JOURNAL OF MULTIVARIATE ANALYSIS 14, 155–168 (1984)

A Condition for Null Robustness

MORRIS L. EATON*

University of Minnesota

AND

TAKEAKI KARIYA†

University of Pittsburgh and Hitotsubashi University, Japan

Communicated by P. R. Krishnaiah

Sufficient conditions are given that certain statistics have a common distribution under a wide class of underlying distributions. Invariance methods are the primary technical tool in establishing the theoretical results. These results are applied to MANOVA problems, problems involving canonical correlations, and certain statistics associated with the complex normal distribution.

1. INTRODUCTION

Using a geometric argument, Fisher [10] showed that Student's one sample t-statistic has the same null distribution under normality as under the assumption of spherical symmetry (see Efron [8] for a discussion and some related topics). This fact about the t-statistic is due to two things:

(i) The t-statistic is scale-invariant.

(ii) The uniform distribution on S_{n-1} (the sphere of radius 1 in R^n) is the unique spherically symmetric distribution on S_{n-1}.

Basically, we have used this observation to formulate a general method for

Received July 7, 1981; revised December 3, 1982.

AMS 1970 subject classifications: Primary 62E10; Secondary 62H05, 62H10.

Key words and phrases. Null robustness, invariance, MANOVA, canonical correlations, tests of independence, multivariate normal distribution theory, complex multivariate normal.

* This research was supported in part by National Science Foundation Grant MCS 77-25112.

† This work was supported by the Air Force Office of Scientific Research under Contract F49629-82-K-001 and by the Andrew Mellon Fellowship when the author was at the University of Pittsburgh.

0047-259X/84 $3.00

proving similar results about other statistics of interest. Some additional properties of the t-test under spherical symmetry are established in Kariya and Eaton [19]. Recently, multivariate analogues of some of these results have been developed by Dawid [5], Fraser and Ng [11], Kariya [17, 18], and Jensen and Good [15].

Let $(\mathscr{X}, \mathscr{B})$ be a measurable space and suppose X takes values in $\mathscr{X}$. We write $\mathscr{L}(X) = P$ to mean that the distribution of X is P. If $t(X)$ is any statistic, $\mathscr{L}(t(X) \mid P)$ denotes the distribution of $t(X)$ when $\mathscr{L}(X) = P$. Now, suppose that the distribution of $t(X)$ is known when $\mathscr{L}(X) = P_0$ and set

$$\mathscr{P} = \{P \mid \mathscr{L}(t(X) \mid P) = \mathscr{L}(t(X) \mid P_0)\}. \tag{1.1}$$

It would be of interest to supply some useful sufficient conditions which imply that $P \in \mathscr{P}$. Using invariance assumptions, this is what is done in Theorem 2.1. This result is related to a result in Das Gupta [4].

EXAMPLE 1.1. Take $\mathscr{X}$ to be the set of all $n \times p$ real matrices of rank p (so $n \geqslant p$) and suppose $X \in \mathscr{X}$. A number of important statistics which arise in testing problems in MANOVA (see Section 3) can be written as functions of

$$T = t(X) = X(X'X)^{-1}X'$$

which is a random orthogonal projection of rank p. The distributions of these functions of $t(X)$ can often be computed when the elements of X are i.i.d. $N(0, 1)$—let this be P_0. We now assert that if $\mathscr{L}(X) = \mathscr{L}(\Gamma X)$ for each $\Gamma \in \mathscr{O}_n$ (the orthogonal group), then $\mathscr{L}(X) \in \mathscr{P}$ which is defined by (1.1). To see this, first observe that $t(\Gamma X) = \Gamma t(X) \Gamma'$ so

$$\mathscr{L}(T) = \mathscr{L}(\Gamma T \Gamma'), \qquad \Gamma \in \mathscr{O}_n, \tag{1.2}$$

when $\mathscr{L}(X) = \mathscr{L}(\Gamma X)$. However, (1.2) characterizes the distribution of T because: (i) $\mathscr{O}_n$ is compact and (ii) $\mathscr{O}_n$ acts transitively on the set of $n \times n$ rank p orthogonal projections (see Nachbin [21, Chap. 3]). The conclusion is that for any function f,

$$\mathscr{L}(f(t(X)) \mid P_0) = \mathscr{L}(f(t(X)) \mid P),$$

so long as $\mathscr{L}(X) = \mathscr{L}(\Gamma X)$ when $\mathscr{L}(X) = P$.

