AN ABSTRACT OF THE THESIS OF

Mohamed Yahia ElBassiouni for the degree of Doctor of Philosophy

in Statistics presented on July 8, 1977

Title: HYPOTHESIS TESTING FOR THE PARAMETERS OF A COVARIANCE MATRIX HAVING LINEAR STRUCTURE

Abstract approved: Redacted for Privacy

Justus F. Seely

We consider the general family of multivariate normal distributions where the mean vector lies in an arbitrary subspace and the covariance matrix Σ_θ has a linear structure.

When the minimal sufficient statistic is complete, we derive test statistics having optimal properties for testing statistical hypotheses about θ. The important special case where the matrices in the covariance structure commute is emphasized. In this case χ^2-tests and F-tests are obtained. Several examples are worked out to indicate the wide range of applications covered by our investigations.

In the process of constructing similar tests some results concerning the completeness of families of product measures are established. When one family is boundedly complete and the other is strongly complete the resulting product family is boundedly complete. Bounded completeness is strengthened to completeness provided that the strongly complete family is an exponential family.

Hypothesis testing when the minimal sufficient statistic is not complete is also discussed. The case where there are two parameters describing the covariance structure is considered first. Several tests are suggested of which the easiest to carry out is Wald's test. The power function of Wald's test is then studied in an attempt to determine the circumstances under which Wald's test would be a reasonable one. The unbalanced random one-way classification model is worked out as an example of our considerations. Finally, the case of p parameters is briefly considered.

Hypothesis Testing for the Parameters of a Covariance Matrix Having Linear Structure

by

Mohamed Yahia ElBassiouni

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed July 1977

Commencement June 1978

APPROVED:

Redacted for Privacy

Associate Professor of Statistics

in charge of major

Redacted for Privacy

Head of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented July 8, 1977

Typed by Clover Redfern for Mohamed Yahia ElBassiouni

ACKNOWLEDGMENT

The author expresses his great appreciation to Dr. Justus Seely who suggested the initial problem and whose guidance has been very precious all throughout this research. Also, without the background provided by his approach to general linear models this thesis would never have been written.

The author also wishes to thank Dr. Lyle Calvin who supervised the author for the first two years of the author's stay at the Department of Statistics. Besides his constant encouragement and help, Dr. Calvin's expertise in applied statistics had a great impact on the author's educational experience.

The author wishes to express a special note of appreciation to Dr. David Birkes for many stimulating and rewarding conversations during the preparation of this thesis.

Final thanks go to the many members of the Department of Statistics with whom the author has had valuable course work and discussions, especially to Dr. Donald Pierce, Dr. Fred Ramsey and Dr. H. D. Brunk.

TABLE OF CONTENTS

HYPOTHESIS TESTING FOR THE PARAMETERS OF A COVARIANCE MATRIX HAVING LINEAR STRUCTURE

I. INTRODUCTION

This dissertation is concerned with tests of hypotheses about the parameters of the covariance matrix of a multivariate normal distribution. The starting point of the research is a paper by Seely (1976) where the minimal sufficient statistic for a general family of multivariate normal distributions was derived and where necessary and sufficient conditions were given for the minimal sufficient statistic to be complete. Given these general results, it is then natural to consider hypothesis testing problems that arise in such multivariate normal families--both when the minimal sufficient statistic is complete and when it is not complete.

There are two general types of hypothesis testing problems that can arise: problems involving the mean vector and problems involving the covariance matrix. We have concentrated solely on hypothesis testing problems involving the covariance matrix. We have assumed that the covariance matrix has a linear structure. This is essentially no loss in generality since any family of multivariate normal distributions can be reparametrized to have a linear covariance structure. However, we have also assumed that the class of covariance matrices is rich enough so that some "nonempty interior" requirements are met. This should be satisfied by almost any problem of practical interest.

Because of the generality we allow for our model, the work of Scheffé (1956), Herbach (1959), Imhof (1960), Anderson (1971) and Graybill (1976) can be viewed as special cases of our treatment of the case when the sufficient statistic is complete.

In the process of constructing similar tests some results on the completeness of families of product measures are established. They can be regarded as generalizations and refinements of the work of Lehmann and Scheffé (1955), Fraser (1957), Gautschi (1959), Imhof (1960) and Seely (1976).

When the minimal sufficient statistic is not complete, a test procedure due to Wald (1947) appears to be very tractable. Several authors, e.g., Thompson (1955b), Spjotvoll (1967, 1968). Portnoy (1973), Hultquist and Thomas (1975) and Thomsen (1975), have applied it to unbalanced variance-component models. Here again we treat the problem in a more general setting. We consider first the case where there are two parameters describing the covariance structure. The case of more than two parameters is then briefly discussed. In this latter case, Wald's procedure has its limitations in the sense that we may have to assume that certain effects play no role in order to be able to apply it in some hypothesis testing situations.

The tests considered in this thesis are those having optimum properties, e.g., similar tests, unbiased tests and invariant tests. We also give some references to approximate test procedures, e.g., Cochran (1951) and Naik (1974a, b). Likelihood ratio tests, however, are not considered since they are fully discussed by Srivastava (1966), Hartley and Rao (1967) and Jensen (1975). We mention here that even in the balanced random (or mixed) experimental design models the likelihood ratio test is not equivalent to the usual F-test, see Herbach (1959).

Different kinds of examples are given throughout the thesis to illustrate the wide range of applications covered by our methods, e.g., variance-component models, correlation models, the two sample problem, the balanced incomplete block design and models where some of the parameters describing the covariance structure are known to satisfy certain relationships.

The results on the (bounded) completeness of families of product measures, which are needed for the construction of similar tests, are given in Chapter II.

In Chapter III we give the model and develop a minimal sufficient statistic that takes the form (T, R) where T is a vector of linear forms and R is a vector of quadratic forms.

The problem of hypothesis testing is considered in three cases: (a) The family induced by (T, R) is complete. (b) The family induced by R is complete but the one induced by T is not complete. (c) Neither family is complete. Cases (a) and (b) are discussed in Chapter IV where UMPU tests are derived. In case (b) these tests are UMPU among (location) invariant tests. Case (c) is discussed in Chapter VI where we focus on the power function of Wald's test.

In Chapter V we discuss the case where the matrices in the covariance structure commute. In this case the null distributions of the test statistics are greatly simplified, e.g., for testing about ratios of variance components the F-test is obtained. It appears that while completeness buys similarity, it is commutativity which produces simple test procedures.

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 vectors $a, b \in R^n$, $a'b$ and $a \cdot b$ are used interchangeably to denote the usual inner product. For a linear transformation or matrix A, $\underline{R}(A)$, $\underline{N}(A)$, $\underline{r}(A)$, $|A|$ and $tr(A)$ denote the range, null space, rank, determinant and trace of A respectively. $\underline{R}(A)^{\perp}$ is used to denote the orthogonal complement, wrt (with respect to) the usual inner product, of $\underline{R}(A)$. For a function g and for a set B, $g[B]$ denotes the image of B and $g^{-1}[B]$ denotes the inverse image. We denote random variables by capital letters, e.g., Y, T, R; and denote their realizations by small letters, e.g., y, t, r. The family of probability measures induced by a random variable Y is denoted by $\mathcal{P}_Y$. Further, a.e. $[\mathcal{P}]$ written after a statement means that the statement holds except on a set N with $P(N) = 0$, $\forall P \in \mathcal{P}$. For a random variable Y, $f_Y(y|\theta)$ denotes the density function of Y indexed by a parameter θ. Some important density functions are denoted by different symbols. In particular, $N_n(\mu, \Sigma)$ denotes an n-dimensional normal distribution with mean vector μ and covariance matrix Σ and χ_n^2 denotes a chi-square distribution with n degrees of freedom. We use φ to denote the critical function of a given test.

II. COMPLETENESS OF FAMILIES OF PRODUCT MEASURES

In this chapter we investigate completeness for families of product measures. The results are needed for the construction of similar tests in subsequent chapters. We begin by giving a lemma about the joint measurability of a bivariate function. Then we develop the main theorems of the chapter. We conclude by discussing the connection between our results and those already in the literature.

The following notation is used throughout. For $X \subset R^m$, $\beta(X)$ denotes the Borel sets in X. For a bivariate function $g(x,y)$, g_x denotes the x-section of g. Likewise, for a set E of pairs (x,y), E_x denotes the x-section of E. Also $\mathcal{P}_X$ denotes the family of probability measures induced by the random variable X.

<u>Lemma 2.1.</u> Consider two measurable spaces $(\mathcal{X}, \mathcal{A})$ and $(\mathcal{Y}, \beta(\mathcal{Y}))$, where $\mathcal{Y} \subset R^m$. Suppose $g: \mathcal{X} \times \mathcal{Y} \to R$ has the following properties:

(a) $g_x : \mathcal{Y} \to R$ is continuous (wrt the usual metrics on R^m and R) $\forall\ x \in \mathcal{X}$,

(b) $g_y : \mathcal{X} \to R$ is $\mathcal{A}$-measurable $\forall\ y \in \mathcal{Y}$.

Then g is $\mathcal{A} \times \beta(\mathcal{Y})$ measurable.

<u>Pf</u>. For each integer n choose $\{A_{k,n}\}_{k=1}^{\infty} \subset \beta(\mathcal{Y})$ s.t. $\mathcal{Y}$ is the disjoint union $\cup_{k=1}^{\infty} A_{k,n}$ and diameter $(A_{k,n}) \leq 1/n$.

Choose $y_{k,n} \in A_{k,n}$. Define $g_n : \mathcal{X} \times \mathcal{Y} \to R$ by $g_n(x,y) = g(x, y_{k,n})$, if $y \in A_{k,n}$. For $B \in \beta(R)$ we have

$$g_n^{-1}[B] = \cup_{k=1}^{\infty} (g_{y_{k,n}}^{-1}[B] \times A_{k,n}).$$

By (b) we have $g_{y_{k,n}}^{-1}[B] \in \mathcal{A}$ so that $g_n^{-1}[B] \in \mathcal{A} \times \beta(\mathcal{Y})$. Hence g_n is $\mathcal{A} \times \beta(\mathcal{Y})$ measurable. Now fix $y \in \mathcal{Y}$. Note that for each n there is a k s.t. $y \in A_{k,n}$. Hence, $y_{k,n} \to y$ as $n \to \infty$ because diameter $(A_{k,n}) \to 0$. Thus by (a) we see that $g_n(x,y) \to g(x,y)$ as $n \to \infty$. Therefore g is $\mathcal{A} \times \beta(\mathcal{Y})$ measurable. []

We now turn to the main results of the chapter. In the sequel three families of probability measures will play a central role. These families are $\mathcal{P}_T = \{P_{\lambda,\omega} : (\lambda, \omega) \in \Lambda \times \Omega\}$, where Λ and Ω are two sets of parameters, with an associated measurable space $(\mathcal{J}, \mathcal{B})$, $\mathcal{P}_S = \{Q_\omega : \omega \in \Omega\}$ with an associated measurable space $(\mathcal{S}, \mathcal{C})$, and $\mathcal{P}_{(T,S)} = \{P_{\lambda,\omega} \times Q_\omega : (\lambda, \omega) \in \Lambda \times \Omega\}$. We shall investigate the completeness of the family $\mathcal{P}_{(T,S)}$. Let $\Omega \subset R^m$ and let L denote Lebesgue measure on $\beta(\Omega)$. If the condition $\int_{\mathcal{S}} f(s) dQ_\omega(s) = 0$, a.e. $[L]$, implies that $f(s) = 0$, a.e. $[\mathcal{P}_S]$, then the family $\mathcal{P}_S$ is said to be strongly complete wrt L. We assume also that $\mathcal{P}_S$ is dominated by a σ-finite measure μ. Let

$(dQ_\omega/d\mu)(s) = q(s,\omega)$, so $q(s,\omega)$ is a particular Radon-Nikodym derivative. We have the following

<u>Theorem 2.2.</u> Suppose the following conditions are satisfied:

(a) the family $\{P_{\lambda,\omega} : \lambda \in \Lambda\}$ is boundedly complete $\forall\ \omega \in \Omega$,

(b) the members of the family $\{P_{\lambda,\omega} : \omega \in \Omega\}$ are all equivalent (i.e., they share the same null sets) $\forall\ \lambda \in \Lambda$,

(c) the family $\mathcal{P}_S$ is strongly complete wrt L,

(d) $q(s,\omega)$ is $\mathcal{C} \times \beta(\Omega)$ measurable.

Then the family $\mathcal{P}_{(T,S)}$ is boundedly complete.

<u>Pf</u>. Consider

$$I(\lambda,\omega) = \int_{\mathcal{J}\times\mathcal{S}} h(t,s)d(P_{\lambda,\omega} \times Q_\omega)(t,s),$$

where h is any real bounded statistic. We want to show that $I(\lambda,\omega) = 0$ $\forall (\lambda,\omega) \in \Lambda \times \Omega \Rightarrow h(t,s) = 0$, a.e. $[\mathcal{P}_{(T,S)}]$. For each $(t,\omega) \in \mathcal{J} \times \Omega$, define

$$F(t,\omega) = \int_{\mathcal{S}} h(t,s)dQ_\omega(s) = \int_{\mathcal{S}} h(t,s)q(s,\omega)d\mu(s) .$$

Since h is bounded, the integral always exists, so F is well defined. Further, $h(t,s)q(s,\omega)$ is $\mathcal{B} \times \mathcal{C} \times \beta(\Omega)$ measurable. So Tonelli's theorem, see Bartle (1966), page 118, applied to the positive and the negative parts of h implies that F is $\mathcal{B} \times \beta(\Omega)$-measurable. Moreover, by Fubini's theorem

$$(2.3) \qquad I(\lambda,\omega) = \int_{\mathcal{J}} F(t,\omega)dP_{\lambda,\omega}(t) .$$

Suppose $I(\lambda,\omega) = 0, \ \forall \ (\lambda,\omega) \in \Lambda \times \Omega$. Then for $\omega \in \Omega$

$$\int_{\mathcal{J}} F(t,\omega)dP_{\lambda,\omega}(t) = 0, \quad \forall \ \lambda \in \Lambda .$$

Hence by (a) $\exists \ N_\omega \in \mathcal{B}$ s.t. $P_{\lambda,\omega}(N_\omega) = 0, \ \forall \ \lambda \in \Lambda$, and $F(t,\omega) = 0, \ \forall \ t \notin N_\omega$. Let $G = \{(t,\omega) : F(t,\omega) \neq 0\}$, $G_t = \{\omega : (t,\omega) \in G\}$ and $G_\omega = \{t : (t,\omega) \in G\}$. Since F is $\mathcal{B} \times \beta(\Omega)$ measurable, then $G \in \mathcal{B} \times \beta(\Omega) \Rightarrow G_t \in \beta(\Omega)$ and $G_\omega \in \mathcal{B}$. But $t \notin N_\omega \Rightarrow t \notin G_\omega \Rightarrow G_\omega \subset N_\omega \Rightarrow P_{\lambda,\omega}(G_\omega) = 0, \ \forall \ (\lambda,\omega) \in \Lambda \times \Omega$. Now for $(\lambda,\omega) \in \Lambda \times \Omega$, $P_{\lambda,\omega}(G_\gamma) = P_{\lambda,\gamma}(G_\gamma) = 0, \ \forall \ \gamma \in \Omega$, so $(P_{\lambda,\omega} \times L)(G) = 0 \Rightarrow \exists \ M_{\lambda,\omega} \in \mathcal{B}$ s.t. $P_{\lambda,\omega}(M_{\lambda,\omega}) = 0$, and $L(G_t) = 0, \ \forall \ t \notin M_{\lambda,\omega}$. Fix an arbitrary $(\lambda_0,\omega_0) \in \Lambda \times \Omega$ and consider M_{λ_0,ω_0}. Then $t \notin M_{\lambda_0,\omega_0} \Rightarrow F_t(\omega) = 0$, a.e. $[L]$, i.e., $\forall \ \omega \notin G_t$. So (c) $\Rightarrow h_t(s) = 0$, a.e. $[\mathcal{P}_S]$, that is, $t \notin M_{\lambda_0,\omega_0}$ implies that $\exists \ B_t \in \mathcal{C}$ s.t. $Q_\omega(B_t) = 0, \quad \forall \ \omega \in \Omega$, and $h_t(s) = 0, \ \forall \ s \notin B_t$. Let $H = \{(t,s) : h(t,s) \neq 0\}$ and $H_t = \{s : (t,s) \in H\}$. Since h is $\mathcal{B} \times \mathcal{C}$ measurable, then $H \in \mathcal{B} \times \mathcal{C} \Rightarrow H_t \in \mathcal{C}$. But $s \notin B_t \Rightarrow s \notin H_t \Rightarrow H_t \subset B_t \Rightarrow Q_\omega(H_t) = 0$,

$\forall\, \omega \in \Omega$, $\forall\ t \notin M_{\lambda_0,\omega_0}$. Now $P_{\lambda_0,\omega_0}(M_{\lambda_0,\omega_0}) = 0$, and for $t \notin M_{\lambda_0,\omega_0}$, $Q_{\omega_0}(H_t) = 0$, so that $(P_{\lambda_0,\omega_0} \times Q_{\omega_0})(H) = 0, \forall$ $(\lambda_0, \omega_0) \in \Lambda \times \Omega$. []

Remark 2.4. Several remarks are in order regarding the conditions given in the above theorem:

(a) Condition (b) is going to be satisfied if the respective family of measures constitutes an exponential family.

