In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

7/25/68

AN INVESTIGATION OF TESTS EMPLOYING SATTERTHWAITE'S SYNTHETIC MEAN SQUARES WHEN CERTAIN ANALYSIS OF VARIANCE ASSUMPTIONS ARE VIOLATED

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

by

John Joseph McKinney

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in the School of Industrial and Systems Engineering

Georgia Institute of Technology

June, 1970

AN INVESTIGATION OF TESTS EMPLOYING
SATTERTHWAITE'S SYNTHETIC MEAN SQUARES
WHEN CERTAIN ANALYSIS OF VARIANCE
ASSUMPTIONS ARE VIOLATED

Approved:

Chairman

Date approved by Chairman: June 1, 1970

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to his thesis advisor, Dr. Douglas C. Montgomery, for his valuable assistance, continued guidance, and encouragement throughout this study. A special word of thanks is due Dr. Harrison M. Wadsworth and Dr. David E. Fyffe for their assistance and service on the Thesis Reading Committee.

Thanks are also extended to Mr. Michael P. Deisenroth for his advice on computer programming, and to Mr. James G. Knight for his assistance in verifying several components used in the computer simulation.

A special debt of gratitude is owed to my wife for her infinite patience and encouragement throughout this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF ILLUSTRATIONS

SUMMARY

When an experimenter is interested in testing for the existence of certain effects (or variance components) in a multiway Analysis of Variance, the situation often arises in which it is impossible to construct an exact test statistic consisting of a ratio of two independent mean squares. In such situations, the analyst usually constructs a linear combination of certain mean squares, chosen so that, under the null hypothesis, the expected value of the statistic is the effect in question. This research investigated the results of departures from the usual analysis of variance assumptions of normality of errors, statistical independence of errors, and homoscedasticity when utilizing Satterthwaite's synthetic mean squares. The analysis was conducted with respect to the size (frequency of incorrect rejection) and power (frequency of correct rejection) of the test. Also, for a given experimental situation, it is often possible to construct two statistics for each main effect which, under the null hypothesis, have expectation equal to the desired effect. As a result, one test statistic may be more appropriate than another when violation of the assumptions occurs.

This research investigated both test statistics and concluded that the size and power of the test employing Satterthwaite's synthetic mean squares are not significantly influenced by departures from the analysis of variance normality assumption where the error variance is of the same magnitude or smaller than the other variance components. Also, highly correlated error terms cause significant departures from the advertised

size and power of the test. Where there is unequal variance in the error term, but the average variance approximates the variance of the other terms, there will be little effect on the size or power of the test. Finally, a two cubed experiment without replication should not be used to investigate the completely random case.

CHAPTER I

INTRODUCTION

This chapter contains a description of the problem being investigated, a review of the literature, and a discussion of the anticipated results.

Nature of the Problem

Industrial engineers are often concerned with the collection and evaluation of experimental data. The analysis of variance is widely used to investigate and analyze measurements on several effects which operate simultaneously.

The analysis of variance is a statistical technique developed by R. A. Fisher (1) over a half century ago to facilitate the analysis and interpretation of data from field tests and laboratory experiments in agricultural and biological research. This technique was found to have wide applicability, and it is now used in all phases of industry and science. Fisher reduced a complex mathematical problem to an almost mechanical procedure. This procedure, because of its simplicity and wide applicability, has become a valuable tool in experimental statistics.

The main purposes of analysis of variance are:

(i) To estimate certain treatment differences that are of interest.

(ii) To obtain some idea of the accuracy of these estimates, e.g., by attaching to them estimated standard errors, fiducial or confidence limits.

(iii) To perform tests of significance. The most common tests are the F-test of the null hypothesis that a treatment difference is zero, or has some predetermined value.

There are three general types of effects recognized in analysis of variance. Treatment effects are the effects of procedures deliberately introduced by the experimenter. Environmental effects are certain features of the environment which the analysis enables us to measure. Experimental error (the third effect), constitutes all elements of variation that are not included in treatment or environmental effects.

In an experiment a collection of one or more of the same type of treatment is called a factor. If there is only one factor and no environmental effects present in an experiment, a one-way classification is said to be used in performing the analysis of variance. If there is only one factor and one type environmental error, a two-way classification would be used. The number associated with the classification used equals the sum of the factors and the type environmental effects present in an experiment.

Another element must be considered in the discussion of analysis of variance. This is the extent of the population about which we are making inferences. If, from an analysis of variance we make inferences about only those levels of each of the factors which are contained in the experiment, we are said to be performing a test on means. This is referred to as the fixed effects model. If we use those levels in the experiment as a sample to make inferences about a greater population, then we are performing a test on variances. This is referred to as the random effects model. Appropriate combinations of fixed and random factors give mixed effects models.

In order for his technique to be applied to an experimental situation, Fisher made the following assumptions about the experiment:

(i) The treatment or factor effects and environmental effects are additive.

(ii) The experimental errors are independent.

(iii) The experimental errors have a common variance.

(iv) The experimental errors are distributed normally.

The industrial experimenter usually wishes to decide which effects or factors are important and to estimate them. When he is interested in testing for the existence of certain effects (or variance components) in a multiway analysis of variance, the situation often arises in which it is impossible to construct an exact test statistic consisting of a ratio of two independent mean squares. In such situations the analyst usually constructs a linear combination of certain mean squares chosen so that, under the null hypothesis, the expected value of the statistic is the effect in question. Satterthwaite (2) has shown that the ratio of linear combinations of mean squares is distributed approximately as an F, and thus the above technique is widely employed in practice. This research investigates the results of departures from the usual analysis of variance assumptions of normality of errors, statistical independence of errors, and homoscedasticity when utilizing Satterthwaite's synthetic mean squares. The analysis is conducted with respect to the size (frequency of incorrect rejection) and power (frequency of correct rejection) of the test. Also, for a given experimental situation, it is often possible to construct two statistics for each main effect which, under the null hypothesis, have expectation equal to the desired effect. As a result one test statistic may be more appropriate than another when violation of the assumptions occurs. Both statistics will be investigated under each set of violations of the assumptions.

Survey of the Literature

Pearson (3), using a simple one-way classification on empirical data from the Bell Telephone Laboratories, New York, concluded that the effect of violation of the normality assumption is slight on inferences drawn about means. However, he also concluded that inferences drawn about variances could be extremely dangerous.

Box (4), satisfied that general non-normality had little effect on test on means, concentrated his effort on determining the effect of non-normality on tests on variances. He stated that the test on variances is particularly sensitive to changes in Kurtosis from the normal theory value of three. Furthermore, that the sensitivity is even greater when the number of variances to be compared exceeds two. He also discovered, using a one-way classification, that when there is a difference in the group size of the non-normal distributions tested, that the effect of departures from normality will be great. This was done using 20 different trials of two independent groups of five and 20 observations per group. Box also showed that the logarithmic transformation technique, developed by Bartlett and Kendell, will in fact bring non-normally distributed data to a form suitable for the application of the analysis of variance. Using a rectangularly distributed population, he noted that the probability of rejecting a false hypothesis are greater when data is transformed than when it is not transformed.

David and Johnson (5), as a special case, considered a one-way classification in which the observations were normally distributed, but the variances differed from group to group, and group sizes were equal. They found that for this case the effects on tests of variances and means are slight.

Grunow (6) and Welch (7), using different investigative techniques, found that a fixed model, one-way classification, having a different error variance for each level will not significantly affect the size and power of the test when the number of observations at each level are the same.

Box (8) applied some theorems on quadratic forms to determine the effect of group to group inequality of variance in a one-way classification. He concluded that if the groups are of equal size, the inequality of variance does not seriously affect the test on means, but with unequal sized groups much larger discrepancies will appear.

Box (9), using the theorems of (8), studied the effects of inequality of variance and first order serial correlation of errors in the two-way classification on the analysis of variance tests on means. He found that when the approximate null hypothesis is true, inequality of variance from column to column results in an increased chance of exceeding the significance point for the test of homogeneity of column means, and decreased chance for the corresponding test on row means. Also, for moderate differences in variance neither effect is large. First order serial correlation within rows was found to produce a large effect on the "between rows" comparisons of means, but little effect on the "between columns" comparisons of means.

Hudson and Krutchkoff (10) investigated the accuracy of Satterthwaite's method for approximating the number of degrees of freedom associated with a linear combination of independent mean squares. Using a particular three way random effects model, the power and size of the test were studied by generating and analyzing data which conformed to all of the assumptions of analysis of variance. Hudson investigated two statistics which can be constructed to test the null hypothesis that the variance of a main effect is equal to zero. He concluded that Satterthwaite's method for approximating the number of degrees of freedom provided a probability of rejection which seemed adequately close to the apparent significance level except for several special cases.

The preceding authors have done considerable investigation of the assumptions when using one-way classification. Scheffe (11) summarizes much of their work. The assumption of equality of variances for the fixed effect model has received the most attention. The assumption of normally distributed errors and additivity of treatment and environmental effects each received the attention of one author for the fixed effect model. The only assumption studied for the random effects model was that of normally distributed errors.

Research in the two-way classification was restricted to the assumptions of independence of experimental errors and equality of variances for the fixed effect model only.

No research was found which studied the assumptions when using the random effects model in the two-way classification. Furthermore, no research was found which studied the assumptions when using the three-way or higher classification for either the fixed or random model.

Anticipated Results

Prior investigations have illustrated that the analysis of variance is a relatively robust statistical technique. We expect that this investigation will confirm that property for the cases studied, except perhaps, in a few special situations.

The anticipated results of this investigation were a clear identification of some limitations of the use of Satterthwaite's synthetic mean squares in the analysis of variance. This would be accomplished by identifying critical assumptions or determining the circumstances for which the assumptions may be relaxed. Also, it was anticipated that some measure of the relative significance of each assumption will be discovered. This would enable the experimenter to extend with reasonable safety the analysis of variance to areas in which an assumption or assumptions may be invalid.

By clearly appreciating the impact of the underlying assumptions of the analysis of variance, the experimenter can perhaps use this methodology to conduct experiments in a less restrictive scientific or industrial environment.

CHAPTER II

THE MODEL AND ASSUMPTIONS TO BE INVESTIGATED

Model to be Investigated

The case to be considered will be a three-factor factorial in a completely randomized design. Each level of any factor is combined with every level of the other two factors. All effects in the model are assumed to be random. Thus the variables in the model are random samples from the population about which inferences are to be drawn.

The model to be investigated is:

$$Y_{ijk} = \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + \epsilon_{ijk} \qquad (2.1)$$

$i = 1,2,\ldots,a; \quad j = 1,2,\ldots,b; \quad k = 1,2,\ldots,c$

where α_i represents the added effect associated with the i^{th} level of factor A.

β_j represents the added effect associated with the j^{th} level of factor B.

γ_k represents the added effect associated with the k^{th} level of factor C.

$(\alpha\beta)_{ij}$ is the added effect associated with the interaction of the i^{th} level of factor A with the j^{th} level of factor B.

$(\alpha\gamma)_{ik}$ is the added effect associated with the interaction of the i^{th} level of factor A with the k^{th} level of factor C.

ϵ_{ijk} is the added effect resulting from pooling the two-way interaction $(\beta\gamma)_{jk}$ and the three-way interaction $(\alpha\beta\gamma)_{ijk}$, and is used as a measure of experimental error.

Every effect in the model is assumed to be a random variable. This is a special case of the more general three-way random effects model.

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} \qquad (2.2)$$

$$+ (\alpha\beta\gamma)_{ijk} + \epsilon_{(ijk)l}$$

in which the overall mean, $\mu = 0$, the interactions $(\beta\gamma)_{jk} = (\alpha\beta\gamma)_{ijk} = 0$ and there is only one observation per cell. The principal objective in eliminating replication in the model was to reduce the amount of computer time required.

If none of the assumptions of analysis of variance are violated, the components of the model are distributed as follows:

(i) $\alpha_i \sim \text{NID}\ (0, \sigma^2{}_A)$ $i = 1,2,\ldots,a$

(ii) $\beta_j \sim \text{NID}\ (0, \sigma^2{}_B)$ $j = 1,2,\ldots,b$

(iii) $\gamma_k \sim \text{NID}\ (0, \sigma^2{}_C)$ $k = 1,2,\ldots,c$

(iv) $(\alpha\beta)_{ij} \sim \text{NID}\ (0, \sigma^2{}_{AB})$ $i = 1,2,\ldots,2;\ j = 1,2,\ldots b$

(v) $(\alpha\gamma)_{ik} \sim \text{NID}\ (0, \sigma^2{}_{AC})$ $i = 1,2,\ldots,a;\ k = 1,2,\ldots,c$

(vi) $\epsilon_{ijk} \sim \text{NID}\ (0, \sigma^2)$ $i = 1,2,\ldots,a;\ j = 1,2,\ldots b;\ k = 1,2,\ldots,c$

where $\alpha_i \sim \text{NID}\ (0, \sigma^2{}_A)$ means that the $\alpha_1, \alpha_2, \ldots, \alpha_a$ are normally and independently distributed with mean zero and common variance $\sigma^2{}_A$.

The table of expected mean squares for this model is given in Table 2.1

Table 2.1 ANOVA Including Expected Mean Squares for the Case Investigated

Source of Variation	Degrees of Freedom	Expected Mean Square
A	$a-1 = v_1$	$\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC} + bc\sigma^2_{A}$
B	$b-1 = v_2$	$\sigma^2 + c\sigma^2_{AB} + ac\sigma^2_{B}$
C	$c-1 = v_3$	$\sigma^2 + b\sigma^2_{AC} + ab\sigma^2_{C}$
AB	$(a-1)\ (b-1) = v_4$	$\sigma^2 + c\sigma^2_{AB}$
AC	$(a-1)\ (c-1) = v_5$	$\sigma^2 + b\sigma^2_{AC}$
E(Error)	$v_1 v_2 v_3 + v_2 v_3$	σ^2
TOTAL	$abc - 1 = v$	

None of the expected mean squares equal $\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC}$; therefore, it is impossible to have an exact test of the hypothesis H_0: $\sigma^2_A = 0$. However, it is possible to form a linear combination of means squares which have the necessary expected value, and use Satterthwaite's approximation. Since there is only one replication and the interactions $(\alpha\beta\gamma)_{ijk}$ and $(\beta\gamma)_{jk}$ are both assumed to equal zero, these interactions are pooled to provide the error term ϵ_{ijk}. Pooling $(\beta\gamma)_{jk}$ with $(\alpha\beta\gamma)_{ijk}$ provides additional degrees of freedom for the error term and does not affect the approximate test of the null hypothesis H_0: $\sigma^2_A = 0$.

The two ratios which will be considered as possible test statistics are:

$$F_P^* = \frac{MS_A + MS_E}{MS_{AB} + MS_{AC}} \qquad (2.3)$$

$$F_A^* = \frac{MS_A}{MS_{AB} + MS_{AC} - MS_E} \tag{2.4}$$

The ratios are denoted F^* to indicate that they follow an approximate rather than an exact F distribution.

