The Annals of Mathematical Statistics
1969, Vol. 40, No. 2, 549–567

ONE-SIDED TESTING PROBLEMS IN MULTIVARIATE ANALYSIS[1]

BY MICHAEL D. PERLMAN[2]

University of Cambridge

1. Introduction. Suppose one obtains N independent observations from a p-dimensional normal distribution with mean μ and covariance matrix Σ. Under either of the assumptions (a) Σ is known or (b) $\Sigma = \sigma^2\Sigma_0$ with σ^2 unknown and Σ_0 known, the problem of testing $H:\mu = 0$ against the restricted alternative $K:\mu_i \geqq 0$, $i = 1, \cdots, p$, (with at least one inequality strict) has been studied extensively in the past ten years. Bartholomew, Chacko, Kudô, Nüesch, and Shorack have derived the likelihood ratio tests (LRT) and their null distributions and have studied their power functions. Computations of Bartholomew (1961a) and Nüesch (1964) show that the LRT's have substantially higher power than the usual χ^2 or F tests used for testing $\mu = 0$ against the unrestricted alternative $\mu \neq 0$. Abelson and Tukey have proposed simple tests based on the best linear contrast, and their idea has been extended by Schaafsma and Smid. Bartholomew's computations show that these tests are also substantially better than the usual tests, but neither the LRT nor the Abelson-Tukey test is uniformly more powerful than the other.

In this paper we study the above and related testing problems with restricted hypotheses or alternatives, *under the assumption that Σ is completely unknown.* Two procedures are considered: the LRT's and a family of tests based on the notions of Schaafsma and Smid. In Section 5 the LRT is derived for the general problem of testing $H:\mu \in \mathcal{P}_1$ vs. $K:\mu \in \mathcal{P}_2$ where $\mathcal{P}_1$ and $\mathcal{P}_2$ are positively homogeneous sets with $\mathcal{P}_1 \subset \mathcal{P}_2$, and it is shown that the power of the LRT approaches one as the distance from the hypothesis H becomes large. In Section 7 the exact null distribution of the LRT is obtained for the special case where $\mathcal{P}_1 = \{0\}$ and $\mathcal{P}_2 = \{\mu:\mu_i \geqq 0, i = 1, \cdots, p\}$. Since this distribution depends on the unknown matrix Σ, this result in itself cannot be used to obtain the level α rejection region of the LRT. In Section 6, however, sharp upper and lower bounds (as Σ varies) on the null distribution of the LRT statistic are derived for the more general case of a one-sided alternative, where $\mathcal{P}_1 = \{0\}$ and $\mathcal{P}_2 = \mathcal{C}$, a cone (see Section 2 for definitions). These bounds are independent of the particular cone $\mathcal{C}$, and the upper bound provides a simple formula for the level α cutoff point of the LRT. Similar results are given in Section 8 for the problem of testing a one-sided hypothesis against unrestricted alternatives, where $\mathcal{P}_1 = \mathcal{C}$, a cone, and $\mathcal{P}_2$ is the entire space. Cases where only a subset of the components of μ are tested are also discussed (Sections 6, 7, 8).

Received 19 June 1968.

[1] Based in part on the author's doctoral dissertation, written at Stanford University under the support of the National Science Foundation. The present paper was written while the author was a National Science Foundation Postdoctoral Fellow.

[2] Now at the University of Minnesota, Minneapolis.

In Section 9 the family of somewhere most powerful similar tests, defined by Schaafsma and Smid (1966), is investigated for the case $\mathcal{P}_1 = \{0\}$ and $\mathcal{P}_2 = \mathcal{C}$, a cone. It is shown that for this problem all such tests have low power at some alternative points which are arbitrarily far from the null hypothesis $\mu = 0$, whereas the power of the LRT approaches one uniformly as the distance from the null hypothesis becomes large. This contrasts with the situation under assumption (a) (respectively (b)), where there exists a most stringent somewhere most powerful (respectively similar) test whose power compares favorably with that of the corresponding LRT.