Example 1.1 contains the elements of the argument which led to the theorem in Section 2. Once this result is established, the rest of the paper consists of examples from multivariate analysis. In Section 3, we discuss the MANOVA problem. In Section 4, we provide sufficient conditions that the sample canonical correlations have the same distribution as if the variables were independent normals. Some applications to problems involving the complex normal distribution are given in Section 5.

2. Main Result

In this section, we give our main results. To set notation, R^k will denote the k-dimensional Euclidean space of column vectors, x' denotes the transpose of $x \in R^k$, $\mathcal{O}_k$ is the group of $k \times k$ orthogonal matrices and Gl_k is the group of $k \times k$ nonsingular matrices. If X is a random vector, $\mathcal{L}(X)$ will denote the distribution of X. Also, $N(\mu, I_n \otimes \Sigma)$ denotes the normal distribution on the vector space of $n \times p$ matrices. Here, μ is the mean matrix and $I_n \otimes \Sigma$ is the Kronecker product of the $n \times n$ identity matrix with the $p \times p$ positive definite matrix Σ.

Suppose that $(\mathcal{X}, \mathcal{B})$ is a measurable space and G_0 is a group which acts measurably on the left of $\mathcal{X}$. Let $\mathcal{M}(\mathcal{X})$ be the set of all probability measures on $(\mathcal{X}, \mathcal{B})$. If $P \in \mathcal{M}(\mathcal{X})$ and $g \in G_0$, then gP denotes the probability defined by

$$(gP)(B) = P(g^{-1}B), \qquad B \in \mathcal{B}.$$

Now suppose K is a compact group acting on $\mathcal{X}$ and suppose t is a measurable mapping from $(\mathcal{X}, \mathcal{B})$ onto $(\mathcal{Y}, \mathcal{C})$. Let

$$\mathcal{P}_K = \{P \mid P \in \mathcal{M}(\mathcal{X}), P = kP \text{ for all } k \in K\}.$$

The basic result in this paper provides sufficient conditions so that

$$\mathcal{L}(t(X) \mid P) = \mathcal{L}(t(X) \mid P') \qquad \text{for all} \quad P, P' \in \mathcal{P}_K. \tag{2.1}$$

Consider the following two conditions:

(A) For every $k \in K$, $t(x_1) = t(x_2)$ implies $t(kx_1) = t(kx_2)$.

(B) Given $x_1, x_2 \in \mathcal{X}$, there exists $k \in K$ such that $t(x_1) = t(kx_2)$.

Condition (A) allows one to define K acting on $\mathcal{Y}$ (see Hall, Wijsman, and Ghosh [14]) and (B) implies this action on $\mathcal{Y}$ is transitive. Assuming (A) and (B), Das Gupta [4] proved that (2.1) holds. In addition, Das Gupta [4] noted that if a group G acts transitively on $\mathcal{X}$ and if G is a direct product, $G = H \times K$ with K as above, then any H-maximal invariant statistic t satisfies conditions (A) and (B). There are many examples to which Das Gupta's approach applies—see Sections 3 and 5. However, there are also important examples (see Section 4) for which (A) does not hold but (2.1) does hold. Theorem 2.1 gives a different set of conditions which imply that (2.1) holds.

To describe our main result, assume that G is a locally compact topological group which acts transitively on $\mathcal{X}$, and K is a compact subgroup of G. In addition, assume that H is a subgroup of G so that

$$G = K \cdot H \equiv \{kh \mid k \in K, h \in H\}.$$

THEOREM 2.1. *Let t be any H invariant statistic. If either H or K is a normal subgroup of G, then*

$$\mathscr{L}(t(X) \mid P) = \mathscr{L}(t(X) \mid P') \qquad \textit{for} \quad P, P' \in \mathscr{P}_K. \tag{2.2}$$

Proof. The proof will be given for the case that K is normal in G. The proof for the other case is easier. It suffices to show that for any bounded measurable f,

$$\int f(t(x))\, P(dx) = \int f(t(x))\, P'(dx) \qquad \text{for} \quad P, P' \in \mathscr{P}_K. \tag{2.3}$$

