MULTIVARIATE ANALYSIS OF VARIANCE FOR TWO WAY CLASSIFICATION

By

RHETT FEN-SHENG TSAO

Master of Science

Oklahoma State University

Stillwater, Oklahoma

1957

Submitted to the faculty of the Graduate School of
the Oklahoma State University in partial
fulfillment of the requirements for
the degree of
MASTER OF SCIENCE
1959

MULTIVARIATE ANALYSIS OF VARIANCE FOR TWO WAY CLASSIFICATION

Thesis Approved:

Robert D Morrison
Thesis Adviser

Franklin A. Graybill

Dean of Graduate School

ACKNOWLEDGEMENTS

The writer expresses his sincere appreciation to Dr. Robert D. Morrison and Dr. Franklin A. Graybill for their assistance in the preparation of this thesis and also for suggesting the problem.

Acknowledgements are also due to Dr. Wayne W. Huffine, Agronomy Department, Oklahoma State University, for making available the experimental data used in Table V.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The objective of this thesis is to study the multivariate analysis of variance by introducing some conventional matrix notation and to show the computation technique by using illustrations which a researcher without special mathematical training can easily follow. Most of the theory presented in this thesis has been developed, but it is hoped that this paper will make the theory easier to read and apply.

As opposed to the univariate analysis, this study is concerned with the value of n individuals each of which bears the value of p components which should be considered simultaneously. Thus, vectors are used instead of scalars. It is assumed that the errors associated with observation are independent in univariate case. In the multivariate case, the errors within a given component are assumed to be independent while the errors between components may be correlated.

This thesis will be restricted to the two way classification design having fixed effects. First, the maximum likelihood estimates of the parameters are found, then the likelihood ratio test is derived. An oversimplified illustrated example is given in order that one may follow the computing procedures with ease. A few multivariate analyses of variance on data from the Department of Agronomy, Oklahoma State University, are given and the results are compared with results when the data were handled as in univariate cases.

CHAPTER II

MODEL AND NOTATION

Consider the two way classification model with p components having fixed effects:

(2.1) $$y_{ijk} = \mu_i + \beta_{ij} + \alpha_{ik} + e_{ijk}$$

$$i = 1, 2, \ldots, p$$

$$j = 1, 2, \ldots, b$$

$$k = 1, 2, \ldots, t$$

where

y_{ijk} = observation of i^{th} component in j^{th} block k^{th} treatment

μ_i = general mean of all observations in i^{th} component

β_{ij} = effect of j^{th} block in i^{th} component

α_{ik} = effect of k^{th} treatment in i^{th} component

e_{ijk} = random error associated with the observation y_{ijk}.

We shall assume for every i, i' (i, i' = 1, 2, . . . , p) e_{ijk} is correlated with $e_{i'jk}$ and the covariance of e_{ijk} and $e_{i'jk}$ is $\sigma_{ii'}$; $e_{ijk} \sim NID(0, \sigma_{ii})$ for every fixed i.

Let us introduce the following matrix notation:

$$Y_{jk} = \begin{bmatrix} y_{1jk} \\ \vdots \\ y_{ijk} \\ \vdots \\ y_{pjk} \end{bmatrix}$$

is a p x 1 vector.

$$Y = \begin{bmatrix} Y'_{11} \\ \vdots \\ Y'_{jk} \\ \vdots \\ Y'_{bt} \end{bmatrix} = \begin{bmatrix} y_{111} \cdots & y_{i11} \cdots & y_{p11} \\ \vdots & \vdots & \vdots \\ y_{1jk} \cdots & y_{ijk} \cdots & y_{pjk} \\ \vdots & \vdots & \vdots \\ y_{1bt} \cdots & y_{ibt} \cdots & y_{pbt} \end{bmatrix}$$

is an n x p matrix where n = bt.

Z is an n x c design matrix with elements either 0 or 1 where c = 1 + b + t.

$$B = \begin{bmatrix} \mu_1 & \cdots \; \mu_i & \cdots \; \mu_p \\ \beta_{11} \cdots & \beta_{i1} \cdots & \beta_{p1} \\ \vdots & \vdots & \vdots \\ \beta_{1b} \cdots & \beta_{ib} \cdots & \beta_{pb} \\ \alpha_{11} \cdots & \alpha_{i1} \cdots & \alpha_{p1} \\ \vdots & \vdots & \vdots \\ \alpha_{1t} \cdots & \alpha_{it} \cdots & \alpha_{pt} \end{bmatrix}$$

is a c x p matrix.

$$e_{jk} = \begin{bmatrix} e_{1jk} \\ \vdots \\ e_{ijk} \\ \vdots \\ e_{pjk} \end{bmatrix}$$

is a p x 1 vector.

$$e = \begin{bmatrix} e'_{11} \\ \vdots \\ e'_{jk} \\ \vdots \\ e'_{bt} \end{bmatrix} = \begin{bmatrix} e_{111} & \cdots & e_{i11} & \cdots & e_{p11} \\ \vdots & & \vdots & & \vdots \\ e_{1jk} & \cdots & e_{ijk} & \cdots & e_{pjk} \\ \vdots & & \vdots & & \vdots \\ e_{1bt} & \cdots & e_{ibt} & \cdots & e_{pbt} \end{bmatrix}$$

is an n x p matrix.