The degrees of freedom for the numerator of F_A^* are the degrees of freedom associated with mean square A. The degrees of freedom for the numerator and denominator F_P^* and for the denominator of F_A^* must be calculated using Satterthwaite's approximation. Equations (2.5), (2.6) and (2.7) show how these degrees of freedom are calculated:

$$\text{degrees of freedom numerator } (F_P^*) = \frac{(MS_A + MS_E)^2}{\frac{(MS_A)^2}{v_1} + \frac{(MS_E)^2}{v_1 v_2 v_3 + v_2 v_3}} \tag{2.5}$$

$$\text{degrees of freedom denominator } (F_P^*) = \frac{(MS_{AB} + MS_{AC})^2}{\frac{(MS_{AB})^2}{v_4} + \frac{(MS_{AC})^2}{v_5}} \tag{2.6}$$

$$\text{degrees of freedom denominator } (F_A^*) = \frac{(MS_{AB} + MS_{AC} - MS_E)^2}{\frac{(MS_{AB})^2}{v_4} + \frac{(MS_{AC})^2}{v_5} + \frac{(MS_E)^2}{v_1 v_2 v_3 + v_2 v_3}} \tag{2.7}$$

Departures from Assumptions

Using both statistics F_P^* and F_A^* for the same data, the size and power of the test were calculated when certain assumptions were violated. The results of the analysis of variance under the violated assumptions were obtained for the following conditions when the size of the test is under study:

Condition 1

$$\sigma^2_B = \sigma^2_C = \sigma^2_{AB} = \sigma^2_{AC} = \sigma^2 = 1; \quad \sigma^2_A = 0$$

Condition 2

$$\sigma^2_B = \sigma^2_C = \sigma^2_{AB} = \sigma^2_{AC} = 1; \quad \sigma^2 = 0.01; \quad \sigma^2_A = 0$$

When the power of the test is under study, the above conditions apply except that σ^2_A is assigned a variety of non-zero values.

To investigate the effect of violation of the normality assumption the following three distributions were considered:

(i) Uniform (a special case of beta)

(ii) Triangular (a special case of beta)

(iii) Parabolic

These distributions, illustrated in Figure 2.1, were selected because they represent typical violations of the normal distribution often encountered in practice.

To investigate independence of errors two schemes for correlating errors were used. One provides highly correlated errors, while the other provides moderately correlated errors.

In the investigation of equality of variances three methods of changing the variance of the error term ϵ_{ijk} were used:

(i) σ^2 was given a different value for any change in either i, j, or k.

(ii) σ^2 was given a different value for any change in i only.

(iii) σ^2 was given a new value only one time as i increased from 1 to a.

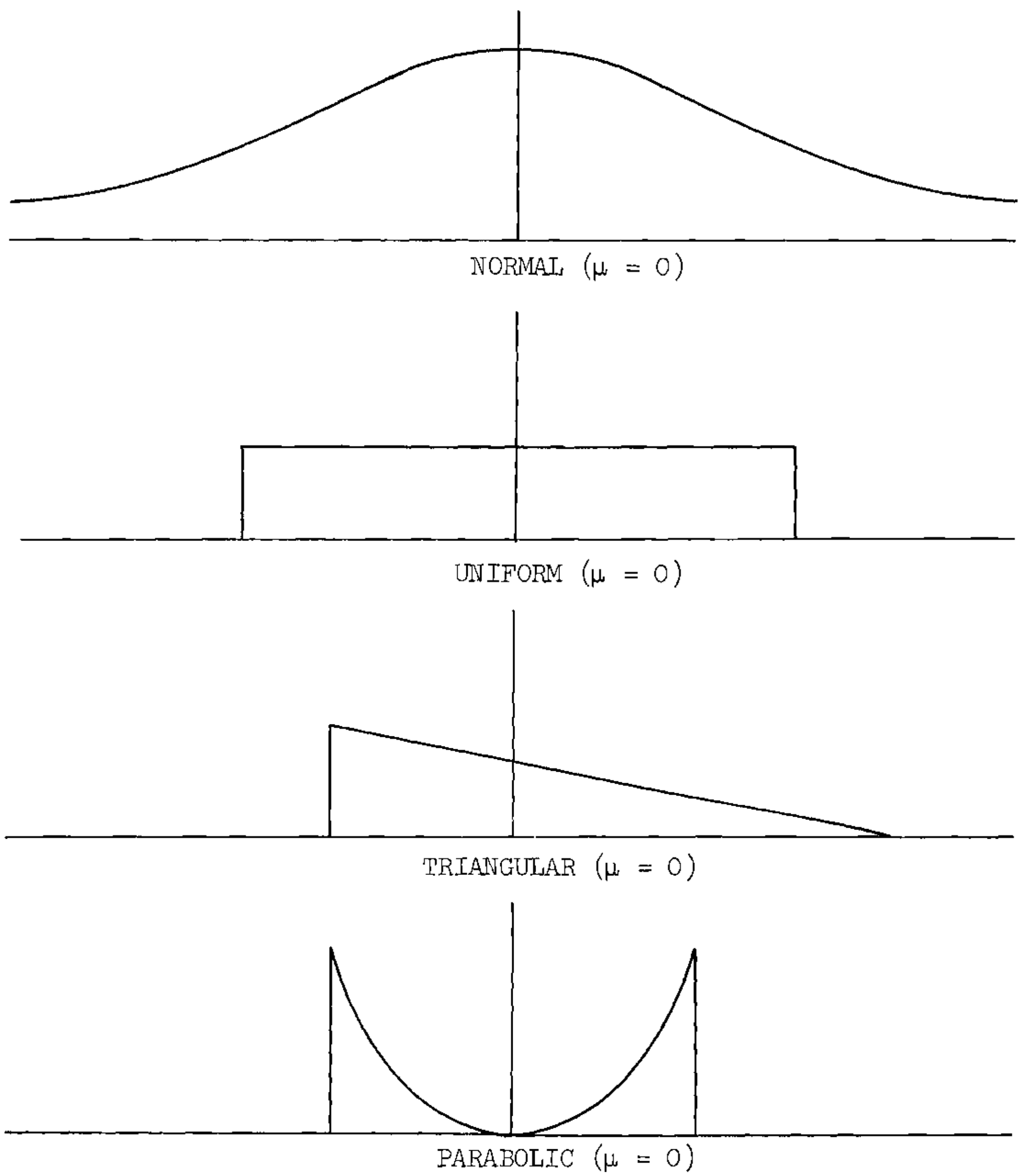

Figure 2.1. Distributions Investigated

CHAPTER III

THE EXPERIMENTAL PROCEDURE

In order to investigate the analysis of variance, it is necessary to obtain data similar to that which would be available in an actual experiment. A Monte Carlo simulation is used to provide this data. For convenience in the sequel, data which conforms to the assumptions of analysis of variance shall be referred to as conforming data, and data which violates the assumptions will be referred to as non-conforming data.

To generate non-conforming data it is necessary to modify the error term ϵ_{ijk} in an appropriate fashion. Since the procedures used to generate non-conforming data differ from those used to generate conforming data only in the error term, and since the procedures used to determine the size and power of the test are used throughout this investigation, generation of conforming data and the determination of size and power will be discussed first. This will be followed by a discussion of the generation of the appropriate error term.

Generation and Testing of Conforming Data

Generation of the Data

To generate conforming data each of the right hand components (except α_i and ϵ_{ijk}) of the model

$$Y_{ijk} = \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{jk} + \epsilon_{ijk}$$

is assumed to be distributed normally with mean equal zero and variance one. ϵ_{ijk} will have variances of 0.01 or 1.0. In order to examine the size of the test it is necessary that the null hypothesis $\sigma^2_A = 0$ be true; therefore, $\alpha_1, \alpha_2, \ldots, \alpha_a$ are set equal to zero.

In generating these random variables use was made of the library subroutine designated RANDU in the Fortran computer language, which generates random numbers which are uniformly distributed between zero and one. In order to obtain the normally distributed random variables it is necessary to use a normal process generator which converts the uniformly distributed random variables provided by RANDU to normally distributed random variables. The normal process generator used was developed by Schmidt and Taylor (12) and is listed below:

$$x = \frac{(r - 0.5)(\sigma)(f(v))}{|r - 0.5|} + \mu \qquad (3.1)$$

$$f(v) = v - \frac{(2.515517 + 0.802853v + 0.010328v^2)}{(1 + 1.432788v + 0.189269v^2 + 0.001308v^3)} \qquad (3.2)$$

$$v = \sqrt{-2 \log_e 0.5(1-|1-2r|)} \qquad (3.3)$$

where x is the normally distributed random variable with variance equal to σ^2, r is the uniformly distributed random variable obtained from RANDU, where $0 < r < 1$ and μ is the mean. The above equation, when properly translated into Fortran computer language, provides normally distributed random variables with whatever mean and variance are desired.

If, for example, it is desired to generate eight values of ϵ_{ijk} from a normal distribution with a mean zero and a variance equal to .01. The following seven Fortran statements will generate eight such values and store them as E_{111}, E_{112}, E_{121}, E_{122}, E_{211}, E_{212}, E_{221}, and E_{222}.

```
(i)   DO 75 I = 1,2
(ii)  DO 75 J = 1,2
(iii) DO 75 K = 1,2
(iv)  R = RDM(MACK)
(v)   V = SQRT(-2.0*ALOG(0.5*(1.0-ABS(-1.0-2.0*R))))
(vi)  FOFV = V-(2.515517 + 0.80285*V + 0.01038*V**2)/(1.0 +
      1.432788*V + 0.189269*V**2 + 0.001308*V**3)
(vii) 75 E(1,J,K) = ((R-0.5)*0.1*FOFV)/ABS(R-0.5)
```

RDM(MACK) is a Fortran subroutine which secures from RANDU an array of uniform random numbers and selects a new random number from this array each time the statement is encountered. (See Appendix C for the Fortran listing.)

Statements similar to those above were used to generate the other components of the model, and the sum of these components provided one observation, denoted Y_{ijk}. This process was repeated N times (where N = a·b·c) until all the Y_{ijk} necessary for the desired analysis of variance were generated. Another Fortran subroutine (See Appendix C) was used to conduct an analysis of variance on these generated Y_{ijk}. Both test statistics were compared with the critical values of the F distribution at $\alpha = .05$.

Size of Test

The process of generating observations, Y_{ijk} analyzing, and accepting or rejecting the null hypothesis was done 2000 times using 20 blocks of 100 iterations each. After each 100 iterations, the number of rejections of the null hypothesis, the average value of the mean squares, degrees of freedom for the numerator and denominator, and both test statistics were printed out. After 2000 iterations, the average number of rejections, mean squares and both test statistics were computed.

By conducting the experiment in blocks it was possible to calculate the standard error of the mean calculated from each set of twenty blocks. (See Appendix D for typical standard errors). The appropriate equation is

$$SE = [\sum_{i=1}^{20} X^2_i - (\sum_{i=1}^{20} X_i)^2/20]/[(19)(20)] \qquad (3.4)$$

where SE represents the standard error of the mean, and X_i represents the number of rejections of the i^{th} block of 100 iterations. Similar computations were made for degrees of freedom and mean squares.

Power of the Test

The investigation of power was handled in a similar fashion. An example of the procedure developed by Hudson and Krutchkoff (10) to determine the power of the test is given below.

2000 iterations of the analysis were performed in 20 blocks of 100 with the number of rejections and other pertinent data printed out after each 100 iterations. Since the power of a test is the probability of rejecting the null hypothesis when it is, in fact, false, values of $\alpha_1, \alpha_2, \ldots, \alpha_a$ were generated so as to reflect various values of σ^2_A.

The degree to which the variance of A differed from zero was measured by the parameter ϕ. This parameter was defined so as to indicate the degree to which the ratio of the expectation of the numerator of the test statistic to the expectation of the denominator of the test statistic differs from unity.

Thus

$$\frac{\text{Expectation of Numerator of Test Statistic}}{\text{Expectation of Denominator of Test Statistic}} = 1 + \phi. \qquad (3.5)$$

Consider the special case of the previously mentioned model with

$a = b = c = 3,$

$$Y_{ijk} = \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + \epsilon_{ijk}$$

$$i = 1, 2, 3; \; j = 1, 2, 3; \; k = 1, 2, 3.$$

The assumptions on this model are, as before,

$\alpha_i \sim \text{NID } (0, \sigma^2_A) \; i = 1, 2, 3$

$\beta_j \sim \text{NID } (0, \sigma^2_B) \; j = 1, 2, 3$

$\gamma_k \sim \text{NID } (0, \sigma^2_C) \; k = 1, 2, 3$

$(\alpha\beta)_{ij} \sim \text{NID } (0, \sigma^2_{AB}) \; i, j = 1, 2, 3$

$(\alpha\gamma)_{ik} \sim \text{NID } (0, \sigma^2_{AC}) \; i, k = 1, 2, 3$

$\epsilon_{ijk} \sim \text{NID } (0, \sigma^2) \; i, j, k = 1, 2, 3$

Table 3.1 Expected Mean Squares for Special Case of Three-Way Random Model ($\sigma^2_{BC} = 0$ $\sigma^2_{ABC} = 0$ and $a = b = c = 3$)

Source	Expected Mean Square
A	$\sigma^2 + 3\sigma^2_{AB} + 3\sigma^2_{AC} + 9\sigma^2_A$
B	$\sigma^2 + 3\sigma^2_{AB} + 9\sigma^2_B$
C	$\sigma^2 + 3\sigma^2_{AC} + 9\sigma^2_C$
AB	$\sigma^2 + 3\sigma^2_{AB}$
AC	$\sigma^2 + 3\sigma^2_{AC}$
E(Error)	σ^2
TOTAL	

Using the test statistic $F_P^* = \frac{MS_A + MS_E}{MS_{AB} + MS_{AC}}$, the ratio

$[E(MS_A + MS_E)]/[E(MS_{AB} + MS_{AC})]$ is

$$\frac{2\sigma^2 + 3\sigma^2_{AB} + 3\sigma^2_{AC} + 9\sigma^2_A}{2\sigma^2 + 3\sigma^2_{AB} + 3\sigma^2_{AC}} = 1 + \frac{9\sigma^2_A}{2\sigma^2 + 3\sigma^2_{AB} + 3\sigma^2_{AC}}$$

$$= 1 + \phi.$$

For the case where $\sigma^2 = \sigma^2_{AB} = \sigma^2_{AC} = 1$, the value of σ^2_A which would result in a ϕ of (say) 3, is found by solving

$$\frac{9\sigma^2_A}{2(1) + 3(1) + 3(1)} = 3.$$

$$\therefore \sigma^2_A = 24/9$$

Therefore, to check the power for this case, namely a = b = c = 3 with $\phi = 3$, $\alpha_1, \alpha_2, \ldots, \alpha_a$ were generated from a normal distribution with mean zero and variance 24/9. This was accomplished by the following two Fortran statements

```
(i)    DO 7 I = 1, 3
(ii)  7 A(I) = (SQRT(24./9.)) *RDM (MACK)
```

which generated α_1, α_2, and α_3 and stored them as A_1, A_2, and A_3.

The block diagram of Figure 3.1 gives the main steps of the basic program. The block diagram 3.2 shows the steps required to obtain the necessary observations. In this diagram a, b, and c denote the number of levels of factors A, B, and C respectively.

