12-1978

A Computer Simulation Study of Seven Pairwise Comparison Procedures: Imbalanced Case

Wayne S. Petroelje
Western Michigan University

A Computer Simulation Study of Seven Pairwise
Comparison Procedures: Imbalanced Case

by

Wayne S. Petroelje

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Doctor of Education

Western Michigan University
Kalamazoo, Michigan
December 1978

ACKNOWLEDGEMENTS

It is a pleasure to express my appreciation to the people who helped bring this dissertation and doctoral program to completion.

To Mary Anne Bunda, my advisor, thanks for her suggestions and guidance.

To Uldis Smidchens, my chairperson, thanks for his inspiration, help, and support.

To Michael Stoline, my third committee member, thanks for his advice, patience, and time.

To Patricia Klimp, thanks for her time in typing.

To Barbara Petroelje, a very special thanks for her comfort, encouragement, and love during my graduate program.

To Patricia, Gregory, Michael, and Michelle Petroelje, a very special thanks to them for being patient with me as a father.

Wayne S. Petroelje

ii

INFORMATION TO USERS

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

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

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

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

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

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

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

7907960

PETROELJE, WAYNE S.
A COMPUTER SIMULATION STUDY OF SEVEN PAIRWISE
COMPARISON PROCEDURES: IMBALANCED CASE.

WESTERN MICHIGAN UNIVERSITY, ED.D., 1979

University
  Microfilms
International   300 N. ZEEB ROAD, ANN ARBOR, MI 48106

TABLE OF CONTENTS

iii

LIST OF TABLES

iv

INTRODUCTION

Numerous pairwise comparison procedures are available to determine if any differences are present among the population means $\mu_1, \ldots, \mu_k$. A common approach taken is to determine which pairs of means are not equal. For this approach the hypothesis $H_0:\mu_i=\mu_j$, is tested for each of the $k(k-1)/2$ pairs of means (μ_i,μ_j) from $\mu_1, \ldots, \mu_k$. In each case the alternative hypothesis is $H_1:\mu_i\neq\mu_j$. With any particular pairwise comparison procedure, if the absolute value of the difference between two sample means is greater than the product of the appropriate critical value and the corresponding standard error of the difference between means, the two corresponding population means are declared to be different; otherwise, the null hypothesis of no difference between the population means is not rejected. Critical value is the value of a statistic that corresponds to a given significance level as determined from its sampling distribution and standard error of the difference between means is the standard deviation of the sampling distribution of the difference of the two involved population means. Because the size of the critical values and the standard errors of the difference between means varies among procedures, results obtained from the application of several different procedures to a given set of data will often differ.

In testing each null hypothesis $H_0:\mu_i=\mu_j$, a correct or an incorrect

1

decision may be made. If an incorrect decision is made, one of two types of errors is committed. A Type I error occurs when the null hypothesis H_0 is falsely rejected. The significance level α is the probability of committing a Type I error. The second error one may make is a Type II error, which occurs when one fails to reject the H_0 when it is false. The probability of making a Type II error is designated by β; whereas, $1-\beta$ is defined as power, which is the probability of accepting a true alternate hypothesis.

The literature contains several computer simulation studies comparing pairwise comparison procedures. Usually these studies have examined which of the numerous pairwise comparison procedures best control for Type I error rate.

The conceptual unit used for the Type I error rate in these studies is called the experimentwise error rate. Kirk (1968) defines the experimentwise error rate (EER) as the probability that one or more Type I errors concerning the entire class of $k(k-1)/2$ pairwise comparisons of means will be made in a given experiment involving the k population means.

$$EER = \frac{\text{number of experiments with at least one statement falsely declared significant}}{\text{total number of experiments}} \qquad (1.1)$$

This error rate is used to guard against making any erroneous statements concerning the pairwise means in an experiment and is based on the premise that it is as serious to make one erroneous statement in an experiment as it is to make, say, four erroneous statements.

It may help to understand the meaning of the experimentwise error rate with an example. Suppose that there are 1,000 experiments, each

with 10 statements of significance, 10,000 statements in all. Of these statements, 90 are actually false, and these false statements are distributed among 70 of the experiments. Then the experimentwise error rate is equal to 70/1000 = .07.

One of the most comprehensive simulation studies for equal sample size was conducted and reported by Carmer and Swanson (1973). In this study ten pairwise comparison procedures were compared as to their experimentwise Type I error rates and power.

The ten compared by Carmer and Swanson were: Fisher's (1949) least significance difference procedure, Scheffé's (1953) S procedure, Tukey's (1953) T procedure, two Waller-Duncan Bayesian procedures established by Waller and Duncan (1969), Newman-Keuls procedure first proposed by Newman (1939), and three protected least significance difference procedures established by Fisher (1949). As a result of this study, one of the protected least significance difference (PLSD) was recommended.

The recommended PLSD is a two stage procedure. The first stage consists of performing an overall test of the null hypothesis $H_0: \mu_1 = \ldots \mu_k$ by means of the F ratio in a one way analysis of variance. If the F ratio is significant, the second stage consists of a procedure analogous to the t test and is used to make all $k(k-1)/2$ pairwise comparisons of means (μ_i, μ_j) from $\mu_1, \ldots, \mu_k$ at the same nominal significance level as that for the F ratio. However, if the F ratio is not significant, no pairwise comparisons of means are made, thus eliminating the possibility of committing Type I errors.

One of the most recent and involved simulation studies with unequal sample sizes has been conducted by Keselman and Rogan (1978). In

this study, Kramer's (1956) K procedure, Scheffé's (1953) S procedure, Spjøtvoll and Stoline's (1973) T' procedure, Hochberg's (1976) T" procedure, and Games and Howell's (1976) GH procedure were compared. These procedures were compared as to their experimentwise Type I error rates and power to varying degrees of sample sizes and variance heterogeneity. The sampling for this study was from both normal and skewed distributions. As a result of this study, the GH procedure was recommended.

The GH is a single stage procedure in which the hypothesis $H_0: \mu_i = \mu_j$ $i \neq j$, is tested for each of the $k(k-1)/2$ pairs of means (μ_i, μ_j) from $\mu_1, \ldots, \mu_k$. The standard error of the difference between means for the GH procedure uses the variances and sample sizes of the samples involved in the contrast. The ordinary studentized ranges are used in calculating the critical values for the GH procedure. This procedure is discussed in greater detail in Chapter II.

A serious weakness of the Keselman and Rogan study was that it did not include the PLSD procedure. This procedure was recommended by Carmer and Swanson with equal sample size. However, the PLSD procedure can be legitimately used with unequal sample sizes. Thus a comprehensive comparison of all potential useful pairwise comparison procedures has not been accomplished with unequal sample sizes.

In order to correct this omission, this study utilizes computer simulation techniques in order to study experimentwise Type I error rates for seven pairwise comparison procedures. The procedures are: protected least significance difference (PLSD) procedure, protected Welch (PW) procedure which is proposed in this study, Spjøtvoll and Stoline's T' procedure, Hochberg's T" procedure, Kramer's K procedure,

Games and Howell's GH procedure, and Tamhane's (1977) TH procedure. Specifications for these pairwise comparison procedures are contained in Chapter II.

CHAPTER II

RELATED LITERATURE

The literature contains numerous pairwise comparison procedures. These procedures are designed for testing the null hypothesis $H_0: \mu_i = \mu_j$. This chapter first contains the specifications of the pairwise comparison procedures that are used in this study. The next three sections review single stage procedures for both equal and unequal sample sizes, review two stage procedures, and discuss related simulation studies that compare these procedures. Finally, a summary to this chapter is presented.

Specifications of the Pairwise Comparison Procedures being Studied

This section presents the definitions to the seven pairwise comparison procedures along with each of their critical values and standard error of the difference between means. The next two sections, immediately following the present section, compares and categorizes the seven pairwise comparison procedures along with several other comparison procedures.

Let X_{ij} represent the ith observation in the jth group, where $i=1,\ldots,n_j$ and $j=1,\ldots,k$. The X_{ij}'s are independent normal variates with expected value μ_j and constant population variance σ^2. The estimate of μ_j is the sample mean

$$\overline{X}_j = \sum_i X_{ij}/n_j. \tag{2.1}$$

The jth sample variance is

6

$$s_j^2 = \sum_i (X_{ij} - \overline{X}_j)^2 / (n_j - 1) \qquad (2.2)$$

which estimates σ^2.

For the pairwise comparison procedures that use a pooled estimate of the population variance, the estimate is called the mean square within and is

$$MSW = \sum_j \sum_i (X_{ij} - \overline{X}_j)^2 / (N-k) \qquad (2.3)$$

where $N = \sum_j n_j$.

Of the seven pairwise comparison procedures being compared in this study, the PLSD and PW are two stage procedures and the remaining five are single stage procedures. For the two stage procedures, a preliminary significance test is based on the observed F ratio of the mean square between divided by the mean square within. This F ratio is used to test the null hypothesis $H_0: \mu_1 = \ldots = \mu_k$. The mean square between is

$$MSB = \sum_j n_j (\overline{X}_j - \overline{X}..)^2 / (k-1) \qquad (2.4)$$

where

$$\overline{X}.. = \sum_j n_j \overline{X}_j / N \qquad (2.5)$$

is the grand mean.

The test of the null hypothesis is made by comparing the obtained F ratio with a value $F_{(\alpha, k-1, N-k)}$ obtained from an F table with k-1 and N-k degrees of freedom, which is the value exceeded by 100α percent of all F ratios obtained under the assumption that the null hypothesis is true. If the obtained F ratio does not exceed the tabled value

$F_{(\alpha,k-1,N-k)}$, then the null hypothesis is not rejected and hence no pairwise comparisons of means are made. However, if the obtained F ratio exceeds the tabled value, then the null hypothesis is rejected and attention is directed to each of the $k(k-1)/2$ pairs of means (μ_i, μ_j) from $\mu_1, \ldots, \mu_k$ in order to attempt to find the particular unequal pairs of means that caused the null hypothesis to be rejected.

The five single stage and the second stage of the PLSD and PW procedures differ either in their standard error of the difference between the two sample means, and/or their critical values. If $|\overline{X}_i - \overline{X}_j|$ is greater than the product of the standard error of the difference between means and the corresponding critical value, then the null hypothesis $H_0:\mu_i=\mu_j$ is rejected. The standard error of the difference between means and critical value for these procedures in this study are described next.

Spjøtvoll and Stoline's T' procedure

In the T' procedure the standard error of the difference between means is

$$MSW^{\frac{1}{2}}[\max((1/n_i)^{\frac{1}{2}},(1/n_j)^{\frac{1}{2}})] \tag{2.6}$$

and the critical value is $q'_{(\alpha,k,N-k)}$ which represents the value exceeded by 100α percent of the studentized augmented range distribution, computed by Stoline (1978), with degrees of freedom k and $N-k$.

Kramer's K procedure

In the K procedure the standard error of the difference between

means is

$$(MSW/2)^{\frac{1}{2}}(1/n_i+1/n_j)^{\frac{1}{2}} \tag{2.7}$$

and the critical value is $q_{(\alpha,k,N-k)}$ which represents the value exceeded by 100α percent of the ordinary studentized range distribution with degrees of freedom k and N-k.

Hochberg's T" procedure

In the T" procedure the standard error of the difference between means is

$$\max\,(s_i/n_i^{\frac{1}{2}},s_j/n_j^{\frac{1}{2}}) \tag{2.8}$$

and the critical value is $q'_{(\alpha,k,g)}$ which represents the value exceeded by 100α percent of the studentized augmented range distribution with degrees of freedom k and $g=\min(n_i-1)$ for $i=1,\dots,k$.

Games and Howell's GH procedure

In the GH procedure the standard error of the difference between means is

$$(1/2)^{\frac{1}{2}}(s_i^2/n_i+s_j^2/n_j)^{\frac{1}{2}} \tag{2.9}$$

and the critical value is $q_{(\alpha,k,h)}$ which represents the value exceeded by 100α percent of the ordinary studentized range distribution with degrees of freedom k and Welch's (1948) approximate estimate for the degrees of freedom

$$h=(s_i^2/n_i+s_j^2/n_j)/[(s_i^2/n_i)^2/(n_i-1)+(s_j^2/n_j)^2/(n_j-1)]. \tag{2.10}$$

Tamhane's TH procedure

In the TH procedure the standard error of the difference between

means is

$$(s_i^2/n_i + s_j^2/n_j)^{\frac{1}{2}} \tag{2.11}$$

and the critical value is $t_{(\alpha', h)}$ which represents the value exceeded by $100\alpha'$ percent of the student's t distribution with

$$\alpha' = \frac{1}{2}[1-(1-\alpha)^{1/\binom{k}{2}}] \tag{2.12}$$

and Welch's approximate estimate for the degrees of freedom h as defined in equation (2.10).

PLSD procedure

In the PLSD procedure the standard error of the difference between means is

$$[MSW(1/n_i + 1/n_j)]^{\frac{1}{2}} \tag{2.13}$$

only if the F ratio is significant at level α and the critical value is $t_{(\alpha/2, N-k)}$ which represents the value exceeded by $100(\alpha/2)$ percent of the Student's t distribution with degrees of freedom N-k.

