AN ABSTRACT OF THE THESIS OF

El-Houssainy Abdelbar Rady for the degree of Doctor of Philosophy

in Statistics presented on May 1, 1986.

Title: Testing Fixed Effects in Mixed Linear Models

Redacted for Privacy

Abstract approved: ______________________

David S. Birkes

Redacted for Privacy

______________________

Justus F. Seely


This dissertation is concerned with hypothesis testing for fixed effects in mixed linear models. Our primary emphasis is on mixed models when the class of covariance matrices has what we call a rich linear structure. Such models include mixed ANOVA models and regression models with heteroscedastic variances. For the majority of our results we have also assumed that the covariance structure is commutative.

We show how to uniquely partition a chi-squared sum of squares into two independent parts, one related to the fixed effects and the other involving only error contrasts. Based on this partition and the commutativity of the covariance structure, we can decompose our model into a number of independent models. When we assume that the model has Zyskind structure, that is, that all estimable linear parameteric functions have BLUEs (best linear unbiased estimators), then this gives full partitions of both the expectation space and the error space. For a model with

Zyskind structure, a convenient minimal sufficient statistics is obtained.

The problem of testing a linear hypothesis about the fixed effects, when the model is commutative and has Zyskind structure, is shown to be invariant under a certain group of transformations. The maximal invariant statistic is derived, generalizing a result of Seifert (1979). The ANOVA test is seen to be uniformly most powerful invariant unbiased.

A uniformly most powerful test within the class of Bartlett-Scheffé tests is considered.

# Testing Fixed Effects in Mixed Linear Models

by

El-Houssainy Abdelbar Rady

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed May 1, 1986

Commencement June 1986

APPROVED:

Redacted for Privacy

Professor of Statistics in charge of major

Redacted for Privacy

Professor of Statistics in charge of major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 1, 1986

Typed by Genevieve J. Downing for El-Houssainy Abdelbar Rady.

## ACKNOWLEDGMENTS

I would like to express my deepest appreciation to Dr. David Birkes and Dr. Justus Seely who served as my major professors.

I would also like to thank the entire faculty, staff and students in the Statistics Department for too much to mention.

Finally, I dedicate this thesis to all those who have contributed to my education, and especially my father whose death occurred while completing this work, my mother, my wife Safia, my son Mohamed and my daughter Noha.

TABLE OF CONTENTS

# TESTING FIXED EFFECTS IN MIXED LINEAR MODELS

## I. INTRODUCTION

In the theory of mixed linear models much research has been done in the area of variance components, but little attention has been paid to inference about fixed effects. This is especially true for hypothesis testing and confidence intervals.

One approach to the problem of testing fixed effects that has been pursued in the literature is the idea of an ANOVA. The initial work here appears to be that of Graybill and Hultquist (1961), with contributions by other authors such as Albert (1976), Hultquist and Atzinger (1972), Brown (1983, 1984) and Harville (1985).

In mixed classification models when the ANOVA method does not lead to an exact test, a Satterthwaite approximation is frequently used and occasionally a likelihood ratio test is used. In addition, a few authors have investigated the possibility of constructing exact tests. For example, see Scheffé (1956), Imhof (1960) and Seifert (1979,1981).

This dissertation is concerned with hypothesis testing for the fixed effects in mixed linear models. Our primary emphasis is on mixed models when the class of covariance matrices has a rich linear structure (defined in Chapter III). This condition is essentially no restriction for problems of practical interest. However, for the majority of our results we have also assumed that the covariance structure is commutative. This condition is not satisfied by all mixed linear models, but it is satisfied by a wide class of models.

Chapter II introduces the terminology and the notation used throughout the thesis. This chapter also provides some algebraic results, which are needed for proving results in later chapters.

In Chapter III we describe the idea of linear structure for a class of covariance matrices and discuss some related topics. In this chapter we also introduce the general model used throughout the thesis. In addition, we derive a minimal sufficient statistic for the model, and discuss when this statistic is complete. This section is adapted from Seely (1977).

Chapter IV is devoted to partitioning sums of squares. We begin by showing how to uniquely partition a chi-squared sum of squares into two independent parts, one related to the fixed effects and the other involving only error contrasts. Based on this partition, assuming commutativity of the covariance structure, we can decompose our model into a number of independent models.

We give a partition of the model estimability space and illustrate it for the case of two variance components. Then we add the condition of Zyskind structure and give the partition under this assumption. In this case, we also obtain a partition of the expectation space and the error space, the minimal sufficient statistic, and a sufficient condition for completeness. The literature related to the ANOVA is reviewed and a necessary and sufficient condition for an ANOVA to exist is introduced.

The partition of Chapter IV is the starting point for our investigation of exact tests in Chapters V and VI.

Chapter V is devoted to hypothesis testing of fixed effects. The problems involved are first reviewed and discussed. Then a maximal invariant statistic under a group of transformations is derived for the special case of a Zyskind partition. In addition, we define an ANOVA test and show that it is a uniformly most powerful invariant unbiased test (UMPIU).

Chapter VI is devoted to the idea of a Bartlett-Scheffé test. In particular, the definition, construction, and optimal properties of such tests are considered. Particular attention is paid to the test statistics when the model has a Zyskind partition. Most of the results in Chapters V and VI are generalizations of work by Seifert (1979) on testing fixed effects in a k-way layout balanced mixed ANOVA model.

# II. LINEAR ALGEBRA RESULTS

In this chapter the notation which will be used throughout the thesis will be given. We state and prove some algebraic relationships which will be used to prove the results given in the thesis.

## 2.1. Notation

Regarding notation, we use $R^n$ to denote an n-dimensional Euclidean space. We use $1_n$, $J_n$ and $I_n$ to denote the $n \times 1$ vector of ones, the $n \times n$ matrix of ones and the $n \times n$ identity matrix. For a matrix A, let $A'$, $\underline{R}(A)$, $\underline{N}(A)$, $\underline{r}(A)$, $|A|$, $tr(A)$ denote the transpose, range, null space, rank, determinant and trace of A respectively. $\underline{R}(A)^{\perp}$ is used to denote the orthogonal complement of $\underline{R}(A)$. A useful relationship that we will frequently use is that $\underline{R}(A)^{\perp} = \underline{N}(A')$. The orthogonal projection on $\underline{R}(A)$ is $P_A$ and $N_A = I - P_A$ is the orthogonal projection on $\underline{R}(A)^{\perp}$. We use the abbreviations p.d. for positive definite and n.n.d. for nonnegative definite. We write $A \leq B$ if $B - A$ is an n.n.d. matrix.

For a subset $\mathbf{F}$ of a vector space let $\mathbf{F}_1$, $\mathbf{F}_0$ and f denote the affine hull of $\mathbf{F}$ ($\mathbf{F}_1 = \text{aff}\mathbf{F}$), the unique subspace of the vector space that is parallel to $\mathbf{F}$ ($\mathbf{F}_0 = \mathbf{F}_1 - \mathbf{F}_1$) and the dimension of $\mathbf{F}_0$ ($f = \dim\mathbf{F}_0$) respectively. Also let $\text{sp}\mathbf{F}$ denote the linear span of $\mathbf{F}$.

The family of probability measures induced by a random vector Y is denoted by $P_Y$. $N_n(\mu,\Delta)$ denotes an n-dimensional normal distribution with mean vector $\mu$ and covariance matrix $\Delta$. $\chi^2_{(n,\lambda)}$ denotes a chi-square distribution with n degrees of freedom and non-centrality parameter $\lambda$. We use the symbol $\underline{\underline{d}}$ to mean equal in distribution.

## 2.2. Some Linear Algebra Results

In this section we will introduce a proposition and some lemmas which will be used later in the thesis to prove most of the results.

Proposition 2.2.1. For an n x n symmetric matrix A and n x k matrix B the following statements are equivalent.

(A) $\underline{R}(AB) \subset \underline{R}(B)$

(B) $P_B A P_B = A P_B$

(C) A commutes with $P_B$

(D) A commutes with $N_B$

(E) $A = E + F$ where $E = E'$, $F = F'$, $\underline{R}(E) \subset \underline{R}(B)$ and $\underline{R}(F) \subset \underline{R}(B)^{\perp}$. Furthermore, E and F are unique.

Proof:

(A) => (B): (A) => $\underline{R}(AP_B) \subset \underline{R}(P_B)$ => $P_B A P_B = A P_B$.

(B) => (C): $P_B A P_B = A P_B = (P_B A P_B)' = (A P_B)' = P_B A$.

(C) <=> (D): $A P_B = P_B A \iff A N_B = A - A P_B = A - P_B A = N_B A$.

(D) => (E): Set $E = P_B A$ and $F = N_B A$. Then E,F have the correct containments and symmetry follows from (C) <=> (D). For uniqueness of E and F, suppose $E^*$ and $F^*$ also satisfy the stated conditions. Then $E^* + F^* = E + F \Rightarrow E^* - E = F^* - F$. Because $\underline{R}(B)$ and $\underline{N}(B')$ are disjoint, we must have $E^* = E$ and $F^* = F$.

(E) => (A): (E) => $AB = EB + FB = EB$. This gives the desired result becasue $\underline{R}(E)$ is contained in $\underline{R}(B)$.[]

Corollary 2.2.2. Any one of the equivalent statements in Proposition 2.2.1 imply $P_A P_B = P_B P_A$.

Proof: By (C) we get $\underline{R}(P_B A) \subset \underline{R}(A)$. Thus, (A) <=> (C) gives $P_B$ commutes with $P_A$.[]

It is interesting to note that the conclusion of the above corollary is not equivalent to the conditions in Proposition 2.2.1. To see this take A nonsingular and B singular. Then $P_A = I$. So, $P_B P_A = P_A P_B$. But it is not true that $P_B A = A P_B$ for all nonsingular A. For example, the choices

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad P_B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

give a counter-example.

Lemma 2.2.3. Let A, E and F be as in Proposition 2.2.1. If $A = A^2$, then $E = E^2$ and $F = F^2$.

Proof: $A = E + F = A^2 = E^2 + F^2$. So, $E - E^2 = F - F^2 = 0$. Because $\underline{R}(E)$ and $\underline{R}(F)$ are disjoint, the result follows.[]

Lemma 2.2.4. Assume $\mathbf{A}$ is affine. If there exists $b \in \mathbf{A}$ and

$\alpha \neq 1 \in R$ such that $\alpha b \in \mathbf{A}$, then $\mathbf{A}$ is a subspace.

Proof: Since $\mathbf{A}$ is affine, $\lambda_1 a_1 + \lambda_2 a_2 \in \mathbf{A}$ whenever $a_1, a_2 \in \mathbf{A}$ and $\lambda_1 + \lambda_2 = 1$. For $\alpha \neq 1$, there is a unique solution for $\lambda_1, \lambda_2$ in the equations

$$\lambda_1 + \lambda_2 = 1$$

$$\lambda_1 + \alpha\lambda_2 = 0 .$$

Thus, $0 = \lambda_1 b + \lambda_2(\alpha b) \in \mathbf{A}$. So, $\mathbf{A}$ is a subspace. []

Lemma 2.2.5. For any $p \times q$ matrix $C$, $C'C \leq I_q \Leftrightarrow CC' \leq I_p$.

Proof: We can find an orthogonal matrix $Q$ such that $Q'C'CQ = D$ where $D$ is a diagonal matrix with diagonal elements equal to the eigenvalues of $C'C$. So $C'C \leq I_q \Leftrightarrow D \leq I_q \Leftrightarrow$ all eigenvalues of $C'C$ are less than one. But $C'C$ and $CC'$ have the same nonzero eigenvalues. []

Now let us introduce the following lemma which is a generalization of Lemma 6.5 in Seely (1977). Let $V$ be a symmetric matrix and set

$$\mathbf{B} = \{\gamma I + \delta V : (\gamma,\delta) \in G\}$$

where $G$ is an open set in $R^2$ such that all $B \in \mathbf{B}$ are p.d. Set

$$\mathbf{C} = \{B^{-1} : B \in \mathbf{B}\}.$$

Let $V = \sum_{i=1}^{d} \alpha_i C_i$ be the spectral decomposition of $V$ where $\sum_i C_i = I$ and the $\alpha_i$ are distinct. Then:

Lemma 2.2.6. $\mathrm{sp}\mathbf{C} = \mathrm{sp}\{C_1,\ldots,C_d\}$.

Proof: For $B = \gamma I + \delta V \in \mathbf{B}$, we can write $B^{-1} = \sum(\gamma + \delta\alpha_i)^{-1}C_i$. It is clear $\mathrm{sp}\mathbf{C} \subset \mathrm{sp}\{C_1,\ldots,C_d\}$. Let $f_i(\gamma,\delta) = (\gamma + \delta\alpha_i)^{-1}$. Note $\mathbf{C} = \{\sum f_i(\gamma,\delta)C_i : (\gamma,\delta) \in G\}$. To show the reverse containment, it suffices to find $(\gamma_1,\delta_1)$, ..., $(\gamma_d,\delta_d) \in G$ such that $\det\{f_i(\gamma_j,\delta_j)\} \neq 0$ .

Since $G$ is open, we can find $(\bar{\gamma},\bar{\delta}) \in G$ and $\varepsilon > 0$ such that $\bar{\delta} \neq 0$ and $(\bar{\gamma} + \eta,\bar{\delta}) \in G$ for all $|\eta| \leq \varepsilon$. Let $(\gamma_j,\delta_j) = (\bar{\gamma} + \varepsilon/j\ ,\ \bar{\delta})$. Then $(\gamma_j,\delta_j) \in G$ and $f_i(\gamma_j,\delta_j)) = (\bar{\gamma} + \varepsilon/j + \bar{\delta}\alpha_i)^{-1} = (\mu_j + \tau_i)^{-1}$ where the $\mu_j$ are distinct and the $\tau_i$ are also distinct. So

$$\det\{\ f_i(\gamma_j,\delta_j)\ \} = \det\{\ (\mu_j + \tau_i)^{-1}\ \}.$$

But this is a Cauchy determinant and hence is nonsingular (e.g., see page 186 in Bellman (1960)).[]

# III. THE MODEL

In this chapter we first present definitions for linear structure, rich linear structure and commutative linear structure. Some examples of covariance matrices having these structures will be given.

We also introduce the model that will be considered throughout the thesis. We develop a minimal sufficient statistic and discuss completeness. The results regarding sufficiency and completeness are adapted from Seely (1977).

## 3.1. Linear Structure

In statistical problems the structure of the covariance matrix is important. Srivastava (1966) defined the notion of a 'pattern' for a covariance matrix; Anderson (1969) dealt with covariance matrices having 'linear structure'; and Rogers and Young (1975) defined 'patterned' covariance matrices. All these definitions are essentially the same and amount to the following notion.

Definition 3.1.1. A set $D$ of covariance matrices is said to have a linear structure if the elements $\Delta$ of $D$ are expressed as

$$\Delta = \sum_{i=0}^{k} \theta_i(\Delta) V_i$$

for a fixed set $V_0, V_1, \ldots, V_k$ of symmetric $n \times n$ matrices. The matrices $V_0, V_1, \ldots, V_k$ are called the generators of the linear structure.