Then (2.1) can be written in matrix notation as

$$Y = ZB + e \tag{2.2}$$

The assumption can be rewritten as

$$e_{jk} \sim N(\phi, V)$$

where ϕ is a vector (or matrix) with every element equal to zero and

$$\underset{pxp}{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

$$\mathrm{Cov}(e_{jk}, e_{j'k'}) = \phi$$

where $j \neq j'$ and/or $k \neq k'$.

$$e \sim N(\phi, V \otimes I)$$

where $V \otimes I$ is the Kronecker product (or direct product)of p x p matrix V and n x n matrix I

$$V \otimes I = \underset{pxp}{V} \otimes \underset{nxn}{\begin{bmatrix} 1 & & & \\ & 1 & & \phi \\ & & \ddots & \\ & \phi & & 1 \end{bmatrix}} = \underset{np \times np}{\begin{bmatrix} V & & & \\ & V & & \phi \\ & & \ddots & \\ & \phi & & V \end{bmatrix}}$$

CHAPTER III

MAXIMUM LIKELIHOOD ESTIMATES

In model $Y = ZB + e$ we shall estimate B and V by the method of maximum likelihood.

Since

$$f(e_{jk}) = \frac{1}{(2\pi)^{\frac{p}{2}} \left|V\right|^{\frac{1}{2}}} \exp.(-\frac{1}{2} e'_{jk} V^{-1} e_{jk})$$

the likelihood function is

$$(3.1) \qquad L = \prod_{jk}^{n} f(e_{jk}) = \frac{1}{(2\pi)^{\frac{np}{2}} \left|V\right|^{\frac{n}{2}}} \exp.(-\frac{1}{2} \sum_{jk}^{n} e'_{jk} V^{-1} e_{jk})$$

Since the exponent is scalar and a scalar is equal to its trace, we may rewrite the exponent as follows: (tr. denotes trace)

$$\begin{aligned} \exp. &= -\frac{1}{2} \sum_{jk}^{n} e'_{jk} V^{-1} e_{jk} \\ &= \mathrm{tr}[-\frac{1}{2} \sum_{jk} e'_{jk} V^{-1} e_{jk}] \\ &= -\frac{1}{2} \sum_{jk} \mathrm{tr}[e'_{jk} V^{-1} e_{jk}] \\ &= -\frac{1}{2} \sum_{jk} \mathrm{tr}[e_{jk} e'_{jk} V^{-1}] \\ &= -\frac{1}{2} \mathrm{tr}[V^{-1} \sum_{jk} e_{jk} e'_{jk}] \end{aligned}$$

$$= -\frac{1}{2} \text{ tr } V^{-1} e'e$$

$$= -\frac{1}{2} \text{ tr } V^{-1} (Y-ZB)'(Y-ZB) .$$

Therefore (3.1) may be written as

$$(3.2) \qquad L = \frac{1}{(2\pi)^{\frac{np}{2}} \left|V\right|^{\frac{n}{2}}} \quad \exp. [-\frac{1}{2} \text{ tr } V^{-1}(Y-ZB)'(Y-ZB)].$$

Using logarithms we get

$$(3.3) \qquad \ln L = -\frac{np}{2} \ln (2\pi) + \frac{n}{2} \ln \left|V^{-1}\right| - \frac{1}{2} \text{ tr } V^{-1} (Y-ZB)'(Y-ZB) .$$

To find the maximum of ln L we state the following definitions and lemmas.

Definition 3.1. *Let X be a p x q matrix with elements x_{ij} (i = 1, 2, . . . p; j = 1, 2, . . . , q). The derivative of a scalar f with respect to the matrix X, which will be written as $D_X f$ will mean the p x q matrix $(\frac{\partial f}{\partial x_{ij}})$ with ij^th element $\frac{\partial f}{\partial x_{ij}}$* .

Definition 3.2. *Two matrices A and B are independent if every element in A is independent of every element in B.*

Lemma 3.1. *Let*

$$f = \text{tr } AXB$$

where

A is n x p matrix

X is p x q matrix

B is q x n matrix

A and B are both independent of X, and the elements of X are independent. Then

$$D_X f = D_X(\text{tr } AXB) = (BA)' = A'B'$$

Proof: Since

$$\text{tr } AXB = \sum_{h=1}^{n} \sum_{i=1}^{p} \sum_{j=1}^{q} a_{hi}\, x_{ij}\, b_{jh}$$

$$D_X f = \frac{\partial \text{ tr } AXB}{\partial x_{ij}}$$

where

$$i = 1, 2, \ldots, p$$

$$j = 1, 2, \ldots, q$$

the typical element of $D_X f$ can be written as

$$(D_X f)_{st} = \frac{\partial \text{ tr } AXB}{\partial x_{st}} = \frac{\partial}{\partial x_{st}} \sum_{h=1}^{n} \sum_{i=1}^{p} \sum_{j=1}^{q} a_{hi} x_{ij} b_{jh}$$

$$= \sum_{h=1}^{n} a_{hs}\, b_{th} = \sum_{h=1}^{n} b_{th} a_{hs}$$

where

$$s = 1, 2, \ldots, p$$

$$t = 1, 2, \ldots, q$$

Therefore,

$$D_X f = (BA)' = A'B' .$$

Lemma 3.2. Let

$$f = \text{tr } AXBX'$$

where

A is a p x p symmetric matrix

X is a p x q matrix

B is a q x q symmetric matrix

A and B are both independent of X and the elements in X are independent. Then

$$D_X f = D_X(\text{tr } AXBX') = 2AXB \quad .$$

Proof:

$$\text{tr } AXBX' = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{k=1}^{q} \sum_{\ell=1}^{q} a_{ij}x_{jk}b_{k\ell}x_{i\ell}$$

