The Annals of Statistics
1980, Vol. 8, No. 6, 1326–1341

UNBIASEDNESS OF INVARIANT TESTS FOR MANOVA AND OTHER MULTIVARIATE PROBLEMS[1]

BY MICHAEL D. PERLMAN[2] AND INGRAM OLKIN[3]

University of Washington and Stanford University

Let $Y: p \times r$ and $Z: p \times n$ be normally distributed random matrices whose $r + n$ columns are mutually independent with common covariance matrix, and $EZ = 0$. It is desired to test $\mu = 0$ vs. $\mu \neq 0$, where $\mu = EY$. Let $d_1, \cdots, d_p$ denote the characteristic roots of $YY'(YY' + ZZ')^{-1}$. It is shown that any test with monotone acceptance region in $d_1, \cdots, d_p$, i.e., a region of the form $\{g(d_1, \cdots, d_p) \leq c\}$ where g is nondecreasing in each argument, is unbiased. Similar results hold for the problems of testing independence of two sets of variates, for the generalized MANOVA (growth curves) model, and for analogous problems involving the complex multivariate normal distribution. A partial monotonicity property of the power functions of such tests is also given.

1. Introduction. It is well-known that the noncentral χ^2- and F-tests have monotone power functions. That is, if $\chi^2_m(\Delta)$ and χ^2_n denote independent χ^2-variates, the former noncentral with noncentrality parameter $\Delta \geq 0$, then

$$P[\chi^2_m(\Delta) > c] \qquad \text{and} \qquad P[\chi^2_m(\Delta)/\chi^2_n > c]$$

each achieves its minimum at $\Delta = 0$ and in fact both are strictly increasing in Δ. We think of these probabilities as the power functions of tests for testing $\Delta = 0$ vs. $\Delta > 0$.

There are (at least) two distinct methods of proving these results: (i) represent $\chi^2_m(\Delta)$ as $[N(\Delta^{1/2}, 1)]^2 + \chi^2_{m-1}$, condition on χ^2_{m-1}, and invoke the symmetry and unimodality of the normal density to deduce that $P[-k \leq N(\Delta^{1/2}, 1) \leq k]$ is decreasing in Δ; (ii) use the facts that the noncentral χ^2 and F densities have strictly monotone likelihood ratios, so that the corresponding distributions are stochastically increasing in Δ. Note also that the unbiasedness and monotonicity of the noncentral F-test either can be proved directly, or deduced from the corresponding properties of the noncentral χ^2-test by conditioning on χ^2_n.

In the multivariate case it is somewhat surprising that no complete analogs of these results have yet been obtained. By using method (i) extended to the multivariate normal distribution, Das Gupta, Anderson, and Mudholkar (1964) obtained partial analogs of these results, as follows.

(a) *The* MANOVA *problem with known covariance matrix.* This is the multivariate version

Received September 1978: revised June 1979.

[1] This research was begun while both authors were visiting professors in the Department of Mathematics, University of British Columbia and was completed while the first author was a visting professor in the Department of Statistics, Stanford University.

[2] This author's research was supported in part by U.S. National Science Foundation Grant No. MCS76-81435, by U.S. Energy Research and Development Administration Contract No. EY-76-S-02-2751, and by the National Research Council of Canada.

[3] This author's research was supported in part by the National Science Foundation Grant No. MPS 75-09450.

AMS 1970 *subject classifications.* Primary 62H10, 62H15; secondary 62H20, 62J05, 62J10.

Key words and phrases. Unbiasedness of invariant multivariate tests, monotonicity of power functions, noncentral Wishart matrix, characteristic roots, maximal invariants, noncentral distributions, hypergeometric function of matrix arguments, stochastically increasing, MANOVA, growth curves model, testing for independence, canonical correlations, complex multivariate normal distribution, FKG inequality, HPKE inequality, positively associated random variables, pairwise total positivity of order two, rectangular coordinates.

of the noncentral χ^2 testing problem. One observes a normally distributed random $p \times r$ matrix X with $EX = \mu$ and whose columns are independent with known common covariance matrix Σ_0, which we take to be the $p \times p$ identity matrix I_p without loss of generality. The problem is to test $\mu = 0$ vs. $\mu \neq 0$. The problem is invariant under orthogonal linear transformations of the form $X \to \Psi X \Gamma'$, where $\Psi: p \times p$ and $\Gamma: r \times r$ are orthogonal matrices. A maximal invariant statistic is $l \equiv (l_1, \cdots, l_t)$, where $t = \min(p, r)$ and $l_1 \geq \cdots \geq l_t > 0$ are the nonzero characteristic roots of XX' (or, more generally, of $XX'\Sigma_0^{-1}$ when $\Sigma_0 \neq I_p$). Any invariant acceptance region $\mathscr{A} \subseteq R^{pr}$ (i.e., $X \in \mathscr{A} \Rightarrow \Psi X \Gamma' \in \mathscr{A}$ for all Ψ, Γ) is equivalent to one of the form $\{h(l) \leq c\}$ for some function h, and the power of the corresponding test depends on μ only through the noncentrality parameters $\lambda \equiv (\lambda_1, \ldots, \lambda_t)$, where $\lambda_1 \geq \cdots \geq \lambda_t \geq 0$ are the nontrivial characteristic roots of $\mu\mu'$. The hypothesis $\mu = 0$ is equivalent to $\lambda_1 = \cdots = \lambda_t = 0$. The method of proof of Das Gupta, Anderson and Mudholkar (1964)—an extension of method (i) to the multivariate case—yields the following result. (In the applications in this paper we consider only *nonrandomized* tests with *nontrivial* acceptance regions $\mathscr{A}$; i.e., both $\mathscr{A}$ and $\mathscr{A}^c$ have positive Lebesgue measure, or equivalently, the significance level α satisfies $0 < \alpha < 1$.)

THEOREM 1.1. *Suppose that the invariant acceptance region $\mathscr{A} \subseteq R^{pr}$ is convex in each column vector of X when the remaining columns are held fixed. Then the power function $P_\mu[X \in \mathscr{A}^c]$ is strictly increasing in each noncentrality parameter λ_j, $1 \leq j \leq t$; in particular, the test is strictly unbiased.*

We remark that our treatment of problem (a) may appear artificial, in the sense that the problem is actually invariant under a larger group, namely the group of all pr-dimensional orthogonal transformations acting on X. We have considered the smaller group so that our discussion of problem (a) is parallel to that of problem (b) below, in that the maximal invariant statistic can be represented as the set of characteristic roots of a random matrix in each problem. In fact, our main results for problem (b) can be deduced from those for problem (a) by a conditioning argument—cf. Theorem 4.1.

(b) *The* MANOVA *problem with unknown covariance matrix.* This problem, the multivariate version of the noncentral F-test problem, is the one actually considered by Das Gupta, Anderson and Mudholkar (1964). Here, in canonical form, one observes the independent, normally distributed random matrices $Y = p \times r$ and $Z: p \times n$ whose columns are mutually independent with common covariance matrix Σ, assumed nonsingular but otherwise unknown, with $EY = \mu$ and $EZ = 0$. The problem of testing $\mu = 0$ vs. $\mu \neq 0$ is invariant under $(Y, Z) \to (AY\Gamma_1', AZ\Gamma_2')$, where $A: p \times p$ is nonsingular and $\Gamma_1: r \times r$ and $\Gamma_2: n \times n$ are orthogonal. If $n + r \leq p$, no nontrivial invariant test exists. Therefore, assume that $n + r > p$, in which case a maximal invariant statistic is $d \equiv (d_{s+1}, \cdots, d_t)$, where $s = \max(p - n, 0)$ and where $1 = d_1 = \cdots = d_s > d_{s+1} > \cdots > d_t > 0$ are the nonzero characteristic roots of $YY'(YY' + ZZ')^{-1}$ [cf. Lehmann (1959), Chapter 7, problems 24 and 25]. Any invariant acceptance region $\mathscr{A} \subseteq R^{p(r+n)}$ is of the form $\{h(d) \leq c\}$ for some h, and the power of the corresponding test depends on μ, Σ only through the noncentrality parameters $\delta \equiv (\delta_1, \cdots, \delta_t)$, where $\delta_1 \geq \cdots \geq \delta_t \geq 0$ are the nontrivial characteristic roots of $\mu\mu'\Sigma^{-1}$. The following result is proved in the same way as Theorem 1.1, conditioning also on Z.