Note that the generators $V_i$ are not required to be in **D**.

When **D** has a linear structure, we parametrize it by the parameter set

$$\boldsymbol{\Theta} = \{\theta(\Delta) \in R^{k+1} : \Delta \in \mathbf{D}\},$$

where $\theta(\Delta) = (\theta_0(\Delta), \ldots, \theta_k(\Delta))'$. Rather than regard $\theta$ as a function of $\Delta$, we usually prefer to regard $\Delta$ as a function $\Delta_\theta$ of $\theta$. that is,

$$\Delta_\theta = \textstyle\sum_{i=0}^{k} \theta_i V_i$$

for $\theta = (\theta_0, \ldots, \theta_k)'$.

Srivastava (1966) did not require the $V_i$'s to be symmetric. He required $k+1 < n(n+1)/2$, to avoid the trivial linear structure in Example 3.1.2, as did Rogers and Young (1975). Anderson (1969) required the generators to be linearly independent.

<u>Example 3.1.2</u>. Any set **D** of covariance matrices has a linear structure using the generators

$$V_{ii} = e_i e_i' \qquad 1 \leq i \leq n$$

$$V_{ij} = e_i e_j' + e_j e_i' \qquad 1 \leq i \leq j \leq n$$

where $e_i$ is the n x 1 vector with the $i^{th}$ component 1 and the other components 0. For any symmetric matrix C, we can write $C = \sum \sum (e_i' C e_j)\ V_{ij}$ . Note that these generators are linearly independent and that there are $n(n+1)/2$ of them.

The preceding linear structure is not very useful. For us, a linear structure is particularly useful if it satisfies the following condition.

Definition 3.1.3. The linear structure is said to be rich if $\boldsymbol{\Theta}$ contains an open set in $R^{k+1}$.

Example 3.1.4. The generating set in Example 3.1.2 gives a rich linear structure for the case when $\mathbf{D} = \mathbf{S}^+$, the set of all p.d. n x n matrices. This is because the generating set in Example 3.1.2 is a vector space basis for $\mathbf{S}$, the vector space of all symmetric n x n matrices, and $\mathbf{S}^+$ is an open set in $\mathbf{S}$.

Example 3.1.5. Consider the model

$$Y_{ij} = \mu + a_i + e_{ij}$$

where $i = 1, 2$ and $j = 1, \ldots, n_i$. Assume $n_1 = 2, n_2 = 1$ and write the model in matrix form as

$$Y = X\beta + Aa + e$$

where

$$X = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Suppose $a \sim N_3(0, \theta_1 I_3)$, $\theta_1 \geq 0$,

$e \sim N_3(0, \theta_0 I_3)$, $\theta_0 > 0$,

$a$ and $e$ independent.

A rich linear structure for Cov(Y) is

$$\text{Cov}(Y) = \theta_0 I + \theta_1 AA'.$$

If we use the generating set in Example 3.1.2, Cov(Y) will have the following linear structure

$$(\theta_0 + \theta_1)\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} + (\theta_0 + \theta_1)\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix} +$$

$$(\theta_0 + \theta_1)\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{pmatrix} + \theta_1\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} +$$

$$0\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix} + 0\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix} .$$

Of course, this structure is not rich.

Definition 3.1.6. A linear structure is said to be commutative if the generators $V_0, \ldots, V_k$ commute.

The main advantage of $\mathbf{D}$ having a commutative linear structure is that $\mathbf{D}$ admits a simultaneous spectral decomposition. That is, there is a collection of matrices $H_0, \ldots, H_m$ which satisfy conditions (C1) and (C2) below. Conditions (C3) and (C4) can be arranged.

(3.1.7) (C1) The $H_j$ are mutually orthogonal projections.

(C2) $V_i = \sum_{j=0}^{m} \lambda_{ij} H_j \quad , \quad i = 0, 1, \ldots, k.$

(C3) $H_0 = I - P_B$ where $B = (V_1, \ldots, V_k)$.

(C4) $m$ is minimal.

We can arrange condition (C3) because $V_0(I - P_B) = I - P_B$ and $V_i(I - P_B) = 0$ for $i \neq 0$, so that the columns of $I - P_B$ are simultaneously eigenvectors for $V_0, \ldots, V_k$. Among all collections of matrices $H_0, \ldots, H_m$ satisfying (C1), (C2) and (C3), we can choose one so that $m$ is minimal. This collection is unique apart from indexing and, discarding $H_0$ if it is 0, is the basis for the commutative quadratic subspace consisting of all polynomials of the $V_i$ (see Seely (1971) Lemma 6).

For the rest of this section we will assume the conditions in (3.1.7).

Lemma 3.1.8. $\underline{R}(H_1, \ldots, H_m) = \underline{R}(V_1, \ldots, V_k)$ .

Proof: It follows from (3.1.7) that $\underline{R}(V_i) \subset \sum_{j=1}^{m} \underline{R}(H_j) = \underline{R}(H_1, \ldots, H_m)$. Hence $\underline{R}(V_1, \ldots, V_k) \subset \underline{R}(H_1, \ldots, H_m)$. For the reverse containment it suffices to show $\bigcap_i \underline{N}(V_i) \subset \bigcap_j \underline{N}(H_j)$.

Suppose $x \in \underline{N}(V_i)$ for all $i$. For each $s$ there exists $r$ such that $\lambda_{rs} \neq 0$ (or else $m$ would not be minimal). Now $0 = V_r x = \sum_j \lambda_{rj} H_j x$. Premultiplying by $H_s$, we find $\lambda_{rs} H_s x = 0$, so $H_s x = 0$. []

Now $\Delta_\theta$ can be expressed as a function of the $H_j$ as follows

(3.1.9) $\Delta_\theta = \sum_{i=0}^{k} \theta_i V_i = \sum_{i=0}^{k} \sum_{j=0}^{m} \theta_i \lambda_{ij} H_j = \sum_{j=0}^{m} \delta_j(\theta) H_j$

where $\delta_j(\theta) = \lambda_j{}'\theta$ and $\lambda_j = (\lambda_{0j}, \ldots, \lambda_{kj})'$ , $j = 0, \ldots, m$. For convenience write $\delta_j(\theta) = \delta_j$.

Lemma 3.1.10. $\lambda_0, \ldots, \lambda_m$ are distinct.

Proof: Suppose $\lambda_j = \lambda_0$, $j \geq 1$. Then $\lambda_{ij} = 0$ for all $i \geq 1$

and so $H_j$ is not needed in (3.1.7), contradicting the minimality of $m$. Suppose $\lambda_j = \lambda_s$, $j \geq 1$, $s \geq 1$. Then $\lambda_{ij} = \lambda_{is}$ for all $i$, so $V_i = \lambda_{i1}H_1 + \ldots + \lambda_{im}H_m = \lambda_{i1}H_1 + \ldots + \lambda_{ij}(H_i + H_s) + \ldots + \lambda_{im}H_m$ for all $i$. But this contradicts the minimality of $m$. []

## 3.2. The Model

We consider a random vector $Y$ distributed according to some distribution in a family $\mathbf{P}_Y$. Throughout the thesis we will make the following six assumptions about the model.

(A1) $E(Y) = X\beta$ where $X$ is a known $n \times p$ matrix and $\beta$ is an unknown vector of parameters ranging over $R^p$.

(A2) $\mathrm{Cov}(Y) = \Delta \in \mathbf{D}$ where $\mathbf{D}$ has a rich linear structure. That is, $\mathbf{D} = \{\Delta_\theta : \theta \in \mathbf{\Theta}\}$ where $\Delta_\theta = \sum_{i=0}^k \theta_i V_i$ and $\mathbf{\Theta}$ contains a nonempty open set $G$ of $R^{k+1}$.

(A3) $V_0 = I \in \mathbf{D}$.

(A4) $\Delta_\theta$ is p.d. for all $\theta \in \mathbf{\Theta}$.

(A5) $\beta$ and $\theta$ are functionally independent.

(A6) $\mathbf{P}_Y$ is a family of multivariate normal distributions.

Besides these six basic assumptions, we will often make one or both of the following assumptions.

(A7) The linear structure in (A2) is commutative.

(A8) $\underline{R}(X)$ is an invariant subspace of $\Delta_\theta$ for all $\theta$.

Mixed linear models satisfy assumptions (A1)-(A5) with $\mathbf{\Theta} = \{(\theta_0, \theta_1, \ldots, \theta_k)' : \theta_0 > 0, \theta_1 \geq 0, \ldots, \theta_k \geq 0\}$. In balanced mixed classification models, assumptions (A7)-(A8) also hold.

## 3.3. A Minimal Sufficient Statistic

Let $\mathbf{D}_1$, $\mathbf{D}_0$, d be defined as in Section 2.1. Define $\mathbf{A} = \{\Delta^{-1} : \Delta \in \mathbf{D}\}$, and let $\mathbf{A}_1$, $\mathbf{A}_0$ and a be defined as indicated in Section 2.1.

Lemma 3.3.1. $\mathbf{D}_1 = \mathbf{D}_0$ and $\mathbf{A}_1 = \mathbf{A}_0$.

Proof: By Lemma 2.2.4 it suffices to find $\Delta \in \mathbf{D}$ and $\alpha \neq 1$ such that $\alpha\Delta \in \mathbf{D}$. According to assumption (A2), $\mathbf{D} = \{\Delta_\theta : \theta \in \mathbf{\Theta}\}$ and $\mathbf{\Theta}$ contains a nonempty open set G. Pick $\bar{\theta} \in G$. Define $\xi : R^1 \longrightarrow R^{k+1}$ by $\xi(\alpha) = \alpha\bar{\theta}$. Note $1 \in \xi^{-1}(G)$. Since $\xi$ is continuous, then $\xi^{-1}(G)$ is open, so it is not a singleton set, so there exists $\alpha \neq 1$ in $\xi^{-1}(G)$. Now $\alpha\bar{\theta} = \xi(\alpha) \in G \quad \mathbf{\Theta}$, so $\alpha\Delta_{\bar{\theta}} = \Delta_{\alpha\bar{\theta}} \in \mathbf{D}$. Note also that $\Delta_{\bar{\theta}}^{-1} \in \mathbf{A}$, $\alpha^{-1}\Delta_{\bar{\theta}}^{-1} = (\alpha\Delta_{\bar{\theta}})^{-1} \in \mathbf{A}$, and $\alpha^{-1} \neq 1$. []

Define $\mathbf{U}$ to be the smallest subspace of $R^n$ such that $\underline{R}(AX) \subset \mathbf{U}$ for all $A \in \mathbf{A}$. Seely (1977) has a useful way of expressing $\mathbf{U}$. In our case, when $\mathbf{A}_1 = \mathbf{A}_0$, we can express $\mathbf{U}$ as

$$\mathbf{U} = \underline{R}(A_1X) + \ldots + \underline{R}(A_aX)$$

where $A_1, \ldots, A_a$ constitute a basis for $\mathbf{A}_0$. Let $u = \dim \mathbf{U}$ and let $U = (U_1, \ldots, U_u)$ be such that $\underline{R}(U) = \mathbf{U}$. Let $T = (T_1, \ldots, T_u)$, where $T_i = U_i'Y$, $i = 1, \ldots, u$. Define S =

$(S_1, \ldots, S_a)$, where $S_i = Y'A_iY$. Theorem (2.5) in Seely (1977) states that $(T,S)$ constitutes a minimal sufficient statistic for $\mathbf{P}_Y$.

Remark 3.3.2. For any matrix $W$ with $\underline{R}(W) = \mathbf{U}$, the statistic $W'Y$ is equivalent to $T$. Note that if $W$ is partitioned as $W = (W_1, \ldots, W_t)$, then $W'Y = (W_1'Y, \ldots, W_t'Y)$.

Another form of the minimal sufficient statistic can be formulated. Let $Z = Q'Y$, where $Q$ is an $n \times q$ ($q = n - \underline{r}(X)$) matrix whose columns form an orthonormal basis for $\underline{R}(X)^{\perp}$. The class of distributions induced by $Z$, say $\mathbf{P}_Z$, is $\{N_q(0,\Delta^*) : \Delta^* \varepsilon \mathbf{D}^*\}$ where $\mathbf{D}^* = \{ Q'\Delta Q : \Delta \varepsilon \mathbf{D}\}$. Let $\mathbf{A}^* = \{\Delta^{*-1} : \Delta^* \varepsilon \mathbf{D}^*\}$ and let $\mathbf{A}_1^*$, $\mathbf{A}_0^*$ and $a^*$ be defined as in Section 2.1. Let $A_1^*, \ldots, A_{a^*}^*$ be a basis for $\mathbf{A}_0^*$. Define $R = (R_1, \ldots, R_{a^*})$ where $R_i = Y'QA_i^*Q'Y$, $i = 1, \ldots, a^*$. Theorem (2.7) in Seely (1977) states that $T$ and $R$ jointly constitute a minimal sufficient statistic for $\mathbf{P}_Y$.

Note: In the rest of this section we will assume (A7); that is, we will assume $\mathbf{D}$ has a commutative linear structure.

Under this assumption, the elements of $\mathbf{D}$ can be expressed as in (3.1.9). In the next lemma it will be shown that there exists $\bar{\theta}$ such that the $\delta_j(\theta)$ are distinct.

Lemma 3.3.3. There exists $\bar{\theta}$ in the interior of $\boldsymbol{\Theta}$ such that $\lambda_0'\bar{\theta}, \ldots, \lambda_m'\bar{\theta}$ are distinct.

Proof. For $i \neq j$ let $M_{ij} = \{t \varepsilon R^{k+1} : (\lambda_i - \lambda_j)'t = 0\}$. By

Lemma 3.1.10, the $\lambda_j$ are distinct and so $M_{ij}$ is a k-dimensional subspace in $R^{k+1}$. Let $M = \cup M_{ij}$. The (k+1)-dimensional Lebesgue measure of a k-dimensional subspace is zero and a finite union of sets of measure zero has measure zero. The measure of any nonempty open set, such as G in (A2), is positive, so $G \setminus M$ is nonempty. Note that $G \setminus M$ is the set of all $\theta \varepsilon G$ for which the $\lambda_j'\theta$ are all distinct. []

When k = 1, Olsen, Seely and Birkes (1976) and Seely (1977) discussed the properties of the model and answered a lot of questions concerning minimal sufficiency, completeness and estimability. In this case

$$\mathbf{D} = \{\theta_0 I + \theta_1 V : (\theta_0, \theta_1)' \ \varepsilon \ \mathbf{\Theta}\}$$

where $\mathbf{\Theta}$ is a subset of $R^2$. Some of their results were dependent on this particular structure of **D**. In the next proposition we detect a relationship between linear structure when k = 1 and linear structure when k > 1.

<u>Propostion 3.3.4</u>. There exist a symmetric matrix $\bar{V}$ and a nonempty open set $\bar{G} \subset R^2$ such that:

(a) $\bar{V} = \sum_{j=0}^{m} \bar{\delta}_j H_j$ with the $\bar{\delta}_j$ distinct.

