NON-CONGRUENCE OF STATISTICAL DISTRIBUTIONS:
HOW DIFFERENT IS DIFFERENT?

by

TERRY LEE APPLEGATE

B.S., Kansas State University, 1977

A MASTER'S THESIS

Submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering
KANSAS STATE UNIVERSITY
Manhattan, Kansas

1978

Approved by:

[Signature]
Major Professor
ACKNOWLEDGEMENTS

I would like to thank my family, the faculty and personnel of the Department of Industrial Engineering at Kansas State, and my fellow students. Each of these groups helped make this segment of my education meaningful. Special thanks is given to Dr. Doris Grosh, whose faith and guidance were instrumental in my completing my course of studies and to my wife, Debora, without whose encouragement and understanding nothing would have been worthwhile.
# TABLE OF CONTENTS

## CHAPTER

<table>
<thead>
<tr>
<th>I.</th>
<th>INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0. Genesis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.1. The Problem</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2. Purpose and Objective</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.3. Method</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.3.1 The Index of Non-Congruity - $\delta$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.3.2. The Procedure</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1.3.3. The McGill Random Number Generator</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II.</th>
<th>THE EXPONENTIAL CASE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.0. Introduction</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>2.1. Determination of the Point of Intersection</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>2.2. The Index of Non-Congruity: Exponential Case</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2.3. Description of Exponential Program Features</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2.3.1. Program Listing</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2.3.2. Definition of Program Variables</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2.3.3. Inputs to the Program</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.3.4. Outputs of the Program</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2.3.5. Special Programming Considerations</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2.4. Results of the Exponential Program</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>III.</th>
<th>THE NORMAL CASE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.0. Introduction</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>3.1. Determination of the Point or Points of Intersection</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>3.2. The Index of Non-Congruity: Normal Case</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>3.3. Description of Normal Program Features</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>3.3.1. Program Listing</td>
<td>34</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>PAGE</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>3.3.2. Definition of Program Variables</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>3.3.3. Inputs to the Program</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>3.3.4. Outputs of the Program</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>3.3.5. Special Programming Considerations</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>3.4. Results of the Normal Program</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>IV. CONCLUSION</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>SELECTED BIBLIOGRAPHY</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>APPENDIX 1</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>APPENDIX 2</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1.</td>
<td>Illustration of the Index of Non-Congruity</td>
<td>4</td>
</tr>
<tr>
<td>1-2.</td>
<td>Distribution divided into Equi-probability Regions</td>
<td>6</td>
</tr>
<tr>
<td>2-1.</td>
<td>The Index of Non-Congruity for Exponential Distributions</td>
<td>11</td>
</tr>
<tr>
<td>2-2.</td>
<td>$\delta$ vs. $\lambda_2/\lambda_1$</td>
<td>19</td>
</tr>
<tr>
<td>2-3.</td>
<td>$\chi^2_{PS}$ vs. $\lambda_2/\lambda_1$</td>
<td>20</td>
</tr>
<tr>
<td>2-4.</td>
<td>$\chi^2_{PS}$ vs. $\delta$</td>
<td>21</td>
</tr>
<tr>
<td>2-5.</td>
<td>$\chi^2_R$ vs. $\chi^2_{PS}$</td>
<td>22</td>
</tr>
<tr>
<td>2-6.</td>
<td>$\hat{\alpha}$ vs. $\lambda_2/\lambda_1$</td>
<td>23</td>
</tr>
<tr>
<td>2-7.</td>
<td>$\hat{\alpha}$ vs. $\delta$</td>
<td>24</td>
</tr>
<tr>
<td>2-8.</td>
<td>Possible Model of the Relationship of $\hat{\alpha}$ and $\delta$</td>
<td>26</td>
</tr>
<tr>
<td>3-1.</td>
<td>The Normal Case with Two Points of Intersection</td>
<td>32</td>
</tr>
<tr>
<td>3-2.</td>
<td>The Normal Case for Equal Standard Deviations</td>
<td>33</td>
</tr>
<tr>
<td>3-3.</td>
<td>Tree Diagram for Sorting Procedure</td>
<td>39</td>
</tr>
<tr>
<td>3-4.</td>
<td>$\chi^2_{PS}$ vs. $\delta$ - Single Variation</td>
<td>42</td>
</tr>
<tr>
<td>3-5.</td>
<td>$\chi^2_{PS}$ vs. $\delta$ - Dual Variation</td>
<td>43</td>
</tr>
<tr>
<td>3-6.</td>
<td>$\delta$ vs. $\eta_1$ - Single Variation</td>
<td>44</td>
</tr>
<tr>
<td>3-7.</td>
<td>$\delta$ vs. $\eta_1$ - Dual Variation</td>
<td>45</td>
</tr>
<tr>
<td>3-8.</td>
<td>$\chi^2_{PS}$ vs. $\eta_1$ - Single Variation</td>
<td>46</td>
</tr>
<tr>
<td>3-9.</td>
<td>$\chi^2_{PS}$ vs. $\eta_1$ - Dual Variation</td>
<td>47</td>
</tr>
<tr>
<td>3-10.</td>
<td>$\chi^2_R$ vs. $\chi^2_{PS}$ - Mean-Variate Data</td>
<td>48</td>
</tr>
<tr>
<td>3-11.</td>
<td>$\chi^2_R$ vs. $\chi^2_{PS}$ - Variance-Variate Data $\sigma_1&lt;\sigma_2$</td>
<td>49</td>
</tr>
<tr>
<td>3-12.</td>
<td>$\chi^2_R$ vs. $\chi^2_{PS}$ - Variance-Variate Data $\sigma_1&gt;\sigma_2$</td>
<td>50</td>
</tr>
<tr>
<td>3-13.</td>
<td>$\chi^2_R$ vs. $\chi^2_{PS}$ - Dual Variation</td>
<td>51</td>
</tr>
<tr>
<td>3-14.</td>
<td>$\hat{\alpha}$ vs. $\delta$ - Mean-Variate Data</td>
<td>52</td>
</tr>
<tr>
<td>3-15.</td>
<td>$\hat{\alpha}$ vs. $\delta$ - Variance-Variate Data $\sigma_1&lt;\sigma_2$</td>
<td>53</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3-16.</td>
<td>$\hat{\alpha}$ vs. $\delta$ - Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>54</td>
</tr>
<tr>
<td>3-17.</td>
<td>$\hat{\alpha}$ vs. $\delta$ - Dual Variation</td>
<td>55</td>
</tr>
<tr>
<td>3-18.</td>
<td>$\hat{\alpha}$ vs. $\eta_1$ - Mean-Variate Data</td>
<td>56</td>
</tr>
<tr>
<td>3-19.</td>
<td>$\hat{\alpha}$ vs. $\eta_1$ - Variance-Variate Data $\sigma_1 &lt; \sigma_2$</td>
<td>57</td>
</tr>
<tr>
<td>3-20.</td>
<td>$\hat{\alpha}$ vs. $\eta_1$ - Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>58</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1.</td>
<td>Results of the Exponential Program</td>
</tr>
<tr>
<td>3-1.</td>
<td>Results of the Normal Program</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

1.0 GENESIS

This study had its origin in an earlier study for the United States Nuclear Regulatory Commission concerned with diesel engine failure data obtained from several nuclear power plants throughout the United States (3). A result of this earlier study was a number of questions concerning differences between statistical distributions with the primary question being "How does one determine statistically significant differences between statistical distributions?". A subsequent study by K. Lakshminarayan (4) applied the methodology used in this study to Beta distributions.

1.1 THE PROBLEM

Statistical significance is usually concerned with comparing different sets of sample data or with making inferences about the populations being sampled. Various parametric and non-parametric techniques therefore exist for making these decisions. However, no such techniques exist for comparing the populations themselves. The problem to be investigated is: given a family of statistical distributions, how much may a pair of distributions from this family differ before they can be detected as being significantly different, or "How different is different?".

1.2 PURPOSE AND OBJECTIVE

There are primarily two reasons for studying differences between similar distributions of the same family. The first reason deals with the theoretical insights which can result from studying the effects that perturbations of distribution parameters have on the "sameness" of family members. With better understanding of the role of distribution parameters and their
relative importance in determining the characteristics of a particular distribution, hopefully more powerful estimating and comparative statistical techniques can be developed. The second and probably more important reason is the practical applications which could result from studying differences between statistical distributions. Applications could include new methods for establishing when sample data from similar sources could be pooled, parametric "goodness of fit" tests, and sample-free hypothesis testing.

With these two broad underlying reasons for studying differences in statistical distributions from the same family, the expressed objective of this study is: to develop a method of comparing differences in statistical distributions from the same family (normal distributions and exponential distributions are the families of statistical distributions studied), to use this technique to examine the effects of varying the parameters of the distributions on their "sameness", and to attempt to draw some conclusions pertaining to the usefulness of this technique in answering the question of "How different is different?".

1.3 METHOD

1.3.1. The Index of Non-Congruity - $\delta$

The following discussion is an adaptation of material presented by Lakshminarayan (3).

Theoretically, two continuous distributions are the same only if their probability density functions are identical and for their probability density functions to be identical the two distributions must have exactly the same parameters. In practical situations however, two distributions whose probability density functions (and therefore parameters) are nearly the same, may produce random samples which are indistinguishable from one
another. It is this type of situation which indicates that merely examining the probability density functions (or the parameters) of two distributions to see if they are identical does not provide enough information to judge if the two distributions are similar enough to consider them practically as being the same, or if they are different enough that they must be considered as different.

One measure that determines differences between continuous distributions is the difference in the areas bounded by each probability density function in the region of interest or the amount of non-overlapping area bounded by the curves. If \( f_1(x) \) and \( f_2(x) \) are the probability density functions (Figure 1-1) of the two distributions being compared, then the amount of non-overlapping area or what we have termed "the index of non-congruity" is given by:

\[
\delta = \int_{-\infty}^{\infty} |f_1(x) - f_2(x)| \, dx
\]

The non-overlapping area is shown by the shaded portion of Figure 1-1 and the total amount of this shaded area equals \( \delta \). To qualify as probability density functions, \( f_1(x) \) and \( f_2(x) \) each must satisfy the criterion:

\[
\int_{-\infty}^{\infty} f(x) \, dx = 1
\]

Therefore the values that can be assumed by the index of non-congruity are \( 0 \leq \delta \leq 2 \). If two distributions are approximately the same then \( \delta \) will be close to zero and if two distributions are radically different the value of \( \delta \) will approach 2.

1.3.2. The Procedure

The general procedure used in this study is to choose a particular
FIGURE 1-1 Illustration of the Index of Non-Congruity
distribution from a family of distributions and then compare it to a similar distribution from the same family. The first distribution will be termed the "model distribution" (Distribution 1) and the similar distribution will be termed the "alternative distribution" (Distribution 2).

The procedure by which the alternative distribution is compared with the model distribution consists of a number of steps. The first step is to calculate the index of non-congruity between the two distributions as explained in Section 1.3.1.

Secondly, the model distribution is divided into ten equi-probability regions. A set of values \{x(i)\} of the independent variable is calculated such that:

\[ \int_{\infty}^{x(i)} f_1(x) dx = \frac{i}{10}, \quad i = 1, \ldots, 10 \]  \hspace{1cm} (3)

The values \{x(i)\} are such that the sample space of the independent variable is divided into regions which have the same area under the curve of the probability density function, as shown in Figure 1-2. After the equi-probability regions for the model distribution have been determined, a "perfect" sample is drawn from the alternative distribution by using the \{x(i)\} from the model distribution as interval boundaries of the alternative distribution. A "pseudo" - \( \chi^2 \) statistic is then calculated from this "perfect" sample. This pseudo-\( \chi^2 \) statistic, \( \chi^2_{ps} \), is:

\[ \chi^2_{ps} = \sum_{i=1}^{10} \left( \frac{[F_2(i)-F_2(i-1)]M - .1(M))^2}{.1(M)} \right) \]  \hspace{1cm} (4)

where

\[ F_2(i) = \int_{\infty}^{x(i)} f_2(x) dx, \]  \hspace{1cm} (5)

\M\ is the sample size and \( F_2(0) = 0 \). We are interested in small sample sizes because small sample sizes are usually encountered in practical
FIGURE 1-2  Distribution divided into Equi-probability Regions
situations, and with very large sample sizes even small differences are discernable. Therefore a value of 50 is used for \( M \) in this study. The "pseudo" \( \chi^2 \) and "perfect" sample are used to reduce the effects introduced by random fluctuations and to better ascertain the basic relationship between the index of non-congruity and differences between the model and alternative distributions.

The third step in the procedure comparing the alternative distribution to the model distribution is that a genuine random sample of size \( M \) is drawn from the alternative distribution and the sample is compared to the model distribution. The comparison is made by an ordinary \( \chi^2 \) goodness of fit test. The random \( \chi^2 \) statistic is calculated to provide a check on the \( \chi^2_{PS} \) results and to provide additional insight into the question of differences between statistical distributions. Once the random \( \chi^2 \), denoted by \( \chi^2_R \), has been determined, the level of significance \( \alpha \) is calculated. Usually when a \( \chi^2 \) statistic is calculated it is then compared to a value obtained from an appropriately chosen \( \chi^2 \) distribution to determine significance at a prescribed confidence level. In this situation this technique is not totally satisfactory since we are not only interested in whether a particular \( \chi^2_R \) value is significant but also in how significant it is. The level of significance \( \alpha \) is the area to the right of the computed (observed) \( \chi^2_R \) of a \( \chi^2 \) distribution with 9 degrees of freedom. There are 9 degrees of freedom since the model distribution is divided into 10 equi-probability regions and the random observations are sorted into these regions for ease of computation. The level of significance gives a more intuitively comprehensible measure of difference than the \( \chi^2_{PS} \) and \( \chi^2_R \) values.

The final step in comparing the alternative distribution with the model distribution is the calculation of parametric indicators which
attempt to quantify the differences between the model distribution and the alternative distribution. It is hoped that a relationship can be discovered between a parametric indicator and the index of non-congruity. Using this relationship combined with knowledge about the relationship between the index of non-congruity and statistical significance it may be possible to find a measure of statistical difference between distributions based solely on the parameters of the distributions. Such a parametric indicator would be of considerable practical importance because it would be easy to calculate.

