I+J-1
120 SUM=SUM+(A(I) -AVER) *(A(IJ) -AVER)
130 R(J)=SUM/FN
COVO=R(1)
RETURN
END
** EQPDIS ROUTINE **

** THIS ROUTINE COMPUTES BOUND VALUES OF EQUAL
PROBABILITIES, NUMBER OF OBSERVED FREQUENCIES AND
CHI-SQUARE VALUES. CHI-SQUARE VALUES WILL BE
RETURNED TO THE CALLING ROUTINE.

** REQUIRED IBM IMSL SUBROUTINES ARE MDBETI, MDNRIS,
AND MDCHI.**

SUBROUTINE EQPDIS (DF, ICAT, VALUE, NO, IDX, CHI, IIfI, ISET)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:**
- ICAT — NUMBER OF CATEGORIES.
- OBS — ARRAY FOR OBSERVED FREQUENCIES FOR CATEGORIES.
- P — PROBABILITIES.
- BOUND — ARRAY FOR BOUND VALUES FOR EQUAL PROBABILITIES.
- DINTV — INTERVAL BETWEEN TWO SUCCESSIVE PROBABILITIES.
- IDX — INDEX.
  1: FOR UNIT UNIFORM INVERSE DIST. PROB. FUNCTION.
  2: FOR NORMAL INVERSE DIST. PROB. FUNCTION.
  3: FOR CHI-SQUARE INVERSE DIST. PROB. FUNCTION.
- DF — DEGREES OF FREEDOM.
- DCHI — CHI-SQUARE VALUES FOR EACH CATEGORY.
- SCHI — ARRAY FOR STORING CHI-SQUARE VALUES FOR KL=50 SETS.
- SMCHI — ARRAY FOR STORING CHI-SQUARE VALUES FOR ISET=50
  SAMPLES.
- CTFT — NUMBERS OF SIGNIFICANT SETS PER SAMPLE.
- TCFT — NUMBERS OF SIGNIFICANT SAMPLES FOR WHOLE SERIES.

* IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION BOUND(100), OBS(100), VALUE(1), SCHI(100)
REAL*4 X, P, DINTV, DF

* DINTV = 1./ICAT - 0.00001

* CLEAR COUNTERS
DO 500 I = 1, ICAT
500 OBS(I) = 0. DO
CHI = 0. DO
TOBS = 0. DO
P = 0.

* COMPUTE BOUND VALUES FOR ICAT CATEGORIES
DO 100 I = 1, ICAT
P = DINTV + P
GOTO (1, 2, 3), IDX

* IBM IMSL INVERSE UNIFORM INVERSE DIST. PROB. ROUTINE
1 CALL MDBETI (P, 1.0, 1.0, X, IER)
GOTO 90

* IBM IMSL INVERSE NORMAL DIST. PROB. ROUTINE
2 CALL MDNRIS (P, X, IER)
GOTO 90

CC  * IBM IMSL INVERSE CHI-SQUARE DISTR. PROB. ROUTINE
3 CALL MDCHI(P,DF,X,IER)
90 BOUND(I) = X
100 CONTINUE

CC  * COUNT OBSERVED FREQUENCIES
DO 200 I=1,NO
DO 210 J=1,ICAT
IF (VALUE(I) .LE. BOUND(J)) GOTO 205
210 CONTINUE
OBS(ICAT) =OBS(ICAT) +1.DO
GOTO 200
205 OBS (J)=OBS (J) +1.DO
GOTO 200

CC  * NEXT OBSERVED VALUE
200 CONTINUE

CC  * COMPUTE CHI-SQUARE VALUES
EXP=(1.DO/DFLOAT(ICAT)) * DFLOAT(NO)
DO 300 I=1,ICAT
DCHI (I)= (OBS (I)-EXP)**2/EXP
CHI=CHI+DCHI (I)
300 CONTINUE
IF (III .LT. ISET .OR. DF .EQ. 0.) GOTO 402