(b) $\{\gamma_1 I + \gamma_2 \bar{V} : (\gamma_1, \gamma_2) \ \varepsilon \ \bar{G}\} \subset \mathbf{D}$.

<u>Proof</u>. Let $\bar{\theta}$ be as in Lemma 3.3.3 and let $\bar{V} = \sum_{i=0}^{k} \bar{\theta}_i V_i = \sum_{j=0}^{m} \bar{\delta}_j H_j$, $\bar{\delta}_j = \delta_j(\bar{\theta})$. The $\bar{\delta}_j$ are distinct. Let G be an open set of $R^{k+1}$ such that $\bar{\theta} \ \varepsilon \ G \subset \mathbf{\Theta}$. Define a function

$\xi : R^2 \dashrightarrow R^{k+1}$ by $\xi(\gamma_1, \gamma_2) = (\gamma_1 + \gamma_2\bar{\theta}_0, \gamma_2\bar{\theta}_1, \ldots, \gamma_2\bar{\theta}_k)$. Let $\bar{G} = \xi^{-1}(G)$. Since $\xi$ is continuous, $\bar{G}$ is open in $R^2$. It is nonempty because $(0, 1) \in \bar{G}$. []

Using Lemma 2.2.6 and Proposition 3.3.4, we obtain the following lemma.

Lemma 3.3.5. Let $\mathbf{H} = \mathrm{sp}\{H_0, \ldots, H_m\}$. Then $\mathbf{A}_0 = \mathbf{H}$.
Proof: Obviously $\mathbf{A}_0 \subset \mathbf{H}$. Thus we need to prove only the reverse containment. Let $\mathbf{W}$ be the subset of $\mathbf{D}$ described in part (b) of Proposition 3.3.4. Note $\mathrm{sp}\,\mathbf{W} \subset \mathbf{H}$. Let $\mathbf{W}^* = \{W^{-1} : W \in \mathbf{W}\}$. Then $\mathbf{W}^* \subset \mathbf{A} \subset \mathbf{A}_1 = \mathbf{A}_0$, so $\mathrm{sp}\,\mathbf{W}^* \subset \mathbf{A}_0$. The expression for $\bar{V}$ in part (a) of Proposition 3.3.4 is a spectral decomposition, and so Lemma 2.2.6 implies that $\mathrm{sp}\,\mathbf{W}^* = \mathbf{H}$. []

From Lemma 3.3.5 we see that $H_0, \ldots, H_m$ (dropping $H_0$ if it is zero) can be used as the basis $A_1, \ldots A_a$ for $\mathbf{A}_0$. Hence $a = m + 1$ (or $m$ if $H_0$ is zero). Furthermore,
$U = \underline{R}(H_0X) + \ldots + \underline{R}(H_mX)$ and $u = \sum_{j=0}^{m} \underline{r}(H_iX)$.

Theorem 3.3.6. Let $W_i$ be such that $\underline{R}(W_i) = \underline{R}(H_iX)$. Then $(W_0'Y, \ldots, W_m'Y, Y'H_0Y, \ldots, Y'H_mY)$ is a minimal sufficient statistic for $\mathbf{P}_Y$.
Proof: The proof follows directly from Theorem 2.5 in Seely (1977) and Remark 3.3.2. []

## 3.4. Complete Statistic

In this section, the assumption (A7) of commutativity is not needed.

Seely (1977) gave necessary and sufficient conditions for the sufficient statistic (T, R) to be complete. We give below some of his results.

Theorem 3.4.1. $\mathbf{P}_{T,R}$ is complete if and only if $\mathbf{P}_T$ is complete and $\mathbf{P}_R$ is complete.

Theorem 3.4.2. The following conditions are all equivalent.

(a) The family $\mathbf{P}_T$ is complete.

(b) $\underline{R}(\Delta X) \subset \underline{R}(X)$ for all $\Delta \varepsilon \mathbf{D}$.

(c) T and Z are independent with respect to every distribution in the family $\mathbf{P}_Y$.

(d) $\dim \mathbf{U} = \underline{r}(X)$.

Theorem 3.4.3. The following conditions are all equivalent:

(a) The family $\mathbf{P}_R$ is complete.

(b) $\mathbf{D}_0^*$ is a quadratic subspace.

(c) $\mathbf{D}_0^* = \mathbf{A}_0^*$ .

Note that Theorem 3.4.2 and 3.4.3 are somewhat simpler than Seely's results because of our assumption (A3).

# IV. PARTITIONING SUMS OF SQUARES

If $k > 0$ the description of the error space with just one term is no longer adequate. Therefore it is important to find ways to partition the sum of squares. In this chapter we will study this problem. In Section 4.2 we add assumption (A7) on the model introduced in Section 3.2, which allows us to develop a useful partition. In Section 4.3 we add a further assumption (A8), which is Zyskind's (1969) condition for the existence of BLUEs (best linear unbiased estimators). Under all the assumptions we develop a partition of the error space and the expectation space which we call the Zyskind partition. An advantage of this partition is its link with ordinary types of the analysis of variance. This development is a modification and polishing of Zyskind (1969). Also we construct a minimal sufficient statistic and give a sufficient condition for completeness. These results regarding sufficiency and completeness are a generalization of Example 6.4 in Seely (1977). Discussion about analysis of variance will be given and some related results will be investigated.

## 4.1. Partitioning a Chi-Squared Sum of Squares

In this section we investigate some conclusions that can be drawn when symmetric matrices have a certain relationship with the covariance structure. Note that in this section the richness assumption in (A2) is not needed.

<u>Theorem</u> <u>4.1.1.</u> For a symmetric matrix $H$ and a real valued function $\pi(\theta)$, the following two statements are equivalent.

(A) $Y'HY \overset{d}{=} \pi(\theta)\ \chi^2_{(h,\lambda)}$ for every $\theta$. In which case $h = tr(H)$ and $\lambda = ||HX\beta||^2$ .

(B) $H\Delta_\theta H = \pi(\theta)H$ for all $\theta \varepsilon \boldsymbol{\theta}$.

<u>Proof</u>: Follows directly from Rao and Mitra (1971).[]

<u>Remark</u> <u>4.1.2</u>. Since $I \varepsilon \mathbf{D}$, it must be true that $HH = cH$ for some $c$, so by properly normalizing $H$ we may suppose $H$ is idempotent.

Let $H$ and $\pi(\theta)$ be as in Theorem 4.1.1 with $H$ idempotent. Suppose also that $H$ satisfies (B) (and hence (A)) in the theorem. Let us see if the quadratic form $Y'HY$ can be partitioned into two independent parts, one related to the mean vector of $Y$ and the other to the error space. Note $HY \sim N_n(HX\beta, \pi H)$, where $\pi = \pi(\theta)$. Let us view this as a linear model with mean vector $HX\beta$ and covariance matrix $\pi H$. Since $\underline{R}(HX)$ is contained in $\underline{R}(H)$, we can partition $HY$ into $EY$ and $FY$, where $E$ is the orthogonal projection operator on $\underline{R}(HX)$ and $F = H - E$. Then we have the following theorem.

<u>Theorem</u> <u>4.1.3</u>. (a) $EY \sim N_n(HX\beta, \pi E)$

(b) $FY \sim N_n(0, \pi F)$

(c) $EY$ and $FY$ are independent.

Proof: Let $U = HY \sim N(HX\beta, \pi H)$. Note $EY = EU$, $FY = FU$.

For (a): $E(EU) = EHX\beta = HX\beta$ and $Cov(EU) = \pi EHE = \pi E$.

For (b): $E(FU) = FHX\beta = FEHX\beta = 0$ because $FE = 0$ and $Cov(FU) = \pi FHF = \pi F$.

For (c): $Cov(EU,FU) = \pi EHF = \pi EF = 0$. Since EU, FU are jointly normal they are independent.[]

Proposition 4.1.4. If $L'\beta$ is estimable in the model for HY and $F \neq 0$, then an F-statistic for testing $L'\beta = 0$ is

$$fY'Q_0Y/q_0Y'FY$$

where $f = trF$, $Q_0 = E - E_0$, $q_0 = trQ_0$, and $E_0$ is the orthogonal projection on $\{HX\beta : L'\beta = 0\}$.

Proof: As in Theorem 4.1.3, $Q_0Y \sim N(Q_0HX\beta, \pi Q_0)$. Note $Q_0HX\beta = 0 \Longleftrightarrow E(HX\beta) \in \underline{R}(E_0) \Longleftrightarrow L'\beta = 0$. Hence $Y'Q_0Y/\pi$ has a chi-squared distribution with $q_0$ degrees of freedom and a noncentrality parameter which is 0 if and only if $L'\beta = 0$. By Theorem 4.1.3, $Q_0Y = Q_0EY$ is independent of $Y'FY/\pi$, which has a central chi-squared distribution with f degrees of freedom.[]

Theorem 4.1.5. H can uniquely be expressed as a sum of two symmetric idempotent matrices E and F such that $H = E + F$ with $\underline{R}(F) \subset \underline{N}(X')$, $\underline{R}(E) = \underline{R}(HX)$ and $EF = 0$.

Proof: Follows from above. []

Remark 4.1.6. The matrices E and F have the following properties with respect to our covariance structure:

$E\Delta_\theta E = \pi(\ )E$ , $F\Delta_\theta F = \pi(\ )F$ , and $F\Delta_\theta E = 0$ for all $\theta \in \boldsymbol{\Theta}$.

Now let us consider a special case for the matrix H. Suppose H is an orthogonal projection operator such that:

(4.1.7) $\Delta_\theta H = \pi(\theta)H$ for all $\theta \in \boldsymbol{\Theta}$.

It is easy to check that this H satisfies (B) of Theorem 4.1.1 so it is indeed a special case of the H we have been considering. It is also interesting to note that H satisfies (4.1.7) if and only if $\underline{R}(H) \subset \{x : \Delta_\theta x = \pi(\theta)x,\ \text{all}\ \theta \in \boldsymbol{\Theta}\}$.

Now suppose W is a matrix satisfying the equivalent statements in Theorem 4.1.1 except with a function $\eta(\theta)$ instead of the function $\pi$. If $WH = 0$, then $W\Delta_\theta H = 0$ for all $\theta \in \boldsymbol{\Theta}$ so that WY and HY are independent. A sufficient condition for this to happen is the existence of a $\theta_0 \in \boldsymbol{\Theta}$ such that $\pi(\theta_0) \neq \eta(\theta_0)$ because, by the usual eigenvalue argument, one concludes $WH = 0$. If no such $\theta_0$ exists, then one can, if desired, combine W and H into a single matrix satisfying Theorem 4.1.1. Notice that when WY and HY are considered separately we can partition each into a fixed part and an error part and that all parts will be independent.

## 4.2. The H-Partition

In this section we add the assumption:

(A7) **D** has a commutative linear structure.

Now we can express the covariance matrix in the form $\Delta_\theta = \sum \delta_i H_i$

as stated in (3.1.9). Notice each $H_i$ in this collection satisifies (4.1.7), so all our statements in Section 4.1 are applicable. Conversely, if we can find an orthonormal set of orthogonal projection operators satisfying (4.1.7) (with possibly different $\pi$ functions) whose sum is the identity, then it is clear that $\Delta_\theta$ can be expressed as in (3.1.9), and hence that Y can be decomposed into $m + 1$ independent linear models.

For each $H_i$ let $E_i$ and $F_i$ be as in Theorem 4.1.5. Set $T_i = E_i Y$ and $R_i = Y' F_i Y$, $i = 0, \ldots, m$. It is easy to see that the vector T composed of $T_0, \ldots, T_m$ and the vector R composed of $R_0, \ldots, R_m$ is a sufficient statistic. In fact T and R form a minimal sufficient statistic.

Theorem 4.2.1. The statistics T and R form a minimal sufficient statistic.

Proof: Follows directly from Theorem 3.3.6. []

Now let us continue to investigate this decomposition. Set $P = \sum_i E_i$ and $N = \sum_i F_i$. Then it is straightforward to conclude that P and N are othrogonal projection operators whose sum is the identity. We observe that $\underline{R}(N) \subset \underline{N}(X')$ which implies $\underline{R}(X) \subset \underline{R}(P)$. One fact we get from this is the estimability partition stated in the next lemma.

Lemma 4.2.2. $\underline{R}(X')$ can be expressed as the sum

$$\underline{R}(X') = \underline{R}(X'E_0) + \ldots + \underline{R}(X'E_m).$$

<u>Proof</u>: Write $X'X = X'H_0X + \ldots + X'H_mX$. Use the fact that the range of the sum of n.n.d. matrices is the sum of the ranges and that $X'H_i = X'(E_i + F_i) = X'E_i$. []

<u>Example 4.2.3</u>. Let the model be

$$y_{ijk} = \mu + \alpha_i + c_j + e_{ijk}$$

where $i = 1,\ldots, a$, $j = 1, \ldots, t$, and $k = 1, \ldots, n_{ij}$. Assume $\mu,\alpha_1,\ldots,\alpha_a$ are fixed and that the $c_j$ and $e_{ijk}$ are random. Write the model in matrix form as $Y = X\beta + Bc + e$. Then

$$\mathrm{Cov}(Y) = \sigma^2 I + \sigma_b^2 BB' \quad .$$

Assume $B = (b_1, \ldots,b_t)$ and let $r_j = n_{.j}$, $j = 1,\ldots,t$. Then

$$b_i'b_j = 0 \quad \text{and} \quad BB' = \textstyle\sum_{j=1}^t b_jb_j'.$$

Now $(1/r_j)b_jb_j' = H_j$ is the orthogonal projection on $\underline{R}(b_j)$. So

$$\mathrm{Cov}(Y) = \sigma^2 I + \sigma_b^2 \textstyle\sum_j r_jH_j \quad .$$

Let $H_0 = I - (H_1 + \ldots + H_t)$. Then

$$\mathrm{Cov}(Y) = \sigma^2 H_0 + \textstyle\sum_j^t (\sigma^2 + r_j\sigma_b^2)H_j \quad .$$

Note that $\sum H_j = I$. If the $r_j$ are distinct, then this representation is minimal. If not, some are equal, and the corresponding $H_j$ can be pooled to give a minimal representation.

As we mentioned before we have $\underline{R}(N) \subset \underline{N}(X')$. In some situations, it can happen that we have equality. This equality condition can be stated in a number of ways and has a number of desirable consequences which we discuss in the next section.

### 4.3. Zyskind Partition

Let us in addition to (A1) ,..., (A7) assume:

(A8) $\underline{R}(V_iX) \subset \underline{R}(X)$ for all $i = 1, \ldots, k$.

This assumption is equivalent to $\underline{R}(X)$ being an invariant subspace of $\Delta_\theta$ for all $\theta$. It is also the Zyskind characterization for the BLUEs.

Lemma 4.3.1. (a) $E_j = P_X H_j = H_j P_X$

(b) $F_j = N_X H_j = H_j N_X$

(c) $\sum E_j = P_X$ and $\sum F_j = N_X$

Proof: A(8) implies $\underline{R}(H_iX) \subset \underline{R}(X)$ for $i = 0,\ldots,m$. So Proposition 2.2.1 and Lemma 2.2.3 give (a) and (b). Part (c) is now immediate because $H_0 + \ldots + H_m = I$.[]

Lemma 4.3.2. The sum in Lemma 4.2.2 is a direct sum.