(b) Strong completeness can be replaced by the weaker requirement of strong bounded completeness since h is bounded.

(c) In most applications $q_s(\omega)$ will be continuous and Lemma 2.1 will then provide an easy way to verify condition (d). []

The special case where the members of the family $\mathcal{P}_T$ are indexed only by the parameter $\lambda \in \Lambda$, is treated in

Corollary 2.5. Suppose the following conditions are satisfied:

(a) the family $\{P_\lambda : \lambda \in \Lambda\}$ is boundedly complete,

(b) the family $\{Q_\omega : \omega \in \Omega\}$ is strongly complete wrt L,

(c) $q(s,\omega)$ is $\mathcal{C} \times \beta(\Omega)$ measurable.

Then the family $\{P_\lambda \times Q_\omega : (\lambda, \omega) \in \Lambda \times \Omega\}$ is boundedly complete.

So far we have investigated bounded completeness which is sufficient for our purposes, that is, for constructing similar tests. However we can obtain a result on completeness when $\mathcal{P}_S$ is an exponential family. Thus let us now suppose $\mathcal{S} \subset R^k$ and take $\mathcal{C}$ as $\beta(\mathcal{S})$. We are now ready for

Theorem 2.6. Suppose the following conditions are satisfied:

(a) the family $\{P_{\lambda,\omega} : \lambda \in \Lambda\}$ is complete $\forall\ \omega \in \Omega$,

(b) the members of the family $\{P_{\lambda,\omega} : \omega \in \Omega\}$ are all equivalent $\forall\ \lambda \in \Lambda$,

(c) $q(s,\omega) = \psi(\omega) \exp\{\alpha(\omega) \cdot s\}$ where $\psi(\omega)$ is a normalizing constant and $\alpha : \Omega \to R^k$,

(d) $\alpha[\Omega]$ has a nonempty interior in R^k.

Then the family $\mathcal{P}_{(T,S)}$ is complete.

Pf. Let $\Gamma \subset \alpha[\Omega]$ be a convex polyhedron with a nonempty interior in R^k. For each $\gamma \in \Gamma$, let ω_γ be an arbitrary but fixed element in $\alpha^{-1}[\gamma]$. We use γ to parametrize the density function q, that is,

$$q(s,\gamma) = \phi(\gamma) \exp\{\gamma \cdot s\}$$

where $\phi(\gamma) = \psi(\omega_\gamma)$ is a normalizing constant. Let $\mathcal{P}_1 = \{P_{\lambda,\gamma} \times Q_\gamma : (\lambda,\gamma) \in \Lambda \times \Gamma\}$. Here we have $P_{\lambda,\gamma} = P_{\lambda,\omega_\gamma}$ and $Q_\gamma = Q_{\omega_\gamma}$. Since the set $\{\omega_\gamma : \gamma \in \Gamma\} \subset \Omega$, then $\mathcal{P}_1 \subset \mathcal{P}_{(T,S)}$. Let $N \in \mathcal{B} \times \beta(\mathcal{S})$ s.t. $(P_{\lambda,\gamma} \times Q_\gamma)(N) = 0$, $\forall\ (\lambda,\gamma) \in \Lambda \times \Gamma$. We want to show that $(P_{\lambda,\omega} \times Q_\omega)(N) = 0$, $\forall\ (\lambda,\omega) \in \Lambda \times \Omega$. Choose any $(\lambda_1,\omega_1) \in \Lambda \times \Omega$. Take $\gamma_0 \in \Gamma$. Then $(P_{\lambda_1,\gamma_0} \times Q_{\gamma_0})(N) = 0$

so that $\exists\, N_{\lambda_1, \gamma_0} \in \mathcal{B}$ s.t. $P_{\lambda_1, \gamma_0}(N_{\lambda_1, \gamma_0}) = 0$, and for $t \notin N_{\lambda_1, \gamma_0}$, $Q_{\gamma_0}(N_t) = 0$. (b) $\Rightarrow P_{\lambda_1, \omega_1}(N_{\lambda_1, \gamma_0}) = 0$, and for $t \notin N_{\lambda_1, \gamma_0}$, $Q_{\omega_1}(N_t) = 0$, because (c) $\Rightarrow \mathcal{P}_S$ is an equivalent family. Hence $(P_{\lambda_1, \omega_1} \times Q_{\omega_1})(N) = 0$. Therefore it is sufficient to show that $\mathcal{P}_1$ is a complete family. To this end let

$$I(\lambda, \gamma) = \int_{\mathcal{J} \times \mathcal{S}} h(t, s)\, d(P_{\lambda, \gamma} \times Q_\gamma)(t, s),$$

where h is any real statistic. We want to show that $I(\lambda, \gamma) = 0$, $\forall\, (\lambda, \gamma) \in \Lambda \times \Gamma \Rightarrow h(t, s) = 0$, a.e. $[\mathcal{P}_1]$. For each $(\lambda, \gamma) \in \Lambda \times \Gamma$, apply Fubini's theorem to obtain a $\mathcal{B}$-measurable function $G(t, \lambda, \gamma)$ and an $N_{\lambda, \gamma} \in \mathcal{B}$ s.t. $P_{\lambda, \gamma}(N_{\lambda, \gamma}) = 0$,

$$G(t, \lambda, \gamma) = \int_{\mathcal{S}} h(t, s)\, dQ_\gamma(s), \quad \forall\ t \notin N_{\lambda, \gamma},$$

and

$$I(\lambda, \gamma) = \int_{\mathcal{J}} G(t, \lambda, \gamma)\, dP_{\lambda, \gamma}(t)\ .$$

For a set B we denote the complement set by B^c. Let

$$N_\gamma = \{t : \int_{\mathcal{S}} h(t, s)\, dQ_\gamma(s) \text{ exists and is finite}\}^c\ .$$

Notice that

$$N_\gamma^c = \{t: \int_{\mathscr{X}} h_t^+(s)dQ_\gamma(s) < \infty\} \cap \{t: \int h_t^-(s)dQ_\gamma(s) < \infty\} .$$

By Tonelli's theorem $N_\gamma^c \in \mathcal{B}$. Further $N_{\lambda,\gamma}^c \subset N_\gamma^c \Rightarrow N_\gamma \subset N_{\lambda,\gamma} \Rightarrow P_{\lambda,\gamma}(N_\gamma) = 0$. Let $\gamma_1, \gamma_2, \ldots, \gamma_\ell$ be the extreme points of Γ. Set $N = \cup_{i=1}^{\ell} N_{\gamma_i}$. Then $N \in \mathcal{B}$ and (b) $\Rightarrow P_{\lambda,\gamma}(N) = 0$, $\forall$ $(\lambda, \gamma) \in \Lambda \times \Gamma$. Further, the integral $\int_{\mathscr{X}} h(t,s)dQ_{\gamma_i}(s)$ exists and is finite for $i = 1, 2, \ldots, \ell$, whenever $t \notin N$. Since $\gamma_1, \gamma_2, \ldots, \gamma_\ell$ are the extreme points of Γ and since $\exp\{\gamma \cdot s\}$ thought of as a function on Γ is convex, it follows that the integral $\int_{\mathscr{X}} h(t,s)dQ_\gamma(s)$ exists and is finite for all $\gamma \in \Gamma$ whenever $t \notin N$. Thus for $\gamma \in \Gamma$ define F by

$$(2.7) \qquad F(t,\gamma) = \int_{\mathscr{X}} h(t,s)q(s,\gamma)d\mu(s), \quad t \notin N,$$
$$= 0, \quad t \in N.$$

We now show that F is $\mathcal{B} \times \beta(\Gamma)$ measurable. Since $\exp\{\gamma \cdot s\}$ thought of as a function on Γ is continuous, and since $\exp\{\gamma \cdot s\} \leq \Sigma_{i=1}^{\ell} \exp\{\gamma_i \cdot s\}$, $\forall$ $\gamma \in \Gamma$, then the Lebesgue dominated convergence theorem implies that $\int_{\mathscr{X}} \exp\{\gamma \cdot s\}d\mu(s) = 1/\phi(\gamma)$ (thought of as a function on Γ) is continuous. So $q_s(\gamma)$ is continuous. Hence Lemma 2.1 implies that $q(s,\gamma)$ is $\beta(\mathscr{X}) \times \beta(\Gamma)$ measurable. Then $h(t,s)q(s,\gamma)$ is $\mathcal{B} \times \beta(\mathscr{X}) \times \beta(\Gamma)$ measurable and Tonelli's theorem applied to the positive and negative parts of h

implies that F restricted to $\mathcal{J} \setminus N \times \Gamma$ is $\mathcal{B} \times \beta(\Gamma)$-measurable which in turn implies that F is $\mathcal{B} \times \beta(\Gamma)$ measurable. Now since $F(t, \gamma) = G(t, \lambda, \gamma)$, $\forall\ t \notin N \cup N_{\lambda, \gamma}$, we have

$$I(\lambda, \gamma) = \int_{\mathcal{J}} F(t, \gamma) dP_{\lambda, \gamma}(t), \quad \forall\ (\lambda, \gamma) \in \Lambda \times \Gamma .$$

The rest of the proof can be made along the same lines given in the proof of Theorem 2.2, keeping in mind that (d) implies the strong completeness of the family $\{Q_\gamma : \gamma \in \Gamma\}$ wrt Lebesgue measure on $\beta(\Gamma)$. This last statement follows from our choice of Γ at the beginning of the proof along with Theorem 7.3 in Lehmann and Scheffé (1955). $\square$

The essential part in the above proof is the construction of the measurable function F as given in (2.7). This is where we use the additional condition (c). We recall that the boundedness of the real statistic h enabled us to avoid this difficulty in the proof of Theorem 2.2.

Remark 2.8. Theorem 2.6 would still hold if conditions (c) and (d) are replaced by the requirement that there are two subfamilies $\mathcal{P}_0 \subset \mathcal{P}_1 \subset \mathcal{P}_S$ s.t. $\mathcal{P}_1$ is an equivalent (to $\mathcal{P}_S$) subfamily which is strongly complete and $\mathcal{P}_0$ is a countable subfamily s.t. $\mathcal{P}_0$-integrability implies $\mathcal{P}_1$-integrability. $\square$

Corollary 2.9. Suppose the family $\{P_\lambda : \lambda \epsilon \Lambda\}$ is complete, and that conditions (c) and (d) of Theorem 2.6 hold. Then the family $\{P_\lambda \times Q_\omega : (\lambda, \omega) \epsilon \Lambda \times \Omega\}$ is complete.

We devote the rest of the chapter to a discussion of these results and their relation to other results in the literature. Lehmann and Scheffé (1955) replaced condition 2.5(a) by two conditions about the family $\{P_\lambda : \lambda \epsilon \Lambda\}$ similar to conditions 2.5(b) and 2.5(c). They concluded that the family of product measures is strongly complete. However, we have not been able to follow their proof. Fraser (1957) assumed 2.5(b) and that the family $\{P_\lambda : \lambda \epsilon \Lambda\}$ is complete. He concluded, Theorem 6.3, page 26, that the family of product measures is complete. Fraser's conditions are weaker yet than those of Lehmann and Scheffé and as a result we can not follow his proof either. A generalization of Lehmann and Scheffé's result to the family $\mathcal{P}_{(T,S)}$ is given by Gautschi (1959). He considered the case where $\Lambda = R$ and assumed 2.6(c), 2.6(d), and that $\mathcal{P}_T$ is a strongly complete exponential family. He then concluded that the family $\mathcal{P}_{(T,S)}$ is strongly complete. There is one measurability step (concerning the set S given in his proof) that we do not think is justifiable. Imhof (1960) generalized Gautschi's work to the case where $\Lambda = R^m$. He referred to the proof given by Gautschi for the univariate case. In 1976 Seely considered a special case of Theorem 2.6 while studying the completeness of minimal sufficient statistics for multivariate normal families. Our proof of Theorem 2.6 is adapted from his proof. The results of Theorem 2.6 are thus seen to be fairly general. They are more general than those given by Lehmann, Scheffé and Fraser in that they allow for the members of $\mathcal{P}_T$ to be indexed, in addition to λ, by the parameters $\omega \in \Omega$ which also index the members of $\mathcal{P}_S$. They are more general than the results of Gautschi and Imhof in that we do not assume that $\mathcal{P}_T$ is a dominated family.

III. PRELIMINARIES

In this chapter we introduce the model that will be considered throughout the thesis. Then we develop a minimal sufficient statistic and discuss completeness. The results regarding sufficiency and completeness are adapted from Seely (1976). We shall be dealing with the vector space R^n and the vector space of real symmetric matrices.

3.1. The Model

We consider a random vector Y distributed according to some distribution in the general family of multivariate normal distributions,

$$\mathcal{P}_Y = \{N_n(\mu, \Sigma_\theta) : \mu \in \mathcal{M}, \ \theta \in \Theta\},$$

where

(A1) $\mathcal{M}$ is a subspace of R^n.

(A2) Θ is a set in R^p having a nonempty interior.

(A3) For each $\theta \in \Theta$ the covariance matrix Σ_θ is positive definite and can be expressed as $\Sigma_\theta = V_0 + \Sigma_{i=1}^p \theta_i V_i$ where $V_0, \ldots, V_p$ are known $n \times n$ real symmetric matrices and $V_1, \ldots, V_p$ are linearly independent.

(A4) $\mathcal{V} = \{\Sigma_\theta : \theta \in \Theta\}$ is such that $I \in \mathcal{V}$.

It should be noticed above that the only relevant assumptions are that $\mathcal{M}$ is a subspace, that Θ has a nonempty interior and that the parameter space is $\mathcal{M} \times \Theta$. The other assumptions are for convenience, as justified by the following remarks:

Remark 3.1.1. Any covariance matrix can be expressed in the form given in assumption (A3). If $V_1, \ldots, V_p$ are not linearly independent then a reparametrization would produce the required linear independence. □

Remark 3.1.2. If $I \notin \mathcal{V}$, then the transformation from Y to $\tilde{Y} = \Sigma_0^{-1/2} Y$ where $\Sigma_0 \in \mathcal{V}$ will result in a family having covariance matrices $\tilde{\mathcal{V}}$ such that $I \in \tilde{\mathcal{V}}$. □

3.2. A Minimal Sufficient Statistic

First we develop some notation. Let $\mathcal{V}_1$, $\mathcal{V}_0$ and v denote the affine hull of $\mathcal{V}$, the unique subspace that is parallel to $\mathcal{V}_1$, and the dimension of $\mathcal{V}_0$ respectively. Then $\mathcal{V}_1 = I + \mathcal{V}_0$, $\mathcal{V}_0 = sp\{V_1, \ldots, V_p\}$, and $v = p$. Let $\mathcal{A} = \{\Sigma_\theta^{-1} : \theta \in \Theta\}$. Define $\mathcal{U}$ to be the smallest subspace of R^n such that $A\mathcal{M} \subset \mathcal{U}$, $\forall$ $A \in \mathcal{A}$. Let $u = \dim \mathcal{U}$ (we assume here that $u > 0$ since the case $u = 0$ is trivial). Let $U_1, \ldots, U_u$ be a basis for $\mathcal{U}$ and define $U = (U_1, \ldots, U_u)$. Let

(3.2.1) $$T = (T_1, \ldots, T_u)' \text{ where } T_i = U_i'Y, \quad i = 1, 2, \ldots, u.$$

Let $Z = Q'Y$ where Q is an $n \times q$ ($q = n - \dim \mathcal{M}$) matrix whose columns form an orthonormal basis for $\mathcal{M}^{\perp}$. Then $Z \sim N_q(0, \Lambda_\theta)$ where $\Lambda_\theta = Q'\Sigma_\theta Q$. Thus the class of distributions induced by Z is

$$\mathcal{P}_Z = \{N_q(0, \Lambda_\theta) : \theta \in \Theta\}.$$

Let $\mathcal{H} = \{\Lambda_\theta : \theta \in \Theta\}$. Let $\mathcal{H}_1$, $\mathcal{H}_0$, and h denote the affine hull of $\mathcal{H}$, the unique subspace that is parallel to $\mathcal{H}_1$, and the dimension of $\mathcal{H}_0$ respectively. Let $W_i = Q'V_iQ$, $i = 0, 1, \ldots, p$. Then $\mathcal{H}_1 = W_0 + \mathcal{H}_0$ and $\mathcal{H}_0 = \mathrm{sp}\{W_1, \ldots, W_p\}$. We mention that $W_1, \ldots, W_p$ is not necessarily a basis for $\mathcal{H}_0$. So, let $H_1, \ldots, H_h$ be such a basis. Since $W_i \in \mathcal{H}_0$, then

(3.2.2) $$W_i = \Sigma_{j=1}^{h} \lambda_{ij} H_j, \quad i = 1, 2, \ldots, p,$$

for some real numbers λ_{ij}. Further, $I_n \in \mathcal{V}$ implies $I_q \in \mathcal{H} \subset \mathcal{H}_1$ so that $W_0 - I \in \mathcal{H}_0$. Thus, we can write

(3.2.3) $$W_0 - I = \Sigma_{j=1}^{h} \alpha_j H_j,$$

for some real numbers α_j.

