The Annals of Statistics
1983, Vol. 11, No. 2, 654–665

MULTIVARIATE TESTS WITH INCOMPLETE DATA

BY MORRIS EATON[1] AND TAKEAKI KARIYA[2]

University of Minnesota and Hitotsubashi University

In the context of a normal model, testing problems with missing data are considered. Tests on means are treated when independent extra data on the first p_1 variates of p variates is available in addition to complete data. For testing that the mean of the first p_1 variates is zero, the LRT is UMP invariant, but for testing that the whole mean is zero, no LMPI (locally most powerful invariant) test exists. Second, tests for independence are treated in similar situations, and an LMPI test for each situation is derived. In some situations it is found that the LRT for independence ignores the extra data.

1. Introduction. Because of their common occurrence in practice, there has been a continuing interest in inference problems where there is missing or extra data. The causes for the data to be missing or extra will not be discussed explicitly in this paper, but will be implicit in our assumptions concerning the likelihood function of the data (see Section 2). For an illuminating discussion of such issues, the reader is referred to Rubin (1963). With the likelihood assumed in (2.3) and (2.4), it is equivalent to think of certain parts of the data as additional or the "complementary" parts of these data are missing.

The problems treated in this paper concern data on p coordinates which are partitioned into two groups of p_1 and p_2 coordinates—so $p_1 + p_2 = p$ and $1 \le p_i < p$, 1, 2. It is assumed that we have n p-dimensional observation vectors and m_i p_i-dimensional observation vectors, $i = 1, 2$. All $n + m_1 + m_2$ vectors are assumed to be independent. Thus, there are n "complete" observations, m_1 "extra" observations on the first p_1 coordinates, and m_2 "extra" observations on the last p_2 coordinates. When m_1 (or equivalently m_2) is zero, then the data is in triangularly partitioned form. Under the assumption of multivariate normality, Bhargava (1962) derived maximum likelihood estimators (MLE's) and likelihood ratio tests (LRT's) for a number of problems when the data has a general triangular form. This triangular form permits the explicit calculation of MLE's and LRT's along with the relevant distribution theory. Morrison and Bhoj (1973) discuss the power of the LRT for testing a mean vector is zero when $m_1 = 0$.

Ordinarily, likelihood methods are proposed for problems with missing data—especially when the normal distribution is involved. However, in some situations, the likelihood equations cannot be solved explicitly. The article by Hartley and Hocking (1971) provides a good overview of the subject and an extensive bibliography. The recent work of Little (1976) is concerned solely with the normal distribution but general patterns of missing data are allowed. Little compares a variety of estimators both asymptotically and numerically.

To illustrate the possible difficulties involved in missing data problems, consider the following case: $p_1 = p_2 = 1$, the "complete" data is a sample of n from a bivariate normal distribution with unknown mean vector $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ and unknown covariance matrix, and the two extra samples are from univariate normal populations with means and variance of the marginal distributions of the bivariate normal. Suppose the problem is to test that $\mu_1 = \mu_2$. If $n = 0$, this is the Behrens-Fisher problem. When $n > 0$, the problem should be no

Received June 1980; revised October 1982.

[1] Supported in part by National Science Foundation Grant MCS 81-00762.

[2] Supported in part by the Ministry of Education of Japan under General C57530008.

AMS 1970 *subject classifications.* Primary 62H15; secondary 62A05.

Key words and phrases. Missing data, locally most powerful invariant tests, testing independence, invariance.

easier than the Behrens-Fisher problem and the work thus far justifies this belief. A comparison of several different proposals to solve this problem is given in Ekbohm (1976).

The purpose of the present paper is to discuss the existence or nonexistence of tests with certain optimum properties. In Section 1, we set notation and derive a canonical form for the data under consideration. It is assumed that n p-dimensional normal (μ, Σ) vectors are available and m_i p_i-dimensional normal (μ_i, Σ_{ii}), $i = 1, 2$, are available where $\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$, $\Sigma = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$ with $\mu_i : p_i \times 1$, $\Sigma_{ij} : p_i \times p_j$, $i, j = 1, 2$ and $p_1 + p_2 = p$ (see (2.1) below). All parameters are assumed unknown.

In Section 3, it is assumed that $m_2 = 0$ so no "extra" data is available on the last p_2-coordinates. For the problem of testing $\mu_1 = 0$ versus $\mu_1 \neq 0$, the LRT is shown to be uniformly most powerful invariant. However, for testing $\mu = 0$ versus $\mu \neq 0$, a locally most powerful invariant test does not exist. A few comments concerning the LRT of $\mu = 0$ are given.

In Section 4, we consider the problem of testing $\Sigma_{12} = 0$ when both m_1 and m_2 are non-negative. In this case we derive a locally most powerful invariant test. When $m_1 = 0$ and $m_2 > 0$ (or $m_1 > 0$ and $m_2 = 0$), this test is different from the LRT. (When $m_1 > 0$ and $m_2 > 0$, the LRT is not known explicitly). In fact, the LRT does not utilize the "extra" data at all and is identical to the LRT when $m_1 = m_2 = 0$. This point is discussed and we propose a possible test statistic for testing $\Sigma_{21} = 0$ which utilizes the additional information. Two examples are presented in Section 5.

The missing data patterns considered in this paper are among the most simple, but our results indicate the variety of possible answers one can obtain when: (i) comparing LRT'S to optimum (in some sense) tests when they both exist and (ii) trying to settle questions concerning the existence of optimum tests. Invariance considerations play a central role in this paper, and we often employ the method of averaging over groups to obtain the density function of a maximal invariant. Of course, we have only been able to employ this technique when the group under consideration acts transitively on the null hypothesis. The representation result we have used is due to Wijsman (1967) though an alternative representation result obtained by Andersson (1982) is available. Some of the details are only sketched as they are similar to those in Schwartz (1967) or Kariya (1978). The proof of Theorem 3.2 is omitted as it is similar to that outlined for Theorem 4.1.

2. Notation and a canonical form. The extra (or missing) data problems to be considered here are among the simplest but illustrate the mathematical problems encountered when dealing with such models. Consider a multivariate normal population of dimension p with a mean vector μ (a column vector) and a $p \times p$ nonsingular covariance matrix Σ. Write $p = p_1 + p_2$ where $1 \leq p_i < p$ for $i = 1, 2$ and partition μ and Σ as

$$\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \quad \Sigma = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix} \tag{2.1}$$

with μ_i being $p_i \times 1$ and Σ_{ij} being $p_i \times p_j$ for $i, j = 1, 2$. It is assumed that we have "complete" observations $z_1, \cdots, z_N$ which are i.i.d. $N_p(\mu, \Sigma)$ and "marginal" observations $z_{i1}, \cdots, z_{iM_i}$, $i = 1, 2$ which are i.i.d. $N_{p_i}(\mu_i, \Sigma_{ii})$ for $i = 1, 2$. In terms of data matrices, the complete sample yields $\tilde{Z}: N \times p$ with rows z_i', $i = 1, \cdots, N$, while the marginal samples yield $\tilde{Z}_i : M_i \times p_i$ with rows z_{ij}', $i = 1, 2, j = 1, \cdots, M_i$. Then

$$\mathscr{L}(\tilde{Z}) = N(e_N\mu', I_N \otimes \Sigma) \quad \text{and} \quad \mathscr{L}(\tilde{Z}_i) = N(e_{M_i}\mu_i', I_{M_i} \otimes \Sigma_{ii}), \quad i = 1, 2, \tag{2.2}$$

where e_k is the vector of ones in R^k. Here, the notation "$\mathscr{L}(\cdot)$" means the distribution of "$\cdot$" and $\otimes$ denotes the Kronecker product.