Sections 2, 3, and 4 contain definitions and preliminary lemmas. Attention is called to Section 3, where some basic facts in multivariate normal distribution theory are presented. In particular, Lemma 3.2 is a key result for the distribution theory in Sections 6, 7, and 8 and, it is felt, is an important result in its own right.

2. Definitions and notation. $\mathcal{E}_p$ denotes p-dimensional Euclidean space. For a point x in $\mathcal{E}_p$ we write $x \geqq 0$ $(x > 0)$ to indicate that each component of x is non-negative (positive). The closed positive orthant is denoted by $\mathcal{O}$, i.e., $\mathcal{O} = \{x : x \geqq 0\}$. The symbol $\mathcal{H}$ is used to denote any $(p-1)$-dimensional subspace in $\mathcal{E}_p$. Associated with any such subspace are two closed halfspaces whose intersection is $\mathcal{H}$; we denote either of these halfspaces by $\mathcal{H}^+$. The symbol $\mathcal{L}$ denotes any semi-infinite halfline, or ray, emanating from the origin. For any set $\mathcal{A}$ and any linear transformation $A : \mathcal{E}_p \to \mathcal{E}_p$, $A(\mathcal{A})$ is the image of $\mathcal{A}$ under A.

A set $\mathcal{P}$ in $\mathcal{E}_p$ is *positively homogeneous* if $x \in \mathcal{P}$ implies $cx \in \mathcal{P}$ for all positive real numbers c. All positively homogeneous sets to be considered are assumed to be closed sets as well. A set $\mathcal{A}$ containing at least one non-zero point is *one-sided* (with respect to the origin) if there exists a non-zero point z such that $a'z > 0$ for all non-zero $a \in \mathcal{A}$. A closed, positively homogeneous, one-sided set $\mathcal{C}$ is called a *cone*. Note that the properties defined here are preserved under linear transformations.

Examples of cones include: any halfline $\mathcal{L}$; the positive orthant $\mathcal{O}$; any right circular cone $\mathcal{C}_\lambda$, $0 < \lambda \leqq 1$, defined by $\mathcal{C}_\lambda = \{x : x'z/(x'xz'z)^{\frac{1}{2}} \geqq \lambda\}$, where z is an arbitrary nonzero point.

2.1°. For any cone $\mathcal{C}$ there exist a halfline $\mathcal{L}$ and a half-space $\mathcal{H}^+$ such that $\mathcal{L} \subset \mathcal{C} \subset \mathcal{H}^+$. $\mathcal{H}^+$ may be chosen such that $\mathcal{C} \cap \mathcal{H}^+ = \{0\}$.

2.2°. For any cone $\mathcal{C}$ there exists a right circular cone $\mathcal{C}_\lambda$, $0 < \lambda < 1$, such that $\mathcal{C} \subset \mathcal{C}_\lambda$.

2.3°. If the cone $\mathcal{C}$ contains an open set, there exists a right circular cone $\mathcal{C}_\nu$, $0 < \nu < 1$, such that $\mathcal{C}_\nu \subset \mathcal{C}$.

2.4°. For any right circular cone $\mathcal{C}_\lambda$ with $0 < \lambda < 1$, there exist a halfline $\mathcal{L}$, a halfspace $\mathcal{H}^+$, and sequences $\{A_n\}$, $\{B_n\}$ of non-singular linear transformations such that

$$A_n(\mathcal{C}_\lambda) \supset A_{n+1}(\mathcal{C}_\lambda), \qquad \bigcap_{n=1}^{\infty} A_n(\mathcal{C}_\lambda) = \mathcal{L},$$

$$B_n(\mathcal{C}_\lambda) \subset B_{n+1}(\mathcal{C}_\lambda), \qquad \bigcup_{n=1}^{\infty} B_n(\mathcal{C}_\lambda) = \text{interior } (\mathcal{H}^+).$$

If V is a square matrix, write $V > 0$ to denote that V is positive definite symmetric. If $V > 0$, the unique positive definite symmetric square root of V is denoted by $V^{\frac{1}{2}}$.