Since $kP = P$ for $P \in \mathscr{P}_K$, we have

$$\int f(t(x))\, P(dx) = \int f(t(kx))\, P(dx). \tag{2.4}$$

Let μ be the unique invariant probability measure on K. Integrating both sides of (2.4) and using the Fubini theorem entail

$$\int f(t(x))\, P(dx) = \iint f(t(kx))\, \mu(dk)\, P(dx). \tag{2.5}$$

Fix $x_0 \in \mathscr{X}$ and use the transitivity of $G = K \cdot H$ to write $x = k_1 h x_0$ for a fixed $x \in \mathscr{X}$. Then the invariance of μ and the invariance of t under H give

$$\int f(t(kx))\, \mu(dk) = \int f(t(kk_1 hx_0))\, \mu(dk) = \int f(t(h^{-1}khx_0))\, \mu(dk). \tag{2.6}$$

Since the map $k \to h^{-1}kh$ is a continuous isomorphism of K, the uniqueness of μ implies that μ is invariant under this map. Applying this to (2.6) yields

$$\int f(t(kx))\, \mu(dk) = \int f(t(kx_0))\, \mu(dk), \tag{2.7}$$

for all $x \in \mathscr{X}$ and all bounded measurable f. Substituting (2.7) into (2.5) gives

$$\int f(t(x))\, P(dx) = \int f(t(kx_0))\, \mu(dk),$$

for all $P \in \mathscr{P}_K$. This establishes (2.3) and the theorem.

Remark 2.1. If H is normal in G, and t is a maximal H invariant function, then it is easy to show that conditions (A) and (B) hold so Das Gupta's result implies our result (for any H invariant statistic). However, when K is normal in G and H is not, there are interesting examples where condition (A) does not hold—see the canonical correlation example in Section 4.

We end this section with a description of the elements of $\mathscr{P}_K$. In what follows, it is assumed that

$\mathscr{S}$ is a measurable subset of $(\mathscr{X}, \mathscr{B})$ such that $S \cap \{kx \mid k \in K\}$ consists of exactly one point—say $s(x)$—and $x \to s(x)$ from $(\mathscr{X}, \mathscr{B})$ to $(\mathscr{S}, \mathscr{B}_0)$ is measurable. Here, $\mathscr{B}_0$ is the σ-algebra on $\mathscr{S}$ inherited from $(\mathscr{X}, \mathscr{B})$. (2.8)

This assumption simply means that there exists a measurable cross section in $\mathscr{X}$. It is well known that the statistic $s(X)$ is sufficient for $\mathscr{P}_K$ (see Farrell [9]). Let R denote a probability measure on $(\mathscr{S}, \mathscr{B}_0)$ and $\bar{R}$ denote the extension of R to $(\mathscr{X}, \mathscr{B})$, that is,

$$\bar{R}(B) = R(B \cap \mathscr{S}), \qquad B \in \mathscr{B}.$$

Given any such R, define P by

$$P(B) = \int \bar{R}(k^{-1}B)\,\mu(dk), \tag{2.9}$$

where μ is invariant probability measure on K. It is easily verified that $P \in \mathscr{P}_K$. However, it is not too difficult to show that if $P \in \mathscr{P}_K$, then (2.9) holds for some R. Hence (2.9) gives a representation of all the elements of $\mathscr{P}_K$. In terms of random variables, the representation (2.9) is expressed as follows. Let the random group element $U \in K$ have the "uniform" distribution μ and be independent of $S \in \mathscr{S}$. With $X = US$ (the group element U acting on $S \in \mathscr{X}$), it is clear that $\mathscr{L}(X) \in \mathscr{P}_K$. Conversely, if $\mathscr{L}(X) \in \mathscr{P}_K$, one can construct (using (2.9)) an $S \in \mathscr{S}$ which is independent of U and $\mathscr{L}(X) = \mathscr{L}(US)$. The representation (2.9) is discussed in Eaton [7].

Remark 2.2. Returning to the general case, suppose K is compact, acts on $\mathscr{X}$, and t is a statistic on $\mathscr{X}$. With $\mathscr{P}_K$ as above, suppose we want to give a condition so that

$$\mathscr{L}(t(X) \mid P) = \mathscr{L}(t(X) \mid P') \qquad \text{for} \quad P, P' \in \mathscr{P}_K. \tag{2.10}$$

Using the representation of elements of $\mathscr{P}_K$ given previously, (2.10) holds iff the distribution of $t(Ux)$ does not depend on $x \in \mathscr{X}$. (Here, U is uniform on K.) It is exactly this condition which Das Gupta's conditions and our conditions imply.