Proof: It is sufficient to show that if

$$z = \sum_j X'E_j u_j = 0,$$

then $X'E_j u_j = 0$ for all j. If $z = 0$, then

$$\sum_j E_j u_j \varepsilon \underline{R}(X) \cap \underline{N}(X') = \{0\}.$$

It is easy to check that this implies each $E_j u_j = 0$.[]

Lemma 4.3.3. If $\underline{r}(X, V_1, \ldots, V_k) < n$, then $F_0 = I - P_{X,B}$ where $B = (V_1, \ldots, V_k)$ and $tr(F_0) = n - \underline{r}(X,B)$.

Proof: Immediate.[]

Example 4.3.4. Suppose we have two separate experiments, each being a balanced two-way layout without interaction and having one factor fixed and the other random. The model for the combined experiment is

$$E(Y) = X\beta \quad \text{and} \quad Cov(Y) = \theta_0 I + \theta_1 V$$

where

$$X = \begin{bmatrix} 11000 \\ 11000 \\ 10100 \\ 10100 \\ 10010 \\ 10010 \\ 10001 \\ 10001 \end{bmatrix} \qquad \beta = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}$$

$$V = \begin{bmatrix} 1010\ 0000 \\ 0101\ 0000 \\ 1010\ 0000 \\ 0101\ 0000 \\ 0000\ 1010 \\ 0000\ 0101 \\ 0000\ 1010 \\ 0000\ 0101 \end{bmatrix}.$$

Then $P_V = (1/2)V$ and $\Delta_\theta$ can be expressed as

$$\Delta_\theta = \theta_0(I-P_V) + (\theta_0 + 2\theta_1)P_V$$

This is a minimal representation and we also have (A8) satisfied. The table below summarizes the Zyskind partition:

| $i$ | $\delta_i$ | $H_i$ | $E_i$ | $F_i$ |
|---|---|---|---|---|
| 0 | $\theta$ | $I-P_V$ | $(I-P_V)P_X$ | $(I-P_V)(I-P_X)$ |
| 1 | $\theta_0+2\theta_1$ | $P_V$ | $P_V P_X$ | $P_V(I-P_X)$ |

Inference concerning any linear function of $(\alpha_1-\alpha_2, \alpha_3-\alpha_4)'$ should be done via $E_0Y$ and $F_0Y$ and inference about any linear function of $(\alpha_1+\alpha_2, \alpha_3+\alpha_4)'$ should be done via $E_1Y$ and $F_1Y$.

The statistics $E_jY$, $F_jY$ are the building blocks for any inference about the model. In fact as in Theorem 4.2.2 they form a minimal sufficient statistic for the model.

Let $I = \{i : E_i \neq 0\}$, $J = \{j : F_j \neq 0\}$ and $K = I \cap J$. It is convenient to rearrange the indices so that $J = \{0,1,\ldots,s\}$. According to Lemma 4.3.3 the original index 0 is in J if $\underline{r}(X,V_1, \ldots,V_k) < n$, which is usually true, in which case, we let $F_0$ be the same as it was before reindexing. Let T be the random vector with components $E_iY$, $i \varepsilon I$, and let $R = (Y'F_0Y,\ldots,Y'F_sY)'$.

We make one final remark in this section: Zyskind (1969) stated that if $F_i = 0$, then without further knowledge the quantities $\delta_i$ and $\beta'X'E_iX\beta$ cannot be separately estimated unbiasedly. This statement is not in general true, as we will see in Chapter VI.

## 4.4. Minimal Sufficiency and Completeness Under Zyskind Structure

Following Theorem 4.2.1 the statistics T and R jointly form a minimal sufficient statistic. Let $Z = Q'Y$, as defined in Section 3.3. Note that $Q'Q = I$ and $QQ'$ is the orthogonal projection on $\underline{N}(X')$.

Let $\mathbf{H}^* = sp\{Q'H_0Q, \ldots, Q'H_mQ\}$ which can also be expressed as $sp\{Q'F_0Q, \ldots, Q'F_sQ\}$. Recall the definition of $\mathbf{A}_0^*$ from

Section 3.3.

Lemma 4.4.1. $\mathbf{A}_0^* = \mathbf{H}^*$.

Proof: Starting with $\mathbf{W}_* = \{\theta_0 I + \gamma Q'\bar{V}Q : (\theta_0,\gamma) \varepsilon G\}$, we can use the same argument as in Lemma 3.3.6. []

Note that $QQ'F_j = F_j$ because $QQ' = N_X$. From this observation it can be established that $Q'F_0Q, \ldots, Q'F_sQ$ is a basis for $\mathbf{A}_0^*$, and that

$$Z'Q'F_jQZ = Y'F_jY \quad \text{for} \quad j = 0,\ldots,s.$$

Now using the results in Section 3.4 we can state the following theorems.

Theorem 4.4.2. T is complete for $\mathbf{P}_T$.

Proof: The proof follows directly from Theorem 3.4.2 because of assumption (A8). []

Now define $W_i = Q'V_iQ$ for $i = 0,\ldots,k$. Then (3.1.7) implies that we can write

$$W_i = \sum_j \lambda_{ij}Q'H_jQ = \sum_j \lambda_{ij}Q'F_jQ .$$

Let $G_j = Q'F_jQ$. Recall that $G_0, G_1, \ldots, G_s$ are nonzero. Let us express the linear relationships above as

$$\begin{pmatrix} W_0 \\ W_1 \\ \cdot \\ \cdot \\ \cdot \\ W_k \end{pmatrix} = \begin{pmatrix} \lambda_{00} & \cdots & \lambda_{0s} \\ \lambda_{10} & \cdots & \lambda_{1s} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \lambda_{k0} & \cdots & \lambda_{ks} \end{pmatrix} \begin{pmatrix} G_0 \\ G_1 \\ \cdot \\ \cdot \\ \cdot \\ G_s \end{pmatrix}$$

or in compact form as $W = AG$. Let $r + 1$ be the maximal number of linearly independent matrices among $W_0,\ldots,W_k$. Reindex so that the linearly independent ones are $W_0,\ldots,W_r$. Write $\bar{W} = (W_0,\ldots,W_r)$. Then $W = D\bar{W}$ for some k x (r+1) matrix D. Also $\bar{W} = \bar{A}G$ for a (r+1) x (s+1) submatrix $\bar{A}$ of A. If $c'\bar{A} = 0$, then $c'\bar{A}G = 0$. So, $c'\bar{W} = 0$. But the components of $\bar{W}$ are linearly independent, so $c = 0$. Hence $\bar{A}$ has full row rank, so $r \leq s$.

<u>Theorem</u> <u>4.4.3</u>. R is complete if and only if $s = r$, where $s + 1$ is the number of nonzero $F_j$, $j = 0,\ldots,m$ and $r + 1$ is the maximal number of linearly independent $Q'V_iQ$, $i = 0,\ldots,k$.

<u>Proof</u>: $s = r \Rightarrow \bar{A}$ is invertible which implies

$$sp\{G_0,\ldots,G_s\} = sp\{W_0,\ldots,W_1\} = sp\{W_0,\ldots,W_k\}.$$

Note that $sp\{G_0,\ldots,G_s\}$ is a quadratic subspace. Using Theorem 2, Seely (1971), this is equivalent to $(Z'G_0Z,\ldots,Z'G_sZ)' = R$ being complete.

Now suppose $s \neq r$. Since we have shown $r \leq s$, we must have $r < s$. So there exists $a \neq 0$ such that $\bar{A}a = 0$, which implies $Aa = D\bar{A}a = 0$. Let $R^*$ be

$$\left( \frac{Y'F_0Y}{f_0}, \ldots, \frac{Y'F_sY}{f_s} \right)', \quad \text{where } f_i = \underline{r}(F_i).$$

Now $E(R^*) = A'\theta = E(a'R^*) = a'A'\theta = 0$. But $a'R^* \neq 0$ so R is not complete. []

<u>Corollary</u> <u>4.4.4</u>. If R is complete, then $k \geq r = s$.

<u>Theorem 4.4.5</u>. (T,R) is complete if and only if $s = r$.

<u>Proof</u>: Use Theorems 4.4.2, 4.4.3 and 3.4.1. []

## 4.5. Some Distribution Results

In this section the distribution of $E_iY$ and $F_jY$ will be summarized assuming our usual model with assumptions (A1)-(A7). The distribution of the quadratic forms involving $E_i$ and $F_i$ and the ratio between them also will be summarized.

<u>Lemma 4.5.1</u>. For $j = 0,\dots,m$ we can state the following:

(a) $E_jY \sim N_n(E_jX\beta,\ \delta_jE_j)$

(b) $Y'E_jY/\delta_j \sim$ as chi-square with degrees of freedom $e_j = tr(E_j)$ and with noncentrality parameter $\lambda_j = \beta'X'E_jX\beta/\delta_j$.

(c) $F_jY \sim N_n(0,\delta_jF_j)$

(d) $Y'F_jY/\delta_j \sim$ as central chi-square with degrees of freedom $f_j = tr(F_j)$ .

Furthermore, $E_0Y ,\dots, E_mY$ , $F_0Y ,\dots, F_mY$ are all mutually independent.

<u>Lemma 4.5.2</u>. If $E_i$ and $F_i$ are nonzero, then

$$f_iY'E_iY/e_iY'F_iY$$

is distributed as F with $e_i$ and $f_i$ degrees of freedom and noncentrality parameter $||E_iX\beta||^2/\delta_i$ .

## 4.6. Another Parametrization

Let us suppose in this section that we have available a

Zyskind partition. That is, we assume (A1)-(A8). Then the identity $P_X = \sum_i E_i$ gives an orthogonal partition of the expectation space. Let us reparametrize our model to incorporate this partition. Let I be defined as in Section 4.3.

For $i \varepsilon I$ let $A_i$ be a matrix whose columns form an orthonormal basis for $\underline{R}(E_i)$. Then $A_iA_i' = E_i$ and $A_i'A_i$ is an identity matrix. Also, $A_i'A_j = 0$ for $i \neq j$. So,

$$\underline{R}(X) = \sum_i \underline{R}(A_i)$$

and the subspaces in the sum are all mutually orthogonal. Thus, we can reparametrize $E(Y)$ as

$$E(Y) = \sum_{i \varepsilon I} A_i \gamma_i \ . \qquad (4.6.1)$$

It is worth mentioning that inference about $\gamma_i$ is equivalent to inference about $E_iX\beta$.

Lemma 4.6.2. For $i \varepsilon I$, let $g_i$ be the BLUE for $\gamma_i$. Then

(a) $g_i = A_i'Y$ .

(b) $g_i \sim N_r(\gamma_i, \delta_i I_r)$ where $r = e_i$ .

Proof: Part (a) follows because $\sum A_j g_j = P_X Y$ and (b) is immediate.[]

Example 4.6.3. Consider Example 4.3.4. Then

$$A_0 = 1/2 \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ -1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & -1 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad A_1 = 1/2 \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

It is interesting to note that $\underline{R}(A_i)$ is the intersection of $\underline{R}(H_i)$ and $\underline{R}(X)$ for $i = 0,1$. The gamma vectors in the parametrization (4.6.1) are related to the beta vector in the following manner

$$\gamma_0 = 2\begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_3 - \alpha_4 \end{pmatrix} \qquad , \ \gamma_1 = 2\begin{pmatrix} 2\mu + \alpha_1 + \alpha_2 \\ 2\mu + \alpha_3 + \alpha_4 \end{pmatrix}$$

Inference concerning $\gamma_0$ and $\gamma_1$ can be viewed via $g_0$ and $g_1$ respectively.

## 4.7. Analysis of Variance

The analysis of variance (ANOVA) was introduced and developed by Fisher (1918). Graybill and Hultquist (1961) presented a general ANOVA theory for the random effect model. In the case of mixed models, the problem of determining when an ANOVA (defined below) exists is considered by several authors.

Albert (1976) introduced a general theorem on the distribution of quadratic forms that can be used to check whether a given partition of Y'Y forms an ANOVA. Brown (1983, 1984) gave a necessary and sufficient condition for a given partition of Y'Y to form an ANOVA for a mixed models. Harville (1985) derived a generalization of Albert's (1976) result.

In this section we will define ANOVA and partial ANOVA as in Brown (1984). We also introduce the notion of an independent chi-squared partition. We then give necessary and sufficient conditions for a quadratic form to admit an independent chi-squared partition, and we give some related results. We suppose thoughout this section that we have our standard model and that (A1)-(A6) are true.

Definition 4.7.1. An analysis of variance (ANOVA) is defined as a partition of $Y'Y$ into symmetric quadratic forms which

(a) are mutually independent,

(b) when suitably scaled are distributed as (possibly noncentral) chi-squared variables,

(c) have expectations which are different (beyond a known multiplicative constant) parametric functions.

Definition 4.7.2. If $G$ is a known n.n.d. matrix such that $Y'GY < Y'Y$ with probability one, then a partition of $Q = Y'GY$ with the properties of an ANOVA will be called a partial ANOVA.

Definition 4.7.3. For a symmetric idempotent matrix $G$, we say $Q = Y'GY$ admits an independent chi-squared partition if and only if $Q$ can be partitioned into symmetric quadratic forms that are independent and when suitably scaled are chi-squared variables.

Suppose $Q = Y'GY$ admits an independent chi-squared partition. Then $Q$ can be expressed as $Q = Q_1 + \ldots + Q_d$ where $Q_i = Y'G_iY$ and for some positive function $\pi_i(\theta) = \pi_i$ we have $Q_i/\pi_i$ is distributed as a chi-squared random variable. Combining the results of Section 4.1 with Definition 4.7.3, the following statements can be made:

(4.7.4)
(a) $G_1, \ldots, G_d$ are symmetric and indempotent
(b) $G = G_1 + \ldots + G_d$
(c) $G_i \Delta_\theta G_i = \pi_i G_i$ for all $i$
(d) $G_i \Delta_\theta G_j = 0$ for all $i \neq j$

where the conditions in (c) and (d) are true for all $\Delta \in \mathbf{D}$.

Lemma 4.7.5. Suppose $Q = Y'GY$ admits an independent chi-squared partition as described in the previous paragraph. Then the following statements can be made:

(a) Each $Q_i$ can be partitioned as $Q_i = B_i + C_i$ where $B_i$ is associated with the fixed effects and $C_i$ is associated with the error space.

(b) $B_1$, $C_1$, $B_2$, $C_2$ ... are mutually independent.

Proof: Use (4.7.4a), (4.7.4b), and Theorem 4.1.5.[]

For a matrix $B$ with $n$ rows, let $\mathbf{D}_B = \{B'\Delta B : \Delta \in \mathbf{D}\}$. That is, $\mathbf{D}_B$ is the covariance structure for $B'Y$.