PW procedure

In the PW procedure the standard error of the difference between means is

$$(s_i^2/n_i + s_j^2/n_j)^{\frac{1}{2}} \tag{2.14}$$

only if the F ratio is significant at level α and the critical value is $t_{(\alpha/2, h)}$ which represents the value exceeded by $100(\alpha/2)$ percent of the student's t distribution with Welch's approximate estimate for the degrees of freedom h as defined in equation (2.10).

The K and PLSD procedures use a common estimate of the error variance and the sample sizes involved in the contrast; they differ

in their critical values and in the fact that K is a single stage procedure and the PLSD is a two stage procedure. The T' procedure also uses a pooled estimate of variance but substitutes the smaller of the two sample sizes for each n_i. The GH, TH, and PW procedures do not use a pooled estimate of variance. They use the variances and sample sizes involved in the contrast. The GH and TH are both single stage procedures, while PW is a two stage procedure. The single stage procedure T" also uses the variances and sample sizes of the groups involved in the contrast, but substitutes the larger of the two standard error estimates of $\overline{X}_i$ and $\overline{X}_j$.

This completes the section that specifies the standard errors and the critical values for each of the seven procedures included in this study. The next two sections contain a further discussion of these seven pairwise comparison procedures along with other frequently mentioned comparison procedures.

Single Stage Pairwise Comparison Procedures

Single stage procedures determine which pairs of population means are different without first using an over all test of significance. First, single stage procedures that require equal sample sizes are discussed and secondly, single stage procedures that do not require equal sample sizes are discussed.

Equal sample size

The following multiple comparison procedures are those most commonly used in the case of equal sample size. Each of these procedures

utilize the studentized range distribution. The first procedure discussed is Tukey's T procedure. This procedure is designed for testing all $k(k-1)/2$ null hypotheses among the pairwise differences of the k population means. Critical values for this procedure are based upon the ordinary studentized range distribution. The assumptions of normality and homogeneity of variance are required for the T procedure.

A somewhat different approach to pairwise comparisons is adopted with the Newman-Keuls procedure. This procedure is based on a step-wise approach. The critical values used for differences between means for this procedure vary depending on the number of means observed between the two means under consideration. Another procedure, which is similar to the Newman-Keuls procedure is the Duncan (1955) procedure. Duncan's significance level is dependent on the number of means under consideration. Special studentized ranges tabulated by Duncan (1955) are used for this procedure; whereas, the ordinary studentized range tables are used for the Newman-Keuls procedure.

The exact use of these three procedures is limited to those cases with an equal number of observations for each sample. The next section reviews procedures that do not require equal sample sizes.

Unequal sample sizes

Each of these procedures are for the most part based on Tukey's T procedure. These procedures are divided into two types. The two types are approximate procedures and conservative procedures.

Approximate procedures. Each of these pairwise comparison procedures is an approximation of the T procedure and may be used for the

case of unequal sample sizes. No mathematical justification is presented by the authors of these procedures.

One of these procedures is Kramer's K procedure. The K procedure replaces the common sample size in the formula for the T procedure with the harmonic mean of the two sample sizes of the specific means involved in the comparison. Another approximation is due to Winer (1969). Winer's procedure uses the harmonic mean of all k samples for all the pairwise comparisons. Miller (1966) also has an approximation. This procedure uses the mean value of the sample sizes. Each of these procedures utilizes the ordinary studentized range distribution as the critical value.

Two additional approximations are the GH and TH procedures. The GH procedure uses the studentized range distribution, whereas the TH procedure uses the student t distribution in determining critical values. Both the GH and TH procedures use Welch's approximate estimate for degrees of freedom.

Conservative procedures. The T' and T" procedures are extensions of Tukey's procedure to the unequal sample size cases. The S procedure is designed to make all possible, not only pairwise, comparisons among the population means and is based on the F distribution. It can be proven that each of these three procedures is conservative for pairwise comparisons. A procedure is conservative when the Type I error rate is less than the nominal α level.

Both the T' and T" procedures utilize the studentized augmented range distribution in calculating their critical values. The studentized augmented range distribution is tabulated by Stoline (1978).

Stoline claims that Tukey's approximation of the upper α points of the studentized augmented range with the corresponding upper α point of ordinary studentized range for $k \geq 3$ and $\alpha \leq .05$ holds as well, if not better, for those cases where $k \geq 4$ and $\alpha \leq .20$.

In summary, the T, Newman-Keuls, and Duncan procedures are single stage procedures that require equal sample size. In contrast, the single stage procedures that do not require equal sample sizes include the eight procedures of K, Winer, Miller, GH, TH, S, T', and T". The next section is a review of two stage procedures, opposed to the aforementioned single stage procedures. The two stage procedures discussed here do not require equal sample sizes.

Two Stage Comparison Procedures

The two stage pairwise comparison procedures of PLSD and PW are discussed in this section. Each of these procedures utilize the F ratio in a one way analysis of variance as the first stage.

For the PLSD procedure, if the F ratio is significant at a pre-determined α level, the appropriate t test is used to make each single pairwise comparisons among the k population means. The comparisons are made at the predetermined α level with N-k degrees of freedom.

All $k(k-1)/2$ pairwise comparisons can be tested at the second stage, for the PLSD procedure, if the first stage is significant. According to Miller (1966), it can happen that the F test gives a sig-nificant value for the first stage, but none of the pairwise differences at the second stage are significant.

A second two stage procedure similar to the PLSD procedure is being

proposed in this study. The proposed procedure is referred to as the protected Welch (PW). If the F ratio is significant at a predetermined α level, all pairwise comparisons among the k means are performed. The critical value for the second stage is read from the student's t distribution at level α with Welch's approximate estimate for the degrees of freedom. Welch's estimate is dependent on the size and variances for each of the two samples being compared.

The justification for proposing the PW procedure is based upon Scheffé (1959) and Lee and Gurland (1975). Scheffé reports that the F test and t test are robust with respect to unequal variances if one has equal sample size or a small amount of imbalance among sample sizes. However, these tests are not robust with respect to unequal variances if one has a large amount of imbalance among sample sizes. Lee and Gurland reported that the t test with Welch's approximate estimate for degrees of freedom is more robust than several other approximate 2 sample tests with respect to unequal variances.

This completes the discussion of the single and two stage procedures. The following section describes various simulation studies that compare experimentwise Type I error rates and power of some of these discussed procedures.

Simulation Studies

As was previously mentioned, Carmer and Swanson conducted a comprehensive simulation study in which all samples were equal in size. This study compared ten pairwise compairson procedures as to their experimentwise Type I error rates and power. As a recommended procedure, Carmer and Swanson proposed the PLSD procedure.

Supporting evidence for this recommendation is given in Carmer and Swanson, in the form of experimentwise Type I error rates based on 4000 simulations for k=5, 10, and 20 means for the null hypothesis $H_0; \mu_1 = \ldots = \mu_k$. Table 1 presents part of this evidence at the 5 percent significance level.

TABLE 1

Carmer and Swanson (1973, p. 70) Observed Experimentwise
Type I Error Rates in Percentages

Procedures	Number of Treatments		
	k = 5	k = 10	k = 20
Scheffé	2.4	.3	.00
Tukey	5.0	4.8	4.7
Newman-Keuls	5.7	5.0	4.8
Duncan	18.2	37.3	62.6
PLSD	4.8	5.4	5.2

Scheffé's procedure is too conservative (Type I error rate is less than the nominal α level) and Duncan's procedure is too liberal (Type I error rate is greater than the nominal α level) for k=5, 10 and 20. Carmer and Swanson recommended that Scheffé's procedure be used only for nonpairwise comparisons. Clearly, the Tukey, Newman-Keuls, and PLSD procedures have the most stable 5 percent Type I error rates. In addition, the power studies in Carmer and Swanson indicated that the PLSD procedure produced comparable, if not greater power for more cases than did the Tukey and Newman-Keuls procedure.

Smith (1971) compared the approximate procedures of K, Miller, and Winer. These procedures were compared using unequal sample sizes whose ratios were of the order of up to 3:1, while assuming homogeneous

population variances. This study showed that Kramer's procedure consistently provided actual experimentwise Type I error rates in close agreement with the corresponding nominal probabilities.

Smith's work was extended by Keselman, Murray, and Rogan (1976), who compared the Winer and Kramer procedures for unequal sample sizes of various ratios for varying numbers of treatment levels. The observed experimentwise Type I error rates rarely exceeded their nominal significance levels by more than a percentage point, even when group sizes differed by a ratio of 40:1. Typically, the Winer observed experimentwise Type I error rates were larger than the nominal levels of significance, whereas the Kramer estimates were found to be less than the nominal significance levels. Keselman et al. also recommended Kramer's procedure due to its conservative experimentwise Type I error control.

Games and Howell (1976) examined the robustness of the Winer procedure when unequal samples were combined with various patterns of heterogeneous variances. Their simulated experimentwise Type I error rates were less than the nominal value when the smallest sample was paired with the smallest variance, but exceeded the nominal value when the smallest sample was combined with the largest variance.

As was previously mentioned, a study by Keselman and Rogan (1978) compared the S, K, T', GH, and T" procedures. As a result of this study, the S, T', and K procedures were not recommended. It was concluded that these three procedures were often adversely affected by combining unequal sample sizes with certain heterogeneous variances. Only the T" and GH procedures maintained simulated significance levels

close to the nominal level when unequal sample sizes and unequal variances were inversely paired. However, Keselman and Rogan recommended the GH procedure opposed to the T" procedure because it maintained simulated significance levels closer to the nominal level than the T" values. The T" values were consistently less than the nominal values.

Based on the aforementioned review of literature, there is no complete consensus as to which procedure should be used for pairwise comparison between means. According to the Carmer and Swanson study, the PLSD procedure is one of the best pairwise comparison procedures for controlling experimentwise Type I error in the equal sample size case. As far as the Tukey type procedures, the K and GH approximation procedures seem to perform better than the somewhat conservative procedures T' and T". The S procedure is extremely conservative for pairwise comparison of means. Finally, no simulation study which includes the PLSD procedure for unequal sample sizes exists.

Summary

Chapter II included the specifications for the seven procedures being compared in this study, reviewed several additional comparison procedures, and discussed several simulation studies that compared some of these procedures. It is unclear from these studies as to which of these procedure(s) should be used in the unequal sample size case under the assumption of normality and equal variance. This uncertainty is due in part, to the fact that all these procedures have not been compared in a single study. Chapter III describes the preliminary work that is needed to be accomplished prior to this study.

PRELIMINARY WORK

This chapter focuses on the preliminary procedures that were nec-
essary for this simulation study. Specifically discussed in this chap-
ter is the testing of random number generators, the testing and the
selection of an optimal method for transforming generated random digits
into normal variables, and finally in order to add confirmatory evi-
dence to the selected data simulator, a partial replication of the
Carmer and Swanson study is discussed.

Introduction

The computer is admirably suited to the exploratory problem
solving technique of simulation. Millions of trials may be generated
with reasonable speed, and the computer may be used to classify and
count results as it generates them.

An important element in this simulation study is the generation
of random numbers. A random number generator is a procedure for pro-
ducing, by use of a computer, a sequence of numbers $U_1, U_2, U_3 \ldots$ which
is to represent a sequence of independent random numbers. There are
many computer methods for generating such numbers, but all have in
common the generation of a strictly deterministic sequence of digits,
which are therefore more properly called "pseudo-random" numbers. That
is, the digit strings that are produced are not truly random, since the
same strings are reproducible each time the generation routine is used.

19

In order for this simulation study to have credibility, the accuracy of the computer program in generating random numbers is investigated. The next section investigates two such random number generators.

Testing of Random Number Generators

The simulation studies of Carmer and Swanson and Keselman and Rogan used the Marsaglia, MacLaren, and Bray (1964) procedure to obtain random numbers. MacLaren and Marsaglia (1965) conducted a study, using statistical tests described below, comparing the Marsaglia, MacLaren, and Bray procedure with seven other random number generators. On the basis of this aforementioned study, the authors recommended the Marsaglia, MacLaren, and Bray procedure for generating random numbers.

Western Michigan University Computer Center operates a time sharing computer system. The Dec-System Program Data Processing-10 Computer (PDP-10) is utilized in the Computer Center. The PDP-10 random number generator is based on a method by Payne, Rabring, and Bogyo (1969).

A Fortran computer program (see Appendix A) was written for the random number generator described by Marsaglia, MacLaren, and Bray. Additionally, a Fortran computer program (see Appendix B) was written to test the aforementioned two random number generators using the statistical procedures mentioned in MacLaren and Marsaglia.

The statistical tests used for both methods were χ^2 tests applied to the distributions of the random numbers, pairs of random numbers, triples of random numbers, and various simple functions of several random numbers. These χ^2 tests were used primarily to determine if the

generated numbers were indeed random.