THEOREM 1.2. (Das Gupta, Anderson, and Mudholkar (1964)). *Suppose that the invariant acceptance region $\mathscr{A} \subseteq R^{p(r+n)}$ is convex in each column vector of Y when Z and the remaining column vectors of Y are held fixed. Then the power function $P_{\mu,\Sigma}[(Y, Z) \in \mathscr{A}^c]$ is strictly increasing in each noncentrality parameter δ_j and the test is strictly unbiased.*

REMARK 1.3. Das Gupta, Anderson, and Mudholkar (1964) do not state explicitly that the power function is *strictly* increasing, but a close examination of the proofs of their Theorems 1 and 2 (especially the latter) shows that this is true in our Theorems 1.1 and 1.2. The conclusion can also be reached by a different argument. For the remainder of this remark, we drop the order restrictions $\lambda_1 \geq \cdots \geq \lambda_t$ in problem (a) and $\delta_1 \geq \cdots \geq \delta_t$ in problem (b). Thus the power of any invariant test may be thought of as a symmetric function of $\lambda_1, \cdots, \lambda_t$ or $\delta_1, \cdots, \delta_t$ on R_+^t (see (1.2)). Once it is established that the power function is nondecreasing in $\lambda_1(\delta_1)$, say, when the remaining parameters are fixed, then since the power is an analytic function, it is either strictly increasing or else constant in $\lambda_1(\delta_1)$. A sufficient condition to rule out the latter possibility is that the power approach 1 as $\lambda_1(\delta_1)$ approaches ∞ with the remaining parameters fixed. Since the convex sections of the invariant acceptance regions $\mathscr{A}$ in Theorems 1.1 and 1.2 are symmetric about the origin in R^p and $0 < P[\mathscr{A}] < 1$, these sections must be bounded in at least one direction with positive probability, so this sufficient condition is satisfied. We add that it does not seem to be generally known that this condition fails for the invariant acceptance region $\mathscr{A}_1$ in (1.1) below, an often-advocated test for problem (b) [cf. Anderson and Perlman (1978)].

Theorems 1.1 and 1.2 have been referred to as only partial analogs of the univariate χ^2 and F results for reasons discussed now. In these two univariate cases, the monotone likelihood ratio property implies that the class of admissible tests for testing $\Delta = 0$ vs. $\Delta > 0$ consists of all tests with acceptance regions of the form $\{\chi_m^2(\Delta) \leq c\}$ or $\{\chi_m^2(\Delta)/\chi_n^2 \leq c\}$. Thus, in the univariate cases of problems (a) and (b) (i.e., when $\min(p, r) = 1$), all admissible invariant tests are unbiased and have monotone power functions. In the multivariate cases, however, Theorems 1.1 and 1.2 are not sufficiently broad in scope to yield the corresponding results, i.e., there are admissible invariant acceptance regions to which the assumptions of Theorems 1.1 and 1.2 do not apply. In problem (b) the most striking example is the acceptance region

$$(1.1) \qquad \mathscr{A}_1 = \{(Y, Z) \mid \operatorname{tr} YY'(YY' + ZZ')^{-1} \leq c_1\}$$
$$\equiv \{(d_{s+1}, \cdots, d_t) \mid \textstyle\sum_{i=s+1}^t d_i \leq c_1 - s\}.$$

This invariant acceptance region yields an admissible test [Schwartz (1967b)] which is in fact proper Bayes when $p \leq n$ [Kiefer and Schwartz (1965)], which is locally best invariant [Schwartz (1967a)], and which has certain desirable power and robustness properties [Schatzoff (1966), Pillai and Jayachandran (1967), Fujikoshi (1970), Lee (1971), Olson (1974)]. (For a possibly undesirable property of the power of this test, though, see the final sentence in Remark 1.3.) It has been pointed out in Perlman (1974), however, that $\mathscr{A}_1$ does not satisfy the convexity condition in Theorem 1.2 unless $c_1 \leq \max(1, p - n)$. (When $p \leq n$ this restriction is simply $c_1 \leq 1$, which corresponds to *large* significance levels α, those not usually of interest; when $p > n$ this restriction is equivalent to $c_1 \leq s$, which implies that the test is trivial, i.e., $\alpha = 1$.) This fact has often been overlooked in the literature [cf. Kiefer and Schwartz (1965, page 759), Pillai and Jayachandran (1967, page 209), Schwartz (1967a, pages 348, 357)]. Thus the unbiasedness and monotonicity of the admissible invariant test (1.1) has remained an open question. (As remarked by Schwartz (1967a, page 346), since this test is the essentially unique locally best invariant test, it follows that it is locally strictly unbiased.)

In this paper, *complete* multivariate analogs of the univariate *unbiasedness* results are obtained for the following testing problems: (a) MANOVA with known covariance matrix, (b) MANOVA with unknown covariance matrix, (c) testing for independence of two sets of variates (canonical correlations), and (d) the generalized MANOVA (growth curves) model. Our method is broad in scope and applies to the complex analogs of these problems (see Section 5(e)) as well as to problems for which no unbiasedness results have been obtained previously [see Andersson and Perlman (1979)]. The method, developed in Section 2, is a multivariate extension of method (ii): the noncentral densities of the maximal invariant statistics, (l, d, etc.) are studied directly, rather than the normal distributions of the underlying observations.

By "complete multivariate analogs" we mean the following. For each positive integer q define

$$(1.2) \qquad R_+^q = \{(x_1, \cdots, x_q) \mid x_i > 0, \quad 1 \leq i \leq q\},$$

and

$$G^q = \{g \mid g: R_+^q \to R^1, g \text{ is measurable, nondecreasing in each argument, and } g^{-1}(\{c\}) \neq R_+^q \text{ for each } c \in R^1\}. \tag{1.3}$$

In Section 3 it is shown that for problem (a), if $\lambda \neq 0$ then $E_\lambda g(l) > E_0 g(l)$ for every $g \in G^t$ such that the expectations are not both $+\infty$ or both $-\infty$. This may be restated as follows: the characteristic roots of the noncentral ($\mu \neq 0$) random Wishart matrix XX' are strictly stochastically larger than those of the corresponding central ($\mu = 0$) Wishart matrix. In particular, any nontrivial invariant test with monotone acceptance region in l—i.e., one of the form $\{g(l) \leq c\}$ for some $g \in G^t$—is strictly unbiased for problem (a). Similar results are obtained for problem (b) in Section 4 and for other problems in Section 5. It is easy to see that the classes of invariant tests with monotone acceptance regions in l (for problem (a)) or in d (for problem (b)) are strictly larger than the classes of tests to which Theorems 1.1 and 1.2 apply. Furthermore and most important, by applying the argument of Schwartz (1967b, Theorem 2) it can be shown that a necessary (but not sufficient) condition for admissibility of an invariant acceptance region in problem (a) is that it be monotone in l, with similar necessary conditions for the other problems (see Remark 4.5). Thus, for the problems considered in this paper, our results show that *all nontrivial admissible invariant tests are strictly unbiased.* In particular, test (1.1) is strictly unbiased for problem (b), as are its analogs for problems (c), (d), and (e).

A complete multivariate analog of the univariate *monotonicity* result for problem (a), for example, would be that $E_\lambda g(l)$ is increasing in each λ_j, with similar statements for the other problems. Although we strongly conjecture that this is true and that the same method of proof should apply (cf. Proposition 2.6 (ii) and Remarks 3.2 and 4.4), we have been unable to carry out the details. By yet a third method of proof, however, partial monotonicity results are obtained for alternatives of rank one—see Theorems 3.5, 4.3, and Section 5.

2. Stochastic comparison of multivariate distributions with monotone likelihood ratio. This section concerns applications of the FKG inequality (2.9), due to Fortuin, Ginibre and Kasteleyn (1971), more precisely its extension (2.5) due to Holley (1974), Preston (1974), Kemperman (1977), and Edwards (1978), called the HPKE inequality. An excellent exposition of this subject, including related results in probability theory, appears in Kemperman (1977).

Although the FKG inequality and its extension may be stated for probability distributions defined on a general measure space with a partial ordering, for our applications it suffices to consider probability distributions defined on a measurable rectangle B in R^q (q-dimensional Euclidean space) endowed with the usual component-wise partial ordering, as in Section 6 of Kemperman (1977). For each $i = 1, \cdots, q$, let B_i be a Borel subset of R^1 and ν_i a σ-finite measure on B_i; B denotes the "rectangle" $\prod_1^q B_i$ and $d\nu$ the product measure $\prod_1^q d\nu_i$. For two points $x \equiv (x_1, \cdots, x_q)$ and $y \equiv (y_1, \cdots, y_q)$ in R^q write $x \leq y$ if $x_i \leq y_i$ for each $i = 1, \cdots, q$, and define

$$\begin{aligned} x \wedge y &= (x_1 \wedge y_1, \cdots, x_q \wedge y_q), \\ x \vee y &= (x_1 \vee y_1, \cdots, x_q \vee y_q), \end{aligned} \tag{2.1}$$

where $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$. Let φ (with or without subscripts) denote a probability density on B with respect to ν.

DEFINITION 2.1. The density φ satisfies the FKG *condition* on B if

$$\varphi(x)\varphi(y) \leq \varphi(x \wedge y)\varphi(x \vee y) \tag{2.2}$$

for every $x, y \in B$. The ordered pair (φ_1, φ_2) satisfies the HPKE *condition* on B if

$$\varphi_1(x)\varphi_2(y) \leq \varphi_1(x \wedge y)\varphi_2(x \vee y) \tag{2.3}$$

for every $x, y \in B$.