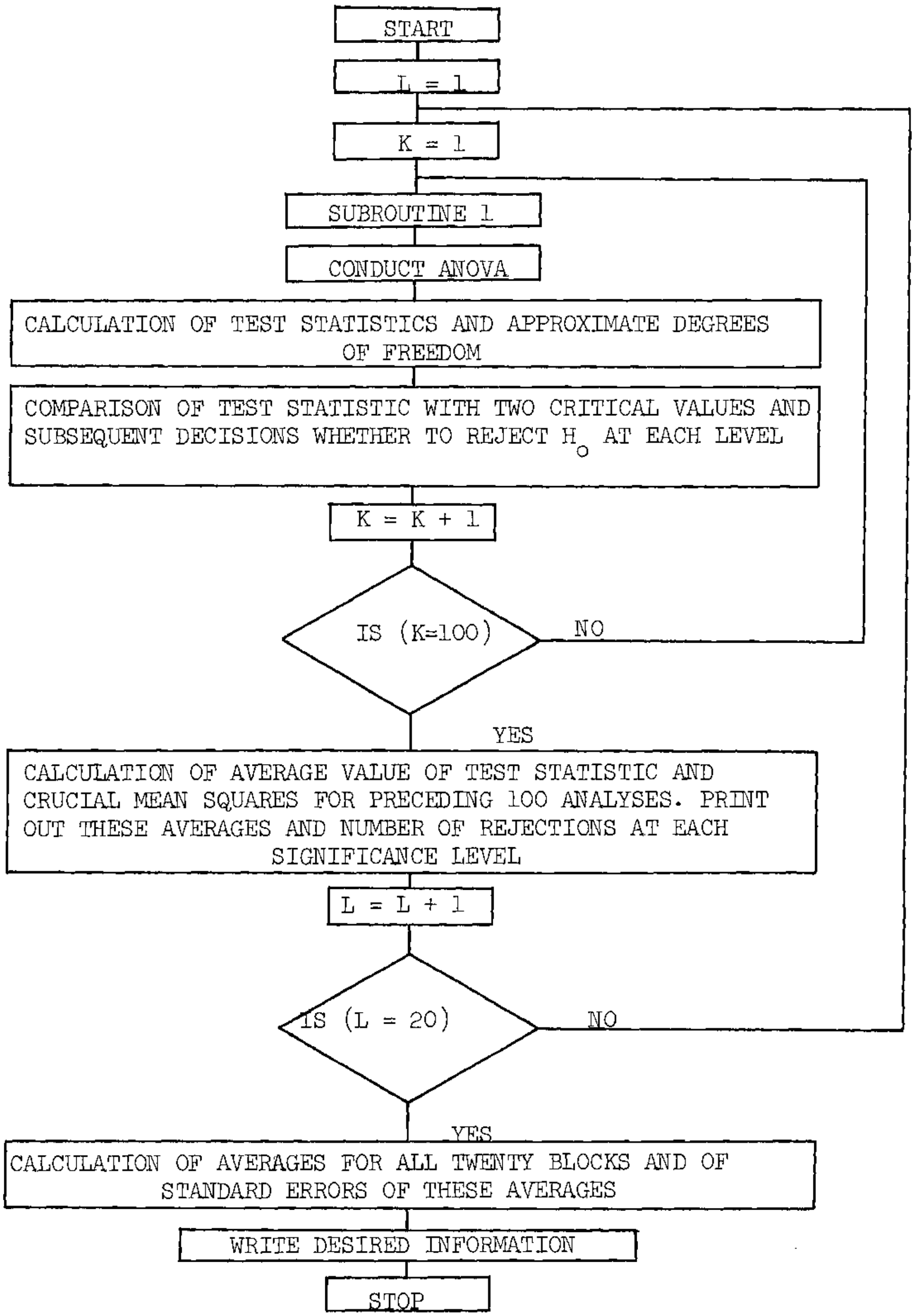

Figure 3.1. Block Diagram for Basic Program

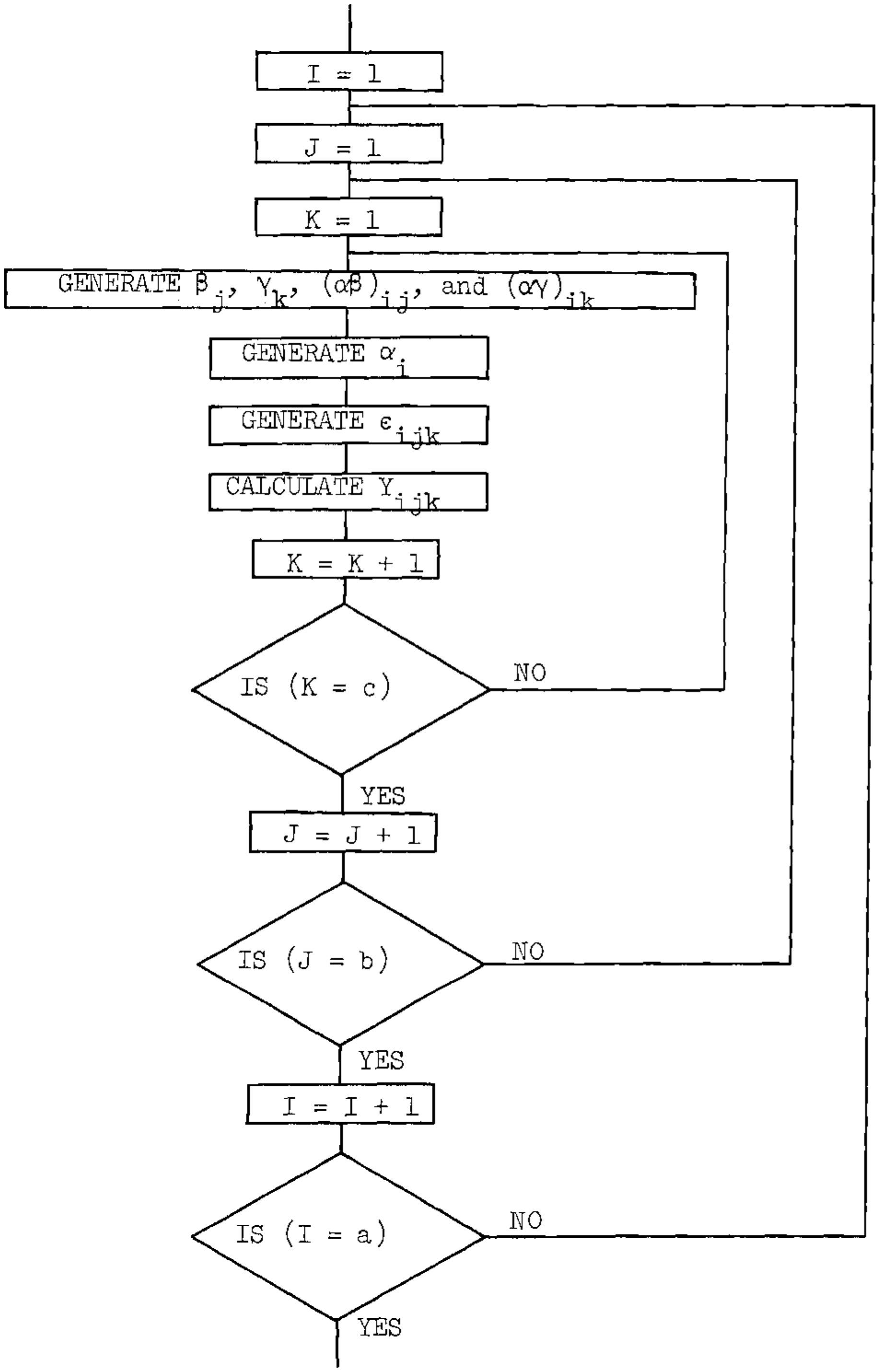

Figure 3.2. Block Diagram for Subroutine 1

Generation of Non-Normal Data

To generate the three non-normal distributions investigated, it was necessary to develop three different random process generators. The procedure used to develop these generators is identical. So that a realistic comparison could be made between the data with error terms drawn from the non-normal distributions and those with error terms drawn from the normal distribution, the non-normal random process generators were developed to provide random variates from the desired distribution with a mean equal to zero and the variance equal to 1 or 0.01.

An example of the development of the random process generator which provides a parabolic distribution with a mean equal to zero and variance equal to one will be given. It was first necessary to find a probability density function. The equation of a parabola centered at zero is

$$f(x) = Kx^2 \tag{3.6}$$

The value of the constant K is unknown, but in order for f to be a probability density function the following relation must be true:

$$\int_{-\alpha}^{\alpha} f(x)dx = 1, \tag{3.7}$$

therefore

$$\int_{-\alpha}^{\alpha} Kx^2 dx = 1$$

$$Kx^3/3 \Big|_{-\alpha}^{\alpha} = 1$$

$$K2\alpha^3/3 = 1,$$

hence

$$K = 3/2\alpha^3 \tag{3.8}$$

and

$$f(x) = 3x^2/2\alpha^3 \tag{3.9}$$

Since the parabola is centered at zero the mean of the distribution already equals zero. It is now necessary to determine the values of α which will provide variances equal to 1.0 or 0.01. We present the case where it is required that $\sigma^2 = 1.0$.

Since

$$\mathrm{VAR}\,[X] = E[X^2] - (E[X])^2 \tag{3.10}$$

where

$$E[X^2] = \int_{-\alpha}^{\alpha} X^2 f(X)dx \tag{3.11}$$

$$E[X] = 0 \tag{3.12}$$

and in this case it is desired that VAR [X] = 1, the following equation can be used to determine the appropriate value of α:

$$\mathrm{VAR}\,[X] = \int_{-\alpha}^{\alpha} X^2 f(X)dx = 1, \tag{3.13}$$

or

$$\mathrm{VAR}\,[X] = \int_{-\alpha}^{\alpha} X^2 (3X^2/2\alpha^3)dx = 1.$$

After integrating we are left with

$$\mathrm{VAR}\,[X] = 3\,\alpha^2/5 = 1. \tag{3.14}$$

Therefore, the value of α that provides a variance of one is

$$\alpha = \sqrt{5/3}. \tag{3.15}$$

In order to verify that the density function of equation (3.9) meets the necessary prerequisites, the cumulative distribution function was developed.

Since
$$F_X(x) = P(X \leq x) \tag{3.16}$$

$$= K \int_{-\alpha}^{x} t^2 dt$$

$$= \frac{K}{3} t^3 \Big|_{-\alpha}^{x}$$

Substituting K from equation (3.8)

$$F_X(x) = (x^3 + \alpha^3)/2\alpha^3, \tag{3.17}$$

Now let $x = -\alpha$

$$F_X(-\alpha) = ((-\alpha)^3 + \alpha^3)/2\alpha^3 = 0,$$

letting $x = 0$

$$F_X(0) = (0 + \alpha^3)/2\alpha^3 = .5,$$

and letting $x = \alpha$

$$F_X(\alpha) = ((\alpha)^3 + \alpha^3)/2\alpha^3 = 1.$$

Once the density function was obtained, the probability integral theorem provided the basis for the development of the random process generator. This theorem may be stated informally as: "Let f be the density function of a continuous random variable x. Then

$$F(t) = \int_{-\infty}^{t} f(x)dx \tag{3.18}$$

is distributed uniformly over the interval (0.1)." Since RDM(MACK) using RANDU could provide a source of uniformly distributed random variables in the interval (0,1), the following substitution into equation (3.18) can be made

$$\text{RDM(MACK)} = \int_{-\infty}^{t} f(x)dx$$

Furthermore, since the density function for the parabolic distribution is known, and since $F_X(x) = 0$ for values of x less than or equal to $-\alpha$ equation (3.18) can then be written as:

$$\text{RDM(MACK)} = \int_{-\alpha}^{x} f(x)dx = \int_{-\alpha}^{x} 3t^2/2\alpha^3 \, dt.$$

By carrying out the integration we obtain

$$\text{RDM(MACK)} = (x^3 + \alpha^3)/2\alpha^3 \tag{3.19}$$

To obtain the random variables from a population with a parabolic distribution solve equation (3.19) for x in terms of α and RDM(MACK) and obtain

$$x = \sqrt[3]{2\alpha^3 \ \text{RDM(MACK)} - \alpha^3} \tag{3.20}$$

To generate eight random variables from a parabolic distribution with a mean equal to zero and a variance of one would require the following four Fortran statements (where $Q = \alpha$):

(i) DO 75 I = 1,2

(ii) DO 75 J = 1,2

(iii) DO 75 K = 1,2

(iv) 75 E_{ijk} = CBRT(2*(Q**3*RDM(MACK))-Q**3

The above statements would, as in the case of the normally distributed errors, generate and store E_{111}, E_{112}, E_{121}, E_{211}, E_{212}, E_{221}, E_{222}, but in this case E_{ijk} is taken from a parabolic distribution.

Since the procedures used to develop the Triangular and uniform processes generators were identical to those above, and quite well-known, the development of these two generators is not included. However, the density functions are shown in equations (3.21) and (3.22) below:

Triangular:

$$f(x) = 2[1 - (x-a)/(b-a)]/(b-a), \quad a \leq x \leq b \tag{3.21}$$

Uniform:

$$f(x) = 1/2\alpha, \quad -\alpha \leq x \leq \alpha. \tag{3.22}$$

Generation of Correlated Data

To generate correlated data, use is made of another normal process generator described by Schmidt and Taylor (12). The equation for this generator is:

$$x = \sigma(\sum_{i=1}^{N} r_i - N/2)\sqrt{N/12} + \mu \tag{3.23}$$

where: x is the random variable drawn from a normal distribution

σ^2 is the desired variance.

μ is the desired mean.

r_i is a random variable drawn from a population uniformly distributed between zero and one. r_i will never equal zero or one.

N is the number of times a sample is drawn from the uniform distribution.

If in equation (3.23) μ is set equal to zero and N is set equal to twelve, equation (3.23) reduces to:

$$x = \sigma(\sum_{i=1}^{12} r_i - 6) \tag{3.24}$$

where x represents one observation on the error term, E_{ijk}.

