1976

Linear Comparisons in Multivariate Analysis of Variance

Hsin-Ming Tzeng
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Part of the Applied Statistics Commons

Recommended Citation
Tzeng, Hsin-Ming, "Linear Comparisons in Multivariate Analysis of Variance" (1976). All Graduate Plan B and other Reports. 1181.
https://digitalcommons.usu.edu/gradreports/1181

This Report is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
LINEAR COMPARISONS IN MULTIVARIATE ANALYSIS OF VARIANCE

by

Hsin-Ming Tzeng

A report submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Applied Statistics

Plan B

Approved:

UTAH STATE UNIVERSITY

Logan, Utah

1976
# TABLE OF CONTENTS

| LIST OF FIGURES                      | ii   |
| LIST OF TABLES                      | iii  |

Chapter

I. INTRODUCTION                      | 1   |
II. MATHEMATICS OF MULTIVARIATE ANALYSIS OF VARIANCE | 4   |
   Mathematics of Multivariate Analysis of Variance | 4   |
   Linear Comparisons in Univariate Analysis of Variance | 12  |
III. PROCEDURES FOR THE MULTIVARIATE ANALYSIS OF VARIANCE | 17  |
   Description                             | 17  |
   The Two Way Analysis of Variance        | 17  |
   Linear Comparisons                      | 23  |
IV. ILLUSTRATIVE EXAMPLE OF MULTIVARIATE ANALYSIS OF VARIANCE | 26  |
   Description                             | 26  |
   Explanation of the Result of FCTCVR     | 30  |
   Data Analysis                           | 30  |
LITERATURE CITED                     | 41  |
APPENDICES                            | 42  |
   Appendix A. Complete Numerical Data of 72 Observations of Pederson's Alfalfa Seed Experiment | 43  |
   Appendix B. Input and Output of FCTCVR | 46  |
   Appendix C. Input and Output of MATRIXT | 53  |
VITA                                  | 58  |
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1.</td>
<td>Research situation for simple ANOVA design</td>
<td>1</td>
</tr>
<tr>
<td>1-2.</td>
<td>Research situation for a bivariate dependent variable</td>
<td>2</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Test for Wilks' $\Lambda$ for special cases</td>
<td>8</td>
</tr>
<tr>
<td>2.</td>
<td>Two way design classification</td>
<td>18</td>
</tr>
<tr>
<td>3.</td>
<td>Matrix elements for the two way analysis of variance</td>
<td>22</td>
</tr>
<tr>
<td>4.</td>
<td>Greatest root distribution parameters</td>
<td>23</td>
</tr>
<tr>
<td>5.</td>
<td>Pederson's data</td>
<td>28</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The analysis of variance was created by Ronald Fisher in 1923. It is most widely used and basically useful approach to study differences among treatment averages.

Most basic statistical books present analysis of variance methods for studies involving a single dependent variable measured on several samples that are suspected of arising from different populations. Figure 1-1 depicts the research situation of the simple ANOVA (analysis of variance) design.

![Figure 1-1. Research situation for simple ANOVA design.](image)

In this report we are concerned with the multivariate generalization of analysis of variance and specifically the use of multivariate linear comparisons to test particular relationships among treatments. The MANOVA (multivariate analysis of variance) is the study of group differences in a multidimensional measurement space, the distinctive multivariate nature
of MANOVA is that the dependent variable is assumed to be multivariate normal in distribution, with the same variance-covariance matrix for each population.

The research use of MANOVA is concerned with the "realness" of the difference among the population centroids, or the mean vectors, that is to say, the research objective is to study whether some or all of the populations are centered at different locations in the measurement space spanned by the dependent vector variable. Figure 1-2 depicts the research situation of the simple MANOVA design with these populations for a bivariate dependent variable.

Figure 1-2. Research situation for a bivariate dependent variable.

When we reject some hypothesis on the parameters of the multivariate linear model, it does not indicate which treatment or treatment
combinations are different and which could be considered as coming from
the same population. Sometimes, in the multivariate model we would
like to know which response or response-treatment combinations may have
led to the rejection, so in this study we also will discuss linear
comparisons.

In most experiments involving several treatments, the researcher
may be interested in certain specific comparisons among the treatment
means. To aid in making such comparisons, the statistician finds it is
convenient to discuss them in terms of "contrasts". A linear contrast
among means $m_1, m_2, \ldots, m_s$ is a linear combination
$$c_1 m_1 + c_2 m_2 + \ldots + c_s m_s$$
such that
$$c_1 + c_2 + \ldots + c_s = 0$$

Usually we have two kinds of contrasts, nonorthogonal contrasts
and orthogonal contrasts. A contrast is orthogonal to another contrast
if the inner product of the coefficients is zero. While orthogonal
contrasts are independent, they may not be testing sensible hypotheses.
A set of contrasts which make sense should be chosen, rather than an
orthogonal set for the sense of orthogonality. The nature of a linear
comparison in multivariate analysis of variance is that $m_1, m_2, \ldots, m_s$
are mean vectors consisting of several variables.
CHAPTER II

MATHEMATICS OF MULTIVARIATE ANALYSIS OF VARIANCE

Mathematics of multivariate analysis of variance

The linear model of MANOVA may be written as:

$$X_{ki} = M + (A_k - M) + (X_{ki} - A_k).$$

where $X_{ki}$ is the dependent vector variable for the $i$th subject in the $k$th sample, $k = 1, 2, \ldots, g$ where $g$ is the number of populations under study, $M$ is the grand centroid, or vector of total sample means, $A_k$ is the centroid for sample $k$, however we usually are not interested in the grand centroid, and it is conventional to subtract it from both sides of the equation, casting the linear model in terms of individual subjects' deviations from the grand centroid.

Let $x_{ki} = X_{ki} - M$, then $x_{ki} = (A_k - m) + (X_{ki} - A_k)$.

In this partition of the deviation of the $i$th subject in group $k$ from the grand centroid, the first term represents the hypothesis effect (difference in locations of groups) and the second term represents the error effect (deviations of subjects from their sample centroids).

$$
\sum_{k=1}^{g} \sum_{i=1}^{N_k} x_{ki}'x_{ki} = \sum_{k=1}^{g} \sum_{i=1}^{N_k} (A_k - M) (A_k - M) + \sum_{k=1}^{g} \sum_{i=1}^{N_k} (X_{ki} - A_k)
$$

By the fundamental partition sum of squares we know that the term on the left of the equal sign is the matrix of sums of squares and cross
products of all subjects from the grand centroid. Call this matrix \( T \) for "Total,"

\[
T = \sum_{k=1}^{g} \sum_{i=1}^{N_k} x_{ki}x_{ki}'.
\]

The first partition term is the matrix of weighted squares and products of deviations of group centroids from the grand centroid which we'll call \( A \) for "among samples,"

\[
A = \sum_{k=1}^{g} \sum_{i=1}^{N_k} (A_k - M)(A_k - M)'.
\]

\[
= \sum_{k=1}^{g} N_k (A_k - M)(A_k - M)'.
\]

The second partition term is the matrix of squares of subjects from their group centroids, pooled over all groups which we will call \( W \) for "within groups,"

\[
W = \sum_{k=1}^{g} \sum_{i=1}^{N_k} (x_{ki} - A_k)(x_{ki} - A_k)'.
\]

Note that \( T = A + W \).

For example, suppose we would like to compare the final test scores of Math and English of two classes in a high school. From class 1 we choose 4 students, from class 2 we choose 5 students. Their final test scores are as follows:

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Math</th>
<th>English</th>
<th>Class 2</th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9</td>
<td>8</td>
<td></td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8</td>
<td></td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>9</td>
<td></td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>7</td>
<td></td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>7</td>
<td></td>
<td>9</td>
<td>7</td>
</tr>
</tbody>
</table>
Class 1  \( N_1 = 4 \)  \( A_1 = (9, 8) \)

Class 2  \( N_2 = 5 \)  \( A_2 = (8, 7) \)

Total  \( N = 9 \)  \( M = (8.4, 7.4) \)

\[ t_{11} = (9-8.4)(9-8.4) + \ldots + (9-8.4)(9-8.4) = 6.44 \]

\[ t_{12} = (9-8.4)(8-7.4) + \ldots + (9-8.4)(7-7.4) = -2.08 \]

\[ t_{22} = (8-7.4)(8-7.4) + \ldots + (7-7.4)(7-7.4) = 14.24 \]

\[ T = \begin{pmatrix} 6.44 & -2.08 \\ -2.08 & 14.24 \end{pmatrix} \]

\[ a_{11} = 4(9-8.4)^2 + 5(8-7.4)^2 = 3.24 \]

\[ a_{12} = 4(9-8.4)(8-7.4) + 5(8-8.4)(7-7.4) = 2.24 \]

\[ a_{22} = 4(8-8.4)^2 + 5(7-7.4)^2 = 1.44 \]

\[ A = \begin{pmatrix} 3.24 & 2.24 \\ 2.24 & 1.44 \end{pmatrix} \]

\[ W = T - A \]

\[ W = \begin{pmatrix} 3.20 & -4.32 \\ -4.32 & 12.80 \end{pmatrix} \]

Each of the partition's mean sum of squares is an independent estimator of the common population variance. When the null hypothesis holds, the estimator based on the group mean vectors is

\[ D_A = \frac{A}{g-1} \]

Let \( N = N_k \)

The estimator based on the pooled within groups deviations is

\[ D_W = \frac{W}{N-g} \]
The null hypothesis is that \( u_k = u \) for \( k = 1, 2, \ldots, g \) and if it holds the best estimator of the common population centroid \( u \) is of course \( M \), the grand centroid

\[
M = \frac{1}{N} \sum_{k=1}^{g} \sum_{i=1}^{N_k} x_{ki}.
\]