REMARK 2.2 If (φ_1, φ_2) and $(\varphi_1^*, \varphi_2^*)$ both satisfy the HPKE condition, then so does $(\varphi_1\varphi_1^*, \varphi_2\varphi_2^*)$. If $\varphi_1 = \varphi$ and $\varphi_2 = F\varphi$, where φ satisfies the FKG condition and F is nonnegative and nondecreasing in each argument on $B \cap \{\varphi > 0\}$, then $(\varphi, F\varphi)$ satisfies the HPKE condition. If F and φ both satisfy the FKG condition, then so does their product $F\varphi$. If $\varphi(x) = \prod_1^q \alpha_i(x_i)$ with $\alpha_i \geq 0$ on B_i, then φ trivially satisfies the FKG condition. Finally, suppose that $\gamma: R_+^1 \to R_+^1$ and $\beta_i: B_i^* \to B_i$ $(1 \leq i \leq q)$ are nondecreasing, where B_i^* is also a Borel subset of R^1. If $\varphi(x_1, \cdots, x_q)$ satisfies the FKG condition on $\prod B_i$, then so does $\varphi^*(x) \equiv \gamma(\varphi(\beta_1(x_1), \cdots, \beta_q(x_q)))$ on $\prod B_i^*$.

REMARK 2.3. If φ satisfies the FKG condition, then φ is *pairwise* TP_2 (i.e., totally positive of order 2 in each pair of arguments) as studied by Barlow and Proschan (1975, page 149), and the converse is true when $q = 2$. When $q > 3$, however, the converse need not be true unless $\varphi > 0$ on B [cf. Kemperman (1977, page 329)]. Also see Remark 2.9.

PROPOSITION 2.4. (Holley, Preston, Kemperman, Edwards). (i) *If* (φ_1, φ_2) *satisfies the HPKE condition* (2.3), *then the multivariate distribution determined by* φ_2 *is stochastically larger than that determined by* φ_1. *That is, if h is a measurable function on B such that*

$$\{x \leq y, \varphi_1(x) > 0, \varphi_2(y) > 0\} \Rightarrow h(x) \leq h(y), \tag{2.4}$$

then

$$\int h\varphi_1 \, d\nu \leq \int h\varphi_2 \, d\nu \tag{2.5}$$

provided that the integrals exist.

(ii) *Suppose, in addition, that*

$$\{x \leq y, x \neq y, \varphi_1(x) > 0, \varphi_2(y) > 0\} \Rightarrow h(x) < h(y) \tag{2.6}$$

and that $\varphi_1 \, d\nu$ *and* $\varphi_2 \, d\nu$ *determine distinct probability measures on B. Then*

$$\int h\varphi_1 \, d\nu < \int h\varphi_2 \, d\nu \tag{2.7}$$

provided that the integrals exist and are not both $+\infty$ *or both* $-\infty$.

PROOF: Part (i) is given by Kemperman (1977, Theorem 5) for nonnegative h which are nondecreasing everywhere on B. By a result of Strassen [cf. Kemperman (1977), Remark, page 316], this implies that there exists a probability measure η on $B \times B$ whose marginal distributions are $\varphi_1 \, d\nu$ and $\varphi_2 \, d\nu$ and such that $\eta\{x \leq y\} = 1$; clearly, $\eta\{\varphi_1(x) > 0, \varphi_2(y) > 0\} = 1$ as well. Therefore, when h satisfies the hypotheses of part (i),

$$\int h\varphi_1 \, d\nu = \int\int h(x) \, d\eta(x, y) \leq \int\int h(y) \, d\eta(x, y) = \int h\varphi_2 \, d\nu. \tag{2.8}$$

Under the additional assumptions of part (ii), $\eta\{x \leq y, x \neq y\} > 0$, so strict inequality holds in (2.8).

REMARK 2.5. Suppose that φ satisfies the FKG condition, that g, h are nondecreasing in each argument on $B \cap \{\varphi > 0\}$, and that g, h, gh are integrable with respect to φ. Then

$$\int gh\varphi \, d\nu \geq \left(\int g\varphi \, d\nu\right)\left(\int h\varphi \, d\nu\right), \tag{2.9}$$

which is a version of the FKG *inequality*. For nonnegative g this follows by applying Proposition 2.4(i) with $(\varphi_1, \varphi_2) = (\varphi, g\varphi)$ (see Remark 2.2) and the general case follows by a truncation argument [see Kemperman (1977), Corollary 2]. In the terminology of Barlow and Proschan (1975), (2.9) can be restated as follows: if $x \equiv (x_1, \cdots, x_q)$ is a random vector distributed according to $\varphi \, d\nu$ on B, and if φ satisfies the FKG condition, then $x_1, \cdots, x_q$ are (positively) *associated* random variables. This should be compared to their Corollary 4.15 (page 142) which states that if φ is pairwise TP_2 (see Remark 2.3) then $x_1, \cdots, x_q$ are associated. As pointed out by Kemperman (1977, page 329), however, when $q \geq 3$ this may be false unless $\varphi > 0$ everywhere on the "rectangle" B (also see Remark 2.9(i)).

Proposition 2.4 can be applied in a very general context to deduce unbiasedness and monotonicity of the power functions of tests with monotone acceptance regions. For the applications in this paper, however, it suffices to consider testing problems with a simple null hypothesis, and such that each of the alternative distributions is absolutely continuous with respect to the null distribution. In this case, simple sufficient conditions for unbiasedness and monotonicity can be stated in terms of the null density and the likelihood ratio.

Let $x = (x_1, \cdots, x_q)$ be a random vector distributed according to φ_θ, where $\{\varphi_\theta \mid \theta \in \Theta\}$ is a family of probability densities on the rectangle B (with respect to the product measure $d\nu$) indexed by a vector parameter $\theta \equiv (\theta_1, \cdots, \theta_m) \in \Theta \subset R^m$. Assume that φ is of the form

$$\varphi_\theta(x) = F_\theta(x)\varphi_0(x), \tag{2.10}$$

where $F_\theta \geq 0$ and satisfies $F_0(x) = 1$ for all x. Note that F_θ is simply the likelihood ratio. Consider the problem of testing $\theta = 0$ vs. $\theta \geq 0$, $\theta \neq 0$. (Of course, 0 could be replaced by any other point θ^0; more generally, one could treat the case where θ assumes values in an arbitrary partially ordered set.) A nonrandomized test is said to have a *monotone acceptance region* in x if its acceptance region is of the form $\{g(x) \leq c\}$ for some function g nondecreasing in each argument on $B \cap \{\varphi_0 > 0\}$. A test is called *nontrivial* if its significance level α satisfies $0 < \alpha < 1$. The next result is essentially a generalization of the implication $(C) \Rightarrow (B)$ in Theorem 1 of Lehmann (1955), where it is also assumed that $x_1, \cdots, x_q$ are independent random variables.

PROPOSITION 2.6. (i) (Unbiasedness). *Suppose that φ_0 satisfies the* FKG *condition on B and that $F_\theta(x)$ is nondecreasing in each x_i on $B \cap \{\varphi_0 > 0\}$. Then*

$$\theta \geq 0 \Rightarrow E_\theta g \geq E_0 g \tag{2.11}$$

whenever g is nondecreasing in each argument on $B \cap \{\varphi_0 > 0\}$, provided that the expectations exist. In particular, every test with monotone acceptance region is unbiased for testing $\theta = 0$ vs. *$\theta \geq 0$, $\theta \neq 0$. If, in addition, $F_\theta(x)$ is strictly increasing in each x_i on $B \cap \{\varphi_0 > 0\}$ whenever $\theta \geq 0$, $\theta \neq 0$, and if*

$$P_0\{x \mid g(x) = c\} < 1 \tag{2.12}$$

for every constant c, then

$$\theta > 0, \theta \neq 0 \Rightarrow E_\theta g > E_0 g, \tag{2.13}$$

provided that the expectations exist and are not both $+\infty$ or both $-\infty$. In particular, every nontrivial test with monotone acceptance region is strictly unbiased.

(ii) (Monotonicity). *Assume that for every $\theta \geq 0$, $F_\theta(x) > 0$ on $B \cap \{\varphi_0 > 0\}$. Suppose that φ_0 and every F_θ satisfy the* FKG *condition on B, and that $F_{\theta'}(x)/F_\theta(x)$ is nondecreasing in each x_i on $B \cap \{\varphi_0 > 0\}$ whenever $\theta' \geq \theta \geq 0$. Then*

$$\theta' \geq \theta \geq 0 \Rightarrow E_{\theta'} g \geq E_\theta g \tag{2.14}$$

if g is nondecreasing in each argument on $B \cap \{\varphi_0 > 0\}$ and the expectations exist. If, in addition, $F_{\theta'}(x)/F_\theta(x)$ is strictly increasing in each x_i on $B \cap \{\varphi_0 > 0\}$ whenever $\theta' \geq \theta \geq 0$, $\theta' \neq \theta$, and if g satisfies (2.12), *then*

$$\theta' \geq \theta \geq 0, \theta' \neq \theta \Rightarrow E_{\theta'} g > E_\theta g \tag{2.15}$$

provided that the expectations exist and are not both $+\infty$ or both $-\infty$. In particular, the power function of every nontrivial test with monotone acceptance region is strictly increasing in each θ_j, $1 \leq j \leq m$.