We have seen above that $I_q \in \mathcal{H}_1$ so that $\mathcal{H}_1 = I + \mathcal{H}_0$. Hence for $\theta \in \Theta$ the covariance matrix Λ_θ has the representation

$$\Lambda_\theta = I + \Sigma_{j=1}^{h} \nu_j(\theta) H_j ,$$

where

$$\nu_j(\theta) = \alpha_j + \Sigma_{i=1}^{p} \lambda_{ij}\theta_i, \quad j = 1, 2, \ldots, h. \tag{3.2.4}$$

To see this we note that $\Lambda_\theta \in \mathcal{H}$ implies that $\Lambda_\theta = W_0 + \Sigma_{i=1}^{p} \theta_i W_i$. Equation (3.2.4) now follows upon making use of (3.2.2), (3.2.3) and the fact that $H_1, \ldots, H_h$ is a basis for $\mathcal{H}_0$.

Let $\mathcal{C} = \{\Lambda_\theta^{-1} : \theta \in \Theta\}$. As usual we define $\mathcal{C}_1$, $\mathcal{C}_0$, and c to be the affine hull of $\mathcal{C}$, the subspace parallel to $\mathcal{C}_1$, and the dimension of $\mathcal{C}_0$. Since $I \in \mathcal{H}$ we have $I \in \mathcal{C} \subset \mathcal{C}_1$ so that $\mathcal{C}_1 = I + \mathcal{C}_0$. Let $C_1, \ldots, C_c$ be a basis for $\mathcal{C}_0$. Then $\Lambda_\theta^{-1} \in \mathcal{C}$ has the representation

$$\Lambda_\theta^{-1} = I + \Sigma_{i=1}^{c} \delta_i(\theta) C_i , \tag{3.2.5}$$

where the δ_i are functions from Θ to R^1 and are defined implicitly by this representation. Now let us define

$$R = (R_1, \ldots, R_c)' \text{ where } R_i = Y'QC_iQ'Y, \; i = 1, 2, \ldots, c. \tag{3.2.6}$$

Theorem 2.7 in Seely (1976) states that T and R jointly constitute a minimal sufficient statistic for $\mathcal{P}_Y$. Seely also gave necessary and sufficient conditions for the sufficient statistic (T,R) to be complete. His results are applicable when $\mathcal{V}$ has a nonempty interior relative to its affine hull $\mathcal{V}_1$ which can be shown to be equivalent to Θ having a nonempty interior in R^p. For reference purposes we give below some of his results. We include only the conditions which are directly applicable for our purposes.

<u>Theorem 3.2.7.</u> $\mathcal{P}_{(T,R)}$ is complete if and only if $\mathcal{P}_T$ is complete and $\mathcal{P}_R$ is complete.

<u>Theorem 3.2.8.</u> Let $\mathcal{B}$ be any fixed set of matrices such that the affine hull of $\mathcal{B}$ is $\mathcal{V}_1$. Then the following conditions are all equivalent:

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

(b) $V\mathcal{M} \subset \mathcal{M}$ for all $V \in \mathcal{B}$.

(c) T and Z are independent wrt every distribution $P \in \mathcal{P}_Y$.

(d) $\dim \mathcal{U} = \dim \mathcal{M}$.

Before we state the next theorem we define a quadratic subspace. A subspace $\mathcal{B}$ of the vector space of real symmetric matrices with the property that $B \in \mathcal{B}$ implies $B^2 \in \mathcal{B}$ is said to be a quadratic subspace. The term quadratic subspace is used by Seely (1971) who investigated several properties of such subspaces. We mention also that the notion of a quadratic subspace is equivalent to that of a Jordan algebra.

Theorem 3.2.9. The following conditions are all equivalent:

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

(b) $\mathcal{H}_0$ is a quadratic subspace.

(c) $\mathcal{C}_1 = \mathcal{H}_1$.

(d) $\mathcal{C}$ has a nonempty interior relative to $\mathcal{C}_1$.

Finally we mention that Theorems 3.2.8 and 3.2.9 are a little simpler than Seely's results because of our assumption (A4).

IV. HYPOTHESIS TESTING WHEN THE SUFFICIENT STATISTIC IS COMPLETE

In this and the following chapters we deal with the problem of testing different hypotheses about the parameter θ. It seems logical to expect to be able to construct optimal tests for the functions $\delta_i(\theta)$, $i = 1,2,\ldots,c$, given by (3.2.5), because of the way they appear in the density function of the minimal sufficient statistic (T,R) as given below. We are going to develop a theory for testing hypotheses about the $\delta_i(\theta)$'s and hope that we may be able to convert some interesting hypotheses about θ to hypotheses about these functions.

We begin by treating the case where the family $\mathcal{P}_{(T,R)}$ is complete. Then we consider the case where the family $\mathcal{P}_R$ is complete but $\mathcal{P}_T$ is not. The case where neither $\mathcal{P}_R$ nor $\mathcal{P}_T$ is complete is dealt with in a subsequent chapter.

4.1. The Case Where $\mathcal{P}_{(T,R)}$ is Complete

Throughout this section the family $\mathcal{P}_{(T,R)}$ is assumed to be complete. Then Theorem 3.2.7 implies that $\mathcal{P}_R$ is complete. Hence by Theorem 3.2.9(c) we have $c = h$ and we can choose $C_i = H_i$, $i = 1,2,\ldots,h$. We shall assume that $h > 1$. (For the case $h = 1$, see the paragraph after Theorem 4.1.9.) Now Equations (3.2.5) and (3.2.6) can be written as

(4.1.1) $$\Lambda_\theta^{-1} = I + \Sigma_{i=1}^{h} \delta_i(\theta)H_i \ ,$$

(4.1.2) $$R = (R_1, \ldots, R_h)' \quad \text{where } R_i = Y'QH_iQ'Y, \quad i = 1, 2, \ldots, h.$$

Let $f_T(t|\mu,\theta)$ denote the density function of $T \sim N_u(U'\mu, U'\Sigma_\theta U)$ wrt Lebesgue measure on R^u. It is easily established that

$$f_T(t|\mu,\theta) = \psi(\mu,\theta)\exp\{t'\Delta(\theta)t + t'\pi(\mu,\theta)\} \ ,$$

where

$$\Delta(\theta) = -\frac{1}{2}(U'\Sigma_\theta U)^{-1} \ ,$$

(4.1.3) $$\pi(\mu,\theta) = (U'\Sigma_\theta U)^{-1}U'\mu \ ,$$

and $\psi(\mu,\theta)$ is a normalizing constant.

It is also easily established that the density function of R wrt some measure on R^h is given by

$$f_R(r|\theta) = \Phi(\theta)\exp\{r'\gamma(\theta)\} \ ,$$

where

$$\gamma(\theta) = (\gamma_1(\theta), \ldots, \gamma_h(\theta))' \ ,$$

(4.1.4) $$\gamma_i(\theta) = -\frac{1}{2}[1+\delta_i(\theta)], \quad i = 1, 2, \ldots, h,$$

and $\Phi(\theta)$ is a normalizing constant. We mention that in most practical situations $[1+\delta_i(\theta)]$ are the natural parameters of interest. The $-1/2$ factor in (4.1.4) is for notation convenience.

Theorem 3.2.7 implies $\mathcal{P}_T$ is complete. Hence Theorem 3.2.8(c) implies that T and R are independent wrt every distribution $P \in \mathcal{P}_Y$. Then

$$f_{(T,R)}(t,r|\mu,\theta) = f_T(t|\mu,\theta)f_R(r|\theta) \,,$$

$$= \psi(\mu,\theta)\Phi(\theta) \exp\{t'\Delta(\theta)t+t'\pi(\mu,\theta)+r'\gamma(\theta)\}.$$

Let $\Omega = \mathcal{M} \times \Theta$ and consider testing

$$H1 : \gamma_1(\theta) \leq \gamma_0 \quad \text{vs} \quad K1 : \gamma_1(\theta) > \gamma_0 \,.$$

The choice of $\gamma_1(\theta)$ is for notational convenience only and the theory to be developed is applicable to any $\gamma_k(\theta)$, $k = 1,2,\ldots,h$. We suppose that γ_0 is selected such that the class of distributions under H1 and that under K1 are nonvoid. From now on such assumptions as this one will be made without explicit mention.

Let $\Theta_{H1} = \{\theta \in \Theta : \gamma_1(\theta) \leq \gamma_0\}$ and $\Theta_{K1} = \{\theta \in \Theta : \gamma_1(\theta) > \gamma_0\}$. For a set B let $\bar{B}$ denote the closure of B. Hence the set of common accumulation points of Θ_{H1} and Θ_{K1} is given by $\Theta_{B1} = \bar{\Theta}_{H1} \cap \bar{\Theta}_{K1}$. Since the function $\gamma_1(\theta)$ is continuous, then $\Theta_{B1} \subset \{\theta \in \Theta : \gamma_1(\theta) = \gamma_0\}$. Now the parameter space under H1 is given by $\Omega_{H1} = \mathcal{M} \times \Theta_{H1}$, and that under K1 is given by $\Omega_{K1} = \mathcal{M} \times \Theta_{K1}$. Let Ω_{B1} denote the set of common accumulation points of Ω_{H1} and Ω_{K1}. Then

$$\Omega_{B1} = \bar{\Omega}_{H1} \cap \bar{\Omega}_{K1} = (\overline{\mathcal{M} \times \Theta_{H1}}) \cap (\overline{\mathcal{M} \times \Theta_{K1}})$$

$$= (\bar{\mathcal{M}} \times \bar{\Theta}_{H1}) \cap (\bar{\mathcal{M}} \times \bar{\Theta}_{K1})$$

$$= \bar{\mathcal{M}} \times (\bar{\Theta}_{H1} \cap \bar{\Theta}_{K1}) = \mathcal{M} \times \Theta_{B1} ,$$

where the third equality follows by Dugundji (1966), page 99.

Let

(4.1.5) $$S = (R_2, \ldots, R_h)' .$$

Then (T,S) is a sufficient statistic for $\mathcal{P}_{(T,R)}$ on Ω_{B1}, where we use the terminology "$\mathcal{P}_{(T,R)}$ on Ω_{B1}" to denote the subfamily of $\mathcal{P}_{(T,R)}$ indexed by Ω_{B1}. We want to show that, on Ω_{B1}, the sufficient statistic (T,S) is (boundedly) complete. We do this using the results of Chapter II. But first some preparations. Let $\Gamma = \{\gamma(\theta) : \theta \in \Theta\}$ and $\Gamma_{B1} = \{\phi(\theta) : \theta \in \Theta_{B1}\}$, where $\phi(\theta) = (\gamma_2(\theta), \ldots, \gamma_h(\theta))'$. Theorem 3.2.9(d) implies that Γ has a nonempty interior in R^h. Unfortunately, this does not in general imply:

(4.1.6) Γ_{B1} has a nonempty interior in R^{h-1} .

However, we notice the following

Remark 4.1.7.

(a) In most practical applications (4.1.6) is true.

(b) If there is $\phi \in \Gamma_{B1}$ such that (γ_0, ϕ) is an interior point of Γ, then (4.1.6) holds true.

(c) If Θ is open in R^p, then Γ is open in R^h and (4.1.6) holds true. $\square$

In the light of Remark 4.1.7, the assumption that (4.1.6) holds true seems to be justifiable.

Lemma 4.1.8. If (4.1.6) holds true, then the family $\mathcal{P}_{(T,S)}$ is complete on Ω_{B1}.

Pf. Since T and S are independent, the family $\mathcal{P}_{(T,S)}$ can be viewed as a family of product measures. So we prove the lemma by verifying conditions (a)-(d) of Theorem 2.6

(a) Since $\mathcal{P}_T$ is complete on Ω, then Theorem 3.2.8(d) implies that $\dim \mathcal{U} = \dim \mathcal{M}$. But $\mathcal{M} \subset \mathcal{U}$ because $I \in \mathcal{V}$. Hence $\mathcal{U} = \mathcal{M}$, so that $\Sigma_\theta \mathcal{U} = \mathcal{M}$, $\forall\ \theta \in \Theta$. Fix $\theta_0 \in \Theta_{B1} \subset \Theta$, and recall (4.1.3). Then

$$\begin{aligned}\{\pi(\mu, \theta_0) : \mu \in \mathcal{M}\} &= \{(U'\Sigma_{\theta_0} U)^{-1} U'\mu : \mu \in \mathcal{M}\},\\ &= \{(U'\Sigma_{\theta_0} U)^{-1} U'\mu : \mu \in \Sigma_{\theta_0} \mathcal{U}\},\\ &= \{(U'\Sigma_{\theta_0} U)^{-1} U'\Sigma_{\theta_0} Ua : a \in R^u\} = R^u .\end{aligned}$$

Thus, for every $\theta \in \Theta_{B1}$, the family $\{f_T(t|\mu, \theta) : \mu \in \mathcal{M}\}$ is complete.

(b) This condition is satisfied because, for every $\mu \in \mathcal{M}$, the family $\{f_T(t|\mu, \theta) : \theta \in \Theta_{B1}\}$ is an exponential family.

(c) It is easily established that the density function of S wrt some measure on R^{h-1} is given by

$$f_S(s|\theta) = \psi(\theta) \exp\{s'\phi(\theta)\},$$

where $\psi(\theta)$ is a normalizing constant.

(d) $\phi[\Theta_{B1}] = \Gamma_{B1}$ has a nonempty interior in R^{h-1} by (4.1.6). []

The above lemma implies that all tests similar on Ω_{B1} have Neyman structure; see Lehmann (1959), Theorem 2, page 134. This leads to the following

Theorem 4.1.9. If (4.1.6) holds true, then a UMPU α-test for H1 vs K1 is given by

(4.1.10)
$$\varphi_1(r_1, s) = \begin{cases} 1, & r_1 > c(s), \\ d(s), & r_1 = c(s), \\ 0, & \text{elsewhere}, \end{cases}$$

where $c(s)$ and $d(s)$ are determined by

$$E_{\gamma_0}[\varphi_1(R_1, S)|s] = \alpha.$$

Pf. Using an argument like the one given in the proof of Theorem 3, page 136, in Lehmann (1959), we conclude that a UMPU α-test for H1 vs K1 is given by

$$\varphi(r_1, t, s) = \begin{cases} 1, & r_1 > c(t, s), \\ d(t, s), & r_1 = c(t, s), \\ 0, & \text{elsewhere}, \end{cases}$$

where $c(t, s)$ and $d(t, s)$ are determined by

$$E_{\gamma_0}[\varphi(R_1, T, S) | t, s] = \alpha .$$

The fact that T and R are independent obtains the test (4. 1. 10). []

When $h = 1$, (4. 1. 6) does not hold. However, Theorem 4. 1. 9 is still true.

We treated above the hypothesis H1 vs K1. Other hypotheses can also be considered, e. g.,

H2: $\gamma_1(\theta) \leq \gamma_1$ or $\gamma_1(\theta) \geq \gamma_2$ vs K2: $\gamma_1 < \gamma_1(\theta) < \gamma_2$,

H3: $\gamma_1 \leq \gamma_1(\theta) \leq \gamma_2$ vs K3: $\gamma_1(\theta) < \gamma_1$ or $\gamma_1(\theta) > \gamma_2$,

H4: $\gamma_1(\theta) = \gamma_0$ vs K4: $\gamma_1(\theta) \neq \gamma_0$.

Using similar techniques we get the following results:

(4.1.11) A UMPU α-test for H2 vs K2 is given by

$$\varphi_2(r_1, s) = \begin{cases} 1, & c_1(s) < r_1 < c_2(s), \\ d_i(s), & r_1 = c_i(s),\ i = 1, 2, \\ 0, & \text{elsewhere}, \end{cases}$$

where the c's and d's are determined by

$$E_{\gamma_1}[\varphi_2(R_1, S) | s] = E_{\gamma_2}[\varphi_2(R_1, S) | s] = \alpha .$$

(4.1.12) A UMPU α-test for H3 vs K3 is given by

$$\varphi_3(r_1, s) = \begin{cases} 1, & r_1 < c_1(s) \text{ or } r_1 > c_2(s), \\ d_i(s), & r_1 = c_i(s),\ i = 1, 2, \\ 0, & \text{elsewhere}, \end{cases}$$

where the c's and d's are determined by

$$E_{\gamma_1}[\varphi_3(R_1, S) | s] = E_{\gamma_2}[\varphi_3(R_1, S) | s] = \alpha .$$

(4.1.13) A UMPU α-test for H4 vs K4 is given by

$$\varphi_4(r_1, s) = \begin{cases} 1, & r_1 < c_1(s) \text{ or } r_1 > c_2(s), \\ d_i(s), & r_1 = c_i(s),\ i = 1, 2, \\ 0, & \text{elsewhere}, \end{cases}$$

where the c's and d's are determined by

$$E_{\gamma_0}[\varphi_4(R_1, S) | s] = \alpha, \quad \text{and}$$

$$E_{\gamma_0}[R_1 \varphi_4(R_1, S) | s] = \alpha E_{\gamma_0}[R_1 | s] .$$

We mention here the work of Anderson (1971). His investigation of time series satisfying a stochastic difference equation led him to consider multivariate normal models where the <u>inverse</u> of the covariance matrix has a linear structure. In these models the parameters in the linear structure have relatively simple expressions in terms of the parameters of the time series. Besides defining the model through the inverse of the covariance matrix, Anderson assumed the natural parameter space as his underlying parameter space. He treated first the case of zero mean and then the case where $\mathcal{M} = R(X)$ with the columns of the matrix X being the eigenvectors of the covariance matrix. This latter condition is equivalent to the assumption that the family $\mathcal{P}_T$ is complete. For more details the reader is referred to Chapter 6 in Anderson.