In summary, the comparison procedure given a model distribution and an alternative distribution is:

1) Calculate $\delta$, the index of non-congruity
2) Calculate $x^2_{PS}$, the "pseudo" chi-square statistic from a "perfect" sample
3) Calculate $x^2_R$ from a random sample from the alternative distribution
4) Calculate $\alpha$, the level of significance for $x^2_R$
5) Calculate various parametric indicators

1.3.3 The McGill Random Number Generator

This study is primarily based on the use of a computer to perform the comparison procedure outlined in Section 1.3.2. One of the major problems in the development of a program to perform this procedure is the generation of a random sample from the alternative distribution to be used in calculating $x^2_R$. The McGill Random Number Generator developed by members of the School of Computer Science of McGill University seemed particularly well suited for the requirements of this study. The McGill RNG has several features which led to its selection. First, the use of the McGill RNG
is FORTRAN compatible and the rest of the program will be written in FORTRAN. Second, the McGill RNG is called as a FORTRAN function rather than as a subroutine, which is advantageous in terms of computation time. Third, the previous value returned is maintained internally by the McGill RNG which leads to easier programming. Fourth, the McGill RNG has special procedures for generating random samples from normal and exponential distributions, thereby eliminating the need to program a transformation for converting a uniform distribution to either of these distributions. Finally, the McGill RNG is included in the subroutine library of the Kansas State University computing system, eliminating the need to include an additional subprogram for the random number generator in the index of non-congruity program.
2.0 INTRODUCTION

This chapter describes the specific procedure for calculating the index of non-congruity for exponential distributions, the program developed to perform the comparison procedure outlined in Section 1.3.2. for exponential distributions, and the results of using this procedure to compare several pairs of exponential distributions.

2.1 DETERMINATION OF THE POINT OF INTERSECTION

Let \( f_1(x) \) be the probability density function of an exponential distribution with parameter \( \lambda_1 \) and let \( f_2(x) \) be the probability density function of an exponential distribution with parameter \( \lambda_2 \). Assume that \( \lambda_2 > \lambda_1 \). Consider Figure 2-1 which shows two exponential distributions fulfilling these requirements. The shaded area represents the index of non-congruity for this pair of distributions. This area difference is given by Equation (1) which is repeated again for clarity.

\[
\delta = \int_{-\infty}^{\infty} |f_1(x) - f_2(x)| \, dx
\]

(1)

To facilitate calculation of \( \delta \), this integral can be divided into 2 components.

\[
\delta = \int_{-\infty}^{\infty} |f_1(x) - f_2(x)| \, dx = \int_0^{X_A} [f_2(x) - f_1(x)] \, dx + \int_{X_A}^{\infty} [f_1(x) - f_2(x)] \, dx
\]

(6)

\( X_A \) is the point of intersection of the two probability density functions. Finally note that this expression applies for the case where \( \lambda_2 > \lambda_1 \) and for the case where \( \lambda_1 > \lambda_2 \) the limits of integration for the component integrals would have to be exchanged to insure the proper sign for \( \delta \).
FIGURE 2-1 The Index of Non-Congruity for Exponential Distributions
The following calculation demonstrates the determination of the point of intersection $X_A$. At $X = X_A$

$$
\lambda_1 e^{-\lambda_1 X_A} = \lambda_2 e^{-\lambda_2 X_A} \tag{7}
$$

$$
e^{(\lambda_2 - \lambda_1) X_A} = \frac{\lambda_2}{\lambda_1} \tag{8}
$$

$$
(\lambda_2 - \lambda_1) X_A = \ln \frac{\lambda_2}{\lambda_1} \tag{9}
$$

$$
X_A = \left[\frac{1}{\lambda_2 - \lambda_1}\right] \ln \frac{\lambda_2}{\lambda_1} \tag{10}
$$

### 2.2 The Index of Non-Congruity: Exponential Case

Having determined the point of intersection $X_A$, the expression for the index of non-congruity can be modified. Substituting into Equation (6) we obtain:

$$
\delta = \int_0^{X_A} [\lambda_2 e^{-\lambda_2 X} - \lambda_1 e^{-\lambda_1 X}] \, dx + \int_{X_A}^\infty [\lambda_1 e^{-\lambda_1 X} - \lambda_2 e^{-\lambda_2 X}] \, dx \tag{11}
$$

$$
= \int_0^{X_A} \lambda_2 e^{-\lambda_2 X} \, dx - \int_0^{X_A} \lambda_1 e^{-\lambda_1 X} \, dx \tag{12}
$$

$$
+ \int_{X_A}^\infty \lambda_1 e^{-\lambda_1 X} \, dx - \int_{X_A}^\infty \lambda_2 e^{-\lambda_2 X} \, dx \tag{13}
$$

$$
= -e^{-\lambda_2 X} |_0^{X_A} - e^{-\lambda_1 X} |_0^{X_A}
$$

$$
+ -e^{-\lambda_1 X} |_{X_A}^\infty - e^{-\lambda_2 X} |_{X_A}^\infty \tag{14}
$$

$$
= [-e^{-\lambda_2 X_A} + 1] - [-e^{-\lambda_1 X_A} + 1]
$$

$$
+ [0 + e^{-\lambda_1 X_A}] - [0 + e^{-\lambda_2 X_A}] \tag{15}
$$

$$
= 2e^{-\lambda_1 X_A} - 2e^{-\lambda_2 X_A}
$$

Therefore

$$
\delta = 2[e^{-\lambda_1 X_A} - e^{-\lambda_2 X_A}] \tag{16}
$$

This development is for the case where $\lambda_2 > \lambda_1$. For the case where
\( \lambda_1 > \lambda_2 \) the index of non-congruity is:

\[ \delta = 2[e^{-\lambda_2 x_A} - e^{-\lambda_1 x_A}] \tag{15a} \]

and:

\[ x_A = \frac{1}{\lambda_1 - \lambda_2} \ln \frac{\lambda_1}{\lambda_2} \tag{10a} \]

2.3 DESCRIPTION OF EXPONENTIAL PROGRAM FEATURES

2.3.1 Program Listing

A complete listing of the computer program to perform the comparison procedure for exponential distributions is given in Appendix 1.

2.3.2. Definition of Program Variables

AHAT: the level of significance, \( \alpha \)

CADTR: function subroutine to determine \( \alpha \)

DIFFL: absolute difference of \( \lambda_1 \) and \( \lambda_2 \), \(|\lambda_1 - \lambda_2|\)

DELTA: the index of non-congruity, \( \delta \)

EI: the expected frequency in the equi-probability regions, M/10

FREQ(10): the array containing the frequency counts of the random sample sorted into the equi-probability regions

F2SUM: the sum of the components of the array FREQ squared

ISEED: one of the seeds for the McGill RNG

JSEED: the other seed for the McGill RNG

K: the index for the array FREQ. K can take on integer values from 1 to 10.

LAMDA1: the parameter of the model exponential distribution

LAMDA2: the parameter of the alternative exponential distribution

M: the sample size

NU: the degrees of freedom for the \( \chi^2 \) distribution which \( \chi^2_{PS} \) and \( \chi^2_R \) are compared with
P2(10): the array containing the cumulative probability of the alternative distribution at the region boundaries \{x(i)\}

RATIO: the ratio of \( \lambda_1 \) and \( \lambda_2 \), if \( \lambda_2 > \lambda_1 \)

\[
RATIO = \frac{\lambda_2}{\lambda_1} \quad \text{and if } \quad \lambda_1 > \lambda_2 \quad RATIO = \frac{\lambda_1}{\lambda_2}
\]

REXP: subroutine to generate a random deviate from an exponential distribution with mean \( \lambda = 1 \)

RSTART: subroutine to initialize the McGill Random Number Generator

SAMPL(200): the array containing the random sample from the alternative distribution

T1: the point of intersection \( x_A \)

X1(10): the equi-probability region boundaries \{x(i)\}

X2ACT: the random \( \chi^2 \) statistic, \( \chi^2_R \)

X2PS: the pseudo \( \chi^2 \) statistic, \( \chi^2_{PS} \)

X2SUM: the sum of \([P2(10) - X1(10)]^2\) used in the calculation of \( X2PS \)

Z1: variable equal to the negative of the product of \( \Lambda MDA1 \) and \( T1 \)

Z2: variable equal to the negative of the product of \( \Lambda MDA2 \) and \( T1 \)

2.3.3. Inputs to the Program

The variables required as input to the program are:

\( \Lambda MDA1, \Lambda MDA2, M, ISEED, JSEED \)

The input is given on two separate cards with the indicated format.

\( \Lambda MDA1, \Lambda MDA2, M \) \quad 1 card \quad (2E10.4, I5)

\( ISEED, JSEED \) \quad 1 card \quad (2I5)

Multiple runs of the program can be made by supplying additional input cards (two per replication) containing the information described above. Program completion is indicated by a blank card.
2.3.4. Outputs of the Program

The program provides two types of output. The first type is an echo check of the input. The second type is information calculated by the program. The following information is produced as output of the second type: the array X1, the array P2, the first M components of the array SAMPL, the array FREQ, DELTA, X2PS, X2ACT, and AHAT. A sample output is shown in Appendix 1.

2.3.5. Special Programming Considerations

Using the McGill Random Number Generator The use of the McGill RNG is accomplished by the two subroutines RSTART and REXP. RSTART initializes the RNG. The arguments of RSTART are ISEED and JSEED. The transfer from the main program to the RSTART subroutine is made by the statement Call RSTART (ISEED, JSEED). The RNG provides default values if RSTART is not used. REXP generates a random exponential deviate from an exponential population with parameter \( \lambda = 1 \). This random deviate is transformed into a random deviate from an exponential population with parameter \( \lambda_A \) by dividing by \( \lambda_A \) e.g. \( z = x / \lambda_A \) where \( z \) is the random deviate from the desired distribution and \( x \) is the generated random deviate. The argument of REXP is a dummy integer constant which is ignored by the program. In other words use of REXP(10) or REXP(98765) produces the same effect i.e. the generation of an exponential random deviate. Use of the function subroutine REXP is accomplished by using REXP(1) in an arithmetic function e.g. \( \text{SAMPL}(I) = \text{REXP}(1) / \text{LAMDA2} \).

Sorting the Random Sample Observations In designing a sorting procedure the objective is to minimize the expected number of sorting trials for a sample set while maintaining a level of simplicity in the programming. The sorting procedure used in the program consists of a loop containing a
set of test statements which compares the random deviate with the equal probability region boundaries sequentially until the deviate is less than the boundary value. The deviate is then placed in the frequency region which has the boundary value as its upper bound. This sorting procedure would minimize the expected number of trials for the case where \( \lambda_2 > \lambda_1 \). However for the case where \( \lambda_1 \gg \lambda_2 \) this procedure would result in a high expected number of trials. This was not considered a significant problem since we are primarily concerned with cases where \( \lambda_1 \) and \( \lambda_2 \) are nearly equal.

Calculating \( \hat{\alpha} \) The level of significance, \( \hat{\alpha} \), is calculated by the function subroutine CADTR which is a slightly modified version of the CDTR subroutine contained in IBM's Scientific Subroutine Package. The modifications include changing the subprogram from a subroutine subprogram to a function subprogram and modifying the inputs and outputs of the subprogram.

2.4. RESULTS OF THE EXPONENTIAL PROGRAM

Four different sets of values of random number generator seeds were used in investigating the exponential case. Values of distributions compared varied from \( \lambda_2/\lambda_1 = 1/3 \) to \( \lambda_2/\lambda_1 = 3 \). For \( \lambda_1 > \lambda_2 \) the value of \( \lambda_2 \) was set to equal 10 and for \( \lambda_2 > \lambda_1 \) the value of \( \lambda_1 \) was set equal to 10. Therefore in every pair of distributions compared the smallest parameter was equal to 10. This was done for computation convenience since only the ratio is pertinent (as seen from Equations (10) and (16)), rather than the absolute size of \( \lambda_1 \) and \( \lambda_2 \). The sample size used in the comparison was set equal to 50 in all cases. The results of the various comparison runs are summarized in Table 2-1.

Various relationships between comparison indices are graphically presented in Figures 2-2 to 2-7. Examination of these figures indicates that there is good reason to believe that there is a strong relationship
## TABLE 2-1

Results of the Exponential Program

<table>
<thead>
<tr>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>RATIO</th>
<th>RNG Seed</th>
<th>$\delta$</th>
<th>$\chi^2_{PS}$</th>
<th>$\chi^2_{R}$</th>
<th>$\hat{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>10</td>
<td>1/3</td>
<td>A</td>
<td>.7698</td>
<td>76.80</td>
<td>86.00</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>84.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>120.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>75.20</td>
<td>.0000</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>1/2</td>
<td>A</td>
<td>.50</td>
<td>28.67</td>
<td>24.40</td>
<td>.0037</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>39.20</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>48.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>41.60</td>
<td>.0000</td>
</tr>
<tr>
<td>17.5</td>
<td>10</td>
<td>4/7</td>
<td>A</td>
<td>.4064</td>
<td>17.93</td>
<td>20.00</td>
<td>.0179</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>46.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>37.20</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>36.00</td>
<td>.0000</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>2/3</td>
<td>A</td>
<td>.2963</td>
<td>8.88</td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>28.80</td>
<td>.0007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>21.60</td>
<td>.0102</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>18.40</td>
<td>.0508</td>
</tr>
<tr>
<td>12.5</td>
<td>10</td>
<td>4/5</td>
<td>A</td>
<td>.1638</td>
<td>2.48</td>
<td>4.80</td>
<td>.8514</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>14.40</td>
<td>.1088</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>10.40</td>
<td>.3191</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>14.00</td>
<td>.1223</td>
</tr>
<tr>
<td>10</td>
<td>12.5</td>
<td>5/4</td>
<td>A</td>
<td>.1638</td>
<td>2.00</td>
<td>10.40</td>
<td>.3191</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>12.40</td>
<td>.1917</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>10.00</td>
<td>.3505</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>3/2</td>
<td>A</td>
<td>.2963</td>
<td>6.13</td>
<td>10.40</td>
<td>.3191</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>13.60</td>
<td>.1373</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>10.80</td>
<td>.2897</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>12.80</td>
<td>.1719</td>
</tr>
<tr>
<td>10</td>
<td>17.5</td>
<td>7/4</td>
<td>A</td>
<td>.4064</td>
<td>11.12</td>
<td>17.60</td>
<td>.0401</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>20.00</td>
<td>.0179</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>18.80</td>
<td>.0269</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>18.80</td>
<td>.0269</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>2/1</td>
<td>A</td>
<td>.50</td>
<td>16.50</td>
<td>19.60</td>
<td>.0205</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>23.20</td>
<td>.0058</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>18.80</td>
<td>.0269</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>26.40</td>
<td>.0018</td>
</tr>
</tbody>
</table>
Table 2-1 continued

<table>
<thead>
<tr>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>RATIO</th>
<th>RNG SEED</th>
<th>$\hat{\delta}$</th>
<th>$\chi^2_{PS}$</th>
<th>$\chi^2_{R}$</th>
<th>$\hat{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>30</td>
<td>3/1</td>
<td>A</td>
<td>.7698</td>
<td>39.50</td>
<td>48.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>40.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>43.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>47.60</td>
<td>.0000</td>
</tr>
</tbody>
</table>

$M = 50$, RNG = A(51562, 62155) B(62155, 51562) C(50020, 11292) D(11292, 50020)
**Figure 2-2** \( \delta \) vs. \( \lambda_2 / \lambda_1 \)
FIGURE 2-3 $x_{PS}^2$ vs. $\lambda_2/\lambda_1$
Figure 2-4: $\chi^2_{PS}$ vs. $\delta$
\[ x_R^2 = 0.95465 x_{PS}^2 + 9.36853 \]

**Figure 2-5**  \( x_R^2 \) vs. \( x_{PS}^2 \)
FIGURE 2-6 $\hat{\alpha}$ vs. $\lambda_2/\lambda_1$
FIGURE 2-7 \( \hat{\alpha} \) vs. \( \delta \).
between all of the comparison indices plotted. It is only possible at present to make very tentative estimates of most of these relationships. The comparison variable pairs where a clear relationship exists are $\delta$ vs $\lambda_2/\lambda_1$, $\chi^2_{PS}$ vs $\delta$, and $\chi^2_{PS}$ vs $\lambda_2/\lambda_1$. These relationships do not vary with different random samples. However, for the relationships involving variables which are affected by different random samples-namely $\chi^2_R$ and $\hat{\alpha}$ - it is only possible to make tentative estimates of the relationships. To obtain better insight into these relationships, it would be necessary to increase the number of replications (using different random number seed values each time). Increasing the number of replications would increase the knowledge of the distribution of the values assumed by $\chi^2_R$ and $\hat{\alpha}$ at fixed values of the non-effected variables ($\delta$, $\lambda_2/\lambda_1$, and $\chi^2_{PS}$) and would also increase the accuracy of the estimate of the mean value of the affected variables. With sufficient replications it would be possible to obtain very reliable figures of the type shown in Figure 2-8 which illustrates a possible model of the relationship of $\hat{\alpha}$ and $\delta$. The knowledge of these tentative relationships is believed to be sufficient for the purposes of this study but it is recognized that further research should proceed in the effort to better quantify these relationships.