PROOF. In part (i), (2.11) follows by applying Proposition 2.4(i) with $(\varphi_1, \varphi_2) = (\varphi_0, \varphi_\theta)$ $(\equiv(\varphi_0, F_\theta\varphi_0))$ and $h = g$. To prove (2.13), first assume that g is nonnegative and apply Proposition 2.4(ii) with $(\varphi_1, \varphi_2) = (\varphi_0, g\varphi_0/\int g\varphi_0\, d\nu)$ and $h = F_\theta$. For general g, write

$$g = [(g + N) \vee 0] + [g \wedge (-N)] \equiv g_1 + g_2,$$

where N is chosen so large that (2.12) holds with g replaced by g_1. Since g_1 and g_2 are nondecreasing on $B \cap \{\varphi_0 > 0\}$ and g_1 is nonnegative, g_1 satisfies (2.13) and g_2 satisfies (2.11). Part (ii) is proved in a similar manner.

REMARK 2.7. Suppose that for every $\theta \geq 0$, $F_\theta(x) > 0$ for all $x \in B$. Then by Remark 2.3, in order to verify the assumptions on F_θ in the first (second) half of Proposition 2.6(ii) it is sufficient to show that each F_θ is pairwise TP_2 in x and that $F_\theta(x)$ is TP_2 (strictly TP_2) in (x_i, θ_j) for each $i = 1, \cdots, q$ and $j = 1, \cdots, m$. Furthermore, these conditions are equivalent to

$$\frac{\partial^2 \log F_\theta(x)}{\partial x_i \partial x_k} \geq 0, \qquad 1 \leq i < k \leq q, \tag{2.16}$$

and

$$\frac{\partial^2 \log F_\theta(x)}{\partial x_i \partial \theta_j} \geq 0\ (> 0), \qquad 1 \leq i \leq q,\ 1 \leq j \leq m, \tag{2.17}$$

provided that the derivatives exist.

In the remainder of the paper Proposition 2.6 is applied with $B_i = R^1_+$, $B = R^q_+$, $d\nu = dx$ = Lebesgue measure, and

$$\varphi_0(x) = k \cdot \textstyle\prod_{i=1}^q \alpha_i(x_i) \prod_{i<j} [\beta(x_i) - \beta(x_j)]^\gamma_+. \tag{2.18}$$

Here the α_i are nonnegative functions on R^1_+, β is a nondecreasing function on R^1_+, $\gamma \geq 0$, $a_+ \equiv a \vee 0$, and $k > 0$ is a normalizing constant chosen such that $\int \varphi_0\, dx = 1$. Note that

$$\{\varphi_0 > 0\} \subseteq \{x_1 > x_2 > \cdots > x_q > 0\} \equiv R^q_{+0}. \tag{2.19}$$

LEMMA 2.8. *Let* φ_0 *be given by* (2.18). *Then* φ_0 *satisfies the* FKG *condition on* R^q_+.

PROOF. By Remark 2.2 it suffices to show that

$$\varphi(x) \equiv \textstyle\prod_{i<j} (x_i - x_j)_+$$

satisfies the FKG condition (2.2) on R^q_+. Fix $x, y \in R^q_+$. We may assume that both $x, y \in R^q_{+0}$, for if not then the left-hand side of (2.2) is zero. Therefore, it suffices to show that

$$(x_i - x_j)(y_i - y_j) \leq [(x_i \wedge y_i) - (x_j \wedge y_j)][(x_i \vee y_i) - (x_j \vee y_j)] \tag{2.20}$$

whenever $x_i > x_j$ and $y_i > y_j$. If $x_i - y_i$ and $x_j - y_j$ have the same sign or at least one is zero then equality holds in (2.20); if they have opposite signs then the right-hand side of (2.20) exceeds the left-hand side by $(x_i - y_i)(y_j - x_j)$ which is positive. This completes the proof.

Therefore, when φ_θ is of the form (2.10) with φ_0 given by (2.18), in order to apply Proposition 2.6 it is enough to demonstrate the appropriate monotonicity properties of the likelihood ratio F_θ. In each of the applications treated in subsequent sections, F_θ is a hypergeometric function ${}_aF_b$ of two matrix arguments [cf. James (1964)] which appears in the noncentral density of the maximal invariant statistic (l, d, etc.) under consideration, and φ_0 is the central density of this statistic. In each case, φ_0 is of the form (2.18), and ${}_aF_b$ is everywhere strictly positive on R^q_+, so only the monotonicity properties of ${}_aF_b$ need be established.

REMARK 2.9. (i) If $x \equiv (x_1, \cdots, x_q)$ is a random vector with density φ_0 of the form (2.18), then Lemma 2.8 and Remark 2.5 imply that $x_1, \cdots, x_q$ are associated. This conclusion is also stated (for a special case) in the theorem of Dykstra and Hewett (1978, page 236). Their proof, however, relies on the implication (i) $\Rightarrow$ (ii) in Theorem 2 of Dykstra, Hewett, and Thompson

(1973), which is not true in general. A counterexample is provided on page 330 of Kemperman (1977) as follows. Take $q = 3$ and let P, Q be the disjoint cubes $\{1 < x_1 < 2, 0 < x_2 < 1, 0 < x_3 < 1\}$, $\{0 < x_1 < 1, 1 < x_2 < 2, 1 < x_3 < 2\}$, respectively. Let the random vector $x \equiv (x_1, x_2, x_3)$ be uniformly distributed over $P \cup Q$. Then, in the terminology of Dykstra, Hewett, and Thompson (1973), (x_1, x_2, x_3) are positively likelihood ratio dependent, but they are not associated. Their implication (i) $\Rightarrow$ (ii) is true, of course, under the additional assumption that the support of the joint density is a measurable "rectangle" B, for in this case "positive likelihood ratio dependence" reduces to the joint density being pairwise TP_2, and the FKG inequality holds—see Remarks 2.3 and 2.5.

(ii) There is also a slight gap at the top of page 237 of Dykstra and Hewett (1978) concerning the application of Theorem 5.1 of Karlin (1968, page 123) to their function $f(x_1, x_2, x_3)$. The hypothesis of Karlin's theorem requires that the support of f be a measurable "rectangle" B, and Kemperman's counterexample shows that this requirement cannot be removed entirely. However, although $f(x_1, x_2, x_3)$ does not satisfy this requirement, it does satisfy the following condition:

$$\{f(x_1, x_2, x_3) = 0 \quad \text{and} \quad x_1^* < x_1\} \Rightarrow f(x_1^*, x_2, x_3) = 0.$$

With a slight modification, the proof of Karlin's theorem remains valid in this case.

3. The MANOVA problem with known covariance matrix. Proposition 2.6 is now applied to problem (a) of Section 1, with $(x, \theta) = (l, \lambda)$ and $q = m = t$.

THEOREM 3.1. *For every $g \in G^t$,*

$$\lambda \geq 0, \lambda \neq 0 \Rightarrow E_\lambda g(l) > E_0 g(l),$$

provided that the expectations exist and are not both $+\infty$ or both $-\infty$. In particular, every nontrivial invariant test with monotone acceptance region $\{g(l) \leq c\}$ is strictly unbiased for the MANOVA *problem with known covariance matrix.*

PROOF. Since the nonzero characteristic roots of XX' coincide with those of $X'X$, we can assume that $p \leq r$ without loss of generality; so $t = p$. The noncentral density of l is given by

$$\varphi_\lambda(l) = \exp(-\tfrac{1}{2} \operatorname{tr} \Lambda)\, {}_0F_1(\tfrac{1}{2} r; \tfrac{1}{4} \Lambda, L)\varphi_0(l), \tag{3.1}$$

where

$$\varphi_0(l) = k(p, r) \prod_{i=1}^{p} [l_i^{1/2(r-p-1)} e^{-1/2l_i}] \prod_{i<j} (l_i - l_j)_+, \tag{3.2}$$

and $\Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_p)$, $L = \operatorname{diag}(l_1, \cdots, l_p)$ [cf. James (1961, 1964)]. Since, under φ_0, each open set in R^t_{+0} is assigned positive probability, each $g \in G^t$ satisfies (2.12). Furthermore, (3.1) and (3.2) are precisely of the form (2.10) and (2.18), respectively, so Proposition 2.6(i) is applicable if it can be shown that ${}_0F_1(\tfrac{1}{2}r; \tfrac{1}{4}\Lambda, L)$ is strictly increasing in each l_i.