3. First Applications

In this section we apply the techniques described in the previous sections to two classical problems in multivariate analysis—namely, the MANOVA problem and the problem of testing for the equality of two covariance matrices. A canonical form of the MANOVA model can be written

$$X = \begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \\ 0 \end{pmatrix} + E,$$

where X is $n \times p$, U and B_1 are $n_1 \times p$, V and B_2 are $n_2 \times p$, W is $n_3 \times p$, and E, the matrix of errors, is $n \times p$. We assume $n_3 \geqslant p$. The MANOVA problem is to test $H_0 : B_2 = 0$ versus $H_1 : B_2 \neq 0$. The sample space for this example is $\mathscr{X}$—the space of $(n_1 + n_2 + n_3) \times p$ matrices

$$x = \begin{pmatrix} u \\ z \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \qquad u: n_1 \times p, \qquad z: (n_2 + n_3) \times p,$$

such that z has rank p. When E is $N(0, I_n \otimes \Sigma)$ with Σ positive definite and unknown, a standard invariance argument (see Eaton [6, Chap. 9]) shows that all fully invariant tests are based on $t_0(X)$ which is the vector of the ordered nonzero eigenvalues of

$$V(V'V + W'W)^{-1}V' = V(Z'Z)^{-1}V',$$

where

$$Z = \begin{pmatrix} V \\ W \end{pmatrix}.$$

Let P_0 denote the $N(0, I_n \otimes I_p)$ distribution on $\mathscr{X}$. It is well known that $\mathscr{L}(t_0(X) \mid P_0)$ is the same as the distribution of $t_0(X)$ when

$$\mathscr{L}(X) = N\left[\begin{pmatrix} B_1 \\ 0 \\ 0 \end{pmatrix}, I_n \otimes \Sigma\right],$$

for any B_1 and any Σ.

We will now apply Theorem 2.1 to obtain a larger class of distributions for which the distribution of $t_0(X)$ is the same as when $\mathscr{L}(X) = P_0$. The technique in Dawid [5] will also yield our results for this example. To apply

Theorem 2.1, consider the group G whose elements are (Γ, A, C), where $\Gamma \in \mathcal{O}_{n_2+n_3}$, $A \in Glp$, and C is an $n_1 \times p$ matrix. The action of G on $\mathscr{X}$ is

$$(\Gamma, A, C)\begin{pmatrix} u \\ z \end{pmatrix} = \begin{pmatrix} uA' + C \\ \Gamma z A' \end{pmatrix},$$

and the group operation is

$$(\Gamma_1, A_1, C_1)(\Gamma_2, A_2, C_2) = (\Gamma_1\Gamma_2, A_1A_2, C_2A_1' + C_1).$$

The action of G is transitive. With

$$H = \{(\Gamma, A, C) \in G \mid \Gamma = I_{n_2+n_3}\}$$

and

$$K = \{(\Gamma, A, C) \in G \mid A = I_p, C = 0\},$$

it follows that $G = K \cdot H$, H is normal in G and K is compact.

Let $\mathscr{Y}$ be the set of all $(n_2+n_3) \times (n_2+n_3)$ rank p orthogonal projections on $R^{n_2+n_3}$ and equip $\mathscr{Y}$ with the usual topology. The function t on $\mathscr{X}$ to $\mathscr{Y}$ defined by

$$t(x) = z(z'z)^{-1}z', \qquad x = \begin{pmatrix} u \\ z \end{pmatrix},$$

is measurable and is H invariant. Note that t_0 defined earlier is a function of t since the upper left $n_2 \times n_2$ block of $t(x)$ is $v(z'z)^{-1}v' = v(v'v + w'w')^{-1}v'$. By Theorem 2.1, $\mathscr{L}(t(X) \mid P) = \mathscr{L}(t(X) \mid P_0)$ for any $P \in \mathscr{P}_K$. In particular, if the distribution of X satisfies

$$\mathscr{L}(X) = \mathscr{L}\left(\begin{pmatrix} U \\ Z \end{pmatrix}\right) = \mathscr{L}\left(\begin{pmatrix} U \\ \Gamma Z \end{pmatrix}\right), \tag{3.1}$$

for $\Gamma \in \mathcal{O}_{n_2+n_3}$, then the distribution of $t(X)$ is the same as if X is $N(0, I_n \otimes I_p)$. Since t_0 is a function of t, the same conclusion holds for t_0.