It is convenient to transform the data $\tilde{Z}$, $\tilde{Z}_1$ and $\tilde{Z}_2$ into what will be called the canonical form. Let Γ be an $N \times N$ orthogonal matrix with first row $e_N'/\sqrt{N}$. Then the transpose of the first row of the matrix $\Gamma\tilde{Z}$ has a $N(\sqrt{N}\,\mu, \Sigma)$ distribution and is independent of the remaining $(N - 1)$ rows which are i.i.d. $N(0, \Sigma)$. Let $Y \in R^p$ be the transpose of the first row of $\Gamma\tilde{Z}$ multiplied by $1/\sqrt{N}$ and let $V: (N-1) \times p$ be the remaining $(N-1)$ rows of $\Gamma\tilde{Z}$. Then Y and V are independent with

$$\mathscr{L}(Y) = N(\mu, c\Sigma) \quad \text{and} \quad \mathscr{L}(V) = N(0, I_n \otimes \Sigma), \tag{2.3}$$

where $c = 1/N$ and $n = N - 1$. Transforming $\tilde{Z}_i$ in a similar manner leads to $X_i \in R^{p_i}$ and $V_i: (M_i - 1) \times p_i$ which are independent and satisfy

$$\mathscr{L}(X_i) = N(\mu_i, c_i\Sigma_{ii}) \quad \text{and} \quad \mathscr{L}(V_i) = N(0, I_{m_i} \otimes \Sigma_{ii}), \quad i = 1, 2, \tag{2.4}$$

where $c_i = 1/M_i$ and $m_i = M_i - 1$. In summary, the complete and partial data $\tilde{Z}$, $\tilde{Z}_1$ and $\tilde{Z}_2$ can be relabeled to yield mutually independent Y, X_1, X_2 and V, V_1, V_2 with the given distributions. Observations which are represented in the forms (2.3) and (2.4) will be said to be in canonical form, where c, c_1 and c_2 are known positive constants. For some of the problems treated below, the data is assumed to be in canonical form. However, to motivate the examples in Section 5, it is necessary to describe the original problem before transforming it to the forms (2.3) and (2.4).

In some cases, parts of the data in (2.4) will be missing. For example, if there is no marginal sample on the last p_2-coordinate, then both X_2 and V_2 are missing in (2.4) and $m_2 = 0$. In fact, this will be the case considered in the next section where we take up the problem of testing hypotheses about μ. The full generality of (2.4) is used in Section 4 for testing that $\Sigma_{12} = 0$.

Throughout this paper, $n \geq \max(p_1, p_2)$ is assumed and, except in Section 5, the notation

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \begin{matrix} p_1 \\ p_2 \end{matrix}, \quad V'V = S = \overset{\begin{matrix} p_1 & p_2 \end{matrix}}{\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}} \begin{matrix} p_1 \\ p_2 \end{matrix}, \quad V_i'V_i = W_{ii}, \quad i = 1, 2, \tag{2.5}$$

and $k^{-1} = c^{-1} + c_1^{-1}$, is used. Further $G\ell_p$ denotes the group of $p \times p$ nonsingular real matrices.

3. Tests on means. Throughout this section, we consider data Y, X_1, V and V_1 in the canonical form (2.3) and (2.4) where X_2 and V_2 are not present. As remarked earlier, this means that no "extra" data was available on the last p_2 coordinates of our basic sample. Of course, it is assumed that Y, X_1, V and V_1 are mutually independent and

$$\begin{aligned} &\mathscr{L}(Y) = N(\mu, c\Sigma), \quad \mathscr{L}(V) = N(0, I_n \otimes \Sigma), \\ &\mathscr{L}(X_1) = N(\mu_1, c_1\Sigma_{11}), \quad \mathscr{L}(V_1) = N(0, I_{m_1} \otimes \Sigma). \end{aligned} \tag{3.1}$$

Based on the data (3.1), we now want to discuss the problem of testing $H_1 : \mu_1 = 0$ versus $H_1 : \mu_1 \neq 0$. This testing problem is invariant under a group of transformations acting on the sample space. In particular, consider the group G whose elements are $g = (A, a)$ with $A \in G\ell_p$ and $a \in R^p$ where

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}, \quad A_{ii} \in G\ell_{p_i}, \quad i = 1, 2 \quad \text{and} \quad a = \begin{pmatrix} 0 \\ a_2 \end{pmatrix}, \quad a_2 \in R^{p_2}. \tag{3.2}$$

The action of $g = (A, a)$ on (Y, X_1, V, V_1) and (μ, Σ) is

$$g(Y, X_1, V, V_1) = (AY + a, A_{11}X_1, VA', V_1A_{11}') \quad \text{and} \quad g(\mu, \Sigma) = (A\mu + a, A\Sigma A'), \tag{3.3}$$

respectively. The composition of two group elements is $(A, a)(B, b) = (AB, Ab + a)$. It is now a routine matter to check that the testing problem is invariant under the group G. A maximal invariant in the parameter space is $\delta_1 = \mu_1'\Sigma_{11}^{-1}\mu_1$. In terms of δ_1, the null hypothesis is $H_0 : \delta_1 = 0$ and the alternative is $H_1 : \delta_1 > 0$. The next result will allow us to derive a uniformly most powerful invariant (UMPI) test of H_0 versus H_1.

THEOREM 3.1. *Let $\delta_1 = \mu_1'\Sigma_{11}^{-1}\mu_1$. The test of $H_0 : \delta_1 = 0$ versus $H_1 : \delta_1 > 0$ which rejects* *for large values of*

$$t = k\left(\frac{Y_1}{c} + \frac{X_1}{c_1}\right)'\left(S_{11} + W_{11} + \frac{Y_1Y_1'}{c} + \frac{X_1X_1'}{c_1}\right)^{-1}\left(\frac{Y_1}{c} + \frac{X_1}{c_1}\right) \tag{3.4}$$

is UMPI. *Here,* S_{11} *and* W_{11} *are given in* (2.5) *and* $k^{-1} = c^{-1} + c_1^{-1}$.