All the tests had the same general form. A sequence of n varia-bles $X_1,...,X_n$ was computed from the sequence of uniform numbers. If the uniform numbers were actually independent uniform random variables, the X_i's would be independent identically distributed random variable. The range of the X_i's was divided into m cells of equal probability $p=1/m$ and the number of occurences k_i in each cell counted. The Chi Square statistic

$$\chi^2 = \sum_1^m (k_i - np)^2 / np \qquad (3.1)$$

was first computed, and then was converted to a percentage which re-presented the probability that a random sample would give a larger value of χ^2 than that observed.

All tests were made simultaneously, using the same sequence $U_1, U_2,$ .... As a programming convenience, the uniform numbers were taken in sets of ten. From a set of ten, five pairs were obtained for the pairs test, three triples for the triples test, and one each of the various one-dimensional variables. Thus, from a set of ten uniform numbers, only one pair was used for the test of the maximum of two uniforms, one triple for the maximum of three, etc. Successive tests start at the point in the generated sequence where previous tests finished.

Statistical tests

There are five statistical tests used to compare the random number generators. The details of each test are described below.

Uniformity. This was included for the sake of completeness. The unit interval was divided into 100 equal cells. For each run 10,000

numbers were used, the exact sequence being $U_1, U_{11}, U_{21}, \ldots$. The occurences in each cell were counted, and the χ^2 statistic computed, which was used to test the null hypothesis that the generated digits were "uniformly" distributed along the unit interval (0,1).

Pairs. Successive pairs of uniform numbers were taken as the coordinates of a point in the unit square. The unit square was divided into 100 equal squares. For each run a total of 50,000 pairs was generated. A χ^2 statistic was computed which was used to test the null hypothesis that the pairs of digits were "uniformly" distributed in a unit plane.

Triples. Successive triples of uniform numbers were taken as the coordinates of a point in the unit cube. However, every tenth uniform number was skipped. The cube was divided into 100 equal cells. For each run a total of 30,000 triples was generated. A χ^2 statistic was computed which was used to test the null hypothesis that the triples of digits were "uniformly" distributed in a unit cube.

Maximum of n. Let $V_1, V_2, \ldots, V_n$ be independent uniform variates from the unit interval (0,1), then $W = \max(V_1, \ldots, V_n)$ should have the distribution $P(W < a) = F(a) = a^n$ for $0 < a < 1$, and hence $F(W) = [\max (V_1, \ldots, V_n)]^n$ should be uniformly distributed over (0,1). This assertion was proved in Hogg and Craig (1970). The χ^2 test was obtained by dividing the unit interval into 100 equal subintervals with samples of 10,000 U's. The first n-tuple was obtained from the first 10 U's, the second n-tuple from the second 10 U's, and so on. Thus, 100,000 U's were generated for each of the tests with n=2,3,4,10. This χ^2 statistic was used to test the hypothesis that the "maximum of n" distribution

function was "uniformly" distributed on the unit interval.

Minimum of n. This test was the same as that for the maximum of n except that the minimum of n-tuples, $W=\min(V_1,\dots V_n)$, was taken and the corresponding distribution function $G(W)=[\min(V_1,\dots V_n)]^n$ was used. Again 100,000 U's were generated for each of the χ^2 tests with n=2,3,4, 10. The χ^2 statistic obtained was used to test the hypothesis that the "minimum of n" distribution function was "uniformly" distributed on the unit interval, as it should be if the data generated was genuinely random and independent.

Test results

Table 2 contains the results of 33 tests conducted for each of the two generators. The tabled value is the probability, expressed as a percentage, that the Chi Square variate will exceed the observed test value, assuming the data is uniformly distributed and independently generated. Thus, the 33 values for each generator should behave some-what like a set of 33 numbers chosen uniformly from the unit interval. There may be occasional small or large values. Any preponderance of such values would be suspect.

On the basis of these tests, which may be viewed as preliminary, but certainly indicative, both generators behave well on these tests, except for the MacLaren, Marsaglia, and Bray generator under the uni-form test. For example, under the uniform test the percentages for the MacLaren, Marsaglia, and Bray generator are 2, 8, and 33. Conse-quently, as a result of these tests the PDP-10 generator looked at least as good as the MacLaren, Marsaglia, and Bray generator, the

TABLE 2

Summary of Tests on the Two
Random Number Generators in Percentages

Tests	Run					
	PDP-10 Generator			MacLaren, Marsaglia & Bray Generator		
	1	2	3	1	2	3
Uniformity	80	92	73	02	08	33
Pairs	56	09	55	32	73	71
Triples	09	79	76	11	49	46
Maximum of 2	60	17	54	21	47	29
Maximum of 3	22	76	77	88	50	70
Maximum of 5	90	38	51	99	15	33
Maximum of 10	96	43	70	85	50	70
Minimum of 2	35	64	44	68	17	70
Minimum of 3	25	16	72	99	22	65
Minimum of 5	83	27	30	16	32	14
Minimum of 10	67	03	25	19	22	31

generator recommended by MacLaren and Marsaglia, and used by Carmer and Swanson and Keselman and Rogan. If one had to choose a generator based on these tests alone, the PDP-10 generator would probably be chosen. However, further testing was conducted prior to the selection of the generator.

Generation of Normal Deviates

In order to conduct the pairwise comparison simulation studies, the generated random numbers need to be converted into deviates that satisfy certain distribution assumptions. This particular simulation study was conducted under the basic analysis of variance assumptions of normality, independent random samples, and constant variance for the populations.

Box and Muller (1958) developed a method of generating normal deviates from independent random numbers. This method uses an algebraic transformation technique. Let U_1, U_2 be independent random variables from the same rectangular density function on the unit interval (0,1). Then

$$X_1 = [-2\log_e(U_1)]^{\frac{1}{2}}\cos(2\pi U_2) \qquad (3.2)$$

$$X_2 = [-2\log_e(U_1)]^{\frac{1}{2}}\sin(2\pi U_2) \qquad (3.3)$$

represents a pair of independent random variables from the same normal distribution with mean zero, and variance one. This has been proved in Box and Muller.

This section described the Box and Muller method of generating normal data from random numbers. The next section describes the results of the testing that was conducted in order to select the

optimal method of generating normal deviates.

Generator Selection

In this section one of the two alternate methods of generating normal deviates with mean of zero and standard deviation of one is selected. The first method is the combination of the PDP-10 generator with Box and Mueller (DBM), the second method is the combination of the MacLaren, Marsaglia, Bray generator with Box and Mueller (MMBM).

A Fortran computer program (see Appendix C) was written to test the two aforementioned methods. A sequence of n normal deviates $X_1,\ldots,X_n$ were generated by the two generator methods DBM and MMBM. If these deviates were actually distributed as normal variables with mean zero and standard deviation one, they should possess the properties of the normal distribution. To test this hypothesis, the real line was divided into 16 disjoint line segments $(L_i,L_{i+1})i=1,\ldots,16,$ where the probability $p_i=Pr[L_i<z<L_{i+1}]$, $i=1,\ldots,16,$ was known and the number of occurences in each was counted. These 16 segments were selected in order to examine a reasonable range of normal deviates.

Test results

Table 3 contains the results of the DBM and MMBM methods of generating normal deviates. The table contains six columns. The first column identifies each of the 16 line segments with a corresponding number. The second column identifies the specific line segments. The first two line segments represent .68 and -.68 standard deviations from zero, respectively. This pattern continues through the 16 seg-

ments, where the 15 and 16 segments represent 3.0 to a positive

TABLE 3

Summary of Tests on Generating Normal Deviates

Interval Number	Line Segments	Chi Square		Probability of χ^2 Expressed as Percentages	
		DBM	MMBM	DBM	MMBM
1	(0,.68)	.38	2.36	100	100
2	(-.68,0)	1.23	1.69	100	100
3	(.68,1.28)	.94	17.58	100	29
4	(-1.28,-.68)	.49	2.25	100	100
5	(1.28,1.65)	.81	13.49	100	57
6	(-1.65,-1.28)	2.10	15.86	100	39
7	(1.65,1.96)	1.22	42.82	100	0
8	(-1.96,-1.65)	.25	39.64	100	0
9	(1.96,2.33)	.03	105.86	100	0
10	(-2.33,-1.96)	2.58	116.49	100	0
11	(2.33,2.58)	.00	1.75	100	100
12	(-2.58,-2.33)	.21	7.14	100	95
13	(2.58,3.0)	.50	3553.10	100	0
14	(-3.0,-2.58)	.03	3575.70	100	0
15	(3.0,+∞)	5.32	1456.00	99	0
16	(-∞,-3.00)	.06	1456.00	100	0

infinity and -3.0 to a negative infinity standard deviations from zero, respectively. Columns three and four contain the Chi Square statistics that were computed for each interval. Finally, the entries in columns five and six are the probabilities, expressed as a percentage, for the observed Chi Square values for the two methods.

On the basis of these Chi Square tests it is concluded that the DBM method of generating normal deviates is superior to the MMBM method. The MMBM method becomes inferior to the DBM method when the line segments are about 1.65 standard deviations from zero. The only exception to this is for the two intervals 11 and 12. Based on these tests, the PDP-10 random number generator followed by the trans- formation of uniform random numbers to normal deviates by using the Box and Muller method was selected as the data simulator for this simulation study. In the next section of this chapter, this data simulator was used to replicate a portion of the Carmer and Swanson study.

Replication of Portions of the Carmer Swanson Study

In order to gather evidence about the accuracy of the DBM data simulator, a Fortran program was written (see Appendix C) in an attempt to replicate portions of the Carmer and Swanson study. The procedures used were the Scheffé, Tukey, and PLSD under the equal sample size case. The DBM method was used to generate normal deviates. An outline of this simulation study is as follows. First, the signi- ficance level of 5 percent was selected because it was used by Carmer and Swanson. Additionally, the significance levels of .1, .5, 1.0,

10.0, and 20.0 percent were selected to further examine the DBM data simulator over a wider range of possible α levels. Second, Carmer and Swanson's number of treatments of 5, 10, and 20 were used. Third, sample size values of 4, 6, 8, and 14 were selected. These were selected in order to examine a reasonable range of sample sizes. Carmer and Swanson used sample size values ranging from 5 to 15. Fourth, for each combination of number of treatments and sample size, 1000 replications of the experiment were performed. This number was selected since Carmer and Swanson used 1000 replications.

Test results

Table 4 contains the results of this simulation study for the experimentwise Type I error rates, assuming the null hypothesis that the 5, 10, or 20 population means are all equal, respectively. For all three groups of treatments, the single stage Tukey procedure and two stage PLSD procedure have observed experimentwise error rates in close agreement with the nominal levels. For example, at the nominal 10 percent significance level with k=5, 10, and 20, the observed experimentwise error rates for Tukey and PLSD are (11.1, 9.8, 10.1) and (10.9, 9.9, 10.0), respectively. For the Scheffe procedure, the observed experimentwise error rates are not in close agreement with the nominal levels. At the nominal 10 percent significance level with k=5, 10, and 20, Scheffe's observed experimentwise error rates are 4.7, .80, and .00. respectively.

TABLE 4

Observed Experimentwise Type I
Error Rates in Percentages

Procedures	Significance Levels					
	20%	10%	5%	1%	.5%	.1%
			k = 5 Treatments			
Tukey	21.3	11.1	5.1	.90	.31	.08
Scheffé	12.2	4.7	2.2	.30	.10	.02
PLSD	21.9	10.9	5.2	.95	.48	.10
			k = 10 Treatments			
Tukey	20.4	9.8	4.7	.90	.45	.05
Scheffé	2.4	.80	.28	.30	.00	.00
PLSD	20.0	9.9	4.7	.95	.48	.02
			k = 20 Treatments			
Tukey	20.2	10.1	4.9	.85	.28	.08
Scheffé	.08	.00	.00	.00	.00	.00
PLSD	20.1	10.1	5.0	1.2	.52	.12

Table 5 contains the observed experimentwise error rates at the 5 percent significance level for a portion of the Carmer and Swanson study and this replication study. The values obtained in this study are consistent with and do not diverge widely from the Carmer and Swanson data. Four values are below and five values are above or equal to the 5 percent values obtained by Carmer and Swanson.

TABLE 5

Observed Experimentwise Type I
Error Rates Obtained by Carmer
and Swanson (1973, P. 70) and
this Replication Study at the
5 Percent Significance Level

Procedures	Treatments		
	k=5	k=10	k=20
Tukey	5.1 (5.0)	4.7 (4.8)	4.9 (4.7)
PLSD	5.2 (4.8)	4.7 (5.4)	5.0 (5.2)
Scheffé	2.2 (2.4)	.28 (.30)	.00 (.00)

(Parenthensized values are the results obtained by Carmer and Swanson.)

This simulation study has replicated a portion of the Carmer and Swanson study. This is an important feature because it adds confirmatory evidence to the choice of the PDP-10 random number generator followed by the use of the Box and Muller procedure to obtain normal deviates for the simulation study reported in the remainder of this

dissertation.

 This chapter described the preliminary work that was performed prior to the simulation study itself. Chapter IV contains a report and discussion of the simulation study.

CHAPTER IV

SIMULATION STUDY RESULTS AND DISCUSSION

Chapter IV is organized as follows: First, the parameters for the simulation study are presented; second, the results of the study are reported and discussed; and third, areas of further study are discussed.