We now turn to a brief discussion of testing for the equality of two covariance matrices. For simplicity, the case of zero means is treated—the general case can be handled by a similar argument. For this problem, consider two independent data matrices X_i: $n_i \times p$, $i = 1, 2$. When $\mathscr{L}(\mathscr{X}_i) = N(0, I_{n_i} \otimes \Sigma_i)$ and we wish to test H_0: $\Sigma_1 = \Sigma_2$ versus H_1: $\Sigma_1 \neq \Sigma_2$, fully invariant tests are based on the p nonzero eigenvalues of $X_1(X_1'X_1 + X_2X_2')^{-1}X_1'$. Set

$$Z = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

and

$$t(Z) = Z'(Z'Z)^{-1}Z'.$$

Proceeding as in the MANOVA problem (with U absent) shows that the distribution of $t(Z)$ when Z is $N(O, I_n \otimes I_p)$ is the same as when $\mathscr{L}(Z) = P \in \mathscr{P}_K$, where $K = \mathscr{O}_{n_1+n_2}$ for this problem. Thus the distributions of fully invariant tests will be the same under a $N(O, I_n \otimes I_p)$ distribution as under any distribution for Z which satisfies $\mathscr{L}(Z) = \mathscr{L}(\Gamma Z)$, $\Gamma \in \mathscr{O}_{n_1+n_2}$.

Remark 3.1. These two examples provide results which are slightly stronger than the corresponding results in Dawid [5] and Jensen and Good [15]. However, the underlying argument in these two papers is very close to that given here.

Remark. The generalized MANOVA problem was introduced in Potthoff and Roy [22] and discussed at length in Gleser and Olkin [12] and Kariya [16]. The techniques used on the MANOVA problem above can also be used on the GMANOVA problem to yield corresponding results. The details are omitted.

4. Canonical Correlations

In this section, we discuss the distributions of canonical correlations. Without essential loss of generality we consider the mean zero case. The sample space for this section is $\mathscr{X}$, the set of $n \times p$ matrices of rank p. Consider $Z \in \mathscr{X}$ and partition Z as $Z = (Z_1, Z_2)$, where Z_i is $n \times p_i$, $i = 1, 2$. The orthogonal projections

$$Q_i = Z_i(Z_i'Z_i)^{-1}Z_i', \qquad i = 1, 2,$$

are elements of $\mathscr{Y}_{n,p_i}$, $i = 1, 2$, where $\mathscr{Y}_{n,p}$ is the space of $n \times n$ orthogonal projections of rank p. The *squared canonical correlations* are defined to be the $r \equiv \min\{p_1, p_2\}$ largest eigenvalues of Q_1Q_2. To see that this definition agrees with the standard definition in terms of the sample covariance matrix $S = Z'Z$, partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \qquad S_{ij}: p_i \times p_j, \tag{4.1}$$

for $i, j = 1, 2$. Classically, the squared canonical correlations are defined to be the r-largest eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$. But $S_{ij} = Z_i'Z_j$, so

$$Q_1Q_2 = Z_1 S_{11}^{-1}S_{12}S_{22}^{-1}Z_2'.$$

However, the nonzero eigenvalues of $Z_1 S_{11}^{-1} S_{12} S_{22}^{-1} Z_2'$ (of which there are at most r) are the same as the nonzero eigenvalues of

$$S_{11}^{-1} S_{12} S_{22}^{-1} Z_2' Z_1 = S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}.$$

Thus, our definition coincides with the usual definition.

Given $Z \in \mathscr{X}$, let $t(Z)$ be the vector of the r largest eigenvalues (arranged in order) of $Q_1 Q_2$. When Z is $N(O, I_n \otimes I_p) \equiv P_0$, the density of $t(Z)$ is known (see Anderson [1, (1958), Chap. 13]). Here we will describe a large class of distributions under which $t(Z)$ has the same distribution as when $\mathscr{L}(Z) = P_0$. Consider the group G whose elements are (ψ, Γ, A, B) with $\psi, \Gamma \in \mathscr{O}_n$, $A \in Gl_{p_1}$, and $B \in Gl_{p_2}$. The action of G on $\mathscr{X}$ is

$$(\psi, \Gamma, A, B)(z_1, z_2) = (\psi z_1 A', \Gamma z_2 B'),$$

and G acts transitively on $\mathscr{X}$. The group operation is

$$(\psi_1, \Gamma_1, A_1, B_1)(\psi_2, \Gamma_2, A_2, B_2) = (\psi_1 \psi_2, \Gamma_1 \Gamma_2, A_1 A_2, B_1 B_2).$$