Examination of Figures 2-2 to 2-7 provides some valuable insights and also produces some interesting observations. First, as shown in Figure 2-2, $\delta$ changes at a faster rate for a given change in the lambda ratio for $\lambda_2/\lambda_1 < 1$ than for $\lambda_2/\lambda_1 > 1$. This indicates that if an alternative distribution is compared to a model distribution having a smaller parameter, of size $\lambda_2 - \varepsilon$ say, it is more likely that the two distributions can be considered as equivalent (because of a smaller $\delta$ value) than if the alternative distribution was compared to a model distribution having a correspondingly larger parameter of $\lambda_2 + \varepsilon$.

Second, as indicated in Figures 2-3 and 2-4, $\chi^2_{PS}$ (and therefore
FIGURE 2-8 Possible model of the Relationship of $\alpha$ and $\delta$
presumably $\chi^2_R$ and $\hat{\alpha}$) is more sensitive to differences in the distributions being compared for $\lambda_2/\lambda_1 < 1$ than for $\lambda_2/\lambda_1 > 1$. This sensitivity is in addition to the effect on $\delta$ of the relative magnitude difference of the parameters of the distributions. This indicates that in addition to the added likelihood that an alternative distribution and a model distribution can be considered equivalent for $\lambda_2/\lambda_1 > 1$ due to the magnitude effect, there is also an additional effect caused by the reduced sensitivity to difference for $\lambda_2/\lambda_1 > 1$.

Figure 2-5 indicates that $\chi^2_{PS}$ in general tends to be smaller than $\chi^2_R$. It also appears that there is a linear trend between $\chi^2_{PS}$ and $\chi^2_R$ if the point $\chi^2_{PS} = 76.8$, $\chi^2_R = 120.40$ is not considered. A least-squares line was calculated for this relationship. The equation of this line is $\chi^2_R = .95465 \cdot \chi^2_{PS} + 9.36853$.

Figure 2-5 and Table 2-1 also produce two interesting observations. First, there is a tighter grouping of $\chi^2_R$ from cases where $\lambda_2/\lambda_1 > 1$. Second, RNG seed A seems to produce peculiar results for cases where $\lambda_2/\lambda_1 < 1$. The underlying causes of these two phenomena are not fully understood.

From Figure 2-7 it appears that $\delta$ tends to be significant at the .05 level above values of .4. At values between $.3 \leq \delta \leq .4$ the result is ambiguous if the judgment is to be based on $\hat{\alpha} (\chi^2_R)$. If $\chi^2_{PS}$ is used to judge the two distributions for equivalence the range of uncertainty for $\delta$ can be determined from Table 2-1. Significance occurs at $\chi^2$ values around 17 at the .05 level. Therefore $\chi^2_{PS}$ would indicate significance at the .05 level above 8 since $.95465(8) + 9.36853 = 17$. It appears that if $\chi^2_{PS}$ is used as the judgment criterion, then for $\delta$ to indicate significance its value must be greater than .3 for $\lambda_2/\lambda_1 < 1$ and greater than .35 for $\lambda_2/\lambda_1 > 1$. Using $\chi^2_{PS}$ as the judgment criterion the uncertainty range for
\( \delta \) appears to be

\[
.2 \leq \delta \leq .3 \quad \lambda_2/\lambda_1 < 1
\]

\[
.25 \leq \delta \leq .35 \quad \lambda_2/\lambda_1 > 1
\]

These \( \delta \) values imply that if \( \lambda_2/\lambda_1 \) is to be used to test for significance as opposed to \( \delta \) then the uncertainty regions are

\[
\begin{align*}
\{ & .57 \leq \lambda_2/\lambda_1 \leq .67 \quad \lambda_2/\lambda_1 < 1 \} \ \hat{\alpha} \ \text{basis} \\
& 1.5 \leq \lambda_2/\lambda_1 \leq 1.75 \quad \lambda_2/\lambda_1 > 1 \\
\{ & .65 \leq \lambda_2/\lambda_1 \leq .75 \quad \lambda_2/\lambda_1 < 1 \} \ \chi_{PS}^2 \ \text{basis} \\
& 1.45 \leq \lambda_2/\lambda_1 \leq 1.6 \quad \lambda_2/\lambda_1 > 1
\end{align*}
\]

and significance is indicated for values of

\[
\lambda_2/\lambda_1 \leq .57 \quad \text{or} \quad \lambda_2/\lambda_1 \geq 1.75 \ \hat{\alpha} \ \text{basis}
\]

\[
\lambda_2/\lambda_1 \leq .65 \quad \text{or} \quad \lambda_2/\lambda_1 \geq 1.6 \ \chi_{PS}^2 \ \text{basis}
\]
Chapter 3
THE NORMAL CASE

3.0 INTRODUCTION

This chapter describes the specific procedure for calculating the index of non-congruity for normal distributions, the program developed to perform the comparison procedure outlined in Section 1.3.2. for normal distributions, and the results of using the procedure to compare several pairs of normal distributions.

3.1 DETERMINATION OF THE POINT OR POINTS OF INTERSECTION

Let \( f_1(x) \) be the probability density function of a normal distribution with mean \( u_1 \) and standard deviation \( \sigma_1 \). Let \( f_2(x) \) be the probability density function of a normal distribution with mean \( u_2 \) and standard deviation \( \sigma_2 \).

At a point of intersection we have,

\[
\frac{1}{\sqrt{2\pi} \sigma_1} e^{-1/2\left(\frac{x-u_1}{\sigma_1}\right)^2} - \frac{1}{\sqrt{2\pi} \sigma_2} e^{-1/2\left(\frac{x-u_2}{\sigma_2}\right)^2} = 0,
\]

so that

\[
e^{-1/2\left[\left(\frac{x-u_1}{\sigma_1}\right)^2 - \left(\frac{x-u_2}{\sigma_2}\right)^2\right]} = \frac{\sigma_1}{\sigma_2}
\]

and

\[
\left[\left(\frac{x-u_1}{\sigma_1}\right)^2 - \left(\frac{x-u_2}{\sigma_2}\right)^2\right] = -2\ln\frac{\sigma_1}{\sigma_2}.
\]

Then

\[
(\sigma_2^2 - \sigma_1^2)x^2 - 2(\sigma_2^2 u_1 - \sigma_1^2 u_2) + [\sigma_2^2 u_1^2 - \sigma_1^2 u_2^2 + 2\sigma_1 \sigma_2^2 u_1 u_2 - \ln\frac{\sigma_1}{\sigma_2}] = 0
\]

Equation (20) is of the form

\[
A x^2 + B x + C = 0
\]

The number of roots of equations of this form can be determined from the discriminant. If \( B^2 - 4AC \) is
negative  
there are no real roots

zero  
one real root $X_A = \frac{-B}{2A}$

positive  
two real roots, $X_{A1} = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$  
$X_{A2} = \frac{-B - \sqrt{B^2 - 4AC}}{2A}$

Denoting the discriminant by $R$, we find that

$$R = 4(c_2^2 u_1 - c_1^2 u_2)^2 - 4(c_2^2 - c_1^2)[c_2^2 u_1^2 - c_1^2 u_2^2 + 2c_2^2 c_1^2 n \frac{c_1}{c_2}]$$ (22)

which reduce to the conditions

$R$ negative  
yields no real roots

$R$ zero  
yields one root $X_A = \frac{c_2^2 u_1 - c_1^2 u_2}{c_2^2 - c_1^2}$

$R$ positive  
yields two roots $X_{A1} = X_A + \frac{\sqrt{R}}{2(c_2^2 - c_1^2)}$  
$X_{A2} = X_A - \frac{\sqrt{R}}{2(c_2^2 - c_1^2)}$

Now in the special case where $c_2 = c_1 = \sigma$, then Equation (17) reduces to

$$\frac{1}{\sqrt{2\pi} \sigma} e^{-1/2\left(\frac{x-u_1}{\sigma}\right)^2} - \frac{1}{\sqrt{2\pi} \sigma} e^{-1/2\left(\frac{x-u_2}{\sigma}\right)^2} = 0$$ (17a)

which yields $(x-u_1)^2 = (x-u_2)^2$ (24)

$x^2 - 2xu_1 + u_1^2 = x^2 - 2xu_2 + u_2^2$ (25)

$2xu_2 - 2xu_1 = u_2^2 - u_1^2$ (26)

$x = \frac{u_2^2 - u_1^2}{2(u_2 - u_1)}$ (27)

$X_A = \frac{u_2 + u_1}{2}$ (28)
3.2 THE INDEX OF NON-CONGRUITY: NORMAL CASE

The two most common situations with normal distributions, where there are two points of intersection and \( \sigma_1 = \sigma_2 \), are shown in Figures 3-1 and 3-2. The shaded area in these figures equals the index of non-congruity.

The index of non-congruity is given by Equation (6).

\[
\delta = \int_{-\infty}^{\infty} |f_1(x) - f_2(x)| \, dx
\]  

(6)

This expression can be modified according to the number of intersection points. If there is no point of intersection then obviously \( \delta = 2 \). If there is one point of intersection, \( X_A \), the integral can be decomposed into two terms. Assume that \( f_1(x) > f_2(x) \) for \( X < X_A \) then

\[
\delta = \int_{-\infty}^{X_A} (f_1(x) - f_2(x)) \, dx + \int_{X_A}^{\infty} (f_2(x) - f_1(x)) \, dx
\]  

(29)

\[
\delta = \int_{-\infty}^{X_A} f_1(x) \, dx - \int_{-\infty}^{X_A} f_2(x) \, dx + \int_{X_A}^{\infty} f_2(x) \, dx - \int_{X_A}^{\infty} f_1(x) \, dx
\]  

(30)

\[
\delta = F_1(X_A) - F_2(X_A) + [1 - F_2(X_A)] - [1 - F_1(X_A)]
\]  

(31)

\[
\delta = 2[F_1(X_A) - F_2(X_A)]
\]  

(32)

where \( F_1(x) \) is the cumulative probability function of \( f_1(x) \). In general, for the case where there is one point of intersection

\[
\delta = 2|F_1(X_A) - F_2(X_A)|
\]  

(33)

If there are two points of intersection (see Figure 3-2) then the index of non-congruity integral can be separated into three parts. Let \( C_1 \) and \( C_2 \) be the points of intersection.

Assume that \( f_1(x) > f_2(x) \) for \( X < C_1 \) then

\[
\delta = \int_{-\infty}^{C_1} (f_1(x) - f_2(x)) \, dx + \int_{C_2}^{\infty} (f_2(x) - f_1(x)) \, dx + \int_{C_1}^{C_2} (f_1(x) - f_2(x)) \, dx
\]  

(34)
FIGURE 3-1 The Normal Case with Two Points of Intersection
FIGURE 3-2 The Normal Case for Equal Standard Deviations
\[ \delta = F_1(C_1) - F_2(C_1) + [F_2(C_2) - F_2(C_1) - F_1(C_2) + F_1(C_1)] \\
+ [1 - F_1(C_2) - 1 + F_2(C_2)] \]  
(35)

\[ \delta = 2[(F_2(C_2) - F_2(C_1)) - (F_1(C_2) - F_1(C_1))] \]  
(36)

In general for the case where there are two points of intersection

\[ \delta = |2[(F_2(C_2) - F_2(C_1)) - (F_1(C_2) - F_1(C_1))]| \]  
(37)

3.3 DESCRIPTION OF NORMAL PROGRAM FEATURES

3.3.1. Program Listing

This section describes the essential features of the program for comparing normal distributions. A complete listing of the program is given in Appendix 2.