Parameters for the Simulation Study

Under the null hypothesis $H_0: \mu_1 = \ldots = \mu_k$, the experimentwise Type I error rate was investigated for each of the seven pairwise comparison procedures. The parameters of the simulation study are:

1. The error rate levels of 1, 5, 10, and 20 percent were selected in order to examine a reasonably wide range of α levels which would cover most levels commonly used in practice.

2. The k values of 3, 5, and 10 were used as the number of treatments in this study. These values were selected because it was felt the range of k=3 to 10 covered most situations arising in practice, at least those arising in the educational and social science areas.

3. There were four categories of sample sizes used in this simulation study. These categories were formed as follows. Let $v = \max(n_1, \ldots, n_k)/\min(n_1, \ldots, n_k)$, where $n_1, \ldots, n_k$ represent the sample sizes for each of the k values 3, 5, and 10, respectively. The first category was of equal sample size (v=1); the second category was of small imbalance (1<v<2); the third category was of moderate imbalance (2≤v<2.5); and the fourth category was of large imbalance (v=5). Table 6 lists the specific sample sizes selected for each of the four categories for each of the k values. For example, for the moderately imbalanced category with k=5, the four sample sizes were (6,6,6,15,15), (6,6,15,15,15), (15,15,30,30,30), and (6,15,15,15,15).

A Fortran program, with the aforementioned parameters, was written (see Appendix D) for the simulation of data generated from the

33

TABLE 6

Sample Sizes for each of the Treatments

Sample Imbalance	Treatments																	
	k=3			k=5					k=10									
Equal Sample size (v=1)	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Small (1<v<2)	6	8	10	6	6	8	10	10	6	6	6	6	6	8	8	8	8	10
	6	6	10	6	6	6	10	10	6	6	6	6	6	6	10	10	10	10
	6	10	10	6	6	10	10	10	6	6	6	6	6	8	8	8	10	10
	8	8	10	8	8	8	10	10	6	6	6	6	6	6	6	6	8	8
	6	6	8	15	15	25	25	25	15	15	25	25	25	25	25	25	25	25
Moderate (2≤v<2.5)	6	6	15	6	6	15	15	15	6	15	15	15	6	6	15	15	15	15
	6	15	15	6	15	15	15	15	6	30	30	30	6	6	30	30	30	30
	15	15	30	15	15	30	30	30	6	10	10	10	10	10	15	15	15	15
	15	30	30	6	15	15	15	15	6	10	10	10	10	10	15	15	15	15
Large (v=5)	6	15	30	6	15	30	30	30	6	15	30	30	30	30	30	30	30	30
	6	6	30	6	6	6	6	30	6	6	6	6	6	6	30	30	30	30
	6	30	30	6	30	30	30	30	6	6	6	6	6	6	6	6	6	30

one way fixed effects analysis of variance model,

$$X_{ij} = \mu + \varepsilon_{ij} \; ; \; i=1,\ldots,n_j, \; j=1,\ldots,k \qquad (4.1)$$

where the ε_{ij} are independent normal random errors with mean zero and constant variance. The DBM method was utilized to obtain these independent normal random errors ε_{ij}.

For each experiment (involving k=3, 5, and 10 and a specific sample in each of the categories) 1000 replications were performed. For example, with k=3 under the equal sample size case, 1000 replications were performed for (6,6,6), (15,15,15), and (30,30,30). The observed experimentwise Type I error rate equals the number of replications with at least one pairwise comparison declared significant at level α divided by 1000.

Simulation Results

This section contains the results of the simulation study. The results are reported in a series of five tables. The first four tables contain the observed experimentwise Type I error rates for the four imbalanced cases. The last table reports the ranges for these error rates for only the PLSD, PW, K, and GH procedures.

Type I observed error rates: balanced case

Table 7 contains the observed experimentwise Type I error rates for the equal sample size case. Under the equal sample size case the PLSD, T', K, GH, and PW procedures all exhibited experimentwise Type I error rates close to the nominal significance levels of 1, 5, 10, and 20 percent with the empirical estimates of the T' and K procedures

TABLE 7

Observed Experimentwise Type I Error Rates
in Percents: Equal Sample Size

	Treatments											
	k=3				k=5				k=20			
	Significance Levels											
Procedures	1%	5%	10%	20%	1%	5%	10%	20%	1%	5%	10%	20%
PLSD	.90	5.2	9.7	20.9	1.1	5.4	10.5	21.3	1.0	4.5	8.8	19.3
T'	1.1	5.1	10.2	20.7	1.2	5.0	10.2	20.8	.93	4.8	9.8	19.6
T''	.20	2.3	5.8	14.4	.2	1.6	5.3	13.0	.13	1.5	4.2	12.1
K	1.1	5.1	10.2	20.7	1.2	5.0	10.2	20.8	.93	4.9	9.8	19.6
GH	.77	4.4	10.6	21.2	.63	5.7	11.1	22.8	1.1	5.3	11.3	23.6
TH	.63	3.7	8.7	17.2	.57	3.7	8.0	15.3	.77	3.0	5.8	12.0
PW	.87	5.2	9.7	20.9	1.1	5.4	10.5	21.2	1.0	4.5	8.8	19.3

deviating least from the nominal significance levels. The T" and TH procedures yielded conservative experimentwise error rates. The degree of conservativeness varied somewhat as a function of the number of treatments. For instance, at the nominal 5 percent significance level with k=3,5, and 10 the observed error rates of T" and TH are (2.3, 1.6, 1.5) and (3.7, 3.7, 3.0), respectively.

Type I observed error rates: small imbalanced cases

Observed experimentwise Type I error rates are presented in Table 8 for the small imbalanced case. This table shows that the PLSD, K, GH, and PW procedures have observed experimentwise error rates close to the nominal significance levels of 1, 5, 10, and 20 percent. However, as the number of treatments increase from k=3 to k=10 the GH observed error rates tend to become somewhat liberal. For example, at the nominal 20 percent significance level, the experimentwise error rates for GH extend from about 19 percent to 26 percent. The T', T", and TH procedures are somewhat conservative in the small imbalanced cases. For instance, at the nominal 10 percent significance level the observed experimentwise error rates for T' and TH are about 7 percent and for T" about 3 percent.

Type I observed error rates: moderate and large imbalanced cases

Tables 9 and 10 contain the observed experimentwise Type I error rates for the moderate and large imbalanced cases, respectively, The results in both of these cases are similar. Very low experimentwise error rates are found for the T" procedure for all the nominal signi-

TABLE 8

Observed Experimentwise Type I Error Rates
in Percents: Small Imbalance in Sample Size

Procedures	Treatments											
	k=3				k=5				k=20			
	Significance Levels											
	1%	5%	10%	20%	1%	5%	10%	20%	1%	5%	10%	20%
PLSD	.86	5.0	9.7	19.4	1.1	4.9	10.2	20.2	.97	4.8	9.7	20.1
T'	.54	3.7	7.6	16.1	.50	3.6	7.7	16.4	.68	3.8	7.5	16.4
T"	.08	.90	3.2	10.6	.05	.87	3.1	10.1	.02	.58	2.4	9.1
K	.86	4.9	9.7	9.5	.80	5.0	10.3	19.8	.95	4.9	9.1	20.1
GH	.78	4.7	9.8	19.4	.95	5.5	11.3	21.6	1.1	6.3	12.7	26.1
TH	.64	3.9	7.8	15.4	.60	3.8	7.3	14.1	.43	3.1	6.1	11.7
PW	.68	4.8	9.6	19.4	1.0	4.9	10.2	20.2	.93	4.8	9.7	20.1

TABLE 9

Observed Experimentwise Type I Error Rates
in Percents: Moderate Imbalance in Sample Size

Procedures	Treatments											
	k=3				k=5				k=20			
	Significance Levels											
	1%	5%	10%	20%	1%	5%	10%	20%	1%	5%	10%	20%
PLSD	.90	4.8	10.1	21.4	.86	5.0	9.9	20.7	1.1	5.8	10.6	21.5
T'	.40	2.8	6.3	15.0	.60	3.0	6.4	14.2	.43	3.1	6.9	14.1
T"	.10	1.4	4.5	12.3	.04	1.1	3.1	9.6	.07	.80	3.2	10.1
K	.90	4.8	9.9	21.2	.94	4.9	9.6	19.9	.90	5.3	10.6	20.3
GH	1.1	5.5	10.7	21.4	1.2	5.5	10.8	22.0	1.6	7.5	13.9	26.5
TH	.90	4.7	8.9	17.2	.92	4.0	7.5	14.7	1.0	4.4	8.1	14.0
PW	.82	4.6	10.0	21.3	.80	4.9	9.9	20.7	1.1	5.8	10.6	21.5

TABLE 10

Observed Experimentwise Type I Error Rates
in Percents: Large Imbalance in Sample Size

	Treatments											
Procedures	k=3				k=5				k=20			
	Significance Levels											
	1%	5%	10%	20%	1%	5%	10%	20%	1%	5%	10%	20%
PLSD	1.2	5.6	11.0	21.5	1.1	5.2	10.7	21.2	.93	5.4	10.3	20.0
T'	.33	2.5	5.4	12.7	.50	3.1	6.5	13.9	.75	3.3	6.9	14.7
T"	.1	1.5	4.1	10.6	.02	.68	2.9	9.1	.00	.63	2.0	7.2
K	1.1	5.2	10.8	21.0	.90	5.2	10.1	20.1	.93	4.4	8.9	18.7
GH	1.9	6.3	11.8	22.4	1.5	6.5	11.6	22.6	1.6	6.6	12.2	24.8
TH	1.6	5.5	9.6	18.5	1.2	4.7	8.3	14.5	1.2	3.9	7.0	12.0
PW	.80	4.8	10.3	21.2	.82	5.1	10.6	21.2	.90	5.3	10.3	20.0

40

cance levels. For instance, at the nominal 5 percent significance level the observed experimentwise error rates for both cases are about 1 percent. The error rates for T' and TH are also low. However, these are not generally as low as the T" rates. Rates for the GH procedure tend to be somewhat high. For instance, at the nominal 5 percent significance level with k=10, the GH observed experimentwise error rates for the moderate and large imbalanced cases are about 7 percent. Finally, the PLSD, K, and PW procedures quite consistently maintain their experimentwise error rates close to the nominal significance levels for both of these rather imbalanced cases.

Ranges for error rates

Table 11 contains the ranges for the observed error rates for the PLSD, PW, K, and GH procedures. Only these four procedures are included in this table because they most consistently maintain the experimentwise Type I error rates at or close to nominal significance levels over all four sample size cases; although, the GH procedure yields some liberal error rates as k becomes large.

The ranges of the error rates for the PLSD and PW procedures are similar over all levels of treatment size. They usually vary from about .3 to 2.0. The ranges for all four of these procedures are not as similar as the ranges for the PLSD and PW; however, none of the procedures have ranges that are consistently lower or higher than the other procedures. The only real exception to this is for the GH procedure with k=10 and nominal significance level of 20 percent. The GH ranges in this case are all about 5; whereas, the ranges for the

TABLE 11

Ranges for the Observed Experimentwise
Type I Error Rates for the
PLSD, PW, K, and GH Procedures

Procedures	Equal				Small				Moderate				Large			
								Significance Levels								
	1%	5%	10%	20%	1%	5%	10%	20%	1%	5%	10%	20%	1%	5%	10%	20%
							k = 3 Treatments									
PLSD	.3	.9	2.3	2.6	.7	2.0	2.4	2.6	.3	1.4	1.2	1.9	.5	.4	.5	3.8
PW	.3	.9	2.3	2.6	.4	1.5	2.3	2.6	.1	1.0	1.0	1.9	.5	.5	.6	3.9
K	.4	.3	1.4	2.7	.9	1.3	2.4	2.0	.6	1.2	1.7	1.8	.5	.3	1.9	3.5
GH	.3	1.2	1.7	2.7	.4	3.2	2.0	3.1	.7	2.1	1.5	2.4	.6	1.4	1.1	3.3
							k = 5 Treatments									
PLSD	.4	1.5	2.0	2.6	1.2	2.9	2.1	3.4	.8	.5	2.6	3.6	.6	1.5	1.9	3.5
PW	.4	1.5	2.0	2.6	.7	2.9	2.1	3.4	.8	.5	2.6	3.6	1.0	1.4	1.9	3.6
K	.5	1.0	2.0	1.8	.7	2.8	2.4	4.5	1.1	1.8	2.1	3.5	.8	1.4	2.7	3.1
GH	.6	1.7	1.9	4.1	1.1	2.5	3.2	4.2	.5	1.8	2.4	3.6	1.4	1.3	2.1	1.5
							k = 10 Treatments									
PLSD	.2	1.4	.9	.3	1.1	2.7	3.8	3.9	.5	1.3	1.0	1.4	.1	1.8	1.9	1.5
PW	.2	1.4	.9	.3	1.0	2.7	3.8	3.9	.5	1.3	1.0	1.4	.2	1.8	1.9	1.5
K	.8	1.2	.3	1.4	.9	2.6	3.6	3.3	.5	1.6	.7	1.7	.8	1.4	2.3	2.9
GH	.6	.4	2.3	4.6	.8	2.1	3.5	6.3	.5	1.5	2.4	3.2	1.8	1.4	1.1	4.2

Note: Under "Sample Imbalance" the groups are: Equal, Small, Moderate, Large.

other three procedures are usually less than 2.0.