Theorem 4.7.6. Consider $Q = Y'GY$ where $G$ is idempotent and symmetric. Then $Q$ admits an independent chi-squared partition if and only if $\mathbf{D}_G$ is commutative.

Proof: Suppose $Q$ admits an independent chi-square partition. Let the partition be as described in (4.7.4). Then

$$G\Delta_\theta G = \pi_1 G_1 + \ldots + \pi_d G_d \ .$$

From this, it is easy to check for $\theta,\eta \in \boldsymbol{\Theta}$ that

$$(G\Delta_\theta G)(G\Delta_\eta G) = (G\Delta_\eta G)(G\Delta_\theta G) \quad .$$

Conversely, suppose $\mathbf{D}_G$ is commutative. Then $GV_0G,\ldots,GV_kG$ commute so there exists an orthogonal matrix $R$ such that $R$ simultaneously diagonalizes the $GV_iG$ for all $i$. Then $Z = RGY$ is normal with a diagonal covariance matrix. Thus $Z_1,\ldots,Z_n$ are independent normals and $Y'GY = Z'Z$ which shows $Q$ can be partitioned into independent chi-squared random variables.[]

<u>Corollary 4.7.7</u>. Suppose $A$ is such that $A'A = I$ and $G = AA'$. Then $\mathbf{D}_G$ is commutative if and only if $\mathbf{D}_A$ is commutative.

<u>Proof</u>: $\mathbf{D}_A$ communitative implies $A'\Delta_\theta A\ A'\Delta_\eta A = A'\Delta_\eta AA'\Delta_\theta A$. Pre and post multiply by $A$ and $A'$ respectively then $G\Delta_\theta G\Delta_\eta G = G\Delta_\eta G\Delta_\theta G$. So, $D_G$ is commutative. Now suppose $\mathbf{D}_G$ is commutative. Then $G\Delta_\theta G\Delta_\eta G = G\Delta_\eta G\Delta_\theta G$. Pre and post multiply by $A'$ and $A$ respectively and the result follows. []

<u>Corollary 4.7.8</u>. A sufficient condition for $\mathbf{D}_G$ to be commutative is $G\Delta = \Delta G$ , $\Delta \in \mathbf{D}$, and $\mathbf{D}$ is commutative.

<u>Corollary 4.7.9</u>. If $Q = Y'GY$ has an independent chi-squared partition, then each $Q_i$ is a function of $GY$.

<u>Proof</u>: For each $i$, $Q_i = Y'G_iY = (GY)'G_i(GY)$. []

<u>Lemma 4.7.10</u>. If $N$, $M$ are two orthogonal projections such that $Y'NY$ and $Y'MY$ each has an independent chi-square partition, then

$Y'(N+M)Y$ has an independent chi-square partition if and only if $M\Delta_\theta N = 0$ for all $\theta \varepsilon \mathbf{\Theta}$.

<u>Proof</u>: If $M\Delta_\theta N = 0$, then $MY$ and $NY$ are independent. The result follows from Corollary 4.7.9. Conversely, it must be true that $Y'MY$ and $Y'NY$ are independent which implies $M\Delta_\theta N = 0$ for all $\theta \varepsilon \mathbf{\Theta}$.[]

Now let $G = M + N$ where $M$ and $N$ are orthogonal projections such that $MN = 0$. Consider the following conditions:

(C1) $\mathbf{D}_M$ is commutative

(C2) $\mathbf{D}_N$ is commutative

(C3) $M\Delta_\theta N = 0$ for all $\theta \varepsilon \mathbf{\Theta}$.

These conditions can be viewed, where taken together, as a necessary and sufficient condition for the existence of an independent chi-squared partition of $Y'GY$ with the property that part of the partition is an independent chi-squared partition of $Y'MY$ and the remaining part is an independent chi-squared partition of $Y'NY$. Also (C3) implies by Lemma (4.7.10) that any independent chi-squared partition of $Y'MY$ and of $Y'NY$ jointly constitite an independent chi-squared partition of $Y'GY$.

<u>Proposition</u> <u>4.7.11</u>. The conditions (C1), (C2) and (C3) are true if and only if $MV = VM$ for all $V \varepsilon \mathbf{D}_G$ and $\mathbf{D}_G$ is commutative.

Proof: Necessity is obvious. Now consider sufficency. Since $M$ commutes with all $V \in \mathbf{D}_G$, it is easy to show

(1) $N\Delta_\theta M = M\Delta_\theta N = 0 \Rightarrow$ (C3)

(2) $M\Delta_\theta M = M\Delta_\theta G$

Now using the fact that $\mathbf{D}_G$ is commutative and (2) we get (C1). A similar argument using $N\Delta_\theta N = N\Delta_\theta G$ gives (C2).[]

Remarks 4.7.12.

(1) Theorem 4.7.6 gives a necessary and sufficient condition for a partial ANOVA to exist. When $G = I$, it gives a necessary and sufficient condition for an ANOVA to exist.

(2) Under (A1),...(A8) a partition of $Y'Y$ can be obtained by using Proposition 4.7.11 with $M = P_X$ and $N = N_X$.

(3) Brown (1984) introduced the following procedure:

(a) Define $T = Y'P_XY$ , $S = Y'N_XY$ and partition $\beta$ into some given set of subvectors, say $\beta_1,\ldots,\beta_t$. Write $T = T_1 + \ldots + T_t$ where $T_1$ is the (least square) sum of squares for $\beta_1$ and $T_i$, for $i > 1$ is the sum of squares for $\beta_i$ adjusted for $\beta_1,\ldots,\beta_{i-1}$. Let $P_1,\ldots,P_t$ be the orthogonal projections such that $T_i = Y'P_iY$ , $i = 1, \ldots t$. Partition $S$ into quadratic forms, $S = S_1 + \ldots + S_s$.

(b) If $\{T_1,\ldots,T_t, S_1,\ldots S_s\}$ forms an ANOVA, then it is refered to as ANOVA $(\beta_1,\ldots,\beta_t,\theta)$.

(4) Theorem 1 in Brown (1984) stated that, given $T_1,\dots,T_t$ there exists and ANOVA $(\beta_1,\dots,\beta_t,\theta)$ if and only if

(a) $P_1,\dots,P_t$ , $V_1,\dots,V_k$ commute.

(b) For each $i = 1,\dots, t$ the nonzero eigenvalues of $P_i \Delta P_i$ are all the same.

Brown requires the quadratic forms $T_1,\dots,T_t$ to come from a particular $(\beta_1,\dots,\beta_t)$ of $\beta$. In Example 4.7.13 we see that this can be too restrictive. In this example, no partition of the $\beta$ vector leads to an ANOVA. But by reparemetrizing the model, an ANOVA can be obtained.

<u>Example</u> <u>4.7.13</u>. Reparametrize the model in Example 4.3.4 by letting $\gamma_i = \mu + \alpha_i$. Then

$$X = \begin{bmatrix} 1000 \\ 1000 \\ 0100 \\ 0100 \\ 0010 \\ 0010 \\ 0001 \\ 0001 \end{bmatrix} \qquad \beta = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix}$$

Following the Brown procedure, partition $\beta$ into $\beta_1 = \gamma_1$, $\beta_2 = \gamma_2$, $\beta_3 = (\gamma_3, \gamma_4)'$. In this case $P_1 V \neq VP_1$. So ANOVA$(\beta_1, \beta_2, \beta_3, \theta)$ does not exist.

Now let us consider another partition for $\beta$. Let $\beta_1 = (\gamma_1, \gamma_2)'$ and $\beta_2 = (\gamma_3, \gamma_4)'$. In this case $P_1 \Delta P_1$ does not satisfy (b) of (4) above. So ANOVA$(\beta_1, \beta_2, \theta)$ does not exist.

Now reparametrize again such that $\gamma_1^f = \gamma_1 + \gamma_2$, $\gamma_2^f = \gamma_3 + \gamma_4$, $\gamma_3^f = \gamma_1 - \gamma_2$ and $\gamma_4^f = \gamma_3 - \gamma_4$. Then

$$X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 \end{bmatrix} \qquad \beta = \begin{bmatrix} \gamma_1^f \\ \gamma_2^f \\ \gamma_3^f \\ \gamma_4^f \end{bmatrix}$$

Let $\beta_1 = (\gamma_1^f, \gamma_2^f)'$ and $\beta_2 = (\gamma_3^f, \gamma_4^f)'$. In this case ANOVA$(\beta_1, \beta_2, \theta)$ exists.

# V. THE PROBLEM

The basic problem considered in the remainder of this thesis is testing linear hypotheses (about fixed effects) in the model introduced in Section 3.2. In addition to the assumptions in Section 3.2, we assume throughout that the class of covariance matrices is commutative and for certain parts we also assume Zyskind structure.

This chapter is divided into four sections. The first section is devoted to a review of some related work and the second section discusses the basic problem. The third section discusses the testing procedure via invariance and the last section is devoted to what we refer to as the ANOVA test.

## 5.1. Related Review

In mixed models, tests about the fixed effects often lead to difficulties. A few authors have investigated this problem in the sense of constructing exact tests.

In (1965) Scheffé developed useful computational procedures for finding exact tests based on the $T^2$ statistic in the case of a two-way layout with one random effect and the other fixed. Imhof (1960) extended the results of Scheffé to the complete three-way layout with two random effects. He found exact tests and a multiple comparison method, both of which were based on the $T^2$ statistic. Seifert (1978) constructed exact tests for certain linear hypotheses in a general balanced normal mixed model and discussed some of their properties. Seifert (1981) introduced explicit formula for exact tests in a mixed balanced ANOVA models.

Some work that is closely related to testing linear hypotheses is the work that has been done to try and extend the idea of an ANOVA in the standard fixed effects model to mixed models. The initial work here seems to be that of Graybill and Hultquist (1961) with later contributions by Albert (1976), Brown (1983, 1984) and Harville (1985).

## 5.2. Definition of the Problem

The problem we consider in the remainder of the thesis is to test a linear hypothesis. In particular, we wish to test a linear hypothesis of the form $H: L'\beta = 0$. With regard to this linear hypothesis, we will be primarily interested in two questions. (i) Does an exact F-test exist for testing the linear hypothesis? (ii) Provided there exists an exact F-test, does there exist an exact F-test that has any optimal properties?

In attempting to answer the above questions, several types of problems can arise. Let us suppose in what follows the model of Section 4.2 (i.e., the model with commutativity) and the notation introduced in that section. The simplest situation that occurs for which an exact F-test exists is when $\underline{R}(L)$ is contained in $\underline{R}(X'E_v)$ for some $v$ with $F_v$ nonzero. In this case, Lemma 4.5.2 implies that an exact F-test exists for testing the linear hypothesis. Even in this case, however, it is frequently the case that other exact F-tests can be constructed. The following example, which also has Zyskind structure, illustrates the non-uniqueness of exact F-tests.

<u>Example 5.2.1</u>. Consider $E(Y) = X\beta$, $\Delta = \theta_0 I + \theta_1 V$ where

$$Y = \begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{31} \\ Y_{32} \end{pmatrix}, \qquad X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

and $V = \mathrm{diag}(0,0,1,1,2,2)$. Clearly the covariance structure is commutative and, in addition, we have Zyskind structure because $\underline{R}(VX)$ is contained in $\underline{R}(X)$. The Zyskind partition is described in the table below:

| $j$ | $\delta_j(\theta)$ | $H_j$ | $E_j$ | $F_j$ |
|---|---|---|---|---|
| 0 | $\theta_0$ | $\begin{matrix} 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{matrix}$ | $\frac{1}{2}\begin{matrix} 1&1&0&0&0&0 \\ 1&1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{matrix}$ | $\frac{1}{2}\begin{matrix} 1&-1&0&0&0&0 \\ -1&1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{matrix}$ |
| 1 | $\theta_0+\theta_1$ | $\begin{matrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&1&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{matrix}$ | $\frac{1}{2}\begin{matrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&1&1&0&0 \\ 0&0&1&1&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{matrix}$ | $\frac{1}{2}\begin{matrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&1&-1&0&0 \\ 0&0&-1&1&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{matrix}$ |
| 2 | $\theta_0+2\theta_1$ | $\begin{matrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&1 \end{matrix}$ | $\frac{1}{2}\begin{matrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&1&1 \\ 0&0&0&0&1&1 \end{matrix}$ | $\frac{1}{2}\begin{matrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&1&-1 \\ 0&0&0&0&-1&1 \end{matrix}$ |

Note that $E_jY$ is equivalent to the simpler statistic $U_j$ shown below. The word equivalence here means that there is a one to one linear transformation between the two statistics. Similarly $F_jY$ is equivalent to $Z_j$. It is worth mentioning that $U_j$ is associated with the fixed effects and $Z_j$ is an error contrast.

| j | $U_j$ | $E(U_j)$ | $Var(U_j)$ | $Z_j$ |
|---|---|---|---|---|
| 0 | $U_1 = Y_{11} + Y_{12}$ | $2\beta_1$ | $2\delta_0$ | $Z_1 = Y_{11} - Y_{12}$ |
| 1 | $U_2 = Y_{21} + Y_{22}$ | $2\beta_2$ | $2\delta_1$ | $Z_2 = Y_{21} - Y_{22}$ |
| 2 | $U_3 = Y_{31} + Y_{32}$ | $2\beta_3$ | $2\delta_1$ | $Z_3 = Y_{31} - Y_{32}$ |

A test statistic for H: $\beta_2 = 0$ is $U_2/\sqrt{Z_2^2}$ which has a t-distribution with one degree of freedom. Notice that this test statistic came from the $v = 1$ line of the above table and so it is like the first test described in the previous paragraph. However, a better test can be constructed as follows: Note that

$$Z = (Z_1 + Z_3)/\sqrt{2} \sim N(0,2\delta_1).$$

Thus, $2U_2 \;/\; \sqrt{(Z_2^2 + Z^2)}$ can also be used as a test statistic for H: $\beta_2 = 0$. And this last test is better than the first becasue it has a t-distribution with two degrees of freedom.

The next simplest situation that occurs in attempting to test the linear hypothesis H: $L'\beta = 0$ is when $\underline{R}(L)$ is contained in $\underline{R}(X'E_v)$ for some v, but the associated $F_v$ is zero. In this situation we are not in general guaranteed that an exact F-test exists. For example, consider the previous example except suppose that $F_2$ is zero (i.e., there is no $Z_2$) and assume we still wish to test $H: \beta_2 = 0$. In this case one can still form a t-test with one degree of freedom by using $U_2$ and $Z = (Z_1 + Z_3)/\sqrt{2}$. However, if additionally, $F_1$ was also zero, then we could not find an exact F-test for testing $H: \beta_2 = 0$. This problem will be further discussed in Chapter VI.

The most difficult situation to determine anything about the existence of an exact F-test is when $\underline{R}(L)$ is not contained in any one $\underline{R}(X'E_v)$. In this situation, exact tests are hard to construct even in very special cases. When L is a vector, some results will be given in Chapter VI.