The typical element of $D_X f$ can be written as

$$(D_X f)_{st} = \frac{\partial \text{ tr } AXBX'}{\partial x_{st}} \, , = \frac{\partial}{\partial x_{st}} \sum_{i}^{p} \sum_{j}^{p} \sum_{k}^{q} \sum_{\ell}^{q} a_{ij}x_{jk}b_{k\ell}x_{i\ell}$$

where

$$s = 1, 2, \ldots, p$$

$$t = 1, 2, \ldots, q$$

When we take the partial derivative of f with respect to x_{st}, the terms that do not involve x_{st} will vanish. We shall find only the terms which involve x_{st}. That is

$$a_{ss} x_{st} b_{tt} x_{st} + \sum_{\substack{i \\ i\ell \neq st}} \sum_{\ell} a_{is} x_{st} b_{t\ell} x_{i\ell} + \sum_{\substack{j \\ jk \neq st}} \sum_{k} a_{sj} x_{jk} b_{kt} x_{st}$$

$$(D_X f)_{st} = 2 a_{ss} x_{st} b_{tt} + \sum_{\substack{i \\ i\ell \neq st}} \sum_{\ell} a_{is} b_{t\ell} x_{i\ell} + \sum_{\substack{j \\ jk \neq st}} \sum_{k} a_{sj} x_{jk} b_{kt}$$

$$= 2 a_{ss} x_{st} b_{tt} + \sum_{\substack{i \\ i\ell \neq st}} \sum_{\ell} a_{si} x_{i\ell} b_{\ell t} + \sum_{\substack{j \\ jk \neq st}} \sum_{k} a_{sj} x_{jk} b_{kt}$$

$$= 2 a_{ss} x_{st} b_{tt} + 2 \sum_{\substack{j \\ jk \neq st}} \sum_{k} a_{sj} x_{jk} b_{kt}$$

$$= 2 \sum_{j} \sum_{k} a_{sj} x_{jk} b_{kt}$$

Therefore,

$$D_X f = 2AXB \ .$$

<u>Lemma 3. 3.</u> <u>Let</u>

$$f = \frac{n}{2} \ln |R| - \frac{1}{2} \text{tr } ARA'$$

<u>where</u>

<u>R is a p x p nonsingular symmetric matrix</u>

<u>A is an n x p matrix</u>

<u>n is a scalar constant</u>

<u>A is independent of R</u>

then $D_R f = \phi$ implies $R^{-1} = \frac{1}{n} A'A.$

Proof: Let r_{ss} be the s^{th} diagonal element in R.

$$\frac{\partial f}{\partial r_{ss}} = \frac{n}{2} \frac{1}{|R|} \frac{\partial}{\partial r_{ss}} |R| - \frac{1}{2} \frac{\partial}{\partial r_{ss}} \sum_h^n \sum_i^n \sum_j^p a_{hi} r_{ij} a_{hi}$$

$$= \frac{n}{2} \frac{1}{|R|} R_{ss} - \frac{1}{2} \sum_{h=1}^{n} a_{hs}^2$$

where R_{ij} is the cofactor of r_{ij} in R.

Let r_{st} be the st^{th} element in R and $s \neq t$. By symmetry,

$$\frac{\partial f}{\partial r_{st}} = \frac{n}{2} \frac{1}{|R|} 2R_{st} - \frac{1}{2} \, 2 \sum_{h=1}^{n} a_{hs} a_{ht}$$

$$\frac{\partial f}{\partial r_{ss}} = 0 \text{ implies } \frac{\partial f}{\partial r_{ss}} = n \frac{1}{|R|} R_{ss} - \sum_{h=1}^{n} a_{hs}^2$$

$$\frac{\partial f}{\partial r_{st}} = 0 \text{ implies } \frac{\partial f}{\partial r_{st}} = n \frac{1}{|R|} R_{st} - \sum_{h=1}^{n} a_{hs} a_{ht}$$

These two partials when set equal to zero may be written in matrix form as

$$D_R f = n \frac{1}{|R|} \text{Adjoint } R - A'A = \phi$$

Therefore

$$nR^{-1} = A'A$$

$$R^{-1} = \frac{1}{n} A'A$$

The maximum likelihood estimates of B and V are the solutions to the following equations:

$$\frac{\partial \ln L}{\partial b_{ij}} = 0 \tag{3.4}$$

where b_{ij} is ij^th^ element of B and $i = 1, 2, \ldots, c$; $j = 1, 2, \ldots, p$;

$$\frac{\partial \ln L}{\partial \sigma_{hk}} = 0 \tag{3.5}$$

where σ_{hk} is hk^th^ element of V and $h, k = 1, 2, \ldots, p$.

(3.4) and (3.5) can be summarized in matrix form as:

$$D_B(\ln L) = \phi \tag{3.6}$$

$$D_V(\ln L) = \phi \tag{3.7}$$

and denote B in (3.6) by $\widetilde{B}$.

From (3.3)

$$\begin{aligned} D_B(\ln L) &= -D_B[\frac{1}{2} \mathrm{tr}\ V^{-1}(Y-ZB)'(Y-ZB)] \\ &= -\frac{1}{2} D_B[\mathrm{tr}\ V^{-1}Y'Y - 2\mathrm{tr}\ V^{-1}Y'ZB + \mathrm{tr}\ V^{-1}B'Z'ZB] \\ &= D_B(\mathrm{tr}\ V^{-1}Y'ZB) - \frac{1}{2} D_B(\mathrm{tr}\ V^{-1}B'Z'ZB) \\ &= D_B(\mathrm{tr}\ Y'ZBV^{-1}) - \frac{1}{2} D_B(Z'ZBV^{-1}B'). \end{aligned}$$

Applying the lemmas (2.1) and (2.2) we have

$$D_B(\ln L) = Z'YV^{-1} - Z'Z\widetilde{B}V^{-1} = \phi$$

$$(Z'Y - Z'Z\widetilde{B})V^{-1} = \phi$$

since V^{-1} is positive definite

$$Z'Y - Z'Z\widetilde{B} = \phi .$$

Therefore,

$$Z'Z\widetilde{B} = Z'Y . \tag{3.8}$$

From the structure of the Z matrix we find that Z'Z is c x c matrix of rank c - 2, i.e., there are exactly c - 2 linearly independent rows in Z'Z. Since the rank of Z'Z is equal to the rank of Z'Z augumented by Z'Y, there are infinitely many $\widetilde{B}$ which will satisfy (3.8).