PROOF. We only sketch the proof. Starting with the sufficient statistic (Y, X_1, S, W_{11}), act with translations $g = (I, a)$ and $a = (0, a_2')'$ and then let G act with transformations of the form $g = (A, 0)$ where A is given by (3.2) with $A_{11} = I$. It is easy to see that this produces a maximal invariant $Y_1, X_1, S_{11}, W_{11})$. An equivalent maximal invariant is (u, v, S_{11}, W_{11}), where $u = c^{-1}Y_1 + c_1^{-1}X_1$ and $v = (c + c_1)^{-1/2}(Y_1 - X_1)$. Note that u and v are independent, and v is $N(0, \Sigma_{11})$. Now a sufficient statistic based on (u, v, S_{11}, W_{11}) is $(u, S_{11} + W_{11} + vv') = (u, R)$, say, and the action of G is $u \to A_{11}u$ and $R \to A_{11}RA_{11}'$. Under this action, a maximal invariant here is a constant times Hotelling's T^2 or $r \equiv u'R^{-1}u$. Hence it suffices to point out that r and t in (3.4) are in 1-1 correspondence. Since the distribution of r has a monotone likelihood ratio, the test based on r or t is UMPI. Noting that the sufficiency and invariance reductions used here commute, Theorem 3.1 is proved.

It is not difficult to show that the likelihood ratio test (LRT) of $H_0: \mu_1 = 0$ versus $H_1: \mu_1 \neq 0$ is equivalent to the test which rejects for large values of t. Standard arguments show that the statistic t has a central Beta distribution under H_0 and a non-central Beta distribution under H_1. See Bhargava (1962) for details.

We now turn to the problem of testing $H_0: \mu = 0$ versus the alternative $H_1: \mu \neq 0$. In this case, the situation is substantially different from the first case considered. Again, we take the data of the problem to be given by (3.1). This testing problem is invariant under the group G_0, a sub-group of G, defined by

$$G_0 = \{g = (A, a) \mid g \in G, \quad a = 0\}.$$

The action on the sample space and parameter space are as before with $a \equiv 0$. A direct calculation shows that a maximal invariant parameter is $\delta = \begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix}$ where

$$\delta_1 = \mu_1' \Sigma_{11}^{-1}\mu_1, \quad \delta_2 = (\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1)' \Sigma_{22\cdot 1}^{-1}(\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1).$$

In terms of δ, the problem is to test that $\delta = 0$ versus $\delta \neq 0$. For this problem, it will be shown that there is no UMPI (under G_0) test of H_0 versus H_1.

THEOREM 3.2. *Let* P_δ *denote the probability measure of a maximal invariant at the parameter value* δ. *Then the Radon-Nikodym derivative* dP_δ/dP_0 *is given by* $R(t_1, t_2, t_3 \mid \delta) = H(\delta)F_1(t_1\delta_2)F_2(t_2\delta_2/c)F_3(t_3\delta_1/k)$, *where*

$$H(\delta) = \exp[-(\delta_2/2c) - (\delta_1/2k)] \quad \textit{with} \quad k^{-1} = c^{-1} + c^{-1},$$

$$t_1 = Y_1'T_{11}^{-1}Y_1/c,$$

$$t_2 = (Y_2 - T_{21}T_{11}^{-1}Y_1)'T_{22\cdot 1}^{-1}(Y_2 - T_{21}T_{11}^{-1}Y_1)/c,$$

$$t_3 = k\left(\frac{Y_1}{c} + \frac{X_1}{c_1}\right)'\left(W_{11} + T_{11} + \frac{X_1X_1'}{c_1}\right)^{-1}\left(\frac{Y_1}{c} + \frac{X_1}{c_1}\right).$$

Here $v'v + 1/c\ YY' \equiv T$ *is partitioned as* $T = (T_{ij})$ *with* $T_{ij}: p_i \times p_j$, $F_1(x) = \exp(x/2c)$, $F_2(x) = F(x: p_2, (n - p_1 + 1)/2)$, $F_3(x) = F(x: p_1, (n + m_1 + 2)/2)$ *where*

$$F(x: \alpha, \beta) = \Sigma_{j=0}^{\infty} \frac{2^j}{(2j)!} \frac{\Gamma(\alpha/2)}{\Gamma\left(\frac{1}{2}\right)} \frac{\Gamma\left(j + \frac{1}{2}\right)}{\Gamma(j + \alpha/2)} \frac{\Gamma(\beta + j)}{\Gamma(\beta)} x^j.$$

PROOF. A proof of this is similar to the proof of Theorem 4.1 which is outlined in Appendix II.

Expanding $R(t_1, t_2, t_3 \mid \delta)$ about $\delta_1 = \delta_2 = 0$, the linear approximation to $R(t_1, t_2, t_3 \mid \delta)$ is

$$R(t_1, t_2, t_3 \mid \delta) = 1 + u_1\delta_1 + u_2\delta_2 + o(t_1, t_2, t_3, \delta),$$

where
$$u_1 = \frac{1}{k}\left(h_1 t_3 - \frac{1}{2}\right) \quad \text{and} \quad u_2 = \frac{1}{c}\left(d_1 t_2 + t_1 - \frac{1}{2}\right)$$

with
$$h_1 = \frac{n + m_1 + 2}{2p_1} \quad \text{and} \quad d_1 = \frac{n - p_1 + 1}{2p_2}.$$

The remainder term is uniform in t_1, t_2 and t_3 since $0 \le t_i \le 1$ for $i = 1, 2, 3$. This implies that if ϕ is an invariant level α test of H_0 versus H_1, then the power function of ϕ, for small δ, is

$$\mathscr{E}_\delta\phi = \alpha + \mathscr{E}_0\phi\{u_1\delta_1 + u_2\delta_2\} + o(\delta),$$

where the error term $o(\delta)$ is uniform in δ. Now consider testing $H_0: \delta_1 = \delta_2 = 0$ versus $\tilde{H}_1: \delta_1 = \gamma\delta_2 > 0$ where γ is a known positive constant. An easy application of the Generalized Neyman-Pearson Lemma shows that the level α test which rejects for large values of $\gamma u_1 + u_2$ is a LMPI test of H_0 versus $\tilde{H}_1$. Since this test depends on γ, there can be no LMPI test of H_0 versus H_1.