Partitioning V as $V = (V_{ij})$ $(i, j = 1, \cdots, k)$ with each V_{ii} a square matrix, we set $V_{ii \cdot j} = V_{ii} - V_{ij}V_{jj}^{-1}V_{ji}$.

Let $Z:p \times 1$ be partitioned as $Z' = (Z_1', Z_2')$ with $Z_1:q \times 1$ and let $V:p \times p$ be partitioned as $V = (V_{ij})$ $(i, j = 1, 2)$ with $V_{11}:q \times q$. Assume $V > 0$. Then

$$(2.1) \quad \begin{aligned} Z'V^{-1}Z &= (Z_1 - V_{12}V_{22}^{-1}Z_2)'V_{11\cdot 2}^{-1}(Z_1 - V_{12}V_{22}^{-1}Z_2) + Z_2'V_{22}^{-1}Z_2 \\ &= (Z_2 - V_{21}V_{11}^{-1}Z_1)'V_{22\cdot 1}^{-1}(Z_2 - V_{21}V_{11}^{-1}Z_1) + Z_1'V_{11}^{-1}Z_1 . \end{aligned}$$

The multivariate normal distribution with mean μ and positive definite covariance matrix Σ is denoted by $\mathfrak{N}(\mu, \Sigma)$. The Wishart distribution with n degrees of freedom and expectation $n\Sigma$ is denoted by $\mathfrak{W}(n, \Sigma)$. The dimensionality of these distributions will be clear from the context. A chi-square variate with n degrees of freedom is denoted by χ_n^2. In this paper, whenever the quotient of two chi-square variates appears, e.g. χ_n^2/χ_m^2, the numerator and denominator are independent. Throughout this paper, c and d denote arbitrary positive constants.

3. Some results in multivariate distribution theory. We now present several basic and useful results in multivariate normal distribution theory. Several of these are well-known but others (in particular 3.6° and 3.7°) have not appeared in a published work, to the author's knowledge. These facts and their proofs may be found in a set of unpublished notes of Stein (1966).

Let $X:p \times 1$ and $S:p \times p$ be independent random variables with $X \sim \mathfrak{N}(\mu, \Sigma)$ and $S \sim \mathfrak{W}(n, \Sigma)$. Partition X as $X' = (X_1', X_2')$ with $X_1:k \times 1$ $(1 \leq k < p)$ and partition μ, S, and Σ accordingly. Let $\gamma = (X_1 - S_{12}S_{22}^{-1}X_2)/(1 + X_2'S_{22}^{-1}X_2)^{\frac{1}{2}}$ and $\delta = \gamma'S_{11\cdot 2}^{-1}\gamma/\gamma'\Sigma_{11\cdot 2}^{-1}\gamma$. Then

3.1°. Conditional on X_2, $X_1 \sim \mathfrak{N}[\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2), \Sigma_{11\cdot 2}]$.

3.2°. $S_{11\cdot 2}$ is independent of (S_{12}, S_{22}) and $S_{11\cdot 2} \sim \mathfrak{W}(n - p + k, \Sigma_{11\cdot 2})$.

3.3°. Conditional on S_{22}, $S_{12}S_{22}^{-\frac{1}{2}} \sim \mathfrak{N}(\Sigma_{12}\Sigma_{22}^{-1}S_{22}^{\frac{1}{2}}, \Sigma_{11\cdot 2} \otimes I)$, where $\otimes$ denotes the Kronecker product. That is, conditional on S_{22} the columns of $S_{12}S_{22}^{-\frac{1}{2}}$ are independent and the jth column is distributed as $\mathfrak{N}(\xi_j, \Sigma_{11\cdot 2})$, where ξ_j is the jth column of $\Sigma_{12}\Sigma_{22}^{-1}S_{22}^{\frac{1}{2}}$.