Now we shall make the restrictions:

$$\sum_{j=1}^{b} \beta_j = \begin{bmatrix} \beta_{11} \\ \vdots \\ \beta_{i1} \\ \vdots \\ \beta_{p1} \end{bmatrix} + \begin{bmatrix} \beta_{12} \\ \vdots \\ \beta_{i2} \\ \vdots \\ \beta_{p2} \end{bmatrix} + \dots + \begin{bmatrix} \beta_{1b} \\ \vdots \\ \beta_{ib} \\ \vdots \\ \beta_{pb} \end{bmatrix} = \phi$$

and

$$\sum_{k=1}^{t} \alpha_k = \begin{bmatrix} \alpha_{11} \\ \vdots \\ \alpha_{i1} \\ \vdots \\ \alpha_{p1} \end{bmatrix} + \begin{bmatrix} \alpha_{12} \\ \vdots \\ \alpha_{i2} \\ \vdots \\ \alpha_{p2} \end{bmatrix} + \dots + \begin{bmatrix} \alpha_{1t} \\ \vdots \\ \alpha_{it} \\ \vdots \\ \alpha_{pt} \end{bmatrix} = \phi$$

These two sets of restrictions require the sum of block effects and the sum of treatment effects within each component to be equal to 0.

If we let

$$\underset{2 \times c}{W} = \begin{bmatrix} 0 & 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & \underbrace{0 \quad 0 \quad \dots \quad 0}_{b \text{ columns}} & & & & \underbrace{1 \quad 1 \quad \dots \quad 1}_{t \text{ columns}} & & & \end{bmatrix}$$

and B be the parameter matrix as defined in Chapter II, the set of restrictions in α_i's and β_i's can be written in matrix form as

(3.9) $$WB = \phi$$

Combining (3.8) and (3.9), we can find the matrix $\hat{B}$, which will satisfy the following matrix equation

$$\begin{bmatrix} Z'Z \\ W \end{bmatrix} \hat{B} = \begin{bmatrix} Z'Y \\ \phi \end{bmatrix}$$

Therefore, $\hat{B}$ is then the maximum likelihood estimate of B under the restriction $WB = \phi$. $\hat{B}$ is uniquely determined since $\begin{bmatrix} Z'Z \\ W \end{bmatrix}$ is full rank.

Now let us find the maximum likelihood estimate of V. Due to the invariant properties of maximum likelihood estimates, the maximum of ln L in (3.1) with respect to V is equal to the maximum of ln L with respect to V^{-1} and the maximizing value of V is the inverse of the maximizing value of V^{-1}. We shall first find $D_{V^{-1}}(\ln L)$. From (3.3)

$$D_{V^{-1}}(\ln L) = D_{V^{-1}}\left[\frac{n}{2}\ln\left|V^{-1}\right| - \frac{1}{2}\operatorname{tr} V^{-1}(Y-ZB)'(Y-ZB)\right]$$

$$= D_{V^{-1}}\left[\frac{n}{2}\ln\left|V^{-1}\right| - \frac{1}{2}\operatorname{tr}(Y-ZB)V^{-1}(Y-ZB)'\right] .$$

By applying lemma (3. 3) we have

$$D_{V^{-1}}(\ln L) = \frac{n}{2}(\hat{V}^{-1})^{-1} - \frac{1}{2}(Y-ZB)'(Y-ZB) = \phi$$

$$\hat{V}^{-1} = n[(Y-ZB)'(Y-ZB)]^{-1}$$

where $\hat{V}^{-1}$ is the maximum likelihood of V^{-1}. Therefore, the maximum likelihood of V is

$$\hat{V} = \frac{1}{n}(Y-ZB)'(Y-ZB) .$$

By the invariant property of maximum likelihood estimates

$$\hat{V} = \frac{1}{n}(Y-Z\hat{B})'(Y-Z\hat{B}) . \qquad (3.10)$$

CHAPTER IV

TESTING HYPOTHESIS

To test the hypothesis that treatment effects are equal we shall partition matrices B and Z in the following manner:

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

where

$$\underset{(b+1)xp}{B_1} = \begin{bmatrix} \mu_1 & \cdots & \mu_p \\ \beta_{11} & \cdots & \beta_{p1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \beta_{1b} & \cdots & \beta_{pb} \end{bmatrix}$$

$$\underset{t \times p}{B_2} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_{1t} & \cdots & \alpha_{pt} \end{bmatrix}$$

and $Z = (Z_1 \; Z_2)$ such that the multiplications of Z_1B_1 and Z_2B_2 are defined.

Then the model $Y = XB + e$ can be written as

$$Y = (Z_1, Z_2) \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + e$$

$$= Z_1B_1 + Z_2B_2 + e \qquad (4.1)$$

In order to test that the treatment effects are equal, one may make an equivalent test, H_o: $B_2 = \phi$. This test is set up as follows. Let the parameter space be Ω and the parameter space under the null hypothesis be ω; and where Ω is defined on $p(2c + p + 1)/2$ dimensional space $E_{p(2c + p + 1)/2}$, and

i) V is positive definite,

ii) the elements in B_1 and B_2 range from $-\infty$ to ∞,

iii) $\sum_{j=1}^{b} \beta_j = \phi$ and $\sum_{k=1}^{t} \alpha_k = \phi$.