UMPU tests are also available about linear combinations of the $\gamma_i(\theta)$'s. Let $b(\theta) = \Sigma_{i=1}^{h} a_i \gamma_i(\theta)$, where the a_i's are real numbers. Assume that $a_1 \neq 0$. Then

$$\gamma_1(\theta) = [b(\theta) - \Sigma_{i=2}^{h} a_i \gamma_i(\theta)]/a_1 .$$

Consider the transformation

$$g_1 = r_1/a_1 \quad \text{and} \quad g_i = r_i - a_i r_1/a_1, \quad i = 2, \ldots, h.$$

It is easily established that the density function of

$G = (G_1, \ldots, G_h)'$ wrt some measure on R^h is given by

$$f_G(g \mid \theta) = h(\theta) \exp\{g_1 b(\theta) + \Sigma_{i=2}^h g_i \gamma_i(\theta)\} .$$

Thus, for example, when testing

$$H5: b(\theta) \leq b_0 \quad \text{vs} \quad K5: b(\theta) > b_0 ,$$

Theorem 4.1.9 is applicable.

Let us now suppose $\gamma_h(\theta) < 0$, $\forall$ $\theta \in \Theta$, where the choice of γ_h is for notational convenience only. This assumption is going to be satisfied in most practical applications and thus seems to be justifiable. When this is the case we can get UMPU tests for the ratios $\gamma_i(\theta)/\gamma_h(\theta)$. We reparametrize to obtain

$$f_R(r \mid \theta) = \Phi(\theta) \exp\{\gamma_h(\theta)\Sigma_{i=1}^h r_i \eta_i(\theta)\} ,$$

where

(4.1.14) $\quad \eta_i(\theta) = \gamma_i(\theta)/\gamma_h(\theta) = [1+\delta_i(\theta)]/[1+\delta_h(\theta)], \quad i = 1, 2, \ldots, h.$

Consider the following hypothesis

$$H6: \eta_1(\theta) \geq \eta_0 \quad \text{vs} \quad K6: \eta_1(\theta) < \eta_0 .$$

Let Θ_{H6}, Θ_{K6}, Θ_{B6}, Ω_{H6}, Ω_{K6}, and Ω_{B6} be defined as usual, i.e., $\Omega_{B6} = \mathcal{M} \times \Theta_{B6}$ and Θ_{B6} is the set of common accumulation points of Θ_{H6} and Θ_{K6}. Observe that $\Theta_{B6} \subset \{\theta \in \Theta : \eta_1(\theta) = \eta_0\}$. Let

$$G = (R_2, \ldots, R_{h-1}, R_h + \eta_0 R_1)' . \tag{4.1.15}$$

Then (T, G) is a sufficient statistic for $\mathcal{P}_{(T,R)}$ on Ω_{B6}. Let $\Gamma_{B6} = \{\phi(\theta) : \phi \in \Theta_{B6}\}$. Corresponding to assumption (4.1.6) we assume the following

(4.1.16) Γ_{B6} has a nonempty interior in R^{h-1}.

Using Theorem 2.6, as we did in Lemma 4.1.8, it is straightforward to prove the following

Lemma 4.1.17. If (4.1.16) holds true, then the family $\mathcal{P}_{(T,G)}$ is complete on Ω_{B6}.

We now give another important result.

Theorem 4.1.18. Assuming that $\gamma_h(\theta) < 0$, $\forall\ \theta \in \Theta$, and that (4.1.16) is true, a UMPU α-test for H6 vs K6 is given by

$$\varphi_6(r_1, g) = \begin{cases} 1, & r_1 > c(g), \\ d(g), & r_1 = c(g), \\ 0, & \text{elsewhere}, \end{cases} \tag{4.1.19}$$

where $c(g)$ and $d(g)$ are determined by

$$E_{\theta_0}[\varphi_6(R_1, G) | g] = \alpha, \quad \text{for a given } \theta_0 \in \Theta_{B6}.$$

Pf. Using an argument like the one given on page 137 in Lehmann (1959), we are to show that the test

(4.1.20) $$\varphi(r_1, t, g) = \begin{cases} 1, & r_1 > c(t,g), \\ d(t,g), & r_1 = c(t,g), \\ 0, & \text{elsewhere}, \end{cases}$$

where $c(t, g)$ and $d(t, g)$ are determined by

(4.1.21) $$E_{(\mu_0, \theta_0)}[\varphi(R_1, T, G | t, g] = \alpha ,$$

(μ_0, θ_0) an arbitrary but fixed element of Ω_{B6}, has the property of maximizing the conditional power against any $(\mu_1, \theta_1) \in \Omega_{K6}$. It is easily established that the density function of R_1 given $(T, G) = (t, g)$, wrt some measure on R^1 is given by

$$f_{R_1 | (t,g)}(r_1 | \theta) = \psi_{t,g}(\theta) \exp\{\gamma_h(\theta)(\eta_1(\theta) - \eta_0) r_1\}.$$

Let $(\mu_0, \theta_0) \in \Omega_{B6}$ and $(\mu_1, \theta_1) \in \Omega_{K6}$. Applying the Neyman-Pearson lemma, the MP α-test of $H: (\mu, \theta) = (\mu_0, \theta_0)$ vs $K: (\mu, \theta) = (\mu_1, \theta_1)$ rejects when

$$f_{R_1 | (t,g)}(r_1 | \theta_1) / f_{R_1 | (t,g)}(r_1 | \theta_0) > C ,$$

where C is determined so that we have a level α test. But this test is equivalent to (4.1.20) and (4.1.21) because the quantity

$\gamma_h(\theta)(\eta_1(\theta)-\eta_0)$ is positive on Ω_{K6} and is zero on Ω_{B6}. We now show

(4.1.22) $$E_{(\mu,\theta)}[\varphi(R_1, T, G) | t, g] \leq \alpha, \quad \forall\ (\mu, \theta) \in \Omega_{H6}.$$

To this end let $(\mu_2, \theta_2) \in \Omega_{H6}$ and apply the Neyman-Pearson lemma to get a MP test of $H: (\mu, \theta) = (\mu_2, \theta_2)$ vs $K: (\mu, \theta) = (\mu_0, \theta_0)$ with power α at (μ_0, θ_0). This test rejects when $f_{R_1|(t,g)}(r_1 | \theta_0)/f_{R_1|(t,g)}(r_1 | \theta_2) > C_1$, for some constant C_1. But this is also equivalent to (4.1.20) because the quantity $\gamma_h(\theta)(\eta_1(\theta)-\eta_0)$ is nonpositive on Ω_{H6}. Hence by Corollary 1, page 67, in Lehmann (1959), we have

$$E_{(\mu_2, \theta_2)}[\varphi(R_1, T, G) | t, g] \leq E_{(\mu_0, \theta_0)}[\varphi(R_1, T, G) | t, g] = \alpha,$$

and (4.1.22) is established. (When the above quantity is zero, we have the test $\varphi \equiv \alpha$.) Now the class of tests satisfying (4.1.22) is contained in the class of tests satisfying $E_{(\mu_0, \theta_0)}[\varphi(R_1, T, G) | t, g] \leq \alpha$, because $\Omega_{B6} \subset \Omega_{H6}$. Since the test (4.1.20) and (4.1.21) maximizes the condition power at (μ_1, θ_1) within this wider class, it also maximizes the conditional power at (μ_1, θ_1) subject to (4.1.22); since it is independent of the particular alternative (μ_1, θ_1) chosen in Ω_{K6}, it is UMP against K6. Hence the test (4.1.20) and (4.1.21) is UMPU α-test for H6 vs K6. Upon making use of the independence of

T and R, the test (4.1.19) is obtained. []

Remark 4.1.23.

(a) The assumption that $\gamma_h(\theta) < 0, \forall \theta \in \Theta$, is used to establish that the MP test is UMP. So, if $\gamma_h(\theta)$ changes sign on Θ, we have to specify an alternative in order to be able to perform the test. As we mentioned before, the assumption is going to be satisfied in most practical applications. However if $\gamma_h(\theta) > 0, \forall \theta \in \Theta$, the theorem still holds with the inequality in (4.1.19) reversed.

(b) Other hypotheses about $\eta_1(\theta)$, similar to H2, H3, or H4, can also be considered. []

We conclude this section by noting that some of the UMPU tests given above, e.g., (4.1.10), (4.1.11), (4.1.12), and (4.1.19), are also UMP among all tests in the wider class of similar tests.

4.2. The Case Where $\mathcal{P}_R$ is Complete but $\mathcal{P}_T$ is Not

Suppose now that the family $\mathcal{P}_R$ is complete but the family $\mathcal{P}_T$ is not complete. We begin by noting that the family $\mathcal{P}_Y$ is invariant under the group of transformations $\mathcal{G} = \{g_\mu : \mu \in \mathcal{M}\}$, where $g_\mu(y) = y + \mu$. Seely (1972) showed that Z is a maximal invariant wrt this group. Further, R is a minimal sufficient statistic for $\mathcal{P}_Z$. Thus, using a location invariant argument, we are led to consider $\mathcal{P}_R$. Applying the same techniques used in Section 4.1, we arrive at the same tests as above. They are to be interpreted now as UMPU among (location) invariant tests.

<u>Remark 4.2.1.</u> One may be tempted to go with invariance all the way. That is, instead of appealing to unbiasedness, one might reduce by scale invariance and hope to get UMP invariant tests based on the maximal invariant statistic. However, this procedure does not work in general as shown by Herbach (1959). The trouble is that both the maximal invariant and the maximal invariant induced in the parameter space may be vector-valued. So, while testing about one of the components, the others play the role of nuisance parameters. []

V. HYPOTHESIS TESTING WHEN $\mathcal{H}_0$ IS A COMMUTATIVE QUADRATIC SUBSPACE

Recall that $\mathcal{H}_0 = \mathrm{sp}\{W_1, \ldots, W_p\}$ where $W_i = Q'V_iQ$, $i = 1, 2, \ldots, p$. Since $W_1, \ldots, W_p$ is not necessarily a basis for $\mathcal{H}_0$, $H_1, \ldots, H_h$ is used to denote such a basis. When $\mathcal{H}_0$ is a commutative quadratic subspace (CQS), the previous results can be greatly simplified. As a result, we devote this chapter to the problem of hypothesis testing when $\mathcal{H}_0$ is a CQS. We begin by obtaining simple versions of the results in Chapter IV. Then we give some examples to illustrate the wide range of applications covered by these models. We note that the assumption that $\mathcal{H}_0$ is a quadratic subspace is equivalent to the assumption that the family $\mathcal{P}_R$ is complete (see Theorem 3.2.9).

A necessary and sufficient condition for $\mathcal{H}_0$ to be a commutative subspace is that the matrices $W_1, \ldots, W_p$ commute. It appears then that we need to compute the matrix Q in order to be able to verify commutativity. The following proposition shows that this need not be the case.

<u>Proposition 5.1</u>. Suppose that $\mathcal{P}_T$ is complete. Then

(a) $QQ'V_i = V_iQQ'$, $i = 1, 2, \ldots, p$.

(b) $W_1, \ldots, W_p$ commute whenever $V_1, \ldots, V_p$ commute.

Pf. Theorem 3.2.8(b) implies $\Sigma \mathcal{M} \subset \mathcal{M}$ for all $\Sigma \in \mathcal{V}_1$. Since $\mathcal{V}_1 = I + \mathcal{V}_0$, we have $V_i \mathcal{M} \subset \mathcal{M}$, $i = 1, 2, \ldots, p$. Since QQ' is the orthogonal projection operator on $\mathcal{M}^{\perp}$, Theorem 2 of Zyskind (1967) implies that QQ' commutes with each V_i; and (a) is proved. To see (b), observe that for any $i, j = 1, 2, \ldots, p$

$$W_i W_j = Q'V_i QQ'V_j Q = Q'V_i V_j Q = Q'V_j V_i Q = W_j W_i. \quad \square$$

The importance of commutative quadratic subspaces in our work arises from the following lemma due to Seely (1971).

Lemma 5.2. A necessary and sufficient condition for a subspace $\mathcal{B}$ to be a CQS is the existence of a basis $B_1, B_2, \ldots, B_k$ for $\mathcal{B}$ such that each B_i is idempotent and such that $B_i B_j = 0$ for $i \neq j$. Moreover, apart from indexing, such a basis for a CQS is unique.

We now take the basis $H_1, \ldots, H_h$ for $\mathcal{H}_0$ to satisfy

(5.3) $$H_i H_j = H_i, \quad i = j,$$
$$= 0, \quad i \neq j.$$

We also define

(5.4) $$k_i = tr(H_i) = \underline{r}(H_i), \quad i = 1, 2, \ldots, h.$$

The first consequence of (5.3) is the following

Proposition 5.5. Suppose $H_1, \ldots, H_h$ satisfy (5.3). Then, for every $\theta \in \Theta$ and $j = 1, 2, \ldots, h$, we can conclude

(a) $[1+\delta_j(\theta)] = [1+\nu_j(\theta)]^{-1}$.

(b) $[1+\delta_j(\theta)] > 0$.

Pf. Using (5.3) and $I - \Lambda_\theta \Lambda_\theta^{-1} = 0$, we can write

$$\Sigma_{i=1}^{h} [\nu_i(\theta)+\delta_i(\theta)+\nu_i(\theta)\delta_i(\theta)]H_i = 0.$$

Since $H_1, \ldots, H_h$ is a basis, it follows that

$$\nu_i(\theta) + \delta_i(\theta) + \nu_i(\theta)\delta_i(\theta) = 0, \quad i = 1, 2, \ldots, h;$$

and (a) is established. To prove (b), let $x \in \underline{R}(H_j)$ s.t. $x \neq 0$ and $x'x = 1$. Because Λ_θ is positive definite, we get

$$0 < x'\Lambda_\theta^{-1}x = x'x + \Sigma_{i=1}^{h} \delta_i(\theta)x'H_i x$$

$$= x'x + \delta_j(\theta)x'x$$

$$= 1 + \delta_j(\theta). \quad \square$$

Recalling Equation (3.2.4), Proposition 5.5 gives a simple relationship between the $\delta_i(\theta)$ functions and the parameters

$\theta_1, \ldots, \theta_p$. Further, because part (b) of the proposition implies that $\gamma_i(\theta) = -\frac{1}{2}[1+\delta_i(\theta)] < 0$, $\forall\ \theta \in \Theta$, we can construct UMPU tests about ratios, i.e., recall Remark 4.1.23(a). Moreover, part (a) of the proposition implies that the $\gamma_i(\theta)$'s are rational functions. Using this fact and the following lemma, we can strengthen some of the results back in Chapter IV.

Lemma 5.6. Let $\mathcal{X}$ be a set having a nonempty interior in R^k. Let $g: \mathcal{X} \to R^1$ be a nonconstant rational function. For $c \in R^1$, let $G_1 = g^{-1}(-\infty, c]$, $G_2 = g^{-1}(c, \infty)$, and $G_3 = g^{-1}[c]$. Then we can conclude

(a) G_3 has no interior in $\mathcal{X}$.

(b) $G_3 = \bar{G}_1 \cap \bar{G}_2$.

Pf. Write $g(x) = P(x)/Q(x)$, where $P(x)$ and $Q(x)$ are two polynomials in x. Let A be an open subset of $\mathcal{X}$ and suppose $g(x) = c$, $x \in A$. Then $P(x) - cQ(x) = 0$, $x \in A$. It then follows from Theorem 7.4, page 19 in Walker (1950), that $P(x) - cQ(x) = 0$, for all $x \in \mathcal{X}$. Thus, $g = c$ and (a) is proved. To establish (b), we notice that the interior of G_1 is given by $g^{-1}(-\infty, c)$. Hence

$$\begin{aligned}\bar{G}_2 &= \cap\{B : B \text{ is a closed set s.t. } B \supset G_2\}\\ &= (\cup\{A : A \text{ is an open set s.t. } A \subset G_1\})^c\\ &= g^{-1}[c, \infty);\end{aligned}$$

and (b) follows. $\square$

Recall the hypothesis

$$H1: \gamma_1(\theta) \leq \gamma_0 \text{ vs } K1: \gamma_1(\theta) > \gamma_0 .$$

We have shown in Chapter IV that $\Theta_{B1} \subset \{\theta \in \Theta : \gamma_1(\theta) = \gamma_0\}$. Now that the $\gamma_i(\theta)$'s are rational functions we use Lemma 5.6 to conclude that

$$\Theta_{B1} = \{\theta \in \Theta : \gamma_1(\theta) = \gamma_0\} .$$

Likewise, for testing the hypothesis

$$H6: \eta_1(\theta) \geq \eta_0 \text{ vs } K6: \eta_1(\theta) < \eta_0 ,$$

we have

$$\Theta_{B6} = \{\theta \in \Theta : \eta_1(\theta) = \eta_0\} .$$

Another consequence of (5.3) has to do with the distributional properties of the random variables $R_i = Z'H_iZ$, $i = 1, 2, \ldots, h$.