When the null hypothesis is rejected the estimator of the research hypothesis effects or the treatment effects is the matrix of deviations of group means from the grand means, each column of which is formed as \( A_k - M \). Let \( H_0 : u_k = u \), then \( \Lambda = \frac{|W|}{|T|} \) was formed by Wilks in 1932 in terms of the distribution of a ratio of determinants as a test of the effects of the treatments represented by the groups. The distribution of \( \Lambda \) is a family of three parameter curves, with parameters based on the number of groups, the number of subjects, and the number of elements in the vector variable. The general utility of the statistic depends on the availability of its distribution or approximation thereto. Rao derived a \( F \) approximation in 1952. For Rao's \( F \) approximation it is necessary to compute a set of functions of the design parameters \( P \) (number of variables), \( g \) (number of groups) and \( N \) (total number of subjects, in all groups).

\[
S = \sqrt{\frac{p^2 (g-1)^2 - 4}{p^2 + (g+1)^2 - 5}}
\]

\[
n_1 = px(g-1)
\]

\[
n_2 = S \left[ (N-1) - \frac{p + (g-1) + 1}{2} \right] - \frac{p(g-1)-2}{2}
\]

Now let \( Y = \Lambda^{1/S} \)

then we get \( F = \left( \frac{1 - Y}{Y} \right) \left( \frac{n_2}{n_1} \right) \)
Where $n_1$ is the degree of freedom associated with the numerator and $n_2$ the denominator of this variance ratio. For certain small values of $p$ or $q$ it is also possible to transform $\Lambda$ into an exact $F$ variate. The following table summarizes this exact $F$ test for special cases in terms of $p$ and $q$ were $q = g - 1$.

Table 1. Test for Wilks' $\Lambda$ for special cases.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$p$</th>
<th>$q$</th>
<th>$F$-ratio</th>
<th>$n_1$</th>
<th>$n_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>any</td>
<td>1</td>
<td>$\frac{1}{\Lambda}$</td>
<td>$(N - p - 1)$</td>
<td>$p$</td>
<td>$N - p - 1$</td>
</tr>
<tr>
<td>any</td>
<td>2</td>
<td>$\frac{1}{\Lambda^2}$</td>
<td>$(N - p - 2)$</td>
<td>$2p$</td>
<td>$2(N - p - 2)$</td>
</tr>
<tr>
<td>1</td>
<td>any</td>
<td>$\frac{1}{\Lambda}$</td>
<td>$(N - q - 1)$</td>
<td>$q$</td>
<td>$N - q - 1$</td>
</tr>
<tr>
<td>2</td>
<td>any</td>
<td>$\frac{1}{\Lambda^2}$</td>
<td>$(N - q - 2)$</td>
<td>$2q$</td>
<td>$2(N - q - 2)$</td>
</tr>
</tbody>
</table>

Now, we describe a different approach. Suppose we have a linear model, say

$$x_{ih} = a_{i1}s_{1h} + \ldots + a_{im}s_{mh} + e_{ih}.$$  

The response of interest can be described by the $p$ dimensional multivariate normal random variable $X$. $N$ observation vectors $x_1, x_2, \ldots, x_n$ have been collected on $x$ under some experimental design. Each response in the $i$th vector has the same coefficients $a_{ij}$ so that the same design matrix $A$ holds for all dimensions. The residual variables $e_{i1}, e_{i2}, \ldots$.
e_{ip} of the ith observation are distributed according to the multinormal law with null mean vector and covariance matrix of full rank p. The model for all observations is

\[ X = AS + E, \]

where the N x p matrix X has the N observation vectors for rows and A is the appropriate design matrix.

\[
\begin{pmatrix}
S_{11} & \ldots & S_{1p} \\
\ldots & \ldots & \ldots \\
S_{q1} & \ldots & S_{qp}
\end{pmatrix}
\]

is the matrix of unknown parameters and the N x p matrix E contains the residual variates e_{ih}.

The multivariate extension of the general hypothesis is

\[ H_0^\prime : \text{CSM} = 0 \quad (2-1) \]

where C is the hypothesis matrix of dimension g x q and rank g \leq r. M permits the generation of hypotheses among the different response parameters and its dimension is p x u with rank u \leq p. M is usually taken as an appropriate identity matrix.

Now, let us sketch the derivation of the union-intersection test of the hypothesis (2-1). That multivariate hypothesis is true if and only if the univariate hypothesis

\[ H_0^\prime : \text{CSMa} = 0 \quad (2-2) \]

holds for all nonnull u-component vectors a. The test statistic for any one of these univariate hypothesis is

\[
F(a) = \frac{(N-r) a'M'X'A^{-1} C(C(A'A)^{-1}C')^{-1} C(A'A)^{-1}A'M'}{ga'M'X'(I-A(A'A)^{-1}A')X'Ma}
\]
For a univariate test of level \( \beta \) we would accept (2-2) at some other level \( \alpha \), say if \( \mathcal{A}(a) \leq F_{\beta}, \mathcal{N} \) for all nonnull \( a \). For if the greatest F ratio falls in the accept region, so must those of all other compound vectors. The maximum value of \( \mathcal{A}(a) \) can be shown to be proportional to the greatest root of the determinantal equation

\[
\| H - E \| = 0
\]

where

\[
H = M'X'\left(A'\right)^{-1}C \left[C\left(A'\right)^{-1}C\right]^{-1}C\left(A'\right)^{-1}A'XM
\]

\[
E = M'X' \left[\mathcal{I} - A\left(A'A\right)^{-1}A'\right] XM
\]

Within \( H \) and \( E \) both kinds of forms have the same coefficient matrix, from \( \mathcal{A}(a) \) we see that the F ratio can be generalized to the multivariate situation merely by replacing the F ratio sums of squares by their matrix extensions \( H \) and \( E \).

Let us denote the greatest root of (2-3) by \( c_s \). We accept the multivariate hypothesis (2-2) at the \( \alpha \) level if \( c_s \leq c(\alpha) \) and reject otherwise, \( c(\alpha) \) is the \( 100\alpha \) percentage point of the greatest root distribution when the hypothesis is true. Comprehensive charts of the upper percentage points of the greatest root distribution have been prepared by Heck. Heck chose to work with the distribution of the largest root of \( | H - \lambda (H+E) | = 0 \) rather than that of (2-3), and for that reason his charts must be entered with the test statistic

\[
\theta_s = \frac{c_s}{1 + c_s}
\]

If the null hypothesis is true, the parameters of the distribution of \( \theta_s \) are
\[ s = \min(g, u) \]
\[ m = \frac{|g - u| - 1}{2} \]
\[ n = \frac{N - r - u - 1}{2} \]

For an \( \alpha \)-level test of (2-1) the accept region is \( \theta_s \leq x_\alpha \), \( s, m, n \) where \( x \) is the upper 100\( \alpha \) percentage point obtained from the appropriate Heck charts. The zero roots of \( |H - \lambda E| = 0 \) are equal to the nonzero characteristic roots of \( HE^{-1} \).

Now, let us use the numerical data in page 5 to give an illustration for a union-intersection test. If \( g=2 \) and \( u=2 \) we have the following matrices:

Hypothesis matrix \( C = \begin{bmatrix} 1 & -1 \end{bmatrix} \)

Design matrix \( A, A' = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \)

Observation matrix \( X, X' = \begin{bmatrix} 9 & 10 & 8 & 9 & 8 & 6 & 8 & 9 & 9 \\ 8 & 8 & 9 & 7 & 5 & 8 & 9 & 6 & 7 \end{bmatrix} \)

\( M \) as a 2x2 identity matrix, \( M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \)

By using formula (2-4), we get \( H \) and \( E \) as follows

\[ H = \begin{bmatrix} 107.84 & 94.08 \\ 94.08 & 82.085 \end{bmatrix}, \quad E = \begin{bmatrix} -0.00689 & -0.05795 \\ -0.05795 & 0.02544 \end{bmatrix} \]

\[ HE^{-1} = \begin{bmatrix} -6.195 & -3.855 \\ -5.405 & -3.363 \end{bmatrix} \]

\[ \begin{vmatrix} -6.195-\lambda & -3.855 \\ -5.405 & -3.363-\lambda \end{vmatrix} = \lambda^2 + 9.558\lambda - 0.002, \]

since \( \lambda_1 = -19.1 \) and \( \lambda_2 = 0.0005 \)
then $c_s = 0.0005$

$$\theta_s = \frac{c_s}{1 + c_s} = \frac{0.0005}{1.0005} = 0.00049$$

$s = \min (g, u)$

$$= \min (2, 2) = 2$$

$m = \frac{|g - u| - 1}{2} = \frac{-1}{2}$

$$n = \frac{N - r - u - 1}{2} = \frac{9 - 2 - 2 - 1}{2} = 2$$

Here $r$ is the number of rank in design matrix.

$X^2_{0.01, 2, -1/2, 2} \leq 0.025$

Since $\theta_s \leq X_{0.01}$, so we accept the null hypothesis, that is $u_1 = u_2$.

If we apply Rao's $F$ approximation in this data, then we have $g=2$, $p=2$, and $s=1$ which gives $n_1=2$, $n_2=6$.

$$\Lambda = \frac{|W|}{|T|} = \frac{22.2976}{87.3792} = 0.255$$

$$Y = \Lambda^{1/2} = 0.255$$

$$F = \frac{1 - 0.255}{0.255} \cdot (\frac{6}{2}) = 8.76$$

$F_{0.01, 2, 6}=10.9$

We also accept the null hypothesis.