Concluding remarks

The underlying basis of the work reported here is on examination of the appropriateness of the seven aforementioned pairwise comparison procedures under the assumption of equal variances. The examination is based upon the control of experimentwise Type I errors for the equal sample size case and the three imbalanced sample size cases. For the equal sample size case, the T' and K procedures have very stable observed experimentwise Type I error rates that are close to their nominal values. The PLSD, GH, and PW procedures are close second choices. The two remaining procedures T" and TH are too conservative.

For the three imbalanced cases, T" and TH are again too conservative. In addition, the T' procedure is also too conservative for these imbalanced cases. The optimal choices for controlling Type I error with unequal sample sizes are the PLSD, K, and PW procedures, with the GH procedure a second choice. Consequently, if one agrees with the notion that an experimenter should want to use a procedure having the most stable experimentwise Type I α level; then, based on this study it is recommended the PLSD, PW, or K procedures be used for pairwise comparisons.

The PLSD, PW, and K procedures are recommended under the assumption that the person performing the pairwise comparison has homogeneity of population variances. However, the experimenter may not have homogeneity of population variances, and consequently the recommended pro-

cedures may be adversely affected by combining unequal sample sizes with unequal variances. In the Keselman and Rogan study, the K procedure is not robust with respect to unequal variances when one has imbalanced sample sizes. As to the two remaining recommended procedures, there exist no emperical data as to their robustness with respect to unequal variances under the imbalanced case. Although, as is discussed in Chapter II, the PW may be more robust than the PLSD procedure with respect to unequal variances especially as the models become more imbalanced. Consequently, additional work needs to be done comparing these pairwise comparison procedures in imbalanced cases with unequal variances.

Areas for Further Study

An extension of this study would be a comparison of the PLSD, PW, K, GH procedures and possibly a few other procedures where the assumption of the homogeneity of variance has been relaxed under the imbalanced cases.

The factors that may be varied in such a study are degree of sample size imbalance, degree of variance heterogeneity, and pattern of variance heterogeneity. It is evident that such a study is more wide ranging than when the homogeneity of variance assumption is satisfied. The reason for this is that there exists a very wide variety of combinations of the levels of the factors described above.

For example, suppose that three sets of unequal sample sizes are selected, corresponding to a small, medium, and large sample size imbalance. Also, suppose small, medium, and large degrees of variance

heterogeneity are selected. In addition, suppose that for each of the three variance heterogeneity cases, two patterns of heterogeneous variances are selected; one pattern in which all variances are unequal and one pattern where all but one of the variances are equal. This study would then contain 18 ($18=3 \cdot 3 \cdot 2$) combinations of unequal variance cases to be analyzed for each pairwise comparison procedure included in the study.

The above example illustrates that a study involving heterogeneity of variance under the imbalanced case may be very involved. However, such a study is clearly called for.

REFERENCES

Box, G.E.P., and Muller, M.E. A note on the generation of normal deviates. _Annals of Mathematics Statistics_, 1958, 29, 610-611.

Carmer, S.G., and Swanson, M.R. An evaluation of ten pairwise multiple comparison procedures by monte carlo methods. _Journal of the American Statistical Association_, 1973, 68, 66-74.

Duncan, D.B. Multiple range and multiple F tests. _Biometrics_, 1955, 11, 1-42.

Fisher, R.A. _The Design of Experiments_. Edinburgh: Oliver and Boyd Ltd., 1949.

Games, P.A., and Howell, J.F. Pairwise multiple comparison procedures with unequal n's and/or variances: A monte carlo study. _Journal of Educational Statistics_, 1976, 1, 113-125.

Hochberg, Y. A modification of the T-method of multiple comparisons for a one-way layout with unequal variances. _Journal of the American Statistical Association_, 1976, 71, 200-203.

Hogg, R.V., and Craig, A.T., _Introduction to Mathematical Statistics_. Toronto: The Macmillan Company, 1970.

Keselman, H.J., Murry R., and Rogan, J. Effect of very unequal group sizes on Tukey's multiple comparison test. _Education and Psychological Measurement_, 1976, 36, 263-270.

Keselman, H.J., and Rogan, J. A comparison of the modified-Tukey and Scheffé methods of multiple comparisons for pairwise contrasts. _Journal of the American Statistical Association_, 1978, 73, 47-52.

Kirk, R.E. _Experimental Design: Procedures for the Behavior Sciences_. Belmont: Wadsworth Publishing Company, Inc., 1968.

Kramer, C.Y. Extension of multiple range tests to group means with un-equal number of replications. _Biometrics_, 1956, 12, 307-310.

Lee, A.F.S., and Garland, J. Size and power of tests for equality of means of two normal populations with unequal variances. _Journal of the American Statistical Association_, 1975, 70(352), 933-941.

MacLaren, M.D., and Marsaglia, G. Uniform random number generators. _Journal of the Association for Computing Machinery_, 1965, 12(1), 83-89.

Marsaglia, G., MacLaren, M.D., and Bray, T.A. A fast procedure for
    generation normal random variables. Communications of the ACM,
    1964, 7, 4-10.

Miller, R.G. Simultaneous Statistical Inference. New York: McGraw-
    Hill, 1966.

Newman, D. The distribution of the range in samples from a normal
    population, expressed in terms of an independent estimate of
    standard deviation. Biometrics, 1939, 31, 28-30.

Payne, W.H., Rabring, J.R., and Bogyo, T.P. Coding the Lehmer pseudo
    random number generator. Communications of the ACM, 1969, 12,
    85-86.

Scheffé, H. A method for judging all contrasts in the analysis of
    variance. Biometrics, 1953, 40, 87-104.

Scheffé, H. The Analysis of Variance. New York: John Wiley and Sons,
    Inc. 1959.

Smith, R.A. The effect of unequal group sizes on Tukey's HSD procedure.
    Psychometrika, 1971, 36, 31-34.

Spjøtvoll, E., and Stoline, M.R. An extension of the T-method of multi-
    ple comparison to include the cases with unequal sample sizes.
    Journal of the American Statistical Association, 1973, 68, 975-978.

Stoline, M.R. Tables of the studentized augmented range and applica-
    tions to problems of multiple comparison. Journal of the American
    Statistical Association, 1978, 73, 656-660.

Tamhane, A.C. Multiple comparisons in model I one-way ANOVA with un-
    equal variances. Communication Statistics, 1977, 1, 15-32.

Tukey, J.W. The Problem of Multiple Comparisons. Princeton University,
    1953.

Waller, R.A., and Duncan, D.B. A Bayes rule for the symmetric multiple
    comparisons problem. Journal of the American Statistical
    Association, 1969, 64, 484-503.

Welch, B.L. Further note on Mrs. Aspin's tables and on certain approxi-
    mations to the tabled functions. Biometrics, 1947, 56, 293-295.

Winer, B.J. Statistical Principles in Experimental Design. (2nd ed.)
    New York: McGraw-Hill, 1969.

APPENDIX A

Computer Program to Generate Random Numbers

48

```
C•••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••
C
C       RANDOM.FOR
C       DATE:           OCTOBER 1977
C
C       THIS REAL FUNCTION CALCULATES A RANDOM NUMBER, X,
C       SUCH THAT 0<=X<1.
C
C       THIS ROUTINE IS BASED ON THE ARTICLE
C       UNIFORM RANDOM NUMBER GENERATORS
C       M. DONALD MACLAREN AND GEORGE MARSAGLIA
C       JOURNAL OF THE ACM, VOL. 12, NO.1 (JAN.1965) PP.83-89
C
C•••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••
        REAL FUNCTION RANDOM(ISEED1,ISEED2,IFLAG)
        DIMENSION TABLE(0/127)
        EQUIVALENCE (INT,FLOAT)
C
C       SUPPRESS ERROR MESSAGES FOR INTEGER OVERFLOW
C
        CALL ERRSET(0)
C
C       IF FLAG IS NOT SET THEN INITIALIZE THE TABLE
C
        IF(IFLAG.NE.0) GO TO 1016
        IFLAG=1
        DO 1004 I=0,127
        ISEED1=(2**17+3)*ISEED1
        ISEED1=ISEED1.AND."377777777777
        CALL SHIFT(ISEED1,-8,INT)
        INT=INT.OR."200000000000
        TABLE(I)=FLOAT
1004    CONTINUE
C
C       CALCULATE RANDOM NUMBER FROM 2ND GENERATOR
C
1016    ISEED2=(2**7+1)*ISEED2+1
        ISEED2=ISEED2.AND."377777777777
        CALL SHIFT(ISEED2,-28,INDEX)
        RANDOM=TABLE(INDEX)
C
C       RESET TABLE VALUE WITH 1ST GENERATOR
C
        ISEED1=(2**17+3)*ISEED1
        ISEED1=ISEED1.AND."377777777777
        CALL SHIFT(ISEED1,-8,INT)
        INT=INT.OR."200000000000
1020    TABLE(INDEX)=FLOAT
C
C       RESET ERROR MESSAGES
C
        CALL ERRSET(2)
        RETURN
        END
```

```
      REAL FUNCTION RANDOM(ISD1,ISD2)
      DATA IFLAG/0/
C
      IF(IFLAG.EQ.0) CALL SETRAN(ISD1)
      IFLAG=1
      RANDOM=RAN(0)
      RETURN
      END
```