ω is defined on $p(2b + p + 3)/2$ dimensional space, $E_{p(2b + p + 3)/2}$, where

i) V is positive definite

ii) the elements in B_1 range from $-\infty$ to ∞,

iii) $B_2 = \phi$,

iv) $\sum_{j=1}^{b} \beta_j = \phi$.

Let $\hat{\Omega}$ be the point in $E_{p(2c + p + 1)/2}$ (i.e., the particular set of values of V_1, B_1, B_2 in Ω) such that $L(\Omega)$ will be maximum and denote this maximum value by $L(\hat{\Omega})$.

Let $\hat{\omega}$ be the point in $E_{p(2b + p + 3)/2}$ (i.e., the particular set of values of V_1 and B_1 in ω) such that $L(\omega)$ will be the maximum and denote this maximum value by $L(\hat{\omega})$. The test criterion is

$$L = \frac{L(\hat{\omega})}{L(\hat{\Omega})} . \tag{4. 2}$$

Now we have to find $L(\hat{\Omega})$ and $L(\hat{\omega})$. From (3. 2) and the invariant property of maximum likelihood estimates, we have

$$L(\hat{\Omega}) = \frac{1}{(2\pi)^{\frac{np}{2}} |\hat{V}|^{\frac{n}{2}}} \exp.[-\frac{1}{2} \mathrm{tr}\ \hat{V}^{-1} (Y-Z\hat{B})'(Y-Z\hat{B})] .$$

From (3.10)

$$\hat{V} = \frac{1}{n}(Y-Z\hat{B})'(Y-Z\hat{B}) .$$

We have

$$L(\hat{\Omega}) = \frac{\exp.\left[-\frac{1}{2}\mathrm{tr}\left\{n\,[(Y-Z\hat{B})'(Y-Z\hat{B})]^{-1}[(Y-Z\hat{B})'(Y-Z\hat{B})]\right\}\right]}{(2\pi)^{\frac{np}{2}}\left[\frac{1}{n}\left|(Y-Z\hat{B})'(Y-Z\hat{B})\right|\right]^{\frac{n}{2}}}$$

$$= (2\pi)^{-\frac{np}{2}} \frac{1}{n}\left|(Y-Z\hat{B})'(Y-Z\hat{B})\right|^{-\frac{n}{2}} \exp.\ [-\frac{1}{2}\mathrm{tr}\ (nI)] .$$

Since I is p x p identity matrix, then

$$L(\hat{\Omega}) = (2\pi)^{-\frac{np}{2}}\left[\frac{1}{n}\left|(Y-Z\hat{B})'(Y-Z\hat{B})\right|\right]^{-\frac{n}{2}} \exp.\ -\frac{1}{2}np . \tag{4. 3}$$

When $B_2 = \phi$ in parameter space ω, the model in (4.1) can be reduced to

(4.4) $$Y = Z_1 B_1 + e .$$

The likelihood function of reduced model is

(4.5) $$L(\omega) = \frac{1}{(2\pi)^{\frac{np}{2}} \left|V_\omega\right|^{\frac{n}{2}}} \exp.\left[-\frac{1}{2} \operatorname{tr} V_\omega^{-1} (Y - Z_1 B_1)'(Y - Z_1 B_1)\right] .$$

By the same token we used to find $\hat{B}$ and $\hat{V}$ in $L(\hat{\Omega})$, we obtain

(4.6) $$Z_1' Z_1 \hat{B}_\omega' = Z_1' Y$$

and

(4.7) $$\hat{V}_\omega = \frac{1}{n} (Y - Z_1 \hat{B}_\omega)'(Y - Z_1 \hat{B}_\omega)$$

where $\hat{B}_\omega$ and $\hat{V}_\omega$ are the maximum likelihood estimates of B_1 and V_ω respectively, under the restriction

$$\sum_{j=1}^{t} \beta_j = \phi .$$

Therefore,

$$L(\hat{\omega}) = \frac{1}{(2\pi)^{\frac{np}{2}} \left|\hat{V}_\omega\right|^{\frac{n}{2}}} \exp.\left[-\frac{1}{2} \operatorname{tr} \hat{V}_\omega^{-1} (Y - Z_1 \hat{B}_\omega)'(Y - Z_1 \hat{B}_\omega)\right]$$

$$= \frac{\exp.\left[-\frac{1}{2}\,\mathrm{tr}\left\{n\,[(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)]^{-1}\,[(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)]\right\}\right]}{(2\pi)^{\frac{np}{2}}\left[\frac{1}{n}\left|(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)\right|\right]^{\frac{n}{2}}}$$

$$(4.8) \qquad L(\omega) = (2\pi)^{-\frac{np}{2}}\left[\frac{1}{n}\left|(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)\right|\right]^{-\frac{n}{2}} \exp.\ -\frac{1}{2}\,np\ .$$

By substituting (4. 3) and (4. 8) in (4. 2) we obtain the likelihood ratio test criterion

$$L = \frac{(2\pi)^{-\frac{np}{2}}\left[\frac{1}{n}\left|(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)\right|\right]^{-\frac{n}{2}} \exp.\ -\frac{np}{2}}{(2\pi)^{-\frac{np}{2}}\left[\frac{1}{n}\left|(Y-Z\hat{B})'(Y-Z\hat{B})\right|\right]^{-\frac{n}{2}} \exp.\ -\frac{np}{2}}$$

$$(4.9) \qquad L = \left[\frac{\left|(Y-Z\hat{B})'(Y-Z\hat{B})\right|}{\left|(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)\right|}\right]^{\frac{n}{2}} = \lambda^{\frac{n}{2}}$$

where

$$\lambda = \frac{\left|(Y-Z\hat{B})'(Y-Z\hat{B})\right|}{\left|(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega)\right|}\ .$$