**Linear comparisons in univariate analysis of variance**

In most experiments involving several treatments, the researcher may be interested in certain specific comparisons among the treatment means. To aid in making such comparisons, the statistician finds it is
convenient to discuss them in terms of "contrast". Algebraically, a contrast among the quantities $T_1, T_2, \ldots, T_k$ (where $T_i$ is the sum of $n_i$ observations) is defined by

$$C_j = c_{1j}T_1 + c_{2j}T_2 + \ldots + c_{kj}T_k$$

Where $\Sigma n_i c_{ij} = 0$

In each $n_i = n$, that is, if each $T_i$ is the sum of the same number of observations, then the necessary condition for a contrast reduces to $\Sigma c_{ij} = 0$

Consider an experiment involving batteries in which four treatments are to be investigated. The four treatments happen to be four different electrolytes. However, it is noted that electrolytes No. 1 and No. 2 are quite similar in composition, that No. 3 and No. 4 are also similar, but that No. 1 and No. 2 differ considerably from No. 3 and No. 4. Then, it would be reasonable to plan comparisons of

(1) treatments 1 and 2 vs treatments 3 and 4
(2) treatment 1 vs treatment 2
(3) treatment 3 vs treatment 4

Assuming that 20 batteries are used and that they are allocated to the treatments in the ratio 4:2:5:9, denoting the treatment totals by $T_i$ ($i = 1, 2, 3, 4$), then the desired contrasts are:

$$C_1 = (7) (T_1 + T_2) + (-3) (T_3 + T_4)$$
$$= 7T_1 + 7T_2 - 3T_3 - 3T_4$$

$$C_2 = (1) T_1 + (-2) T_2 + (0) T_3 + (0) T_4$$

$$C_3 = (0) T_1 + (0) T_2 + (-9) T_3 + (-5) T_4$$
If the sample sizes are the same in each treatment then the three contrasts become

\[ C_1 = T_1 + T_2 - T_3 - T_4 \]
\[ C_2 = T_1 - T_2 \]
\[ C_3 = T_3 - T_4. \]

It may be verified that the sum of squares associated with a particular contrast is given by

\[ (C_j)_{yy} = \frac{C_j^2}{\sum_i n_i c_{ij}^2} \]
\[ = \frac{(\sum_i c_{ij} T_i)^2}{\sum_i n_i c_{ij}^2} \quad \text{(2-5)} \]

If each treatment total \( (T_i) \) is the sum of the same number of observations. Equation (2-5) simplifies to

\[ (C_j)_{yy} = \frac{C_j^2}{n \sum_i c_{ij}^2} \]
\[ = \frac{(\sum_i c_{ij} T_i)^2}{n \sum_i c_{ij}^2} \]

We can use the MS (mean square) error divided into the MS comparison to form a F test to test the hypothesis, that a contrast is equal to zero.

Let \( L_1 \) and \( L_2 \) be two linear functions of the \( n \) response measurements associated with an experiment \( y_1, y_2, \ldots, y_n \):

\[ L_1 = a_1 y_1 + a_2 y_2 + \ldots + a_n y_n \]
\[ L_2 = b_1 y_1 + b_2 y_2 + \ldots + b_n y_n \]

Then \( L_1 \) and \( L_2 \) are said to be orthogonal if \( \sum_i a_i b_i = 0 \), in other words,

\[ \text{Cov}(L_1, L_2) = 0. \]
Given \( t \) quantities \( T_i \), it is possible to form infinitely many sets of \( (t-1) \) mutually orthogonal contrasts. However, the experimenter must choose those contrasts which have reasonable interpretation and are relevant to specific questions. Consider quantities \( T_i \), for every \( m \) and for each \( c_{mi} \) of a set of orthogonal contrasts \( Q_m \), let

\[
a_{mi} = \sqrt{n_i} \frac{c_{mi}}{\sqrt{n_i}}.
\]

Then the matrix

\[
P = \begin{pmatrix}
\frac{n_1}{n_i} & \frac{n_2}{n_i} & \cdots & \frac{n_t}{n_i} \\
\frac{n_1}{n_i} & \frac{n_2}{n_i} & \cdots & \frac{n_t}{n_i} \\
& & & \\
& & & \\
& & & \\
& & & \\
& & & \\
\end{pmatrix}
\]

is orthogonal, since the sum of squares of the elements in any row equals 1 and the sum of products of the elements in any row with the corresponding elements in any other row equal zero. Then the orthogonal matrix \( P \) will transform

\[
(T_1/\sqrt{n_1}, T_2/\sqrt{n_2}, \ldots, T_t/\sqrt{n_t})
\]

into

\[
(\Sigma T_i/\sqrt{\Sigma n_i}, \Sigma a_{1i} T_i/\sqrt{n_i}, \ldots, \Sigma a_{(t-1)i} T_i/\sqrt{n_i})
\]

furthermore, we have

\[
\Sigma \left( \frac{T_i}{\sqrt{n_i}} \right)^2 = \left( \Sigma \frac{T_i}{\sqrt{\Sigma n_i}} \right)^2 + \left( \Sigma \frac{a_{1i} T_i}{\sqrt{n_i}} \right)^2 + \ldots + \left( \Sigma \frac{a_{(t-1)i} T_i}{\sqrt{n_i}} \right)^2
\]

and each term at the right of the equality has a single of degree of freedom. After subtracting the first term at the right from both
sides, the left side becomes the sum of squares for the t levels of a factor with (t-1) degrees of freedom which have been partitioned into components with a single degree of freedom each, corresponding to the orthogonal contrasts. We note the sums of squares

\[ SS_{Qm} = \left( \sum_i \frac{a_{mi}T_i}{\sqrt{n_i}} \right)^2 = \frac{\left( \sum_i \left( c_{mi}T_i/n_i \right) \right)^2}{\sum_i \left( c_{mi}^2/n_i \right)} \]

are distributed independently as \( \sigma^2 x^2(1) \) under the null hypothesis \( Q_m = 0 \) and the assumption that the \( y_i \) are distributed normally. If \( \sigma^2 \) is estimated by MSE with \( v \) degrees of freedom, then, under the null hypothesis, \( SS_{Qm}/MSE \) is distributed as \( F(1,v) \). When all \( n_i \) are equal to say \( r \), we have

\[ a_{mi} = \frac{c_{mi}}{\sqrt{\sum_i c_{mi}^2}}, \quad SS_{Qm} = \frac{\left( \sum_i c_{mi}T_i \right)^2}{r \sum_i c_{mi}^2} \]

and the first row of the orthogonal matrix \( P \) becomes

\[ \frac{1}{\sqrt{t}}, \frac{1}{\sqrt{t}}, \ldots, \frac{1}{\sqrt{t}} \]

Orthogonal contrasts may be preferable if one wishes the estimates derived from the different contrasts to be uncorrelated. However, occasionally it is desirable to design an experiment with the expressed intent of analyzing a set of nonorthogonal contrasts. In such a case, the probability statements accompanying the associated tests of significance are of an ambiguous nature (due to the correlation between the contrasts), and much care should be exercised in interpreting the experimental results. One final remark that needs to be made is that only those contrasts which are meaningful to the researcher should be analyzed.
CHAPTER III

PROCEDURES FOR THE MULTIVARIATE ANALYSIS OF VARIANCE

Description

In this chapter we will discuss the procedure of the two factor experimental MANOVA with a completely random design. Consider a situation in which it is of interest to study the effect of two factors, A and B on some response. For example, in a chemical experiment we would like to simultaneously vary the reaction pressure and reaction time and study the effect of each on the yield. In a biological experiment it might be interesting in studying the drug effect on the loss of weight of a strain of mice in the first and second week. In each of these cases it is important not only to determine if the two factors have an influence on the response but also if there is a significant interaction between the two factors, so we form the following questions:

1. Are the A factor mean vectors equal?
2. Are the B factor mean vectors equal?
3. Does the A factor interact with the B factor to produce some effect on the result of A and B?

In this study, if A or B factor mean vectors are not equal then we also will check the linear comparisons to test the particular relationships among them.

The two way analysis of variance

Let $X_{ijk}$ denote the kth observation on the rth response obtained under the ith treatment of the first (or row) way of classification and
the jth treatment of the second (or column) set. Then if we let
\( X'_{ijk} = (x'_{ijkl}, \ldots, x'_{ijkp}) \), the data obtained from a two way
design with r rows and c columns can be arranged as follows

Table 2. Two way design classification.

<table>
<thead>
<tr>
<th>row treatment</th>
<th>column treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>:</td>
</tr>
<tr>
<td>1</td>
<td>( x'_1l1 )</td>
</tr>
<tr>
<td></td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>( x'_1l2 )</td>
</tr>
<tr>
<td>2</td>
<td>( x'_2l1 )</td>
</tr>
<tr>
<td></td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>( x'_2l2 )</td>
</tr>
<tr>
<td></td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>( x'_r1l1 )</td>
</tr>
<tr>
<td></td>
<td>:</td>
</tr>
<tr>
<td>r</td>
<td>( x'_rl1 )</td>
</tr>
</tbody>
</table>

The linear model for the general observation is

\[ X'_{ijkh} = u + a_{ih} + b_{jh} + (ab)_{ijh} + e_{ijkh}. \]

- **u** = general level parameter of the response.
- **a** = effect of the ith row treatment on the response.
- **b** = effect of the jth column treatment on the hth response.
- **ab** = effect of interaction of ith and jth treatment on hth response.
The parameter matrix $S$ can be written as
$$
S = \begin{pmatrix}
\begin{array}{cccc}
 a_{11} & a_{r1} & b_{11} & b_{c1} & a_{111} & a_{1c1} & a_{r11} & a_{rc1} & u_1 \\
 a_{12} & a_{r2} & b_{12} & b_{c2} & a_{112} & a_{1c2} & a_{r12} & a_{rc2} & u_2 \\
 . & . & . & . & . & . & . & . & . \\
 a_{1p} & a_{rp} & b_{1p} & b_{cp} & a_{11p} & a_{1cp} & a_{r1p} & a_{rcp} & u_p
\end{array}
\end{pmatrix}
$$

The design matrix can be expressed in partitioned form as
$$
A = \begin{pmatrix}
\begin{array}{cccc}
 j & . & j0j & . & . & . & . & . & 0j \\
\end{array}
\end{pmatrix}
\begin{pmatrix}
1 \\
1 \\
1_{nx1}
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
0_{nx1}
\end{pmatrix}
$$

When we want to test the row effect, column effect or interaction effect, the hypotheses are as follows:

$H'_0 :$ some or all of the row treatment mean vector are not equal.