APPENDIX B

Computer Program to Test Random Number Generators

51

```
C**********************************************************************
C
C       RANTST.FOR
C
C       DATE:           OCTOBER 1977
C
C       THIS PROGRAM TESTS A RANDOM NUMBER GENERATOR IN THE MANNER
C       DESCRIBED IN
C
C       UNIFORM RANDOM NUMBER GENERATORS
C       M. DONALD MACLAREN AND GEORGE MARSAGLIA
C       JOURNAL OF THE ACM, VOL. 12, NO. 1 (JAN. 1965) PP.83-89
C
C**********************************************************************
        DIMENSION UNIF(100),FAIR(100),TRIP(1000),TITLE(16)
        DIMENSION NMAX(2/10),NMIN(2/10),NSUM(2/10)
        REAL CMAX(2/10),CMIN(2/10),CSUM(2/10)
        REAL CPMAX(2/10),CPMIN(2/10),CPSUM(2/10)
        REAL MAX(2/10,100),MIN(2/10,100),SUM(2/10,100),TEMP(100)
        REAL R(10)
        DOUBLE PRECISION FILE
        DATA IOUT/1/
C
C       GET OUTPUT FILE AND TITLE
C
        TYPE 1000
1000    FORMAT(' ENTER OUTPUT FILE:',$)
        ACCEPT 1001,FILE
1001    FORMAT(A10)
        OPEN(UNIT=IOUT,DEVICE='DSK',ACCESS='SEQOUT',FILE=FILE)
        TYPE 1002
1002    FORMAT(' ENTER TITLE')
        ACCEPT 1003,TITLE
1003    FORMAT(16A5)
C
C       INPUT PROGRAM PARAMETERS
C
        TYPE 1004
1004    FORMAT(' ENTER THE NUMBER OF TESTS TO BE RUN:',$)
        ACCEPT 1005,NTESTS
1005    FORMAT(I)
        TYPE 1006
1006    FORMAT(' ENTER THE NUMBER OF SETS OF 10 TO BE USED:',$)
        ACCEPT 1005,NSETS
        TYPE 1007
1007    FORMAT(' ENTER THE SEED FOR THE 1ST RANDOM NUMBER GENERATOR:',$)
        ACCEPT 1005,ISEED1
        TYPE 1008
1008    FORMAT(' ENTER THE SEED FOR THE 2ND RANDOM NUMBER GENERATOR:',$)
        ACCEPT 1005,ISEED2
        TYPE 1009
1009    FORMAT(' DO YOU WANT ONLY THE SUMMARY (Y OR N)?',$)
        ACCEPT 1010,ANS
1010    FORMAT(A1)
        ISUMM=0
        IF(ANS.EQ.'Y') ISUMM=1
        ISD1=ISEED1
        ISD2=ISEED2
```

```
C
C          NTESTS PER RUN, NSETS OF 10 PER RUN.
C
           DO 2600 NTEST=1,NTESTS
           ISEED1=ISD1
           ISEED2=ISD2
C
C          ZERO COUNTERS
C
           NUNIF=0
           NPAIR=0
           NTRIP=0
           DO 1012 I=2,10
           NMAX(I)=0
           NMIN(I)=0
           NSUM(I)=0
1012       CONTINUE
           DO 1013 I=1,100
           UNIF(I)=0
           PAIR(I)=0
1013       CONTINUE
           DO 10131 I=1,1000
           TRIP(I)=0
10131      CONTINUE
           DO 1015 I=2,10
           DO 1014 J=1,100
           MAX(I,J)=0
           MIN(I,J)=0
           SUM(I,J)=0
1014       CONTINUE
1015       CONTINUE
C
           DO 1050 NSET=1,NSETS
C
C          CALCULATE SETS OF 10 RANDOM NUMBERS.
C
           DO 1016 I=1,10
           R(I)=RANDOM(ISD1,ISD2,IFLG)
1016       CONTINUE
C
C          UNIFORMITY TEST
C
           ICELL=R(1)*100+1
           IF(ICELL.LT.1.OR.ICELL.GT.100) GO TO 1018
           NUNIF=NUNIF+1
           UNIF(ICELL)=UNIF(ICELL)+1
C
C          PAIRS TEST
C
1018       DO 1020 I=1,5
           ICELL=R((I-1)*2+1)*10+1
           JCELL=R((I-1)*2+2)*10+1
           INDEX=(ICELL-1)*10+JCELL
           IF(INDEX.LT.1.OR.INDEX.GT.100) GO TO 1020
           PAIR(INDEX)=PAIR(INDEX)+1
           NPAIR=NPAIR+1
1020       CONTINUE
C
C          TRIPLES TEST
C
```

```
1022        DO 1024 I=1,3
            ICELL=R((I-1)*3+1)*10+1
            JCELL=R((I-1)*3+2)*10+1
            KCELL=R((I-1)*3+3)*10+1
            INDEX=(ICELL-1)*100+(JCELL-1)*10+KCELL
            IF(INDEX.LT.1.OR.INDEX.GT.1000) GO TO 1026
            TRIP(INDEX)=TRIP(INDEX)+1
            NTRIP=NTRIP+1
1024        CONTINUE
C
C           MAXIMUM OF 2-10
C
1026        DO 1032 I=2,10
            CALL MAXIMUM(R,I+0,W)
            FW=W**I
            ICELL=FW*100+1
            IF(ICELL.LT.1.OR.ICELL.GT.100) GO TO 1032
            MAX(I,ICELL)=MAX(I,ICELL)+1
            NMAX(I)=NMAX(I)+1
1032        CONTINUE
C
C           MINIMUM OF 2-10
C
            DO 1040 I=2,10
            CALL MINIMUM(R,I+0,W)
            FW=1-(1-W)**I
            ICELL=FW*100+1
            IF(ICELL.LT.1.OR.ICELL.GT.100) GO TO 1040
            MIN(I,ICELL)=MIN(I,ICELL)+1
            NMIN(I)=NMIN(I)+1
1040        CONTINUE
C
C           SUM OF 2-10
C
            DO 1048 I=2,2
            SUMR=R(1)
            DO 1044 N=2,I
            SUMR=SUMR+R(N)
            IF(SUMR.LT.1) FSUM=SUMR**2/2
            IF(SUMR.GE.1) FSUM=1-(2-SUMR)**2/2
            SUMR=FSUM
1044        CONTINUE
            ICELL=FSUM*100+1
            IF(ICELL.LT.1.OR.ICELL.GT.100) GO TO 1048
            SUM(I,ICELL)=SUM(I,ICELL)+1
            NSUM(I)=NSUM(I)+1
1048        CONTINUE
C
C           END SETS LOOP
C
1050        CONTINUE
C
C           CALCULATE CHI-SQUARE STATISTIC AND ASSOCIATED PROBABILITY.
C
            CUNIF=CHI(UNIF,100,NUNIF)
            CALL CHIPRB(CUNIF,99,CPUNIF)
            CPAIR=CHI(PAIR,100,NPAIR)
            CALL CHIPRB(CPAIR,99,CPPAIR)
            CTRIP=CHI(TRIP,1000,NTRIP)
            CALL CHIPRB(CTRIP,999,CPTRIP)
```

```
C
      DO 1104 I=2,10
      DO 1100 J=1,100
      TEMP(J)=MAX(I,J)
1100  CONTINUE
      CMAX(I)=CHI(TEMP,100,NMAX(I))
      CALL CHIPRB(CMAX(1),99,CPMAX(1))
1104  CONTINUE
C
      DO 1112 I=2,10
      DO 1108 J=1,100
      TEMP(J)=MIN(I,J)
1108  CONTINUE
      CMIN(I)=CHI(TEMP,100,NMIN(I))
      CALL CHIPRB(CMIN(1),99,CPMIN(1))
1112  CONTINUE
C
      DO 1120 I=2,2
      DO 1116 J=1,100
      TEMP(J)=SUM(I,J)
1116  CONTINUE
      CSUM(I)=CHI(TEMP,100,NSUM(I))
      CALL CHIPRB(CSUM(1),99,CPSUM(1))
1120  CONTINUE
C
C     GENERATE OUTPUT
C
C     UNIFORM
C
      IF(ISUMM.EQ.1) GO TO 2062
      WRITE(IOUT,2004) TITLE
2004  FORMAT('1',16A5//' UNIFORMITY TEST'//)
      WRITE(IOUT,2006) NTEST,ISEED1,ISEED2
2006  FORMAT(' RUN=',I2,' SEED 1 =',I15,'    SEED 2=',I15//)
      WRITE(IOUT,2008) UNIF
2008  FORMAT(' CELL FREQUENCIES1'//(10FB,0))
      WRITE(IOUT,2012) NUNIF,CUNIF,CPUNIF
2012  FORMAT(//
     &  ' NUMBER OF OBSERVATIONS=',T30,I10/
     &  ' CHI-SQUARE=',T30,F10.3/
     &  ' PROBABILITY=',T30,F10.3//)
C
C     PAIRS
C
      WRITE(IOUT,2016) TITLE
2016  FORMAT('1',16A5//' PAIRS TEST'//)
      WRITE(IOUT,2006) NTEST,ISEED1,ISEED2
      WRITE(IOUT,2008) PAIR
      WRITE(IOUT,2012) NPAIR,CPAIR,CPPAIR
C
C     TRIPLES
C
      DO 2024 I=1,10
      WRITE(IOUT,2020) TITLE,I
2020  FORMAT('1',16A5//' TRIPLES TEST I=',I2//)
      WRITE(IOUT,2006) NTEST,ISEED1,ISEED2
      WRITE(IOUT,2008) (TRIP(J),J=(I-1)*100+1,I*100)
2024  CONTINUE
      WRITE(IOUT,2012) NTRIP,CTRIP,CPTRIP
C
```

```
C
C         MAXIMUMS
C
          DO 2040 I=2,10
          WRITE(IOU1,2030) TITLE,I
2030      FORMAT('1',16A5//' MAXIMUM TEST OF ',I2//)
          WRITE(IOU1,2006) NTEST,ISEED1,ISEED2
          WRITE(IOU1,2008) (MAX(I,J),J=1,100)
          WRITE(IOU1,2012) NMAX(I),CMAX(I),CPMAX(I)
2040      CONTINUE
C
C         MINIMUMS
C
          DO 2050 I=2,10
          WRITE(IOU1,2044) TITLE,I
2044      FORMAT('1',16A5//' MINIMUM TEST OF ',I2//)
          WRITE(IOU1,2006) NTEST,ISEED1,ISEED2
          WRITE(IOU1,2008) (MIN(I,J),J=1,100)
          WRITE(IOU1,2012) NMIN(I),CMIN(I),CPMIN(I)
2050      CONTINUE
C
C         SUM
C
          DO 2060 I=2,2
          WRITE(IOU1,2054) TITLE,I
2054      FORMAT('1',16A5//' SUM TEST OF ',I2//)
          WRITE(IOU1,2006) NTEST,ISEED1,ISEED2
          WRITE(IOU1,2008) (SUM(I,J),J=1,100)
          WRITE(IOU1,2012) NSUM(I),CSUM(I),CPSUM(I)
2060      CONTINUE
C
C         OUTPUT SUMMARY
C
2062      WRITE(IOU1,2064) TITLE
2064      FORMAT('1',16A5//' SUMMARY'///)
          WRITE(IOU1,2006) NTEST,ISEED1,ISEED2
          WRITE(IOU1,2068)
2068      FORMAT(T25,'OBSERVATIONS   CHI-SQUARE  PROBABILITY'/
     &    T25,36('-'))
          WRITE(IOU1,2072) NUNIF,CUNIF,CPUNIF
2072      FORMAT(' UNIFORM TEST',T25,I12,2F12.4)
          WRITE(IOU1,2076) NPAIR,CPAIR,CPPAIR
2076      FORMAT(' PAIR TEST',T25,I12,2F12.4)
          WRITE(IOU1,2080) NTRIP,CTRIP,CPTRIP
2080      FORMAT(' TRIPLES TEST',T25,I12,2F12.4)
          DO 2092 I=2,10
          WRITE(IOU1,2088) I,NMAX(I),CMAX(I),CPMAX(I)
2088      FORMAT(' MAXIMUM OF ',I2,T25,I12,2F12.4)
2092      CONTINUE
          DO 2100 I=2,10
          WRITE(IOU1,2096) I,NMIN(I),CMIN(I),CPMIN(I)
2096      FORMAT(' MINIMUM OF ',I2,T25,I12,2F12.4)
2100      CONTINUE
          DO 2108 I=2,2
          WRITE(IOU1,2104) I,NSUM(I),CSUM(I),CPSUM(I)
2104      FORMAT(' SUM OF ',I2,T25,I12,2F12.4)
2108      CONTINUE
C
C         END LOOP
C
2000      CONTINUE
```

```
        STOP
        END
C*************************************************************************
C
C       REAL FUNCTION CHI
C
C       THIS FUNCTION CALCULATES THE CHI-SQUARE STATISTIC.
C
C       INPUT VARIABLES:
C
C               X--A VECTOR CONTAINING THE NUMBER OF OCCURANCES IN
C                  EACH CELL
C               M--LENGTH OF VECTOR X
C               N--SUM OF ALL CELLS IN VECTOR X
C
C*************************************************************************
        REAL FUNCTION CHI(X,M,N)
        DIMENSION X(1)
C
        P=1./M
        CHI=0
        DO 3000 I=1,M
        CHI=CHI+(X(I)-N*P)**2/(N*P)
3000    CONTINUE
        RETURN
        END
```

APPENDIX C

Computer Programs to Duplicate Portions
of the Carmer-Swanson Study and Test
Normal Data