3.4°. X and $X'S^{-1}X/X'\Sigma^{-1}X$ are independent, and $X'S^{-1}X/X'\Sigma^{-1}X \sim 1/\chi_{n-p+1}^2$.

3.5°. $X'S^{-1}X \sim \chi_p'^2(\mu'\Sigma^{-1}\mu)/\chi_{n-p+1}^2$, where the numerator and denominator are independent chi-square variates, the numerator non-central.

3.6°. Conditional on $(S_{12}, S_{22}, X_1, X_2)$, $\delta \sim 1/\chi_{n-p+1}^2$. Therefore δ is independent of $(S_{12}, S_{22}, X_1, X_2)$ and hence δ is independent of γ.

3.7°. Conditional on (S_{22}, X_2), X_1 and $S_{12}S_{22}^{-1}X_2$ are independent, $X_1 \sim \mathfrak{N}[\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2), \Sigma_{11\cdot 2}]$ (by 3.1°), and $S_{12}S_{22}^{-1}X_2 \sim \mathfrak{N}[\Sigma_{12}\Sigma_{22}^{-1}X_2, (X_2'S_{22}^{-1}X_2)\Sigma_{11\cdot 2}]$ (by 3.3°).

Therefore, conditional on (S_{22}, X_2),

$$\gamma \sim \mathfrak{N}[(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2)/(1 + X_2'S_{22}^{-1}X_2)^{\frac{1}{2}}, \Sigma_{11\cdot 2}].$$

In particular, if $\mu = 0$ then γ is independent of (S_{22}, X_2),

$$\gamma \sim \mathfrak{N}(0, \Sigma_{11\cdot 2}), \quad \text{and} \quad \gamma'\Sigma_{11\cdot 2}^{-1}\gamma \sim \chi_k^2.$$

The following lemma concerning orthogonally invariant multivariate distributions is similar to results appearing in Bartholomew (1961a) and Kudo (1963).

LEMMA 3.1. *Let Z be a random vector in $\mathcal{E}_p$ $(p \geqq 2)$ which possesses a density with respect to Lebesgue measure. Suppose that the distribution of Z is orthogonally invariant,* i.e., *for any orthogonal transformation Γ, Z and ΓZ are identically distributed. Then*

(i) *$Z'Z$ is independent of $Z/(Z'Z)^{\frac{1}{2}}$, the unit vector lying along Z.*

(ii) *$Z/(Z'Z)^{\frac{1}{2}}$ is distributed uniformly over $\mathcal{S} = \{x : x'x = 1\}$, the unit sphere in $\mathcal{E}_p$.*

(iii) *For any (measurable) positively homogeneous set $\mathcal{P}$, $P\{Z \varepsilon \mathcal{P}\} = m(\mathcal{P} \cap \mathcal{S})/m(\mathcal{S})$, where m is Lebesgue measure over $\mathcal{S}$.*

(iv) *For any non-singular linear transformation $A : \mathcal{E}_p \to \mathcal{E}_p$, let $\psi_p(A) = m(A(\mathcal{O}) \cap \mathcal{S})/m(\mathcal{S})$, where $\mathcal{O}$ is the positive orthant. Then $P\{AZ \geqq 0\} = \psi_p(A^{-1})$.*

REMARKS. As a consequence of (i), the events $\{Z'Z \geqq d\}$ and $\{Z \varepsilon \mathcal{P}\}$ are independent.

For the degenerate cases $p = 1$ and $p = 0$, we set $\psi_p(A) = \frac{1}{2}$ and 1, respectively.

The conditions of the lemma are satisfied in the following two cases: $Z = \Sigma^{-\frac{1}{2}}X$ and $Z = \Sigma^{\frac{1}{2}}S^{-1}X$, where $X \sim \mathfrak{N}(0, \Sigma)$, $S \sim \mathcal{W}(n, \Sigma)$, and S and X are independent. This implies that the events $\{X\Sigma^{-1}X \geqq d\}$ and $\{X \varepsilon \mathcal{P}\}$ are independent, and also that the events $\{X'S^{-1}\Sigma S^{-1}X \geqq d\}$ and $\{S^{-1}X \varepsilon \mathcal{P}\}$ are independent. These facts are in a sense analogous to the following results, which are essential for the distribution theory of Sections 6, 7, 8, and which are important in their own right.