$H'_1 :$ some or all of the column treatment mean vector are not equal.
\( H'''' : \begin{pmatrix} (ab)_{111} \\ \vdots \\ (ab)_{11p} \end{pmatrix} = \begin{pmatrix} (ab)_{rcl} \\ \vdots \\ (ab)_{rcp} \end{pmatrix} \)

\( H''' : \) some or all of the interaction effect vector are not equal.

So the hypothesis matrix is

\[
C_1 \text{ for testing } H'''
\begin{pmatrix}
1 & \ldots & 0 & \ldots & 0 & \ldots & -1 & \ldots & 0 \\
0 & \ldots & 1 & \ldots & 0 & \ldots & -1 & \ldots & 0 \\
& & \cdots & & \cdots & & \cdots & & \cdots \\
0 & \ldots & 0 & \ldots & 1 & \ldots & -1 & \ldots & 0
\end{pmatrix}
\]

\[
C_2 \text{ for testing } H'''
\begin{pmatrix}
0 & \ldots & 0 & \ldots & 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & -1 & \ldots & 0 \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 1 & \ldots & 0 & \ldots & 0 & \ldots & -1 & \ldots & 0 \\
& & \cdots & & \cdots & & \cdots & & \cdots & & \cdots \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 & \ldots & -1 & \ldots & 0
\end{pmatrix}
\]

here \( c \) is equal to

\[
c \quad \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & -1 \\
0 & \ldots & 1 & \ldots & 0 & \ldots & 0 & \ldots & -1 \\
& & \cdots & & \cdots & & \cdots & & \cdots \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 1 & \ldots & -1 \end{pmatrix}
\]

so the null hypothesis can be written as

\[
H' = C_1 S = 0 \\
H'' = C_2 S = 0 \\
H''' = C_3 S = 0
\]
From the preceding matrices and the general expressions of the last section, the hypothesis and error matrix can be calculated for the two way multivariate analysis of variance. However, it is far easier to obtain those quantities as the matrix generalizations of the sum of squares of the univariate analysis. We define the following totals of the observations on each response.

1. cell : \[ C_{ijh} = \sum_{k=1}^{h} X_{ijkh} \]

2. column treatment : \[ T_{jh} = \sum_{i=1}^{r} \sum_{k=1}^{n} X_{ijkh} \]

3. row treatment : \[ R_{ih} = \sum_{j=1}^{c} \sum_{k=1}^{n} X_{ijkh} \]

4. grand total : \[ G_{h} = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} X_{ijkh} \]

These totals are computed for all combinations of the treatment and response subscripts, so the matrices of error and hypothesis sums of squares are computed as follows in Table 3.

For the tests of the three hypotheses it is first necessary to invert the \( E \) (error) matrix and then to form the products \( AE^{-1} \), \( BE^{-1} \), \( ABE^{-1} \). From these matrices and greatest characteristic roots \( c_{1s} \), \( c_{2s} \), \( c_{3s} \) are extracted and the statistics \( c_{is}/(1+c_{is}) = \theta_{s} \) is referred to the Heck charts described in section 2-1. From the Heck charts we
Table 3. Matrix elements for the two way analysis of variance.

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SUM OF SQUARES AND PRODUCTS</th>
<th>GENERAL ELEMENTS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MATRIX</td>
<td>AE^{-1}</td>
<td></td>
</tr>
<tr>
<td>row treatment</td>
<td>A</td>
<td>$a_{uv} = \frac{1}{cn} \sum_{i=1}^{r} R_{iu} R_{iv} - \frac{G G_v}{rcn}$</td>
<td></td>
</tr>
<tr>
<td>column</td>
<td>B</td>
<td>$b_{uv} = \frac{1}{rn} \sum_{j=1}^{c} T_{ju} T_{jv} - \frac{G G_v}{rcn}$</td>
<td>BE^{-1}</td>
</tr>
<tr>
<td>interaction</td>
<td>AB</td>
<td>$(ab)<em>{uv} = t</em>{uv} - a_{uv} - b_{uv} - e_{uv}$</td>
<td>ABE^{-1}</td>
</tr>
<tr>
<td>error</td>
<td>E</td>
<td>$\frac{1}{n} \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} x_{ijku} x_{ijkv}$</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>T</td>
<td>$G_{uv}$</td>
<td>$\frac{G_{UV}}{rcn}$</td>
</tr>
</tbody>
</table>
can get the values of \( x_{\alpha,s,m,n} \). If \( x_{\alpha} > \theta \) we reject the hypothesis, otherwise accept it.

The parameters of the relevant greatest characteristic root distributions are summarized in Table 4.

Table 4. Greatest root distribution parameters.

<table>
<thead>
<tr>
<th>Source</th>
<th>Statistic</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>rows</td>
<td>( c_{1s} ) ( 1 + c_{1s} ) ( \min(r-1,p) ) ( \frac{r-1-p-1}{2} ) ( \frac{\text{rc}(n-1)-p-1}{2} )</td>
<td></td>
</tr>
<tr>
<td>columns</td>
<td>( c_{2s} ) ( 1 + c_{2s} ) ( \min(c-1,p) ) ( \frac{c-1-p-1}{2} ) ( \frac{\text{rc}(n-1)-p-1}{2} )</td>
<td></td>
</tr>
<tr>
<td>interactions</td>
<td>( c_{3s} ) ( 1 + c_{3s} ) ( \min(r-1)(c-1),p ) ( \frac{(r-1)(c-1)-p-1}{2} ) ( \frac{\text{rc}(n-1)-p-1}{2} )</td>
<td></td>
</tr>
</tbody>
</table>

Linear comparisons

Rejection of some hypothesis on the multivariate linear model does not indicate which treatment or treatment combinations are different. In the multivariate model we would like to determine whether the linear combinations among several treatments are equal to zero or not. We can use orthogonal contrasts to do the analysis, but it should not override the nature of the design. Suppose we have two factors each at three levels and the vector variable has three elements, then we can get the individual averages and total values for each element just as what we do in the univariate case. Suppose we like to compare the three levels of factor A and we know that level 1 is a new kind and level 2 and 3
are old ones, then it is reasonable to compare level 1 versus level 2 and 3. We can list the raw data as follows:

Factor A, level 1 vs level 2 and 3.

<table>
<thead>
<tr>
<th>level</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>vector</td>
<td>$x_1$</td>
<td>$x_2$</td>
<td>$x_3$</td>
</tr>
<tr>
<td>average</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>total</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Coefficient 2 -1 -1

the contrasts are

$Q_{x_i} = \sum c_i x_i$

$Q_{xixj} = Q_{x_1} Q_{x_2}$

let

$y_i^2 = \frac{Q_{x_i}^2}{r E c_i^2}$

$y_i y_j = \frac{Q_{xixj}}{r E c_i^2}$, here $r$ is the number of observations and we assume all the sample sizes are the same.

If we put them in matrix form then we get the matrix of sums of squares and products for the multivariate linear comparison, in this way we can get any linear comparison matrix. For the three element cases we have:

$$C = \begin{pmatrix}
  y_1^2 & y_1 y_2 & y_1 y_3 \\
  y_2^2 & y_2 y_3 \\
  y_3^2 \\
\end{pmatrix}$$

1 vs 2 and 3
to form the F test we just need to multiply C by the inverse of the error matrix (E) and do the same steps as in the test of main effects.
ILLUSTRATIVE EXAMPLE OF MULTIVARIATE ANALYSIS OF VARIANCE

Description

The raw data adopted in this chapter is taken from the F field of Pederson's 1960 alfalfa seed experiment, it includes six treatments, three varieties and four replications. In this illustrative example I chose three variables to study the relationships among the multivariate data. The three variables are:

\( x_1 = \) blotch.
\( x_2 = \) chaff yield sample A ton/acre.
\( x_3 = \) seed weight per acre.

The mathematical model for each variable is:

\[ X_{ijk} = u + a_i + B_j + ab_{ij} + e_{ijk} \]

where:

\( a_i = \) effect of treatment. \( i = 1, 2, 3, 4, 5, 6 \)
\( b_j = \) effect of varieties. \( j = 1, 2, 3 \)
\( c_k = \) effect of replications. \( k = 1, 2, 3, 4 \)
\( ab_{ij} = \) effect of interactions of treatment and variety.
\( e_{ijk} = \) error of measurements of variables.