3.3.2. Definition of Program Variables

AHAT: the level of significance, \( \hat{\alpha} \)

CI: the lower point of intersection, \( C_1 \)

C2: the upper point of intersection, \( C_2 \)

CADTR: function subroutine to determine \( \hat{\alpha} \)

D: an output parameter of subroutine NDTR which is not used in the main program

DELTA: the index of non-congruity, \( \delta \)

EI: the expected frequency in the equi-probability regions, \( M/10 \)

F1: the cumulative probability \( F_1(X) \) at \( X_A \)

F1C1: the cumulative probability \( F_1(X) \) at \( C_1 \)

F1C2: the cumulative probability \( F_1(X) \) at \( C_2 \)

F1DIF: \( F_1C2 - F_1C1 \)

F2: the cumulative probability \( F_2(X) \) at \( X_A \)

F2C1: the cumulative probability \( F_2(X) \) at \( C_1 \)

F2C2: the cumulative probability \( F_2(X) \) at \( C_2 \)
F2DIF: F2C2 - F2C1
F2SUM: the sum of the components of the array FREQ squared
FREQ(IO): The array containing the frequency counts of the random sample sorted into the equi-probability regions
FRSFAC: $4(\sigma_2^2 u_1 - \sigma_2^2 u_2)^2$
IER: error indicator used in subroutine NDTRI
IND: indicator showing the number of intersection points
ISEED: one of the seeds for the McGill RNG
JSEED: the other seed for the McGill RNG
K: the index for the array FREQ
MU1: population mean $u_1$
MU2: population mean $u_2$
M: the sample size
NDTR: subroutine to calculate $F_1(z)$
NDTRI: subroutine to calculate the standard normal deviate $z$, given $F_1(z)$
P: input parameter to subroutine NDTRI containing $F_1(z)$
P2(IO): the array containing the cumulative probability of the alternative distribution at the region boundaries $\{x(i)\}$
RAD: $R$ of Equation (22)
RADPRT: $\sqrt{R}/2(\sigma_2^2 - \sigma_1^2)$
RATIO: the ratio of population standard deviations, $\sigma_1/\sigma_2$
RNOR: McGill RNG function to generate standard normal deviate
RSTART: subroutine to initialize the McGill RNG
SAMPL(200): the array containing the random sample from the alternative distribution
SIGMA1: population standard deviation, $\sigma_1$
SIGMA2: population standard deviation, $\sigma_2$
SNDFAC: $4(\sigma_2^2 - \sigma_1^2)[\sigma_2^2 u_1^2 - \sigma_1^2 u_2^2 + 2\sigma_1^2 \sigma_2^2 \ln \frac{\sigma_1}{\sigma_2}]$

SQMU1: $u_1^2$

SQMU2: $u_2^2$

VAR1: $\sigma_1^2$

VAR2: $\sigma_2^2$

VARDIF: VAR2 - VAR1

VXM12: $\sigma_1^2 u_2$

VXM21: $\sigma_2^2 u_1$

VXMD2: $(\sigma_2^2 u_1 - \sigma_1^2 u_2)^2$

VXMDIF: $(\sigma_2^2 u_1 - \sigma_1^2 u_2)$

VXS12: $\sigma_1^2 u_2^2$

VXS21: $\sigma_2^2 u_1^2$

X1(10): the equi-probability region boundaries \{X(i)\}

X2ACT: the random $\chi^2$ statistic, $\chi_R^2$

X2PS: the pseudo $\chi^2$ statistic, $\chi_{PS}^2$

X2SUM: the sum of [P2(10) - X1(10)] used in the calculation of X2P

XLNFACT: $2 \sigma_2^2 \sigma_1^2 \ln \frac{\sigma_1}{\sigma_2}$

XX: the single point of intersection, $X_A$

XX1: one of two points of intersection, $X_{A1}$

XX2: the other of two points of intersection, $X_{A2}$

Z: a standard normal deviate used in calculating X1(10)

Z1: standard normal equivalent of $X_A$ for distribution 1

Z1C1: standard normal equivalent of $C_1$ for distribution 1

Z1C2: standard normal equivalent of $C_2$ for distribution 1

Z2: standard normal equivalent of $X_A$ for distribution 2

Z2C1: standard normal equivalent of $C_1$ for distribution 2

Z2C2: standard normal equivalent of $C_2$ for distribution 2

Z2PS(10): array containing standard normal equivalent of X1(10) for distribution 2
3.3.3. Inputs to the Program

The inputs to the program are the distribution parameters \( u_1, \sigma_1, u_2, \sigma_2 \); the sample size \( M \); and the RNG seed values - ISEED, JSEED. The input is given on two separate cards with the indicated format.

\[
\begin{align*}
&MU_1, SIGMA_1, MU_2, SIGMA_2, M & 1 \text{ Card} & (4E10.4, I5) \\
&ISEED, JSEED & 1 \text{ Card} & (2I5)
\end{align*}
\]

Multiple runs are possible by supplying additional input cards (two per replication). Program completion is indicated by a blank card.

3.3.4. Outputs of the Program

The output of the normal program is almost identical with the output of the exponential program. There are two types of output of the normal program. The first type is an echo check of the inputs to the program and the second type is the set of values calculated by the program. As in the exponential program, the second type of output for the normal program consists of the values of \( \delta, x_{PS}^2, x_R^2, \alpha, X_1(10), P_2(10) \), the first \( M \) elements of SAMPL(200), and FREQ(10). In addition to these second type outputs, the normal program also prints the number of intersection points and their values. A sample output is contained in Appendix 2.

3.3.5. Special Programming Considerations

Definition of Non-Overlapping Distributions  In actuality any two normal distributions will intersect in at least one point, since both span the interval from \(-\infty\) to \(+\infty\) and both have unit area. However, for widely separated distributions the amount of overlapping area is small and at some point could be considered zero. In the program, if \( u_S + 5\sigma_S \) is less than \( u_L - 5\sigma_L \) (where \( u_S \) refers to the smaller mean, \( u_L \) refers to the larger mean, and \( \sigma_S \) and \( \sigma_L \) refer to the corresponding standard deviations) then the distributions are defined as non-overlapping and \( \delta \) is set equal
Using the McGill RNG  Use of the McGill RNG is accomplished in this program through the use of two subprograms - RSTART and RNOR. The use of RSTART is described in section 2.3.4. RNOR is used to generate a sample from a standard normal distribution. Its use is similar to the use of REXP described in Section 2.3.4. The sample from a standard normal distribution is transformed to a sample from a normal distribution with mean u and standard deviation σ by the equation

\[ X = u + \sigma z \]  

(38)

where \( z \) is the sample from a standard normal distribution and \( X \) is the observation from the desired distribution.

Sorting the Random Sample Observations  The normal index of non-congruity program uses a different sorting scheme than the exponential program did because of the different shapes of the two distributions. The sorting scheme is shown in the tree diagram shown in Figure 3-3. This scheme is designed to search the middle equi-probability regions first. The efficiency of this sorting scheme is dependent on the nature of the alternative distribution and, therefore, the scheme does not minimize the expected number of tests in all situations. However, the scheme does reduce the maximum number of tests to 5, as compared with 9 in the scheme used in the exponential program.

Calculating the Normal Cumulative Probability of \( X \)  The cumulative probability for an argument \( X \) is calculated by first converting the number to standard form by the transformation,

\[ z = (x - u)/\sigma \]  

(39)

The cumulative probability is then calculated by the IBM Scientific Subroutine Package subroutine NDTR which uses the following approximation taken from Hastings
FIGURE 3-3 Tree Diagram for Sorting Procedure
\[ F(z) = 1 - f(z)(b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5) \]  \hspace{1cm} (40)

where

\[ f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad t = \frac{1}{1 + rz}, \quad r = 0.2316419, \]

\[ b_1 = 0.31938153, \quad b_2 = -0.356563782, \quad b_3 = 1.7181477937, \]

\[ b_4 = -1.821255978, \quad b_5 = 1.33027449 \]

This approximation has a maximum error of 7.5 \times 10^{-8} and is valid only for \( z \geq 0 \). For \( z < 0 \) the complement of \( F(-z) \) gives the desired value.

Calculating \( X \) given a Normal Cumulative Probability

A value \( X \) from a \( N(\mu, \sigma^2) \) population can be calculated from a given normal cumulative probability \( P \) by first calculating the value \( Z \) from a standard normal distribution with cumulative probability \( P \) and then applying the transformation given in Equation (38). The value \( Z \) is calculated by the IBM Scientific Subroutine Package subroutine NDTRI which uses the following approximation taken from Hastings,

\[ z = w - \sum_{i=0}^{2} a_i w^i/\sum_{i=0}^{3} b_i w^i \] \hspace{1cm} (41)

where \( w = \sqrt{\ln(1/p^2)} \), \( a_0 = 2.515517, \quad a_1 = 0.802853, \quad a_2 = 0.010328, \]

\[ b_0 = 1, \quad b_1 = 1.432788, \quad b_2 = 0.189269, \quad b_3 = 0.001308 \]

This approximation has a maximum error of 4.5 \times 10^{-4} and is valid only for \( P \leq 0.5 \). For \( P > 0.5 \), \( Z \) of \( 1-P \) is calculated and then the sign of \( Z \) is changed.

3.4 RESULTS OF THE NORMAL PROGRAM

Four different sets of values for random number generator seeds were used in investigating the normal case. The standard normal distribution was chosen as the model distribution. Three different types of alternative distributions were considered. In the first type the mean of the alternative
distribution was different from zero and the standard deviation was equal to one. This set of alternative distributions is referred to as the mean-variate set. In the second type the mean of the alternative distribution was equal to zero and the standard deviation was different from one. This set of alternative distributions is referred to as the variance-variate set. In the third type of alternative distribution considered, the mean of the alternative distribution was different from zero and the standard deviation was different from one. This set is referred to as the mean-variance-variate set. The sample size used in the comparisons was 50 in all cases.

One of the steps in the comparison procedure outlined in Section 1.3.2 is the calculation of various parametric indicators. The parametric indicator which was chosen in the normal case was,

$$\eta_1 = |u_1 - u_2| + |\sigma_1 - \sigma_2|$$  (42)

Various relationships between comparison indices are graphically presented in Figures 3-4 to 3-20. There appears to be a strong relationship between all of the comparison indices plotted. Examination of these figures provides some valuable information. Figure 3-4 shows the relationship between $\chi^2_{PS}$ and $\delta$. This figure indicates that for the cases of the mean-variate and variance-variate sets of alternative distributions $\chi^2_{PS}$ and $\delta$ are strongly related. The figure also indicates that two distributions with a particular $\delta$ value are more easily detected as being significantly different if their means are different than if their standard deviations are different. The figure also shows a greater difference in $\chi^2_{PS}$ for a given $\delta$ value for $\sigma_1 < \sigma_2$ than for $\sigma_1 > \sigma_2$. (Recall that in the exponential case this type of a relationship existed; $\chi^2_{PS}$ for $\lambda_2/\lambda_1 < 1$ was greater than $\chi^2_{PS}$ for $\lambda_2/\lambda_1 > 1$. This is a similar result since the standard deviation of an exponential distribution is $1/\lambda$ and therefore
FIGURE 3-4 $\chi^2_{PS}$ vs. $\delta$ - Single Variation
FIGURE 3-5 $\chi^2_{PS}$ vs. $\delta$ - Dual Variation

$\chi^2_{PS} = 0.0157333 \delta^2$
FIGURE 3-6  $\delta$ vs. $\eta_1$ - Single Variation
\[ \delta = 0.571 \eta_1 + 0.015733 \]
$x_{PS}^2 = 29.5171103 \eta_1^2$

**Figure 3-9** $x_{PS}^2$ vs. $\eta_1$ - Dual Variation
FIGURE 3-10

\[ \chi^2_R = 0.90935 \chi^2_{PS} + 6.55919 \]

\( \chi^2_R \) vs. \( \chi^2_{PS} \) - Mean-Variate Data
\[ x_R^2 = 0.92777 \, x_{PS}^2 + 4.95750 \]

**Figure 3-11** \( x_R^2 \) vs. \( x_{PS}^2 \) - Variance-Variate Data \( \sigma_1 \leq \sigma_2 \)
**Figure 3-12** $\chi^2_R$ vs. $\chi^2_{PS}$ - Variance-Variate Data $\sigma_1 > \sigma_2$

The relationship is given by the equation:

$$\chi^2_R = 0.85908 \chi^2_{PS} + 3.29251$$
$\chi^2_R = 0.580767 \chi^2_{PS} + 5.83267$

**Figure 3-13** $\chi^2_R$ vs. $\chi^2_{PS}$ - Dual Variation
FIGURE 3-14  $\hat{\alpha}$ vs. $\delta$ - Mean-Variate Data
Figure 3-15  \( \hat{\alpha} \) vs. \( \delta \) - Variance-Variate Data \( \sigma_1 < \sigma_2 \)
Figure 3-16: $\hat{\alpha}$ vs. $\delta$ - Variance-Variate Data $\sigma_1 > \sigma_2$
FIGURE 3-17  $\hat{\alpha}$ vs. $\delta$ - Dual Variation
FIGURE 3-18 \( \hat{\alpha} \) vs. \( \eta \) - Mean-Variate Data
FIGURE 3-19  $\hat{\alpha}$ vs. $\eta_1$ - Variance-Variate Data $\sigma_1 < \sigma_2$
\[ \hat{\alpha} \text{ vs. } \eta_1 \quad \text{Variance-Variate Data } \sigma_1 \times \sigma_2 \]
the ratio of the standard deviations is \(1/\lambda_2/1/\lambda_1 = \lambda_1/\lambda_2\). Hence if \(\lambda_2/\lambda_1 < 1\) then \(\sigma_{\text{exp1}} < \sigma_{\text{exp2}}\).

Figure 3-6 indicates the relationship between \(\delta\) and \(\eta_1\) for the mean-variate and variance-variate data. The most sensitive case is for the variance-variate case with \(\sigma_1 > \sigma_2\). This sensitivity is partially offset by the relationship between \(\chi^2_{ps}\) and \(\delta\) for this case (which is not as sensitive as the other two cases shown) as shown in Figure 3-4. However, it is not completely offset as seen in Figure 3-8 which shows that the case of variance-variate data with \(\sigma_1 > \sigma_2\) produces a much higher \(\chi^2_{ps}\) value for a given \(\eta_1\) than the other two cases (which produce comparable values).

As in the exponential case \(\chi^2_{ps}\) tends to be smaller than \(\chi^2_R\), as seen in Figures 3-10 to 3-13. There also appears to be a linear trend between \(\chi^2_R\) and \(\chi^2_{ps}\) in each of these figures. A least-squares line was calculated for each of these cases. The obviously outlying points of Figures 3-11 and 3-13 were omitted from the calculations. The derived lines are

- **Figure 3-10** Mean-Variate Data
  \[
  \chi^2_R = .90935 \chi^2_{ps} + 6.55919
  \]
- **Figure 3-11** Variance-Variate Data \(\sigma_1 < \sigma_2\)
  \[
  \chi^2_R = .92777 \chi^2_{ps} + 4.95750
  \]
- **Figure 3-12** Variance-Variate Data \(\sigma_1 > \sigma_2\)
  \[
  \chi^2_R = .85908 \chi^2_{ps} + 3.29251
  \]
- **Figure 3-13** Mean-Variance-Variate Data
  \[
  \chi^2_R = .580767 \chi^2_{ps} + 5.83267
  \]

From Figures 3-14 to 3-16 it appears that (based on \(\hat{\alpha}\) as the criterion) \(\delta\) tends to be significant at the .05 level as indicated below:

<table>
<thead>
<tr>
<th>Figure</th>
<th>Type</th>
<th>Significance Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-14</td>
<td>Mean-Variate Data</td>
<td>(\delta &gt; .47)</td>
</tr>
<tr>
<td>Figure</td>
<td>Type</td>
<td>Significance Range</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>3-15</td>
<td>Variance-Variate Data $\sigma_1 &lt; \sigma_2$</td>
<td>$\delta \geq .39$</td>
</tr>
<tr>
<td>3-16</td>
<td>Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>$\delta \geq .6$</td>
</tr>
</tbody>
</table>

If $x^2_{PS}$ is used as the criterion, $x^2_{PS}$ would indicate significance at the following values calculated from the regression equations to correspond to a $x^2_R$ value of 17 ($\alpha = .05$).