Let

$$H = \{(\psi, \Gamma, A, B) \in G \mid \psi = \Gamma\}$$

and

$$K = \{(\psi, \Gamma, A, B) \in G \mid \Gamma = I_n, A = I_{p_1}, B = I_{p_2}\}.$$

Then $G = K \cdot H$, K is compact and K is normal in G. Thus Theorem 2.1 is applicable and we have

$$\mathscr{L}(t(X) \mid P_0) = \mathscr{L}(t(Z) \mid P),$$

for all $P \in \mathscr{P}_K$ since $P_0 \in \mathscr{P}_K$. To describe $\mathscr{P}_K$, first observe that elements of K act on $\mathscr{X}$ by

$$(\psi, I_n, I_{p_1}, I_{p_2})(z_1, z_2) = (\psi z_1, z_2).$$

Thus $\mathscr{L}(Z) \in \mathscr{P}_K$ iff $Z = (Z_1, Z_2)$ has the same distribution as $(\psi Z_1, Z_2)$, $\psi \in \mathscr{O}_n$. This certainly occurs when

$$\mathscr{L}(Z) = N(O, I_n \otimes \bar{\Sigma}),$$

where

$$\bar{\Sigma} = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}, \qquad \Sigma_{ii}: p_i \times p_i,\ i = 1, 2.$$

When Z has a density on $\mathscr{X}$ which can be written as

$$f(z_1, z_2) = f_1(z_1' z_1, z_2), \tag{4.2}$$

for some f_1, then it is easy to show that

$$\mathscr{L}(Z_1, Z_2) = \mathscr{L}(\psi Z_1, Z_2), \tag{4.3}$$

for $\psi \in \mathscr{O}_n$. (When Z has a density, condition (4.1) is almost necessary for $\mathscr{L}(Z_1, Z_2) = \mathscr{L}(\psi Z_1, Z_2)$—measure-theoretic difficulties being the only problem). It should be noted that, in this example, H is not normal in G and condition (A) fails to hold for the statistic $t(Z)$. Thus, it makes no sense to speak of K including a transitive action on the space of t values. This example seems to be new.

For comparative purposes, we now discuss a recent result due to Jensen and Good [15] concerning canonical correlations. Consider a random matrix $Z: n \times p$ with a density of the form

$$f(z) = |\Sigma|^{-n/2} f_2(\operatorname{tr} z'z\Sigma^{-1}), \tag{4.4}$$

where $\Sigma: p \times p$ is positive definite. Partition Z as above into $Z_i: n \times p_i$, $i = 1, 2$, and let $t(Z)$ be the vector of $r \equiv \min\{p_1, p_2\}$ canonical correlations. Also, partition Σ as S is partitioned in (4.1) and let $\theta(\Sigma)$ be the vector of population canonical correlations based on Σ. The argument used by Jensen and Good [15] shows that the distribution of $t(Z)$ does not depend on f_2 and depends on Σ only through $\theta(\Sigma)$. The essence of this argument is that when Z has a density given by (4.4), then $\mathscr{L}(Z) = \mathscr{L}(X\Sigma^{1/2})$, where $X: n \times p$ has density $f_2(\operatorname{tr} x'x)$. But the distribution of X is invariant under *all* orthogonal transformations on np-dimensional coordinate space so $U = (\operatorname{tr} X'X)^{-1/2} X$ has the same distribution as if the elements of X were i.i.d. $N(0, 1)$. Since the statistic t is scale-invariant, we have

$$\mathscr{L}(t(Z)) = \mathscr{L}(t(X\Sigma^{1/2})) = \mathscr{L}(t(U\Sigma^{1/2})),$$

so $\mathscr{L}(t(Z))$ does not depend on f_2. This argument shows that when Z has density (4.4), then the distribution of $t(Z)$ is the same as if Z is $N(O, I_n \otimes \Sigma)$. Thus, with a strong assumption on the density of Z, namely (4.4), Jensen and Good show that both the null ($\theta(\Sigma) = 0$) and the nonnull distributions of $t(Z)$ are the same as if Z is normal. When $r = 1$, this result implies that the test of $H_0: \theta(\Sigma) = 0$ versus $H_1: \theta(\Sigma) > 0$ which rejects for large values of $t(Z)$ is UMP invariant for all densities f_2. However, none of these results are valid for the more general class of densities given in (4.2)—only the null distribution of $t(Z)$ is independent of f_1.