We now turn to a brief discussion of the likelihood ratio test of $H_0: \delta = 0$ versus $H_1: \delta \neq 0$. A direct calculation shows that the LRT of $H_0^{(1)}: \mu_1 = 0$ rejects $H_0^{(1)}$ if $\lambda_1^{2/(m_1+n+2)} \equiv 1 - t$ is too small where t is defined by (3.4). Furthermore, the LRT of $H_0^{(2)}: \mu_2 = 0$, $\mu_1 = 0$ versus $H_1^{(2)}: \mu_2 \neq 0$, $\mu_1 = 0$ rejects for small values of

$$\lambda_2^{2/(n+1)} \equiv \frac{|S_{22} - S_{21}S_{11}^{-1}S_{12}|}{\left|S_{22} + \frac{Y_2Y_2'}{c} - \left(S_{21} + \frac{Y_2Y_1'}{c}\right)\left(S_{11} + \frac{Y_1Y_1'}{c}\right)^{-1}\left(S_{21} + \frac{Y_2Y_1'}{c}\right)'\right|},$$

where the S_{ij}'s are given in (2.5). Now, the LRT of $H_0: \delta = 0$ versus $H_1: \delta \neq 0$ rejects for small values of $\lambda_1\lambda_2$. In addition, under H_0, λ_1 and λ_2 are independent, $\lambda_1^{2/(m_1+n+2)}$ has a Beta distribution and $\lambda_2^{2/(n+1)}$ has a Beta distribution. (See Morrison and Bhoj, 1973). But, this does not yield the exact null distribution of $\lambda_1\lambda_2$ under H_0 expressed in terms of a tabled distribution. It should be mentioned that this type of decomposition of likelihood ratio statistics for testing normal means occurs in other contexts. For example, see Hogg (1961) for univariate normal example, Eaton (1972) for the MANOVA case, and Karíya (1974) for the application of these ideas to the multivariate linear growth curve model.

4. Testing for independence. In this section, we consider the problem of testing for independence based on data in canonical form. The canonical form will be of the type described by (2.1), but for simplicity our main discussion will be concerned with the following data: Consider three independent random matrices $V: n \times p$, $V_1: m_1 \times p_1$ and $V_2: m_2 \times p_2$ with $p_1 + p_2 = p$ satisfying

$$\mathscr{L}(V) = N(0, I_n \otimes \Sigma) \quad \text{and} \quad \mathscr{L}(V_i) = N(0, I_{m_i} \otimes \Sigma_{ii}), \quad i = 1, 2. \tag{4.1}$$

Here, the unknown covariance matrix Σ has been partitioned into $\Sigma_{ij}: p_i \times p_j$ for $i, j = 1, 2$. The data (4.1) arises from the data described in Section 2 by assuming that the mean vector μ is known. The problem is to test $H_0: \Sigma_{12} = 0$ versus the alternative $H_1: \Sigma_{12} \neq 0$. After describing our results for this problem, we will state some corresponding results for this testing problem when the data is given by (2.3) and (2.4) (and some minor variations of (2.3) and (2.4)).

The testing problem is invariant under the group G_2 whose elements g are

$$g = A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{12} \end{pmatrix} \text{ with } A_{ii} \in G\ell_{p_i}, \quad i = 1, 2. \tag{4.2}$$

The action of g on (v, v_1, v_2) and Σ is respectively

$$g(V, V_1, V_2) = (VA', V_1A'_{11}, V_2A'_{22}) \quad \text{and} \quad g(\Sigma) = A\Sigma A'. \tag{4.3}$$

A maximal invariant parameter is the vector $\delta = (\delta_1^2, \cdots, \delta_q^2)'$ with $q = \min(p_1, p_2)$, where $\delta_1^2 \geq \cdots \geq \delta_q^2$ are the q-largest eigenvalues of $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}$ ($\delta_1, \cdots, \delta_q$ are the canonical correlations). The main result of this section is that there exists a LMPI test of the H_0 versus H_1. To describe this result, let D_α^I be all the level α G_2-invariant test functions.

THEOREM 4.1. *Let $\tau = \Sigma_1^q \delta_i^2$. For $\phi \in D_\alpha^I$, the power function of ϕ at δ, say $\pi(\phi, \delta)$, has the form*

$$\pi(\phi, \delta) = \alpha + B(\phi)\,\tau + o(\delta, \varphi), \tag{4.4}$$

where $\lim_{\delta\to 0}\sup_\phi o(\delta, \phi) = 0$, $B(\phi) = \mathscr{E}_0(\tfrac{1}{2}\phi\psi_0)$ *and in the notation of* (1.5),

$$\psi_0 = \frac{(n + m_1)(n + m_2)}{p_1p_2}\operatorname{tr}(S_{11} + W_{11})^{-1}S_{12}(S_{22} + W_{22})^{-1}S_{21} + n - \Sigma_{i=1}^2 \left(\frac{n + m_i}{p_i}\right)\operatorname{tr}(S_{ii} + W_{ii})^{-1}S_{ii}. \tag{4.5}$$

The level α test which rejects for $\psi_0 > k$ is a LMPI *level α test.*

PROOF. The representation (4.4) is established in the Appendix. That rejecting for $\psi_0 > k$ gives a LMPI test follows immediately from (4.4) by maximizing $B(\phi)$ and applying the generalized Neyman-Pearson Lemma.

In the discussion below, the situation treated by Theorem 4.1 for the data (4.1) will be called Case (0). We now turn to a brief discussion of some other cases of interest.

CASE (i). This refers to Case (0) when $m_1 = 0$ so that the data matrix V_1 is not available. A direct analogue of Theorem 4.1 shows that the test which rejects for large values of

$$\psi_1 \equiv \frac{(n + m_2)n}{p_2p_1}\operatorname{tr} S_{11}^{-1}S_{12}(S_{22} + W_{22})^{-1}S_{21} - \frac{n + m_2}{p_2}\operatorname{tr}(S_{22} + W_{22})^{-1}S_{22} \tag{4.6}$$

is a LMPI test for testing $H_0 : \Sigma_{12} = 0$ versus $H_0 : \Sigma_{12} \neq 0$.

CASE (ii). In this case we consider the data given in (2.1). Let

$$U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \equiv \begin{pmatrix} (c + c_1)^{-1/2}(Y_1 - X_1) \\ (c + c_2)^{-1/2}(Y_2 - X_2) \end{pmatrix}$$

and set $b^2 = c^2/(c + c_1)(c + c_2)$. Define the statistic ψ_2 by

$$\begin{aligned} \psi_2 = {} & \frac{(n + m_1 + 1)(n + m_2 + 1)}{p_1p_2} \\ & \cdot \operatorname{tr}(S_{11} + W_{11} + U_1U_1')^{-1}S_{12}(S_{22} + W_{22} + U_2U_2')^{-1}S_{21} \\ & + b^2\operatorname{tr}(S_{11} + W_{11} + U_1U_1')^{-1}U_1U_2'(S_{22} + W_{22} + U_1U_2')^{-1}U_2U_1 \\ & + n + b^2 - \frac{n + m_1 + 1}{p_1}\operatorname{tr}(S_{11} + W_{11} + U_1U_1')^{-1}(S_{11} + b^2U_1U_1') \\ & - \frac{n + m_2 + 1}{2}\operatorname{tr}(S_{22} + W_{22} + U_2U_2')^{-1}(S_{22} + b^2U_2U_2'). \end{aligned} \tag{4.7}$$

Rejecting $H_0 : \Sigma_{12} = 0$ for large values of ψ_2 is a LMPI test.