Highly correlated data can be generated by changing only one r_i each time a new observation is generated. For example, consider the sequence

$$\epsilon_{111} = \sigma(\sum_{i=1}^{12} r_i - 6),$$

$$\epsilon_{112} = \sigma(\sum_{i=2}^{13} r_i - 6),$$

$$\epsilon_{113} = \sigma(\sum_{i=3}^{14} r_i - 6), \text{ etc.}$$

Thus each new observation is composed of eleven of the twelve random variables used to generate the previous observation. The Fortran statements necessary to generate highly correlated observations drawn from a normal distribution with a mean of zero and a variance equal to one are:

```
(i)    DO 77 KK = 1, 12
(ii)   77 R(KK) = RDM(MACK)
(iii)  DO 75 I = 1,A
(iv)   DO 75 J = 1,B
(v)    DO 75 K = 1,C
(vi)   E(I,J,K) = 0.0
(vii)  DO 78 KK = 1,12
(viii) 78 E(I,J,K) = E(I,J,K) + R(KK)
(ix)   R(13) = RDM(MACK)
(x)    DO 75 KK = 1,12
```

(xi) 79 R(KK) = R(KK) + 1

(xii) 75 CONTINUE

To generate moderately correlated data the previous technique is used except that six of the variables drawn from the uniform population are changed each time a new normally distributed random variable is required. For example, consider the sequence

$$\epsilon_{111} = \sigma(\sum_{i=1}^{12} r_i - 6),$$

$$\epsilon_{112} = \sigma(\sum_{i=7}^{18} r_i - 6),$$

$$\epsilon_{113} = \sigma(\sum_{i=13}^{24} r_i - 6), \text{ etc.}$$

The Fortran statement necessary to generate N observations drawn from a normal distribution with a mean of zero and a variance equal to 0.01 are

(i) DO 77 KK = 1,12

(ii) 77 R(KK) = RDM(MACK)

(iii) DO 75 I = 1,A

(iv) DO 75 J = 1,B

(v) DO 75 K = 1,C

(vi) E(I,J,K) = 0.0

(vii) DO 78 KK = 1,12

(viii) 78 E(I,J,K) = E(I,J,K) + R(KK) - 0.5

(ix) DO 93 KK = 13, 18

(x) 93 R(KK) = RDM(MACK)

(xi) DO 79 KK = 1,12

(xii) 79 R(KK) = R(KK + 6)

(xiii) E(I,J,K) = E(I,J,K)*0.1

(xiv) 75 CONTINUE

Generation of Data with Unequal Error Variance

The investigation of data with unequal variances used three different schemes to generate the error term, ϵ_{ijk}. The first scheme provides a new value for the variance of the error term for each observation of Y_{ijk}. This was accomplished by selecting an initial low value for σ^2 at ϵ_{111}. Then adding a $\Delta\sigma^2$ for each new ϵ_{ijk} encountered until the maximum σ^2 is obtained at $\epsilon_{a,b,c,}$ where a, b and c represent the number of levels of factors A, B, and C. To accomplish this we use the same normal process generator developed by Schmidt and Taylor, except that the value of σ must be changed with each observation. Let LVAR be the value of σ^2 at ϵ_{111}, and UVAR be the value of σ^2 at $\epsilon_{a,b,c}$. To insure an equal incrementation of σ^2 for each new observation we let

$$\Delta\sigma^2 = \frac{\text{UVAR - LVAR}}{\text{N - 1}}, \tag{3.25}$$

where $N = a \cdot b \cdot c$ and $\sigma^2 = SE$. Therefore, to generate twenty-seven observations of ϵ_{ijk} with $\sigma^2 = 1$ at ϵ_{111} and $\sigma^2 = 7.5$ at ϵ_{333}, the following Fortran statements are used:

(i) A = 3

(ii) B = 3

(iii) C = 3

(iv) N = A*B*C

(v) LVAR = 1

(vi) UVAR = 7.5

(vii) DIF = UVAR - LVAR

```
(viii)  COUNTER = 0.0
(ix)    DO 75 I = 1,A
(x)     DO 75 J = 1,B
(xi)    DO 75 K = 1,C
(xii)   SE = SQRT((LVAR + DIF*COUNTER)/N-1)
(xiii)  V = SQRT(-2.0*ALOG(0.5*(1.0-ABS(-1.0-2.0*R))))
(xiv)   FOFV = V-(2.551527 + 0.80285*V + ).01038*V**2)/1.0 +
        1.432788*V + 0.189269*V**2+0.0013808*V**3)
(xv)    COUNTER = COUNTER + 1
(xvi)   75 E(I,J,K) = ((R-0.5*SE*FOFV)/ABS(R-0.5)
```

The average σ^2 for the N terms generated is, of course, LVAR + DIF/2. For identification purposes, the data generated using the preceding procedure will be referred to as class 1.

The second scheme provides a different value for the variance of the error term for each level of factor A. As before, an initial low value of σ^2 was selected for ϵ_{ijk}. Data generated under this scheme will be called class 2. The same Fortran statements to generate class 1 data were used with the following exceptions:

(i) $\sigma = \sqrt{(\text{LVAR} + (\text{DIF})\ (\text{COUNTER}))/(\text{A}-1)}$ (3.26)

(ii) The relative location of the σ and counter statements were changed as follows:

```
(a) DO 75 I = 1,A
(b) SE = SQRT((LVAR + DIF*COUNTER)/(A-1)
(c) COUNTER = COUNTER + 1
(d) DO 75 j = 1,B
```

The third scheme changed the value of σ^2 only one time during the generation of the N terms ϵ_{ijk}. This was done usually for the highest level of factor A. For example, if a = b = c = 3, the low value for σ^2 would be assigned ϵ_{111}, ϵ_{121}, ..., ϵ_{222} and the high value for σ^2 would be assigned to ϵ_{311}, ϵ_{312}, ..., ϵ_{333}. The only change to the Schmidt and Taylor normal process generator was the insertion of a Fortran "IF" statement which would cause the proper value of σ to be selected. Data generated in this manner will be called class 3.

CHAPTER IV

RESULTS OF THE EXPERIMENT

The calculated values of the size of the test, the power of the test, selected distributions of the degrees of freedom and average values of the degrees of freedom numerator and denominator are listed in appendices A, B, and E. A study of the tables in those appendices provided the information below.

The Size of the Test

For the three non-normal distributions investigated, it was found that the size of the test is not influenced by the distribution of the error term ϵ_{ijk}. For a set of levels (a, b, c), a particular F^*, and σ^2, the size of the test for error terms drawn from normal, uniform, parabolic, and triangular distributions is approximately equal.

The investigation of correlated data showed the following:

(i) If the variance of the error term is of the same magnitude as the variance of the other components of Y_{ijk}(i.e., where $\sigma^2 = 1.0$) there is little difference between the size of the test of uncorrelated data and moderately correlated data; however, the size of the test is significantly higher than both of these when high correlation of the error term exists.

(ii) Where the variance of the error term is much smaller than the variance of the other terms of Y_{ijk}, the size of the test for uncorrelated, moderately correlated, and highly correlated data is approximately equal. Furthermore, little difference can be noted using F_P^* or F_A^*.

The investigation of unequal variance showed that for class 1 data the size of the test differs little from the size of test with equal error variance. The investigation of class 2 data showed no conclusive results on the size of the test when the average variance of the error term was greater than the variance of the other components of the model. The size of the test followed no discernable pattern with regard to F^*, sets of levels, or range of error variance. However, when the average value of the error variance was less than the variance of the other components of the model, the size of the test using F_A^* was greater.

For class 3 data it is noted that when the statistic F_P^* is used, the size of the test increases for an increase in the average value of the variance of the error term. Also, if the average value of the variance of the conforming error term, the size of test is essentially the same for a given statistic F^* and set of levels (a,b,c).

The number and arrangement of levels also influences the size of the test. It can be noted from all the tables in Appendix A, that for both conforming and non-conforming data, the size of the test using the statistic F_A^* will be greater than when using F_P^* except when sets of levels (3,3,3), (2,4,4), and (5,2,2) occur with the variance of the error term equal to or greater than the variance of the other terms of the model.

The Power of the Test

The investigation of the power of the test consistently showed that as ϕ increases, the power of the test increases. This was true for both conforming data and all types of non-conforming data. For any given set of levels (a,b,c), statistics F^*, and σ^2, the power increases as N increases, where $N = a \cdot b \cdot c$. It should be noted that the power of the test was investigated only for the balanced sets of levels (2,2,2), (3,3,3), (4,4,4), and (5,5,5). The violations of the assumptions of normality, equality of errors or independence show that several factors definitely interact to influence the power of the test. A particular F^*, set of levels, error variance, and type of violation will produce a certain power of the test. A change in any one of these factors while holding the others constant will result in a change in the power of the test.

The investigation of the non-conforming distributions showed the following:

(i) For a particular ϕ, σ^2, and F^* the type distribution has little effect on the power of the test.

(ii) For sets of levels (4,4,4) and (5,5,5) with σ^2 either 1.0 or 0.01, the power of the test is greater when F_A^* is used.

(iii) For sets of levels (3,3,3) and $\sigma^2 = 1.0$ the power of the test is greater when F_P^* is used.

The investigation of data with correlated error variances yielded the following:

(i) The power of the test is greater for highly correlated data than for moderately or uncorrelated data, except where $\phi = 10$, $\sigma^2 = 1.0$, and the levels are (4,4,4), or (5,5,5). Under these circumstances the powers of the test on highly correlated, moderately correlated and uncorrelated data are not significantly different.

(ii) For a particular σ^2, ϕ, set of levels and type of correlation, the power of the test using F_P^* is approximately equal to that using F_A^*.

The investigation of unequal error variances showed:

(i) For a particular ϕ, range of error variance, and set of levels, the power of the test is greater for F_A^* than F_P^*.

(ii) For class 3 data the power of the test is less for a particular set of levels, ϕ, and F^* than it is for conforming data.

(iii) Also, for class 3 data with all other parameters equal the power of the test decreases as the average value of the error variance increases.

Degrees of Freedom

The average degrees of freedom for the numerator and denominator were found to be influenced more by the number and arrangement of the levels (a,b,c) than by any type of non-conforming data. With certain exceptions, caused by the arrangement of levels, which are discussed in Appendix E, both the degrees of freedom for the numerator and the denominator increase as the total number of observations increase. The degrees of freedom for the numerator for F_A^* will change only with changes in the number of levels of factor A. Also, with two exceptions, discussed in Appendix E, the degrees of freedom for the numerator are less than the degrees of freedom for the denominator for a particular σ^2, F^*, and set of levels (a,b,c).

It was also noted that when the average value of a class 3 error term was increased, there was an increase in the degrees of freedom in both numerator and denominator.

Other Observations

At a = b = c = 2 the average value of the alternate test statistic F_A^* was found to be negative in several cases. The negative F_A^* occurred when the variance of the error term was greater than or equal to the variance of the other components of Y_{ijk}. Specifically, it occurred with conforming data, the three types of non-normally distributed data, and class 3 data. One case of moderately correlated data having a high error variance also produced many negative F_A^*. Where the error variance approximated the variance of the other components of the model, less than five percent of the F_A^* were negative for either type of correlation. With levels a = b = c = 3 approximately 15 percent of the F_A^* were negative for tests on conforming data, the three types of non-normally distributed data, class 3 data, and the one case of moderately correlated data with a high error variance. For all other cases investigated at a = b = c = 3 less than five percent of the F_A^* were negative. At a = b = c = 4 and any other combination of levels where the total N was greater than or equal to 64, no cases of a negative F_A^* were noted.

It should also be noted that at a = b = c = 2 the probability of making a type I error is zero for almost all cases investigated, but this apparent advantage is offset by having a probability of committing a type II error of greater than 0.88 for all cases investigated. When the F_P^* statistic is used, this probability is equal to 1.0.

Another factor worthy of consideration is that for class 1, class 2, and class 3 data, the mean squares for B and C tended to differ from their expected value to a greater extent than did these same mean squares for all other types of data. This occurred in cases where the average value of the error variance was greater than the variance of the other components of Y_{ijk}.

Discussion

The increase in power for increases in N and the parameter ϕ is not surprising. As N increases, the number of degrees of freedom calculated for the numerator and denominator increase. As both these degrees of freedom increase, the corresponding critical value of F will decrease, thereby increasing the probability that F^* is greater than $F\alpha$ and enabling the null hypothesis to be rejected. Furthermore, since increases in ϕ cause $\sigma^2{}_A$ to increase, the probability of correctly rejecting the null hypothesis will increase.

Two factors influence the fact that F_A^* provides greater power for the sets of levels (4,4,4) and (5,5,5). The first is that the expected value of F_A^* is greater than the expected value of F_P^* (see equations 4.1 and 4.2). The second is that the degrees of freedom calculated using equations 2.5, 2.6, and 2.7 select critical F's corresponding to F_A^* and F_P^* which are not significantly different. Table 4.1 shows the average degrees of freedom, F^*, and corresponding critical F for the levels (4,4,4) and (5,5,5).

Table 4.1. Calculated Degrees of Freedom with Corresponding Critical F.

σ^2	LEVELS	F^*	DF NUMERATOR	DF DENOMINATOR	CRITIAL F
1.0	4,4,4	F_A^*	3	12	3.49
1.0	4,4,4	F_P^*	5	15	2.90
.01	4,4,4	F_A^*	3	15	3.29
.01	4,4,4	F_P^*	3	15	3.29
1.0	5,5,5	F_P^*	4	25	2.76
1.0	5,5,5	F_P^*	5	30	2.53
0.01	5,5,5	F_A^*	4	30	2.69
0.01	5,5,5	F_P^*	4	30	2.69

By substituting the appropriate expected value into equations 2.3 and 2.4, we get equations 4.1 and 4.2 below

$$E[F_P^*] = \frac{2\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC} + bc\sigma^2_A}{2\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC}} = 1 + \frac{bc\sigma^2_A}{2\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC}} \qquad (4.1)$$

$$E[F_A^*] = \frac{\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC} + bc\sigma^2_A}{\sigma^2 + c\sigma^2_{AB} + b\sigma^2_A} = 1 + \frac{bc\sigma^2_A}{\sigma^2 + c\sigma^2_{AB} + b\sigma^2_{AC}} \qquad (4.2)$$

By inserting the values of a, b, c, σ^2_{AB} and σ^2_{AC} for levels (4,4,4) and (5,5,5), equations 4.1 and 4.2 may be reduced to equations 4.3, 4.4, 4.5, and 4.6 below. Where $a = b = c = 4$

$$E[F_A^*] + 1 + \frac{16\sigma^2{}_A}{\sigma^2 + 8} \tag{4.3}$$

$$E[F_P^*] = 1 + \frac{16\sigma^2{}_A}{2\sigma^2 + 8} \tag{4.4}$$

where a = b = c = 5

$$E[F_A^*] = 1 + \frac{25\sigma^2{}_A}{\sigma^2 + 10} \tag{4.5}$$

$$E[F_P^*] = 1 = \frac{25\sigma^2{}_A}{2\sigma^2 + 10} \tag{4.6}$$

From the above it is obvious that $E[F_A^*]$ will be greater than $E[F_P^*]$. Since the critical F's corresponding to F_A^* and F_P^* are not significantly different and since the $E[F_A^*]$ is greater than $E[F_P^*]$, F_A^* has a higher frequency of exceeding critical F than F_P^*. Therefore, when F_A^* is used, the false hypothesis that $\sigma^2{}_A = 0$ was rejected more frequently.

It was noted that at levels (3,3,3) and $\sigma^2 = 1.0$ the power of the test with F_P^* was higher than it was with F_A^*. This results from the fact that $E[F_P^*]$ and $E[F_A^*]$ are not significantly different for a particular ϕ, and that the corresponding critical F's are of different magnitude. The degrees of freedom were determined by equations 2.5 and 2.6, and those associated with F_A^* are (2,4) and with F_P^* are (4,6). These provide the following critical F values: $F_{.05,2,4} = 6.94$ and $F_{.05,4,6} = 4.53$, using equations 4.1 and 4.2 and substituting the values of a,b,c, $\sigma^2{}_{AB}$, $\sigma^2{}_{AC}$ and σ^2 equations 4.7 and 4.8 result..

$$E[F^*_A] = 1 + \frac{9\sigma^2_A}{7} \qquad (4.7)$$

$$E[F^*_P] = 1 + \frac{9\sigma^2_A}{8} \qquad (4.8)$$

For $\phi = (1,2,3,10)$, $\sigma^2_A = (8/9,\ 16/9,\ 24/9,\ 80/9)$ and the resulting expected values of F^*_A and F^*_P are shown in Table 4.2.

Table 4.2. Expected Values of F^* ($a = b = c = 3$ and $\sigma^2 = 1.0$)

	ϕ				Critical
	1	2	3	10	F
$E[F^*_P]$	2	3	4	11	4.53
$E[F^*_A]$	2.14	3.29	4.43	12.43	6.94

Table 4.2 shows that the expected values of F^*_A and F^*_P are approximately the same for various values of ϕ. With the expected values of F^*_A and F^*_P being approximately the same and the critical F associated with each being of different magnitude, and the critical value of F corresponding to F^*_P being the lower, the frequency of rejecting a false hypothesis using F^*_P is greater.