<u>Lemma 5.7</u>. Suppose $H_1, \ldots, H_h$ satisfy (5.3). Then

(a) $R_1, R_2, \ldots, R_h$ are independent.

(b) $[1+\delta_i(\theta)]R_i \sim \chi^2_{k_i}$, $i = 1, 2, \ldots, h$.

<u>Pf</u>. For $j = 1, 2, \ldots, h$ we can write $R_j = (H_jZ)'(H_jZ)$ which says that R_j is a function of H_jZ. So it is sufficient to show

that $H_1Z, \ldots, H_hZ$ are independent. For $i, j = 1, 2, \ldots, h$, it is easy to show, using (5.3), that $\mathrm{cov}(H_iZ, H_jZ) = H_i\Lambda_\theta H_j$ reduces to

$$\begin{aligned} \mathrm{cov}(H_iZ, H_jZ) &= [1+\nu_i(\theta)]H_i, & i &= j, \\ &= 0, & i &\neq j. \end{aligned}$$

Since $H_1Z, \ldots, H_hZ$ are jointly normal, this proves (a). Part (b) follows immediately from Proposition 5.5(a) and the fact that $H_iZ \sim N_q(0, [1+\nu_i(\theta)]H_i)$. []

Lemma 5.7 is the key to the simple tests we give below. The distributional properties given by the lemma result in test procedures which are much easier to carry out than the ones we obtained previously. So, combining Lemma 5.7 and Theorem 4.1.9, we immediately obtain

<u>Theorem 5.8</u>. Let c_1 be s.t. $\Pr\{\chi_1^2 > c_1\} = \alpha$ where χ_1^2 has a chi-square distribution with k_1 degrees of freedom. Set $c = c_1/(-2\gamma_0)$ and suppose the following conditions are satisfied:

(a) the family $\mathcal{P}_T$ is complete,

(b) the matrices $H_1, \ldots, H_h$ satisfy (5.3),

(c) assumption (4.1.6) holds true.

Then the test

$$\varphi(r_1) = \begin{cases} 1, & r_1 > c, \\ 0, & \text{elsewhere}, \end{cases} \tag{5.9}$$

is a UMPU α-test of H1 vs K1.

We also have the following:

Theorem 5.10. Let c_1 be s.t. $\Pr\{\hat{F} > c_1\} = \alpha$, where $\hat{F}$ has an F-distribution with k_1 and k_h degrees of freedom. Set $c = c_1/\eta_0$ and suppose the following conditions are satisfied:

(a) the family $\mathcal{P}_T$ is complete,

(b) the matrices $H_1, \ldots, H_h$ satisfy (5.3),

(c) assumption (4.1.16) holds true.

Then the test

$$\varphi(r_1, r_h) = \begin{cases} 1, & \dfrac{r_1/k_1}{r_h/k_h} > c, \\ 0, & \text{elsewhere}, \end{cases} \tag{5.11}$$

is a UMPU α-test of H6 vs K6.

Pf. Under our assumptions Theorem 4.1.18 is applicable. So, we consider the test (4.1.19). Recall Equation (4.1.15). Given $G = g$, we have

$$r_1 \geq c(g) \Longleftrightarrow [\eta_0 r_1/(\eta_0 r_1 + r_h)] \geq c_1(g),$$

for some $c_1(g)$. Let $X = \eta_0 R_1/(\eta_0 R_1 + R_h)$. Then the test

$$\varphi(x,g) = \begin{cases} 1, & x > c_1(g), \\ d_1(g), & x = c_1(g), \\ 0, & \text{elsewhere}, \end{cases} \tag{5.12}$$

where $c_1(g)$ and $d_1(g)$ are determined by

$$E_{\theta_0}[\varphi(X,G)|g] = \alpha, \quad \text{for a given } \theta_0 \in \Theta_{B6},$$

is a UMPU α-test of H6 vs K6. We want to show

(5.13) X and G are independent on Ω_{B6}.

Let

$$X_1 = Z'([1+\delta_h(\theta)]\eta_0 H_1)Z$$

and

$$X_2 = Z'([1+\delta_h(\theta)]\eta_0 H_1 + [1+\delta_h(\theta)]H_h)Z.$$

Then $X = X_1/X_2$. The joint distribution of X_1 and X_2 is given by the characteristic function

$$\phi(a,b) = |I-2i\Lambda_\theta[(a+b)[1+\delta_h(\theta)]\eta_0 H_1 + b[1+\delta_h(\theta)]H_h]|^{-1/2}.$$

Using (5.3), this reduces to

$$\phi(a,b) = |I-2i[(a+b)\eta_0 H_1/\eta_1(\theta) + bH_h]|^{-1/2}.$$

Hence the distribution of X depends only on $\eta_1(\theta)$. Now Theorem 2, page 162 in Lehmann (1959), implies that X and (T, G) are independent on Ω_{B6}; and (5.13) follows. Therefore, the test (5.12) reduces to

$$\varphi(x) = \begin{cases} 1, & x > c_2, \\ d_2, & x = c_2, \\ 0, & \text{elsewhere}, \end{cases}$$

where c_2 and d_2 are determined by

$$E_{\eta_0}[\varphi(X)] = \alpha .$$

But

$$x = [\eta_0 r_1/(\eta_0 r_1 + r_h)] \geq c_2 \iff \eta_0 r_1/r_h \geq c_3 ,$$

for some c_3; and the F-test (5.11) is obtained by appealing to Lemma 5.7. $\square$

Remark 5.14.

(a) It should be noticed that these simplifications carry over to the tests given by (4.1.11)-(4.1.13) as well as any tests about the ratios corresponding to H2, H3, and H4. For instance, the two sided F-test is UMPU for testing

$$H: \eta_1(\theta) = \eta_0 \text{ vs } K: \eta_1(\theta) \neq \eta_0 .$$

(b) If the family $\mathcal{P}_T$ is not complete, the invariance argument developed in Section 4.2 can be used to interpret the tests as UMPU among (location) invariant tests. □

We devote the rest of the chapter to examples. The examples are chosen to cover a wide range of applications. Moreover we hope to show, by way of these examples, how to use the above results to test some interesting hypotheses concerning the parameters $\theta_1, \ldots, \theta_p$.

Example 5.15. Suppose $Z \sim N_q(0, I+\theta W)$ where $\theta \geq 0$ and $W = W' = W^2$. Then $\mathcal{H}_1 = I + \mathcal{H}_0$ where $\mathcal{H}_0 = \text{sp}\{W\}$. Since $W = W^2$, it follows that $\mathcal{H}_0$ is a CQS. Since $h = 1$, we take $H_1 = W$ so that $R_1 = Z'WZ$. Now $\Lambda_\theta = I + \theta W$. So, $\nu_1(\theta) = \theta$. Thus, $[1+\delta_1(\theta)] = (1+\theta)^{-1}$ which implies

$$\gamma_1(\theta) = -\frac{1}{2}[1+\delta_1(\theta)] = -[2(1+\theta)]^{-1} .$$

Consider testing the hypothesis

$$H: \theta = 0 \text{ vs } K: \theta > 0.$$

This hypothesis can equivalently be written as

$$H: \gamma_1(\theta) = -1/2 \text{ vs } K: \gamma_1(\theta) > -1/2.$$

Hence we can use Theorem 5.8 to construct a UMPU α-test. The test is a chi-square test that rejects iff $r_1 > c$ where c is the appropriate cutoff value. It should be noticed that the chi-square test would also be obtained if $W = W' = aW^2$ where a is some known constant. []

Example 5.16. Consider the model $Y \sim N_n(X\beta, I+\theta V)$ where β is an unknown vector of parameters, $\theta \geq 0$, $V = V' = V^2$, and $\underline{R}(VX) \subset \underline{R}(X)$. We can reduce this model to the one in the previous example. To see this set $Z = Q'Y$ where Q is such that $\underline{R}(Q) = \underline{N}(X')$ and $Q'Q = I$. Then $Z \sim N_q(0, I+\theta W)$ where $q = \underline{r}(Q)$ and $W = Q'VQ$. Now the condition $\underline{R}(VX) \subset \underline{R}(X)$ implies that $QQ'V = VQQ'$, by Proposition 5.1(a). Thus, $W^2 = W$. So in fact we get the model in the previous example. We mention that the present model would arise from a balanced completely random one-way classification model with known error variance equal to one. []

Example 5.17. Let $Z \sim N_q(0, \sigma^2(I+\rho W))$ where $\sigma^2 > 0$, $-1 < \rho < 1$, and $W = W' = W^2$. It is easy to see that $\mathcal{H}_1 = \mathcal{H}_0 = \mathrm{sp}\{I, W\}$ is a CQS. Assuming $0 < \underline{r}(W) < q$, we get that $h = 2$ and we can take $H_1 = I - W$ and $H_2 = W$ to form a basis satisfying (5.3). Then $R_1 = Z'Z - Z'WZ$ and $R_2 = Z'WZ$. Let $\theta = (\sigma^2, \rho)'$. Then $\Lambda_\theta = \sigma^2 I + \rho\sigma^2 W$ can be written as

$$\Lambda_\theta = I + (\sigma^2 - 1)H_1 + (\sigma^2(1+\rho) - 1)H_2 .$$

Thus $\nu_1(\theta) = \sigma^2 - 1$ and $\nu_2(\theta) = \sigma^2(1+\rho) - 1$ so that $1 + \delta_1(\theta) = 1/\sigma^2$ and $1 + \delta_2(\theta) = 1/[\sigma^2(1+\rho)]$. Thus $\eta_1(\theta) = 1 + \rho$. Let us consider testing the hypothesis

$$H: \rho = 0 \text{ vs } K: \rho \neq 0.$$

This hypothesis can equivalently be stated as

$$H: \eta_1(\theta) = 1 \text{ vs } K: \eta_1(\theta) \neq 1.$$

To see if the two sided version of the F-test (5. 11) can be used we must check if the assumption corresponding to (4. 1. 16) is satisfied. Now observe that

$$\begin{aligned} \Gamma_B &= \{\gamma_2(\theta) : \theta \in \Theta_B\} \\ &= \{-1/2\sigma^2(1+\rho) : \sigma^2 > 0,\ -1 < \rho < 1,\ 1 + \rho = 1\}, \\ &= \{-1/2\sigma^2 : \sigma^2 > 0\} , \end{aligned}$$

which clearly shows that the two sided F-test is UMPU. Let $F = R_1 k_2 / R_2 k_1$ where $k_1 = \underline{r}(I-W) = q - \underline{r}(W)$ and $k_2 = \underline{r}(W)$. Hence the test is given by

$$\varphi(f) = \begin{cases} 1, & f < c_1 \text{ or } f > c_2, \\ 0, & \text{elsewhere,} \end{cases}$$

where c_1 and c_2 are determined by

$$E_{\eta_1(\theta)=1}[\varphi(F)] = \alpha$$

and

$$E_{\eta_1(\theta)=1}[F\varphi(F)] = \alpha E_{\eta_1(\theta)=1}(F).$$

Transforming the F-distribution into a beta distribution, tables of the incomplete beta function can be used to determine c_1 and c_2. []

Example 5.18. Suppose $Y_{ij} \sim N(\beta_i, \theta_i)$ where $\beta_i \in R^1$ and $\theta_i > 0$ for $i = 1, 2$, $j = 1, \ldots, n_i$. Also, suppose that the Y_{ij} are independent. This model can be written in matrix form as $Y \sim N(X\beta, \theta_1 V_1 + \theta_2 V_2)$ where

$$X = \begin{pmatrix} 1_{n_1} & 0 \\ 0 & 1_{n_2} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad V_1 = \begin{pmatrix} I_{n_1} & 0 \\ 0 & 0 \end{pmatrix}, \quad V_2 = \begin{pmatrix} 0 & 0 \\ 0 & I_{n_2} \end{pmatrix}.$$

Let Q be such that $\underline{R}(Q) = \underline{N}(X')$ and $Q'Q = I$. Set $W_i = Q'V_iQ$, $i = 1, 2$. It is easily verified that $\underline{R}(V_iX) \subset \underline{R}(X)$ so that $\mathcal{P}_T$ is complete so that Proposition 5.1 is applicable. Hence $\mathcal{H}_1 = \mathcal{H}_0 = \text{sp}\{W_1, W_2\}$ is a CQS. We note that W_1 and W_2 satisfy (5.3) so we can take $H_i = W_i$, $i = 1, 2$. Then, partitioning Y into Y_1 and Y_2, we have

$$R_i = Z'W_iZ = Y'QQ'V_iQQ'Y$$
$$= Y'(I-P_X)V_i(I-P_X)Y$$
$$= Y'V_iY - (P_XY)'V_i(P_XY)$$
$$= Y_i'Y_i - n_i\bar{Y}_{i.}^2 ,$$

where P_X is the orthogonal projection operator on $\underline{R}(X)$ and $\bar{Y}_{i.}$ is the average of the observations which comprise Y_i, $i = 1,2$. Further $k_i = tr(H_i) = tr(V_i) - tr(P_XV_i) = n_i - 1$, $i = 1, 2$. Let $\Sigma_\theta = \theta_1V_1 + \theta_2V_2$. Then $\Lambda_\theta = Q'\Sigma_\theta Q = \theta_1H_1 + \theta_2H_2$ can be written as $\Lambda_\theta = I + (\theta_1-1)H_1 + (\theta_2-1)H_2$, Thus, $\nu_i(\theta) = (\theta_i-1)$ so that $1 + \delta_i(\theta) = 1/\theta_i$, $i = 1, 2$. So $\eta_1(\theta) = \theta_2/\theta_1$. Consider testing the hypothesis:

$$H:\theta_1 = \theta_2 \text{ vs } K:\theta_1 \neq \theta_2.$$

This hypothesis can equivalently be stated as

$$H:\eta_1(\theta) = 1 \text{ vs } K:\eta_1(\theta) \neq 1.$$

As we have done in Example 5.17 we need to check if Γ_B has a nonempty interior in R^1. Observe that

$$\Gamma_B = \{\gamma_2(\theta) : \theta \in \Theta_B\},$$
$$= \{-1/2\theta_2 : \theta_1 > 0,\ \theta_2 > 0,\ \theta_1 = \theta_2\},$$
$$= \{-1/2\theta_2 : \theta_2 > 0\}.$$

So, the usual two sided F-test is UMPU. []

Example 5.19. Consider a balanced incomplete block design with treatments fixed and blocks random. Let v, b, k, r, and λ denote the parameters associated with the design and express the $n = rv = bk$ observations $Y_{ij} = \mu + \gamma_i + \tau_j + e_{ij}$ in matrix form as $Y = 1\mu + B\gamma + X\tau + e$. Assume that $\gamma \sim N_b(0, \sigma_b^2 I)$, $e \sim N_n(0, \sigma^2 I)$ and that γ and e are independent. Then $Y \sim N_n(1\mu + X\tau, \sigma^2 I + \sigma_b^2 BB')$. The case $k = v$ corresponds to the balanced complete block design which is a special case of the models to be treated in the next example. So we consider here the case $k < v$. In this latter case Seely (1971) showed that it is no longer true that $\underline{R}(BB'X) \subset \underline{R}(X)$. As a result $\mathcal{P}_T$ is not complete. Hence we reduce by (location) invariance. Consider $Z = Q'Y$ where Q is such that $\underline{R}(Q) = \underline{N}(X')$ and $Q'Q = I$. Then $Z \sim N_q(0, \sigma^2 I + \sigma_b^2 Q'BB'Q)$ where $q = \underline{r}(Q) = n - v$. So, $\mathcal{H}_1 = \mathcal{H}_0 = sp\{I, Q'BB'Q\}$. Now suppose that $b = v$. The reader is referred to Seely (1971) for a proof of the results we give below. When $b = v$, $\mathcal{H}_0$ is a CQS and $H_1 = (r/\lambda v)Q'BB'Q$ along with $H_2 = I - H_1$ form a basis satisfying (5.3). Further, $R_1 = Z'H_1Z$ is the sum of squares for blocks eliminating treatments, and $R_2 = Z'H_2Z$ is the sum of squares for intrablock error. Moreover, $k_1 = tr(H_1) = v - 1$ and $k_2 = tr(H_2) = n - 2v + 1$. Let

$\theta = (\sigma^2, \sigma_b^2)'$. Then $\Lambda_\theta = \sigma^2 I + \sigma_b^2 Q'BB'Q$ can be written as

$$\Lambda_\theta = I + (\sigma^2 + \lambda v \sigma_b^2/r - 1)H_1 + (\sigma^2 - 1)H_2 .$$

Thus $\nu_1(\theta) = \sigma^2 + \lambda v \sigma_b^2/r - 1$ and $\nu_2(\theta) = \sigma^2 - 1$ so that $1 + \delta_1(\theta) = 1/(\sigma^2 + \lambda v \sigma_b^2/r)$ and $1 + \delta_2(\theta) = 1/\sigma^2$. Hence $\eta_1(\theta) = 1/(1 + \lambda v \rho/r)$ where $\rho = \sigma_b^2/\sigma^2$. Now that $\eta_1(\theta)$ is a function of ρ, we can test about the ratio σ_b^2/σ^2. One interesting hypothesis is

$$H: \sigma_b^2 = 0 \text{ vs } K: \sigma_b^2 > 0,$$

which can equivalently be stated as

$$H: \eta_1(\theta) = 1 \text{ vs } K: \eta_1(\theta) < 1.$$

It is easy to check that the F-test (5. 11) is a UMPU test among (location) invariant tests. □

Let us now consider the model for the complete m-way classification with p random effects. This model can be expressed as

(5. 20) $$Y \sim N(X\beta, \Sigma_{i=1}^p \theta_i V_i) .$$

Several special cases are treated in the literature:

(a) The balanced completely random one (and two) way classification models are dealt with by Herbach (1959).