CASE (iii). Again consider the data as given in (2.1) but assume that the mean of X_i is unrelated to the mean of Y, $i = 1, 2$. In this case, the LMPI test of $H_0 : \Sigma_{12} = 0$ rejects for large values of ψ_0 given by (4.5).

For the remainder of this section, we will be concerned only with the data given by (4.1), and the problem of testing $H_0 : \Sigma_{12} = 0$ versus $H_1 : \Sigma_{12} \neq 0$. When both m_1 and m_2 are positive, we have been unable to calculate the likelihood ratio test (LRT) of H_0 versus H_1. However, in Case (i) when $m_2 = 0$, the likelihood ratio is not difficult to derive.

PROPOSITION 4.1. *With the data of Case* (i), *the* LRT *of H_0 versus H_1 rejects for small values of*

$$\Lambda = | I - S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} | \tag{4.8}$$

PROOF. This is a routine calculation and is omitted.

It is rather surprising that the LRT ignores the data V_1 in testing H_0. Of course, when $m_1 = m_2 = 0$, rejecting for small Λ values gives the LRT. However, when m_2 is very large, the value of Σ_{22} is essentially known but the likelihood ratio criterion ignores this information. Indeed, if Σ_{22} is known, the LRT, based on S alone, for testing $H_0 : \Sigma_{12} = 0$ is also that given in Proposition 4.1.

It is not clear what to do in practice for testing H_0 versus H_1. One possibility is to ignore the two ancillary statistics $\operatorname{tr}(S_{ii} + W_{ii})^{-1} S_{ii}$, $i = 1, 2$, in (4.5) and reject H_0 for large values of

$$\tilde{\psi}_1 = \operatorname{tr}(S_{11} + W_{11})^{-1} S_{12} (S_{22} + W_{22})^{-1} S_{21}.$$

The null distribution of ψ_0 and $\tilde{\psi}_1$ is not known. Letting n and m_i tend to ∞ with $m_i/n \to \beta_i$, $i = 1, 2$, it is not too hard to show that $n\tilde{\psi}_1$ converges in distribution to a random variable with a scaled Chi squared distribution. In particular,

$$n\tilde{\psi}_1 \to_d (1 + \beta_1)(1 + \beta_2)\chi^2_{p_1 p_2}.$$

When n is large, this provides one possible method of testing H_0 versus H_1.

Now, consider the special case when $p_2 = 1$ and $m_1 = 0$. The problem of testing for independence is similar in structure to a mean testing problem discussed by Giri (1968). Even though the maximal invariant parameter is one dimensional, a uniformly most powerful invariant test does not exist and the LRT is not the locally best test (Theorem 4.1). As in the problem treated by Giri (1968), a natural maximal invariant in the sample space is two dimensional say (ξ_1, ξ_2), and ξ_2 is an ancillary statistic. The LRT rejects for large values of ξ_1 while the locally best test involves both ξ_1 and ξ_2. The details of this are given in Eaton and Kariya (1974). A related reference is Marden (1978).

Finally, consider the special case of $p_1 = p_2 = 1$ so $p = 2$ and the data is given by (4.1). A minimal sufficient statistic is (S, W_{11}, W_{22}) where $S = V'V$ and $W_{ii} = V_i'V_i$ $i = 1, 2$. In this case, the problem is to test that $\rho = 0$ where ρ is the bivariate correlation coefficient. The testing problem is invariant under scale changes and a maximal invariant statistic is $T = (t_1, t_2, t_3)$ where

$$t_1 = S_{12}^2 / S_{11} S_{22}, \quad t_2 = W_{11}/S_{11}, \quad t_3 = W_{22}/S_{22}.$$

When $m_1 = 0$ (so W_{11} is not present and t_2 is not present), the LRT rejects for large values of t_1 while the LMPI test involves both t_1 and t_3. Since t_3 is ancillary, it may be most reasonable to condition on t_3 and test $\rho = 0$ conditionally. But, when both m_1 and m_2 are positive there is a complication. The statistics t_2 and t_3 are marginally ancillary but (t_2, t_3) is not an ancillary statistic. It is not clear how to condition in this case, but rejecting for large t_1 is not appropriate.

5. Examples. In this section, we present two problems which are special cases of the problems discussed in Section 4. The notation used in these two examples is independent of the notation used earlier.

EXAMPLE 5.1. In this example, we consider the problem of testing independence in a model of covariate discriminant analysis. Suppose $X_1, \cdots, X_M$ are i.i.d. $N_p(\mu, \Sigma)$ and $Y_1, \cdots, Y_N$ are i.i.d. $N_p(\nu, \Sigma)$ and write the dimension parameter p as $p = p_1 + p_2$. Partitioning the data and parameters, we have

$$X_i = \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix}, \quad Y_j = \begin{pmatrix} Y_j^{(1)} \\ Y_j^{(2)} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \nu = \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Of course, $X_i^{(\alpha)}$, and $Y_j^{(\alpha)}$, μ_α and ν_α are $p_\alpha \times 1$ and $\Sigma_{\alpha\beta}$ is $p_\alpha \times p_\beta$, $\alpha, \beta = 1, 2$. It is assumed that $\mu_2 = \nu_2$. Discrimination problems in this situation have been considered by Cochran and Bliss (1948), Rao (1949), Cochran (1964), Rao (1966) and Memon and Okamoto (1970). A survey of this situation is given in Kshirsargar (1972, page 200–203). When Σ is known and $\Sigma_{12} \neq 0$, Cochran and Bliss (1948) constructed a discriminant function based on all the data which is more efficient than the usual discriminant function based on $X_i^{(1)}$, $i = 1, \cdots, M$ and $Y_j^{(1)}, j = 1, \cdots, N$. When Σ is unknown, Cochran and Bliss (1948) proposed a discriminant function in which Σ is replaced by an estimate. However, when Σ_{12} is close to zero, this discriminant function does not seem to be better than the usual one based on $X_i^{(1)}$ and $Y_j^{(1)}$, $i = 1, \cdots, M$, $j \equiv 1, \cdots, N$. Of course, when $\Sigma_{12} = 0$, it seems most reasonable to base discrimination solely on the basis of $X_i^{(1)}$ and $Y_j^{(1)}$, $i = 1, \cdots, M$ and $j = 1, \cdots, N$. This motivates the problem of testing $\Sigma_{12} = 0$ in this situation.