Those interactions among treatments and replications and those among varieties and replications have no interest and for ease of explaining the data, I supposed that there were no interactions among them. The hypotheses are:
$H_1$: There are no differences among the six treatment effects.
\[
\begin{pmatrix}
a_{11} \\
a_{12} \\
a_{13}
\end{pmatrix}
= \ldots =
\begin{pmatrix}
a_{61} \\
a_{62} \\
a_{63}
\end{pmatrix}
\]

$H_2$: There are no differences among the three variety effects.
\[
\begin{pmatrix}
b_{11} \\
b_{12} \\
b_{13}
\end{pmatrix}
= \ldots =
\begin{pmatrix}
b_{31} \\
b_{32} \\
b_{33}
\end{pmatrix}
\]

$H_3$: There are no differences among the interactions of treatments and varieties.
\[
\begin{pmatrix}
ab_{11} \\
ab_{12} \\
ab_{13} \\
\vdots \\
ab_{61} \\
\vdots \\
ab_{63}
\end{pmatrix}x_1 = \ldots =
\begin{pmatrix}
ab_{11} \\
ab_{12} \\
ab_{13} \\
\vdots \\
ab_{61} \\
\vdots \\
ab_{63}
\end{pmatrix}x_3
\]

After the test of main effects and interaction effects we make multivariate linear comparisons. First compare the effect among the varieties then the treatments.

The required computations needed for explaining this example are produced by the computer statistical package, STATPAC/FCTCVR and STATPAC/MATRIXT. The computer printout from these programs is included in appendices B and C. These two programs were written by Rex L. Hurst at Utah State University. Table 5 is the data form of Pederson's data.
Table 5. Pederson's data.

<table>
<thead>
<tr>
<th>PEDERSON</th>
<th>Pederson Alfalfa Seed</th>
<th>Field 1960</th>
<th>(72 Records)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1-2)</td>
<td>Treat</td>
<td>(01-06)</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>Rep</td>
<td>(1-4)</td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>Variety</td>
<td>(1-3)</td>
<td>1=Ranger, 2=Lahontan, 3=Syn C</td>
</tr>
<tr>
<td>(5)</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(6-8)</td>
<td>Soil Moisture Atm</td>
<td>XX+X</td>
<td>May 13</td>
</tr>
<tr>
<td>(9-11)</td>
<td></td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>(12-14)</td>
<td></td>
<td></td>
<td>June 6</td>
</tr>
<tr>
<td>(15-17)</td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>(18-20)</td>
<td></td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>(21-23)</td>
<td></td>
<td></td>
<td>July 5</td>
</tr>
<tr>
<td>(24-25)</td>
<td>% Bloom June 13</td>
<td>XX†</td>
<td></td>
</tr>
<tr>
<td>(26-27)</td>
<td>% Bloom June 21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(28)</td>
<td>Blotch</td>
<td>(1-9)</td>
<td></td>
</tr>
<tr>
<td>(29)</td>
<td>Frost Damage June 21</td>
<td>(1-4)</td>
<td></td>
</tr>
<tr>
<td>(30-31)</td>
<td>Tripped Flower per Raceme</td>
<td>X+X</td>
<td></td>
</tr>
<tr>
<td>(32-34)</td>
<td>Racemes/stem</td>
<td></td>
<td>XXXX</td>
</tr>
<tr>
<td>(35-38)</td>
<td>Stems/Acre (thousands)</td>
<td></td>
<td>XXX+X</td>
</tr>
<tr>
<td>(39-41)</td>
<td>Racemes/Acre (millions)</td>
<td></td>
<td>XXX+X</td>
</tr>
<tr>
<td>(42-45)</td>
<td>Seed/Acre Sample A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(46-49)</td>
<td>Seed/Acre Sample B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(50-52)</td>
<td>Chaff Yield Sample A Ton/Acre</td>
<td>X+XX</td>
<td></td>
</tr>
<tr>
<td>(53-55)</td>
<td>Chaff Yield Sample B</td>
<td></td>
<td>X+XX</td>
</tr>
<tr>
<td>(56-59)</td>
<td>Seed Weight Per Acre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(60-62)</td>
<td>Weight/Seed mg</td>
<td>X+XX</td>
<td></td>
</tr>
<tr>
<td>(63-65)</td>
<td>Seed/Pods</td>
<td>X+XX</td>
<td></td>
</tr>
<tr>
<td>(66-68)</td>
<td>Flowers/Raceme</td>
<td>XX+X</td>
<td></td>
</tr>
<tr>
<td>(69-71)</td>
<td>Pods/Raceme</td>
<td>X+XX</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. (Continued)

<table>
<thead>
<tr>
<th>Period</th>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(72-75)</td>
<td>Flowers/Acre Millions</td>
<td>XXX+X</td>
</tr>
<tr>
<td>(76-78)</td>
<td>Sugar Conc.</td>
<td>XX+X</td>
</tr>
<tr>
<td>(79-80)</td>
<td>Bees/Sq. Yd.</td>
<td>X+X</td>
</tr>
</tbody>
</table>
Explanation of the result of FCTCVR

The FCTCVR computer printout is in appendix B. The first 72 numbers are the values of 72 observations of $x_1, x_2, x_3$.

The numerical data under line 2 represents the correction terms $Y_{.../abc}$ (average of all observations for variables $x_1, x_2, x_3$ respectively), line number 3 represents the average values of the six levels of factor A (treatment 1 to 6 for all varieties and replications), as $Y_{A1../bc}, \ldots, Y_{A6../bc}$. Line number four's represent the average values of the three levels of varieties for all treatments and replications as $Y_{B1../ac}, Y_{B2../ac}, Y_{B3../ac}$. Line number sixes are the values of 18 interactions among the six treatments and three varieties.

By using the linear combination function cards we get the analysis of variance table for each variable. By using the linear function to combine the uncorrected matrices we get the corrected matrices, such as error matrix(E), treatment + error matrix (T+E), variety + error matrix (V+E) and treatment time variety + error matrix (TV+E), by using the STATPAC/MATRIXT we can get those matrices that we need, such as T, V, TV, and $E^{-1}$ etc.

Data analysis

By the rule of two way multivariate analysis of variance, there the first way of classification consisted of $r=3$ varieties and the second involved $c=6$ treatments. Under the original homogeneity of each variety assigned to each treatment. The total and error sum of squares and products matrices are:
\[ \begin{pmatrix} 634.653 & -37.3826 & -26072.9 \\ 46.1012 & 3578.5 \\ -2252500 \end{pmatrix} \]

**Total**

\[ \begin{pmatrix} 55.54167 & -4.000694 & -677.4722 \\ 24.20871 & 1064.709 \\ 518996.6 \end{pmatrix} \]

**E**

The hypotheses matrices for testing the treatment, variety and the interactions effects are:

\[ \begin{pmatrix} 80.56943 & -23.23098 & 1623.5000 \\ 16.15536 & 810.3570 \\ 225248.1 \end{pmatrix} \]

**T**

\[ \begin{pmatrix} 438.8611 & -10.77098 & -26626.79 \\ 0.6564700 & 622.0560 \\ 1618035 \end{pmatrix} \]

**V**

\[ \begin{pmatrix} 57.97223 & -0.2656950 & -481.1248 \\ 3.902680 & 1041.236 \\ 698360.4 \end{pmatrix} \]

**TxV**

The inverse of the error matrix is:

\[ \begin{pmatrix} 0.01840000 & 0.00218121 & 530.000 \\ 0.04566258 & -0.00009082844 \\ 0.000002138639 \end{pmatrix} \]
The matrix product for the treatment test is:

\[
\begin{bmatrix}
1.463535 & -1.032507 & 0.007156741 \\
-0.3763743 & 0.6134202 & -0.0001883251 \\
36.04215 & 20.08526 & 0.4398502
\end{bmatrix}
\]

The matrix for the test of variety is:

\[
\begin{bmatrix}
7.531164 & 2.883889 & -0.04738980 \\
-0.1845968 & -0.05001807 & 0.001060222 \\
-456.9537 & -176.6376 & 2.883506
\end{bmatrix}
\]

The matrix product for the interaction test is:

\[
\begin{bmatrix}
1.056707 & 0.1580173 & 0.000128174 \\
0.0239734 & 0.08305306 & 0.00186714 \\
7.067029 & -16.93490 & 1.389564
\end{bmatrix}
\]

By using the RTS operator we get the greatest characteristic root for the \( TxE^{-1} \) is 1.389567 which implies that:

\[
\theta_s = \frac{1.389567}{1 + 1.389567} = 0.5815141
\]

at \( \alpha = 0.05 \) level we check the Heck chart when the other parameters are:

\[ s = \min (10, 3) = 3 \]
\[ m = \frac{7-1}{2} = 3 \]
\[ n = \frac{8(4)-3-1}{2} = 25 \]

\[ x_{0.05, 3, 3, 25} = 0.74. \text{ Since } \theta_s < x_{0.05, 3, 3, 25}. \]

We accept \( H_3 \), that is there is no difference or no interactions between the treatments and the varieties.
We are now ready to test the main effects. First we test the treatments. The greatest characteristic root is 2.155076.

\[ \theta_s = \frac{2.155076}{1+2.155076} = 0.6830504 \]

\[ s = \min (c-1,p) = \min (6-1,3) = 3 \]

\[ m = \frac{|r-1-p|-1}{2} = \frac{1}{2} \]

\[ n = \frac{rc(n-1)-p-1}{2} = \frac{18(3)-3-1}{2} = 25 \]

\[ x^{0.05, 3, 1/2, 25} = 0.72 \]

\[ \theta_s < x^{0.05} \text{ So we accept } H_1 \text{ that is to say there are no differences among the six treatment effects.} \]

From the VE \(^{-1}\) matrix we get the maximum characteristic root 8.503845.

\[ \theta_s = \frac{8.503845}{1+8.503845} = 0.8947794 \]

\[ s = \min (3-1,3) = 2 \]

\[ m = \frac{|r-1-p|-1}{2} = \frac{3-1-3-1}{2} = 0 \]

\[ n = 25 \]

\[ x^{0.05, 2,0, 25} = 0.655 \]

\[ \theta_s > x^{0.05} \text{ so we reject the hypothesis. That is there exists some difference between the three varieties.} \]