The method used to generate highly correlated data produced a significantly greater power of the test than the power of the test with uncorrelated data for levels (2,2,2) and (3,3,3) with $\phi = (1,2,3)$. This may have resulted from the method used to generate highly correlated data. From equation 3.24 it can be noted that the error component of any observation will contain 11 of the 12 random variables used to generate the previous observation, thereby providing consistency in the variance of the error term. Since the power of the test is the probability of rejecting the hypothesis that $\sigma^2_A = 0$ when it is false, and since highly correlated data provided a more consistent error variance than uncorrelated data, tests on correlated data would be more sensitive to changes in σ^2_A. However, as ϕ increases, the power increases for tests on both correlated and uncorrelated data and decreases the difference between the power with highly correlated and uncorrelated data.

Although the possibility of a negative F^* was discussed by Hudson and Krutchkoff (10) and a negative F^*_A occurred for levels (2,2,2) and (3,3,3) in this investigation, it is unlikely to occur in practice. Normally in a scientific or industrial experiment the experimenter will make his test on the interaction containing the most factors first, following this, he will test the interactions with the next largest number of factors and continue testing decreasing factor interactions until all interactions are tested. He then will test main effects. When testing, if any interaction is found to be not significant, that interaction is considered to be non-existent and the variance associated with this interaction is assumed to be zero. All components of the expected mean squares corresponding to variance of this interaction are then also assumed to be zero. If all interactions are zero, then the table of expected mean squares will show that main effects can be compared directly with the error mean square.

In this investigation no attempt was made to eliminate the variance components corresponding to those interactions which were found to be non-existent; therefore, where $MS_{AB} + MS_{AC}$ was less than MS_E, an approximate test was made and a negative F_A^* appeared for levels (2,2,2) and (3,3,3). In practice if either MS_{AB} or MS_{AC} or both were less than MS_E, a direct test of the null hypothesis, H_o: $\sigma^2{}_A = 0$, could be made. Neither F_A^* or F_P^* would be used as the test statistic, but a test involving one of the ratios MSA/MSAC, MSA/MSAC, or MSA/MSE with the appropriate degrees of freedom would be made.

The exceedingly high probability of type II error occurring at $a = b = c = 2$ results from the few degrees of freedom available. Appendix E shows that, out of 2000 samples taken, 1225 samples had one degree of freedom numerator and one degree of freedom denominator and 775 samples had two degrees of freedom numerator and one degree of freedom denominator. The critical values of F (with $\alpha = 0.05$ for these degrees of freedom combinations) are 161.0 and 200.0 respectively. When using the statistic F_P^*, where $a = b = c = 2$, F_P^* was less than 110 for $\phi = 1$, 2, and 3, but did exceed this value at times where $\phi = 10$. Since the low values of the degrees of freedom dictate that a number consistently higher than F^* be selected as critical F, we must accept the null hypothesis, even when it is false.

The low probability of type I error occurring at $a = b = c = 2$ results from the same large critical F values with the degree of freedom sets (1,1) and (2,1). The critical F is so large that the null hypothesis, H_o: $\sigma^2{}_A = 0$, cannot be rejected.

The distortion noted in the mean squares for data with unequal error variances can be attributed to the relative magnitude of the error variance to the mean squares for the other components of Y_{ijk}. The differences between the expected values of the mean squares due to A and to B and their actual values increased as the average value of the error variance increased.

Hudson and Krutchkoff (10), working with conforming data, noted that when the levels were balanced, the actual size of the test closely approximated the advertised size. This was particularly true where the levels were (4,4,4) or higher. In this investigation an attempt was made to determine if using balanced levels would cause greater agreement between conforming and non-conforming data. However, an examination of the appendices shows that this was not the case. In some cases balanced levels provide greater agreement and in some cases unbalanced levels provide greater agreement.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Based on the results of this investigation, a number of pertinent conclusions can be drawn. First, the size and power of tests employing Satterthwaite's synthetic mean squares are not significantly influenced by departures from the analysis of variance normality assumption. This conclusion is apparently valid for experimental error variances of the same magnitude or smaller than the other variance components.

Second, we have noticed that highly correlated error terms cause significant departures from the advertised size of the test, and corresponding irregularities in the power. This would tend to indicate that care in designing experiments or collecting data to minimize correlation should be exercised.

Third, if the data does not conform to the assumption of homoscedasticity, but the average value of σ^2 approximates the other variance components, there will be little effect on the size and power of the test. However, distortions of the mean squares will occur if the average error variance is greater than that of the other components of Y_{ijk}, and an experimenter is cautioned against estimating components of variance with these mean squares.

Fourth, in almost all cases the primary statistic F_p^* provides a smaller probability of type I error than does the alternate statistic,

F_A^*. However, the power using F_A^* is almost always higher than the power using F_P^*. Since a negative value of F_A^* will not occur in practice, since the experimenter can determine the amount of type I error he is willing to accept, and since the power of the test is greater using F_A^* we recommend the use of F_A^*.

Fifth, for the cases a = b = c = 2 considerable distortion in size and power frequently occurred. This is due, perhaps in part, to the low number of available degrees of freedom, and experimenters should be cautioned about employing Satterthwaite's method in those cases.

Last, the use of Satterthwaite's synthetic mean squares in analysis of variance appears to be quite robust, except in the special cases noted.

Recommendations for Further Study

This investigation has indicated that the following points are worthy of further study:

(i) As only highly correlated error terms significantly influence both the size and power of the test, some research to ascertain the degree of correlation at which these effects become apparent might be indicated. Also different types of correlation structure could be investigated. A logical part of this study would be an analysis of transformations on the original data to reduce the effects of correlation.

(ii) Compound sensitivity studies may be appropriate in some circumstances, that is, a study of the departures form the assumptions which occur in pairs (or other combinations).

(iii) This method of sensitivity analysis should be applied to statistical models and test for which the assumptions are less well understood, such as the multivariate analysis of variance and factor analysis.

APPENDICES

APPENDIX A

Appendix A contains seven tables which display the calculated size of the test under the conditions investigated. The size as shown in the tables is the proportion of times that a hypothesis is incorrectly rejected.

Table A1 provides for the comparison of statistics F_P^* and F_A^* for seven sets of levels, four distributions, and two variances. This table shows that, for the distributions investigated, the size of test is approximately equal for the four distributions for a given F^*, set of levels, and variance. Sign tests were made to determine if a particular F^* provided a consistently lower probability of type I error. Where $\sigma^2 = 0.01$ the probability of type I error is consistently lower when F_P^* is used; however, where $\sigma^2 = 1.0$ neither F_A^* nor F_P^* provides a consistently lower value of type I error.

Table A2 displays the size of the test for moderately correlated, highly correlated, and uncorrelated data using seven sets of levels, both F_P^* and F_A^*, and two variances. A Wilcoxon signed rank test was used to determine if differences existed between the size of the test with conforming data and the size of the test with correlated data. When the variance of the error term equals the variance of the other terms (i.e., where $\sigma^2 = 1.0$) there is no significant difference between the size of the test with uncorrelated or moderately correlated data. When the data is highly correlated, the size is significantly higher than both the moderately correlated and uncorrelated cases. Where ($\sigma^2 = 0.01$) the variance of the error term is much smaller than the variance of the other terms, there is no significant difference between the size of the test with uncorrelated, moderately correlated, or highly correlated data. Furthermore, there is no significant difference between the size of the test using F_P^* or F_A^*.

Table A3 displays the size of the test when the variance of the error term differs for each ϵ_{ijk}. A Wilcoxon signed rank test was made on the data in this table to determine if differences between the size of the test using class 1 data and the size of the test using conforming data exist. For a particular F^*, set of levels, and error variance there was no significant difference between the size of the test with class 1 data and the size of the test with conforming data.

Tables A4 and A5 display the size of the test when class 2 data was used. A Wilcoxon signed rank test was used to determine if differences in the size of the test resulted from using F_A^* or F_P^*. There is no significant difference between the size of the test using F_P^* or F_A^* when the average value of the variance of the error term is greater than the variance of the other components of the model. When the average value of the error variance is less than the variance of the other components of the model, the size of the test using F_A^* was found to be consistently greater.

Tables A6 and A7 are similar to Table A4 except that the error term is class 3. Table A6 shows that generally the size of the test increases for an increase in the average value of the variance of the error when using F_P^*. Both tables show that when the average values of the variance of the class 3 error term approaches the variance of the conforming error term, the size of the test is approximately the same for a given F^* and set of levels.

By scanning all seven tables the following facts may be noted. Generally the size of the test is greater using F_A^* than when using F_P^*. The exception to this generalization occurs where the variance of the error term is equal or greater than one and with set of levels (3,3,3), (2,4,4), and (5,2,2).

Table A1 Non-Normality: Size

(Proportion of Incorrect Rejections in 2000 Trials)

Run	No. of Levels	Distributions	$\sigma^2 = 1.0$ Size Using F_P^*	$\sigma^2 = 1.0$ Size Using F_A^*	$\sigma^2 = 0.01$ Size Using F_P^*	$\sigma^2 = 0.01$ Size Using F_A^*
	a = 2	Normal	.0000	.0000	.0000	.0000
1	b = 2	Uniform	.0000	.0000	.0000	.0000
	c = 2	Triangular	.0000	.0000	.0000	.0000
		Parabolic	.0000	.0010	.0000	.0005
	a = 3	Normal	.0275	.0200	.0435	.0500
2	b = 3	Uniform	.0285	.0215	.0440	.0510
	c = 3	Triangular	.0305	.0215	.0440	.0505
		Parabolic	.0315	.0205	.0445	.0515
	a = 4	Normal	.0405	.0420	.0445	.0485
3	b = 4	Uniform	.0385	.0420	.0450	.0485
	c = 4	Triangular	.0380	.0430	.0445	.0485
		Parabolic	.0415	.0455	.0440	.0490
	a = 5	Normal	.0495	.0560	.0455	.0455
4	b = 5	Uniform	.0460	.0555	.0450	.0460
	c = 5	Triangular	.0435	.0560	.0450	.0455
		Parabolic	.0450	.0525	.0465	.0465
	a = 2	Normal	.0575	.0350	.1340	.1365
5	b = 4	Uniform	.0570	.0350	.1350	.1375
	c = 4	Triangular	.0580	.0350	.1355	.1345
		Parabolic	.0545	.0345	.1375	.1385
	a = 2	Normal	.0440	.0515	.0510	.0530
6	b = 6	Uniform	.0430	.0510	.0505	.0540
	c = 6	Triangular	.0445	.0485	.0510	.0545
		Parabolic	.0455	.0510	.0500	.0535
	a = 5	Normal	.0315	.0065	.0165	.0230
7	b = 2	Uniform	.0260	.0040	.0155	.0220
	c = 2	Triangular	.0285	.0065	.0165	.0235
		Parabolic	.0270	.0110	.0150	.0215

Table A2. Correlation: Size

(Proportion of Incorrect Rejections in 2000 Trials)

			$\sigma^2 = 1.0$		$\sigma^2 = 0.01$	
Run	No. of Levels	Type of Corre-lation	Size Using F^*_P	Size Using F^*_A	Size Using F^*_P	Size Using F^*_A
	a = 2					
9	b = 2	High	.0000	.0075	.0000	.0010
17	c = 2	Moderate	.0000	.0040	.0000	.0010
1		None	.0000	.0000	.0000	.0000
	a = 3					
10	b = 3	High	.1005	.1170	.0515	.0615
18	c = 3	Moderate	.0385	.0295	.0420	.0495
2		None	.0275	.0200	.0435	.0500
	a = 4					
11	b = 4	High	.1535	.1625	.0545	.0590
19	c = 4	Moderate	.0430	.0460	.0555	.0570
3		None	.0405	.0420	.0450	.0485
	a = 5					
12	b = 5	High	.1655	.1720	.0445	.0445
20	c = 5	Moderate	.0440	.0495	.0400	.0410
4		None	.0495	.0560	.0455	.0455
	a = 2					
13	b = 4	High	.0860	.0900	.1280	.1390
21	c = 4	Moderate	.0545	.0275	.1365	.1390
5		None	.0575	.0350	.1340	.1365
	a = 2					
14	b = 6	High	.0755	.0795	.0380	.0405
22	c = 6	Moderate	.0360	.0435	.0410	.0425
6		None	.0440	.0515	.0510	.0530
	a = 5					
15	b = 2	High	.0800	.0890	.0115	.0170
23	c = 2	Moderate	.0410	.0180	.0145	.0225
7		None	.0315	.0065	.0165	.0230

Table A3 Unequal Variance Class 1: Size

(Proportion of Incorrect Rejections in 2000 Trials)

Run	Number of Levels	Range of σ^2	Actual Size F_P^*	Actual Size F_A^*
25	a = 2 b = 2 c = 3	1. to 5. 1.0	.0000 .0000	.0000 .0000
26	a = 3 b = 3 c = 3	1. to 5. 1.0	.0275 .0275	.0190 .0200
27	a = 4 b = 4 c = 4	1. to 5. 1.0	.0395 .0405	.0410 .0420
28	a = 5 b = 5 c = 5	1. to 5. 1.0	.0475 .0495	.0565 .0565
29	a = 2 b = 2 c = 2	.01 to .05 .01	.0000 .0000	.0000 .0003
30	a = 3 b = 3 c = 3	.01 to .05 .01	.0410 .0435	.0495 .0500
31	a = 4 b = 4 c = 4	.01 to .05 .01	.0415 .0445	.0425 .0485
32	a = 5 b = 5 c = 5	.01 to .05 .01	.0480 .0455	.0495 .0455

Table A4. Unequal Variance Class 2: Size

(Proportion of Incorrect Rejections in 2000 Trials)

Run	Number of Levels	Range of σ^2	Actual Size F_P^*	Actual Size F_A^*
		1 to 4	.0000	.0000
33	a = 2	1 to 5	.0000	.0000
	b = 2	1 to 6	.0000	.0000
	c = 2	1 to 7	.0435	.0515
		1.0	.0000	.0000
		1 to 4	.0275	.0080
34	a = 3	1 to 5	.0350	.0080
	b = 3	1 to 6	.0335	.0045
	c = 3	1 to 7	.0285	.0475
		1.0	.0275	.0200
		1 to 4	.0355	.0335
35	a = 4	1 to 5	.0350	.0280
	b = 4	1 to 6	.0335	.0255
	c = 4	1 to 7	.0235	.0430
		1.0	.0405	.0420
		1 to 4	.0460	.0575
36	a = 5	1 to 5	.0440	.0550
	b = 5	1 to 6	.0440	.0525
	c = 5	1 to 7	.0405	.0565
		1.0	.0495	.0560
	a = 5	1 to 6	.0530	.0005
37	b = c = 2	1.0	.0315	.0065
	a = 2	1 to 6	.0345	.0150
38	b = c = 4	1.0	.0575	.0350
	a = 2	1 to 6	.0385	.0335
39	b = c = 6	1.0	.0440	.0515

Table A5. Unequal Variance Class 2: Size

(Proportion of Incorrect Rejections in 2000 Trials)

Run	Number of Levels	Range of σ^2	Actual Size F_P^*	Actual Size F_A^*
	a = 2	.01 to .02	.0000	.0005
40	b = 2	.01 to .1	.0000	.0000
	c = 2	.01 to .8	.0000	.0000
		.01	.0000	.0000
	a = 3	.01 to .02	.0480	.0550
41	b = 3	.01 to .1	.0415	.0495
	c = 3	.01 to .8	.0370	.0390
		.01	.0435	.0500
	a = 4	.01 to .02	.0460	.0480
42	b = 4	.01 to .1	.0465	.0485
	c = 4	.01 to .8	.0405	.0430
		.01	.0445	.0480
	a = 5	.01 to .02	.0440	.0445
43	b = 5	.01 to .1	.0460	.0470
	c = 5	.01 to .8	.0505	.0430
		.01	.0455	.0455