After a reduction by invariance, we will show that testing $\Sigma_{12} = 0$ is a special case of the problem described in Section 4. As demonstrated in Section 2, the data $X_1, \cdots, X_M$ is equivalent (via a linear transformation) to (U_1, V_1) where $U_1 : p \times 1$ and $V_1 : (M - 1) \times p$ are independent and

$$\mathscr{L}(U_1) = N\left(\mu, \frac{1}{M}\Sigma\right), \quad \mathscr{L}(V_1) = N(0, I_{M-1} \otimes \Sigma).$$

Similarly $Y_1, \cdots, Y_N$ is equivalent to (U_2, V_2) which are independent and

$$\mathscr{L}(U_2) = N\left(\nu, \frac{1}{N}\Sigma\right), \quad \mathscr{L}(V_2) = N(0, I_{N-1} \otimes \Sigma).$$

The problem of testing $\Sigma_{12} = 0$ is obviously invariant under translations of U_1 and U_2 given by

$$U_1 \to U_1 + \begin{pmatrix} a \\ c \end{pmatrix}, \quad U_2 \to U_2 + \begin{pmatrix} b \\ c \end{pmatrix}$$

with $a, b \in R^{p_1}$ and $c \in R^{p_2}$. A maximal invariant is $W_2 \equiv k(U_1^{(2)} - U_2^{(2)})$, $k = (1/M + 1/N)^{-1/2}$. After this reduction by invariance, the data is W_2 and $V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ where

$$\mathscr{L}(W_2) = N(0, \Sigma_{22}), \quad \mathscr{L}(V) = N(0, I_{M+N-2} \otimes \Sigma)$$

and the problem is to test $\Sigma_{12} = 0$. In this form, the results of Theorem 4.1 are applicable (with $m_1 = 0$ and $W_{11} = 0$), so a locally most powerful invariant test exists and is given in Theorem 4.1. It is interesting to note that the LRT based on all the data of $H_0 : \Sigma_{12} = 0$ versus $H_1 : \Sigma_{12} \neq 0$ is the same as the LRT based only on the data matrix V. In other words, the LRT described in Proposition 4.1 ignores the extra information that $\mu_2 = \nu_2$. However, the LRT would not ignore the information that $\mu = \nu$.

Our second example concerns the growth curve model.

EXAMPLE 5.2. Consider a data matrix $Y : N \times p$ such that

$$\mathscr{L}(Y) = N(X_1 B X_2, I_N \otimes \Omega)$$

where X_1 is $N \times r$ of rank r, X_2 is $q \times p$ of rank q, and both are known (see Potthoff and

Roy, 1964). Also, $B: r \times q$ is a matrix of unknown parameters and Ω is a $p \times p$ positive definite matrix. The estimation of B is of concern here and of course the structure of Ω affects the estimation. Let Z_2 be a $(p-q) \times p$ matrix of rank $p-q$ such that $X_2 Z_2' = 0$. If Ω has the form

$$\Omega = X_2' \psi_1 X_2 + Z_2' \psi_2 Z_2 \tag{5.1}$$

where $\psi_1 : q \times q$ is positive definite and $\psi_2 : (p-q) \times p - q$ is positive definite, then the least squares (acting as if $\Omega = I_p$) estimator of B is also the Gauss-Markov and maximum likelihood estimator of B. This claim and its converse follow from results in Eaton (1970), or a modification of a result due to Rao (1967). The covariance structure (5.1) is known as Rao's covariance structure (Rao, 1967) and has been discussed in Geisser (1970). Further Lee and Geisser (1972) derived the LRT for testing that Ω has the form (5.1) versus arbitrary alternatives. We will show that, after a reduction by invariance, testing Ω has the form (5.1) is a special case of testing for independence with additional information. First, a transformation to a canonical form will simplify certain calculations. Let $Z_1 : N \times (N-r)$ and satisfy $Z_1' X_1 = 0$. Then, let

$$\Gamma_1 = [X_1(X_1'X_1)^{-1/2}, Z_1(Z_1'Z_1)^{-1/2}], \quad \Gamma_2 = [X_2'(X_2X_2')^{-1/2}, Z_2'(Z_2Z_2')^{-1/2}],$$

so Γ_1 is $N \times N$ and orthogonal and Γ_2 is $p \times p$ and orthogonal. Now, let

$$W = \Gamma_1' Y \Gamma_2, \quad \Sigma = \Gamma_2' \Omega \Gamma_2, \quad \mu = (X_1'X_1)^{1/2} B (X_2X_2')^{1/2}$$

and partition W and Σ as

$$W = \overset{\quad q \qquad p-q}{\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}} \begin{matrix} r \\ N-r \end{matrix}, \qquad \Sigma = \overset{\quad q \qquad p-q}{\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}} \begin{matrix} q \\ p-q \end{matrix}.$$

With this relabeling, we have

$$\mathscr{L}(W) = N\left(\begin{pmatrix} \mu & 0 \\ 0 & 0 \end{pmatrix}, (I_N \otimes \Sigma) \right).$$

The null hypothesis that Ω has the form (5.1) becomes $H_0 : \Sigma_{12} = 0$ when expressed in terms of Σ. This testing problem is invariant under the translations $W_{11} \to W_{11} + a$, $a : r \times q$ and a maximal invariant under this group of translations is $\{W_{12}, (W_{21}, W_{22})\}$. Clearly W_{12} is independent of (W_{21}, W_{22}),

$$\mathscr{L}(W_{12}) = N(0, I_r \otimes \Sigma_{22}) \quad \text{and} \quad \mathscr{L}((W_{21}, W_{22})) = N(0, I_{N-r} \otimes \Sigma).$$

Based on this data, testing H_0 is a special case of the problem treated in Case (i) following Theorem 4.1. Also, the result described in Proposition 4.1 shows that the LRT based on $\{W_{12}, (W_{21}, W_{22})\}$ ignores W_{12} and the LRT is different from the locally best test. Furthermore, this LRT is the same as the LRT based on all the data W which Lee and Geisser (1972) derived.

APPENDIX: PROOFS OF THEOREMS 3.2 AND 4.1.