<table>
<thead>
<tr>
<th>Type</th>
<th>Significance Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Variate Data</td>
<td>$x^2_{PS} \geq 11.5$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &lt; \sigma_2$</td>
<td>$x^2_{PS} \geq 13.0$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>$x^2 \geq 16.0$</td>
</tr>
</tbody>
</table>

Based on these values and consulting Figure 3-4, $\delta$ would indicate significance at the .05 level as given below.

<table>
<thead>
<tr>
<th>Type</th>
<th>Significance Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Variate Data</td>
<td>$\delta \geq .36$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &lt; \sigma_2$</td>
<td>$\delta \geq .4$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>$\delta \geq .55$</td>
</tr>
</tbody>
</table>

The uncertainty regions can be estimated as

<table>
<thead>
<tr>
<th>Type</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Variate Data</td>
<td>$0.38 \leq \delta \leq 0.47$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &lt; \sigma_2$</td>
<td>$0.34 \leq \delta \leq 0.39$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>$0.48 \leq \delta \leq 0.6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Variate Data</td>
<td>$0.32 \leq \delta \leq 0.36$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &lt; \sigma_2$</td>
<td>$0.34 \leq \delta \leq 0.4$</td>
</tr>
<tr>
<td>Variance-Variate Data $\sigma_1 &gt; \sigma_2$</td>
<td>$0.45 \leq \delta \leq 0.55$</td>
</tr>
</tbody>
</table>

These values for $\delta$ imply that if $\eta_1$ is to be used as a test for significance instead of $\delta$, then the uncertainty regions can be estimated from Figure 3-6 to be
Mean-Variate Data  \( .48 \leq \eta_1 \leq .6 \)

Variance-Variate Data  \( \sigma_1 < \sigma_2 \)  \( .43 \leq \eta_1 \leq .5 \)  \( \hat{\alpha} \) basis

Variance-Variate Data  \( \sigma_1 > \sigma_2 \)  \( .4 \leq \eta_1 \leq .46 \)

Mean-Variate Data  \( .4 \leq \eta_1 \leq .44 \)

Variance-Variate Data  \( \sigma_1 < \sigma_2 \)  \( .42 \leq \eta_1 \leq .52 \)  \( \chi^2_{PS} \) basis

Variance-Variate Data  \( \sigma_1 > \sigma_2 \)  \( .38 \leq \eta_1 \leq .42 \)

These regions compare favorably with results which can be taken from Figures 3-18 to 3-20 for the \( \hat{\alpha} \) basis and Figure 3-8 for the \( \chi^2_{PS} \) basis. Significance is therefore indicated for \( \eta_1 \) values of

Mean-Variate Data  \( \eta_1 \geq .6 \)

Variance-Variate Data  \( \sigma_1 < \sigma_2 \)  \( \eta_1 \geq .5 \)  \( \hat{\alpha} \) basis

Variance-Variate Data  \( \sigma_1 > \sigma_2 \)  \( \eta_1 \geq .46 \)

Mean-Variate Data  \( \eta_1 \geq .44 \)

Variance-Variate Data  \( \sigma_1 < \sigma_2 \)  \( \eta_1 \geq .52 \)  \( \chi^2_{PS} \) basis

Variance-Variate Data  \( \sigma_1 > \sigma_2 \)  \( \eta_1 \geq .42 \)

Figures 3-5, 3-7, 3-9, 3-13 and 3-17 show the various relationships between indices for the mean-variance-variate data. Indications of strong relationships between the various indices are shown by these figures. However, it is believed that the knowledge of these relationships is too limited to draw any satisfactory results. Further research, in which the range of the comparison indices is expanded, is needed to better quantify these relationships.
### TABLE 3-1

Results of the Normal Program

<table>
<thead>
<tr>
<th>DISTRIBUTION 2 (Alternative)</th>
<th>RNG</th>
<th>δ</th>
<th>$\chi^2_{PS}$</th>
<th>$\chi^2_{R}$</th>
<th>$\hat{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(.1, 1)</td>
<td>A</td>
<td>.0798</td>
<td>.48</td>
<td>7.60</td>
<td>.5749</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2.80</td>
<td>.9717</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.40</td>
<td>.9835</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>12.00</td>
<td>.2133</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.2, 1)</td>
<td>A</td>
<td>.1593</td>
<td>1.94</td>
<td>8.40</td>
<td>.4944</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2.00</td>
<td>.9915</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>6.80</td>
<td>.6579</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>12.80</td>
<td>.1719</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.3, 1)</td>
<td>A</td>
<td>.2385</td>
<td>4.41</td>
<td>13.20</td>
<td>.1538</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>10.00</td>
<td>.3505</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>8.00</td>
<td>.5341</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>12.40</td>
<td>.1917</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.4, 1)</td>
<td>A</td>
<td>.3170</td>
<td>7.97</td>
<td>13.60</td>
<td>.1373</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>8.00</td>
<td>.5341</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>10.00</td>
<td>.3505</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>20.00</td>
<td>.0179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.5, 1)</td>
<td>A</td>
<td>.3948</td>
<td>12.70</td>
<td>20.00</td>
<td>.0179</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>11.60</td>
<td>.2368</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>16.00</td>
<td>.0669</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>25.20</td>
<td>.0028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.6, 1)</td>
<td>A</td>
<td>.4716</td>
<td>18.70</td>
<td>30.00</td>
<td>.0004</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>19.20</td>
<td>.0235</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>18.00</td>
<td>.0352</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>33.20</td>
<td>.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.7, 1)</td>
<td>A</td>
<td>.5473</td>
<td>26.08</td>
<td>34.00</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>22.00</td>
<td>.0089</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>30.40</td>
<td>.0004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>40.80</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.8, 1)</td>
<td>A</td>
<td>.6217</td>
<td>34.94</td>
<td>43.20</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>26.40</td>
<td>.0018</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>38.00</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>44.00</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.9, 1)</td>
<td>A</td>
<td>.6946</td>
<td>45.38</td>
<td>62.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>40.40</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>40.80</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>53.20</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3-1 continued

<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>RNG SEED</th>
<th>( \delta )</th>
<th>( \chi^2_{\text{PS}} )</th>
<th>( \chi^2_{\text{R}} )</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>((1.0, 1))</td>
<td>A</td>
<td>0.7659</td>
<td>57.45</td>
<td>66.40</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>48.40</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>49.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>63.20</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.1, 1))</td>
<td>A</td>
<td>0.8354</td>
<td>71.19</td>
<td>83.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>67.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>66.80</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>77.60</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.2, 1))</td>
<td>A</td>
<td>0.9030</td>
<td>86.58</td>
<td>97.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>76.00</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>78.80</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>92.80</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.3, 1))</td>
<td>A</td>
<td>0.9686</td>
<td>103.55</td>
<td>137.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>85.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>89.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>106.00</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.4, 1))</td>
<td>A</td>
<td>1.0321</td>
<td>121.99</td>
<td>143.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>95.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>111.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>123.60</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.5, 1))</td>
<td>A</td>
<td>1.0935</td>
<td>141.70</td>
<td>172.40</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>118.80</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>126.00</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>126.00</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.6, 1))</td>
<td>A</td>
<td>1.1526</td>
<td>162.46</td>
<td>184.00</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>147.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>141.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>156.40</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.7, 1))</td>
<td>A</td>
<td>1.2093</td>
<td>184.00</td>
<td>206.40</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>163.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>155.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>180.00</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.8, 1))</td>
<td>A</td>
<td>1.2638</td>
<td>206.00</td>
<td>246.40</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>173.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>175.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>180.80</td>
<td>0.0000</td>
</tr>
<tr>
<td>((1.9, 1))</td>
<td>A</td>
<td>1.3158</td>
<td>228.15</td>
<td>274.80</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>197.20</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>235.60</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>206.80</td>
<td>0.0000</td>
</tr>
<tr>
<td>DISTRIBUTION 2 (Alternative)</td>
<td>RNG SEED</td>
<td>$\delta$</td>
<td>$\chi^2_{PS}$</td>
<td>$\chi^2_{R}$</td>
<td>$\hat{\alpha}$</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>(2.0, 1)</td>
<td>A</td>
<td>1.3654</td>
<td>250.11</td>
<td>290.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>226.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>261.20</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>220.80</td>
<td>.0000</td>
</tr>
<tr>
<td>(0, .1)</td>
<td>A</td>
<td>1.5964</td>
<td>194.35</td>
<td>182.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>174.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>192.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>181.20</td>
<td>.0000</td>
</tr>
<tr>
<td>(0, .2)</td>
<td>A</td>
<td>1.2942</td>
<td>117.35</td>
<td>95.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>125.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>113.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>88.80</td>
<td>.0000</td>
</tr>
<tr>
<td>(0, .3)</td>
<td>A</td>
<td>1.0435</td>
<td>67.04</td>
<td>63.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>58.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>75.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>50.00</td>
<td>.0000</td>
</tr>
<tr>
<td>(0, .4)</td>
<td>A</td>
<td>.8300</td>
<td>40.56</td>
<td>39.60</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>40.80</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>42.00</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>26.40</td>
<td>.0018</td>
</tr>
<tr>
<td>(0, .5)</td>
<td>A</td>
<td>.6453</td>
<td>24.72</td>
<td>25.20</td>
<td>.0028</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>22.00</td>
<td>.0089</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>32.40</td>
<td>.0002</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>15.60</td>
<td>.0757</td>
</tr>
<tr>
<td>(0, .6)</td>
<td>A</td>
<td>.4840</td>
<td>14.48</td>
<td>14.40</td>
<td>.1088</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>14.40</td>
<td>.1088</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>20.00</td>
<td>.0179</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td>(0, .7)</td>
<td>A</td>
<td>.3416</td>
<td>7.63</td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>16.00</td>
<td>.0669</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>10.40</td>
<td>.3191</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>6.40</td>
<td>.6993</td>
</tr>
<tr>
<td>(0, .8)</td>
<td>A</td>
<td>.2151</td>
<td>3.19</td>
<td>4.80</td>
<td>.8514</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>5.20</td>
<td>.8165</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>11.60</td>
<td>.2368</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>2.40</td>
<td>.9835</td>
</tr>
<tr>
<td>(0, .9)</td>
<td>A</td>
<td>.1019</td>
<td>.75</td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>2.40</td>
<td>.9835</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>8.00</td>
<td>.5341</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>4.40</td>
<td>.8832</td>
</tr>
<tr>
<td>DISTRIBUTION 2 (Alternative)</td>
<td>RNG SEED</td>
<td>( \delta )</td>
<td>( \chi^2_{PS} )</td>
<td>( \chi^2_R )</td>
<td>( \hat{\alpha} )</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>((0, 1.1))</td>
<td>A</td>
<td>.0922</td>
<td>.65</td>
<td>5.60</td>
<td>.7792</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>6.80</td>
<td>.6579</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>4.00</td>
<td>.9114</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>10.80</td>
<td>.2897</td>
</tr>
<tr>
<td>((0, 1.2))</td>
<td>A</td>
<td>.1760</td>
<td>2.42</td>
<td>6.40</td>
<td>.6993</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>7.20</td>
<td>.6163</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>18.80</td>
<td>.0269</td>
</tr>
<tr>
<td>((0, 1.3))</td>
<td>A</td>
<td>.2525</td>
<td>5.05</td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>10.00</td>
<td>.3505</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>7.60</td>
<td>.5749</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>22.00</td>
<td>.0089</td>
</tr>
<tr>
<td>((0, 1.4))</td>
<td>A</td>
<td>.3226</td>
<td>8.29</td>
<td>15.20</td>
<td>.0856</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>11.20</td>
<td>.2622</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>14.40</td>
<td>.1088</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>32.00</td>
<td>.0002</td>
</tr>
<tr>
<td>((0, 1.5))</td>
<td>A</td>
<td>.3872</td>
<td>11.97</td>
<td>17.20</td>
<td>.0457</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>19.60</td>
<td>.0205</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>16.00</td>
<td>.0669</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>35.60</td>
<td>.0000</td>
</tr>
<tr>
<td>((0, 1.6))</td>
<td>A</td>
<td>.4467</td>
<td>15.93</td>
<td>18.80</td>
<td>.0269</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>20.40</td>
<td>.0156</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>22.40</td>
<td>.0077</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>47.20</td>
<td>.0000</td>
</tr>
<tr>
<td>((0, 1.7))</td>
<td>A</td>
<td>.5019</td>
<td>20.07</td>
<td>19.60</td>
<td>.0205</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>24.40</td>
<td>.0037</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>24.40</td>
<td>.0037</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>45.60</td>
<td>.0000</td>
</tr>
<tr>
<td>((0, 1.8))</td>
<td>A</td>
<td>.5531</td>
<td>24.29</td>
<td>27.60</td>
<td>.0011</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>26.80</td>
<td>.0015</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>25.20</td>
<td>.0028</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>45.60</td>
<td>.0000</td>
</tr>
<tr>
<td>((0, 1.9))</td>
<td>A</td>
<td>.6008</td>
<td>28.52</td>
<td>27.60</td>
<td>.0011</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>32.80</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>28.40</td>
<td>.0008</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>55.60</td>
<td>.0000</td>
</tr>
<tr>
<td>((0, 2.0))</td>
<td>A</td>
<td>.6453</td>
<td>32.73</td>
<td>30.40</td>
<td>.0004</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>40.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>40.40</td>
<td>.0000</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>62.40</td>
<td>.0000</td>
</tr>
</tbody>
</table>
Table 3-1 continued