James (1961, Equation (8)) gives the integral representation

$$\begin{aligned} {}_0F_1(\tfrac{1}{2}r; \tfrac{1}{4}\Lambda, L) &= \int_{O(p)} \int_{O(r)} \exp(\operatorname{tr} D'_\lambda \Psi D_l \Gamma')\, d\Psi\, d\Gamma \\ &= \int_{O(p)} \int_{O(r)} \exp(\textstyle\sum_{i=1}^{t} \sum_{j=1}^{t} \lambda_i^{1/2} l_j^{1/2} \psi_{ij}\gamma_{ij})\, d\Psi\, d\Gamma, \end{aligned} \tag{3.3}$$

where $\Psi \in O(p)$ = the group of $p \times p$ orthogonal matrices, $\Gamma \in O(r)$, $d\Psi$ and $d\Gamma$ denote the Haar probability measures on $O(p)$ and $O(r)$, respectively, and the $p \times r$ matrix D_l is defined by

$$\begin{aligned} (D_l)_{ij} &= l_i^{1/2} \quad &\text{if } i = j \\ &= 0 \quad &\text{if } i \neq j, \end{aligned} \tag{3.4}$$

with $D_\lambda: p \times r$ defined similarly. Since the distribution of Ψ is invariant under sign changes of its columns, the identity

$$\tfrac{1}{2}(e^x + e^{-x}) = \textstyle\sum_{k=0}^{\infty} x^{2k}/(2k)! \tag{3.5}$$

can be used to rewrite ${}_0F_1$ in (3.3) as

$$\int_{O(p)} \int_{O(r)} \prod_{j=1}^{t} \left[\sum_{k=0}^{\infty} \frac{1}{(2k)!} l_j^k \left(\sum_{i=1}^{t} \lambda_i^{1/2} \psi_{ij} \gamma_{ij}\right)^{2k} \right] d\Psi \, d\Gamma, \tag{3.6}$$

from which it is immediate that ${}_0F_1$ is strictly increasing in each l_j. This completes the proof.

In the preceding proof, in order to prove that ${}_0F_1$ is increasing in each l_i it is apparently easier to appeal to the zonal polynomial expansion

$${}_0F_1(\tfrac{1}{2} r; \tfrac{1}{4} \Lambda, L) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(\tfrac{1}{4}\Lambda) C_\kappa(L)}{(\tfrac{1}{2} r)_\kappa C_\kappa(I_p) k!}, \tag{3.7}$$

given in James (1964). Here, the real zonal polynomial $C_\kappa(L)$ is a homogeneous symmetric polynomial in $l_1, \cdots, l_p$. It is a (nontrivial) property of these zonal polynomials that their coefficients are all positive [cf. James (1968, Section 8) or Farrell (1976, problem 13.1.13)]. Therefore, ${}_0F_1(\tfrac{1}{2} r; \tfrac{1}{4}\Lambda, L)$ is strictly increasing in each l_i. This argument, however, relies heavily on properties of the real zonal polynomials $C_\kappa(L)$ and (in our opinion) is less satisfactory than the direct argument in the final paragraph of the proof of Theorem 3.1, for three reasons. First, there is pedagogical interest in a direct proof since the theory of zonal polynomials is not simple. Furthermore, it was noted by S. A. Andersson that for some multivariate testing problems other than those considered here, zonal polynomial expansions of the density of the maximal invariant statistic either are not known or else involve new zonal polynomials whose properties are not known. For example, no polynomial expansion has yet been obtained in the problem of testing for the reality of the covariance matrix of a complex multivariate normal distribution, considered by Khatri (1965). Here, however, an integral representation similar to (3.3) is available from which the desired monotonicity can be deduced [cf. Andersson and Perlman (1979)]. Finally, if one attempts to apply Proposition 2.6(ii) to deduce the stronger monotonicity property for the power functions of invariant tests for our problem (a), the zonal polynomial expansion of ${}_0F_1$ seems difficult to work with, whereas the integral representation of (3.3) is more promising. See also Remark 3.2.

REMARK 3.2. In order to apply Proposition 2.6(ii), it suffices to show that

$$\frac{\partial^2 \log {}_0F_1(\tfrac{1}{2} r; \tfrac{1}{4}\Lambda, L)}{\partial l_i \partial l_j} \geq 0, \qquad i \neq j, \tag{3.8}$$

and

$$\frac{\partial^2 \log {}_0F_1(\tfrac{1}{2} r; \tfrac{1}{4}\Lambda, L)}{\partial l_i \partial \lambda_j} > 0, \qquad \text{all } i, j, \tag{3.9}$$

(see Remark 2.7). By symmetry, it suffices to take $i = 1, j = 2$ in (3.8) and $i = j = 1$ in (3.9). If these inequalities are established, then for $g \in G^t$ it follows that $E_\lambda g(l)$ is strictly increasing in each λ_j, provided the expectation is finite, and hence that the power function of any nontrivial invariant test with monotone acceptance region $\{g(l) \leq c\}$ is strictly increasing in each λ_j.

Although we have been unable to prove the general monotonicity property suggested in Remark 3.2, a different argument yields a partial result when $EX \equiv \mu$ is of rank one, i.e., when $\lambda = (\lambda_1, 0, \cdots, 0)$. After a preliminary lemma (Lemma 3.3, which contains an apparently new property of the lower triangular (Cholesky) decomposition of a positive definite matrix), this result is stated in Theorem 3.5.

The following notation is used for Lemma 3.3. Let $\mathscr{T}$ be the collection of all real lower triangular $p \times p$ matrices $T \equiv (t_{ij})$ ($t_{ij} = 0$ if $i < j$) with positive diagonal elements; $\mathscr{T}$ can be identified with an open subset of $R^{p(p+1)/2}$. Let $l_1(T) \geq \cdots \geq l_p(T) > 0$ denote the ordered characteristic roots of TT' and let

$$h_T(x) \equiv \prod_{i=1}^{p} (x - l_i(T)) \equiv \sum_{i=0}^{p} (-1)^i \mathrm{tr}_i(TT')x^{p-i} \tag{3.10}$$

be the characteristic polynomial of TT'. In (3.10),

$$\mathrm{tr}_i(TT') = \mathrm{tr}[(TT')_{[i]}] \tag{3.11}$$

is the ith elementary symmetric function of $l_1(T), \cdots, l_p(T)$, where $(TT')_{[i]}$ denotes the ith compound matrix of TT' [cf. Karlin (1968, page 1)]. By convention, $\mathrm{tr}_0 = 1$. For $T \in \mathscr{T}$ let T_0: $(p-1)\times(p-1)$ be the submatrix obtained by deleting the first row and first column of T. Let $m_1(T_0) \geq \cdots \geq m_{p-1}(T_0)$ denote the ordered characteristic roots of T_0T_0' and let

$$h_{T_0}(x) \equiv \prod_{i=1}^{p-1} (x - m_i(T_0)) \equiv \sum_{i=0}^{p-1} (-1)^i \mathrm{tr}_i(T_0T_0')x^{p-1-i} \tag{3.12}$$

be the characteristic polynomial of T_0T_0'. Since the $m_i \equiv m_i(T_0)$ are also the characteristic roots of $T_0'T_0$, which is a principal submatrix of $T'T$, and since the $l_i \equiv l_i(T)$ are also the characteristic roots of $T'T$, the Poincaré Separation Theorem [cf. Bellman (1970), Theorem 4, page 117] states that

$$l_1 \geq m_1 \geq l_2 \geq m_2 \geq \cdots \geq l_{p-1} \geq m_{p-1} \geq l_p. \tag{3.13}$$

Define

$$\mathscr{T}^* = \{T \in \mathscr{T} \mid l_1(T) > m_1(T_0) > l_2(T) > \cdots > m_{p-1}(T_0) > l_p(T)\}, \tag{3.14}$$

an open subset of $\mathscr{T}$. Finally, for each $T \equiv (t_{ij}) \in \mathscr{T}$ we write $T = (t_{11}, \tilde{T})$, where

$$\tilde{T} = (t_{21}, t_{22}, t_{31}, t_{32}, t_{33}, \cdots, t_{p1}, \cdots, t_{pp}) \in R^{[p(p+1)/2]-1}, \tag{3.15}$$

and define

$$\tilde{\mathscr{T}} = \{\tilde{T} \mid T \in \mathscr{T}\}. \tag{3.16}$$

LEMMA 3.3. (i) *Let $\tilde{T} \in \tilde{\mathscr{T}}$. Then either $(t_{11}, \tilde{T}) \in \mathscr{T}^*$ for all $t_{11} > 0$ or else $(t_{11}, \tilde{T}) \notin \mathscr{T}^*$ for all $t_{11} > 0$.*

(ii) *Define*

$$\tilde{\mathscr{T}}^* = \{\tilde{T} \in \tilde{\mathscr{T}} \mid (t_{11}, \tilde{T}) \in \mathscr{T}^* \quad \text{for all} \quad t_{11} > 0\}. \tag{3.17}$$

Then $\tilde{T} \in \tilde{\mathscr{T}}^* \Rightarrow l_i(t_{11}, \tilde{T}))$ is strictly increasing in t_{11} for each $i = 1, \cdots, p$.