5. Some Complex Normal Problems

In this section we discuss two problems related to some recent results of Andersson and Perlman [2]. To describe the situation, suppose that we have a random sample with

$$\mathscr{L}\left[\begin{pmatrix} X_i \\ Y_i \end{pmatrix}\right] = N(O, \Sigma), \qquad i = 1,..., n,$$

with $X_i \in R^p$ and $Y_i \in R^p$, and partition Σ as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \qquad \Sigma_{ij}: p \times p.$$

Consider the following three classes of $(2p) \times (2p)$ covariances:

$$\mathscr{C}_1 = \{\Sigma \mid \text{ is positive definite}\},$$
$$\mathscr{C}_2 = \{\Sigma \mid \Sigma \in \mathscr{C}_1, \Sigma_{11} = \Sigma_{22}, \Sigma_{12} = -\Sigma_{21}\},$$
$$\mathscr{C}_3 = \{\Sigma \mid \Sigma \in \mathscr{C}_1, \Sigma_{11} = \Sigma_{22}, \Sigma_{12} = \Sigma_{21} = 0\}.$$

Khatri [20] considered the problem of testing $H_0^{(3)}$: $\Sigma \in \mathscr{C}_3$ versus $H_1^{(3)}$: $\Sigma \in \mathscr{C}_2$. Elements of $\mathscr{C}_2$ are usually said to have "complex structure" while those in $\mathscr{C}_1$ and $\mathscr{C}_3$ are said to have "real structure"—see Goodman [13] and Brillinger [3]. The above testing problem can be interpreted as testing that a complex normal random vector is in fact real. In contrast, the problem of testing $H_0^{(2)}$: $\Sigma \in \mathscr{C}_2$ versus $H_1^{(2)}$: $\Sigma \in \mathscr{C}_1$ is testing that a real normal is in fact complex. Both of these problems are discussed in detail in Andersson and Perlman [2]. They reduce both problems via invariance and establish many properties of invariant tests. In what follows, we apply the results of Section 2 to show that the null distribution of all invariant tests is the same under normality as under a wider class of distributions. The result and techniques are similar to those in Section 3.

First, we treat testing $H_0^{(3)}$ versus $H_1^{(3)}$. Write the data in matrix from to yield W: $(2n) \times p$ whose first n rows are $X_1',..., X_n'$ and whose second n rows are $Y_1',..., Y_n'$. Under $H_0^{(3)}$,

$$\mathscr{L}(W) = N(O, I_{2n} \otimes \Sigma_{11}).$$

The sample space for W is taken to be $\mathscr{X}$—the set of real $(2n) \times p$ matrices of rank p. Take G to be the group whose elements are (Γ, A) with $\Gamma \in \mathscr{O}_{2n}$, $A \in Gl_p$, and group operation $(\Gamma_1, A_1)(\Gamma_2, A_2) = (\Gamma_1\Gamma_2, A_1A_2)$. The action of G on $\mathscr{X}$ is $(\Gamma, A)W = \Gamma W A'$, so G is transitive on $\mathscr{X}$. Also, take

$$H = \{(\Gamma, A) \in G \mid \Gamma = I_{2n}\}$$

and

$$K = \{(\Gamma, A) \mid A = I_p\},$$

so K is compact and both H and K are normal in G. The function

$$t(W) = W(W'W)^{-1}W',$$

is a maximal invariant under H.

Theorem 2.1 implies that the distribution of $t(W)$ under a $N(O, I_{2n} \otimes I_p)$ distribution for W is the same as under any distribution for W which satisfies $\mathscr{L}(W) = \mathscr{L}(\Gamma W)$ for $\Gamma \in \mathscr{O}_{2n}$. Now, all the tests of $H_0^{(3)}$ versus $H_1^{(3)}$ discussed in Andersson and Perlman [2] are invariant under the group H and are thus functions of $t(W)$. Hence the null distribution of all these tests is the same as when $\mathscr{L}(W) = \mathscr{L}(\Gamma W)$ for $\Gamma \in \mathscr{O}_{2n}$.