CC  * PRINT BOUND VALUES, NUMBERS OF OBSERVED FREQ.
CC  AND CHI-SQUARE VALUES OF EACH CATEGORY FOR
CC  LAST SAMPLE
WRITE (6,10) (BOUND (I),I=1,ICAT)
10 FORMAT (10X,'BOUND VALUES=',10F10.5)
WRITE (6,11) (OBS(I),I=1,ICAT)
11 FORMAT (10X,'NO. OF OBS=',10F10.5)
WRITE (6,12) (DCHI(I),I=1,ICAT)
12 FORMAT (10X,'DCHI=',10F10.5)
402 CONTINUE
RETURN
END
** ** THIS ROUTINE WILL BE CALLED BY RUNAB AND RUNUD
FOR COMPUTING TOTAL NO. OF RUNS, NO. OF RUNS
BY LENGTH AND CHI-SQUARES OF NO. OF RUNS BY
LENGTH FOR EACH SET. THESE COMPUTED VALUES
WILL BE RETURNED TO THE CALLING ROUTINE.

SUBROUTINE SRUNLT (ICOT,N,EXPTL,EXPL,NOLTH,TOL,CS,
1 IIII,ISET)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:
TOL--TOTAL NUMBER OF RUNS.
RLNTH--NUMBER OF RUNS BY LENGTH.
RUNLTH--ARRAY FOR NO. OF RUNS BY LENGTH.
NOLTH--NUMBER OF LENGTHS.
EXPL--EXPECTED NO. OF RUNS BY LENGTH.
TRUN--RATIO OF TOTAL NO. OF RUNS OVER TOTAL NO. OF
EXPECTED RUNS.
AEXPL--ADJUSTED EXPECTED NO. OF RUNS BY LENGTH.
CS--CHI-SQ. VALUE BY TESTING NO. OF RUNS BY LENGTH.

IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION ICOT (1),RUNLTH(20),PRUN (20),AEXPL(20)

1 starred variables

* CLEAR COUNTERS FOR NEXT SET OF DATA
DO 500 I=1,NOLTH
RUNLTH (I)=0.DO
500 CONTINUE
TRUN=0.DO
CS=0.DO
TOL=0.DO
RLNTH=1.DO

* COMPUTE NO. OF TOTAL RUNS AND NO. OF RUNS BY LENGTH
DO 100 I=1,N
IF (ICOT(I) .EQ. ICOT(I+1)) GOTO 110
TOL=TOL+1.DO
DO 200 J=1,NOLTH
IF (RLNTH .NE. J) GOTO 200
RUNLTH(J)=RUNLTH(J)+1.DO
GOTO 120
200 CONTINUE
RUNLTH(NOLTH)=RUNLTH(NOLTH)+1.DO
120 CONTINUE

RUNLTH=0.DO
110 RLNTH=RLNTH+1.DO
100 CONTINUE
DO 400 I=1,NOLTH
TRUN=TRUN+RUNLTH(I)
400 CONTINUE
TRUN=TRUN/EXPTL
DO 300 I=1,NOLTH
AEXPL(I) = EXPL(I) * RUN

CONTINUE

* CALCULATE CHI-SQUARE VALUE

DO 410 I = 1, NOLTH

DD(I) = ((RUNLTH(I) - AEXPL(I)) ** 2) / AEXPL(I)

CS = CS + DD(I)

CONTINUE

IF (IIIII .EQ. ISET) GOTO 600

RETURN

CONTINUE

* WRITE NO. OF RUNS AND TOTAL NO. OF RUNS FOR L. SAMP.

IF (NOLTH .EQ. 5) GOTO 602

WRITE (6, 603) TOL, EXPTL, (RUNLTH(I), I = 1, NOLTH),
(AEXPL(I), I = 1, NOLTH), (DD(I), I = 1, NOLTH)

RETURN

600 CONTINUE

601 FORMAT (1X, 'AT=', F6.0, ' EXTL=', F6.0, ' LTH=', F6.1,
AEXPL=', F7F6.1 /1X,7F7.2)

RETURN

602 WRITE (6, 603) TOL, EXPTL, (RUNLTH(I), I = 1, NOLTH),
(AEXPL(I), I = 1, NOLTH), (DD(I), I = 1, NOLTH)

RETURN

603 FORMAT (1X, 'UT=', F6.0, ' EXTL=', F6.0, ' LTH=', F5F6.1,
AEXPL=', F5F6.1 , 5F7.2)

RETURN

END
*** RMEAN ROUTINE ***

** THIS ROUTINE COMPUTES MEAN, MEAN SQUARE AND
STANDARD DEVIATION.

SUBROUTINE RMEAN (X, N, SUM, SUM2, SS, SSE, TPS)

** KEY VARIABLES IN THIS ROUTINE ARE AS FOLLOWS:
SUM—SAMPLE MEAN.
SSE—MEAN SQUARE ERROR.
SS—SAMPLE VARIANCE.
SUM2—SAMPLE MEAN SQUARE.

IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION X(1)

* CLEAR COUNTERS
SSE=0.DO
SUM=0.DO
SUM2=0.DO
SS=0.DO

* COMPUTE SUM, SUM OF SQUARE ERROR AND SUM OF SQUARE
DO 10 I=1,N
SUM=X(I) +SUM
SSE=SSE+(X(I)-TPS)**2
10 SUM2=X(I)**2 +SUM2

* COMPUTE MEAN AND MEAN SQUARE
SUM=SUM/DFLOAT(N)
SUM2=SUM2/DFLOAT(N)

* COMPUTE SUM OF SQUARE ERROR
DO 20 I=1,N
SS=(X(I)-SUM)**2 +SS
20 CONTINUE

* COMPUTE SAMPLE VARIANCE AND VARIANCE
SS=SS/DFLOAT(N-1)
SSE=SSE/DFLOAT(N)
RETURN
END
*** FFT ROUTINE ***

** This routine transforms autocovariance functions over time domain into spectra over frequency domain and returns spectra to the calling routine (Fast Fourier Transform).

SUBROUTINE FFT(FR,FI,K)

* Fast Fourier transform using time decomposition with input bit reversal.
Data is in FR (real) and FI (imaginary) arrays.
Computation is in place, output replaces input.
Number of points must be N = 2**K.
FR(N) and FI(N) must be dimensioned in main program.

IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION FR(1),FI(1)
N=2**K
MR=0
NN=N-1
DO 2 M=1,NN
L=N
1 L=L/2
IF (MR+L.GT.NN) GOTO 1
MR=MOD(MR,L)+L
IF (MR.LE.M) GOTO 2
TR=FR(M+1)
FR(M+1)=FR(MR+1)
FR(MR+1)=TR
TI=FI(M+1)
FI(M+1)=FI(MR+1)
FI(MR+1)=TI
2 CONTINUE
L=1
3 IF (L.GE.N) RETURN
ISTEP=2*L
EL=DFLOAT(L)
DO 4 M=1,L
A=3.1415926535D0 *DFLOAT(1-M)/EL
WR=DCOS(A)
WI=DSIN(A)
DO 4 I=M,N,ISTEP
J=I+L
TR=WR*FR(J)-WI*FI(J)
TI=WR*FI(J)+WI*FR(J)
FR(J)=FR(I)-TR
FI(J)=FI(I)-TI
FR(I)=FR(I)+TR
4 FI(I)=FI(I)+TI
L=ISTEP
GOTO 3
END
VITA

Win-yeu Winnie Chen was born in February 3, 1941 in Taipei, Taiwan. She received her Bachelor degree in Accounting and Banking from the Department of Business Administration of National Taiwan University. She received three Master degrees in the United States. These are one Master of Business Administration from University of Cincinnati in August, 1973, one Master of System Science from the Department of Computer Science of Louisiana State University in August, 1979, and the other one Master of Science from the Department of Quantitative Methods of Louisiana State University in May, 1980. Presently, she is a candidate from the Doctor of Philosophy degree from the Department of Quantitative Methods of Louisiana State University.
EXAMINATION AND THESIS REPORT

Candidate: Win-Yeu Winnie Chen

Major Field: Quantitative Methods

Title of Thesis: Quality Evaluation of Some Combinations of Unit Uniform Random Number Generators and Unit Normal Transformation Algorithms

Approved:

[Signatures of Major Professor and Chairman, Dean of the Graduate School]

EXAMINING COMMITTEE:

[Signatures of committee members]

Date of Examination: November 12, 1980