Our attention is restricted to the proof of Theorem 4.1 since Theorem 3.2 is proved similarly. To prove Theorem 4.1, we apply Wijsman's (1967) representation theorem concerning the density function of a maximal invariant. Let $\mathscr{X} = \mathscr{X}_0 \times \mathscr{X}_1 \times \mathscr{X}_2$ denote the sample space of the data given by (4.1) where $\mathscr{X}_i$ is the linear space of $n_i \times p_i$ real matrices $(i = 0, 1, 2)$. Here the notation

$$n_0 = n, \quad n_1 = m_1, \quad n_2 = m_2 \quad \text{and} \quad \Sigma_{00} = \Sigma$$

is used, and we shall write $x = (x_0, x_1, x_2) \in \mathscr{X}$ for $X = (V, V_1, V_2) \in \mathscr{X}$. The Lebesque measure on $\mathscr{X}$ will be denoted by $dx = dx_0 dx_1 dx_2$ and from (4.1), the density of x is given by

$$f(x \mid \Sigma) = \prod_{i=0}^{2} (\sqrt{2\pi})^{-n_i/2} |\Sigma_{ii}|^{-n_i/2} \exp(-\tfrac{1}{2} \operatorname{tr} x_i \Sigma_{ii}^{-1} x_i'). \tag{A.1}$$

The group G_2 defined in Section 4 acts on $x \in \mathscr{X}$ by $gx = (x_0A', x_1A_1', x_2A_2')$ as in (4.3) where $g = A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ is given by (4.2) with $A_i \in G\ell p_i$. A left invariant measure on G_2 is

$$\nu(dg) = \nu_1(dA_1)\nu_2(dA_2), \quad \text{with} \quad \nu_i(dA_i) = |A_iA_i'|^{-p_i/2}\, dA_i, \quad i = 1, 2. \tag{A.2}$$

To establish Theorem 4.1, we will use the following well known argument. For any invariant test function φ, the power function of φ at a maximal invariant parameter point δ is

$$\pi(\varphi, \delta) = \int \varphi\, dP_\delta = \int \varphi(dP_\delta/dP_0)\, dP_0 \tag{A.3}$$

where P_δ is the probability measure of a maximal invariant under δ. The ratio dP_δ/dP_0 is obtained by the following lemma, which is an expression of Theorem 4 in Wijsman (1967) in the present problem. Assume $p_1 \geq p_2$ without loss of generality.

LEMMA. *The ratio dP_δ/dP_0 is given by*

$$r_\delta(x) = \int_{G_2} f(gx \mid \Sigma(\delta))\chi_0(g)\nu(dg) \Big/ \int_{G_2} f(gx \mid \Sigma(0))\chi_0(g)\nu(dg), \tag{A.4}$$

where

$$\chi_0(g) = \prod_{i=1}^{2} |A_iA_i'|^{(n+m_i)/2}, \quad \Sigma(\delta) = \begin{pmatrix} I_1 & \Delta \\ \Delta' & I_2 \end{pmatrix} \tag{A.5}$$

and Δ is $p_1 \times p_2$ with $\Delta_{ii} = \delta_i$ for $i = 1, \cdots, p_2$ and $\Delta_{ij} = 0$ for $i \neq j$.

Theorem 4 in Wijsman (1967) states the conditions for which (A.4) holds. But checking the conditions is included in the proof of Lemma 5.1 in Kariya (1978) and so it is omitted here.

To evaluate $r_\delta(x)$ in (A.4), let $x_0'x_0 = S = (S_{ij})$ and $x_i'x_i = W_{ii} (i = 1, 2)$ be as before (see 2.5)) and let

$$\begin{aligned} T &= \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \\ &\equiv \begin{pmatrix} (S_{11} + W_{11})^{-1/2} & 0 \\ 0 & (S_{22} + W_{22})^{-1/2} \end{pmatrix} S \begin{pmatrix} (S_{11} + W_{11})^{-1/2} & 0 \\ 0 & (S_{22} + W_{22})^{-1/2} \end{pmatrix}. \end{aligned} \tag{A.6}$$

Note that $0 < T < I_p$ in the sense of positive definiteness. Let $\mathscr{E}_{k,\alpha}$ denote expectation with respect to the distribution on $G\ell_k$ given by

$$P(dg) = c(\alpha, k)|gg'|^{\alpha/2} \exp(-\tfrac{1}{2} \operatorname{tr} gg')\nu_\ell(dg),$$

with $\nu_\ell(dg) = |g'g|^{-k/2} dg$, where $\alpha > 0$, and $c(\alpha, k)$ is a normalizing constant. A bit of algebra and a change of variable show that

$$r_\delta(x) = |\Sigma(\delta)|^{-n/2} \mathscr{E}_{p_1, n+m_1} \mathscr{E}_{p_2, n+m_2}(\exp[-\tfrac{1}{2} \operatorname{tr} TA'\{\Sigma^{-1}(\delta) - I_p\}A]), \tag{A.7}$$

where $\mathscr{E}_{p_i, n+m_i}$ is expectation on A_i, $i = 1, 2$. Define $\gamma: p \times p$ by

$$\gamma \equiv \Sigma^{-1}(\delta) - I_p = \begin{pmatrix} (I_1 - \Delta\Delta')^{-1} - I_1 & -(I_1 - \Delta\Delta')^{-1}\Delta \\ -(I_2 - \Delta'\Delta)^{-1}\Delta' & (I_2 - \Delta'\Delta)^{-1} - I_2 \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix},$$

where γ_{ij} is $p_i \times p_j$, $i = 1, 2$. Then we have

$$\operatorname{tr} TA'\{\Sigma^{-1}(\delta) - I_p\}A = \operatorname{tr} T_{11}A_1'\gamma_{11}A_1 + 2 \operatorname{tr} T_{12}A_2'\gamma_{21}A_1 + \operatorname{tr} T_{22}A_2'\gamma_{22}A_2.$$

We now make the following claim: For δ small, $\tau = \Sigma_1^{p_2} \delta_i^2$, and all T satisfying $0 < T < I_p$,

$$\text{(A.8)} \quad \begin{cases} \exp(-\tfrac{1}{2}\operatorname{tr} T_{ii}A_i'\gamma_{ii}A_i) = 1 - \tfrac{1}{2}\operatorname{tr} T_{ii}A_i'\Delta\Delta'A_i + R_i(T_{ii}, A_i, \Delta), \quad i = 1, 2, \\ \exp(-\operatorname{tr} T_{12}A_2'\gamma_{21}A_1) = 1 + \operatorname{tr} T_{12}A_2'\Delta'A_1 + \tfrac{1}{2}(\operatorname{tr} T_{12}A_2'\Delta'A_1)^2 \\ \quad + R_3(T_{12}, A_1, A_2, \Delta), \end{cases}$$

where the error terms R_1, R_2 and R_3 satisfy

$$\text{(A.9)} \quad \begin{cases} |R_i(T_{ii}, A_i, \Delta)| \le H_i(A_i)\eta_i(\tau), \quad i = 1, 2, \\ |R_3(T_{12}, A_1, A_2, \Delta)| \le H_3(A_1)H_4(A_2)\eta_3(\tau). \end{cases}$$