(b) The balanced completely random q-way classification model is treated by Graybill (1976).

(c) Scheffé (1956) treated the mixed model for the complete two-way classification with one random-effects factor. However, under his assumptions $\mathcal{H}_0$ is not a CQS. For a discussion on this point see Hocking (1973).

(d) The generalization of Scheffé's work to the mixed model for the complete three-way classification with two random-effects factors is done by Imhof (1960).

In the following example we use a model of the form (5.20) to illustrate our methods.

Example 5.21. 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 that μ and α_i are fixed while $\beta_j,\ldots,(\alpha\beta\gamma)_{ijk}$ are assumed to be independent and normally distributed with zero means and variances $\sigma_b^2,\ldots,\sigma_{abc}^2$ respectively. In matrix form we have

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

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

$$\mathrm{cov}(Y) = ac\sigma_b^2 P_B + c\sigma_{ab}^2 P_{AB} + ab\sigma_c^2 P_C + b\sigma_{ac}^2 P_{AC} + a\sigma_{bc}^2 P_{BC} + \sigma_{abc}^2 I ,$$

where we have written P_{AB}, for example, to denote the orthogonal projection on $\underline{R}(U_{AB})$ for notational convenience. In the above expression for $\mathrm{cov}(Y)$ we have used facts like $BB' = acP_B$. It is clear that this model satisfies (5.20) where X, β, θ_i's and V_i's are defined in an obvious fashion, e.g., $\theta_1 = ac\sigma_b^2$, $V_1 = P_B$. In this balanced situation the above orthogonal projections satisfy some interesting relationships that prove to be useful. For instance,

(5.22) $$P_A P_1 = P_A P_B = P_A P_{BC} = P_1 \quad \text{and} \quad P_{AB} P_{AC} = P_A ,$$

and roles of A, B or A, C can be interchanged. Let Q be such that $\underline{R}(Q) = \underline{N}(A')$ and $Q'Q = I$. Set $W_i = Q'V_iQ$, $i = 1, 2, \ldots, 6$. Using (5.22) it is easy to verify that $\mathcal{V}_T$ is complete and that $\mathcal{H}_0 = \mathrm{sp}\{W_1, \ldots, W_6\}$ is a CQS. We want to find a basis $H_1, \ldots, H_6$ for $\mathcal{H}_0$ satisfying (5.3). To this end consider

$$G_1 = P_B - P_1 ,$$
$$G_2 = P_{AB} - P_A - P_B + P_1 ,$$
$$G_3 = P_C - P_1 ,$$
$$G_4 = P_{AC} - P_A - P_C + P_1 ,$$
$$G_5 = P_{BC} - P_B - P_C + P_1 ,$$
$$G_6 = I - P_{AB} - P_{AC} - P_{BC} + P_A + P_B + P_C - P_1 .$$

We note that $G_i^2 = G_i$ and $G_i G_j = 0$, $i \neq j$. Moreover, for $i = 1,2,\ldots,6$ we have $QQ'G_i = (I-P_A)G_i = G_i$. Hence $H_i = Q'G_iQ$, $i = 1,2,\ldots,6$, form the required basis. Thus $R_i = Z'H_iZ = Y'QQ'G_iQQ'Y = Y'G_iY$, $i = 1,2,\ldots,6$. We mention that the R_i's are the sums of squares appearing in the usual analysis of variance tables. For example, $R_2 = Y'(P_{AB}-P_A-P_B+P_1)Y$ is the sum of squares for the interaction $(\alpha\beta)$. To get $\nu_1(\theta),\ldots,\nu_6(\theta)$ as given by (3.2.4), we observe that $W_1 = H_1$, $W_2 = H_1 + H_2$, $W_3 = H_3$, $W_4 = H_3 + H_4$, $W_5 = H_1 + H_3 + H_5$ and $W_6 = \Sigma_{i=1}^6 H_i$. From these observations we can conclude

$$
\begin{aligned}
1 + \nu_1(\theta) &= ac\sigma_b^2 + c\sigma_{ab}^2 + a\sigma_{bc}^2 + \sigma_{abc}^2 , \\
1 + \nu_2(\theta) &= c\sigma_{ab}^2 + \sigma_{abc}^2 , \\
1 + \nu_3(\theta) &= ab\sigma_c^2 + b\sigma_{ac}^2 + a\sigma_{bc}^2 + \sigma_{abc}^2 , \\
1 + \nu_4(\theta) &= b\sigma_{ac}^2 + \sigma_{abc}^2 , \\
1 + \nu_5(\theta) &= a\sigma_{bc}^2 + \sigma_{abc}^2 , \\
1 + \nu_6(\theta) &= \sigma_{abc}^2 .
\end{aligned}
$$

We see that the usual F-tests are UMPU for testing about the interaction effects. They are also UMPU for testing about the main effects provided we are willing to assume that certain interactions do not play any role, e.g., to test about σ_b^2 we must assume that either $\sigma_{ab}^2 = 0$ or $\sigma_{bc}^2 = 0$. []

As we have seen in the above example, there may not be available exact F-tests for some of the hypotheses of interest. However, we may be able to use an approximation due to Satterthwaite (1946) to get an approximate F-test as suggested by Cochran (1951). Cochran showed that this approximate F-test is asymptotically equivalent to the likelihood ratio test. To study the adequacy of the approximation, he considered the case of three parameters. He investigated the distribution of the test statistic as well as the power function and concluded that the proposed F-test is quite satisfactory when three parameters are involved. An improvement of this approximation has been given by Howe and Myers (1970) but their method is too complicated to be useful in practice. This led Davenport and Webster (1972) to establish criteria for the use of Cochran's test. They gave a chart for acceptable combinations of the degrees of freedom. The cases where the approximation is poor were found to be those when the descrepancy in the degrees of freedom is large. Naik (1974a, b) gave test procedures which control the size of the tests very effectively. He showed that his tests have monotone power functions and are most powerful in a certain class of tests.

VI. HYPOTHESIS TESTING WHEN THE SUFFICIENT STATISTIC IS NOT COMPLETE

In this chapter we consider the case where neither the family $\mathcal{P}_R$ nor the family $\mathcal{P}_T$ is complete. We begin by giving a brief review of the literature. Then we consider the case where we have two parameters, $p = 2$. Different tests are discussed and some examples are given. Of these tests, the most practical appears to be Wald's test. As a result we next investigate the power function of Wald's test. The unbalanced random one-way classification model is worked out as an important example of our considerations. We conclude with a brief discussion of the case where we have more than two parameters describing the covariance structure.

6.1. Review of the Literature

In 1947 Wald suggested a procedure that can be used to test hypotheses about ratios in many mixed linear models. In essence, one simply acts as if one has a fixed-effects model and computes the sums of squares in the usual way, i.e., using the reduction sum of squares principle. Many statisticians have used this procedure ever since. Thompson (1955b) applied it to incomplete block designs, while Spjotvoll (1968) and Thomsen (1975) treated unbalanced variance components models for two-way layouts. Both Spjotvoll and Thomsen assumed that the variance component of the interaction effect is zero when testing about main effects. Thomsen's tests differ from those of Spjotvoll in that he pooled the interaction sum of squares and the error sum of squares when testing about main effects. Portnoy (1973) considered experimental designs with randomized blocks. He suggested an improved test based on the recovery of intrablock information. Hultquist and Thomas (1975) re-examined the consideration of Portnoy in a more general formulation of the problem. It turns out that Portnoy's test and many of the tests of Hultquist and Thomas are just the ones we would obtain by applying Wald's procedure.

Although Wald's test procedure is a general one, other tests have been suggested for dealing with specific models. Thompson (1955a) derived a test for the ratio of variances in a mixed incomplete block model which maximizes the minimum power among all invariant tests.

Spjotvoll (1967) derived a most powerful invariant and similar test for the ratio of variances in an unbalanced random one-way classification model. He also discussed the optimum properties of his test. A jackknifed version of Spjotvoll's test which is not sensitive to departures from normality has been proposed by Arvesen and Layard (1975). An exact procedure for testing that the ratio is zero is given by Green (1975). However, his test is not symmetric in the observations.

Tietjen (1974) dealt with the completely random three stage nested model. He considered testing about the variance component of the top stage factor. He compared a conventional F-test (ignoring the imbalance) with another approximate test based on a Satterthwaite-like procedure. A Monte Carlo study revealed that the conventional F-test behaves far better than the other test. Tietjen also found that the statistic of the conventional F-test follows approximately an F-distribution when $.88 \leq r_1 \leq 1.13$, where the quantity r_1 is available from the design and is equal to one in the balanced case.

6.2. The Case of Two Parameters

In the present section we suppose that

$$\mathcal{P}_Y = \{N_n(X\beta, \theta_1 I + \theta_2 V) : \beta \in R^m, \theta = (\theta_1, \theta_2) \in \Theta\},$$

where V is a nonnegative definite matrix and Θ is a set having a nonempty interior in R^2 such that the covariance matrix, $\Sigma_\theta = \theta_1 I + \theta_2 V$, is positive definite for each $\theta \in \Theta$. We also assume that $\underline{r}(X, V) < n$. With regard to our previous notation we have $\mathcal{M} = \underline{R}(X)$, $p = 2$, $V_0 = 0$, $V_1 = I$ and $V_2 = V$.

We shall treat here the case where neither the family $\mathcal{P}_R$ nor the family $\mathcal{P}_T$ is complete. However, the unlikely case where $\mathcal{P}_R$ is not complete but $\mathcal{P}_T$ is complete is still covered by our discussion although we need a different interpretation. The tests to be derived in this section are going to be most powerful similar and (location) invariant tests. The reason for reducing by invariance is to avoid conditioning on T. The independence of T and R, when $\mathcal{P}_T$ is complete, takes care of this problem. As a result, the restriction to invariant tests is not needed when $\mathcal{P}_T$ is complete.

Let Q be such that $\underline{R}(Q) = \underline{N}(X')$ and $Q'Q = I$. We have seen before that $Z = Q'Y$ is a maximal invariant wrt the group $\mathcal{G}$ of translations defined in Section 4.2. The covariance matrix of Z is given by

$$\Lambda_\theta = Q'\Sigma_\theta Q = \theta_1 I + \theta_2 W,$$

where $W = Q'VQ$. Then $h = 2$ and $\mathcal{H}_1 = \mathcal{H}_0 = \mathrm{sp}\{I, W\}$. Let $W = \Sigma_{i=0}^{s} \lambda_i E_i$ denote the spectral decomposition of the matrix W. Also suppose that $\lambda_0 < \lambda_1 < \ldots < \lambda_s$. The assumption that $\underline{r}(X, V) < n$ implies that $\lambda_0 = 0$ (see Olsen, Seely and Birkes (1976)). It can be shown that $\mathcal{G}_1 = \mathcal{G}_0 = \mathrm{sp}\{E_0, \ldots, E_s\}$ and $c = s + 1$. Thus a minimal sufficient statistic for $\mathcal{P}_Z$ is given by

(6.2.1) $$R = (R_0, \ldots, R_s)' \quad \text{where} \quad R_i = Y'QE_iQ'Y, \; i = 0, 1, \ldots, s.$$

When $\mathcal{P}_R$ is complete we have $c = h$, i.e., $s = 1$. Since this case has been considered, we assume that $s > 1$. Observe that $E_0, \ldots, E_s$ satisfy (5.3) so that $\mathcal{G}_0$ is a CQS. Thus for

$i = 0, 1, \ldots, s$ we have the following results:

(6.2.2) $$1 + \delta_i(\theta) = 1/(\theta_1 + \lambda_i \theta_2) ,$$

(6.2.3) $$\theta_1 + \lambda_i \theta_2 > 0,$$

(6.2.4) $$R_i/(\theta_1 + \lambda_i \theta_2) \sim \chi^2_{k_i} , \quad \text{where} \quad k_i = tr(E_i) ,$$

(6.2.5) $$R_0, R_1, \ldots, R_s \quad \text{are independent.}$$

Since $\lambda_0 = 0$, then $\theta_1 > 0$ and we can define $\rho = \theta_2/\theta_1$. It is easily established that the density function of R wrt some measure on R^{s+1} is given by

$$f_R(r|\theta) = a(\theta) \exp\{-1/2 \Sigma_{i=0}^s r_i/(\theta_1 + \lambda_i \theta_2)\} ,$$

$$= a(\theta) \exp\{(-1/2\theta_1)[r_0 + \Sigma_{i=1}^s r_i/(1+\lambda_i \rho)]\} ,$$

where $a(\theta)$ is a normalizing constant.

Consider testing the hypothesis

$$H7: \rho \leq \rho_0 \text{ vs } K7: \rho > \rho_0 ,$$

where ρ_0 is selected such that the class of distributions under H7 and that under K7 are nonvoid. The set of common accumulation points of Θ_{H7} and Θ_{K7} is given by

$$\Theta_{B7} = \bar{\Theta}_{H7} \cap \bar{\Theta}_{K7} = \{\theta \in \Theta : \rho = \rho_0\} ,$$

as can easily be verified using Lemma 5.6. Let

$$G = R_0 + \Sigma_{i=1}^{s} R_i /(1+\lambda_i \rho_0) .$$

Assuming that

(6.2.6) $\{\theta_1 : (\theta_1, \rho_0 \theta_1) \in \Theta_{B7}\}$ has a nonempty interior in R^1,

one can easily establish that

(6.2.7) G is a complete sufficient statistic for $\mathcal{P}_R$ on Θ_{B7}.

From now on we change the notation a little and use T to denote the test statistic

$$(6.2.8) \qquad T = \frac{\Sigma_{i=1}^{s}[1/(1+\lambda_i \rho_0) - 1/(1+\lambda_i \rho_1)]R_i}{R_0 + \Sigma_{i=1}^{s} R_i /(1+\lambda_i \rho_0)} ,$$

where $\rho_1 > \rho_0$. Going along the lines given in the proof of Theorems 4.1.18 and 5.10, we arrive at

<u>Theorem 6.2.9.</u> Suppose that the assumption (6.2.6) holds true. Then the test

$$\varphi(t) = \begin{cases} 1, & t > c, \\ d, & t = c, \\ 0, & t < c, \end{cases} \tag{6.2.10}$$

where c and d are determined by

$$E_{\rho_0}[\varphi(T)] = \alpha,$$

is a MP (against a given $\rho_1 > \rho_0$) similar and (location) invariant α-test for H7 vs K7.

Remark 6.2.11.

(a) Let $\beta(\rho|\rho_1)$ denote the power function of the test (6.2.10). Then $\beta(\rho|\rho_1)$ is an increasing function of ρ. This can be shown using an argument similar to that given by Spjotvoll (1967) who treated the special case of the unbalanced random one-way classification model. In particular it is seen that the test (6.2.10) is unbiased.

(b) The test (6.2.10) can also be considered as a MP (location and scale) invariant α-test. To show this we reduce by scale invariance instead of appealing to similarity. This leads to $(R_1/R_0, \ldots, R_s/R_0)$ as a maximal invariant. The distribution of the maximal invariant depends only on ρ. Hence any test which depends on the maximal invariant must be similar on Θ_{B7}. Thus the class of (location and scale) invariant tests is contained in the class of similar and (location) invariant tests. But the test (6.2.10), which is optimum in the larger class, depends only on the maximal invariant and the claim is established. []

It is seen that T depends on ρ_1 and hence no UMP test exists. To avoid specifying an alternative ρ_1 several ways are possible:

Locally MP Test. Let $\rho_1 = \rho_0 + \delta$ and $\delta \to 0$. This leads (see Equation (36), page 107, in Cox and Hinkley (1974)) to the locally most powerful test statistic

$$T_L = \frac{\Sigma_{i=1}^{s} \lambda_i R_i/(1+\lambda_i \rho_0)^2}{R_0 + \Sigma_{i=1}^{s} R_i/(1+\lambda_i \rho_0)} \, .$$

Wald's Test. Let $\rho_1 \to \infty$. This leads to Wald's test statistic

$$T_W = [\Sigma_{i=1}^{s} R_i/(1+\lambda_i \rho_0)]/R_0 .$$

β-Optimal Test. Let $\bar{\rho}$ be s.t. $\beta(\bar{\rho}|\bar{\rho}) = \beta$, where $0 \leq \alpha < \beta \leq 1$. This leads (see Davies (1969)) to the β-optimal test statistic

$$T_\beta = T \text{ evaluated at } \bar{\rho} .$$

It should be noticed that the above three tests are nearly equivalent when the λ_i's are nearly equal. In fact, all three of them coincide in the balanced case where $s = 1$. In the unbalanced case, however, they are different. The one to use depends on what type of alternatives one wishes to guard against. That is, to guard against close alternatives T_L is better than the others whereas T_W is better against large alternatives.