58

```
C***************************************************************
C       TCS2.FON
C
C       DATE: JANUARY 1978
C
C       THIS PROGRAM DUPLICATES PARTS OF THE CARMER AND
C       SWANSON STUDY, JOURNAL OF THE ASA, VOL. 68, NO.341
C       (MARCH 1973) PP. 66-74.
C       IN ADDITION TO THE ABOVE, IT CAN USE AN ALTERNATE
C       RANDOM NUMBER GENERATOR, DEC-10, AND TESTS NORMAL
C       DATA.
C
C***************************************************************
C
        DIMENSION A(8),Q(8),F(8),T(8),E(16),NC(16),CE(6,6)
        DIMENSION EE(6,8),C(6,8),CK(6,4),ID(6,8)
        DIMENSION X(20,20),CH(16),SUMTR(20),XM(20),SSQ(20)
        DIMENSION CT(6),CS(6),CLSD(6),DXM(20,20)
        ISD1=1
        ISD2=8
5       TYPE 10
10      FORMAT (' ENTER NUMBER OF TREATMENTS')
        ACCEPT 15,N1
15      FORMAT (I)
        TYPE 20
20      FORMAT (' ENTER NUMBER IN EACH TREATMENT')
        ACCEPT 25,NIN
25      FORMAT (I2)
        TYPE 30
30      FORMAT (' ENTER NUMBER OF EXPERIMENTS')
        ACCEPT 35,NEX
35      FORMAT (I)
        TYPE 40
40      FORMAT (' ENTER NUMBER OF SIG. LEVELS')
        ACCEPT 45,NS
45      FORMAT (I)
        TYPE 50
50      FORMAT (' ENTER THE SIG. LEVELS')
        DO 60 I=1,NS
        ACCEPT 55,A(I)
55      FORMAT (F6.5)
60      CONTINUE
        TYPE 70
70      FORMAT (' ENTER Q-VALUES')
        DO 80 I=1,NS
        ACCEPT 75,Q(I)
75      FORMAT (F6.4)
80      CONTINUE
        TYPE 90
90      FORMAT (' ENTER F-VALUES')
        DO 100 I=1,NS
        ACCEPT 95,F(I)
95      FORMAT (F6.4)
100     CONTINUE
        TYPE 110
110     FORMAT (' ENTER T-VALUES')
        DO 120 I=1,NS
        ACCEPT 115,T(I)
115     FORMAT (F6.4)
120     CONTINUE
```

```
        IFLAG=0
        NM=NT*NIN*NEX
        E(1)=.2517*NM
        E(3)=.148*NM
        E(5)=.0500*NM
        E(7)=.0245*NM
        E(9)=.0151*NM
        E(11)=.005*NM
        E(13)=.0036*NM
        E(15)=.0013*NM
        DO 150 I=1,15,2
        E(I+1)=E(I)
150     CONTINUE
        DO 160 I=1,16
        NC(I)=0
160     CONTINUE
        DO 170 J=1,4
        DO 170 I=1,NS
        CE(I,J)=0
        EE(I,J)=0
170     CONTINUE
        DO 180 J=1,8
        DO 180 I=1,NS
        C(I,J)=0
180     CONTINUE
C
C       THE OUTER LOOP
        DO 400 M=1,NEX
        DO 190 J=1,4
        DO 190 I=1,NS
        CK(I,J)=0
        ID(I,J+4)=0
190     CONTINUE
C       GENERATE NORMAL DATA
        DO 200 J=1,NT
        DO 200 I=1,NIN,2
        R1=RANDOM(ISD1,ISD2,IFLAG)
        R2=RANDOM(ISD1,ISD2,IFLAG)
        U=SQRT(-2*ALOG(R1))
        X(I,J)=U*COS(2*3.14159*R2)
        X(I+1,J)=U*SIN(2*3.14159*R2)
200     CONTINUE
C       TEST OF NORMAL DATA
        DO 210 J=1,NT
        DO 210 I=1,NIN
        IF (X(I,J).GE.0.AND.X(I,J).LT..68) IND=1
        IF (X(I,J).GE..68.AND.X(I,J).LT.1.28) IND=3
        IF (X(I,J).GE.1.28.AND.X(I,J).LT.1.65) IND=5
        IF (X(I,J).GE.1.65.AND.X(I,J).LT.1.96) IND=7
        IF (X(I,J).GE.1.96.AND.X(I,J).LT.2.33) IND=9
        IF (X(I,J).GE.2.33.AND.X(I,J).LT.2.58) IND=11
        IF (X(I,J).GE.2.58.AND.X(I,J).LT.3.) IND=13
        IF (X(I,J).GE.3.,) IND=15
        IF (X(I,J).GE.-.68.AND.X(I,J).LT.0.0) IND=2
        IF (X(I,J).GE.-1.28.AND.X(I,J).LT.-.68) IND=4
        IF (X(I,J).GE.-1.65.AND.X(I,J).LT.-1.28) IND=6
        IF (X(I,J).GE.-1.96.AND.X(I,J).LT.-1.65) IND=8
        IF (X(I,J).GE.-2.33.AND.X(I,J).LT.-1.96) IND=10
        IF (X(I,J).GE.-2.58.AND.X(I,J).LT.-2.33) IND=12
        IF (X(I,J).GE.-3.0.AND.X(I,J).LT.-2.58) IND=14
```

```
                IF (X(I,J).LT.-3.0) IND=16
                NC(IND)=NC(IND)+1
210        CONTINUE
C          CALCULATE MEAN FOR TREATMENTS
           DO 240 J=1,NT
           SUMTR(J)=0
240        CONTINUE
           DO 250 J=1,NT
           DO 250 I=1,NIN
           SUMTR(J)=SUMTR(J)+X(I,J)
250        CONTINUE
           DO 260 J=1,NT
           XM(J)=SUMTR(J)/NIN
260        CONTINUE
C          CALCULATE WITHIN MEAN SUM OF SQUARES
           SUMXS=0
           DO 270 J=1,NT
           DO 270 I=1,NIN
           SUMXS=SUMXS+(X(I,J))**2
270        CONTINUE
           SUMTS=0
           DO 280 J=1,NT
           SSQ(J)=SUMTR(J)**2
           SUMTS=SUMTS+SSQ(J)
280        CONTINUE
           A1=SUMTS/NIN
           WMS=(SUMXS-A1)/(NT*(NIN-1))
C          CALCULATE BETWEEN MEAN SUM OF SQUARES
           SUMG=0
           DO 290 J=1,NT
           SUMG=SUMG+SUMTR(J)
290        CONTINUE
           SUMGS=(SUMG**2)/(NT*NIN)
           BMS=(A1-SUMGS)/(NT-1)
C          F-VALUE
           FV=BMS/WMS
C          CALCULATE ALL CRITICAL VALUES
           SC=SQRT((WMS/NIN)*2)
           DO 300 I=1,NS
           CT(I)=(SQRT(WMS/NIN))*Q(I)
           CS(I)=(SQRT((NT-1)*F(I)))*SC
           CLSD(I)=SC*T(I)
300        CONTINUE
C          CALCULATE DIFFERENCE BETWEEN MEANS
           DO 310 I=1,NT-1
           DO 310 J=I+1,NT
           DXM(I,J)=ABS(XM(I)-XM(J))
310        CONTINUE
C          MULTIPLE COMPARISON TESTS
           DO 330 L=1,NS
           DO 330 I=1,NT-1
           DO 330 J=I+1,NT
           DO 320 K=1,4
           ID(L,K)=0
320        CONTINUE
           IF (DXM(I,J).GT.CT(L)) ID(L,1)=1
           IF (DXM(I,J).GT.CS(L)) ID(L,2)=1
           IF (DXM(I,J).GT.CLSD(L)) ID(L,3)=1
           IF (DXM(I,J).GT.CLSD(L).AND.FV.GT.F(L)) ID(L,4)=1
           DO 330 K1=1,4
```

```
        C(L,K1)=C(L,K1)+ID(L,K1)
        CK(L,K1)=CK(L,K1)+ID(L,K1)
330     CONTINUE
C
        DO 400 I=1,NS
        DO 400 J=5,8
        IF (CK(I,J=4).GE.1.0) ID(I,J)=1
        C(I,J)=C(I,J)+ID(I,J)
400     CONTINUE
C
        CHT=0
        DO 410 I=1,16
        CH(I)=((E(I)-NC(I))**2)/E(I)
        CHT=CHT+CH(I)
410     CONTINUE
        NEF=((NI*(NI-1))/2)*NEX
        DO 420 I=1,NS
        DO 420 J=1,4
        CE(I,J)=C(I,J)/NEF
        EE(I,J+4)=C(I,J+4)/NEX
420     CONTINUE
        TYPE 430,NT,NIN
430     FORMAT ('0',10X,'NO. TREA1=',I2,5X,'NO. IN EACH=',I2)
        TYPE 440
440     FORMAT ('0',15X,'TUKEY',12X,'SCHEFFE',10X,'LSD',13X,'FLSD')
        DO 490 I=1,NS
        TYPE 450,A(I)
450     FORMAT ('0',32X,'LEVEL OF SIG=',F6.5)
        TYPE 470,CE(I,1),CE(I,2),CE(I,3),CE(I,4)
470     FORMAT ('0','CMP ERROR',5X,F8.6,9X,F8.6,9X,F8.6,8X,F8.6)
        TYPE 480,EE(I,5),EE(I,6),EE(I,7),EE(I,8)
480     FORMAT ('0','EXP ERROR',5X,F8.6,9X,F8.6,9X,F8.6,8X,F8.6)
490     CONTINUE
        TYPE 500
500     FORMAT ('0','CELL',8X,'NUMBER',8X,'EXP',9X,'CHI')
        DO 520 I=1,16
        TYPE 510,I,NC(I),E(I),CH(I)
510     FORMAT ('0',I2,6X,I9,4X,F8.2,3X,F8.2)
520     CONTINUE
        TYPE 530,CHT
530     FORMAT ('0','CHI-SQUARE WITH 15 D.F. =',F8.2)
        TYPE 540
540     FORMAT (' TYPE 1 TO CONTINUE 0 TO EXIT')
        ACCEPT 550,IC
550     FORMAT (I)
        IF (IC.GT.0) GO TO 5
        STOP
        END
```