In the above discussion we have been considering when an exact F-test exists for testing our linear hypothesis. This is not the common approach to the problem. The commonly used approach is to use a Satterthwaite approximation. In terms of our notation, the Satterthwaite approach is basically as follows: Find a linear combination of the $Y'E_iY$ whose expectation involves the norm of $L'\beta$ plus some linear combination $\pi$ of the $\delta_i$. Then try to find a linear combination of the $Y'F_iY$ whose expectation is $\pi$. Then approximate the ratio of these two linear combinations by an F-distribution. This technique does not in general lead to a test statistic whose distribution under the null hypothesis is an F-distribution.

## 5.3. Maximal Invariant Statistic

In this section we assume the model and notation of Section 4.2. That is, we have our usual model with assumptions (A1)–A(8). The purpose of this section is to consider testing the linear hypothesis $H: L'\beta = 0$ using invariance. Let $P_0$ be the orthogonal projection on $\{X\beta : L'\beta = 0\}$. The group of transformations we consider is the group $\mathbf{G}$ of affine transformations $g: Y \longrightarrow CY + d$ where $C$ and $d$ satisfy

(G1) $d \in \underline{R}(P_0)$,

(G2) $C'\Delta_\theta C = \Delta_\theta$ for all $\theta \in \boldsymbol{\Theta}$,

(G3) $C'P_X C = P_X$,

and (G4) $C'P_0 C = P_0$.

To investigate the group $\mathbf{G}$, it is convenient to first consider some lemmas. With regard to the C matrices describing the transformation group, notice that they are all orthogonal matrices because of (G2) and because the identity is one of the possible covariance matrices. The next lemma gives some equivalent ways of stating (G3) and (G4).

Lemma 5.3.1. If $C$ is an $n \times n$ orthogonal matrix and $A$ is an $n \times s$ matrix, then the following statements are equivalent:

(A) $\underline{R}(CA) = \underline{R}(A)$.

(B) $\underline{R}(C'A) = \underline{R}(A)$.

(C) $C'P_A C = P_A$ .

Proof: Straightforward.[]

Lemma 5.3.2. Let C be an $n \times n$ orthogonal matrix. The the set of matrices $\{B : B = B', C'BC = B\}$ is a quadratic subspace.

Proof: Check that the set is a subspace, and then check that $B^2$ is in the set whenever B is in the set.[]

Now let us introduce some notation. Let **C** denote the set of all C matrices that satisfy (G2)-(G4) and let **A** denote the set of all symmetric matrices A for which $C'AC = A$ for all C in **C**. The primary result in this section is a description of the maximal invariant under the group **G**. To investigate the maximal invariant, it is handy to know some facts about the set **A**. The primary facts we need are contained in the following lemma.

Lemma 5.3.3. The set **A** is a quadratic subspace. Moreover, $E_0, \ldots, E_m$, $F_0, \ldots, F_m$, $P_X$, and $P_0$ are all members of the set **A**.

Proof: By using Lemma 5.3.2 with conditions (G2)-(G4), and realizing the intersection of quadratic subspaces is a quadratic subspace, it follows that **A** is a quadratic subspace. Let $a \in \boldsymbol{\Theta}$ be such that the $d_j = \delta_j(a)$ are distinct. Since $\Delta_a \in \mathbf{D}$ and since $\Delta_a = \sum_j d_j H_j$ is a spectral representation with the $d_j$ distinct and nonzero, it follows (because **A** is a quadratic subspace) that $H_0, \ldots, H_m \in \mathbf{A}$. Clearly $P_X$ and $P_0$ are also in $\mathbf{A}$. Thus, $P_X H_i P_X$ is in $\mathbf{A}$. Because of (A8), this last quantity is $E_i$. Similarly, the $F_i$ are in $\mathbf{A}$ because $N_X$ is in $\mathbf{A}$.[]

For convenience in expression, let $M = P_X - P_0$. Note that $M$ is in $\mathbf{A}$ and hence that $MAM$ is in $\mathbf{A}$ whenever $A$ is in $\mathbf{A}$ because of the quadratic subspace property. In dealing with invariance we frequently will insist that our model have one additional assumption. This assumption is:

(A9) $\underline{R}(V_i P_0) \subset \underline{R}(P_0)$ for $0 = 1, \ldots, k$.

Notice that this assumption is the same as assumption (A8), except applied to the linear model under the null hypothesis. That is, this assumption means that the linear model under the null hypothesis has Zyskind structure.

Lemma 5.3.4. Assume (A9). Then we can write

$$E_j = E_j M + H_j P_0 , \quad j = 0, \ldots, m,$$

where $E_j M = ME_j$ and $H_j P_0 = P_0 H_j$ are orthogonal projection operators. Furthermore, these two orthogonal projection operators are orthogonal to one another.

Proof: Because of (A9), Proposition 2.2.1 implies $H_j$ can be expressed as

$$P_0 H_j + (N_X + M) H_j = P_0 H_j + MH_j + F_j .$$

And the result follows because $MH_j = M(E_j + F_j) = ME_j$. []

Lemma 5.3.5: If $\Delta_\theta A = \pi(\theta)A$ for all $\theta \varepsilon \boldsymbol{\Theta}$ and $A = A'$, then A commutes with **D**. Moreover, if A,B are symmetric and commute with **D**, then any linear combination of A and B commutes with **D**.

Proof: Note that $\Delta_\theta = \pi(\theta)A$ implies commutativity because A is symmetric. Commutativity of a linear combination is straightforward to verify.[]

Let $T_1(Y)$ and $T_2(Y)$ be the random vectors composed of the quadratic forms

$$Y'ME_jMY \qquad \text{and} \qquad Y'F_jY$$

respectively where $j = 0,\ldots,m$. Set $T(Y) = (T_1(Y),T_2(Y))$.

Lemma 5.3.6. T(Y) is invariant under **G**.

Proof: For d in the affine transformation, note that (G1) implies $F_jd = 0$ and $Md = 0$. Using this observation and Lemma 5.3.3, we have for each $j = 0,\ldots,m$

$$(CY + d)'F_j(CY + d) = Y'C'F_jCY = Y'F_jY,$$

$$(CY + d')ME_jM(CY + d) = Y'ME_jMY.$$

Therefore, T(Y) is invariant under **G**.[]

Theorem 5.3.7. Assume (A9). Then T(Y) is a maximal invariant statistic under **G**.

Proof: By Lemma 5.3.6 T(Y) is invariant under **G**. Now suppose $T(y_1) = T(y_2)$. Then for $j = 0,\ldots,m$,

$$||E_jMy_1||^2 = ||E_jMy_2||^2$$

and $\quad ||F_j y_2||^2 = ||F_j y_1||^2$ .

Because of Lemma 5.3.4 there is an orthogonal transformation C mapping $\underline{R}(E_j M)$, $\underline{R}(H_j P_0)$ and $\underline{R}(F_j)$ onto themselves such that $E_j M y_2 = C E_j M y_1$ and $F_j y_2 = C F_j y_1$ for $j = 0,\ldots,m$.

Again using Lemma 5.3.4, we can write

$$y_1 = I y_1 = \textstyle\sum_j E_j M y_1 + P_0 y_1 + \sum_j F_j y_1.$$

So, $C y_1 = \sum_j C E_j M y_1 + C P_0 y_1 + \sum_j C F_j y_1 = y_2 - P_0 y_2 + C P_0 y_1$.

Thus, $y_2 = C y_1 + d$ where $d = P_0 y_2 - C P_0 y_1$.

Now using Lemmas 5.3.1 and 5.3.4 along with the construction of C, we can conlude that

$$C'AC = A \quad \text{for} \quad A = E_j M,\ H_j P_0,\ \text{and}\ F_j.$$

Thus, also for $A = E_j$ $(= E_j M + H_j H_0)$ and $A = H_j$ $(= E_j + F_j)$.

Therefore,

$$C' P_0 C = \textstyle\sum_{j=0}^m C' H_j P_0 C = \sum_{j=0}^m H_j P_0 = P_0 .$$

This shows (G1) and (G4) are satisfied. To show (G3), consider

$$C' P_X C = \textstyle\sum_{j=0}^m C' E_j C = \sum_{j=0}^m E_j = P_X.$$

To show (G2) consider

$$C' \Delta_\theta C = \textstyle\sum_{i=0}^m \delta_i C' H_i C = \sum \delta_i H_i = \Delta_\theta.$$

Therefore $Y \to CY + d$ is in G and $y_1 \to C y_1 + d = y_2$. Thus, we have established that $T(y)$ is a maximal invariant statistic.[]

## 5.4. The ANOVA Test

In this section we discuss the ANOVA method for testing the linear hypothesis $H: L'\beta = 0$. We assume the same model and notation as in Section 3. It is possible to discuss the ANOVA test without assumption (A8). However, because our final results require the assumption, it convenient to simply make the assumption throughout our discussion.

The basic idea of the ANOVA test is to take the chi-square test statistic which would arise if the covariance matrix were known to be an identity, and then look for a linear combination of the $Y'F_iY/f_i$ with the same expectation under the null hypothesis. To discuss the test in more detail, let $Q_H = Y'MY/r$ where $M = P_X - P_0$ (as defined in Section 3) and where $r = tr(M)$. The quadratic form $Q_H$ is the chi-squared test statistic which would be used if the covariance matrix were the identity matrix. By taking the expectation of $Q_H$ as a quadratic form in $Y$, we find that

$$E(Q_H) = \lambda + b'\theta$$

where $\lambda = ||MX\beta||^2/r$ and where the components of $b$ are $b_i = tr(MV_i)/r$ for $i = 0, \ldots, k$. A fact we will sometimes use is that $\lambda = 0$ if and only if $L'\beta = 0$.

It is convenient from here on to slightly change our notation. Suppose the indicies $j = 0,\ldots,m$ are reordered in such a way that $R_j = Y'F_jY$ for $j = 0,\ldots,s$ are all nonzero and that $F_j$ for $j = s+1,\ldots,m$ are zero. Now let us concentrate on the parametric function $b'\theta$. Set $R = (R_0, \ldots, R_s)'$. Then $E(R) = A'\theta$ where $A$ is the $(k+1)x(s+1)$

matrix defined in Section 4.4. Note that $A'\theta = \delta(\theta)$. Let us assume that $A'$ has full column rank. Then we can always find a $g$ such that $g'R$ is unbiased for $b'\theta$. From here on let $g$ be such that $g'R$ is unbiased for $b'\theta$. Some facts we use that can be found in the literature are outlined below:

(a) $\theta$ is estimable if $Q'V_iQ$, $i = 0, \ldots, k$, are linearly independent. Under Zyskind structure, this linear independence is necessary and sufficient for $A'$ to have full column rank.

(b) The choice of the vector $g$ is typically not unique. It will be unique if we have Zyskind structure and if sp**D** is a quadratic subspace.

We are now in a position to define an ANOVA test. We define the test as follows: Let

$R_- = -$ [sum of $g_iR_i$ such that $g_i$ is negative]

$R_+ =$ [sum of $g_iR_i$ such that $g_i$ is positive]

Set $T_A = [Q_H + R_-]/R_+$. The ANOVA test is then based on $T_A$. The distribution of the test statistic is not easy to find. As a result, a critical region $T_A \geq C_\alpha$ is typically obtained by choosing $C_\alpha$ in some approximate manner such as the Satterthwaite method.

Up to the present, we have not utilized the complete Zyskind structure. From here on, however, we do need the entire Zyskind structure. In addition, we assume the following is satisfied in the remainder of our results:

(A10) $\underline{R}(M) \subset \underline{R}(E_s)$.

The subscript s is used here for convenience. The basic requirement we need is that $\underline{R}(M)$ is contained in $\underline{R}(E_v)$ for some v such that $F_v$ is nonzero. Because of (A10) we have $b'\theta = \delta_s$. So, $g'R = R_s$ is one choice to use in the construction of the ANOVA test. Hereafter, we assume this choice. Thus, the ANOVA test statistic is

$$F_A = f_s Y'MY \,/\, rR_s \quad .$$

By Lemma 4.5.1, $F_A$ has an F-distribution with r and $f_s$ degrees of freedom and noncentrality parameter $\beta'X'E_sX\beta/\delta_s$.

<u>Lemma</u> <u>5.4.2</u>. The test statistic $F_A$ provides an unbiased test for H: $L'\beta = 0$.

<u>Proof</u>: Straightforward.[]

<u>Lemma</u> <u>5.4.3</u>. Assume (A10) and s = k. Let $R_A$ denote the random vector consisting of $Y'MY + R_s$, $R_0, \ldots, R_{s-1}$. Then under the null hypothesis, $R_A$ is a complete sufficient statistic for the family of distributions induced by $R_M = (R', Y'MY)'$.

<u>Proof</u>: Under the null hypothesis, $R_s/\delta_s$ and $Y'MY/\delta_s$ both have central chi-squared distributions. Because $\delta_s$ is common to both, it is easily seen that $R_A$ is sufficient by writing out the density of $R_M$. The completeness follows because when s = k, the vector $\delta$ is a nonsingular linear transformation of $\boldsymbol{\theta}$ so that an open set of $\boldsymbol{\theta}$ vectors is mapped to an open set of

$\delta$ vectors.[]

Theorem 5.4.4. Assume (A10) and $s = k$. Then the ANOVA test based on $F_A$ is a Uniformaly Most Powerful Invariant Unbiased test where invariance is with respect to the group **G** discussed in Section 3.

Proof: By using (A10), it is easy to check that $P_0$ commutes with each $H_i$ which implies that assumption (A9) is satisfied. Thus, Theorem 5.3.7 says T of Section 3 is a maximal invariant for the group **G**. But (A10) implies T and $R_M$ defined in Lemma 5.4.3 are the same statistic. Thus, Lemma 5.4.3 implies $R_A$ is a complete sufficient statistic (for the family of the maximal invariant) under the null hypothesis. Hence every invariant unbiased test has Neyman structure with respect to $R_A$ (see Lehmann Chapter 4). Using Basu's theorem $F_A$ is independent of $R_A$ because $F_A$ has an F-distribution with w and $f_s$ degrees of freedom. Because of the monotone likelihood of the F-distribution, the ANOVA test is the best test in this family.[]

# VI. BARTLETT-SCHEFFÉ TESTS

The Bartlett-Scheffé procedure is a method to get unbiased tests by constructing two independent normal variates with the same variance and with the appropriate expectations for testing the hypothesis. This procedure is an analogue to the construction of unbiased tests for the Behrens-Fisher problem by Bartlett (1943) and Scheffé (1943). Such tests are called tests of Bartlett-Scheffé type.

A Bartlett test deals with the case of two independent samples of equal size: $(X_1, \ldots, X_n)$, $X_i \sim N(a_1, \alpha_1^2)$ and $(Y_1, \ldots, Y_n)$, $Y_i \sim N(a_2, \alpha_2^2)$. To test $a_1 = a_2$, Bartlett's test statistic is of the form

$$\frac{[n(n-1)]^{1/2} \, |\bar{X}-\bar{Y}|}{\sum_i [(X_i-\bar{X}) - (Y_i-\bar{Y})]^2} .$$

Scheffé's test does not require equality of the sample sizes. Both tests are based on the t-distribution.