Further, the inequalities in (A.9) hold for all T, $0 < T < I_p$, the functions H_i are integrable (A_1, A_2) and $\lim_{\tau\to 0} \eta_i(\tau)/\tau = 0$, $i = 1, 2, 3$. The arguments leading to (A.8) and (A.9) are similar to those in Schwartz (1967) and Kariya (1978) and are omitted. The following identities are used in the evaluation of (A.7):

$$\text{(A.10)} \quad \begin{cases} \mathscr{E}_{p_i, n+m_i} \operatorname{tr} T_{ii}A_i'\Delta\Delta'A_i = \dfrac{n + m_i}{p_i} (\operatorname{tr} T_{ii})\tau, \quad i = 1, 2, \\ \mathscr{E}_{p_1, n+m_1}\mathscr{E}_{p_2, n+m_2} \operatorname{tr} T_{12}A_2'\Delta'A_1 = 0, \\ \mathscr{E}_{p_1, n+m_1}\mathscr{E}_{p_2, n+m_2} (\operatorname{tr} T_{12}A_2'\Delta'A_1)^2 = \dfrac{n + m_1}{p_1}\dfrac{n + m_2}{p_2} (\operatorname{tr} T_{12}T_{12}')\tau. \end{cases}$$

These identities are proved in a similar manner as in Kariya (1978) (see equation (5.10) there). Note that $|\Sigma(\delta)|^{-n/2} = 1 + (n/2)\tau + o(\delta)$ where $\lim_{\delta\to 0} o(\delta)/\tau = 0$. Substituting this and the expressions in (A.8) and (A.7) leads to the expression

$$\text{(A.11)} \quad r_\delta(x) = 1 + \tfrac{1}{2}\psi_0\tau + o(T, \delta),$$

where ψ_0 is defined in (4.5). The remainder term is uniformly bounded in T, $0 < T < I_p$, and satisfies $\lim_{\delta\to 0}\sup_{0<T<I_p} o(T, \delta)/\tau = 0$. The identities in (A.10) and the results expressed in (A.9) are used to establish (A.11).

Now, let φ be any level α invariant test of $H_0: \delta = 0$ versus $H_1: \delta \ne 0$. Substituting (A.11) into (A.3) yields

$$\pi(\varphi, \delta) = \int \varphi \frac{dp_\delta}{dP_0}\, dP_0 = \int \varphi \left[1 + \frac{1}{2}\psi_0\tau + o(T, \delta)\right] dP_0 = \alpha + \frac{1}{2}(\mathscr{E}_0\varphi\psi_0)\tau + o(\varphi, \delta)$$

with

$$\lim_{\delta\to 0}\sup_\varphi \frac{o(\varphi, \delta)}{\tau} = 0.$$

This proves Theorem 4.1.

REFERENCES

ANDERSSON, S. A. (1978). Invariant measures. Stanford University Technical Report No. 129.

ANDERSSON, S. A. (1982). Distributions of maximal invariants using quotient measures. *Ann. Statist.* **10** 955–961.

BHARGAVA, R. N. P. (1962). Multivariate tests of hypotheses with incomplete data. Technical Report No. 3, Applied Mathematics and Statistical Laboratories, Stanford University.

COCHRAN, W. G. and BLISS, C. I. (1948). Discriminant functions with covariance. *Ann. Math. Statist.* **19** 151.

COCHRAN, W. G. (1964). Comparison of two methods of handling covariates in discriminatory analysis. *Ann. Inst. Statist. Math.* **16** 43.

EATON, M. L. (1970). Gauss-Markov estimation for multivariate linear models: A coordinate free approach. *Ann. Math. Statist.* **44** 528–538.

EATON, M. L. (1972). *Multivariate Statistical Analysis.* Institute of Math. Statist., University of Copenhagen, Copenhagen, Denmark.

EATON, M. L. and KARIYA, T. (1975). Testing for independence with additional information. University of Minnesota Technical Report No. 238.

EATON, M. L. and KARIYA, T. (1975). Tests and means with additional information. University of Minnesota Technical Report No. 243.

EKBOHM, G. (1976). On comparing means in the paired case with incomplete data on both responses. *Biometrika* **63** 299–304.

FARRELL, R. (1976). *Techniques of Multivariate Calculation.* Springer Verlag, New York.

GEISSER, S. (1970). Bayesian Analysis of Growth Curves. *Sankhyā Ser. A* **32** 53–64.

GIRI, N. (1968). Locally and asymptotic tests of multivariate problem. *Ann. Math. Statist.* **39** 171–178.

GIRI, N. and KIEFER, J. (1964). Local and asymptotic minimax properties of a normal multivariate testing problem. *Ann. Math. Statist.* **35** 21–35.

HARTLEY, H. O. and HOCKING, R. R. (1971). The analysis of incomplete data. *Biometrics* **27** 783–823.

HOGG, R. V. (1961). On the resolution of statistical hypotheses. *J. Amer. Statist. Assoc.* **56** 978–989.

KARIYA, T. (1975). Hypothesis testing problems in the general MANOVA model. Ph.D. Thesis, University of Minnesota.

KARIYA, T. (1978). The general MANOVA problem. *Ann. Statist.* **6** 200–214.

KSHIRSARGAR, A. M. (1972). *Multivariate Analysis.* Dekker, New York.

LEE, J. C. and GEISSER, S. (1972). Growth curve prediction. *Sankhyā Ser. A* **34** 393–412.

LITTLE, R. J. A. (1976). Inference about means from incomplete multivariate data. *Biometrika* **63** 593–604.

MARDEN, J. (1978). Invariant tests on covariance matrices. *Ann. Statist.* **6** 1258–1266.

MEMON, A. Z. and OKAMOTO, M. (1970). The classification statistic W^* in covariate discriminant analysis. *Ann. Math. Statist.* **41** 1491–1499.

MORRISON, D. F. and BHOJ, D. S. (1973). Power of the likelihood ratio test on the mean vector of the multivariate normal distribution with missing observations. *Biometrika* **60** 365–368.

NACHBIN, L. (1965). *The Haar Integral.* Van Nostrand, New York.

POTTHOFF, R. F. and ROY, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve models. *Biometrika* **51** 313–326.

RAO, C. R. (1966). Covariance adjustment and related problems in multivariate analysis. *Multivariate Analysis I* (p. 87) (edited by P. R. Krishnaish). Academic, New York.

RAO, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. *Proc. 5th Berk. Symp. Math. Statist. and Probab.* **1** 355–372.

RUBIN, D. B. (1976). Inference and missing data. *Biometrika* **63** 581–592.

SCHWARTZ, R. (1967). Local minimax tests. *Ann. Math. Statist.* **38** 340–360.

STEIN, C. (1966). Notes on multivariate analysis. Stanford University. (Notes recorded by M. L. Eaton).

WIJSMAN, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Fifth Berk. Symp. Math. Statist. Probab.* **1** 389–400, University of California.

DEPARTMENT OF THEORETICAL STATISTICS
UNIVERSITY OF MINNESOTA
206 CHURCH STREET S.E.
MINNEAPOLIS, MINNESOTA 55455

THE INSTITUTE OF ECONOMIC RESEARCH
HITOTSUBASHI UNIVERSITY
KUNITACHI, TOKYO
186 JAPAN