Remembering $Z'Z\hat{B} = Z'Y$, let us evaluate the quantities inside these two determinants of λ:

$$(Y-Z\hat{B})'(Y-Z\hat{B}) = Y'Y-\hat{B}'Z'Y-Y'Z\hat{B}+\hat{B}'Z'Z\hat{B}$$

$$= Y'Y-\hat{B}'Z'Y$$

$$= Q_1$$

and

$$(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega) = \left[(Y-Z\hat{B})+(Z\hat{B}-Z_1\hat{B}_\omega)\right]'\left[(Y-Z\hat{B})+(Z\hat{B}-Z_1\hat{B}_\omega)\right]$$

$$= (Y-Z\hat{B})'(Y-Z\hat{B})+(Z\hat{B}-Z_1\hat{B}_\omega)'(Y-Z\hat{B})$$

$$+(Y-Z\hat{B})'(Z\hat{B}-Z_1\hat{B}_\omega)+(Z\hat{B}-Z_1\hat{B}_\omega)'(Z\hat{B}-Z_1\hat{B}_\omega)$$

Applying (3.8) and (4.6) we evaluate the above equation term by term:

(a) $$(Y-Z\hat{B})'(Y-Z\hat{B}) = Q_1$$

(b) $$(Z\hat{B}-Z_1\hat{B}_\omega)'(Y-Z\hat{B}) = \hat{B}'Z'(Y-Z\hat{B})-\hat{B}'_\omega Z_1'Y+\hat{B}'_\omega Z_1'Z\hat{B}$$

$$= \hat{B}'Z'Y-\hat{B}'Z'Z\hat{B}-\hat{B}'_\omega Z_1'Y+\hat{B}'_\omega Z_1'Z\hat{B}$$

$$= \hat{B}'_\omega Z_1'Z\hat{B}-\hat{B}'_\omega Z_1'Y$$

(c) $$(Y-Z\hat{B})'(Z\hat{B}-Z_1\hat{B}_\omega) = (Y-Z\hat{B})'Z\hat{B}-Y'Z_1\hat{B}_\omega+\hat{B}'Z'Z_1\hat{B}_\omega$$

$$= \hat{B}'Z'Z_1\hat{B}_\omega-Y'Z_1\hat{B}_\omega$$

(d) $$(Z\hat{B}-Z_1\hat{B}_\omega)'(Z\hat{B}-Z_1\hat{B}_\omega) = \hat{B}'Z'ZB-\hat{B}'_\omega Z_1'Z\hat{B}-\hat{B}'Z'Z_1\hat{B}_\omega+\hat{B}'_\omega Z_1'Z_1\hat{B}_\omega$$

$$= \hat{B}'Z'Y-\hat{B}'_\omega Z_1'Z\hat{B}-\hat{B}'Z'Z_1\hat{B}_\omega+Y'Z_1\hat{B}_\omega$$

Combining the results of (a), (b), (c), and (d), we have

$$(Y-Z_1\hat{B}_\omega)'(Y-Z_1\hat{B}_\omega) = Q_1+\hat{B}'Z'Y-\hat{B}'_\omega Z_1'Y$$

$$= Q_1 + Q_2$$

where

$$Q_2 = \hat{B}'Z'Y - \hat{B}_\omega' Z_1'Y .$$

Thus

$$\lambda = \frac{|Q_1|}{|Q_1+Q_2|}. \tag{4.10}$$

We observe from (4.9) that L is a monotonic function of λ. From this fact, we see that λ can be used as a test function to test the hypothesis H_o: $B_2 = \phi$.

Under the null hypothesis, λ is distributed as $U_{(p,\ f_1,\ f_2)}$ where f_1 is the degrees of freedom of treatments, (t-1); and f_2 is the degrees of freedom of error, (t-1)(b-1).[1]

To test H_o, one rejects the null hypothesis at α probability level if $\lambda > \lambda_o$ where λ_o is a number such that

$$\int_{\lambda_o}^{\infty} f(U;\ p,\ f_1,\ f_2)dU = \alpha \tag{4.11}$$

Thus, it is desirable to know the distribution of U in order to evaluate the integral in (4.11). In general, this distribution can only be indicated as integrals. However, when p = 1 and p = 2, the distribution of U can be transformed to the Snedecor F as follows:

$$\frac{f_2}{f_1} \cdot \frac{1 - U_{1,\ f_1,\ f_2}}{U_{1,\ f_1,\ f_2}} = F_{f_1,\ f_2} \tag{4.12}$$

---

[1] See Bibliography [1] and [3].

$$(4.13) \qquad \frac{f_2 - 1}{f_1} \cdot \frac{1 - \sqrt{U_{2, f_1, f_2}}}{\sqrt{U_{2, f_1, f_2}}} = F_{2f_1, 2(f_2-1)} .$$

In these cases tabulated F can be used.

CHAPTER V

APPLICATION OF TECHNIQUES AND ILLUSTRATIONS

We shall now consider the technique of analysis which may be used to the best advantage when the model in (1.1) is assumed for the design.

Consider the statistical layout shown in Table I for a two way classification design with b blocks, t treatments, and p components.