Table A6. Unequal Variance Class 3: Size

(Proportion of Incorrect Rejections in 2000 Trails)

Run	Number of Levels	Values of σ^2	Size Using F_P^*	Size Using F_A^*	Average Value of σ^2
		1.0	.0000	.0000	1.0
44	a=b=c=2	1 and 3	.0000	.0000	2.0871
		1 and 4	.0000	.0000	2.6216
		1 and 5	.0000	.0000	3.1547
		1 and 9	.0000	.0000	5.2890
		1 and 16	.0000	.0000	9.0239
		1 and 25	.0000	.0000	13.8259
		1 and 36	.0000	.0000	19.6951
		1.0	.0275	.0200	1.0
45	a=b=c=3	1 and 3	.0360	.0105	2.3349
		1 and 4	.0350	.0050	3.0028
		1 and 5	.0365	.0030	3.6707
		1 and 9	.0390	.0005	6.3423
		1 and 16	.0460	.0005	11.0175
		1 and 25	.0510	.0005	17.0285
		1 and 36	.0555	.0005	24.3154
		1.0	.0405	.0420	1.0
46	a=b=c=4	1 and 3	.0390	.0410	1.5108
		1 and 4	.0370	.0380	1.7623
		1 and 5	.0365	.0360	2.0137
		1 and 9	.0380	.0240	3.0195
		1 and 16	.0460	.0135	4.7797
		1 and 25	.0500	.0085	7.0428
		1 and 36	.0545	.0075	9.8087
		1.0	.0495	.0560	1.0
47	a=b=c=5	1 and 3	.0470	.0565	1.4134
		1 and 4	.0460	.0570	1.6155
		1 and 5	.0465	.0555	1.8176
		1 and 9	.0465	.0600	2.6262
		1 and 25	.0585	.0580	5.8603
		1 and 36	.0655	.0590	8.0837

Table A7. Unequal Variance Class 3: Size

(Proportion of Incorrect Rejections in 2000 Trials)

Run	Number of Levels	Values of σ^2	Size Using F_P^*	Size Using F_A^*	Average Value of σ^2
		1.0	.0000	.0000	1.0
48	a=b=c=2	.5 and 1.5	.0005	.0000	1.00
		1.0	.0275	.0200	1.0
49	a=b=c=3	.6 and 1.8	.0315	.0235	1.01
		1.0	.0405	.0420	1.0
50	a=b=c=4	2/3 and 2	.0390	.0420	1.00
		1.0	.0495	.0560	1.0
51	a=b=c=5	5/7 and 15/7	.0480	.0555	1.01

APPENDIX B

Appendix B contains tables displaying the power of the test.

Tables B1, B2, and B3 include the four distributions and show the following:

(i) The power of the test increases as ϕ increases.

(ii) For a particular ϕ, σ^2, and F^* the type distribution has little effect on the power of the test.

(iii) For levels (4,4,4) and (5,5,5) with σ^2 either 1.0 or 0.01, the power of the test is greater when F_A^* is used.

(iv) For levels (2,2,2) and $\sigma^2 = 1.0$ the power of the test is greater when F_A^* is used.

(v) For levels (3,3,3) and $\sigma^2 = 1.0$ the power of the test is greater when F_P^* is used.

(vi) The power of the test increases as N increases $N = a \cdot b \cdot c$.

Tables B4, B5, and B6 each contain displays of the power of the test for moderately correlated, highly correlated, and uncorrelated data and show the following:

(i) Where $\sigma^2 = 1$ the power of the test is consistently greater for highly correlated data than for moderately correlated or uncorrelated data. The difference between the power of the test with highly correlated data and the power of the test with moderate or uncorrelated data is greatest where $\phi = 1$. This difference decreases as ϕ increases and where $\phi = 10$ the power of the test with highly correlated data, moderately correlated data, and uncorrelated data is approximately equal. The preceeding is true except for levels (2,2,2).

(ii) For a particular F^*, σ^2, set of levels, and type of correlation the power of the test increases as ϕ increases.

(iii) With levels (3,3,3), (4,4,4) or (5,5,5) the power of the test using F_P^* or F_A^* is approximately equal for a particular σ^2, ϕ, and type of correlation. However, the power of the test with F_A^* is consistently higher.

(iv) For a particular F^*, ϕ, σ^2, and type of correlation the power of the test increases as N increases. $N = a \cdot b \cdot c$.

Tables B7, B8, and B9 show the power of the test for class 3 unequal error variances. These tables show the following:

(i) For a particular set of levels F^* and range of error variances, the power of the test increases as ϕ increases.

(ii) For a particular ϕ, range of error variance, and sets of levels (2,2,2), (4,4,4) or (5,5,5) the power using F_A^* is greater than the power using F_P^*.

(iii) For class 3 data the power of the test is less for a particular set of levels, ϕ, and F^* than it is for conforming data. The preceding is not true for levels (3,3,3).

It should be noted that for $a = b = c = 3$, when investigating class 3 data displayed on B7 and B8, the variance of the error terms ϵ_{211} through ϵ_{233} were given the same high σ^2 as were ϵ_{311} through ϵ_{333}. This accounts for the noticeably lower values of the power of the test where $a = b = c = 3$ for class 3 data. It also gives an indication that when all other parameters are equal and the data is class 3, an increase in the average value of the variance will result in a decrease in the power of the test. An investigation of all values on tables one, two, and three will also show this decrease in power. For a given set of circumstances the average value of the error for those statistics on B7 is larger than the average error on B8 which is larger than the average error on B9. The power of the test for a given set of circumstances is greater on B9 than B8 and greater on B8 than on B7.

Table B1. Non-Normality: Power

(Proportion of Correct Rejections in 2000 Trials)

$\alpha = 0.05 \quad a = b = c = 4$

			$\sigma^2 = 1.0$		$\sigma^2 = 0.01$	
Run	ϕ	Distri- butions	Power Using F^*_P	Power Using F^*_A	Power Using F^*_P	Power Using F^*_A
		Normal	.1990	.2085	.2050	.2120
1	1	Uniform	.1940	.2075	.2065	.2120
		Triangular	.1925	.2050	.2070	.2140
		Parabolic	.1910	.2030	.2075	.2130
		Normal	.3730	.3840	.3775	.3850
2	2	Uniform	.3675	.3785	.3770	.3860
		Triangular	.3705	.3835	.3780	.3860
		Parabolic	.3670	.3805	.3775	.3855
		Normal	.5015	.5140	.4875	.5000
3	3	Uniform	.4950	.5115	.4805	.5000
		Triangular	.4975	.5120	.4875	.4995
		Parabolic	.4935	.5150	.4885	.5005
		Normal	.8715	.8770	.8630	.8675
4	10	Uniform	.8675	.8770	.8635	.8570
		Triangular	.8685	.8775	.8640	.8675
		Parabolic	.8695	.8725	.8640	.8675

Table B2. Non-Normality: Power

(Proportion of Correct Rejections in 2000 Trials)

$\alpha = 0.05 \quad a = b = c = 5$

			$\sigma^2 = 1.0$		$\sigma^2 = 0.01$	
Run	ϕ	Distri- butions	Power Using F^*_P	Power Using F^*_A	Power Using F^*_P	Power Using F^*_A
		Normal	.2710	.2970	.2580	.2605
		Uniform	.2680	.2910	.2590	.2590
		Triangular	.2700	.2925	.2580	.2605
5	1	Parabolic	.2705	.2865	.2570	.2585
		Normal	.4900	.5065	.4765	.4780
		Uniform	.4835	.5030	.4760	.4785
		Triangular	.4840	.5070	.4760	.4775
6	2	Parabolic	.4875	.5085	.4750	.4780
		Normal	.6320	.6420	.6155	.6160
		Uniform	.6270	.6405	.6160	.6160
		Triangular	.6260	.6420	.6160	.6160
7	3	Parabolic	.6325	.6440	.6165	.6165
		Normal	.9325	.9325	.9305	.9305
		Uniform	.9300	.9380	.9300	.9305
		Triangular	.9320	.9370	.9300	.9305
8	10	Parabolic	.9305	.9345	.9310	.9310

Table B3. Non-Normality: Power

(Proportion of Correct Rejections in 2000 Trials)

$\sigma^2 = 1.0$

			a = b = c = 2		a = b = c = 3	
Run	ϕ	Distri-butions	Power Using F_P^*	Power Using F_A^*	Power Using F_P^*	Power Using F_A^*
9	1	Normal	.0000	.0065	.1375	.1225
		Uniform	.0000	.0015	.1235	.1150
		Triangular	.0000	.0025	.1265	.1150
		Parabolic	.0000	.0025	.1190	.1040
10	2	Normal	.0000	.0190	.2345	.2255
		Uniform	.0000	.0015	.2235	.2060
		Triangular	.0000	.0140	.2230	.2010
		Parabolic	.0000	.0055	.2190	.2060
11	3	Normal	.0000	.0320	.3330	.3210
		Uniform	.0000	.0150	.3110	.2955
		Triangular	.0000	.0245	.3075	.2905
		Parabolic	.0000	.0120	.3140	.2950
12	10	Normal	.0000	.0945	.6580	.6480
		Uniform	.0000	.0535	.6550	.6395
		Triangular	.0000	.0795	.6580	.6355
		Parabolic	.0000	.0345	.6575	.6330

Table B4. Correlation: Power

(Proportion of Correct Rejection in 2000 Trials)

$\sigma^2 = 1.0$

			a = b = c = 2		a = b = c = 3	
Run	ϕ	Type of Corre-lation	Power Using F^*_P	Power Using F^*_A	Power Using F^*_A	Power Using F^*_A
13		High	.0000	.0455	.1965	.2125
17	1	Moderate	.0000	.0135	.1410	.1430
9		None	.0000	.0065	.1375	.1225
14		High	.0000	.0895	.2970	.3195
18	2	Moderate	.0000	.0265	.2490	.2460
10		None	.0000	.0190	.2345	.2255
15		High	.0000	.1275	.3865	.4225
19	3	Moderate	.0000	.0420	.3185	.3210
11		None	.0000	.0320	.3330	.3210
16		High	.0000	.3365	.6790	.6960
20	10	Moderate	.0000	.1185	.6340	.6340
12		None	.0000	.0945	.6580	.6480

Table B5. Correlation: Power

(Proportion of Correct Rejection in 2000 Trials)

a = b = c = 4

Run	ϕ	Type of Corre-lation	$\sigma^2 = 1.0$ Power Using F_P^*	$\sigma^2 = 1.0$ Power Using F_A^*	$\sigma^2 = 0.01$ Power Using F_P^*	$\sigma^2 = 0.01$ Power Using F_A^*
21	1	High	.3225	.3350	.2175	.2230
25		Moderate	.2150	.2260	.2185	.2260
1		None	.1990	.2085	.2050	.2120
22	2	High	.4620	.4710	.3820	.3905
26		Moderate	.3590	.3790	.3670	.3730
2		None	.3730	.3840	.3775	.3850
23	3	High	.5720	.5815	.4950	.5050
27		Moderate	.4845	.4970	.4830	.4915
3		None	.5015	.5140	.4875	.5000
24	10	High	.8560	.8635	.8670	.8730
28		Moderate	.8495	.8605	.8615	.8655
4		None	.8715	.8770	.8630	.8675

Table B6. Correlation: Power

(Proportion of Correct Rejections in 2000 Trials)

a = b = c = 5

Run	ϕ	Type of Corre-lation	$\sigma^2 = 1.0$ Power Using F_P^*	$\sigma^2 = 1.0$ Power Using F_A^*	$\sigma^2 = 0.01$ Power Using F_P^*	$\sigma^2 = 0.01$ Power Using F_A^*
29	1	High	.3895	.3945	.2730	.2760
33		Moderate	.2575	.2695	.2465	.2490
5		None	.2710	.2970	.2580	.2605
30	2	High	.5340	.5390	.4850	.4870
34		Moderate	.4425	.4605	.4465	.4485
6		None	.4425	.4605	.4465	.4485
31	3	High	.6605	.6650	.6355	.6365
35		Moderate	.5900	.6080	.5995	.6005
7		None	.6320	.6420	.6155	.6160
32	10	High	.9320	.9235	.9310	.9310
36		Moderate	.9290	.9335	.9345	.9355
8		None	.9325	.9325	.9305	.9305

Table B7. Unequal Variance Class 3: Power

(Proportion of Correct Rejections in 2000 Trials)

			σ^2 = 1 and 4		σ^2 = 1.0	
Run	ϕ	Number of Levels	Power Using F_P^*	Power Using F_A^*	Power Using F_P^*	Power Using F_A^*
41	1	a=b=c=2	.0000	.0000	.0000	.0065
	2		.0000	.0015	.0000	.0019
	3		.0000	.0030	.0000	.0320
	10		.0000	.0110	.0000	.0945
42	1	a=b=c=3	.0940	.0375	.1375	.1225
	2		.1680	.0785	.2345	.2255
	3		.2390	.1320	.3330	.3210
	10		.5390	.3690	.6580	.6480
43	1	a=b=c=4	.1760	.1945	.1990	.2085
	2		.3335	.3485	.3730	.3840
	3		.4480	.4645	.5015	.5140
	10		.8340	.8420	.8715	.8770
44	1	a=b=c=5	.2475	.2735	.2710	.2970
	2		.4500	.4730	.4900	.5065
	3		.5900	.6140	.6320	.6420
	10		.9210	.9285	.9325	.9395

Table B8. Unequal Variance Class 3: Power

(Proportion of Correct Rejections in 2000 Trials)

Run	ϕ	Number of Levels	$\sigma^2 = 1$ and 3: Power Using F_P^*	$\sigma^2 = 1$ and 3: Power Using F_A^*	$\sigma^2 = 1.0$: Power Using F_P^*	$\sigma^2 = 1.0$: Power Using F_A^*
45	1	a=b=c=2	.0000	.0000	.0000	.0065
	2		.0000	.0035	.0000	.0019
	3		.0000	.0060	.0000	.0320
	10		.0000	.0230	.0000	.0945
46	1	a=b=c=3	.0980	.0490	.1375	.1225
	2		.1810	.1035	.2345	.2255
	3		.2580	.1535	.3330	.3210
	10		.5735	.4300	.6580	.6480
47	1	a=b=c=4	.1815	.1950	.1990	.2085
	2		.3425	.3585	.3730	.3840
	3		.4575	.4775	.5015	.5140
	10		.8430	.8500	.8715	.8770
48	1	a=b=c=5	.2535	.2790	.2710	.2970
	2		.4550	.4840	.4900	.5065
	3		.6030	.6240	.6320	.6420
	10		.9260	.9305	.9325	.9395

Table B9. Unequal Variance Class 3: Power

(Proportion of Correct Rejection in 2000 Trials)

Run	ϕ	No. of Levels and Values of ^{2}	Power Using F_P^*	Power Using F_A^*	$\sigma^2 = 1.0$ Power Using F_P^*	Power Using F_A^*
		a=b=c=2				
49	1	.5 and 1.5	.0000	.0045	.0000	.0065
	2		.0000	.0130	.0000	.0019
	3		.0000	.0250	.0000	.0320
	10		.0005	.0800	.0000	.0945
		a=b=c=3				
50	1	.6 and 1.5	.1315	.1180	.1375	.1225
	2		.2195	.2070	.2345	.2255
	3		.3135	.2990	.3330	.3210
	10		.6535	.6310	.6580	.6480
		a=b=c=4				
51	1	2/3 and 2	.1950	.2075	.1990	.2085
	2		.3720	.3840	.3730	.3840
	3		.4890	.5060	.5015	.5140
	10		.8650	.8740	.8715	.8770
		a=b=c=5				
52	1	5/7 and 15/7	.2690	.2895	.2710	.2970
	2		.4825	.5035	.4900	.5065
	3		.6245	.6420	.6320	.6420
	10		.9320	.9370	.9325	.9395

APPENDIX C

Appendix C contains listings of the two subroutines which were used in the generation and testing of all data. The main program is not listed because it required modification for each run and these modifications were discussed in Chapter III.