(iii) *The set*

$$\tilde{\mathscr{T}} - \tilde{\mathscr{T}}^* \subset R^{[p(p+1)/2]-1} \tag{3.18}$$

has Lebesgue measure 0.

PROOF. Fix $\tilde{T} \in \tilde{\mathscr{T}}$ and $t_{11}, \hat{t}_{11} \in (0, \infty)$. Let $T = (t_{11}, \tilde{T})$ and $\hat{T} = (\hat{t}_{11}, \tilde{T})$. From the definition of $(TT')_{[i]}$ and the Binet-Cauchy Theorem [cf. Karlin (1968), page 1], one finds that for $1 \leq i \leq p$,

$$\begin{aligned} \mathrm{tr}_i(TT') &= \mathrm{tr}\ T_{[i]}T_{[i]}' \\ &= \text{sum of squares of all } i \times i \text{ minors of } T \\ &= t_{11}^2 \mathrm{tr}_{i-1}(T_0T_0') + Q(\tilde{T}) \end{aligned} \tag{3.19}$$

for some function Q. Since $(\hat{T})_0 = T_0$, it follows from (3.10) and (3.12) that

$$h_T(x) - h_{\hat{T}}(x) = (\hat{t}_{11}^2 - t_{11}^2)h_{T_0}(x). \tag{3.20}$$

Therefore, if $T \equiv (t_{11}, \tilde{T}) \in \mathscr{T}^*$, then also $\hat{T} \equiv (\hat{t}_{11}, \tilde{T}) \in \mathscr{T}^*$, proving (i). Furthermore, if $\hat{t}_{11}$

$\neq t_{11}$, then (3.20) implies that h_T and $h_{\hat{T}}$ cross at $x = m_1(T_0), \cdots, m_{p-1}(T_0)$ and nowhere else. If $\hat{t}_{11} > t_{11} > 0$, it follows that each root of $h_{\hat{T}}$ must be strictly larger than the corresponding root of h_T, i.e., $l_i((\hat{t}_{11}, \tilde{T})) > l_i((t_{11}, \tilde{T}))$ as stated in (ii).

Since $\mathscr{T} = (0, \infty) \times \tilde{\mathscr{T}}$ and $\mathscr{T}^* = (0, \infty) \times \tilde{\mathscr{T}}^*$, in order to prove part (iii) it suffices to show that $\mathscr{T} - \mathscr{T}^* (\subset R^{p(p+1)/2})$ has Lebesgue measure 0. This fact is proved in the same way as the theorem of Okamoto (1973, page 764). Consider the polynomial

$$\pi_T(x) \equiv h_T(x) h_{T_0}(x),$$

where $T \in \mathscr{T}$. The roots of the equation $\pi_T(x) = 0$ are the $2p - 1$ real numbers

$$l_1(T) \geq m_1(T_0) \geq l_2(T) \geq \cdots \geq m_{p-1}(T_0) \geq l_p(T).$$

Let $D(T)$ be the discriminant of $\pi_T(x)$. As pointed out by Okamoto, $D(T)$ is a polynomial in the elements of T, and the roots of π_T are distinct iff $D(T) \neq 0$. Therefore, it suffices to show that $\{T \in \mathscr{T} \mid D(T) = 0\}$ is a Lebesgue-null set. By Okamoto's lemma, however, it suffices to show that $D(T)$ is not identically zero, and for this it is sufficient to show that there exists at least one T such that $D(T) \neq 0$, i.e., such that

$$l_1(T) > m_1(T_0) > l_2(T) > \cdots > m_{p-1}(T_0) > l_p(T).$$

Such a T is easily constructed: for example, let $t_{ii} = i$ and $t_{i1} = 1$, $1 \leq i \leq p$, and set $t_{ij} = 0$ otherwise.

REMARK 3.4. Lemma 3.3 (ii) states that each characteristic root of TT' is an increasing function of t_{11}. This result does *not* follow from the Courant-Fischer Minmax theorem [Bellman (1970), Theorem 3, page 117], since TT' is not increasing in t_{11} with respect to the Loewner ordering ($\Sigma_1 \geq \Sigma_2$ iff $\Sigma_1 - \Sigma_2$ is positive semidefinite).

THEOREM 3.5. *Suppose that* $\lambda = (\lambda_1, 0, \cdots, 0)$. *Then for* $g \in G^t$, $E_{(\lambda_1, 0, \cdots, 0)} g(l)$ *is strictly increasing in* λ_1, *provided that the expectation is finite for all* λ_1; *in particular, the power function of every nontrivial invariant test with monotone acceptance region* $\{g(l) \leq c\}$ *is strictly increasing in* λ_1.

PROOF. The well-known representation of a central Wishart matrix in terms of independent "rectangular coordinates" carries over to the noncentral Wishart matrix XX' when rank $(\mu) = 1$ [cf. Farrell (1976), Section 11.4]. Again assume that $p \leq r$ without loss of generality; furthermore, to study the distribution of l, by invariance it can be assumed that $\mu = D_\lambda$. Then XX' has the same distribution as TT', where $T{:}p \times p$ is now a *random* lower triangular matrix whose elements $T_{ij} (i \geq j)$ are mutually independent with the following distributions:

$$\begin{aligned} T_{11} &\sim [\chi_r^2(\lambda_1)]^{1/2}, & \\ T_{ii} &\sim [\chi_{r-i+1}^2]^{1/2}, & 2 \leq i \leq p \\ T_{ij} &\sim N(0, 1), & i > j. \end{aligned} \tag{3.21}$$

Since T_{11} is stochastically increasing in λ_1 and since the characteristic roots $(l_1, \cdots, l_p)$ of XX' have the same joint distribution as those of TT' Lemma 3.3 implies that $E_{(\lambda_1, 0, \cdots, 0)} g(l)$ is nondecreasing in λ_1.

To see that this expectation is strictly increasing, two additional facts are required whose proofs are straightforward but lengthy, hence omitted. First each $g \in G^p$ has a point of strict increase on R^p_{+0} (see 2.19)), i.e., there exist $x_1^* > \cdots > x_p^* > 0$ such that $x_i' < x_i^* < x_i''$, $1 \leq i \leq p$, implies that $g(x_1', \cdots, x_p') < g(x_1'', \cdots, x_p'')$. Second, given $x_1^* > \cdots > x_p^* > 0$ there exists $T^* \in \mathscr{T}^*$ such that $l_i(T^*) = x_i^*$ for each $i = 1, \cdots, p$. From these two facts and the continuity of the mapping

$$T \to (l_1(T), \cdots, l_p(T)),$$

together with Lemma 3.3, it can be deduced that the set

$$\{\tilde{T} \in \tilde{\mathscr{T}}^* \mid g(l_1((t_{11}, \tilde{T})), \cdots, l_p((t_{11}, \tilde{T}))) \text{ is nondecreasing and somewhere strictly increasing in } t_{11}\}$$

contains a set of positive Lebesgue measure in $R^{[p(p+1)/2]-1}$. Since T_{11} is *strictly* stochastically increasing in λ_1, this implies that $E_{(\lambda_1, 0, \cdots, 0)}g(l)$ is *strictly* increasing in λ_1, and the proof is complete.

If rank $(\mu) \geq 2$, i.e., if $\lambda_1, \lambda_2 > 0$, then the method used to prove Theorem 3.5 does not apply. The reason for this is that in the lower triangular decomposition $XX' = TT'$, the (random) elements of T are no longer independent, and the distributions of the off-diagonal elements, as well as those of the diagonal elements, can depend on the noncentrality parameters λ_j. We have considered the case $p = 2$ in detail, for in this case one has explicit expressions for the two characteristic roots l_1, l_2 of TT' in terms of the elements of T. The method of Theorem 3.5 can be applied to show that l_1 is strictly stochastically increasing in λ_1 and λ_2 (which is also a consequence of Theorem 1.1), and also that l_2 is strictly stochastically increasing in λ_1 and λ_2 (which seems to be a new result), but we have been unable to show that $g(l_1, l_2)$ is stochastically increasing in λ_1 and λ_2 for an arbitrary function $g \in G^2$.

4. The MANOVA problem with unknown covariance matrix. Here we consider problem (b) of Section 1. Just as the unbiasedness and monotonicity results for the noncentral F-test can be derived from the corresponding results for the noncentral χ^2-test by a conditional argument, so can the results for problem (b) be deduced from those of problem (a) by a conditional and invariance argument.

THEOREM 4.1. *Let λ^*, λ^{**} be two fixed values of λ. Suppose it has been shown in problem* (a) *that $E_{\lambda^{**}}g(l) > E_{\lambda^*}g(l)$ for every $g \in G^t$. Then in problem* (b),

$$E_{\delta=\lambda^{**}}g(d) > E_{\delta=\lambda^*}g(d) \tag{4.1}$$

for every $g \in G^{t-s}$.