Now we are going to consider the linear comparisons among the main effects. First we will compare the three varieties, because Ranger and Lahontan are two old varieties, Syn C is a new synthetic variety, it is of interest to compare Syn C versus Ranger and Lahontan. The numerical data obtained from FCTCVR is listed as follows:
Numerical data of FCTCVR.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ranger</th>
<th>Lahontan</th>
<th>Syn C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x_1$</td>
<td>$x_2$</td>
<td>$x_3$</td>
</tr>
<tr>
<td>Average</td>
<td>6.083333</td>
<td>2.536667</td>
<td>604.375</td>
</tr>
<tr>
<td></td>
<td>7.208333</td>
<td>2.686667</td>
<td>521.875</td>
</tr>
<tr>
<td>Total</td>
<td>145.99999</td>
<td>60.880008</td>
<td>14506</td>
</tr>
<tr>
<td>Coefficient</td>
<td>-1</td>
<td>-1</td>
<td>2</td>
</tr>
</tbody>
</table>
\[Q_{x_1} = \sum_{k=1}^{3} c_k x_{1k}\]
\[= (-1) 145.99999 + (-1) 172.99999 + (2) 36 = -246.99998\]
\[Q_{x_2} = \sum_{k=1}^{3} c_k x_{2k}\]
\[= (-1) 60.880008 + (-1) 64.480008 + (2) 66.40992 = 7.45997\]
\[Q_{x_3} = \sum_{k=1}^{3} c_k x_{3k}\]
\[= (-1) \times 14506 + (-1) \times 12525 + (2) \times 19512 = 11994\]

Here we have six treatments and four replications, so we have 6 x 4 = 24 observations for each variety.

Divisor = \[\sum_{i=1}^{3} c_i^2 = 24 \times 6 = 144\]

\[SS_{Qx_1} = \frac{Q_{x_1}^2}{\text{Divisor}} = \frac{(-246.99998)^2}{144} = 423.67354\]

\[SS_{Qx_2} = \frac{Q_{x_2}^2}{\text{Divisor}} = \frac{(7.45997)^2}{144} = 0.3864663\]

\[SS_{Qx_3} = \frac{Q_{x_3}^2}{\text{Divisor}} = \frac{(11994)^2}{144} = 999000.25\]

\[SS_{Qx_1Qx_2} = \frac{Q_{x_1Qx_2}}{\text{Divisor}} = \frac{(-246.99998)(7.45997)}{144} = -12.795919\]

\[SS_{Qx_1Qx_3} = \frac{Q_{x_1Qx_3}}{\text{Divisor}} = \frac{(-246.99998)(11994)}{144} = -20573.039\]

\[SS_{Qx_2Qx_3} = \frac{Q_{x_2Qx_3}}{\text{Divisor}} = \frac{(7.45997)(11994)}{144} = 621.35333\]
Thus, we get the SSSP matrix for the linear comparison as:

\[ C_V = \begin{pmatrix} 
423.67354 & -12.795919 & -20573.039 \\
0.3864663 & 621.35333 & 999000.25 \\
621.35333 & 999000.25 \\
\end{pmatrix} \]

The matrix products for extracting the greatest characteristic root to test the hypothesis: \( 2b_S - b_R - b_L = 0 \) is

\[ C_V E^{-1} = \begin{pmatrix} 
7.365608 & 2.208456 & -0.03455591 \\
-0.2224594 & -0.06670048 & 0.001043668 \\
-357.6644 & -107.2398 & 1.67990 \\
\end{pmatrix} \]

The greatest characteristic root of this matrix is 7.972428 which implies that:

\[ \theta_S = \frac{7.972428}{1 + 7.972428} = 0.8885481 \]

at \( \alpha = 0.05 \) and \( s = \min (r-1,p) = \min (3-1,3) = 2 \)

\[ m = \frac{|r-1-p|-1}{2} = 0 \]

\[ n = \frac{rc(n-1)-p-1}{2} = \frac{(24-1)-3-1}{2} = 32.5 \]

\( X^{0.05, 2, 0, 32.5} = 0.625 \)

\( \theta_S > x \), so we reject the hypothesis: \( H_0: 2b_S - b_R - b_L = 0 \) that is to say there is a difference between Syn C and Ranger and Lahontan.

After we know that there is a difference between Syn C versus Ranger and Lahontan we would like to know if there exists differences between Ranger and Lahontan making the hypothesis:
$H_0: \theta b_S + b_R - b_L = 0$. The comparable coefficients are $(0, 1, -1)$.

We can get $Q_{x1}$ to $Q_{x2x3}$ in the same method as in the last comparison.

The matrix of sums of squares and products is:

$$C_V = \begin{pmatrix} 15.1875 & 2.025 & -111.75 \\ 0.27 & -148.5 \\ 81675 \end{pmatrix}$$

Ranger vs Lahontan

The matrix of products for testing the hypothesis is:

$$C_V^{-1} = \begin{pmatrix} 0.2621002 & 0.2267545 & -0.002269016 \\ 0.0349469 & 0.03023393 & -0.0003025355 \\ -19.22068 & -16.62866 & 0.1663945 \end{pmatrix}$$

the greatest characteristic root of this matrix is $0.4008583$

$$\theta_s = \frac{0.4008583}{1 + 0.4008583} = 0.2861519$$

at $\alpha = 0.05$ and $s = \min (r-1, p) = \min (3-1,")= 2$

$$m = 0$$

$$n = 32.5$$

$$X_{0.05, 2, 0, 32.5} = 0.625$$

$\theta_s < x_2^2$, so we accept the null hypothesis, that is to say Ranger and Lahontan are not significantly different at $\alpha = 0.05$ level.

Now, we turn the test of the linear comparisons of treatments.

An interesting comparison is to compare the first three treatments versus the second three, since in Pederson's data the six treatments are arranged in two ways. The first way they were classified as solid, $24"$, $48"$ and the second way they were classified as not thin and thin.

We can display this as:
To compare solid versus spaced (24" and 48"), the data obtained from FCTCVR is listed as follows:

<table>
<thead>
<tr>
<th>Treatment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>01</td>
<td>02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>thin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The average and total values for each variable from treatment 1 to treatment 6 are listed on the following page.

The SSSP matrix for this comparison is:

\[
C_T = \begin{pmatrix}
0.694445 & -0.7597222 & -74.652902 \\
0.8311348 & 81.670152 & 8025.1881
\end{pmatrix}
\]

The matrix for testing the linear comparison is:

\[
C_T E^{-1} = \begin{pmatrix}
0.009661655 & -0.02639552 & -0.00007707918 \\
-0.01056984 & 0.02887665 & 0.00008432449 \\
-1.038269 & 2.837519 & 0.008286012
\end{pmatrix}
\]

The greatest characteristic root is 0.04735911, so the statistic is:

\[
\theta_s = \frac{0.04735911}{1 + 0.04735911} = 0.0452176
\]

at \( \alpha = 0.05 \) and \( s = \min (c-1, p) = \min (5, 3) = 3 \)

\[
m = \frac{1}{2}
\]

\[
n = \frac{rc(n-1)-p-1}{2} = \frac{6(12-1)-3-1}{2} = 31
\]
Numerical data of FCTCVR.

<table>
<thead>
<tr>
<th>treatment variable</th>
<th>00</th>
<th>01</th>
<th>02</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>4.166667</td>
<td>5.666667</td>
<td>2.640833</td>
</tr>
<tr>
<td>$x_2$</td>
<td>3.538333</td>
<td>732.9167</td>
<td>732.5833</td>
</tr>
<tr>
<td>$x_3$</td>
<td>732.9167</td>
<td>5.5</td>
<td>2.295</td>
</tr>
<tr>
<td>average</td>
<td>50.00004</td>
<td>68.00004</td>
<td>27.54</td>
</tr>
<tr>
<td>total</td>
<td>42.459996</td>
<td>31.68999</td>
<td>66</td>
</tr>
<tr>
<td>average</td>
<td>8795.0004</td>
<td>8790.999</td>
<td>7469.0004</td>
</tr>
<tr>
<td>total</td>
<td>68.00004</td>
<td>66</td>
<td>72.41667</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>treatment variable</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>622.4167</td>
<td>2.916667</td>
<td>2.916667</td>
</tr>
<tr>
<td>$x_2$</td>
<td>5.416667</td>
<td>2.4675</td>
<td>2.4675</td>
</tr>
<tr>
<td>$x_3$</td>
<td>2.0925</td>
<td>628.6667</td>
<td>692.5</td>
</tr>
<tr>
<td>average</td>
<td>622.4167</td>
<td>692.5</td>
<td>2.946667</td>
</tr>
<tr>
<td>total</td>
<td>25.11</td>
<td>7544.0004</td>
<td>7073.0004</td>
</tr>
<tr>
<td>average</td>
<td>71.00004</td>
<td>35.000004</td>
<td>35.000004</td>
</tr>
<tr>
<td>total</td>
<td>29.61</td>
<td>8310</td>
<td>7073.0004</td>
</tr>
</tbody>
</table>
\[ x_{0.05, 3, 1/2, 31} = 0.68 \]

\[ \theta_s < x, \text{ so we can accept the hypothesis that is to say that the linear comparison } -2b_1 + b_2 + b_3 -2b_4 + b_5 + b_6 \text{ is equal to zero, in other words, solid planting is not different than spaced planting.} \]

According to the above analysis, we can conclude that in Pederson's data there are no interactions between varieties and treatments. Three varieties show a difference between Syn C versus Ranger and Lahontan. There is no difference between Ranger and Lahontan. There are no differences among the six treatments. From the linear comparisons of the main effects their results show that they have the same conclusions as in the main effect tests when all tests are performed at 95 percent significance level. Of course, the coincidence is not necessary and not always to be true.
LITERATURE CITED