However, a decisive factor for this issue seems to be the simplicity of the test procedure. Of these three tests, the simplest null distribution is that of T_W. Making use of (6.2.4) and (6.2.5) we have

(6.2.12) when $\rho = \rho_0$, $k_0 T_W / k \sim F_{k, k_0}$, where $k = \Sigma_{i=1}^{s} k_i$.

For T_L and T_β, the null distribution is that of the ratio of two dependent weighted sums of χ^2 random variables. So, even the determination of the cutoff value turns out to be a formidable task.

In the special, but still important, case where $\rho_0 = 0$, the test statistics simplify a little bit. They become

$$T_{L,0} = \Sigma_{i=1}^{s} \lambda_i R_i / \Sigma_{i=0}^{s} R_i ,$$

$$T_{W,0} = \Sigma_{i=1}^{s} R_i / R_0 ,$$

and

$$T_{\beta,0} = [\Sigma_{i=1}^{s}\lambda_i R_i/(1+\lambda_i\bar{\rho})]/\Sigma_{i=0}^{s} R_i .$$

The simplest to compute is again Wald's test.

Remark 6.2.13. Let B be such that $V = BB'$. Then our model is equivalent to the model

$$(6.2.14) \qquad Y = X\beta + Bb + e ,$$

where $b \sim N(0, \theta_2 I)$ independently of $e \sim N(0, \theta_1 I)$. Now think of the model (6.2.14) as a fixed-effects model, i.e., as if b is an unknown parameter. Since QE_0Q' is the orthogonal projection operator on $\underline{R}(X, B)^{\perp}$ (see Olsen Seely and Birkes (1976)) then $R_0 = Y'QE_0Q'Y$ can be interpreted as the residual sum of squares for the full model (6.2.14). Further, QQ' is the orthogonal projection operator on $\underline{R}(X)^{\perp}$ so $Y'QQ'Y$ is the residual sum of squares for the reduced model $Y = X\beta + e$. Hence $\Sigma_{i=1}^{s} R_i = Y'QQ'Y - R_0$ can be interpreted as the reduction sum of squares. Thus we can compute $T_{W,0}$ without having to compute the individual R_i's. Moreover, the degrees of freedom are given by $k_0 = tr(E_0) = n - \underline{r}(X, V)$ and $k = \Sigma_{i=1}^{s} k_i = \underline{r}(X, V) - \underline{r}(X)$. □

Although the above discussion is restricted to H7 vs K7 it should be noticed that other hypotheses can also be tested, e.g., $H: \rho = 0$ vs $K: \rho \neq 0$. The treatment is essentially the same and the modifications are almost obvious.

We conclude the section by commenting on the applicability of the above model. As a first example we mention the unbalanced two variance components model. A variation of the correlation model, Example 5.17, provides a second example. In this latter example we assume now that W is nonnegative definite instead of being idempotent. We also assume, without loss of generality, that the eigenvalues of W are less than or equal to one in order to ensure that the covariance matrix $\sigma^2(I+\rho W)$ is going to be positive definite for $-1 < \rho < 1$. If the largest eigenvalue of W is equal to $M > 1$, one can always scale it back to one. A third example is a slightly different version of the homogeneity of variance problem discussed in Example 5.18. We consider this example in more detail.

Example 6.2.15. Suppose $Y_{ij} \sim N(\mu, \sigma_i^2)$ where $\mu \in R^1$, $\sigma_i^2 > 0$, $i = 1,2$, $j = 1,\ldots,n_i$ and that the Y_{ij}'s are independent. This model can be written in matrix form as

$$Y \sim N(X\mu, \sigma_1^2 V_1 + \sigma_2^2 V_2) ,$$

where $X = 1_n$, $n = n_1 + n_2$ and V_1 and V_2 are defined in Example 5.18. We also assume that $n_i > 1$, $i = 1,2$. The covariance matrix $\sigma_1^2 V_1 + \sigma_2^2 V_2$ does not fall into our framework because $V_1 \neq I$. So we reparametrize. Let $\theta_1 = \sigma_1^2$ and $\theta_2 = \sigma_2^2 - \sigma_1^2$.

Thus $\Theta = \{(\theta_1, \theta_2) : \theta_1 > 0, \theta_2 \in R^1 \text{ and } \theta_1 + \theta_2 > 0\}$. Further, the covariance matrix $\Sigma_\theta = \sigma_1^2 V_1 + \sigma_2^2 V_2 = \theta_1 I + \theta_2 V_2$ satisfies our requirements. It is easily verified that $\underline{R}(V_2X)$ is not a subspace of $\underline{R}(X)$ so that $\mathcal{P}_T$ is not complete. The family $\mathcal{P}_R$ is also not complete. To show this we consider the eigenvalues of the $(n-1) \times (n-1)$ matrix $Q'V_2Q$. Since $n_2 < n-1$, then $\lambda_0 = 0$. The nonzero eigenvalues of $Q'V_2Q$ are the same as those of $V_2QQ'V_2 = V_2 - \frac{1}{n} V_2 J_n V_2$. So we consider the eigenvalues of $I_{n_2} - \frac{1}{n} J_{n_2}$. They satisfy $J_{n_2} x = -n(\lambda-1)x$, where x is the eigenvector associated with λ. Hence, either $-n(\lambda-1) = 0$ or $-n(\lambda-1) = n_2$ so that $\lambda_1 = n_1/n$ and $\lambda_2 = 1$. Thus $s = 2$. It is straightforward to verify that $R_0 = Y_1'Y_1 - n_1\bar{Y}_{1\cdot}^2$, $R_1 = (n_1n_2/n)(\bar{Y}_{1\cdot} - \bar{Y}_{2\cdot})^2$, and $R_2 = Y_2'Y_2 - n_2\bar{Y}_{2\cdot}^2$. The degrees of freedom are given by $k_0 = n_1 - 1$, $k_1 = 1$, and $k_2 = n_2 - 1$. Suppose we want to test the hypothesis

$$H: \sigma_1^2 = \sigma_2^2 \text{ vs } K: \sigma_1^2 \neq \sigma_2^2 .$$

Then a two sided version of any of the above tests can be used. □

6.3. The Power Function of Wald's Test

The discussion in the previous section points out that in comparison with the other tests, Wald's test is by far the simplest to use.

For this reason attention will be focused in this section on the power properties of Wald's test in an attempt to determine the circumstances under which Wald's test will be a reasonable one.

Let c be such that $\Pr\{\hat{F} > c\} = \alpha$, where $\hat{F}$ has an F-distribution with $k = \Sigma_{i=1}^{s} k_i$ and k_0 degrees of freedom. Let Q_i be a random variable having a chi-square distribution with k_i degrees of freedom. The power function of Wald's test depends only on $\rho = \theta_2/\theta_1$ and is given by

$$\beta(\rho) = \Pr\{(\Sigma_{i=1}^{s} \gamma_i(\rho)Q_i)/Q_0 > (kc/k_0)\}, \tag{6.3.1}$$

where

$$\gamma_i(\rho) = (1+\lambda_i\rho)/(1+\lambda_i\rho_0), \quad i = 1, 2, \ldots, s. \tag{6.3.2}$$

The exact power function given in (6.3.1) involves the distribution of the ratio of a weighted sum of chi-square variables to an independent chi-square variable. Unfortunately it is too complicated to allow a useful analytic investigation. As a result we use an approximation due to Satterthwaite (1946). The Satterthwaite approximation to the distribution of $\Sigma_{i=1}^{s} \gamma_i(\rho)Q_i$ is the distribution $g(\rho)Q_{m(\rho)}$ where $Q_{m(\rho)}$ is a chi-square random variable with $m(\rho)$ degrees of freedom and the quantities $g(\rho)$ and $m(\rho)$ are chosen such that the first two moments of the approximate distribution are equal to those of the exact one. This condition leads to the following relationships

$$(6.3.3) \qquad g(\rho) = \Sigma_{i=1}^{s} \gamma_i^2(\rho) k_i / \Sigma_{i=1}^{s} \gamma_i(\rho) k_i ,$$

and

$$(6.3.4) \qquad m(\rho) = (\Sigma_{i=1}^{s} \gamma_i(\rho) k_i)^2 / \Sigma_{i=1}^{s} \gamma_i^2(\rho) k_i .$$

Therefore, we get

$$(6.3.5) \qquad \beta(\rho) \simeq \Pr\{Q_{m(\rho)}/Q_0 > kc/k_0 g(\rho)\}$$

$$= I_{b(\rho)}(k_0/2, m(\rho)/2) ,$$

where $I_x(p, q)$ denotes the incomplete beta ratio, and

$$(6.3.6) \qquad b(\rho) = k_0 g(\rho)/(kc+k_0 g(\rho)) .$$

By examining (6.3.5) it is seen that we get good power properties with high values of k_0, $m(\rho)$ and $b(\rho)$. It should be noted that high values of $b(\rho)$ will result from high values of $g(\rho)$.

We now concentrate on the most common hypothesis

$$H: \rho \leq 0 \text{ vs } K: \rho > 0,$$

and examine the power in a design framework, i.e., as though one has control over the experiment. This will lead to some design considerations which one can employ if one is fortunate enough to be able to design the experiment. It will also tell us when we can expect Wald's test to be good.

Putting $\rho_0 = 0$ in (6.3.2) we obtain

(6.3.7) $$\gamma_i(\rho) = 1 + \lambda_i\rho, \quad i = 1, 2, \ldots, s.$$

Let $a_1 = \Sigma_{i=1}^{s}\lambda_i k_i$ and $a_2 = \Sigma_{i=1}^{s}\lambda_i^2 k_i$. Then substituting (6.3.7) into (6.3.3) and (6.3.4) we get

(6.3.8) $$g_0(\rho) = 1 + \frac{a_1\rho + a_2\rho^2}{a_1\rho + k},$$

and

(6.3.9) $$m_0(\rho) = \frac{a_1^2\rho^2 + k(2a_1\rho + k)}{a_2\rho^2 + (2a_1\rho + k)}.$$

<u>Remark 6.3.10</u>. It should be noticed that $g_0(\rho)$ and $m_0(\rho)$ depend on the λ_i's only through the two quantities a_1 and a_2 which are easy to compute. In fact we have

$$a_1 = \Sigma_{i=1}^{s}\lambda_i k_i = \mathrm{tr}(Q'VQ) = \mathrm{tr}(V(I-P_X)),$$

and

$$a_2 = \Sigma_{i=1}^{s}\lambda_i^2 k_i = \mathrm{tr}(Q'VQ)^2 = \mathrm{tr}(V(I-P_X))^2. \quad \square$$

Consider small values of ρ. Then $m_0(\rho)$ is approximately equal to k. So, we would like k to be large. On the other hand, consideration of $g_0(\rho)$ calls for small k as compared to a_1.

Further, since $k_0 + k = n - \underline{r}(X)$ it follows that small values of k imply large values of k_0, which are desirable. So, we need some sort of compromise. The decisive factor appears to be the degree of skewness of the beta distribution with parameters $k_0/2$ and $k/2$. The measure of skewness (see Pearson (1968)) is given by

$$(6.3.11) \qquad \sqrt{\beta_1} = 2(k-k_0)\sqrt{2k+2k_0+4}/(k+k_0+4)\sqrt{kk_0}\ .$$

If $k \ll k_0$ then the beta distribution would be so severely skewed to the left that the values of the integral $I_b(k_0/2, k/2)$ would not be high, at least for reasonable values of k_0 because then b cannot be close enough to 1. On the other hand, $k \gg k_0$ would cause skewness to the right. However, b would be drastically reduced. In fact, even the symmetric case $k = k_0$ would make $b < 1/2$ so that the power would also be less than $1/2$.

Now consider large values of ρ. We know that Wald's test is "optimal" against such alternatives for a given design. In choosing a design with high power at these alternatives we want k to be small and a_2 to be maximum in order to get large values of $g_0(\rho)$. As for $m_0(\rho)$ we need to maximize a_1^2/a_2. But $a_1^2/a_2 \leq k$. So, the two considerations, for $g_0(\rho)$ and $m_0(\rho)$, work against each other again.

It should be noticed here that $a_1^2/a_2 \simeq k$ when the λ_i's are nearly equal. Thus, the discussion in the previous paragraph suggests that Wald's test will be at its best when the model is "nearly balanced" and k is reasonable. That is, if we believe that ρ is small, i.e., $\theta_2 \ll \theta_1$, and we want to have good power near $\rho_0 = 0$, we should make k_0 large as compared to k. On the other hand, if we think that ρ is large, so that θ_2 is at least as important as θ_1, then we should arrange for enough degrees of freedom to estimate both parameters. So, k should not be small. However, the above analysis suggests that k should still be smaller than k_0.

6.4. The Unbalanced Random One-Way Classification Model

Consider the model

(6.4.1) $$Y_{ij} = \mu + \alpha_i + e_{ij}, \quad j = 1, \ldots, n_i, \quad i = 1, \ldots, a,$$

where μ is an unknown parameter and $\alpha_i \sim N(0, \theta_2)$ independently of $e_{ij} \sim N(0, \theta_1)$. Let us write the model in matrix form as

(6.4.2) $$Y = 1_n\mu + A\alpha + e,$$

where $n = \Sigma_{i=1}^{a} n_i$. By our assumptions above we have

(6.4.3) $$Y \sim N_n(1_n\mu, \theta_1 I + \theta_2 AA') .$$

We consider here the hypothesis

$$H: \theta_2 = 0 \text{ vs } K: \theta_2 > 0,$$

to illustrate the ideas of Section 6.3. With respect to this hypothesis we find

$$k = \underline{r}(1_n, AA') - \underline{r}(1_n) = a - 1,$$

and

$$k_0 = n - \underline{r}(1_n, AA') = n - a.$$

With these values (6.3.11) reduces to

$$\sqrt{\beta_1} = -2\sqrt{2}\,(n-2a+1)\sqrt{n+1}/(n+3)\sqrt{(a-1)(n-a)}\;.$$

Let

(6.4.4) $$N_j = \Sigma_{i=1}^{a} (n_i)^j\;, \quad j = 2, 3.$$

Using Remark 6.3.10 we find

$$a_1 = n - N_2/n\;,$$

and

$$a_2 = N_2 - 2N_3/n + N_2^2/n^2\;.$$

Let us recall from Section 6.3 that to get large values of $m_0(\rho)$ we want k to be large and a_1^2/a_2 to be maximum. To get large values of $g_0(\rho)$ we want k to be small and at small values of ρ we want a_1 to be maximum, while at large values of ρ we want a_2 to be maximum.

Because of the optimality of Wald's test against large values of ρ we now concentrate on the best ways to improve its performance at small values of ρ.

First let us consider maximizing a_1. This is equivalent to minimizing N_2. Given a groups and n units how do we allocate the n units to the a groups such that N_2 is a minimum? Anderson and Crump (1967) proved that the "as balanced as possible" allocation

(6.4.5) $M + 1$ units to each of ℓ groups and M units to each of $a - \ell$ groups where $n = aM + \ell$, $0 \leq \ell < a$,

minimizes N_2. Moreover, they also showed that the allocation (6.4.5) maximizes a_1^2/a_2 which is needed for $m_0(\rho)$ to be large. Fortunately, this allocation leads to at most three different, but adjacent, nonzero eigenvalues, i.e., $s \leq 3$. When $\ell > 1$ the eigenvalues and their multiplicities are given by

(6.4.6) $\lambda_1 = M$, $\lambda_2 = aM(M+1)/n$, $\lambda_3 = M + 1$, $k_1 = a - \ell - 1$, $k_2 = 1$, $k_3 = \ell - 1$.

If $\ell = 1$ we have only λ_1 and λ_2. If $\ell = 0$ we have only λ_1, e.g., see LaMotte (1976). It should be noticed here that, with these λ_i's, Satterthwaite's approximation is fairly accurate.

When the allocation (6.4.5) is used and $M > 2$ one can easily verify that $a_1^2/a_2 \simeq a - 1$. Using this in the expressions (6.3.8) and (6.3.9) we get

(6.4.7) $$g_0(\rho) \simeq 1 + a_1\rho/(a-1),$$

and

(6.4.8) $$m_0(\rho) \simeq a - 1.$$

Now $N_2 = (a-\ell)M^2 + \ell(M+1)^2 = nM + \ell(M+1)$. Thus we get

$$a_1 = n - M - \ell(M+1)/n = (a-1)M + \ell(1-(M+1)/n).$$

Hence, using this expression for a_1, we see that

$$M \leq a_1/(a-1) \leq M+1 .$$

So, we can rewrite (6.4.7) as

(6.4.9) $$g_0(\rho) \simeq 1 + M_0\rho, \quad M \leq M_0 \leq M+1.$$

Now let us consider the power function. Its form is

(6.4.10) $$\beta(\rho) \simeq I_{b(\rho)}\left(\frac{n-a}{2}, \frac{a-1}{2} \right),$$

where $b(\rho) = 1/(1+d(\rho))$ and $d(\rho)$ is given by

(6.4.11) $$d(\rho) = (a-1)c/(n-a)g_0(\rho) ,$$

$$\simeq (a-1)c/(n-a)(1+M_0\rho) .$$

Besides ρ, the power function is seen to depend on the sample size n and the number of groups a. If n is fixed we assume that one has control over a, and vice versa.