TABLE I

STATISTICAL LAYOUT I

		Treatment						
		1	2	. . .	k	. . .	t	Sum
Block	1	Y_{11}	Y_{12}	. . .	Y_{1k}	. . .	Y_{1t}	$Y_{1\cdot}$
	2	Y_{21}	Y_{22}	. . .	Y_{2k}	. . .	Y_{2t}	$Y_{2\cdot}$
	.	.	.		.		.	.
	.	.	.		.		.	.
	.	.	.		.		.	.
	j	Y_{j1}	Y_{j2}	. . .	Y_{jk}	. . .	Y_{jt}	$Y_{j\cdot}$
	.	.	.		.		.	.
	.	.	.		.		.	.
	.	.	.		.		.	.
	b	Y_{b1}	Y_{b2}	. . .	Y_{bk}	. . .	Y_{bt}	$Y_{b\cdot}$
	Sum	$Y_{\cdot 1}$	$Y_{\cdot 2}$	. . .	$Y_{\cdot k}$	. . .	$Y_{\cdot t}$	$Y_{\cdot\cdot}$

where

$$\underset{p\times 1}{Y}_{jk} = \begin{bmatrix} y_{1jk} \\ \cdot \\ \cdot \\ \cdot \\ y_{ijk} \\ \cdot \\ \cdot \\ \cdot \\ y_{pjk} \end{bmatrix}$$

$$Y_{j\cdot} = \sum_{k=1}^{t} Y_{jk}$$

$$Y_{\cdot k} = \sum_{j=1}^{b} Y_{jk}$$

$$Y_{\cdot\cdot} = \sum_{j=1}^{b} \sum_{k=1}^{t} Y_{jk}$$

or

$$= \sum_{j=1}^{b} Y_{j\cdot}$$

or

$$= \sum_{k=1}^{t} Y_{\cdot k} \ .$$

Let

$$\underset{n\times p}{Y} = (Y_{11} \ . \ . \ . \ Y_{jk} \ . \ . \ . \ Y_{bt})'$$

$$\underset{n\times p}{Y_{\alpha}} = (Y_{\cdot 1} \ . \ . \ . \ Y_{\cdot k} \ . \ . \ . \ Y_{\cdot t})'$$

$$\underset{n\times p}{Y_{\beta}} = (Y_{1\cdot} \ . \ . \ . \ Y_{j\cdot} \ . \ . \ . \ Y_{b\cdot})'$$

Compute the following quantities

$$\underset{p \times p}{Y'Y} = \sum_{jk}^{bt} Y_{jk} Y'_{jk}$$

$$\underset{p \times p}{Y'_\alpha Y_\alpha} = \sum_{k}^{t} Y_{\cdot k} Y'_{\cdot k}$$

$$\underset{p \times p}{Y'_\beta Y_\beta} = \sum_{j}^{b} Y_{j\cdot} Y'_{j\cdot}$$

$$\underset{p \times p}{(Y_{\cdot\cdot})(Y_{\cdot\cdot})'} = \left[\sum_{jk}^{bt} Y_{jk}\right] \left[\sum_{jk}^{bt} Y_{jk}\right]' .$$

Reduction due to mean is

$$\frac{1}{bt} (Y_{\cdot\cdot})(Y_{\cdot\cdot})' .$$

Reduction due to blocks is

$$\frac{1}{t}(Y'_\beta Y_\beta) - \frac{1}{bt}(Y_{\cdot\cdot})(Y_{\cdot\cdot})' .$$

Reduction due to treatments (adjusted) is

$$Q_2 = \frac{1}{b}(Y'_\alpha Y_\alpha) - \frac{1}{bt}(Y_{\cdot\cdot})(Y_{\cdot\cdot})' .$$

The error sum of squares is

$$Q_1 = Y'Y - \frac{1}{t}(Y'_\beta Y_\beta) - \frac{1}{b}(Y'_\alpha Y_\alpha) + \frac{1}{bt}(Y_{\cdot\cdot})(Y_{\cdot\cdot})' .$$

These quantities can be put into an analysis of variance (A. O. V.) table.

TABLE II

A.O.V. FOR STATISTICAL LAYOUT IN TABLE I

Source	d.f.	S.S.	Test
Total	$n = bt$	$Y'Y$	
Mean	1	$\frac{1}{bt}(Y_{..})'(Y_{..})'$	
Blocks	$b-1$	$\frac{1}{t}(Y'_{\beta}\ Y_{\beta}) - \frac{1}{bt}(Y_{..})'(Y_{..})'$	
Treatments (adj)	$t-1$	Q_2	$U = \frac{\lvert Q_1 \rvert}{\lvert Q_1 + Q_2 \rvert}$
Error	$(b-1)(t-1)$	Q_1	

For an artificial numerical example, let $p = 2$, $b = 2$, $t = 3$, consider the following statistical layout:

TABLE III

STATISTICAL LAYOUT II

		Treatment			
		1	2	3	Sum
Block	1	1	2	3	6
		2	1	2	5
	2	1	2	1	4
		4	2	3	9
Sum		2	4	4	10
		6	3	5	14

H_o:

$$B_2 = \begin{bmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \\ \alpha_{13} & \alpha_{23} \end{bmatrix} = \phi$$

We shall compute the following quantities:

$$Y'Y = \begin{bmatrix} 1 & 2 & 3 & 1 & 2 & 1 \\ 2 & 1 & 2 & 4 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 2 \\ 1 & 4 \\ 2 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 20 & 21 \\ 21 & 38 \end{bmatrix}$$

$$Y_\alpha' Y_\alpha = \begin{bmatrix} 2 & 4 & 4 \\ 6 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 4 & 3 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 36 & 44 \\ 44 & 70 \end{bmatrix}$$

$$Y_\beta' Y_\beta = \begin{bmatrix} 6 & 4 \\ 5 & 9 \end{bmatrix} \begin{bmatrix} 6 & 5 \\ 4 & 9 \end{bmatrix} = \begin{bmatrix} 52 & 66 \\ 66 & 106 \end{bmatrix}$$

$$(Y_{..})'(Y_{..})' = \begin{bmatrix} 10 \\ 14 \end{bmatrix} [10 \quad 14] = \begin{bmatrix} 100 & 140 \\ 140 & 196 \end{bmatrix}$$