<table>
<thead>
<tr>
<th>DISTRIBUTION 2 (Alternative)</th>
<th>RNG SEED</th>
<th>$\delta$</th>
<th>$\chi^2_{PS}$</th>
<th>$\chi^2_{R}$</th>
<th>$\hat{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(.1, .8)$</td>
<td>A</td>
<td>.2282</td>
<td>3.59</td>
<td>5.60</td>
<td>.7792</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>10.00</td>
<td>.3505</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>10.40</td>
<td>.3191</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>4.40</td>
<td>.8832</td>
</tr>
<tr>
<td>$(.1, .9)$</td>
<td>A</td>
<td>.1262</td>
<td>1.19</td>
<td>6.40</td>
<td>.6693</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>3.60</td>
<td>.9357</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td>$(.1, 1.1)$</td>
<td>A</td>
<td>.1142</td>
<td>1.16</td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>7.20</td>
<td>.6163</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>4.00</td>
<td>.9114</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>14.80</td>
<td>.0966</td>
</tr>
<tr>
<td>$(.1, 1.2)$</td>
<td>A</td>
<td>.1868</td>
<td>2.95</td>
<td>9.60</td>
<td>.3838</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>8.40</td>
<td>.4944</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>9.20</td>
<td>.4190</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>17.20</td>
<td>.0457</td>
</tr>
<tr>
<td>$(.2, .8)$</td>
<td>A</td>
<td>.2659</td>
<td>4.79</td>
<td>12.00</td>
<td>.2133</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>5.60</td>
<td>.7792</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td>$(.2, .9)$</td>
<td>A</td>
<td>.1876</td>
<td>2.52</td>
<td>7.60</td>
<td>.5749</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>10.80</td>
<td>.2897</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td>$(.2, 1.1)$</td>
<td>A</td>
<td>.1698</td>
<td>2.70</td>
<td>8.40</td>
<td>.4944</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>4.80</td>
<td>.8514</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>8.00</td>
<td>.5341</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>16.40</td>
<td>.0590</td>
</tr>
<tr>
<td>$(.2, 1.2)$</td>
<td>A</td>
<td>.2177</td>
<td>4.52</td>
<td>12.80</td>
<td>.1719</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>4.00</td>
<td>.9114</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>8.80</td>
<td>.4559</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>23.60</td>
<td>.0050</td>
</tr>
<tr>
<td>$(.3, .8)$</td>
<td>A</td>
<td>.3230</td>
<td>6.86</td>
<td>13.60</td>
<td>.1373</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>6.80</td>
<td>.6579</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>7.60</td>
<td>.5749</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>10.80</td>
<td>.2897</td>
</tr>
<tr>
<td>$(.3, .9)$</td>
<td>A</td>
<td>.2639</td>
<td>4.81</td>
<td>11.60</td>
<td>.2368</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
<td>6.00</td>
<td>.7399</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td>8.00</td>
<td>.5341</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td>14.00</td>
<td>.1223</td>
</tr>
</tbody>
</table>
Table 3-1 continued

<table>
<thead>
<tr>
<th>DISTRIBUTION 2 (Alternative)</th>
<th>RNG SEED</th>
<th>$\delta$</th>
<th>$\chi^2_{PS}$</th>
<th>$\chi^2_{R}$</th>
<th>$\hat{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(.3, 1.1)</td>
<td>A</td>
<td>.2389</td>
<td>5.29</td>
<td>13.60</td>
<td>.1373</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td>7.20</td>
<td>.6163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td>7.20</td>
<td>.6163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td>21.20</td>
<td>.0118</td>
<td></td>
</tr>
<tr>
<td>(.3, 1.2)</td>
<td>A</td>
<td>.2648</td>
<td>7.16</td>
<td>14.80</td>
<td>.0966</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td>6.40</td>
<td>.6993</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td>7.20</td>
<td>.6163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td>26.80</td>
<td>.0015</td>
<td></td>
</tr>
</tbody>
</table>

$M = 50$, RNG: A(51562, 62155) B(62155, 51562) C(50020, 11292) D(11292, 50020)
Chapter 4
CONCLUSION

In concluding this study, two questions need to be answered. The first question is "How well did this study accomplish its objective?". The second question is "What direction should future research in this area take?".

The first question can be answered by reconsidering the objective of this study which was to develop a comparison method, use this method to investigate the effects of varying distribution parameters, and to evaluate the usefulness of this technique. The first two parts of this objective have already been accomplished and the third part can be completed by a brief review of the results of the study. The technique used in this study to compare statistical distributions appears to have considerable usefulness because of the consistency of results (i.e. \( \delta \) in all cases indicated significance in the range \( .3 \leq \delta \leq .6 \)), the ability to compare distributions for statistical difference in terms of only their parameters (\( \lambda_2/\lambda_1 \) for exponential distributions, \( n_1 \) for most normal distributions), and the relative simplicity of the method.

The reader should recognize that in the normal case, a technique already exists to answer the question of "How different is different?" based on the classical z-test. This technique is statistically sufficient and is therefore more powerful than the method presented in this study. The use of the z-test is demonstrated below for \( M = 50, \alpha = .05 \), and \( \sigma_1 = \sigma_2 = \sigma \).

\[
1.96 \leq \frac{u_1 - u_2}{\sqrt{\sigma^2/50}}, \quad (43)
\]

\[
u_1 - u_2 \geq .277 \sigma \quad (44)
\]
This indicates that for our investigation ($\sigma = 1$) significance would be indicated for a $\eta_1$ value greater than or equal to .277, as compared to a $\eta_1$ value of .60 for the index of non-congruity method.

The second question concerning the direction of future research is easily answered. There are four readily apparent directions for future research. They are:

1. Extension of this research in terms of additional replications and inclusion of more mean-variance-variate comparisons for the normal case as mentioned in Sections 2.4 and 3.4.

2. Application of the methodology used in this study to other continuous distributions such as the Weibull, Log-Normal, or Gamma distribution.

3. Development of a similar methodology which can be applied to the evaluation of statistically significant differences in discrete distributions.

4. Application of the methodology used in this study to study statistical differences of similar distributions which are from different families.
SELECTED BIBLIOGRAPHY


PROGRAM TO EVALUATE "THE INDEX OF NON-CONGRUITY" FOR EXPONENTIAL DISTRIBUTIONS. THIS INDEX IS BASED ON THE AMOUNT OF NON-OVERLAPPING AREA BETWEEN TWO DISTRIBUTIONS. VALUES TO BE SUPPLIED TO THE PROGRAM ARE DISTRIBUTION PARAMETERS, SAMPLE SIZE, AND INITIAL VALUES FOR THE RNG.

FORMAT FOR DATA CARDS: LAMDA1, LAMDA2, M 1 CARD (2E10.4, I5)

If (LAMDA2 = EQ.0) GO TO 201
If (LAMDA2.LT.LAMDA1) GO TO 120
LAMDA2 IS GREATER THAN OR EQUAL TO LAMDA1. EXPLICITLY DETERMINE DELTA, THE INDEX OF NON-CONGRUITY.

RAT1=LAMDA2/LAMDA1
DIFF1=LAMDA2-LAMDA1
T1=ALOG(RAT1)/DIFF1
Z1=-LAMDA1*T1
Z2=-LAMDA2*T1
DELTA=2*(EXP(Z1)-EXP(Z2))
GO TO 121
LAMDA1 IS GREATER THAN LAMDA2. EXPLICITLY DETERMINE DELTA, THE INDEX OF NON-CONGRUITY.

RAT2=LAMDA1/LAMDA2
DIFF2=LAMDA1-LAMDA2
T2=ALOG(RAT2)/DIFF2
Z2=-LAMDA1*T2
Z1=-LAMDA2*T2
DELTA2=2*(EXP(Z2)-EXP(Z1))
GO TO 121

WRITE(I6,98)M, LAMDA1, LAMDA2
98 FORMAT(I16, //, I16, //, I16, //, I16, //, I16)
CALCULATE THE EXPECTED VALUE FOR CELL FREQUENCIES E1=M/10.

Determine equal probability regions for distributions 1 and 2. The cumulative probability associated with these regions for distribution 2.

X1(I1)=ALOG(.49)/LAMDA1
PZ1(I1)=1-EXP(-LAMDA2*X1(I1))
X2SUM=X2SUM+(E1-M*(P2(I1)-P2(I1-1)))**2
GO 1 I=2,9
X1(I1)=ALOG(.1)/LAMDA1
P2(I1)=1-EXP(-LAMDA2*X1(I1))
X2SUM=X2SUM+(E1-M*(P2(I1)-P2(I1-1)))**2
1 CONTINUE

C CALCULATE THE PSEUDO CHI-SQUARE STATISTIC
X2PS=X2SUM/E1

C GENERATE RANDOM SAMPLE FROM SECOND DISTRIBUTION
C READ RANDOM NUMBER GENERATOR SEED VALUES
READ(5,97)JSEED,JSEED

97 FORMAT(2I5)

C ECHO SEED VALUES
WRITE(6,96)JSEED,JSEED

96 FORMAT(3I9,'Seeds for run are',/4I9)

C INITIALIZE RANDOM NUMBER GENERATOR
CALL RSTART(JSEED,JSEED)

C GENERATE SAMPLE
DO 2 I=1,M
SAKPLI(I)=REXP(I1/LAMDA2)
CONTINUE

C SORT RANDCH OBSERVATIONS INTO FREQUENCY CLASSES
DO 100 I=1,M

1 CONTINUE
IF(SAMPI(I).LE.X1(1)) GO TO 101
IF(SAMPI(I).LE.X1(2)) GO TO 102
IF(SAMPI(I).LE.X1(3)) GO TO 103
IF(SAMPI(I).LE.X1(4)) GO TO 104
IF(SAMPI(I).LE.X1(5)) GO TO 105
IF(SAMPI(I).LE.X1(6)) GO TO 106
IF(SAMPI(I).LE.X1(7)) GO TO 107
IF(SAMPI(I).LE.X1(8)) GO TO 108
IF(SAMPI(I).LE.X1(9)) GO TO 109
K=10
GO TO 110

101 K=1
GO TO 110
102 K=2
GO TO 110
103 K=3
GO TO 110
104 K=4
GO TO 110
105 K=5
GO TO 110
106 K=6
GO TO 110
107 K=7
GO TO 110
108 K=8
GO TO 110
109 K=9
110 FRC0(K)=FRECIK)*1
CONTINUE

C CALCULATE THE ACTUAL CHI-SQUARE STATISTIC FOR RANDOM SAMPLE
DO 4 I=1,M
F2SUM=F2SUM+FREQ(I)**2
CONTINUE
X2ACT=F2SUM/E1-M

4 CONTINUE
C CALL FUNCTION TO CALCULATE "ALPHA HAT" FOR THE COMPUTED CHI-SQUARE
C VALUE
C OUTPUT VALUES OF DELTA, PSEUDO CHI-SQUARE, CHI-SQUARE, AND ALPHA HAT


WRITE(6,94)
WRITE(6,92)((XI(I),I=1,10))
WRITE(6,91)((PI(I),I=1,10))
WRITE(6,93)((SAMPLE(I),I=1,1))
WRITE(6,90)((REGION COUNTS,I=1,10))
WRITE(6,91)(SAMPL(I),I=1,1)

94 FORMAT(/11X,'RANDOM SAMPLE',10F11.4)
93 FORMAT('0','REGION COUNTS',10F11.4)
92 FORMAT('0','REGION BOUNDARIES',10F11.4)
91 FORMAT('0','CUMULATIVE PKCB',10F11.4)
90 FORMAT('0','CELL FREQUENCY',10F11.1)
GO TO 200
201 STOP
END
FUNCTION CADTRIX,G)

PURPOSE
COMPUTES P(X) = PROBABILITY THAT THE RANDOM VARIABLE U,
DISTRIBUTED ACCORDING TO THE CHI-SQUARE DISTRIBUTION WITH G
DEGREES OF FREEDOM, IS LESS THAN OR EQUAL TO X. F(G,X), THE
USAGE
PROP=CADTRIX,G)

DESCRIPTION OF PARAMETERS
X - INPUT SCALE FOR WHICH P(X) IS COMPUTED.
G - NUMBER OF DEGREES OF FREEDOM OF THE CHI-SQUARE DISTRIBUTION. G
IS A CONTINUOUS PARAMETER.
IER - RESULTANT ERROR CODE WHERE
IER= 0 --- NO ERROR
IER=-1 --- AN INPUT PARAMETER IS INVALID. X IS LESS THAN 0.0, OR G IS
LESS THAN 0.5 OR GREATER THAN 2*10**(5). P AND D ARE SET TO -1.E75.
IER=+1 --- INVALID OUTPUT. P IS LESS THAN ZERO OR GREATER THAN ONE.
SERIES FOR F(X) HAS FAILED TO CONVERGE. P IS SET TO 1.E75.

SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
DLGAH, ANDDLG
DLGAM, NOTR

DOUBLE PRECISION XX,QLXX,X2,OLX2,GG,G2,OLJ3,THETA,THPI,
1T1,1.SER,CC,XI,FAC,1LCG,TEPH,GA2,A8,C,DT2,GT3,THPI

TEST FOR VALID INPUT DATA
1F(G-1.5-1.E-5) 590,10,10
1F(G>2.E-5) 20,20,590
0 1F(X) 590,30,30

TEST FOR X NEAR 0.0
0 30 1F(G-1.E-6) 40,40,80
0 40 P=0.0
0 0008 1F(G-2.E-5) 50,60,70
0 0009 50 D=1.E75
0 0110 GD TO 610
0 0111 60 D=0.5
0 0112 GD TO 610
0 0113 70 GD=0.0
0 0114 GD TO 610

TEST FOR X GREATER THAN 1.E+6
80 1F(X-1.E+6) 100,100,90
90 D=0.0
0 0117 P=1.0
0 0118 GD TO 610

SET PROGRAM PARAMETERS
100 XX=08LE2(X)
FORTRAN IV G LEVEL 21

0020 DLXX=OLDG(XX)
0021 X2=X/2.D0
0022 DLX2=OLDG(X2)
0023 GG=DBLE(G)
0024 G2=GG/2.D0

C TEST FOR C GREATER THAN 1000.0
0025 IF(CG-1000.0) 160,160,180
0026 160 P=1.0
0027 170 GD TD 610
0028 180 A=OLDG(XX/GG)/3.00
0029 0030 A=DEXP(A)
0031 B=2.00/(5.00*GG)
0032 C=(A-1.00+B)/DSORT(B)
0033 SC=SNGL(C)
0034 CALL NDTR(SC,P,DUMMY)
0035 GD TD 490

C COMPUTE THETA
0036 K= HDIN1(G2)
0037 THEATA=G2-DFLOATK)
0038 IF(THEAT-1.D-6) 200,200,210
0039 200 THEAT=0.00
0040 210 THPL=THEAT+1.D0

C SELECT METHOD OF COMPUTING TI
0041 IF(THEAT1230,230,220
0042 220 IF(X=10.00)260,260,320
0043 CDPUTE TI FOR THEATA EQUALS 0.0
0044 230 IF(X-1.660D2) 250,249,240
0045 240 TI=1.0
0046 250 GD TD 400

C COMPUTE TI FOR THEATA GREATER THAN 0.0 AND X LESS THAN DR EQUAL TO 10.0
0047 260 SER=X2*(1.00/THPL-X2/(THPL+1.D0))
0048 270 CONTINUE