Imhof (1960) suggested the use of Bartlett-Scheffé tests for testing fixed effects in a three-way classification model with two factors random and one fixed: 'The fact that several unknown covariance matrices are involved in our model suggests trying to apply a device similar to the one proposed by Scheffé for solving the Behrens-Fisher problem.' Mazuy and Connor (1965) used the Bartlett-Scheffé procedure in the case of a two-way classification model with unequal variances. Linnik (1968) studied Bartlett-Scheffé tests from a theoretical point of view and gave some properties of these tests. Siefert (1979, 1981) developed tests of the Bartlett-Scheffé type in general balanced mixed classification models and discussed some of their properties.

In this chapter we give a general definition of a Bartlett-Scheffé test statistic, a general method of constructing such tests and properties of the resulting tests. A test statistic for $E_iX\beta$ and for $a'\beta$ under models satisfying Zyskind structure will be constructed and their properties will be given.

## 6.1. Definition and Examples

In this section a definition of a Bartlett-Scheffé test statistic as well as some examples will be given.

Definition 6.1.1. $T$ is a Bartlett-Scheffé test statistic for testing $L'\beta = 0$ if

$$T = bY'A'AY/aY'B'BY$$

where

(a) $AY \sim N_a(AX\beta, \gamma(\theta) I_a)$

(b) $BY \sim N_b(0, \gamma(\theta) I_b)$

(c) $Cov(AY, BY) = 0$

(d) $AX\beta = 0$ if and only if $L'\beta = 0$.

It follows that $T$ has a noncentral F-distribution with a, b degrees of freedom and noncentrality parameter $\lambda = \beta'X'A'AX\beta/\gamma(\theta)$.

Remark 6.1.2. Suppose $Y'GY/\gamma(\theta)$ is a quadratic form which has a noncentral chi-squared distribution with a noncentrality parameter which is zero if and only if $L'\beta = 0$. And suppose $Y'HY/\gamma(\theta)$ is a quadratic form, independent of $Y'GY$, which has a central chi-squared distribution. Then we can find A and B satisfying Definition 6.1.1 such that $Y'GY = Y'A'AY$ and $Y'HY = Y'B'BY$.

Example 6.1.3. Let $Y \sim N_4(E(Y), Cov(Y))$ where $Y = (X_1,X_2,Y_1,Y_2)'$, $E(Y) = (\mu_1,\mu_1,\mu_2,\mu_2)'$ and $Cov(Y) = diag(\theta_1,\theta_1,\theta_2,\theta_2)$. Suppose we want to test $\mu_1 - \mu_2 = 0$.

Let $U = \bar{X} - \bar{Y}$. Then $E(U) = \mu_1 - \mu_2$ and $Var(U) = \sigma_U^2 = (1/2)(\theta_1 + \theta_2)$. Note $U$ is a linear function of $Y$, $U = a'Y$ where $a = (1/2)(1\ 1\ -1\ -1)'$. Let $Z = (1/2)((X_1-X_2) - (Y_1-Y_2))$. Then $E(Z) = 0$ and $Var(Z) = \sigma_U^2$. It is clear that $Z$ also is a linear function of $Y$, $Z = b'Y$ where $b = (1/2)(1\ -1\ 1\ -1)'$.

Now we have $U \sim N(\mu_1-\mu_2, \sigma_U^2)$ and $Z \sim N(0, \sigma_U^2)$. Also note that $U$ and $Z$ are independent, so $T = U^2/Z^2$ is a Bartlett-Scheffé test statistic for testing $\mu_1 - \mu_2 = 0$. In fact, the square root of $T$ is the same test statistic which Bartlett (1943) constructed for the Behrens-Fisher problem in the case $n_1 = n_2 = 2$.

Another Bartlett-Scheffé test statistic can be constructed by letting $Z_1 = (1/2)((X_1-X_2) + (Y_1-Y_2))$. Note $E(Z_1) = 0$, $Var(Z_1) = \sigma_U^2$, and $U$ and $Z_1$ are independent, so $T_1 = U^2/Z_1^2$

is another Bartlett-Scheffé test statistic for testing the same hypothesis $\mu_1 - \mu_2 = 0$.

Example 6.1.4. Consider Example 5.2.1. We want to test $\beta_2 = 0$. A Bartlett-Scheffé test statistic is $T = 2U^2/Z'Z$ where $U = \bar{Y}_2$ and $Z = (Z_1, Z_2)$, $Z_1 = 1/2(Y_{21}-Y_{22})$ and $Z_2 = 1/2(Y_{22}-Y_{21})$.

A linear function $a'Y$ of $Y$ such that $E(a'Y) = 0$ is called an error contrast. (Strictly speaking, it is truly a 'contrast' only if the vector $1 \varepsilon \underline{R}(X)$, but this is usually the case anyway.) The denominator of a Bartlett-Scheffé test statistic is formed from a vector of error contrasts. Sometimes it is advantageous to put error contrasts in the numerator too.

## 6.2. Using Error Contrasts in the Numerator

Suppose we have random vectors $\hat{\eta}$, $Z_1$, ..., $Z_r$ such that:

(a) $\hat{\eta} \sim N_t(\eta, \gamma(\theta) I_t)$ , $Z_j \sim N_{dj}(0, \gamma_j(\theta) I_{dj})$ for $j = 1, \ldots, r$

where $\eta$ and $\theta$ are unknown parameters.

(b) $\gamma(\theta) = \sum_{j=1}^{r} g_j \gamma_j(\theta)$ for known scalars $g_j$.

(c) $\hat{\eta}$, $Z_1$, ..., $Z_r$ are jointly normal and independent.

Define $Z = (Z_1', \ldots, Z_r')'$. Note that under assumptions (A1) - (A8), for any estimable linear parameter function $\eta = a'\beta$ ($t = 1$), we can find $\hat{\eta}$ and $Z$ satisfying the conditions above (see Section 6.4). And for certain estimable linear parameter vectors $\eta = L'\beta$ with $t > 1$ (see Section 6.3), $\hat{\eta}$ and $Z$ can be found.

Given $\hat{\eta}$ and Z satisfing (a) - (d), we can construct a Bartlett-Scheffé test of $\eta = 0$ as follows. Let

$J_- = \{j : g_j$ is negative$\}$

$J_+ = \{j : g_j$ is positive$\}$

$\gamma_-(\theta)$ = the sum of $g_j\delta_j(\theta)$ for $j \varepsilon J_-$

$\gamma_+(\theta)$ = the sum of $g_j\delta_j(\theta)$ for $j \varepsilon J_+$.

Note $\gamma_+(\theta) = \gamma(\theta) - \gamma_-(\theta)$.

We must now assume $t \leq \min\{d_j : j \varepsilon J_-\}$. From each of the vectors $Z_j$, $j \varepsilon J_-$, pick $t$ components to form a $t \times 1$ vector $\bar{Z}_j$. Define $CZ = \sum_- s_j\bar{Z}_j$ where $s_j$ denotes the square root of $|g_j|$ and $\sum_-$ denotes summation over $j \varepsilon J_-$. Its covariance matrix is $\mathrm{Cov}(CZ) = \gamma_-(\theta)I_t$. In the same way, using the vectors $Z_j$, $j \varepsilon J_+$, we can form DZ such that $\mathrm{Cov}(DZ) = \gamma_+(\theta)I_f$ where $f = \min\{d_j : j \varepsilon J_+\}$. Let $U = \eta + CZ$. It is easy to see that $E(U) = \eta$ and that U and DZ are independent. Then

(6.2.1) $\quad T = fU'U/tZ'D'DZ$

is a Bartlett-Scheffé test statistic and $T \sim F_{(t,f,\lambda)}$ where $\lambda = \eta'\eta/\gamma_+(\theta)$.

<u>Lemma</u> <u>6.2.2</u>. Assume that $\gamma_1(\theta), \ldots, \gamma_r(\theta)$ are linearly independent functions of $\theta$. If B is a $q \times d$ matrix, $d = \sum_j d_j$, such that $\mathrm{Cov}(BZ) = c(\theta)I_q$ where $c(\theta) = \sum_j b_j\gamma_j(\theta)$, $b_j \geq 0$, then $q \leq \min\{d_j : b_j > 0\}$.

<u>Proof</u>: We can write $BZ = \sum_j B_jZ_j$, $B_j$ a $q \times d_j$ matrix. Then $[\sum_j b_j\gamma_j(\theta)]I_q = \mathrm{Cov}(BZ) = \sum_j \mathrm{Cov}(B_jZ_j) = \sum_j \gamma_j(\theta)B_jB_j'$. By the

independence of the functions $\gamma_j(\theta)$, we must have $b_j I_q = B_j B_j'$. If $b_j > 0$, then $\underline{r}(B_j) = \underline{r}(B_j B_j') = q$. Hence $q \leq d_j$. []

<u>Theorem</u> <u>6.2.3</u>. Let $T_1 = bY'A'AY/aY'B'BY$ be a test statistic of Bartlett-Scheffé type for testing $L'\beta = 0$. Suppose A is written as $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ where $A_2 X = 0$. Then a better Bartlett-Scheffé test statistic for testing $L'\beta = 0$ is

$$T_2 = (b+a-t)Y'A_1'A_1Y/t(Y'B'BY + Y'A_2'A_2Y)$$

$t = \underline{r}(A_1)$.

<u>Proof</u>: $\mathrm{Cov}(AY) = \mathrm{Cov} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} Y = c(\theta) \begin{pmatrix} I_t & \\ & I_{a-t} \end{pmatrix}$ and $\mathrm{Cov}(BY) = c(\theta) I_b$.

(a) $t \leq a$ since $a = r(A) = r(A') = r(A_1') + r(A_2') - \dim \underline{R}(A_1') \cap \underline{R}(A_2') = t + c$, $c \geq 0$

(b) $(a-t) + b \geq b$, since $(a-t) \geq 0$

(c) $$||AX\beta||^2 = \left[\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} X\beta\right]' \left[\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} X\beta\right] = \begin{pmatrix} A_1X\beta \\ 0 \end{pmatrix}' \begin{pmatrix} A_1X\beta \\ 0 \end{pmatrix}$$

$$= (A_1X\beta)'A_1X\beta = ||A_1X\beta||^2, \text{ so}$$

$$\lambda_{T_1} = \lambda_{T_2}.$$

From (a), (b) and (c) and the montonicity of the power of F-test the assertion follows. []

<u>Theorem</u> <u>6.2.4</u>. If $\hat{\eta}$ is BLUE for $\eta$ and $\gamma(\theta)$ uniquely expressed as $\gamma(\theta) = g_1\delta_1(\theta) + \ldots + g_r\delta_r(\theta)$. Then T defined

in (6.2.1) is UMP test in the class of Bartlett-Scheffé tests.

Proof: Let $b\, Y'A'AY/a\, Y'B'BY$ be a Bartlett-Scheffé test for testing $\eta = L'\beta = 0$. So we have

(a) $r(A) = a$ , $r(B) = b$

(b) $E(AY) = H\beta$ , $\underline{r}(H) = a$ and $H'\beta = 0 \Leftrightarrow L'\beta = 0$.

This implies that $\underline{R}(L) = R(H)$ and $a = t$.

Let $TY$ be BLUE for $L'\beta$, so $E(TY) = TX\beta$ and $L = X'T'$. Then $C'TY$ is a BLUE for $H'\beta$ and $A$ can be written as $A = C'T + A_2$, where $A_2X = 0$. Now $Cov(AY) = Cov(C'TY) + Cov(A_2Y)$ $= \gamma(\theta)C'C + \sum_{j=1}^{r} \delta_j(\theta)R_jR_j'$. Note $A_2Y = RZ = \sum_{o=1}^{r} R_iZ_i$.

Now using the fact that $Y'A'AY/c(\theta) \sim \chi^2_{(t,\lambda)}$, then $t\, c(\theta) = \gamma(\theta)\, tr\, C'C + \sum\delta_j(\theta)\, tr\, R_jR_j'$ and so $c(\theta) = c\gamma(\theta) + \sum h_j\, \delta_j(\theta)$ , $h_j \geq 0$ , $c > 0$ . In the same way $Cov(BY) = \sum b_j\, \delta_j(\theta)\, I_b$ , $b_j \geq 0$, so $c\, \gamma(\theta) + \sum h_j\, \delta_j(\theta) = \sum b_j\, \delta_j(\theta)$, hence $\gamma(\theta) = \sum_j \frac{b_j - h_j}{c}\, \delta_j(\theta) = \sum_j g_j\, \delta_j(\theta)$.

Now by uniqueness $g_j = \frac{b_j - h_j}{c}$ and since now $b_j > 0$ for all $g_j > 0$, hence $\{j : g_j > 0\} \subset \{j : b_j > 0\}$ and $\min\{f_j : b_j > 0\} \leq \min\{f_j : g_j > 0\} = f$. So by Lemma 6.2.2, $b \leq f$.

Now $\lambda_T = \beta'\eta\eta'\beta \,/\, \gamma_+(\theta)$ and $\lambda = \beta'H'H\beta \,/\, \sum_j b_j\, \delta_j(\theta)$.

So

$$\lambda_T - \lambda = \beta'L \left[ \frac{I_t}{\gamma_+(\theta)} - \frac{CC'}{\sum_j b_j \delta_j(\theta)} \right] L'\beta$$

So

$$\gamma_+(\theta) \sum_j b_j \delta_j(\theta) \ [\lambda_T - \lambda] = \beta' L[\sum b_j \delta_j(\theta) I_t - \gamma_p(\theta) CC'] L' \beta$$

But $\gamma(\theta) C'C + \sum \delta_j(\theta) R_j R_j' =$

$$\sum g_j \delta_j(\theta) C'C + \sum \delta_j(\theta) R_j R_j' =$$

$$\sum \delta_j(\theta)[g_j C'C + R_j R_j'] =$$

$$\sum b_j \delta_j(\theta) I_a.$$

Then for all $j$, $g_j C'C \le b_j I$ which implies that $\sum g_j \delta_j(\theta) C'C \le \sum \delta_j(\theta) b_j I_t$. Hence $\lambda_T - \lambda \ge 0$.