Reduction due to mean is

$$\frac{1}{6} \begin{bmatrix} 100 & 140 \\ 140 & 196 \end{bmatrix}$$

Reduction due to blocks is

$$\frac{1}{3} \begin{bmatrix} 52 & 66 \\ 66 & 106 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 100 & 140 \\ 140 & 196 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}$$

Reduction due to treatments (adjusted) is

$$\frac{1}{2} \begin{bmatrix} 36 & 44 \\ 44 & 70 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 100 & 140 \\ 140 & 196 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 8 & -8 \\ -8 & 14 \end{bmatrix}$$

The error sum of squares is

$$\begin{bmatrix} 20 & 21 \\ 21 & 38 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 52 & 66 \\ 66 & 106 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 36 & 44 \\ 44 & 70 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} 100 & 140 \\ 140 & 196 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix}$$

The results of this particular layout are summarized in Table IV as below:

TABLE IV

A.O.V. FOR STATISTICAL LAYOUT IN TABLE III

Source	d.f.	S.S.	Test
Total	6	$\begin{bmatrix} 20 & 21 \\ 21 & 38 \end{bmatrix}$	
Mean	1	$\frac{1}{6}\begin{bmatrix} 100 & 140 \\ 140 & 196 \end{bmatrix}$	
Blocks	1	$\frac{1}{6}\begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}$	
Treatments (adj)	2	$\frac{1}{6}\begin{bmatrix} 8 & -8 \\ -8 & 14 \end{bmatrix}$	U = .0545
Error	2	$\frac{1}{6}\begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix}$	

$$U = \frac{\frac{1}{6}\begin{vmatrix} 8 & 2 \\ 2 & 2 \end{vmatrix}}{\left| \frac{1}{6}\begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix} + \frac{1}{6}\begin{bmatrix} 8 & -8 \\ -8 & 14 \end{bmatrix} \right|} = \frac{\begin{vmatrix} 8 & 2 \\ 2 & 2 \end{vmatrix}}{\begin{vmatrix} 16 & -6 \\ -6 & 16 \end{vmatrix}} = \frac{12}{220} = .0545$$

This result is to be compared with the significance point for $U_{2, 2, 2}$. Using (4.13), we can transform it into F, i.e.,

$$\frac{2 - 1}{2} \cdot \frac{1 - \sqrt{.0545}}{\sqrt{.0545}} = 1.6413$$

is to be compared with the significance point of ${}_{.05}F_{4, 2} = 19.25$. This is not significant at the 5% level. We do not reject the H_0: $B_2 = \phi$.

Data from six experiments were obtained from the Department of Agronomy, Oklahoma State University, and the analyses of variance were computed for each year separately as a univariate case. Each experiment was then combined over two years and analyzed as the multivariate case in which the two components were years. Table V below shows the comparison among the probability levels of computed F values under the null hypothesis for each of the analyses.

TABLE V

PROBABILITY REGION, R, OF THE COMPUTED F VALUES UNDER THE NULL HYPOTHESIS FOR SIX EXPERIMENTS

Experiment	Univariate		Multivariate
	Year 1956	Year 1957	Two Years Combined
1	$F_{5,15} = 1.394$ $.10 < R < .30$	$F_{5,15} = 5.077$ $.005 < R < .01$	$F_{10,28} = 3.425$ $.001 < R < .005$
2	$F_{7,21} = 1.806$ $.10 < R < .30$	$F_{7,21} = .828$ $R > .50$	$F_{14,40} = 1.260$ $.10 < R < .30$
3	$F_{10,30} = .914$ $R > .50$	$F_{10,30} = 5.752$ $R < .0005$	$F_{20,58} = 2.592$ $.001 < R < .005$
4	$F_{3,9} = .753$ $R > .50$	$F_{3,9} = .846$ $R > .50$	$F_{6,16} = .618$ $R > .50$
5	$F_{3,9} = 1.945$ $.10 < R < .30$	$F_{3,9} = 19.820$ $R < .0005$	$F_{6,16} = 5.050$ $.001 < R < .005$
6	$F_{3,9} = .737$ $R > .50$	$F_{3,9} = 9.748$ $.001 < R < .005$	$F_{6,16} = 3.130$ $.025 < R < .05$

BIBLIOGRAPHY

[1] Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, Inc., 1957

[2] Graybill, Franklin A. General Linear Hypothesis Theory. Unpublished Notes, 1959.

[3] Wilks, S. S. Mathematical Statistics. Princeton University Press, Princeton, New Jersey, 1944.

VITA

Rhett Fen - sheng Tsao
candidate for the degree of
Master of Science

Thesis: MULTIVARIATE ANALYSIS OF VARIANCE FOR TWO WAY CLASSIFICATION

Major: Mathematics

Minor: Statistics

Biographical and Other Items:

Born: May 9, 1933 at Foochow, China

Undergraduate Study: Received the degree of Bachelor of Arts in Economics from National Taiwan University, Taipei, Taiwan, China, 1954.

Graduate Study: Received the degree of Masters in Economics from Oklahoma State University, Stillwater, Oklahoma, 1957. Oklahoma State University, Stillwater, Oklahoma, 1957-59

Experiences: Graduate Assistant in Mathematics Department, Oklahoma State University, 1957-59.

Member of the American Statistical Association and Pi Mu Epsilon

Date of Degree: May 24, 1959