PROOF. Recall that $d \equiv (d_{s+1}, \cdots d_t)$ are the nontrivial characteristic roots of $YY'(YY' + ZZ')^{-1}$. To study the distribution of d, it can be assumed that $\Sigma = I$, so that $\delta \equiv (\delta_1, \cdots, \delta_t)$ are the nontrivial characteristic roots of $\mu\mu'$. Thus if $\hat{l} \equiv (\hat{l}_1, \cdots, \hat{l}_t)$ denotes the set of nonzero characteristic roots of YY', with $\hat{l}_1 > \cdots > \hat{l}_t > 0$, then $\hat{l}$ has the same distribution as the characteristic roots $l \equiv (l_1, \cdots, l_t)$ of XX' in problem (a), but with λ replaced by δ. Consider the spectral decomposition

$$YY' = RD_{\hat{l}}R',$$

where $R: p \times p$ is a (random) orthogonal matrix and $D_l = \operatorname{diag}(\hat{l}_1, \cdots, \hat{l}_t, 0, \cdots, 0): p \times p$. Since the characteristic roots of $YY'(YY' + ZZ')^{-1}$ coincide with those of $D_{\hat{l}}(D_{\hat{l}} + R'ZZ'R)^{-1}$, and since Z and $R'Z$ are identically distributed for each fixed R, it follows that d has the same distribution as $\hat{d} \equiv \hat{d}(\hat{l}, Z) \equiv (\hat{d}_{s+1}, \cdots, \hat{d}_t)$, where $1 > \hat{d}_{s+1} > \cdots > \hat{d}_t > 0$ are the nontrivial characteristic roots of $D_{\hat{l}}(D_{\hat{l}} + ZZ')^{-1}$. It is easy to see that each $\hat{d}_j \equiv \hat{d}_j(\hat{l}, Z)$ is a nondecreasing function of each $\hat{l}_i$ for fixed Z. For any $g \in G^{t-s}$ and any fixed value of Z, define $\hat{g}_Z$ on R_+^t by

$$\hat{g}_Z(\hat{l}) = g(\hat{d}(\hat{l}, Z)).$$

Then $\hat{g}_Z$ is nondecreasing in each $\hat{l}_i$, and furthermore it can be shown that $\hat{g}_Z$ is nonconstant on R_+^t, i.e., $\hat{g}_Z^{-1}(\{c\}) \neq R_+^t$ for each real number c, provided that Z is of full rank. Thus, for such Z, $\hat{g}_Z \in G^t$. Since Z is of full rank with probability one, it follows from the hypothesis that

$$
\begin{aligned}
E_{\delta=\lambda^{**}}g(\hat{d}) &= E\{E_{\lambda^{**}}[\hat{g}_Z(\hat{l})\,|\,Z)\} \\
&> E\{E_{\lambda^{*}}[\hat{g}_Z(\hat{l})\,|\,Z]\} \qquad (4.2)\\
&= E_{\delta=\lambda^{*}}g(\hat{d}),
\end{aligned}
$$

which is equivalent to (4.1).

By Theorem 4.1, the results of Theorems 3.1 and 3.5 immediately carry over to problem (b):

THEOREM 4.2. *For every* $g \in G^{t-s}$,

$$\delta \geq 0,\ \delta \neq 0 \Rightarrow E_\delta g(d) > E_0 g(d).$$

In particular, every nontrivial invariant test with monotone acceptance region $\{g(d) \leq c\}$ *is strictly unbiased for the* MANOVA *problem with unknown covariance matrix.*

THEOREM 4.3. *Suppose that* $\delta = (\delta_1, 0, \cdots, 0)$. *Then for* $g \in G^{t-s}$, $E_{(\delta_1,0,\ldots,0)}g(d)$ *is strictly increasing in* λ_1, *provided that the expectation is finite for all* λ_1; *in particular, the power function of every nontrivial invariant test with monotone acceptance region* $\{g(d) \leq c\}$ *is strictly increasing in* δ_1.

REMARK 4.4. If the monotonicity property conjectured for problem (a) in Remark 3.2 is true, then Theorem 4.1 will also imply that the corresponding result is true for problem (b), namely that $E_\delta g(d)$ is strictly increasing in each δ_j.

For the reason mentioned in Remark 4.5 below, it is important to point out that Theorem 4.2 can be proved by a direct application of Proposition 2.6(i), as was Theorem 3.1. Consider first the case where $p \leq r$ and $p \leq n$, so that $s = 0$ and $t = p$. In this case the joint density of $d \equiv (d_1, \cdots, d_p)$ is given by [cf. James (1964), page 486]

$$\varphi_\delta(d) = \exp(-\tfrac{1}{2}\operatorname{tr}\Delta)\,{}_1F_1(\tfrac{1}{2}(r+n);\ \tfrac{1}{2}r;\ \tfrac{1}{2}\Delta, D)\varphi_0(d), \qquad (4.3)$$

where $\Delta = \operatorname{diag}(\delta_1, \cdots, \delta_p)$, $D = \operatorname{diag}(d_1, \cdots, d_p)$, and $\varphi_0(d)$ is of the form (2.18). There are again two ways to show that the hypergeometric function ${}_1F_1$ is strictly increasing in each d_i: either appeal to the integral representation of ${}_1F_1$, or to its zonal polynomial expansion. The former, given by Schwartz (1967a, page 344, Equation (6)), is (omitting the arguments of ${}_1F_1$ and changing to our notation)

$${}_1F_1 = \int_{G(p)} |TT'|^{(r+n)/2}\exp(-\tfrac{1}{2}\operatorname{tr}TT')\left\{\int_{O(p)}\int_{O(r)} \exp(\operatorname{tr}\mu' T\Psi\bar{T}Y\Gamma')\,d\Psi\,d\Gamma\right\} dT, \qquad (4.4)$$

where Ψ and Γ are as in (3.3), $G(p)$ is the group of all $p \times p$ nonsingular lower triangular matrices, dT denotes a (left) Haar measure on $G(p)$, and $\bar{T} \in G(p)$ satisfies $\bar{T}(YY' + ZZ')\bar{T}' = I_p$. Since $d\Psi$ and $d\Gamma$ are invariant under orthogonal transformations on the left and right, the term in brackets in (4.4) can be rewritten as

$$\int_{O(p)}\int_{O(r)} \exp(\operatorname{tr} D'_{\tilde{\delta}}\Psi D_d\Gamma')\,d\Psi\,d\Gamma, \qquad (4.5)$$

where D_d and $D_{\tilde{\delta}}$ are defined as in (3.5), with $\tilde{\delta}_1 \geq \cdots \geq \tilde{\delta}_p \geq 0$ the characteristic roots of $\mu\mu' TT'$. Since (4.5) is precisely of the form (3.3), the argument given in the proof of Theorem 3.1 shows that for each fixed T, (4.5) is strictly increasing in each d_i, hence so is (4.4).

Entirely similar arguments are valid for the case where $p \geq r$, $p \leq n$. When $p > n$ (but still $p < n + r$), however, the zonal polynomial expansion is not available. Nonetheless, the right-hand side of (4.4) remains a valid representation of the likelihood ratio $\varphi_\delta(d)/\varphi_0(d)$ of the maximal invariant statistic d, and the subsequent argument showing that this ratio is strictly increasing in each d_i still holds. (Note that now $d_1 = \cdots = d_s$ are identically 1, only $d_{s+1}, \cdots, d_t$ are varied.) Furthermore, although the central density φ_0 of $d \equiv (d_{s+1}, \cdots, d_t)$ apparently has not appeared in the published literature, it has been derived by Olkin (1951) and is again of the form (2.18). Thus, Proposition 2.6(i) is applicable in all cases.

REMARK 4.5. The fact that the likelihood ratio $\varphi_\delta(d)/\varphi_0(d)$ is strictly increasing in each d_i, $s + 1 \le i \le t$, is stated without proof in Schwartz (1967b, Theorem 2), where it is used to show that for the MANOVA problem with unknown covariance matrix (problem (b)), a necessary condition for admissibility of an invariant acceptance region is that it be monotone in these d_i. We have now given two arguments for the strict monotonicity of the likelihood ratio $\varphi_\delta(d)/\varphi_0(d)$ for problem (b), as well as for that of $\varphi_\lambda(l)/\varphi_0(l)$ for problem (a) in Section 3, and similar arguments yield the strict monotonicity of the likelihood ratios of the maximal invariant statistics for the problems treated in Section 5. For each of these problems, therefore, a similar necessary condition for admissibility of an invariant acceptance region obtains.