C
0049 J=+1
0050 CC=DFLOAT(J)
0051 OD 270 ITI=3,30
0052 X1=DFLOAT(11)
0053 CALL DLGAM(X1,FAC,10K)
0054 TLOG=X1*OLGX2-OLDG(X1*THEAT)
0055 TERM=DEXP(TLOG)
0056 T=SIGN(TERM,CC)
0057 SER=SER+TERM
0058 CC=-CC
0059 IF(DBL6(TERM)-1.D-9) 280,270,270
0060 270 CONTINUE
0061 GD TD 600
IF (SER) GO 600,600,290
CALL DLGAM(THP1,GTH,1CK)
IF (TLCG=1+6.D0) 300,300,310
T1=SNGL(T11)
GO TO 400
C
COMPUTE T1 FOR THETA GREATER THAN 0.0 AND
X GREATER THAN 10.0 AND LESS THAN 2000.0
C
A2=0.0
GO 340 I=1,25
XI=DFLOAT(I)
CALL DLGAM(THP1,GTH,1CK)
T11=-((13.0*XX)/XI +THP1*DLG(13.0*XX/XI) -GTH*DLG(XI)
IF (T11<1.68D02) 340,340,330
T1=0.0
GO TO 400
COMPUTE T1 FOR THETA GREATER THAN 0.0 AND
X GREATER THAN 10.0 AND LESS THAN 2000.0
C
A2=A2+T11
CONTINUE
A1=1.01282051*THETA/150.D0-XX/312.DO
B=DABS(A1)
C=-X2+THP1*DLX2*DLG(B)-GTH-3.951243718581427
IF (C>1.68D02) 370,370,350
C=0.0
GO TO 390
C=DEXP(C)
C=A2*C
T11=1.00*C
T1=SNGL(T11)
SELECT PROPER EXPRESSION FOR P
C
IF (G<2.0) 420,410,410
IF (G<4.0) 450,430,440
CALL CLGAM(THP1,GTH,1CK)
DT2=THE THP1-2*THP1*DLX2+DLG(8)-GTH-3.951243718581427
IF (DT2>1.68D02) 430,430,440
P=T1
GO TO 490
C
COMPUTE P FOR G GREATER THAN ZEROC AND LESS THAN 2.0
C
420 CALL DLGAM(THP1,GTH,1CK)
430 P=T1
GO TO 490
440 DT2=EXP(DT2)
T2=SNGL(DT2)
P=T1*T2+T2
GO TO 490
C
COMPUTE P FOR G GREATER THAN OR EQUAL TO 2.0
C
AND LESS THAN 4.0
C
450 P=T1
GO TO 490
10/03/40
C COMPUTE P FOR G GREATER THAN OR EQUAL TO 4.0
AND LESS THAN OR EQUAL TO 1000.0
C
0107  460 DT3=0.00
0108  GO 480, 13=2, K,
0109  THPI=FLOAT(13)*THETA
0110  CALL DLGAMIT(THPI,GTH,IOK)
0111  DLT3=THPI*DIX2-CLXX-X2-GTH
0112  IF(DLT3+1.6D02) 480, 480, 470
0113  470 DT3=DT3*EXP(DLT3)
0114  480 CONTINUE
0115  T3=SNGL(IO13)
0116  P=11-T3-T3
C SET ERROR INDICATOR
0117  490 IF(P) 500, 520, 520
0118  500 IF(ABS(P)-1.0E-7) 510, 510, 520
0119  510 P=0.0
0120  GO TO 610
0121  520 IF(1.0-P) 530, 550, 550
0122  530 IF(AABS(1.0-P)-1.0E-7) 540, 540, 560
0123  540 P=1.0
0124  GO TO 610
0125  550 IF(1.0-P) 560, 560, 570
0126  560 P=0.0
0127  GO TO 610
0128  570 IF(AABS(P)-1.0E-8) 580, 580, 580
0129  580 P=1.0
0130  GO TO 610
0131  590 IER=1
0132  D=1.0E75
0133  PA=1.0E75
0134  GO TO 620
0135  600 IER=1
0136  P=1.0E75
0137  GO TO 620
0138  610 IER=0
0139  620 CADTR=1.0-P
0140  IF(IER.EQ.1)PRINT 910
0141  910 FORMAT('D', IOX, 'FAILURE TO CONVERGE IN X-SQ FUNCTION')
0142  IF(IER.EQ.-1)PRINT 911
0143  911 FORMAT('D', IOX, 'INVALID INPUT TO X-SQ FUNCTION')
0144  RETURN
0145  END
SUBROUTINE NDTR(X,P,D)
  AX=ABS(X)
  T=1.0/(1.0+0.2316419*AX)
  D=0.3989423*EXP(-AX/2.0)
  P=1.0-0.385378*({1.0-0.3565638+T+0.02515517*10.3193815})
  IF(X.LT.0.0) P=1.0-P
RETURN
END
**The sample size equals**: 50

**Lambda1 equals**: 0.1000e 02  
**Lambda2 equals**: 0.1250e 02

**The seed values for this run are**

- **Iseed equals**: 62155  
- **Jseed equals**: 51562

**The value of the index of non-congruity (delta) equals**: 0.1638

**The value of the pseudo chi-square statistic equals**: 2.00

**The value of the chi-square statistic equals**: 12.40

**The area of the chi-square distribution to the right of the chi-square statistic (alpha) equals**: 0.1917

<table>
<thead>
<tr>
<th>Region Boundaries</th>
<th>0.0105</th>
<th>0.0223</th>
<th>0.0357</th>
<th>0.0511</th>
<th>0.0693</th>
<th>0.0916</th>
<th>0.1204</th>
<th>0.1609</th>
<th>0.2303</th>
<th>0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative Prob</td>
<td>0.1234</td>
<td>0.2434</td>
<td>0.3557</td>
<td>0.4719</td>
<td>0.5796</td>
<td>0.6819</td>
<td>0.7780</td>
<td>0.8663</td>
<td>0.9438</td>
<td>0.0</td>
</tr>
<tr>
<td>Random Sample 1</td>
<td>0.5302</td>
<td>0.0023</td>
<td>0.4790</td>
<td>0.3014</td>
<td>0.0083</td>
<td>0.0028</td>
<td>0.1542</td>
<td>0.0115</td>
<td>0.0736</td>
<td>0.1119</td>
</tr>
<tr>
<td>Random Sample 2</td>
<td>0.0037</td>
<td>0.3412</td>
<td>0.0602</td>
<td>0.0210</td>
<td>0.0250</td>
<td>0.0664</td>
<td>0.0686</td>
<td>0.2834</td>
<td>0.1306</td>
<td>0.0428</td>
</tr>
<tr>
<td>Random Sample 3</td>
<td>0.1768</td>
<td>0.0614</td>
<td>0.0193</td>
<td>0.0702</td>
<td>0.0275</td>
<td>0.0137</td>
<td>0.0085</td>
<td>0.1358</td>
<td>0.4015</td>
<td>0.0403</td>
</tr>
<tr>
<td>Random Sample 4</td>
<td>0.0212</td>
<td>0.0008</td>
<td>0.0145</td>
<td>0.3709</td>
<td>0.2125</td>
<td>0.0620</td>
<td>0.0501</td>
<td>0.0016</td>
<td>0.0265</td>
<td>0.1270</td>
</tr>
<tr>
<td>Random Sample 5</td>
<td>0.1136</td>
<td>0.0519</td>
<td>0.1323</td>
<td>0.1139</td>
<td>0.0296</td>
<td>0.0045</td>
<td>0.0603</td>
<td>0.0596</td>
<td>0.0573</td>
<td>0.0428</td>
</tr>
</tbody>
</table>

| Cell Frequency    | 7.0    | 6.0    | 4.0    | 3.0    | 10.0   | 3.0    | 2.0    | 6.0    | 2.0    | 7.0 |

79
APPENDIX 2
C PROGRAM TO EVALUATE "THE INDEX OF NON-CONGRUITY" FOR NORMAL DISTRIBUTIONS. THIS INDEX IS BASED ON THE AMOUNT OF NON-OVERLAPPING VALUES TO BE SUPPLIED TO THE PROGRAM ARE DISTRIBUTION PARAMETERS, SAMPLE SIZE, AND INITIAL VALUES FOR THE RNG.

C FORMAT FOR DATA CARDS:
M1, SIGMA1, M2, SIGMA2, M
C CASE, SIGMA2.
C MULTIPLE RUNS ARE POSSIBLE BY SUPPLYING ADDITIONAL INPUT CARDS.
C (TWO PER REPLICATION). PROGRAM COMPLETION IS INDICATED BY A BLANK CARD.

DIMENSION XII(10), P2(10), SAMPL(200), FREQ(10), 22P5(10)
REAL M1, M2, MU
C INITIALIZE PROGRAM PARAMETERS.

XSUM=0.
F2SUM=0.
DO 11 I=1, 10
11 FRED(I)=0.
C INPUT VALUES FOR THE PARAMETERS OF THE NORMAL DISTRIBUTIONS AND
C THE SAMPLE SIZE.
READ 99, MU1, SIGMA1, M2, SIGMA2, M
C CHECK FOR PROGRAM COMPLETION.
IF(MU1.EQ.0..AND.SIGMA1.EQ.0.) GO TO 303
C ECHO PARAMETER VALUES AND SAMPLE SIZE.
WRITE(6, 98) MU1, SIGMA1, M2, SIGMA2, M
C DETERMINE IF DISTRIBUTIONS ARE OVERLAPPING.
IF(MU2.LT.MU1) GO TO 21
IF(MU2+5.*SIGMA2.LT.MU1-5.*SIGMA1) GO TO 100
C DETERMINE THE NUMBER OF INTERSECTION POINTS USING THE QUADRATIC EQUATION.
IF(SIGMA1.EQ.SIGMA2) GO TO 1010
VAR2=SIGMA2*SIGMA2
VAR1=SIGMA1*SIGMA1
RATIO=SIGMA1/SIGMA2
SOM1=MU1*MU1
SOM2=MU2*MU2
VXM21=VAR2*MU1
VXM12=VAR1*MU2
VXM21-VXM12
VXMDIF=VXM21-VXM12
VXMD2=VXMDIF*VXMDIF
FRSFAC=4.*VXMD2
VARDIF=VAR2-VAR1
VXS21=VAR2*SOM1
VXS12=VAR1*SOM2
XLFAC=2.*VAR1*VAR2*ALOGRATIO
SNOFAC=4.*VARDIF*VXS21-VXS12*XLFAC
RAD=FRSFAC-SNOFAC
IF(RAD) 100, 101, 102
C THE DISTRIBUTIONS ARE NON-OVERLAPPING. THERE ARE NO INTERSECTION POINTS.
100 WRITE(6, 900)
C\ There is one intersection point

1010 XX=(MU1+MU2)/2.
1011 Z1=(XX-MU1)/SIGMA1
Z2=(XX-MU2)/SIGMA2
CALL NDTR(Z1,F1,0)
CALL NDTR(Z2,F2,0)
DELTA=2.*ABS(F1-F2)
1047
C\ There are two intersection points

102 XX=(VXM21-VXM11)/VARDIF
RADPRT=SQRT(RAD)/2.*VARDIF
XX1=XX+RADPRT
XX2=XX-RADPRT
IF(XX1.LE.XX2) GO TO 1021
C1=XX2
C2=XX1
GO TO 1022
1021 C1=XX1
C2=XX2
Z1C1=(C1-MU1)/SIGMA1
Z2C1=(C1-MU2)/SIGMA2
Z1C2=(C2-MU1)/SIGMA1
Z2C2=(C2-MU2)/SIGMA2
CALL NDTR(Z1C1,F1C1,D)
CALL NDTR(Z1C2,F1C2,D)
CALL NDTR(Z2C1,F2C1,D)
CALL NDTR(Z2C2,F2C2,D)
F2DIF=F2C2-F2C1
F1DIF=F1C2-F1C1
DELTA=2.*ABS(F2DIF-F1DIF)
C\ Determine values for class boundaries for equal probability regions
C\ Of model distribution (Distribution 1)

103 DO 1 I=1,9
P=1.*I
X1(I)=MU1*SIGMA1*Z
1
C\ Determine cumulative probabilities for class boundaries for alternative
C\ distribution (Distribution 2)

105 DO 2 I=1,9
Z2PS(I)=(X1(I)-MU2)/SIGMA2
2 CALL NDTR(Z2PS(I),P2(I),D)
C\ Calculate the expected value for cell frequencies

107 E1=M/10.
C\ Calculate the pseudo chi-square statistic

108 X2SUM=(E1-M*(P2(I)-P2(I-1)))*2
109 DO 3 I=2,9
X2PS=X2SUM+(E1-M*(P2(I)-P2(I-1)))*2
3 X2PS=X2SUM+(E1-M*(P2(I)-P2(I-1)))*2

C\ Read random number generator seed values

READ(5,97)JSEED,JSEED
C

C ECHO SEED VALUES
WRITE(*,96) ISEED, JSEED
C INITIALIZE RANDOM NUMBER GENERATOR
CALL RSTART(ISEED, JSEED)
C GENERATE RANDOM SAMPLE FROM SECOND DISTRIBUTION
DO 4 I=1, M
4 SAMPL(I)=MUZ*SIGM2*RNOR(N)
C SORT RANDOM OBSERVATIONS INTO FREQUENCY CLASSES
DO 5 I=1, M
IF(SAMPL(I).LE.X1I5) GO TO 201
IF(SAMPL(I).LE.X1I6) GO TO 206
IF(SAMPL(I).LE.X1I7) GO TO 207
IF(SAMPL(I).LE.X1I8) GO TO 208
IF(SAMPL(I).LE.X1I9) GO TO 209
K=10
GO TO 5
201 IF(SAMPL(I).GT.X1I4) GO TO 205
206 IF(SAMPL(I).GT.X1I3) GO TO 204
207 IF(SAMPL(I).GT.X1I2) GO TO 203
208 K=6
GO TO 5
209 K=9
5 FREQ(K)=FREQ(K)+1.
C CALCULATE THE ACTUAL CHI-SQUARE STATISTIC FOR RANDOM SAMPLE
DO 6 I=1,10
6 F2SUM=F2SUM+FREQ(I)**2
X2ACT=F2SUM/E1-M
C CALCULATE "ALPHA HAT" FOR THE COMPUTED CHI-SQUARE VALUE
AHAT=CAOTR2XACT,NU
C OUTPUT VALUES OF DELTA, PSEUDO CHI-SQUARE, CHI-SQUARE, AND ALPHA HAT
WRITE(*,93) DELTA, X2PS, X2ACT, AHAT
C OUTPUT VALUES OF THE INDEX OF NON-CONGRUITY (DELTA) E
95 FORMAT(*), I11, 'THE VALUE OF THE INDEX OF NON-CONGRUITY (DELTA) E
1 QUALS', F12.4, '/*/*, I11, 'THE VALUE OF THE PSEUDO CHI-SQUARE STATISTIC 
2 QUALS', F12.2, '/*/*, I11, 'THE VALUE OF THE CHI-SQUARE STATISTIC EQUALS 
3 F12.2, '/*/*, I11, 'THE AREA OF THE CHI-SQUARE DISTIBUTION TO THE RIG 
4 HT OF THE CHI-SQUARE STATISTIC (ALPHA HAT) QUALS', F12.4
C OUTPUT VALUES OF INTERSECTION POINTS OR POINTS
WRITE(*,94)
1 IF(IND.EQ.1) GO TO 301
FORTRAN IV G LEVEL 21