LEMMA 3.2. *Let $X \sim \mathfrak{N}(0, \Sigma)$ and $S \sim \mathcal{W}(n, \Sigma)$ be independent, let $\mathcal{P}$ be a (measurable) positively homogeneous set, and let c be a positive number. Then*

(i) *the events $\{X'S^{-1}X \geqq c\}$ and $\{X \varepsilon \mathcal{P}\}$ are independent;*

(ii) *the events $\{X'S^{-1}X \geqq c\}$ and $\{S^{-1}X \varepsilon \mathcal{P}\}$ are independent.*

PROOF. We may take $\Sigma = I$ without loss of generality.

(i) By 3.4°, X and $v = X'S^{-1}X/X'X$ are independent. Therefore

$$\begin{aligned} P[X'S^{-1}X \geqq c, X \varepsilon \mathcal{P}] &= P[vX'X \geqq c, X \varepsilon \mathcal{P}] \\ &= E\{P[X'X \geqq c/v, X \varepsilon \mathcal{P} \mid v]\} \\ &= E\{P[X'X \geqq c/v \mid v]P[X \varepsilon \mathcal{P}]\} \\ &= P[X'S^{-1}X \geqq c]P[X \varepsilon \mathcal{P}], \end{aligned}$$

where the third equality follows from the remarks after Lemma 3.1.

(ii) Let $X^* = X/(X'X)^{\frac{1}{2}}$ be the unit vector along X. By Lemma 3.1, X^* and $X'X$ are independent. Let $\Gamma = \Gamma(X^*)$ be a random orthogonal matrix $(p \times p)$ defined in such a way that $\Gamma X^* = e_1 \equiv (1, 0, \cdots, 0)'$ and let $W = \Gamma S \Gamma'$.

The conditional distribution of W given X is again $\mathcal{W}(n, I)$ so W and X are independent. Since $X = X^*(X'X)^{\frac{1}{2}} = \Gamma' e_1(X'X)^{\frac{1}{2}}$ it follows that

$$X'S^{-1}X = e'_1 W^{-1} e_1(X'X) = w^{11}(X'X) \tag{3.1}$$

where $W^{-1} = (w^{ij})$. Also,

$$S^{-1}X = \Gamma' W^{-1} \Gamma X = (\Gamma' W^{-1} e_1)\ (X'X)^{\frac{1}{2}} = \Gamma' w^{(1)}(X'X)^{\frac{1}{2}},$$

where $w^{(1)}$ is the first column of W^{-1}. Thus, since $\mathcal{P}$ is positively homogeneous,

$$\{S^{-1}X \ \varepsilon\ \mathcal{P}\} = \{\Gamma' w^{(1)} \ \varepsilon\ \mathcal{P}\}. \tag{3.2}$$

Partition W and W^{-1} as follows:

$$W = \begin{pmatrix} w_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}, \qquad W^{-1} = \begin{pmatrix} w^{11} & W^{12} \\ W^{21} & W^{12} \end{pmatrix},$$

where w_{11} and w^{11} are scalars. Then from well-known relations between W and W^{-1},

$$w^{(1)} = \begin{pmatrix} w^{11} \\ W^{21} \end{pmatrix} = \begin{pmatrix} W_{11\cdot 2}^{-1} \\ -W_{22}^{-1} W_{21} W_{11\cdot 2}^{-1} \end{pmatrix} = \begin{pmatrix} 1 \\ -W_{22}^{-1} W_{21} \end{pmatrix} W_{11\cdot 2}^{-1}.$$