To discuss testing $H_0^{(2)}$ and $H_1^{(2)}$, it is convenient to introduce the complex random vectors

$$Z_j = X_j + iY_j, \qquad j = 1,..., n,$$

and form the data matrix $Z: n \times p$ with rows $Z_1^*,..., Z_n^*$ where $*$ denotes conjugate transpose. Under $H_0^{(2)}$, Z has a complex normal distribution,

$$L(Z) = \mathbb{C}N(O, I_n \otimes H),$$

where $H = \Sigma_{11} + i\Sigma_{12}$ and

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \in \mathscr{C}_2.$$

Of course H is Hermitian and positive definite. The sample space for Z is taken to be $\mathscr{X}$—the space of all $n \times p$ complex matrices of rank p. To apply Theorem 2.1, consider the group G whose elements are (U, A), where U is an element of $\mathscr{U}_n$ (the group of $n \times n$ unitary matrices) and $A \in \mathbb{C}Gl_p$ (the group of $p \times p$ nonsingular complex matrices). The action of G on $\mathscr{X}$ is

$$(U, A)Z = UZA^*,$$

so G is transitive on $\mathscr{X}$. Also take

$$H = \{(U, A) \in G \mid U = I_n\}$$

and

$$K = \{(U, A) \in G \mid A = I_p\},$$

so K is compact and both are normal in G.

A maximal invariant under the action of H on $\mathscr{X}$ is

$$t(Z) = Z(Z^*Z)^{-1}Z^*.$$

Let P_0 denote the $\mathbb{C}N(O, I_n \otimes I_p)$. Theorem 2.1 implies that $\mathscr{L}(t(Z) \mid P_0) = \mathscr{L}(t(Z) \mid P)$ for any probability measure P for which $\mathscr{L}(Z) = \mathscr{L}(UZ)$, $U \in \mathscr{U}_n$. All of the tests discussed in Andersson and Perlman [2] are H invariant and thus functions of $t(Z)$. Hence the null distribution under P_0 is the same as the null distribution under any P for which $\mathscr{L}(Z) = \mathscr{L}(UZ)$, $U \in \mathscr{U}_n$.

References

[1] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[2] Andersson, S. A., and Perlman, M. D. (1979). Two testing problems relating the real and complex multivariate normal distributions. Technical Report No. 95, Department of Statistics, University of Chicago.

[3] Brillinger, D. R. (1974). *Time Series: Data Analysis and Theory.* Holt, Rinehart, & Winston, New York.

[4] Das Gupta, S. (1979). A note on ancillarity and independence via measure-preserving transformations. *Sankhyā Ser. A* **41** 117–123.

[5] Dawid, A. P. (1977). Spherical matrix distributions and a multivariate model. *J. Roy. Statist. Soc. Ser. B* **39** 254–261.

[6] Eaton, M. L. (1972). *Multivariate Analysis.* University of Copenhagen.

[7] Eaton, M. L. (1983). *Multivariate Statistics.* Wiley, New York.

[8] Efron, B. (1969). Student's t-test under symmetry conditions. *J. Amer. Statist. Assoc.* **64** 1278–1302.

[9] Farrell, R. (1962). Representations of invariant measures. *Illinois J. Math.* **6** 447–467.

[10] Fisher, R. A. (1925). Applications of Student's distribution. *Metron* **5** 90–104.

[11] Fraser, D. A. S., and Ng, K. W. (1980). Multivariate regression analysis with spherical error. *Multivariate Analysis* V (P. R. Krishnaiah, Ed.), pp. 369–386. North-Holland, New York.

[12] Gleser, L., and Olkin, I. (1966). A k-sample regression model with covariance. *Multivariate Analysis* (P. R. Krishnaiah, Ed.), pp. 59–72. Academic Press, New York.

[13] Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Statist.* **34** 152–176.

[14] Hall, W. J., Wijsman, R. A., and Ghosh, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[15] Jensen, D. R., and Good, I. J. (1981). Invariant distributions associated with matrix laws under structural symmetry., *J. Roy. Statist. Soc. Ser. B* **43** 327–332.

[16] Kariya, T. (1978). The general MANOVA problem. *Ann. Statist.* **6** 200–213.

[17] Kariya, T. (1981). A robustness property of Hotelling's T^2-test. *Ann. Statist.* **9** 210–214.

[18] Kariya, T. (1981). Robustness of multivariate tests. *Ann. Statist.* **9** 1267–1275.

[19] Kariya, T., and Eaton, M. L. (1978). Robust tests for spherical symmetry. *Ann. Statist.* **5** 206–215.

[20] KHATRI, C. G. (1965). A test for reality of a covariance matrix in a certain complex Gaussian distribution. *Ann. Math. Statist.* **36** 115–119.

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

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