Table C1

Fortran Listing of Subroutine Which Provides Random Variables Drawn From a Population With a Mean Equal to Zero and a Variance Equal to One.

```
      FUNCTION RDM (I)
      DIMENSION ARY (20000)
      IF (I.NE.1) GO TO 100
      ARY(I)=6734269
      CALL RANDU(ARY, 20000)
  100 RDM = ARY(I + 1)
      IF (I.EQ.19999) I=0
      I = I + 1
      RETURN
      END
```

NOTE: By the proper location of RDM(MACK) in the main program, the same array of uniform (0,1) random variables can be reproduced. This provides both the conforming and non-conforming data with identical components for the same observation, Y_{ijk}, except those components which must be different for comparison purposes.

Table C2.

Fortran Analysis of Variance Subroutine

```
      SUBROUTINE ANOVA
      DIMENSIONY(7,7,7),TI(7),TJ(7), TK(7),TIJ(7,7),TIK(7,7)
      COMMON/BLOKB/XMSA,XMSB,XMSC,XMSAB,XMSAC,XMSE
      COMMON/BLOKA/Y,DFA,DFB,DFC,DFE,BCN,ACN,NA,NB,NC,ABN,GN,DFAB,DFAC
      TIS=0.0
      DO 11 I=1,NA
      TI(I)=0.0
      DO 1 J=1,NB
      DO 1 K=1,NC
    1 TI(I)=TI(I)+Y(I,J,K)
   11 TIS=TIS+TI(I)**2
      TJS=0.0
      DO 22 J=1,NB
      TJ(J)=0.0
      DO 2 I=1,NA
      DO 2 K=1,NB
    2 TJ(J)=TJ(J)+Y(I,J,K)
   22 TJS=TJS+TJ(J)**2
      TKS=0.0
      DO 33 K=1,NC
      TK(K)=0.0
      DO 3 I=1,NA
      DO 3 J=1,NB
    3 TK(K)=TK(K)+Y(I,J,K)
   33 TKS=TKS+TK(K)**2
      TIJS=0.0
      DO 44 I=1,NA
      DO 44J=1,NB
      TIJ(I,J)=0.0
      DO 4 K=1,NC
    4 TIJ(I,J)=TIJ(I,J)+Y(I,J,K)
   44 TIJS=TIJS+TIJ(I,J)**2
      TIKS=0.0
      DO 55 I=1,NA
      DO 55 K=1,NC
      TIK(I,K)=0.0
      DO 5 J=1,NB
    5 TIK(I,K)=TIK(I,K)+Y(I,J,K)
   55 TIKS=TIKS+TIK(I,K)**2
      T=0.0
      DO 8 I=1,NA
    8 T=T+TI(I)
      TS=0.0
```

```
      DO 9 I=1,NA
      DO 9 K=1,NB
      DO 9 J=1,NC
    9 TS=TS+Y(I,J,K)**2
      CF=T**2/GN
      SSA=TIS/BCN-CF
      SSB=TJS/ACN-CF
      SSC=TKS/ABN-CF
      SSAB=TIJS/RC-CF-SSA-SSB
      SSAC=TIKS/RB-CF-SSA-SSC
      SSE=TS-CF-SSA-SSB-SSC-SSAB-SSAC
      XMSA=SSA/DFA
      XMSB=SSB/DFB
      XMSC=SSC/DFC
      XMSAB=SSAB/DFAB
      XMSAC=SSAC/DFAC
      XMSE=SSE/DFE
      RETURN
      END
```

APPENDIX D

This appendix includes tables of expected mean squares, and block standard deviations for several sets of levels of A, B, and C. Not all of these statistics investigated are included. Those statistics excluded from the appendix follow the same pattern as those herein, i.e., the average mean squares approximate the expected mean squares, and the block standard deviation for a given statistic is generally one-fourth the value of the average mean square.

Table D1. Examples of Typical Block Standard Deviations

(Average = sum of twenty block averages divided by twenty, where the block average is the sum of a hundred replicates of a particular mean square within a block divided by one hundred).

			Parabolic		Normal	
Number of Levels	Mean Squares and Size	Expected Averages	Actual Average	Block Standard Deviation	Actual Average	Block Standard Deviation
a = 2	MSA	5.0000	5.0885	1.1707	5.1524	1.1851
b = 2	MSAB	3.0000	2.9327	.6790	2.9900	.6920
c = 2	MSAC	3.0000	3.0200	.6964	3.0512	.7036
	MSE	1.0000	1.0056	.2308	1.0205	.2349
	SIZE	.0500	.0000	.0000	.0000	.0000
a = 3	MSA	7.0000	6.9678	1.6118	6.9046	1.5949
b = 3	MSAB	4.0000	3.9887	.9161	4.0075	.9215
c = 3	MSAC	4.0000	4.0987	.9428	4.0929	.9410
	MSE	1.0000	1.0020	.2297	1.0090	.2314
	SIZE	.0500	.0315	.0080	0.0225	0.0075
a = 4	MSA	9.0000	8.9618	2.0567	8.9501	2.0542
b = 4	MSAB	5.0000	4.9760	1.1412	5.0075	1.1482
c = 4	MSAC	5.0000	5.0462	1.1572	5.0537	1.1588
	MSE	1.0000	.9980	.2289	1.0093	.2313
	SIZE	.0500	.0415	.0103	.0405	.0100
a = 5	MSA	11.0000	11.0044	2.5259	11.2196	2.2758
b = 5	MSAB	6.0000	6.0591	1.3890	5.9978	1.3747
c = 5	MSAC	6.0000	6.0334	1.3800	6.0772	1.3929
	MSE	1.0000	1.0025	.2297	1.0091	.2312
	SIZE	.0500	.0450	.0109	.0495	.0129
a = 2	MSA	9.0000	9.5806	2.2119	9.4489	2.1947
b = 4	MSAB	5.0000	5.0339	1.1571	5.0636	1.1639
c = 4	MSAC	5.0000	4.9285	1.1330	4.9105	1.1291
	MSE	1.0000	.9992	.2290	1.0144	.2325
	SIZE	.0500	.0545	.0132	.0575	.0139
a = 2	MSA	13.0000	13.2923	3.0589	13.3059	3.0615
b = 6	MSAB	7.0000	6.9361	1.5910	6.9254	1.5885
c = 6	MSAC	7.0000	6.9252	1.5892	6.9051	1.5849
	MSE	1.0000	.9979	.2289	1.0061	.2305
	SIZE	.0500	.0455	.0110	.0440	.0107

Table D2. Examples of Typical Block Standard Deviations

(Average = sum of twenty block averages divided by twenty, where the block average is the sum of a hundred replicates of a particular mean square within a block divided by one hundred).

			Uniform		Triangular	
Number of Levels	Mean Squares and Size	Expected Averages	Actual Average	Block Standard Deviation	Actual Average	Block Standard Deviation
a = 2	MSA	5.0000	5.1022	1.1735	5.1428	1.1829
b = 2	MSAB	3.0000	2.9596	.6847	2.9712	.6880
c = 2	MSAC	3.0000	3.0404	.7008	3.0439	.7015
	MSE	1.0000	1.0027	.2303	1.0210	.2344
	SIZE	.05	.0000	.0000	.0000	.0000
a = 3	MSA	7.0000	6.9357	1.6039	6.9469	1.6054
b = 3	MSAB	4.0000	4.0097	.9211	4.0112	.9220
c = 3	MSAC	4.0000	4.0932	.9413	4.0999	.9430
	MSE	1.0000	1.0009	.2295	1.0077	.2311
	SIZE	.05	.0285	.0074	.0305	.0081
a = 4	MSA	9.0000	8.9603	2.0564	8.9336	2.0505
b = 4	MSAB	5.0000	4.9884	1.1439	4.9946	1.1453
c = 4	MSAC	5.0000	5.0510	1.1581	5.0541	1.1589
	MSE	1.0000	1.0006	.2293	1.0051	.2303
	SIZE	.05	.0385	.0096	.0380	.0097
a = 5	MSA	11.0000	11.0023	2.5255	11.0072	2.5263
b = 5	MSAB	6.0000	6.0591	1.3890	6.0655	1.3905
c = 5	MSAC	6.0000	6.0438	1.3854	6.0430	1.3853
	MSE	1.0000	1.0018	.2296	1.0064	.2306
	SIZE	.05	.0460	.0110	.0435	.0106
a = 2	MSA	9.0000	9.4811	2.2031	9.4741	2.2014
b = 4	MSAB	5.0000	5.0611	1.1634	5.0669	1.1646
c = 4	MSAC	5.0000	4.9066	1.1281	4.9929	1.1317
	MSE	1.0000	1.0033	.2299	1.0092	.2313
	SIZE	.05	.0570	.0136	.0580	.0141
a = 2	MSA	13.0000	13.2949	3.0596	13.2860	3.0562
b = 6	MSAB	7.0000	6.9369	1.5913	6.9351	1.5910
c = 6	MSAC	7.0000	6.9179	1.5877	6.9158	1.5875
	MSE	1.0000	.0081	.2287	1.0030	0.2298
	SIZE	.05	.0430	.0105	.0445	.0109

APPENDIX E

This appendix includes 14 tables which display the average degrees of freedom for the numerator and the denominator and selected distributions of these degrees of freedom. Those included are representative of all those investigated.

Tables E1 and E2 show the average degrees of freedom numerator and denominator calculated for the normal, uniform, triangular and parabolic distributions. The trends that are apparent on these tables are similar, for correlated data and data with unequal error variances. It is obvious from these tables that the average degrees of freedom numerator and denominator tend to increase as the total number of observations increase. However, this increase is not solely dependent on the total number of observations, but is influenced by the arrangement of the number of levels. For example if a = b = c = 3, n = 27, and the average degrees of freedom numerator equals 4.1134. But if a = 5, b = c = 2, n = 20, and degrees of freedom numerator equals 6.0341. It can also be noted that except for where a = b = c = 2 and a = 2, b = c = 4, the degrees of freedom denominator are greater than the degrees of freedom numerator. This is true for both F_P^* and F_A^* and for $\sigma^2 = 1.0$ or $\sigma^2 = 0.01$.

Tables E3 through E8 display the distribution of the degrees of freedom combinations for 2000 observations under several conditions for a typical set of treatment levels using F_P^*. The levels a = 2, b = c = 4 were selected for display because the spread of the distributions for this combination of treatment levels lies between the extremely tight spread occurring at a = b = c = 2 and the wide spread occurring at a = b = c = 5 (see Table E12). Where a = b = c = 2, 1225 observations have one degree of freedom numerator and one degree of freedom denominator, and 775 observations have two degrees of freedom numerator and two degrees of freedom denominator.

Tables E9, E10, and E11 show the distribution of the degrees of freedom denominator under several different conditions. Tables E9, E10, and E11 show that the distribution degrees of freedom for conforming data and the various types of non-conforming data generally follow the same pattern for a particular set of levels and error variance.

In addition to the general comments made about Tables E1 and E2, a glance at Table E13 shows that for data with class 3 error variance the degrees of freedom numerator and denominator increase with an increase in the average variance of the error term.

An interesting phenomenon occurs with moderately correlated data, Table E14. When F_p^* is used, the average degrees of freedom numerator are greater for moderately correlated data than for uncorrelated or highly correlated data.

Table El. Average Values of Degrees of Freedom

(Average = sum of the degrees of freedom divided by 2000 trials)

$\sigma^2 = 1.0$

Level a,b,c	Type of Distribution	F_P^* Average DF NUM.	F_P^* Average DF DENOM.	F_A^* Average DR NUM.	F_A^* Average DF DENOM.	N
2,2,2	Normal	1.7838	1.5398	1.0000	.9746	8
	Uniform	1.7996	1.5730	1.0000	.9833	
	Triangular	1.7887	1.5454	1.0000	.9580	
	Parabolic	1.8138	1.5212	1.0000	.9047	
3,3,3	Normal	4.1130	6.9166	2.0000	4.8914	27
	Uniform	4.1244	6.9091	2.0000	4.8766	
	Triangular	4.1145	6.9103	2.0000	4.8820	
	Parabolic	4.1592	6.9148	2.0000	4.8614	
4,4,4	Normal	5.3975	15.7611	3.0000	12.3483	64
	Uniform	5.4007	15.7287	3.0000	12.3347	
	Triangular	5.3906	15.7477	3.0000	12.3429	
	Parabolic	5.4143	15.6732	3.0000	12.2813	
5,5,5	Normal	5.7612	30.4540	4.0000	25.2521	125
	Uniform	5.8013	30.4064	4.0000	25.2516	
	Triangular	5.7795	30.4172	4.0000	25.2404	
	Parabolic	5.7446	30.4362	4.0000	25.667	
2,4,4	Normal	4.9640	4.7601	1.0000	3.3669	32
	Uniform	4.9461	4.7604	1.0000	3.3855	
	Triangular	4.9061	4.7652	1.0000	3.3867	
	Parabolic	4.9596	4.7376	1.0000	3.4624	
2,6,6	Normal	6.5212	8.9151	1.0000	7.4352	72
	Uniform	6.7261	8.9199	1.0000	7.4458	
	Triangular	6.7242	8.9218	1.0000	7.4391	
	Parabolic	6.8172	8.9266	1.0000	7.4514	
5,2,2	Normal	6.0341	6.9182	4.0000	4.3365	20
	Uniform	6.0498	6.9237	4.0000	4.3178	
	Triangular	6.0657	6.9171	4.0000	4.3041	
	Parabolic	6.0585	6.9581	4.0000	4.2485	

Table E2. Average Values of Degrees of Freedom:

(Average = sum of the degrees of freedom divided by 2000 trials)

$\sigma^2 = 0.01$

Level a,b,c	Type of Distribution	F_P^* Average DF NUM.	F_P^* Average DF DENOM.	F_A^* Average DF NUM.	F_A^* Average DF DENOM.	N
2,2,2	Normal	1.1043	1.7926	1.0000	1.7780	8
	Uniform	1.1032	1.7933	1.0000	1.7800	
	Triangular	1.1054	1.7934	1.0000	1.7798	
	Parabolic	1.0936	1.7939	1.0000	1.7939	
3,3,3	Normal	2.0572	7.1309	2.0000	7.1007	27
	Uniform	2.0544	7.1307	2.0000	7.1007	
	Triangular	2.0560	7.1319	2.0000	7.1018	
	Parabolic	2.0534	7.1295	2.0000	7.0993	
4,4,4	Normal	3.0282	15.2309	3.0000	15.1759	64
	Uniform	3.0279	15.2306	3.0000	15.1760	
	Triangular	3.0284	15.2303	3.0000	15.1753	
	Parabolic	3.0281	15.2286	3.0000	15.1734	
5,5,5	Normal	4.0156	30.4738	4.0000	30.4099	125
	Uniform	4.0156	30.4751	4.0000	30.4117	
	Triangular	4.0156	30.4753	4.0000	30.4115	
	Parabolic	4.0156	30.4774	4.0000	30.4138	
2,4,4	Normal	1.3839	4.4979	1.0000	4.2727	32
	Uniform	1.3916	4.4982	1.0000	4.2709	
	Triangular	1.3898	4.4970	1.0000	4.2719	
	Parabolic	1.4163	4.4926	1.0000	4.2602	
2,6,6	Normal	1.5618	9.0444	1.0000	9.0264	72
	Uniform	1.5467	9.0457	1.0000	9.0278	
	Triangular	1.5389	9.0455	1.0000	9.0276	
	Parabolic	1.6002	9.0470	1.0000	9.0291	
5,2,2	Normal	4.0811	7.0316	4.0000	6.9861	20
	Uniform	4.0790	7.0350	4.0000	6.9902	
	Triangular	4.0796	7.0310	4.0000	6.9859	
	Parabolic	4.0757	7.0350	4.0000	6.9902	