Note that $w^{11} = W_{11\cdot 2}^{-1}$ is a positive scalar (with probability one) and is independent of $W_{22}^{-1}W_{21}$ (see 3.2°). Then (3.1) and (3.2) become

$$\{X'S^{-1}X \geqq c\} = \{W_{11\cdot 2}^{-1}(X'X) \geqq c\},$$

$$\{S^{-1}X \ \varepsilon\ \mathcal{P}\} = \left\{\Gamma' \begin{pmatrix} 1 \\ -W_{22}^{-1}W_{21} \end{pmatrix} \varepsilon\ \mathcal{P}\right\}.$$

Therefore, setting $\rho = \begin{pmatrix} 1 \\ -W_{22}^{-1}W_{21} \end{pmatrix}$,

$$\begin{aligned} P[X'S^{-1}X \geqq c,\ S^{-1}X \ \varepsilon\ \mathcal{P}] &= P\{W_{11\cdot 2}^{-1}(X'X) \geqq c,\ \Gamma'\rho \ \varepsilon\ \mathcal{P}\} \\ &= E\{P[W_{11\cdot 2}^{-1} \geqq c/X'X,\ \rho \ \varepsilon\ \Gamma(\mathcal{P}) \mid X]\} \\ &= E\{P[W_{11\cdot 2}^{-1} \geqq c/X'X \mid X] P[\rho \ \varepsilon\ \Gamma(\mathcal{P}) \mid X]\} \\ &= E\{P[W_{11\cdot 2}^{-1} \geqq c/X'X \mid X]\} E\{P[\rho \ \varepsilon\ \Gamma(\mathcal{P}) \mid X]\} \\ &= P[X'S^{-1}X \geqq c] P[S^{-1}X \ \varepsilon\ \mathcal{P}]. \end{aligned}$$

Here, the third equality follows from the independence of W and X and the independence of $W_{11\cdot 2}$ and ρ, while the fourth equality follows from the independence of $X'X$ and $\Gamma = \Gamma(X^*)$. □

4. Orthogonal projection onto a positively homogeneous set. If $\Sigma: p \times p$ is a positive definite symmetric matrix, the inner product $\langle \cdot, \cdot \rangle_\Sigma$ on $\mathcal{E}_p$ is defined by $\langle x, y \rangle_\Sigma = x'\Sigma^{-1}y$. The associated norm $\|\cdot\|_\Sigma$ is defined by $\|x\|_\Sigma^2 = \langle x, x \rangle_\Sigma$. The distance $\|x - \mathcal{A}\|_\Sigma$ from the point x to the set $\mathcal{A}$ is given by $\inf_{\{z \varepsilon \mathcal{A}\}} \|x - z\|_\Sigma$.

Definition. Let $\mathcal{P}$ be a closed positively homogeneous set in $\mathcal{E}_p$. For any point x and any $\Sigma > 0$, the *orthogonal projection* $\pi_\Sigma(x; \mathcal{P})$ *of* x *onto* $\mathcal{P}$ *with respect to the inner product* $\langle \cdot, \cdot \rangle_\Sigma$ is any point in $\mathcal{P}$ (not necessarily uniquely determined) which minimizes $\|x - z\|_\Sigma$ among all $z \varepsilon \mathcal{P}$. Any such minimizing point is called a *version* of $\pi_\Sigma(x; \mathcal{P})$.

If $x \varepsilon \mathcal{P}$ then $\pi_\Sigma(x; \mathcal{P})$ is uniquely determined and is equal to x. If $\mathcal{P}$ is convex then $\pi_\Sigma(x; \mathcal{P})$ is uniquely determined for any x. Note that $\|x - \mathcal{P}\|_\Sigma = \|x - \pi_\Sigma(x; \mathcal{P})\|_\Sigma$.

Lemma 4.1. *For any closed positively homogeneous set* $\mathcal{P}$ *and any point* x,

(i) $\langle \pi_\Sigma(x; \mathcal{P}), x - \pi_\Sigma(x; \mathcal{P}) \rangle_\Sigma = 0$,

(ii) $\|x\|_\Sigma^2 = \|\pi_\Sigma(x; \mathcal{P})\|_\Sigma^2 + \|x - \pi_\Sigma(x; \mathcal{P})\|_\Sigma^2$.