By examining $\beta(\rho)$ it is clear that high power will result from small values of d and positive values of $\sqrt{\beta_1}$. As we have seen before these two requirements cannot be achieved simultaneously. Indeed d would be smallest when $a = 2$ (the case where $a = 1$ is excluded of course). However taking $a = 2$ leads to severe negative skewness. On the other hand, taking $a > (n+1)/2$, while producing positive skewness, will result in values of $d > c/(1+\rho)$ which may not be as small as we wish. In fact $b(\rho)$ may be less than $1/2$ even at moderate values of ρ.

A numerical study has been carried out to learn about a good compromising value for a. Several sample sizes are considered. For each sample size a wide range of values for a is tried. For $n = 15$ and $n = 30$ the power curves corresponding to different values of a are given in Figures 1 and 2 below. These figures are typical of those obtained for other sample sizes. The plotted curves are seen to support the suggestion that a value for a around $n/4$, i.e., $M = 4$, may be a good compromise. That is, the best power

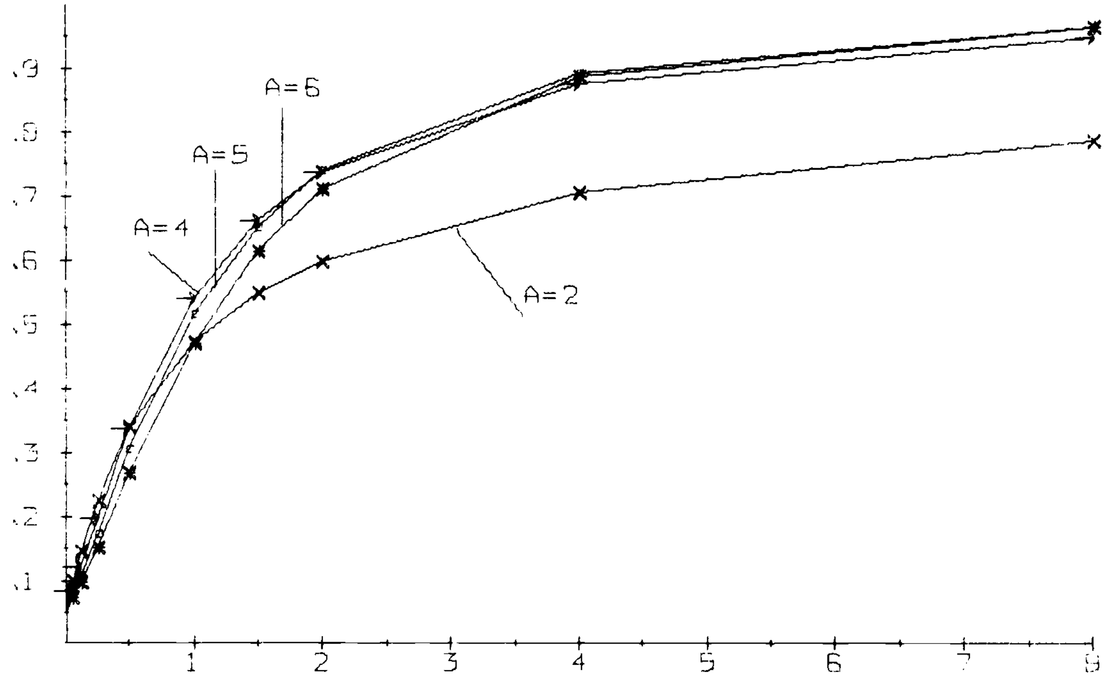

FIG.1. POWER CURVES FOR N=15

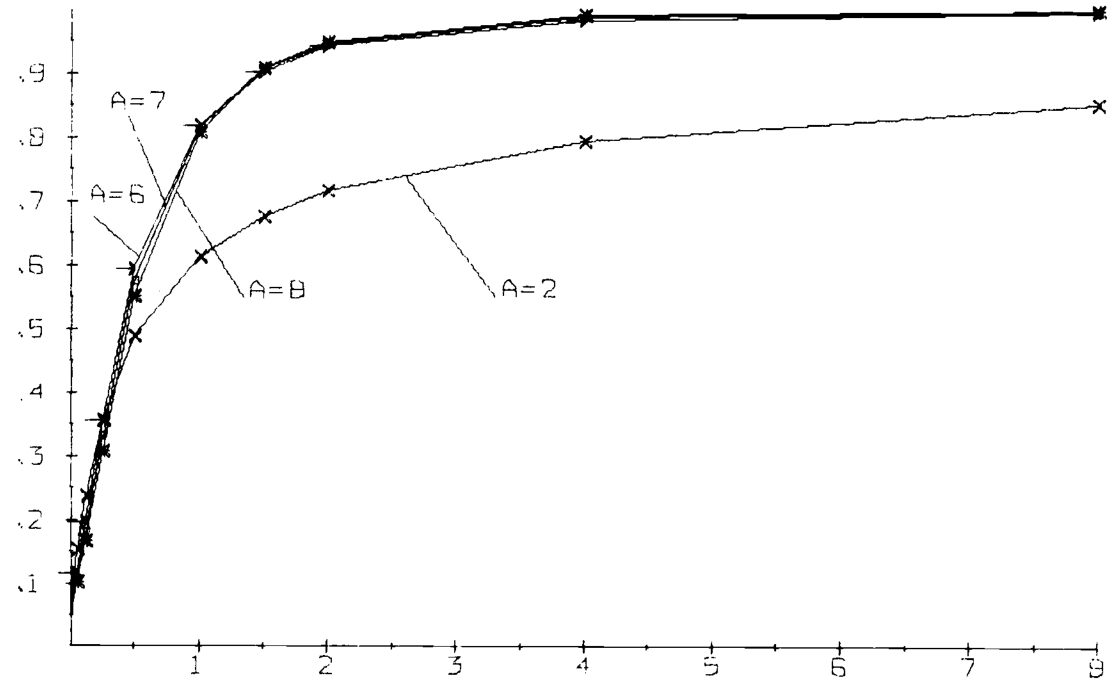

FIG.2. POWER CURVES FOR N=30

properties are obtained when the units are allocated to the groups as evenly as possible with around four units per group.

6.5. The Case of p Parameters

Consider the model

$$(6.5.1) \qquad Y \sim N(X\beta, V_0 + \Sigma_{i=1}^{p} \theta_i V_i) ,$$

where the parameters and the matrices are defined as above.

Suppose we are interested in testing about θ_p where the choice of θ_p is for notational convenience only. The main idea is to expand the concept of invariance to get rid of as many of the nuisance parameters as possible, thus, hopefully reducing the problem to the case of two parameters discussed above.

We proceed along the lines given by Goodnight (1976). Let $D = (X, V_0, \ldots, V_{p-1})$ and consider $Z = Q'Y$ where Q is such that $\underline{R}(Q) = \underline{N}(D')$. When $Q \neq 0$ we have that $Z \sim N(0, \theta_p Q'V_pQ)$ and we can test about θ_p. The problem arises though that there might be an ℓ, $1 \leq \ell \leq p-1$, such that $\underline{R}(V_p) \subset \underline{R}(V_\ell)$ so that $\underline{R}(Q) \subset \underline{N}(V_p)$. In order to avoid a zero covariance matrix $Q'V_pQ$, we have to drop V_ℓ out of D. Hence θ_ℓ must be kept as a nuisance parameter. Let $\mathcal{L} = \{1, \ldots, p-1\}$ and set $U_J = \Sigma_{j \in J} \underline{R}(V_j)$. Define

$$\mathcal{L}_1 = \{\ell \in \mathcal{L} : \exists\, J \subset \mathcal{L} \ \text{ s.t. } \ \ell \in J,\ \underline{R}(V_p) \subset U_J$$
$$\text{and } \ \underline{R}(V_p) \not\subset U_{J_1} \ \ \forall\, J_1 \subsetneq J\},$$

and

$$\mathcal{L}_2 = \{\ell \in \mathcal{L} : \ell \notin \mathcal{L}_1\}.$$

Drop the matrices V_ℓ, $\ell \in \mathcal{L}_1$, from D and denote the resulting matrix by D_1. Let Q_1 be such that

(6.5.2) $$\underline{R}(Q_1) = \underline{N}(D_1') .$$

We follow Goodnight and say that $Z_1 = Q_1'Y$ is maximally invariant wrt θ_ℓ, $\ell \in \mathcal{L}_2$. When $Q_1 \neq 0$, we have that

(6.5.3) $$Z_1 \sim N(0, \Sigma_{\ell \in \mathcal{L}_1} \theta_\ell W_\ell) ,$$

where

(6.5.4) $$W_\ell = Q_1'V_\ell Q_1, \quad \ell \in \mathcal{L}_1 .$$

It might be that the family $\mathcal{P}_{Z_1}$ admits a complete sufficient statistic in which case the results in Chapter IV are applicable. Otherwise we hope that $\mathcal{L}_1$ has only two elements so that we can use the results in this chapter. If this is not the case we may have to assume that certain effects play no role in order to be able to use Wald's test. We illustrate by way of an example.

Example 6.5.5. Consider the model

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} ,$$

where $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, n_{ij}$. We assume that μ is an unknown parameter and that α_i, β_j, γ_{ij} and e_{ijk} are independent normal random variables with means zero and variances σ_a^2, σ_b^2, σ_{ab}^2 and σ^2 respectively.

The above model can be rewritten in matrix form as follows

$$Y = 1_n\mu + A\alpha + B\beta + C\gamma + e,$$

where $n = \Sigma_{i=1}^a \Sigma_{j=1}^b n_{ij}$. For this model it is well known that

(6.5.6) $$\underline{R}(A) \subset \underline{R}(C) \subset \underline{R}(I) = R^n ,$$
$$\underline{R}(B) \subset \underline{R}(C) \subset \underline{R}(I) = R^n .$$

Furthermore

$$Y \sim N_n(1_n\mu, \sigma_a^2 AA' + \sigma_b^2 BB' + \sigma_{ab}^2 CC' + \sigma^2 I).$$

Let us consider testing about σ_{ab}^2. From (6.5.6) we see that $\mathcal{L}_1 = \{3, 4\}$. Thus $D_1 = (1_n, AA', BB')$. Let Q_1 be such that $\underline{R}(Q_1) = \underline{N}(D_1')$ and $Q_1'Q_1 = I$. Then

$$Z_1 \sim N(0, \sigma_{ab}^2 Q_1'CC'Q_1 + \sigma^2 I).$$

Therefore, the results in Sections 6.2 and 6.3 can immediately be applied for testing about σ_{ab}^2/σ^2. In particular we can test the important hypothesis of no interaction, i.e., $\sigma_{ab}^2 = 0$. The invariance argument in the above sense provides a justification for using any of the tests in Section 6.2.

Now consider testing about σ_a^2. (The treatment of σ_b^2 is quite similar.) From (6.5.6) we get $\mathcal{L}_1 = \{1, 3, 4\}$. Thus $D_1 = (1_n, BB')$. Let Q_1 be such that $\underline{R}(Q_1) = \underline{N}(D_1')$ and $Q_1'Q_1 = I$. Then

$$Z_1 \sim N(0, \sigma_a^2 Q_1'AA'Q_1 + \sigma_{ab}^2 Q_1'CC'Q_1 + \sigma^2 I).$$

If we can assume that $\sigma_{ab}^2 = 0$ then we can use the tests in Section 6.2.

As mentioned in Section 6.1, this example has been discussed by Spjotvoll (1968) and Thomsen (1975). They approached the problem differently though. Neither of them claims any optimal properties for the proposed tests. The reader is referred to these papers for more details.

BIBLIOGRAPHY

Anderson, R. L. and Crump, P. P. (1967). Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process. Technometrics, 9, 499-516.

Anderson, T. W. (1971). The Statistical Analysis of Time Series. Wiley, New York.

Arvesen, J. N. and Layard, M. W. J. (1975). Asymptotically robust tests in unbalanced variance component models. Ann. Statist. 3, 1122-1134.

Bartle, R. G. (1966). The Elements of Integration. Wiley, New York.

Cochran, W. G. (1951). Testing a linear relation among variances. Biometrics, 7, 17-32.

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

Davenport, J. M. and Webster, J. T. (1972). Type I error and power of a test involving a Satterthwaite's approximate F-statistic. Technometrics, 14, 555-570.

Davies, R. B. (1969). Beta-optimal tests and an application to the summary evaluation of experiments. JRSSB, 31, 524-538.

Dugundji, J. (1966). Topology. Allyn and Bacon, Boston.

Fraser, D. A. S. (1957). Nonparametric Methods in Statistics. Wiley, New York.

Gautschi, W. (1959). Some remarks on Herbach's paper. Ann. Math. Statist., 30, 960-963.

Goodnight, J. H. (1976). Maximally-invariant quadratic unbiased estimators. Biometrics, 32, 477-480.

Graybill, F. A. (1976). Theory and Application of the Linear Model. Duxbury Press, Massachusetts.

Green, J. J. (1975). Exact testing and estimating procedures for the one-way random effects ANOVA model in the unbalanced case. Presented at the national meeting of the ASA at Atlanta, Georgia.

Hartley, H. O. and Rao, J. N. K. (1967). Maximum-likelihood estimation for the mixed analysis of variance model. Biometrika, 54, 93-108.

Herbach, L. H. (1959). Properties of model II-type analysis of variance tests, A: optimum nature of the F-test for model II in the balanced case. Ann. Math. Statist., 30, 939-959.

Hocking, R. R. (1973). A discussion of the two-way mixed model. American Statistician, 27, 148-152.

Howe, R. B. and Myers, R. H. (1970). An alternative to Satterthwaite's test involving positive linear combinations of variance components. JASA, 65, 404-412.

Hultquist, R. A. and Thomas, J. (1975). Testing hypotheses relative to variance components in unbalanced mixed effects models. Presented at the Rochester 1975 spring meetings, New York.

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

Jensen, S. T. (1975). Covariance hypotheses which are linear in both the covariance and the inverse covariance. Institute of Math. Statist., Univ. of Copenhagen preprint # 1, Denmark.

LaMotte, L. R. (1976). Invariant quadratic estimators in the random one-way ANOVA model. Biometrics, 32, 793-804.

Lehmann, E. L. and Scheffé, H. (1955). Completeness, similar regions, and unbiased estimation - part II. Sankhya, 15, 219-236.

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

Naik, U. D. (1974a). On tests of main effects and interactions in higher-way layouts in the analysis of variance random effects model. Technometrics, 16, 17-25.

Naik, U. D. (1974b). On some procedures for testing a linear relation among variances. JRSS B, 36, 243-257.

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

Pearson, K. (1968). Tables of the Incomplete Beta-Function. Cambridge Univ. Press, Cambridge.

Portnoy, S. (1973). On recovery of intra-block information. JASA, 68, 384-391.

Satterthwaite, F.E. (1946). An approximate distribution of estimates of variance components. Biometrics, 2, 110-114.

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

Seely, J. (1971). Quadratic subspaces and completeness. Ann. Math. Statist., 42, 710-721.

Seely, J. (1972). Completeness for a family of multivariate normal distributions. Ann. Math. Statist., 43, 1644-1647.

Seely, J. (1976). Minimal sufficient statistics and completeness for multivariate normal families. To appear in Sankhyā.

Spjotvoll, E. (1967). Optimum invariant tests in unbalanced variance components models. Ann. Math. Statist., 38, 422-428.

Spjotvoll, E. (1968). Confidence intervals and tests for variance ratios in unbalanced variance components models. Rev. Internat. Statist. Inst., 36, 37-42.

Srivastava, J.N. (1966). On testing hypotheses regarding a class of covariance structures. Psychometrika, 31, 147-164.

Thompson, W.A., Jr. (1955a). The ratio of variances in a variance components model. Ann. Math. Statist., 26, 325-329.

Thompson W.A., Jr. (1955b). On the ratio of variances in the mixed incomplete block model. Ann. Math. Statist., 26, 721-733.

Thomsen, I. (1975). Testing hypotheses in unbalanced variance components models for two-way layouts. Ann. Statist., 3, 257-265.

Tietjen, G. L. (1974). Exact and approximate tests for unbalanced random effects designs. Biometrics, 30, 573-581.

Wald, A. (1947). A note on regression analysis. Ann. Math. Statist., 18, 586-589.

Walker, R. J. (1950). Algebraic Curves. Princeton University Press, Princeton.

Zyskind, G. (1967). On canonical forms, nonnegative covariance matrices and best and simple least squares linear estimators in linear models. Ann. Math. Statist., 38, 1092-1109.