Table E3. Moderate Correlation (F_P^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$$\sigma^2 = 1,\ a = 2,\ b = 4,\ c = 4$$

Degrees of Freedom Numerator	Degrees of Freedom Denominator				
	1	2	3	4	5
1	0	0	410	312	646
2	0	0	58	34	62
3	0	0	29	22	32
4	0	0	16	8	18
5	0	0	9	7	12
6	0	0	12	7	5
7	0	0	9	3	12
8	0	0	11	1	8
9	0	0	5	5	8
10	0	0	3	4	7
11	0	0	2	3	5
12	0	0	3	3	5
13	0	0	3	3	4
14	0	0	3	0	1
15	0	0	3	0	8
16	0	0	9	1	4
17	0	0	1	3	3
18	0	0	45	32	81

Table E4. High Correlation (F_P^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$$\sigma^2 = 1, a = 2, b = 4, c = 4$$

Degrees of Freedom Numerator	Degrees of Freedom Denominator 1	2	3	4	5
1	0	0	932	123	909
2	0	0	2	1	3
3	0	0	1	1	1
4	0	0	1	0	3
5	0	0	2	1	1
6	0	0	3	1	0
7	0	0	0	0	0
8	0	0	1	0	0
9	0	0	0	0	0
10	0	0	0	0	0
11	0	0	0	0	0
12	0	0	1	0	0
13	0	0	1	0	0
14	0	0	1	0	1
15	0	0	0	0	0
16	0	0	2	0	0
17	0	0	1	0	1
18	0	0	4	0	2

Table E5. Normal (F_P^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$$\sigma^2 = 1,\ a = 2,\ b = 4,\ c = 4$$

Degrees of Freedom Numerator	Degrees of Freedom Denominator 1	2	3	4	5
1	0	0	348	247	567
2	0	0	46	44	86
3	0	0	25	25	38
4	0	0	23	7	22
5	0	0	13	11	18
6	0	0	11	3	12
7	0	0	12	6	17
8	0	0	7	3	9
9	0	0	4	4	11
10	0	0	7	7	7
11	0	0	6	4	4
12	0	0	2	3	10
13	0	0	3	5	4
14	0	0	6	2	8
15	0	0	6	2	6
16	0	0	5	5	10
17	0	0	8	10	21
18	0	0	72	53	105

Table E6. Uniform (F_p^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$\sigma^2 = 1,\ a = 2,\ b = 4,\ c = 4$

Degrees of Freedom Numerator	Degrees of Freedom Denominator 1	2	3	4	5
1	0	0	330	256	572
2	0	0	65	43	81
3	0	0	23	16	45
4	0	0	11	6	24
5	0	0	14	13	18
6	0	0	17	9	10
7	0	0	9	4	11
8	0	0	11	6	8
9	0	0	9	7	13
10	0	0	1	3	6
11	0	0	5	2	5
12	0	0	5	4	5
13	0	0	3	4	13
14	0	0	3	2	7
15	0	0	6	5	10
16	0	0	9	5	11
17	0	0	6	5	14
18	0	0	68	52	111

Table E7. Parabolic (F_P^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$$\sigma^2 = 1,\ a = 2,\ b = 4,\ c = 4$$

Degrees of Freedom Numerator	Degrees of Freedom Denominator 1	2	3	4	5
1	0	0	122	137	203
2	0	0	59	65	103
3	0	0	17	30	60
4	0	0	24	26	35
5	0	0	18	20	39
6	0	0	19	23	27
7	0	0	8	19	30
8	0	0	10	9	16
9	0	0	10	8	20
10	0	0	8	6	22
11	0	0	9	7	21
12	0	0	11	14	23
13	0	0	6	11	12
14	0	0	8	10	14
15	0	0	10	12	19
16	0	0	10	9	17
17	0	0	11	16	25
18	0	0	145	173	214

Table E8. Triangular (F_P^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$$\sigma^2 = 1,\ a = 2,\ b = 4,\ c = 4$$

Degrees of Freedom Numerator	Degrees of Freedom Denominator				
	1	2	3	4	5
1	0	0	332	250	562
2	0	0	64	48	87
3	0	0	22	17	48
4	0	0	21	10	25
5	0	0	12	9	16
6	0	0	6	4	9
7	0	0	10	12	11
8	0	0	8	6	10
9	0	0	10	4	11
10	0	0	4	3	9
11	0	0	5	7	11
12	0	0	6	1	8
13	0	0	4	7	7
14	0	0	2	5	10
15	0	0	7	3	6
16	0	0	12	1	8
17	0	0	9	4	16
18	0	0	58	59	104

Table E9. Distribution of Degrees of Freedom (F_A^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$\sigma^2 = 1.0$, a = 2, b = 4, c = 4

Degrees Freedom Denominator	Degrees of Freedom Numerator = 1					
	Distribution					
	Normal	Uniform	Parabolic	Triangular	High Correlation	Moderate Correlation
1	163	165	644	170	24	99
2	152	138	489	131	50	117
3	492	477	439	498	315	456
4	412	445	252	416	593	457
5	499	485	108	496	333	458
6	280	288	22	287	684	409
7			6			
8			8			2
9			10			
10			4			1
11			6		1	1
12			5			
13			3			
14			2			
15			2			
16		1				

Table E10. Distribution of Degrees of Freedom (F_A^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$$\sigma^2 = 0.01, \; a = 2, \; b = 4, \; c = 4$$

Degrees of Freedom Denominator	Degrees of Freedom Numerator = 1 Distribution			
	Normal	Uniform	Parabolic	Triangular
1	14	17	117	18
2	24	25	83	20
3	187	182	380	186
4	824	820	545	828
5	110	102	99	97
6	841	842	770	840
7		1	1	1
8			2	
9			2	
10		1		
11				
12			1	

Table E11. Distribution of Degrees of Freedom (F_A^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

a = b = c = 5

Degrees of Freedom Denominator	Degrees of Freedom Numerator = 4			
	σ^2 = 1.0 Normal	σ^2 = 1.0 High Correlation	σ^2 = 1.0 Moderate Correlation	σ^2 = 1 and 3 Unequal Variance
15				3
16				7
17				7
18	2		2	28
19	14		3	42
20	30	2	13	86
21	46	4	24	121
22	82	24	44	179
23	134	32	76	216
24	193	69	121	287
25	266	85	188	351
26	373	112	264	342
27	416	132	397	236
28	343	174	460	86
29	98	231	313	9
30	3	327	91	
31		545	4	
32		263		

Table E12. Normal (F_p^*)

(Number of times that a particular degrees of freedom combination is calculated in 2000 trials)

$\sigma^2 = 1.0,\ a = 5,\ b = 5,\ c = 5$

Degrees of Freedom Numerator	Degrees of Freedom Denominator										
	21	22	23	24	25	26	27	28	29	30	31
1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
4	2	6	7	10	23	34	48	75	102	172	598
5	0	2	5	4	9	18	26	39	48	85	290
6	0	0	0	3	3	8	5	16	7	36	97
7	0	0	1	2	3	1	2	2	6	7	45
8	0	0	0	0	0	0	2	4	2	13	24
9	0	0	1	0	0	1	1	3	2	5	19
10	0	0	0	0	0	0	0	0	0	6	6
11	0	0	0	0	0	1	0	2	2	2	2
12	0	0	0	0	0	0	0	0	0	2	7
13	0	0	0	0	0	0	0	1	0	3	5
14	0	0	0	0	0	0	0	1	1	2	2
15	0	0	0	0	0	1	0	0	1	1	3
16	0	0	0	0	0	0	0	0	0	1	3
17	0	0	0	0	0	0	0	0	0	1	0
18	0	0	0	0	0	0	0	0	0	0	2
19	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	1
21	0	0	0	0	0	0	0	0	0	0	1
22	0	0	0	0	1	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	2
25	0	0	0	0	0	0	0	0	0	1	0
26	0	0	0	0	0	0	0	0	0	1	0
27	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	1	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	1
31	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	1	1

(Continued on Next Page)

Table E12. Normal (F_P^*) (Continued)

Degrees of Freedom Numerator	Degrees of Freedom Denominator										
	21	22	23	24	25	26	27	28	29	30	31
36	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	1
42	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	1	0
45	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	1
51	0	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0
59	0	0	0	1	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	1
62	0	0	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0
65	0	0	0	0	0	0	0	0	0	0	0
66	0	0	0	0	0	0	0	0	0	0	0
67	0	0	0	0	0	0	0	0	0	0	0
68	0	0	0	0	0	0	0	0	0	0	0
69	0	0	0	0	0	0	0	0	0	0	0
70	0	0	0	0	0	0	0	0	1	0	0

Table E13. Average Values of Degrees of Freedom-Unequal Variance

(Average = sum of the degrees of freedom divided by 2000 trials)

$\sigma^2 = 1.0$

		F_P^*	F_P^*	F_A^*	F_A^*	
Level a,b,c,	Value of Error	Average DF Num.	Average DF Denom.	Average DF Num.	Average DF Denom.	Average Value of Error
2,2,2	1.0	1.1043	1.7926	1.0000	1.7780	1.0
	.5 & 1.5	1.7852	1.4995	1.0000	.9328	1.0
	1 & 3	1.9328	1.4867	1.0000	.7903	2.0876
	1 & 5	2.0104	1.4349	1.0000	.7372	3.0773
3,3,3	1.0	2.0572	7.1309	2.0000	7.1007	1.0
	.6 & 1.8	4.1154	6.9120	2.0000	4.9048	1.0
	1 & 3	5.2980	6.8708	2.0000	3.7174	2.3349
	1 & 5	6.0304	6.8415	2.0000	3.1841	3.6707
4,4,4	1.0	3.0282	15.2309	3.0000	15.1759	1.0
	2/3 & 2	5.2554	15.7735	3.0000	12.3981	1.0
	1 & 3	6.0916	15.8986	3.0000	11.3729	1.5108
	1 & 5	6.7584	15.9574	3.0000	10.5243	2.0137
5,5,5	1.0	4.0156	30.4738	4.0000	30.4099	1.0
	5/7 & 15/7	5.8007	30.3687	4.0000	25.2026	1.0
	1 & 3	6.5005	30.4293	4.0000	23.6686	1.4134
	1 & 5	7.1930	30.3910	4.0000	22.2954	1.8176

Table E14. Average Values of Degrees of Freedom-Correlated Data

(Average = sum of degrees of freedom divided by 2000 trials)

$\sigma^2 = 1.0$

		F_P^*	F_P^*	F_A^*	F_A^*
Level a,b,c	Level Corre-lation	Average DF Num.	Average DF Denom.	Average DF Num.	Average DF Denom.
2,2,2	None	1.1043	1.7926	1.0000	1.7780
	Moderate	1.4968	1.6389	1.0000	1.0000
	High	1.2282	1.6544	1.0000	1.5394
3,3,3	None	2.0572	7.1309	2.0000	7.1007
	Moderate	3.5363	6.8833	2.0000	5.3753
	High	2.2802	7.0053	2.0000	6.7158
4,4,4	None	3.0282	15.2309	3.0000	15.1759
	Moderate	4.6978	15.7626	3.0000	13.1903
	High	3.1806	15.5395	3.0000	15.0110
5,5,5	None	4.0156	30.4738	4.0000	30.4099
	Moderate	5.2771	30.2698	4.0000	26.2877
	High	4.1648	29.4696	4.0000	28.6512

BIBLIOGRAPHY

1. R. A. Fisher, Statistical Methods for Research Workers, Oliver and Boyd Ltd., London, England, 1925.

2. F. E. Satterthwaite, "Synthesis of Variance," Psychometrika, Vol. 6, 1941, pp. 309-316.

3. E. S. Pearson, "The Analysis of Variance in Cases of Non-Normal Variation," Biometrika, Vol. 23, 1931, pp. 113-133.

4. G. E. P. Box, "Non-Normality and Tests on Variances," Biometrika, Vol. 40, 1953, pp. 318-335.

5. F. N. David and N. L. Johnson, "The Effects of Non-Normality on the Power Function of the F-Test in the Analysis of Variance," Biometrika, Vol. 38, 1951, pp. 43-57.

6. D. G. C. Grunow, "Test for Significance of the Difference Between Means in Two Normal Populations Having Unequal Variance," Biometrika, Vol. 38, 1951.

7. B. L. Welch, "The Significance of the Difference Between Means in Two Normal Populations Having Unequal Variances," Biometrika, Vol. 29, 1937, pp. 350-362.

8. G. E. P. Box, "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems: I. Effect of Inequality of Variance Problems: I. Effect of Inequality of Variance in the One Way Classification," Annals of Mathematical Statistics, Vol. 25, 1954a, pp. 290-302.

9. G. E. P. Box, "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems: II. Effect of Inequality of Variance and of Correlation of Errors in the Two Way Classification," Annals of Mathematical Statistics, Vol. 25, 1954b, pp. 484-498.

10. J. D. Hudson and R. G. Krutchkoff, "A Monte Carlo Investigation of the Size and Power of Tests Employing Satterthwaite's Synthetic Mean Squares," Biometrika, Vol. 55, 1968, pp. 431-433.

11. H. Scheffe, The Analysis of Variance, John Wiley and Sons, Inc., New York, 1959.

12. J. W. Schmidt and R. E. Taylor, Simulation and Analysis of Industrial Systems, Richard D. Irwin, Inc., Homewood, Illinois, 1970.

Other References

Bartlett, M. S., "The Use of Transformations," Biometrics, Vol. 3, No. 1, 1947, pp. 39-45.

Box, G. E. P., and Anderson, S. L., "Permutation Theory in the Derivation of Robust Criteria and the Study of Departures from Assumption," Journal of the Royal Statistical Society, Series B, Vol. 17, 1955, pp. 1-34.

Cochran, W. G., "Some Consequences When the Assumptions for the Analysis of Variance are not Satisfied," Biometrics, Vol. 3, No. 1, 1947, pp. 22-38.

Eisenhart, C., "The Assumptions Underlying the Analysis of Variance," Biometrics, Vol. 3, No. 1, 1947, pp. 1-21.

McCracken, D. D., A Guide to Fortran Programming, John Wiley and Sons, Inc. New York, 1966.

McMillan, C., and Gonzalez, R., Systems Analysis A Computer Approach to Decision Models, Richard D. Irwin, Inc., Homewood, Illinois, 1965.

Peng, K. C., The Design and Analysis of Scientific Experiments, Addison Wesley Publishing Company, Reading, Massachusetts, 1947.