APPENDICES
Appendix A
Complete Numerical Data of 72 Observations of Pederson's Alfalfa Seed Experiment
0621 0160 0160 0160 0089 0153 3541 1384 3197 1910 9400 0616 2812 4805 4290 0311 3159 4102 8844 1041
0631 0206 0050 0040 0070 1223 4096 0863 2332 1197 0206 2771 2268 0573 2226 0178 4703 7494 1944
0641 0108 1100 3000 0400 5006 1030 3341 4723 3767 1271 0499 0054 1329 2336 0519 2102 2761 4629 3894 3764
0612 0105 0007 0040 0490 0050 1638 9416 0293 3121 1103 9509 2337 3950 4819 9141 4124 3702 6183 6529
0622 0105 0006 0030 0400 5060 1550 4429 0122 4470 9904 4804 0012 0722 3059 2215 2686 1732 9834 3840 052
0632 1010 0300 0010 0030 0430 0629 0431 4166 9730 7813 9576 0512 3553 0804 6819 6312 1253 2466 2411
0642 0101 2010 0300 0300 0701 1134 5134 6494 2260 1580 4230 0422 2621 2041 5810 9249 1644 1325 9437 638
0651 0070 1423 0330 0300 4100 2579 0142 1574 0364 9811 9412 8892 9205 9842 1434 9168 6421 5013 3604 2
0661 0406 3300 0300 0400 2040 1314 5463 4146 1181 2332 2305 0609 2093 4216 4352 2401 4003 38
0671 0709 0000 0300 0400 5010 3065 1415 2872 7151 4509 1203 3024 5232 2060 1187 3471 5326 5221 8414 34
0681 0404 0403 0300 0400 5010 3014 1879 4238 4205 0909 0813 3292 2580 7702 0442 7154 3653 1603 5344
Appendix B

Input and Output of FCTCVR
<p>| 1 | 0 | 0 | 6 | 3 | 4 | 1 | 1 | 1 | 1 | T2 | 4 | 5 | 4 | 0 | 0 | 0 | 5 | 1 | 0 | 0 |
| 1 | 5,000 | 4,030 | 554,0 | |
| 2 | 4,000 | 4,073 | 623,0 | |
| 3 | 5,000 | 4,190 | 564,0 | |
| 4 | 4,000 | 1,960 | 535,0 | |
| 5 | 7,000 | 2,080 | 596,0 | |
| 6 | 7,000 | 2,700 | 452,0 | |
| 7 | 5,000 | 4,140 | 461,0 | |
| 8 | 9,000 | 3,960 | 457,0 | |
| 9 | 1,000 | 3,420 | 1147 | |
| 10 | 1,000 | 3,230 | 1135 | |
| 11 | 1,000 | 4,100 | 1246 | |
| 12 | 1,000 | 4,270 | 1193 | |
| 13 | 7,000 | 2,640 | 775,0 | |
| 14 | 8,000 | 2,880 | 572,0 | |
| 15 | 8,000 | 1,660 | 637,0 | |
| 16 | 8,000 | 1,520 | 653,0 | |
| 17 | 7,000 | 2,080 | 511,0 | |
| 18 | 8,000 | 3,100 | 655,0 | |
| 19 | 9,000 | 2,260 | 535,0 | |
| 20 | 9,000 | 2,140 | 360,0 | |
| 21 | 1,000 | 2,320 | 757,0 | |
| 22 | 1,000 | 4,010 | 1163 | |
| 23 | 1,000 | 3,210 | 1065 | |
| 24 | 1,000 | 3,370 | 1105 | |
| 25 | 1,000 | 2,130 | 660,0 | |
| 26 | 8,000 | 2,630 | 535,0 | |
| 27 | 7,000 | 1,450 | 526,0 | |
| 28 | 5,000 | 2,570 | 650,0 | |
| 29 | 9,000 | 1,730 | 693,0 | |
| 30 | 8,000 | 3,000 | 535,0 | |
| 31 | 6,000 | 1,420 | 461,0 | |
| 32 | 9,000 | 4,520 | 636,0 | |
| 33 | 1,000 | 2,460 | 904,0 | |
| 34 | 1,000 | 2,040 | 609,0 | |
| 35 | 2,000 | 1,400 | 563,0 | |
| 36 | 3,000 | 1,340 | 757,0 | |
| 37 | 7,000 | 1,960 | 613,0 | |
| 38 | 6,000 | 2,410 | 655,0 | |
| 39 | 6,000 | 1,740 | 729,0 | |
| 40 | 7,000 | 1,360 | 581,0 | |
| 41 | 9,000 | 2,990 | 665,0 | |
| 42 | 8,000 | 2,220 | 525,0 | |
| 43 | 8,000 | 2,750 | 595,0 | |
| 44 | 9,000 | 1,350 | 526,0 | |
| 45 | 1,000 | 2,050 | 711,0 | |
| 46 | 1,000 | 2,760 | 603,0 | |
| 47 | 2,000 | 1,770 | 563,0 | |
| 48 | 3,000 | 1,670 | 572,0 | |
| 49 | 7,000 | 1,940 | 613,0 | |
| 50 | 6,000 | 2,770 | 637,0 | |
| 51 | 7,000 | 2,170 | 603,0 | |
| 52 | 8,000 | 2,490 | 748,0 | |
| 53 | 6,000 | 2,190 | 517,0 | |
| 54 | 9,000 | 2,190 | 492,0 | |
| 55 | 9,000 | 1,850 | 591,0 | |
| 56 | 7,000 | 3,220 | 406,0 | |
| 57 | 3,000 | 3,370 | 979,0 | |
| 58 | 4,000 | 2,730 | 784,0 | |</p>
<table>
<thead>
<tr>
<th>LINE</th>
<th>DF</th>
<th>UNCORRECTED SS</th>
<th>MS</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>72</td>
<td>2385.000</td>
<td>35957750E+08</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1750.347</td>
<td>1750.347</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>1830.917</td>
<td>305.16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2189.203</td>
<td>729.733</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1752.356</td>
<td>438.138</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>2327.750</td>
<td>130.43</td>
<td></td>
</tr>
</tbody>
</table>

**LINEAR FUNCTION CARDS**

<table>
<thead>
<tr>
<th>TOTAL</th>
<th>7</th>
<th>2</th>
<th>1</th>
<th>-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>REPLICATE</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>-2</td>
</tr>
<tr>
<td>TREAT(T)</td>
<td>9</td>
<td>2</td>
<td>3</td>
<td>-2</td>
</tr>
<tr>
<td>VARIETY(V)</td>
<td>10</td>
<td>2</td>
<td>4</td>
<td>-2</td>
</tr>
<tr>
<td>T*V</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>-3</td>
</tr>
<tr>
<td>ERROR</td>
<td>12</td>
<td>4</td>
<td>1</td>
<td>-6</td>
</tr>
</tbody>
</table>

**ANALYSIS OF VARIANCE**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>71</td>
<td>634.6528</td>
<td>0.938772</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>71</td>
<td>46.10123</td>
<td>0.6493131</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>71</td>
<td>3.07950</td>
<td>4.337575</td>
<td>3</td>
</tr>
<tr>
<td>REPLICATE</td>
<td>3</td>
<td>1.705133</td>
<td>0.5694444</td>
<td>1</td>
</tr>
<tr>
<td>REPLICATE</td>
<td>3</td>
<td>1.178015</td>
<td>0.3925718</td>
<td>2</td>
</tr>
<tr>
<td>REPLICATE</td>
<td>3</td>
<td>1.910344</td>
<td>0.6369648</td>
<td>3</td>
</tr>
<tr>
<td>TREAT(T)</td>
<td>5</td>
<td>63.35944</td>
<td>16.11389</td>
<td>1</td>
</tr>
<tr>
<td>TREAT(T)</td>
<td>5</td>
<td>16.15536</td>
<td>3.231071</td>
<td>2</td>
</tr>
<tr>
<td>TREAT(T)</td>
<td>5</td>
<td>225249.2</td>
<td>45047.63</td>
<td>3</td>
</tr>
<tr>
<td>VARIETY(V)</td>
<td>2</td>
<td>432.3551</td>
<td>219.1426</td>
<td>1</td>
</tr>
<tr>
<td>VARIETY(V)</td>
<td>2</td>
<td>0.6564654</td>
<td>0.3282347</td>
<td>2</td>
</tr>
<tr>
<td>VARIETY(V)</td>
<td>2</td>
<td>16190.35</td>
<td>890.1775</td>
<td>3</td>
</tr>
<tr>
<td>T*V</td>
<td>10</td>
<td>57.97222</td>
<td>5.797222</td>
<td>1</td>
</tr>
<tr>
<td>T*V</td>
<td>10</td>
<td>3.962681</td>
<td>0.392681</td>
<td>2</td>
</tr>
<tr>
<td>ERROR</td>
<td>51</td>
<td>553.54167</td>
<td>1.097052</td>
<td>1</td>
</tr>
<tr>
<td>ERROR</td>
<td>51</td>
<td>24.20871</td>
<td>0.4746506</td>
<td>2</td>
</tr>
<tr>
<td>ERROR</td>
<td>51</td>
<td>519996.6</td>
<td>10176.40</td>
<td>3</td>
</tr>
</tbody>
</table>
UNCORRECTED MATRICES

LINE OF ROW SS AND SP
1 72
1 2385.000 908.1300 21050.0
2 556.8753 13137.7
3 * 35054.75E+06
2 11 1750.347 945.5326 236877.9
2 510.7741 127798.7
3 * 3197600E+06
3 6 1 1830.917 922.3017 278201.4
2 526.9294 128609.1
3 * 3220125E+06
4 3 1 2189.209 934.7617 209951.1
2 511.4353 128420.6
3 * 3359442E+06
5 4 1 1752.356 946.4183 236656.9
2 511.9521 127638.9
3 * 3199511E+06
6 18 1 2327.750 911.2650 211893.5
2 531.4886 133272.4
3 * 3451765E+06