APPENDIX D

Computer Program for the
Simulation Study

63

```
C***********************************************************
C         PROGRAM NAME: SIM1.FOR
C
C         DATE: MARCH 1978
C
C         THIS PROGRAM INVOLVES A SIMULATION STUDY
C         COMPARING NINE MULTIPLE COMPARISON
C         TECHNIQUES WITH RESPECT TO EXPERIMENTWISE
C         AND COMPARISONWISE ERROR RATES.
C
C***********************************************************
          DIMENSION IN(20),IH(4),CE(4,9),EE(4,18),C(4,18)
          DIMENSION CK(4,9),IU(4,18),X(52,20),SUMTR(20)
          DIMENSION XM(20),SUMV(20),VAR(20),D(20)
          DIMENSION DAM(20,20),A(20,20),B(20,20),C2(20,20)
          DIMENSION E(20,20),CGH(20,20,4)
          DIMENSION IGH(20,20),FL(4),ICON(4)
          DIMENSION CTV(4),CFV(4),CUV(4),CHV(4),CUPV(4)
          DIMENSION CUPHV(4),CGGHV(20,20,4),CINNV(20,20,4)
          DIMENSION CITHV(20,20,4),CLSC(20,20,4)
          DIMENSION CIP(20,20,4),CTHV(20,20,4)
          DIMENSION CGT2(20,20,4),C1PF(20,20,4),CKR(20,20,4)
          DIMENSION CGH(20,20,4),SICON(4)
          DIMENSION C1H(20,20,4),CNN(20,20,4)
          COMMON /AREA1/XSO(20,4,4),XSUP(12,2,4),XST(34,1,4),XSTH(33,3,4)
          ISD1=1
          ISD2=0
          IFLAG=0
          TYPE 05
05        FORMAT (' INTER 03 IF LINE PRINT 05 IF TERMINAL')
          ACCEPT 10,IUNIT
10        FORMAT (I)
          OPEN (UNIT=1,DEVICE='DSK',ACCESS='SEQIN',FILE='ISO.DAT')
          DO 14 J=1,4
          DO 14 K=1,20
          READ (1,12) (XSO(K,J,I),I=1,4)
12        FORMAT (4F6.3)
14        CONTINUE
          OPEN (UNIT=2,DEVICE='DSK',ACCESS='SEQIN',FILE='ISUP.DAT')
          DO 16 J=1,2
          DO 16 K=1,12
          READ (2,15) (XSUP(K,J,I),I=1,4)
15        FORMAT (4F6.3)
16        CONTINUE
          OPEN (UNIT=4,DEVICE='DSK',ACCESS='SEQIN',FILE='IST.DAT')
          DO 17 K=1,34
          READ (4,15) (XST(K,1,I),I=1,4)
17        CONTINUE
          OPEN (UNIT=9,DEVICE='DSK',ACCESS='SEQIN',FILE='ISTH.DAT')
          DO 18 J=1,3
          DO 18 K=1,33
          READ (9,15) (XSTH(K,J,I),I=1,4)
18        CONTINUE
C         TREATMENT PARAMETERS
20        TYPE 24
24        FORMAT (' ENTER NUMBER OF TREATMENTS')
          ACCEPT 25,N1
25        FORMAT (I)
          TYPE 30
30        FORMAT (' ENTER NUMBER IN TREATMENTS')
```

```
              DO 40 I=1,NT
              ACCEPT 35,IN(I)
35            FORMAT (I)
40            CONTINUE
              TYPE 50
50            FORMAT (' ENTER NUMBER OF EXPERIMENTS')
              ACCEPT 55,NEX
55            FORMAT (I)
              INS=IN(1)
              DO 60 I=2,NT
              IF (IN(I).LT.INS) INS=IN(I)
60            CONTINUE
C             COUNTERS TO ZERO
63            FORMAT (' ',I2)
              DO 65 J=1,9
              DO 65 I=1,4
              CE(I,J)=0
              EE(I,J+9)=0
65            CONTINUE
              DO 70 J=1,18
              DO 70 I=1,4
              C(I,J)=0
70            CONTINUE
              TOTIN=0
              DO 75 I=1,NT
              TOTIN=TOTIN+IN(I)
75            CONTINUE
              FL(1)=.01
              FL(2)=.05
              FL(3)=.10
              FL(4)=.20
              ID2=TOTIN-NT
              IK=(NT*(NT-1))/2
              IP=INS-1
              CON=1./IK
              DO 80 I=1,4
              SICOM(I)=(1.0-FL(I))**CON
              ICON(I)=1.-SICOM(I)
80            CONTINUE
              IF (NT.EQ.5) NT1=1
              IF (NT.EQ.10) NT1=2
              IF (NT.EQ.20) NT1=3
              IF (NT.EQ.3) NT1=4
              DO 85 I1=1,4
              I=I1
              CTV(I)=FVALUE(FL(I),1,ID2)**.5
              CFV(I)=FVALUE(FL(I),NT-1,ID2)
              CWV(I)=SQ(ID2,NT1,I)
              IF (NT.NE.10) GO TO 83
              IF (I.EQ.1.AND.ID2.GT.120) GO TO 81
              GO TO 83
81            XMID=((1./120)-(1./ID2))/(1./120)
              CMV(I)=(3.805)*(1.-XMID)+(3.691)*XMID
              GO TO 85
83            CMV(I)=SMMINV(IK,FL(I),ID2)
85            CONTINUE
87            FORMAT ('0',F6.2)
              IF (NT1.EQ.1) GO TO 95
              DO 90 I1=1,4
              I=I1
```

```
              CQPV(I)=CQV(I)
              CQPPV(I)=SQ(IP,NT1,I)
90            CONTINUE
              GO TO 100
95            I=I1
              DO 100 II=1,4
              CQPV(I)=SQP(ID2,NT1,I)
              CQPPV(I)=SQP(IP,NT1,I)
              TYPE 98,CQPV(I),CQPPV(I)
98            FORMAT (' ',2F8.4)
100           CONTINUE
C             OUTER LOOP
              DO 500 M=1,NEX
              DO 150 J=1,9
              DO 150 I=1,4
              CK(I,J)=0
              ID(I,J+9)=0
150           CONTINUE
C             NORMAL DATA
              DO 155 J=1,NT
              DO 155 I=1,IN(J),2
              R1=RANDOM(ISD1,ISD2,IFLAG)
              R2=RANDOM(ISD1,ISD2,IFLAG)
              U=SQRT(-2*ALOG(R1))
              X(I,J)=U*COS(2*3.14159*R2)
              X(I+1,J)=U*SIN(2*3.14159*R2)
155           CONTINUE
C             TREATMENT MEANS
              DO 160 I=1,NT
              SUMTR(I)=0
160           CONTINUE
              DO 165 J=1,NT
              DO 165 I=1,IN(J)
              SUMTR(J)=SUMTR(J)+X(I,J)
165           CONTINUE
              DO 170 J=1,NT
              XM(J)=SUMTR(J)/IN(J)
170           CONTINUE
C             TREATMENT VARIANCES
              DO 180 J=1,NT
              SUMV(J)=0
180           CONTINUE
              DO 185 J=1,NT
              DO 185 I=1,IN(J)
              SUMV(J)=SUMV(J)+(X(I,J)-XM(J))**2
185           CONTINUE
              DO 188 J=1,NT
              VAR(J)=SUMV(J)/(IN(J)-1)
188           CONTINUE
C             WITHIN MEAN SUM OF SQUARES
              SUMVR=0
              TOTIN=0
              DO 190 J=1,NT
              SUMVR=SUMVR+SUMV(J)
              TOTIN=TOTIN+IN(J)
190           CONTINUE
              WMS=SUMVR/(TOTIN-NT)
C             BETWEEN MEAN SUM OF SQUARES
              SUMGR=0
              SUMTS=0
```

```
            DO 200 J=1,NT
            SUMGR=SUMGR+SUMTH(J)
            SUMTS=SUMTS+(SUMTH(J)**2)/IN(J)
200         CONTINUE
            BMS=(SUMTS-((SUMGR**2)/TOTIN))/(NT-1)
C           DIFFERENCE BETWEEN MEANS
            DO 210 I=1,NT-1
            DO 210 J=I+1,NT
            DXM(I,J)=ABS(XM(I)-XM(J))
210         CONTINUE
C           SUBFORMULAS
            DO 240 I=1,NT-1
            DO 240 J=I+1,NT
            A(I,J)=((1,/IN(I))+(1,/IN(J)))**.5
            B(I,J)=((VAR(I)/IN(I))+(VAR(J)/IN(J)))**.5
            IF (IN(I)/IN(J).GE.1) GO TO 220
            C2(I,J)=(1,/IN(I))**.5
            GO TO 225
220         C2(I,J)=(1,/IN(J))**.5
225         D(I)=(VAR(I)/IN(I))**.5
            D(J)=(VAR(J)/IN(J))**.5
            IF (D(I).GE.D(J)) GO TO 230
            E(I,J)=D(J)
            GO TO 235
230         E(I,J)=D(I)
235         VI=VAR(I)/IN(I)
            VJ=VAR(J)/IN(J)
            VIS=VI**2
            VJS=VJ**2
            IGH(I,J)=((VI+VJ)**2)/(VIS/(IN(I)-1)+VJS/(IN(J)-1))
240         CONTINUE
C           CRITICAL VALUES
            I2=1
            F=BMS/WMS
250         FORMAT (' ',F6,4)
            DO 380 I1=1,NT-1
            I=I1
            DO 380 J1=I+1,NT
            J=J1
            DO 380 K1=1,4
270         FORMAT (' ',I4)
280         FORMAT (' ',F8,4)
            K=K1
            CUGHV(I,J,K)=SU(IGH(I,J),NT1,K)
            CINNV(I,J,K)=ST(IGH(I,J),1,K)
            IF (NT1.EQ.4) GO TO 300
            CTHV(I,J,K)=SIH(IGH(I,J),NT1,K)
            GO TO 380
300         CTHV(I,J,K)=SIH(IGH(I,J),3,K)
C
380         CONTINUE
C           COMPARISON VALUES
            DO 400 K=1,4
            DO 400 I=1,NT-1
            DO 400 J=I+1,NT
            CLSD(I,J,K)=A(I,J)*(WMS**.5)*CTV(K)
            CTP(I,J,K)=C2(I,J)*(WMS**.5)*CGFV(K)
            CGT2(I,J,K)=A(I,J)*(WMS**.5)*CHV(K)
            CTPP(I,J,K)=E(I,J)*CGPPV(K)
            CKR(I,J,K)=A(I,J)*((WMS**.5)/(2**.5))*CGV(K)
```

```
            CGH(I,J,K)=B(I,J)/(2**.5)*COGHV(1,J,K)
            CNN(I,J,K)=B(I,J)*CINNV(I,J,K)
            CTH(I,J,K)=B(I,J)*CTHV(I,J,K)
390     FORMAT (' ',F8.4)
400     CONTINUE
C       COMPARISON TESTS
        DO 450 L=1,4
        DO 450 I=1,NT-1
        DO 450 J=I+1,NT
        DO 410 K=1,9
        ID(L,K)=0
410     CONTINUE
        IF (DXM(I,J).GT.CLSD(I,J,L)) ID(L,1)=1
        IF (DXM(I,J).GT.CLSD(I,J,L).AND.F.GT.CFV(L)) ID(L,2)=1
        IF (DXM(I,J).GT.CTP(I,J,L)) ID(L,3)=1
        IF (DXM(I,J).GT.CGT2(I,J,L)) ID(L,4)=1
        IF (DXM(I,J).GT.CIPP(I,J,L)) ID(L,5)=1
        IF (DXM(I,J).GT.CKR(I,J,L)) ID(L,6)=1
        IF (DXM(I,J).GT.CGH(I,J,L)) ID(L,7)=1
        IF (DXM(I,J).GT.CIH(I,J,L)) ID(L,8)=1
        IF (DXM(I,J).GT.CNN(I,J,L).AND.F.GT.CFV(L)) ID(L,9)=1
        DO 450 K1=1,9
        C(L,K1)=C(L,K1)+ID(L,K1)
        CK(L,K1)=CK(L,K1)+ID(L,K1)
450     CONTINUE
C
        DO 500 I=1,4
        DO 500 J=10,18
        IF (CK(I,J-9).GE.1.0) ID(I,J)=1
        C(I,J)=C(I,J)+ID(I,J)
500     CONTINUE
        DO 550 I=1,4
        DO 550 J=1,9
        CE(I,J)=C(I,J)/(IK*NEX)
        EE(I,J+9)=C(I,J+9)/NEX
550     CONTINUE
C       PRINTING
        WRITE (IUNI1,560) N1
560     FORMAT ('0','NUMBER OF TREATMENTS IS',I2)
        WRITE (IUNI1,570) (IN(I),I=1,NT)
570     FORMAT ('0','NUMBER IN EACH TREATMENT',10I3)
        WRITE (IUNI1,590)
590     FORMAT ('0',19X,'COMP. ERROR',2X,'EXP. ERROR')
        DO 700 I=1,4
        WRITE (IUNI1,600) FL(I)
600     FORMAT ('0',32X,'LEVEL OF SIG. IS',1X,F4.3)
        WRITE (IUNI1,610) CE(I,1),EE(I,10)
610     FORMAT ('0','LSD',15X,F8.6,5X,F8.6)
        WRITE (IUNI1,620) CE(I,2),EE(I,11)
620     FORMAT ('0','PLSD',14X,F8.6,5X,F8.6)
        WRITE (IUNI1,630) CE(I,3),EE(I,12)
630     FORMAT ('0','T-PRIME',11X,F8.6,5X,F8.6)
        WRITE (IUNI1,640) CE(I,4),EE(I,13)
640     FORMAT ('0','GT2',15X,F8.6,5X,F8.6)
        WRITE (IUNI1,650) CE(I,5),EE(I,14)
650     FORMAT ('0','T-2PRIME',10X,F8.6,5X,F8.6)
        WRITE (IUNI1,660) CE(I,6),EE(I,15)
660     FORMAT ('0','KRAMER',12X,F8.6,5X,F8.6)
        WRITE (IUNI1,670) CE(I,7),EE(I,16)
670     FORMAT ('0','GAMES-HOWELL',6X,F8.6,5X,F8.6)
```

```
        WRITE (IUNIT,680) CE(1,8),EE(1,17)
680     FORMAT ('0','IAMHAME',11X,F8.6,5X,F8.6)
685     WRITE (IUNIT,690) CE(1,9),EE(1,18)
690     FORMAT ('0','P=wELCH',11X,F8.6,5X,F8.6)
700     CONTINUE
        TYPE 710
710     FORMAT ('0','TYPE 1 TO CONTINUE & TE STOP')
        ACCEPT 720,ICE
720     FORMAT (I)
        IF (ICE.EQ.1) GO TO 20
        STOP
        END
```

```
C************************************************************
C        FUNCTION   SQP(IDF,NT,IA)
C        SQP1.FOR IS A FUNCTION USED TO FIND THE AUGMENTED
C        STUDENTIZED Q-PRIME-VALUES BY USING SQP(IDF,NT,IA),
C
C************************************************************
         FUNCTION SQP(IDF,NT,IA)
         DIMENSION INDEX(12)
         COMMON /AREA1/XSQ(26,4,4),XSQP(12,2,4),XS1(34,1,4),XSTH(33,3,4)
         DATA INDEX/5,7,10,12,16,20,24,30,40,60,120,1000000/
         IF (IDF.LE.4.OR.IDF.GT.1000000) CALL EXIT
         DO 30 I=1,12
         IF (INDEX(I).LE.IDF) L=I
30       CONTINUE
         XLO=INDEX(L)
         XUP=INDEX(L+1)
         XMID=((1./IDF)-(1./XLOW))/((1./XUP)-(1./XLOW))
         SQP=XSQP(L,NT,IA)*(1-XMID)+XSQP(L+1,NT,IA)*XMID
         RETURN
         END
```

```
C*****************************************************************
C       FUNCTION   SU(IA,NT,IDF)
C       SQ1,FOR IS A FUNCTION USED TO FIND STUDENTIZED
C       Q-VALUES BY USING SU(IA,NT,IDF).
C
C*****************************************************************
        FUNCTION SU(IDF,NT,IA)
        DIMENSION INDEX(26)
        COMMON /AREA1/XSQ(26,4,4),XSQP(12,2,4),XSI(34,1,4),XSTH(33,3,4)
        IF (IDF.LT.1.OR.IDF.GT.1000000) CALL EXIT
        DATA INDEX/1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,
     1  19,20,24,30,40,60,120,1000000/
        DO 30 I=1,26
        IF (INDEX(I).LE.IDF) L=I
30      CONTINUE
        XLOW=INDEX(L)
        XUP=INDEX(L+1)
        XMID=((1./IDF)-(1./XLOW))/((1./XUP)-(1./XLOW))
        SU=XSQ(L,NT,IA)*(1-XMID)+XSQ(L+1,NT,IA)*XMID
        RETURN
        END
```

```
C*********************************************************************
C       FUNCTION   ST(IDF,NT,IA)
C       ST1.FOR IS A FUNCTION USED TO FIND I-VALUES BY
C       USING ST(IDF,NT,IA).
C
C*********************************************************************
        FUNCTION ST(IDF,NT,IA)
        DIMENSION INDEX(34)
        COMMON /AHEA1/XSQ(26,4,4),XSGP(12,2,4),XST(34,1,4),XSTH(33,3,4)
        IF (IDF,LT,1,OR,IDF,GT,1000000) CALL EXIT
        DATA INDEX/1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,
     1  19,20,21,22,23,24,25,26,27,28,29,30,40,60,120,1000000/
        DO 30 I=1,34
        IF (INDEX(I),LE,IDF) L=1
30      CONTINUE
        XLOW=INDEX(L)
        XUP=INDEX(L+1)
        XMID=((1,/IDF)-(1,/XLOW))/((1,/XUP)-(1,/XLOW))
        ST=XST(L,1,IA)*(1,-XMID)+XST(L+1,1,IA)*XMID
        RETURN
        END
```

```
C**************************************************************
C       FUNCTION   STH(IDF,NT,IA)
C       STH1,FOR IS A FUNCTION USED TO FIND T-VALUES FOR
C       THE TAMHAME PROCEDURE BY USING STH(IDF,NT,IA).
C
C**************************************************************
        FUNCTION STH(IDF,NT,IA)
        DIMENSION INDEX(33)
        COMMON /AREA1/XSQ(26,4,4),XSQP(12,2,4),XST(34,1,4),XSTH(33,3,4)
        IF (IDF.LT.2.OR.IDF.GT.1000000) CALL EXIT
        DATA INDEX/2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19
     1  ,20,21,22,23,24,25,26,27,28,29,30,40,60,120,1000000/
        DO 30 I=1,33
        IF (INDEX(I).LE.IDF) L=I
30      CONTINUE
        XLOW=INDEX(L)
        XUP=INDEX(L+1)
        XMID=((1./IDF)-(1./XLOW))/((1./XUP)-(1./XLOW))
        STH=XSTH(L,NT,IA)*(1-XMID)+XSTH(L+1,NT,IA)*XMID
        RETURN
        END
```