Proof. Since $\pi_\Sigma(x; \mathcal{P})$ is not uniquely determined, we show that (i) and (ii) hold for all versions of $\pi_\Sigma(x; \mathcal{P})$. Let z be any such version. If $z = 0$ the results are trivial, so assume $z \neq 0$. Let $\mathcal{L}^* = \{y : y = az, -\infty < a < \infty\}$, a one-dimensional subspace. Since $z \neq 0$, it is readily shown that $z = \pi_\Sigma(x; \mathcal{L}^*)$. Therefore, (i) and (ii) follow from well-known properties of orthogonal projection onto a subspace. □

Remark. By (ii), $\|\pi_\Sigma(x; \mathcal{P})\|_\Sigma$ is a well-determined function even though $\pi_\Sigma(x; \mathcal{P})$ is not.

Lemma 4.2. *If* $A: \mathcal{E}_p \to \mathcal{E}_p$ *is a non-singular linear transformation*

(i) $A\pi_\Sigma(x; \mathcal{P}) = \pi_{A\Sigma A'}(Ax; A(\mathcal{P}))$,

(ii) $\|\pi_\Sigma(x; \mathcal{P})\|_\Sigma = \|\pi_{A\Sigma A'}(Ax; A(\mathcal{P}))\|_{A\Sigma A'}$,

(iii) $\|x - \pi_\Sigma(x; \mathcal{P})\|_\Sigma = \|Ax - \pi_{A\Sigma A'}(Ax; A(\mathcal{P}))\|_{A\Sigma A'}$.

Lemma 4.3. (i) *If* $\mathcal{P}_1 \supset \mathcal{P}_2 \supset \cdots \supset \mathcal{P}_n \supset \cdots$, then $\|x - \pi_\Sigma(x; \mathcal{P}_n)\|_\Sigma \uparrow \|x - \pi_\Sigma(x; \mathcal{P})\|_\Sigma$, *where* $\mathcal{P} = \bigcap_{n=1}^{\infty} \mathcal{P}_n$.

(ii) *If* $\mathcal{P}_1 \subset \mathcal{P}_2 \subset \cdots \subset \mathcal{P}_n \subset \cdots$, *then* $\|x - \pi_\Sigma(x; \mathcal{P}_n)\|_\Sigma \downarrow \|x - \pi_\Sigma(x; \mathcal{P})\|_\Sigma$ *where* $\mathcal{P} =$ *closure of* $\bigcup_{n=1}^{\infty} \mathcal{P}_n$.

5. The likelihood ratio test. Let $Y_1, Y_2, \cdots, Y_N$ be independent random vectors of dimension $p \times 1$, each distributed as $\mathfrak{N}(\eta, \Sigma)$ with unknown mean η and unknown covariance matrix Σ. Assume $N \geqq p + 1$ so that $S = \sum_{i=1}^N (Y_i - \bar{Y})(Y_i - \bar{Y})'$ is positive definite with probability one. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be closed positively homogeneous sets in $\mathcal{E}_p$ such that $\mathcal{P}_1 \subset \mathcal{P}_2$, and consider the problem of testing $H: \eta \varepsilon \mathcal{P}_1$ vs. $K: \eta \varepsilon \mathcal{P}_2$. Since $X = N^{\frac{1}{2}}\bar{Y}$ and S are sufficient statistics for η and Σ, and since $\eta \varepsilon \mathcal{P}_i$ if and only if $\mu = N^{\frac{1}{2}}\eta \varepsilon \mathcal{P}_i$, we treat the problem in the following equivalent form: observe $X \sim \mathfrak{N}(\mu, \Sigma)$ and $S \sim \mathcal{W}(N - 1, \Sigma)$ with X and S independent, and

$$\text{test } H: \mu \varepsilon \mathcal{P}_1 \text{ against } K: \mu \varepsilon \mathcal{P}_2 \tag{5.1}$$

with Σ completely unknown. This general problem includes all cases to be treated later. When $\mathcal{P}_1$ and $\mathcal{P}_2$ are subspaces, this problem reduces to a standard problem in multivariate analysis which has been thoroughly investigated by Giri (see Giri (1968) and the references listed there).