CORRECTED MATRICES

LINE = -1 DF = 31
Error
1 55.54167 -4.000694 -677.4722
2 24.20471 1064.709
3 3189.768
LINE = 2 DF = 41
Treatment
1 136.1111 27.23167 946.0278
2 40.36407 1375.066
3 744.244
+ Error
3 4 1.1 5 1. 3-1. 6
LINE = 3 DF = 53
Variance + Error
1 494.4028 -14.77157 27304.26
2 24.86518 1636.765
3 2137.032
LINE = 4 DF = 61
Treatment + Error
1 115.0149 4.266369 1158.597
2 28.11139 2155.945
3 1217357
3 0 1 1 1
1 2 3

LINE NUMBER
1

INVERSE ROW COL ELEMENT
-1 1 1 1849090E-01
<table>
<thead>
<tr>
<th>REG COEF</th>
<th>X</th>
<th>Y</th>
<th>COEFFICIENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>1</td>
<td>2</td>
<td>-2181215E-02</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>3</td>
<td>1954171E-04</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
<td>2</td>
<td>4546256E-01</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
<td>3</td>
<td>-9352644E-04</td>
</tr>
<tr>
<td>-1</td>
<td>3</td>
<td>3</td>
<td>2138839E-05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DUE TO REG</th>
<th>DF</th>
<th>ADW</th>
<th>CCL</th>
<th>SS AND SP</th>
<th>WS AND HP</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEV FR REG</th>
<th>DF</th>
<th>ADW</th>
<th>CCL</th>
<th>SS AND SP</th>
<th>WS AND HP</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TRT ADJ</th>
<th>DF</th>
<th>ADW</th>
<th>CCL</th>
<th>SS AND SP</th>
<th>WS AND HP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Input and Output of MATRIXT
E = STATPACK/KRATAT
#RUNNING 1469
ENTER OPERATION CODE

1

ENTER (AREA#, MATRIX-TYPE(1=SYN, 2=ASYM))

MATRICES

ENTER ORDER

ENTER MATRIX ELEMENTS, ROWWISE, UPPER TRIANGULAR ONLY

55.54167, -4.66934, -677.4728
24.28571, 1384.789
516796.6

ENTER OPERATION CODE

2

ENTER (AREA#, MATRIX-TYPE(1=SYN, 2=ASYM))

MATRIX OUTPUT

1  83.56943  -23.23438  1623.588
2  16.15536  810.3578
3  225840.1

ENTER OPERATION CODE

3

ENTER (FROM-AREA, DISK#)

ENTER OPERATION CODE

3

ENTER (AREA#, MATRIX-TYPE(1=SYN, 2=ASYM))

ENTER ORDER

3

ENTER MATRIX ELEMENTS, ROWWISE, UPPER TRIANGULAR ONLY

494.4929, -14.7715, -27394.26
24.38515, 1686.795
2137832

ENTER OPERATION CODE

SUB

ENTER (PRE-AREA, POST-AREA, DIFF-AREA)

ENTER OPERATION CODE

3

ENTER AREA#
MATRX OUT,

1 66.56943 -23.23098 1623.518
2 16.15536 618.3573
3 425840.4

ENTER OPERATION CODE

KLT

ENTER (PRE-AREA, POST-AREA, PROD-AREA)
2,1,3

ENTER OPERATION CODE

VT

ENTER AREA#
3

MATRX OUT
1 1.463595 -1.635297 0.715670E-52
2 -0.3763742 0.6130233 -1.8632535-53
3 36.04215 29.88526 0.4398651

ENTER OPERATION CODE

RTS

ENTER (AREA#, AREA#-FOR-ROOTS, #ROOTS)
3,4,3

THE ROOTS ARE 
0.2155076E+01 
0.4393576E+00 
-7812769E-01

ENTER OPERATION CODE

RDX

ENTER (TO-AREA, DISK#)
2,12

ENTER OPERATION CODE

MLT

ENTER (PRE-AREA, POST-AREA, PROD-AREA)
2,1,3

ENTER OPERATION CODE

VT

ENTER AREA#
3

MATRX OUT
1 7.531164 2.883869 -0.4738973E-21
2 -0.1845938 -0.2801637E-01 -0.1862225E-22
3 -0.559.9536 -0.761.6377 0.833363

ENTER OPERATION CODE

RTS

ENTER (AREA#, AREA#-FOR-ROOTS, #ROOTS)
3,4,3

THE ROOTS ARE 
0.333334E+01 
0.283324E+00 
-1.622413E+01

ENTER OPERATION CODE

RD

ENTER (AREA#, MATRIX-TYPE(1=SYM, 2=ASYM))
2,1

ENTER ORDER
3

ENTER MATRIX ELEMENTS, ROWWISE, UPPER TRIANGULAR ONLY
423.57t34-12.795219 -2.573.439 
0.3564663 0.604.58333
999.000.25

ENTER OPERATION CODE

MLT

ENTER (PRE-AREA, POST-AREA, PROD-AREA)
2,1,3

ENTER OPERATION CODE
MATRIX OUTPUT
1 432386111 16.7769 26626.79
2 0.6564766 022.0560
3 10150358
ENTER OPERATION CODE
WTK
ENTER (FROM-AREA, DISK#)
3,12
ENTER OPERATION CODE
RD
ENTER (AREA#, MATRIX-TYPE(1=SYN, 2=ASYN))
2,1
ENTER ORDER
3
ENTER MATRIX ELEMENTS, ROWWISE, UPPER TRIANGULAR ONLY
113.513948.26389 4-1158.597
28.111394.2135.945
1217357
ENTER OPERATION CODE
SUB
ENTER (PRE-AREA, POST-AREA, DIFF-AREA)
2,1,3
ENTER OPERATION CODE
JNV
ENTER (AREA#, BZ-E, END-C, RSH-RHS, #RHS)
1,1,2,3,2
DETERMINANT 15 = 631323262+89
ENTER OPERATION CODE
WT
ENTER AREA#
1
MATRIX OUTPUT
1 13923383E-01 .2131214E-32 19543735E-04
2 .45662538E-01 -.9888844E-04
3 .21363923E-05
ENTER OPERATION CODE
MLT
ENTER (PRE-AREA, POST-AREA, PROD-AREA)
2,1,3
ENTER OPERATION CODE
YT
ENTER AREA#
3
MATRIX OUTPUT
1 2.065676 0.1567073 12817425E-03
2 .20973435E-01 1.330853 18571618E-02
3 .16734933 16.734933 2.3695633
ENTER OPERATION CODE
RTS
ENTER (AREA#, AREA- FOR-ROOTS, #ROOTS)
3,2,3
THE ROOTS ARE 23372362E+01 23517092E+01 12558473E+31
ENTER OPERATION CODE
RDK
ENTER (TO-AREA, DISK#)
2,11
ENTER OPERATION CODE
VT
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.365503</td>
<td>2.203456</td>
<td>-1.345591E-01</td>
</tr>
<tr>
<td>2</td>
<td>-0.2224584</td>
<td>-6.672446E-01</td>
<td>1.245666E-02</td>
</tr>
<tr>
<td>3</td>
<td>-357.5644</td>
<td>-167.2396</td>
<td>1.677920</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**RTS**

**ENTER (AREA#, AREA#=FOR-ROOTS, #ROOTS)**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4/3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**THE ROOTS ARE**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.7973433E+00</td>
<td>1.677361E+01</td>
<td>-1.673450E+30</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**RD**

**ENTER (AREA#, MATRIX-TYPE(1=SYN, 2=ASYM))**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**ENTER ORDER**

**ENTER MATRIX ELEMENTS, ROWWISE, UPPER TRIANGULAR ONLY**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15.1875</td>
<td>6.325</td>
<td>-1113.75</td>
<td>0.277</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**MLT**

**ENTER (PRE-AREA, POST-AREA, PROD-AREA)**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**UT**

**ENTER AREA#**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2621002</td>
<td>0.2267545</td>
<td>-1.2269216E-32</td>
</tr>
<tr>
<td>2</td>
<td>349766E-01</td>
<td>3033393E-01</td>
<td>-1.325355E-03</td>
</tr>
<tr>
<td>3</td>
<td>-19.02058</td>
<td>-16.85865</td>
<td>8.1653945</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**RTS**

**ENTER (AREA#, AREA#=FOR-ROOTS, #ROOTS)**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4/3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**THE ROOTS ARE**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.839583E+03</td>
<td>1.663792E+30</td>
<td>-1.085686E+30</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**RD**

**ENTER (AREA#, MATRIX-TYPE(1=SYN, 2=ASYM))**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**ENTER ORDER**

**ENTER MATRIX ELEMENTS, ROWWISE, UPPER TRIANGULAR ONLY**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.944445</td>
<td>-0.757322</td>
<td>-74.656242</td>
<td>2.311343</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**MLT**

**ENTER (PRE-AREA, POST-AREA, PROD-AREA)**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

**ENTER OPERATION CODE**

**UT**

**ENTER AREA#**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
VITA
Hsin-Ming Tzeng
Candidate for the Degree of
Master of Science

Report: Linear Comparisons in Multivariate Analysis of Variance

Major Field: Applied Statistics

Biographical Information:


Education: Received the Bachelor of Science degree from National Cheng-Chi University, Republic of China, with a major in Statistics in 1971; completed requirements for the Master of Science degree in Applied Statistics, at Utah State University in 1976.