0130 WRITE(6,920) C1,C2
0131 920 FORMAT(1X,'THERE ARE TWO POINTS OF INTERSECTION. THEY ARE',F10.2,5
0132 1X,'AND',F10.2)
0133 GO TO 302
0134 301 WRITE(6,910) XX
0135 C OUTPUT VALUES OF XI, P2, SAMPL, AND FREQ
0136 910 FORMAT(1X,'THERE IS ONE POINT OF INTERSECTION. IT IS',F10.2)
0137 302 WRITE(6,941)
0138 941 FORMAT(1X,3F10.2,F15.5)
0139 WRITE(6,911) P2(1),XI(1),XI(2)
0139 WRITE(6,911) P2(1),XI(1),XI(2)
0140 911 FORMAT(1X,3F10.2,F15.5)
0141 GO TO 300
0142 94 FORMAT(1X,'RANDOM SAMPLE',10F11.41)
0143 92 FORMAT(1X,'REGION BOUNDARIES',9F11.41)
0144 91 FORMAT(1X,'CUMULATIVE PROB',9F11.41)
0145 90 FORMAT(1X,'CELL FREQUENCY',10F11.11)
0146 94 FORMAT(1X,'REGION BOUNDARIES',9F11.41)
0147 300 STOP
0148 ENDO
SUBROUTINE NDTRI(P, X, D, IE)

IE=0
X=.99999E+74
D=X

1 IE=-1
GO TO 12

2 IF(P>1.0) 7, 5, 1
4 X=-.99999E+74
5 D=0.0
GO TO 12

7 D=P
10 IF(P<0.5) 9, 8
11 D=1.0-1
12 T2=ALOG(1.0/(D*D1))
13 T=SQRT(T2)
14 X=T-{2.515517+0.802853*T+0.010328*T2+1.0*1.432788*T+0.189269*T2+10.001300*T*T2)/10.001300*T+12)
15 IF(P<0.5) 10, 10, 11

10 X=-X
11 D=0.3989423*EXP(-X*X/2.0)
12 RETURN
END
FUNCTION CADTRI(G)

PURPOSE
COMPUTES P(X) = PROBABILITY THAT THE RANCCM VARIABLE U,
DISTRIBUTED ACCORDING TO THE CHI-SQUARE DISTRIBUTION WITH G
DEGREES OF FREEDOM, IS LESS THAN OR EQUAL TO X. P(X), THE CDF
USAGE
PKCB=CADTRI(G),

DESCRIPTION OF PARAMETERS
X = INPUT SCALE FOR WHICH P(X) IS COMPUTED.
G = NUMBER OF DEGREES OF FREEDOM OF THE CHI-SQUARE
DISTRIBUTION. G IS A CONTINUOUS PARAMETER.
IER = RESULTANT ERROR CODE WHERE
IER=0 --- NO ERROR
IER=-1 --- AN INPUT PARAMETER IS INVALID. X IS LESS
THAN 0.0, OR G IS LESS THAN 0.5 OR GREATER
IER=-1 --- INVALID OUTPUT. P IS LESS THAN ZERO OR
GREATER THAN ONE. P AND G ARE SET TO -1.E75.
SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
DLGAM, NDTR, DOUBLE PRECISION XX, DLXX, X2, DLX2, G2, DLT3, THETA, THPI,
1T1, SER, CC, XI, FAC, TLSG, TERM, 6TH, A2, A, B, C, E12, DT3, THPI

TEST FOR VALID INPUT DATA
0000 IF(G-(.5-1.E-5)) 500, 10, 10
0001 10 IF(G-2.E+5) 20, 20, 500
0002 20 IF(X) 500, 30, 30

TEST FOR X NEAR 0.0
0003 IF((G-1.5-1.E-5)) 590, 10, 10
0004 IF((G-2.5-1.E-5)) 20, 20, 590
0005 20 IF(X) 590, 30, 30

TEST FOR X GREATER THAN 1.E+6
0006 IF(X-1.E+8) 40, 40, 80
0007 40 P=0.0
0008 IF(G-2.1) 50, 60, 70
0009 50 G=1.E75
0010 GO TO 610
0011 60 G=0.5
0012 GO TO 610
0013 70 G=0.0
0014 GO TO 610

TEST FOR X GREATER THAN 1.E+6
0015 IF(X-1.E+6) 100, 100, 60
0016 100 P=1.0
0017 GO TO 610

SET PROGRAM PARAMETERS
0018

100 XX=0B/4E(X)

0019
FORTRAN IV LEVEL 21

0020 OLXX=OLG(XX) Cru
0021 X2=XX/2.00
0022 CLXX=OLG(X2) Cru
0023 GG=DBLE(G) Cru
0024 G2=GG/2.00
0025 TEST FOR G GREATER THAN 1000.0 Cru
0026 TEST FOR X GREATER THAN 2000.0 Cru
0027 IF(G<1000.) 160,160,180
0028 IF(X<2000.) 190,190,170
0029 GC TO 610
0030 A=OLG(X2/GG)/3.00
0031 B=2.00/49.00*GG
0032 C=(A-1.00+1)/0.50RT(B)
0033 SC=SNGL(C)
0034 CALL NTR(SC,P,DUMMY)
0035 GO TO 490
0036 COMPUTE THETA
0037 190 K= I0INTG(G)
0038 IF(THETA-1.0<0) 200,200,210
0039 200 THETA=0.00
0040 210 THP1=THETA+1.00
0041 SELECT METHOD OF COMPUTING T1
0042 IF(THETA<230,230,220
0043 220 IF(X<1.68002) 250,240,240
0044 240 T1=1.00
0045 GO TO 400
0046 250 T1=-1.00-OLG(-X2)
0047 260 T1=SNGL(T11)
0048 GO TO 400
0049 COMPUTE T1 FOR THETA GREATER THAN 0.0 AND
0050 X LESS THAN OR EQUAL TO 10.0
0051 260 SER=X2*(1.00/THP1-X2/(THP1+1.00))
0052 JC=1
0053 CC=OFLGAT(J)
0054 GO TO 320 I11=3.30
0055 XI=OFLGAT(I11)
0056 CALL OLDA(XI,FAC,10K)
0057 TLOG= XI*OLX2-OLG(XI*THETA)
0058 TERM=OLG(TLOG)
0059 TER=SIGN(TERM,CC)
0060 SER=SER*TERM
0061 CC=CC
0062 IF(ISABS(TERM)-1.0<91) 280,270,270
0063 CONTINUE
0064 GO TO 600
COMPUTE P FOR G GREATER THAN OR EQUAL TO 4.0
AND LESS THAN OR EQUAL TO 1000.0

0107 460 DT3=0.00
0108 480 DO 13=2,K
0109 510 THPI=DFLOAT(THPI,1GK)
0110 530 DLT3=THPI*OLX2-CLXX-2*GTH
0111 540 IF(DLT3+1.68D02) 480,460,470
0112 560 DT3=DT3+DEXP(DLT3)
0113 580 CONTINUE
0114 600 T3=SNGL(DT3)
0115 620 P=T1-T3-T3

C SET ERROR INDICATOR

0117 640 IF(P) 500,520,520
0118 550 IF(SQRT(P)=1.0) 510,510,600
0119 570 P=0.0
0120 590 GO TO 610
0121 610 IF(1-P) 530,550,550
0122 530 IF(SQRT(1-P)=1.0) 540,540,600
0123 550 P=1.0
0124 570 GO TO 610
0125 590 IF(P=1.0) 560,560,570
0126 610 P=0.0
0127 630 GO TO 610
0128 650 IF((1.0-P)-1.E-8) 580,580,610
0129 670 P=1.0
0130 690 GO TO 610
0131 710 IER=-1
0132 730 D=1.E-75
0133 750 P=1.E-75
0134 770 GO TO 620
0135 790 IER=0
0136 810 P=1.E-75
0137 830 GO TO 620
0138 850 IER=0
0139 870 IER=1
0140 890 IER=2
0141 910 FORMAT('00',10X,'FAILURE TO CONVERGE IN X-SQ FUNCTION')
0142 930 IF(IER.EQ.-1)PRINT 911
0143 950 IF(IER.EQ.-1)PRINT 911
0144 970 FORMAT('00',10X,'INVALID INPUT TO X-SQ FUNCTION')
0145 990 RETURN
END
SUBROUTINE NDTR(X,P,D)

AX=ABS(X)

T=1.0/(1.0+0.2316419*AX)

D=0.3989423*EXP(-X*X/2.0)

P=1.0-D*T*(((1.330274*T-1.821256)*T+1.781478)*T-0.2565368)*T*

RETURN

END
SUBROUTINE NDIR(X,P,D)
  AX=ABS(X)
  T=1.0/(1.0+0.2316419*AX)
  D=0.3989423*EXP(-AX*AX/2.0)
  P=1.0-D*T*(1.1*(1.330274*T-1.821256)*T+1.781878)*T-0.3565638*T+
  0.3153815
  IF(X.LT.0.0) P=1.0-P
RETURN
END
MEAN AND STANDARD DEVIATION OF DISTRIBUTION 1: 0.0 0.1000E 01
MEAN AND STANDARD DEVIATION OF DISTRIBUTION 2: 0.3000E 00 0.1200E 01
THE SAMPLE SIZE EQUALS: 50

THE SEED VALUES FOR THIS RUN ARE

ISEED EQUALS: 62155
JSEED EQUALS: 51562

THE VALUE OF THE INDEX OF NON-CONGRUITY (DELTAI) EQUALS: 0.2648
THE VALUE OF THE PSEUDO CHI-SQUARE STATISTIC EQUALS: 7.16
THE VALUE OF THE CHI-SQUARE STATISTIC EQUALS: 4.40
THE AREA OF THE CHI-SQUARE DISTRIBUTION TO THE RIGHT OF THE CHI-SQUARE STATISTIC (ALPHA HAT) EQUALS: 0.6993

THERE ARE TWO POINTS OF INTERSECTION. THEY ARE: 2.05 AND 0.68

<table>
<thead>
<tr>
<th>REGION BOUNDARIES</th>
<th>-1.2817</th>
<th>-0.8415</th>
<th>-0.5240</th>
<th>-0.2529</th>
<th>0.0</th>
<th>0.2529</th>
<th>0.5240</th>
<th>0.8415</th>
<th>1.2817</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUMULATIVE PROB</td>
<td>0.0937</td>
<td>0.1707</td>
<td>0.2661</td>
<td>0.3225</td>
<td>0.4013</td>
<td>0.4844</td>
<td>0.5740</td>
<td>0.6741</td>
<td>0.7934</td>
</tr>
<tr>
<td>RANDOM SAMPLE</td>
<td>0.7770</td>
<td>4.3540</td>
<td>-1.4125</td>
<td>0.4722</td>
<td>-0.2033</td>
<td>-0.7779</td>
<td>-0.3549</td>
<td>1.0924</td>
<td>-0.5331</td>
</tr>
<tr>
<td>RANDOM SAMPLE</td>
<td>0.1506</td>
<td>-0.8829</td>
<td>-0.8850</td>
<td>-0.6316</td>
<td>1.1461</td>
<td>0.7400</td>
<td>-2.2633</td>
<td>0.8774</td>
<td>0.0849</td>
</tr>
<tr>
<td>RANDOM SAMPLE</td>
<td>0.2298</td>
<td>-0.0376</td>
<td>1.5552</td>
<td>0.3529</td>
<td>3.2375</td>
<td>3.7394</td>
<td>2.4928</td>
<td>0.6124</td>
<td>1.6426</td>
</tr>
<tr>
<td>RANDOM SAMPLE</td>
<td>-1.0056</td>
<td>-0.7294</td>
<td>0.9289</td>
<td>-1.0840</td>
<td>-1.0789</td>
<td>0.0053</td>
<td>0.5922</td>
<td>0.9800</td>
<td>0.0062</td>
</tr>
<tr>
<td>RANDOM SAMPLE</td>
<td>0.8669</td>
<td>0.7669</td>
<td>-0.7807</td>
<td>1.6423</td>
<td>-0.1953</td>
<td>1.1252</td>
<td>-2.8222</td>
<td>1.5932</td>
<td>-2.0086</td>
</tr>
<tr>
<td>CELL FREQUENCY</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>3.0</td>
<td>3.0</td>
<td>6.0</td>
<td>2.0</td>
<td>6.0</td>
<td>7.0</td>
</tr>
</tbody>
</table>
NON-CONGRUENCE OF STATISTICAL DISTRIBUTIONS:
HOW DIFFERENT IS DIFFERENT?

by

TERRY LEE APPLEGATE

B.S., Kansas State University, 1977

---

AN ABSTRACT OF A MASTER'S THESIS

Submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering
KANSAS STATE UNIVERSITY
Manhattan, Kansas

1978
This thesis studies differences between statistical distributions of the same family. In particular it studies members of the exponential and normal families of statistical distributions. In theory two distributions are the same only if their probability density functions are identical (which implies that their parameters are identical also). However, in practical situations, two distributions which have closely similar probability density functions may produce random samples of small size which are indistinguishable from one another. This thesis is concerned with studying this situation in an attempt to better understand the question of "How different is different?" in relation to differences in statistical distributions from the same family.

The methodology used to study the difference between a pair of statistical distributions from the same family consists of a number of steps. The first step in the comparison procedure consists of determining the amount of non-overlapping area bounded by the probability density functions of the two distributions being compared. The second step consists of drawing a "perfect" sample from one distribution and comparing it with the other distribution. The third step consists of drawing a random sample from one distribution and comparing it with the other distribution. The final step consists of calculating certain indices from the parameters of the distributions and relating these indices to the other comparison results.

Results of the comparison procedure for a sample size of 50 indicate that in both the exponential case and the normal case statistical significant differences at the .05 level would be indicated for amounts of non-overlapping area in excess of a threshold value occurring somewhere in the region of .3 to .6. In addition to this there appears to be strong relationships between the indices derived from the parameters of the distributions being compared and the various other comparison indices.
These strong relationships would allow the comparison of statistical distributions solely on the basis of their parameters without requiring the use of sampling.

Special topics covered in the study which might be of interest to other researchers are the use of the McGill Random Number Generator developed by members of the School of Computer Science of McGill University and the suggestions for further research in this area.