5. Other multivariate testing problems.

(c) *Testing for independence of two sets of variates (canonical correlations).* The reader is referred to Anderson and Das Gupta (1964a) for a description of this problem. By conditioning on one of the two sets of variates, Anderson and Das Gupta show that this problem reduces to the MANOVA problem (b), and thereby obtain the analog of Theorem 1.2 for the present problem [also see Perlman (1974)]. Exactly the same conditional argument can be used to obtain the analogs of Theorems 4.2 and 4.3 for problem (c). The analog of Theorem 4.2 for problem (c), for example, states that when the population canonical correlations are not all zero, then the sample canonical correlations are strictly stochastically larger than when the population correlations are all zero, and hence any nontrivial invariant acceptance region which is monotone in the sample canonical correlations yields a test for independence which is strictly unbiased. A remark analogous to Remark 4.4 also holds. Furthermore, the analog of Theorem 4.2 for problem (c) also can be proved directly from Proposition 2.6(i) by demonstrating the required monotonicity property of the hypergeometric function ${}_2F_1$ of two matrix arguments [cf. James (1964), Equation (76)]. As before, this monotonicity property of ${}_2F_1$ can be derived in two ways, from either the integral representation or the zonal polynomial expansion, and this property in turn yields a necessary condition for admissibility of invariant tests as discussed in Remark 4.5.

(d) *The generalized* MANOVA *(growth curves) model.* This problem is discussed by Fujikoshi (1973) and Kariya (1978); also see the references therein. Fujikoshi presents a conditional argument which reduces this problem to the MANOVA problem (b), and thereby obtains the analog of Theorem 1.2 for a class of invariant tests for the present problem. Again, the same conditional argument can be applied to obtain the analogs of Theorem 4.2, Theorem 4.3, and Remarks 4.4 and 4.5 for problem (d). Because (at least) two different groups of transformations have been considered for this problem [cf. Kariya (1978)], it should be pointed out that we are considering (in Fujikoshi's terminology) the class of invariant tests based on the characteristic roots of $S_h S_e^{-1}$. By the analog of Theorem 4.2 for problem (d), for example, we mean that any nontrivial invariant acceptance region which is monotone in these roots, yields a strictly unbiased test.

(e) *Testing problems involving the complex multivariate normal distribution.* Pillai and Li (1970) consider the analogs of testing problems (b) and (c) for the complex multivariate normal and complex Wishart distributions. They have shown that Theorem 1.2 and the corresponding result in Anderson and Das Gupta (1964) extend to the complex versions of problems (b) and (c). Similarly, Theorem 1.1 extends to the complex version of problem (a), whereas Fujikoshi's result extends to the complex version of problem (d). Here we point out that all of *our* results and remarks for problems (a)–(d) extend to the complex versions of these problems, using parallel arguments. For example, consider the complex versions of Theorem 3.1 for (a), of Theorem 4.2 for (b), and of the corresponding result for problem (c). The noncentral densities of the maximal invariant statistics for the complex versions of problems (a), (b), and (c) are given by James (1968, Section 8). In each case the density is of the form (2.10), where φ_0 is of the form (2.18) where F_θ is a *complex* hypergeometric function ${}_a\tilde{F}_b$ of two matrix arguments, as defined by James. Thus, Proposition 2.6(i) is applicable if it is shown that ${}_a\tilde{F}_b$ is strictly increasing in each argument. Again, there are two ways to demonstrate this: first, by means of integral representations of ${}_a\tilde{F}_b$ (now the integrals extend over groups of complex unitary transformations, rather than over real orthogonal transformations); second, by means of the expansion of ${}_a\tilde{F}_b$ in terms of the *complex* zonal polynomials C_κ given by James. To carry out the second approach one needs the fact that, as in the real case, the complex zonal polynomials $\tilde{C}_\kappa$ have positive coefficients. This fact has been established (only recently) by Farrell (1980).

Acknowledgment. The authors have benefitted from helpful discussions with R. Farrell, A. W. Marshall, R. Muirhead, L. Rosen, R. Schwartz, and J. V. Zidek on many of the topics in this paper.

REFERENCES

[1] ANDERSON, T. W. and DAS GUPTA, S. (1964). Monotonicity of the power functions of some tests of independence between two sets of variates. *Ann. Math. Statist.* **35** 206–208.
[2] ANDERSON, T. W. and PERLMAN, M. D. (1978). On a consistency property of invariant multivariate tests. Unpublished manuscript.
[3] ANDERSSON, S. A. and PERLMAN, M. D. (1979). Two testing problems relating the real and complex multivariate normal distributions. Unpublished manuscript.
[4] BARLOW, R. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart, and Winston, New York.
[5] BELLMAN, R. (1970). *Introduction to Matrix Analysis*, 2nd Ed. McGraw-Hill, New York.
[6] DAS GUPTA, S., ANDERSON, T. W., and MUDHOLKAR, G. S. (1964). Monotonicity of the power functions of some tests of the multivariate linear hypothesis. *Ann. Math. Statist.* **35** 200–205.
[7] DYKSTRA, R. L. and HEWETT, J. E. (1978). Positive dependence of the roots of a Wishart matrix. *Ann. Statist.* **6** 235–238.
[8] DYKSTRA, R. L., HEWETT, J. E., and THOMPSON, W. A. (1973). Events which are almost independent. *Ann. Statist.* **1** 674–681.
[9] EDWARDS, D. A. (1978). On the Holley-Preston inequalities. *Proc. Royal Soc. Edinburgh*, **78A**, 265–272.
[10] FARRELL, R. (1976). Techniques of Multivariate Calculation. *Lecture Notes in Math.* **520** Springer, Berlin.
[11] FARRELL, R. (1980). Calculation of complex zonal polynomials. In *Multivariate Analysis*-V: *Proceedings Fifth Internat. Conference Multivariate Analysis.* (P. R. Krishnaiah, ed.) 301–320. North Holland, Amsterdam.
[12] FORTUIN, C. M., GINIBRE, J. and KASTELEYN, P. W. (1971). Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* **22** 89–103.
[13] FUJIKOSHI, Y. (1970). Asymptotic expansions of the distributions of test statistics in multivariate analysis. *J. Sci. Hiroshima Univ. A*-1 **34** 73–144.
[14] FUJIKOSHI, Y. (1973). Monotonicity of the power functions of some tests in general MANOVA models. *Ann. Statist.* **1** 388–391.
[15] HOLLEY, R. (1974). Remarks on the FKG inequalities. *Commun. Math. Phys.* **36** 227–231.
[16] JAMES, A. T. (1961). The distribution of noncentral means with known covariance. *Ann. Math. Statist.* **32** 874–882.
[17] JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475-501.
[18] JAMES, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann. Math. Statist.* **34** 1711–1718.
[19] KARIYA, T. (1978). The general MANOVA problem. *Ann. Statist.* **6** 200–214.
[20] KARLIN, S. (1968). *Total Positivity.* Stanford University Press, Stanford, California.
[21] KEMPERMAN, J. H. B. (1977). On the FKG-inequality for measures on a partially ordered space. *Indag. Math.* **39** 313–331.
[22] KHATRI, C. G. (1965). A test for reality of a covariance matrix in a certain complex Gaussian distribution. *Ann. Math. Statist.* **36** 98–114.
[23] KIEFER, J. and SCHWARTZ, R. (1965). Admissible Bayes character of T^2-, R^2-, and other fully invariant tests for classical multivariate normal problems. *Ann. Math. Statist.* **36** 747–770.
[24] LEE, Y. S. (1971). Asymptotic formulae for the distribution of a multivariate test statistic; power comparisons of certain multivariate tests. *Biometrika* **58** 647–651.
[25] LEHMANN, E. L. (1955). Ordered families of distributions. *Ann. Math. Statist.* **26** 399–419.

[26] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.
[27] OKAMOTO, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* **1** 763–765.
[28] OLKIN, I. (1951). On distribution problems in multivariate analysis. Doctoral dissertation, Univ. of North Carolina, Chapel Hill.
[29] OLSON, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *J. Amer. Statist. Assoc.* **69** 894–908.
[30] PERLMAN, M. D. (1974). On the monotonicity of the power functions of tests based on traces of multivariate beta matrices. *J. Multivariate Anal.* **4** 22–30.
[31] PILLAI, K. C. S. and JAYACHANDRAN, K. (1967). Power comparisons of tests of two multivariate hypotheses based on four criteria. *Biometrika* **54** 195–210.
[32] PILLAI, K. C. S. and LI, H. C. (1970). Monotonicity of the power functions of some tests of hypotheses concerning multivariate complex normal distributions. *Ann. Inst. Statist. Math.* **22** 307–318.
[33] PRESTON, C. J. (1974). A generalization of the FKG inequalities. *Commun. Math. Phys.* **36** 233–341.
[34] SCHATZOFF, M. (1966). Sensitivity comparisons among tests of the general linear hypothesis. *J. Amer. Statist. Assoc.* **61** 415–435.
[35] SCHWARTZ, R. (1967a). Locally minimax tests. *Ann. Math. Statist.* **38** 340–359.
[36] SCHWARTZ, R. (1967b). Admissible tests in multivariate analysis of variance. *Ann. Math. Statist.* **38** 698–710.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98105

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305