So we have $a = t$, $b \le f$ and $\lambda_T \ge \lambda$, then we are done.[]

## <u>6.3. Test $E_i X\beta = 0$ for $i \varepsilon I$</u>

Assume that we have Zyskind structure, let us use the same notation as in Section 4.3 and we want to test $E_i X\beta = 0$, $i \varepsilon I$. Recall

$$E_i Y \sim N_n(E_i X\beta, \delta_i E_i), \ i \varepsilon I$$

$$F_j Y \sim N_n(0, \delta_j F_j), \ j \varepsilon J.$$

Consider testing $E_1 X\beta = 0$. Suppose $\delta_1 = \sum_{j \varepsilon J} g_j \delta_j$. Define $\delta_{1-} = \sum_{g_j < 0} g_j \delta_j$ and $\delta_{1+} = \sum_{g_j > 0} g_j \delta_j$. Assume $e_1 \le \min\{f_j = g_j < 0\}$, also let $D_1$ be an $e_1 \times n$ matrix such that $D_1 E_1 D_1' = I_{e_1}$, so

$$D_1 E_1 Y \sim N_{r1}(D_1 E_1 X\beta, \delta_1 I_{r1}),$$

<u>Lemma</u> <u>6.3.1</u>. $E_1 X\beta = 0 \iff ||E_1 X\beta||^2 \iff ||D_1 EX\beta||^2$

<u>Proof</u>: We only need to show that $E_1 = E_1 D_1' D_1 E_1$. We have $D_1 E_1 D_1' = I_{e_1}$, $E_1 D_1' D_1 E_1$ is n.n.d. matrix, so $(E_1 D_1' D_1 E)^2 = E_1 D_1' D_1 E_1$

which implies that $E_1D_1'D_1E_1$ is orthogonal projection on $\underline{R}(E_1D_1'D_1E_1) \subset \underline{R}(E_1)$. But $e_1 = \underline{r}(I_{e_1}) = \underline{r}(D_1E_1D') = r(E_1D') \leq r(E_1) = e_1$, hence $r(E_1) = r(E_1D_1'D_1E_1)$.

Then $E_1D_1'D_1E_1$ is orthogonal projection on $R(E_1)$, so $E_1 = E_1D_1'D_1E_1$. []

Let $C_j$ , $j \in J$ be $f_j$ x n matrices such that $C_jF_jC' = I_{fj}$ then

$$C_jF_jY \sim N_{fj}\ (0,\ \delta_j I_{fj}).$$

By Lemma 6.3.1 test $E_1X\beta = 0$ is equivalent to test $D_1E_1X\beta = 0$. To construct a Bartlett-Scheffé test we proceed as follows:

(a) Let $A_1$ be $e_1$ x n matrix such that $A_1A_1' = I_{r_1}$ and define $L_1 = A_1 \sum_{g_j<0} |g_j| C_jF_jY$. Note $E(L_1) = 0$ and $Cov(L_1) = \delta_{1-} I_{r_1}$, one choice of $C_1$ is $C_1 = (I_{r_1}\ 0)$, so

$L_1 \sim N_{r_1}(\delta_{1-}I_{r_1})$.

(b) Define $R = D_1E_1Y - L_1$, so $E(R) = D_1E_1X\beta$ , $Cov(R) = \delta_{1+}I_{r_1}$.

(c) Let $f = \min\{f_j : g_j > 0\}$ , let $C_2$ be an fxn matrix such that

$S = C_2 \sum_{g_j>0} g_j\ C_jF_jY$ , so $E(S) = 0$ ,

$Cov(S) = \delta_{1+}I_f$ and $S \sim N_f(0,\ \delta_{1+}I_f)$ , one choice of $C_2$

is $C_2 = (I_f\ 0)$. It is easy to see that R and S independent.

Now define

(6.3.2) $T = f_1R'R/e_1S'S$

is a Bartlett-Scheffé test and $T \sim F_{(r_1, f, \lambda_T)}$ where $\lambda_T = \beta'X'E_1X\beta/\delta_{1+}$.

<u>Theoem</u> <u>6.3.3</u>. If there is a complete sufficient statistic in the model, then test (6.3.2) is a UMP-test in the class of all tests of Bartlett-Scheffé type.

<u>Proof</u>: $E_1Y$ is BLUE for $E_1X\beta$ and completeness insure that $\delta_1$ has unique representation as a function of $\delta_j$'s. Then by Theorem 6.2.2 the results follows.[]

<u>Example</u> <u>6.3.4</u>. Consider the following model

$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk}$, $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, c$.

We assume $\mu$ and $\alpha_i$ are fixed while $\beta_j, \ldots, (\alpha\beta\gamma)_{ijk}$ are assumed to be independent and normally distributed with zero mean and variances $\sigma^2_\beta, \ldots, \sigma^2_{\alpha\beta\gamma}$ respectively. In matrix notation we have

$Y = 1_n u + A\alpha + B\beta + U_{AB}(\alpha\beta) + C\gamma + V_{AC}(\alpha\gamma) + W_{BC}(\beta\gamma) + I(\alpha\beta\gamma)$.

Hence $E(Y) = 1_n u + A\alpha$ and

$Cov(Y) = \theta_0 + \theta_1 P_B + \theta_2 P_{AB} + \theta_3 P_C + \theta_4 P_{AC}$

where $\theta_0 = \alpha^2$, $\theta_1 = ac\sigma^2_\beta$, $\theta_2 = c\sigma^2_{\alpha\beta}$, $\theta_3 = ab\sigma^2_\gamma$, $\theta_3 = b\sigma^2_{\alpha\gamma}$, $\theta_5 = a\sigma^2_{\beta\gamma}$ and we have written $P_{AB}$, for example, to denote the orthogonal projection on $\underline{R}(U_{AB})$ for convenience.

| $i$ | $\delta_i(\theta)$ | $H_i$ |
|---|---|---|
| 0 | $\theta_0$ | $I - P_{AB} - P_{AC} - P_{BC} + P_A + P_B + P_C - P_1$ |
| 1 | $\theta_0 + \theta_2 + \theta_4$ | $P_A - P_1$ |
| 2 | $\theta_0+\theta_1+\theta_2+\theta_5$ | $P_B - P_1$ |
| 3 | $\theta_0 + \theta_2$ | $P_{AB} - P_A - P_B + P_1$ |
| 4 | $\theta_0 + \theta_3+\theta_4+\theta_5$ | $P_C - P_1$ |
| 5 | $\theta_0 + \theta_4$ | $P_{AC} - P_A - P_C + P_1$ |
| 6 | $\theta_0 + \theta_5$ | $P_{BC} - P_B - P_C + P_1$ |
| 7 | $\theta_0+\theta_1+\theta_2+\theta_3+\theta_4+\theta_5$ | $P_1$ |

Test $\alpha = 0$. It is clear that $E_1 = P_A - P_1$ and $F_1 = 0$. Testing $\alpha = 0$ is equivalent to test $E_1X\beta = 0$, we have $E_1Y = (P_A - P_1)Y \sim N_n(A\alpha, \delta_1(P_A - P_1))$ and $\delta_1 = \delta_3 + \delta_5 - \delta_0$.

Note $F_0 = H_0$, $f_0 = (a-1)(b-1)(c-1)$, $F_3 = H_3$, $f_3 = (a-1)(b-1)$ and $F_5 = H_5$, $f_5 = (a-1)(c-1)$. Find $C_i$, $i = 0, 3, 5$ such that $C_iF_iY \sim N_{fi}(0, I_{fi})$, also find $D_1$ such that $D_1E_1Y \sim N_{r1}(D_1E_1X\beta, \delta_1 I_{r1})$, $r_1 = a-1$, so we have

(a) $R = D_1E_1Y - A_1C_0F_0Y \sim N_{r1}(D_1E_1X\beta, (\delta_1+\delta_0)I_{r1})$

(b) Let $f = \min\{f_3, f_5\}$, so

$$S = A_2(C_3F_3Y + C_5F_5Y) \sim N_f(0, (\delta_1+\delta_0)I_f)$$

(c) $T = fR'R/e_1S'S$ is a Bartlett-Scheffé test for testing $A\alpha = 0$. It is a UMP-test in this class.

## 6.4. Test $\underline{a}'\underline{\beta} = \underline{0}$

Let us assume that we have the same structure and notation as Section 6.3. Consider the problem is to test $a'\beta = 0$ where $a \varepsilon \underline{R}(X') = \sum \underline{R}(X'E_i)$. For simplicity let $a \varepsilon \underline{R}(X'E_1) + \underline{R}(X'E_2)$. This implies that $a = X't_1 + X't_2$, where $t_1 \varepsilon \underline{R}(E_1)$ and $t_2 = \underline{R}(E_2)$. Note $t'Y$ is BLUE for $a'\beta$ where $t = t_1 + t_2$, so $Cov(t'Y) = c_1\delta_1 + c_2\delta_2$ where $c_1 = t_1't_1$ and $c_2 = t_2't_2$.

Write $\delta_1 = \delta_{1+} + \delta_{1-}$ and, similarly, $\delta_2 = \delta_{2+} + \delta_{2-}$, so $Cov\ t'Y = c_1\delta_{1+} + c_2\delta_{2+} + c_1\delta_{1-} + c_2\delta_{2-} = \delta_+ + \delta_-$, where

$$\delta_+ = c_1\delta_{1+} + c_2\delta_{2+} = \sum b_j\delta_j \ , \ b_j > 0$$
$$\delta_- = c_1\delta_{1-} + c_2\delta_{2-} = \sum b_j\delta_j \ , \ b_j < 0.$$

Now to construct a Bartlett-Scheffé test follow the procedure in Section 6.3. We can find

(a) $R$ such that $E(R) = a'\beta$, $Var(R) = \delta_+$

(b) $S$ such that $E(S) = 0$, $Var(S) = \delta_+ I_f$ where $f = \{\min f_j : b_j > 0\}$, and $S$ , $L$ are independent.

So we can define

(6.4.1) $T = fR^2/S^2$ is a Bartlett-Scheffé test, $T$ is distributed as $F_{(1,f,\lambda)}$ , $\lambda = (a'\beta)^2/\delta_+$ .

<u>Theorem</u> <u>6.4.2</u>. If there is a complete sufficient statistic in the model test (6.4.1) is a UMP-test in the class of Bartlett-Scheffé type.

<u>Proof</u>: Let $T_1 = f_1(m'y)^2/(BY)'(BY)$ be a Bartlett-Scheffé test,

where $E(m'y) = a'\beta$, $f_1 = \underline{r}(B)$, and $Cov(m'y) = \delta_+ + \delta_- + \sum r_j \delta_j = \sum \delta_j k_j$, $r_j \geq 0$ and $k_j \geq 0$, since $m'y$ can be written as $m'y = t'y + h'y$, where $h \varepsilon \underline{R}(X)$ . Now $r_j \geq 0$ for all $j$, so for $b_j > 0$, $k_j \geq b_j$, hence $\sum_{bj>0} b_j \delta_j \leq \sum k_j \delta_j$, then $Var(m'y) \geq Var(R)$, which implies that $\lambda \geq \lambda_{T1} = (\lambda'\beta)^2 / \sum k_j \delta_j$. Also $\{j : b_j > 0\} \subset \{j : k_j > 0\}$ implies that $\min\{f_j : b_j > 0\} \geq \min\{f_j : k_j > 0\}$ so $f \geq f_1$. Now because of the montonicity of the power of the test the result follows. []

Example 6.4.3. Consider Example 4.1.5, we want to test $\mu + \alpha_1 = 0$.

$T = 2\, U^2 / Z'Z$ is UMP-test in the class of Bartlett-Scheffé test, where $U = \bar{Y}_{1\cdot}$ and $Z = (Y_{11}-Y_{12}\ ,\ Y_{31}-Y_{32})'$.

# VII. BIBLIOGRAPHY


Albert, A. 1976. When is a sum of squares an analysis of variance? Ann. Statis. 4:775-778.

Anderson, T. W. 1969. Statistical inference for covariance matrices with linear structure. Multivariate Analysis II, (P. Krishnaiah, Ed.). Academic Press, New York. 55-66.

Bellman, R. 1960. Introduction to Matrix Analysis. McGraw Hill, New York.

Brown, K. 1983. Subbalanced data and the mixed analysis of variance. J. Amer. Statist. Assoc. 78:162-167.

Brown, K. 1984. On analysis of variance in the mixed model. Ann. Statist. 12:1488-1499.

Cox, D. R. and Hinkley, D. V. 1974. Theoretical Statistics. Chapman and Hall, London.

El-Bassiouni, Y. 1978. Hypothesis Testing for the Parameters of a Covariance Matrix having Linear Structure. Ph.D. Thesis. Oregon State University.

El-Bassiouni, Y. and Seely, J. 1980. Optimal tests for certain functions of the parameters in a covariance matrix with linear structure. Sankhya (Series A). 42:64-77.

Graybill, F. and Hultquist, R. A. 1961. Theorems concerning Eisenhart's model II. Ann. Math. Statist. 32:261-269.

Harville, D. A. 1985. A generalized version of Albert's theorem with applications to the mixed linear models. Experimental Design, Statistical Models and Genetic Statistics. Ed. K. Hinkelmann. Marcel Dekker, NY. 231-238.

Hultquist, R. A. and Atzinger, E. M. 1972. The mixed effects model and simultaneous diagonalization of symmetric matrices. Ann. Math. Statist. 43:2023-2030.

Imhof, J. P. 1960. A mixed model for the complete three-way layout with two random effects factors. Ann. Math. Statist. 31:906-928.

Lehmann, E. L. 1959. Testing Statistical Hypotheses. Wiley, New York.

Linnik, J. V. 1968. Statistical Problems with Nuisance Parameters. American Mathematical Society, Providence, Rhode Island.

Mazuy, K. and Connor. S. 1965. Student's t in a two-way classification with unequal variances. Ann. Math. Statist. 36:1248-1255.

Muir, T. 1960. The Theory of Determinants in the Historial Order of Development. Dover Publications, Inc., New York.

Olsen, A., Seely, J. and Birkes, D. 1976. Invariant quadratic unbiased estimation for two variance components. Ann. Statist. 4:878-890.

Rao, C. R. and Mitra, S. K. 1971. Generalized Inverse of Matrices and its Applications. Wiley, New York.

Rogers, G. S. and Young, D. L. 1975. Some likelihood ratio tests when a normal covariance matrix has certain reducible linear structure. Comm. Statist. 4:537-554.

Scheffé, H. 1934. On solutions of the Behrens-Fisher problem based on the t-distribution. Ann. Math. Statist. 14:35-44.

Scheffé, H. 1956. A mixed model for the analysis of variance. Ann. Math. Statist. 27:23-36.

Scheffé, H. 1970. Practical solutions to the Behrens-Fisher problems. J. Amer. Statist. Assoc. 65:1501-1508.

Seely, J. 1971. Quadratic subspace and completeness. Ann. Math. Statist. 27:23-36.

Seely, J. 1977. Minimal sufficient statistics and completeness for multivariate normal families. Sankhya (Series A). 39:170-185.

Seely, J. 1979. General Linear Hypothesis Notes, unpublished. Oregon State University.

Seely, J. 1980. Some remarks on exact confidence intervals for positive linear combinations fo variance components. J. Amer. Statist. Assoc. 65:1501-1508.

Siefert, B. 1978. Note on the UMPU. character of a test of the mean in balanced randomized nested classification. Math. Operationsforsch. Statist., Ser. Statistics. 9:185-189.

Siefert, B. 1979. Optimal testing for fixed effects in general balanced mixed classification models. Math. Operationsforsch. Statist., Ser. Statistics. 10:237-255.

Siefert. B. 1981. Explicit formula of exact tests in mixed balanced ANOVA mdoels. Biom J. 23:535-550.

Sirvastava, J. N. 1966. On testing hypotheses regarding a class of covariance structures. Psychometrika. 31:147-164.

Zyskind, G. 1969. Parametric augmentations and error structures under which certain simple least squares and analysis of variance procedures are also best. J. Amer. Statist. Assoc. 64:1353-1368.