In Theorem 5.1 the likelihood ratio test (LRT) for (5.1) is derived. In Theorem 5.3 it is shown that the power of the LRT approaches one uniformly as the distance from the true mean μ to the set $\mathcal{P}_1$ becomes large.

Throughout the remainder of this paper, $P_{\mu,\Sigma}$ denotes the underlying probability measure when the true parameter values are μ, Σ.

THEOREM 5.1. *The* LRT *for* (5.1) *is the following*: *reject* H *if*

$$(5.2)\quad U(\mathcal{P}_1, \mathcal{P}_2) \equiv [\|X - \pi_S(X; \mathcal{P}_1)\|_S^2 - \|X - \pi_S(X; \mathcal{P}_2)\|_S^2] \cdot [1 + \|X - \pi_S(X; \mathcal{P}_2)\|_S^2]^{-1} \geqq c.$$

PROOF. The likelihood function is proportional to

$$L = |\Sigma|^{-N/2} \exp\{-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}S - \tfrac{1}{2}(X - \mu)'\Sigma^{-1}(X - \mu)\}.$$

For fixed μ it is well-known (e.g. Anderson (1958), Lemma 3.2.2) that

$$\begin{aligned}\max_{\{\Sigma>0\}} L &= |S + (X - \mu)(X - \mu)'|^{-N/2} \\ &= |S|^{-N/2}[1 + (X - \mu)'S^{-1}(X - \mu)]^{-N/2} \\ &= |S|^{-N/2}[1 + \|X - \mu\|_S^2]^{-N/2},\end{aligned}$$

so for $i = 1, 2$ we obtain

$$\max_{\{\mu\varepsilon\mathcal{P}_i, \Sigma>0\}} L = |S|^{-N/2}[1 + \|X - \pi_S(X; \mathcal{P}_i)\|_S^2]^{-N/2}.$$

Hence the LRT statistic is a strictly increasing function of $U(\mathcal{P}_1, \mathcal{P}_2)$. □

REMARKS. By Lemma 4.1 (ii) we can rewrite the LRT statistic as

$$(5.3)\quad U(\mathcal{P}_1, \mathcal{P}_2) = [\|\pi_S(X; \mathcal{P}_2)\|_S^2 - \|\pi_S(X; \mathcal{P}_1)\|_S^2][1 + \|X - \pi_S(X; \mathcal{P}_2)\|_S^2]^{-1}.$$

If $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_1'$, $\mathcal{P}_2'$ are positively homogeneous sets such that $\mathcal{P}_1' \subset \mathcal{P}_1 \subset \mathcal{P}_2 \subset \mathcal{P}_2'$, then

$$(5.4)\quad U(\mathcal{P}_1, \mathcal{P}_2) \leqq U(\mathcal{P}_1', \mathcal{P}_2').$$

We have shown that $\pi_S(X; \mathcal{P})$ is the (not necessarily uniquely determined) MLE of μ under the restriction $\mu \varepsilon \mathcal{P}$ with Σ unknown.

LEMMA 5.2. *Let* $V{:}p \times 1$ *be a random vector*, $T{:}p \times p$ *a random positive definite symmetric matrix, and* c *an arbitrary positive number. Set* $Y = Y(\nu) = V + \nu$, *where* $\nu{:}p \times 1$ *is a vector parameter. For each* $\epsilon > 0$ *there exists* $M(\epsilon) > 0$ *such that for any set* $\mathcal{A}$ *in* $\mathcal{E}_p$,

$$(5.5)\quad \|\nu - \mathcal{A}\|_I \geqq M(\epsilon) \text{ implies } P[\|Y - \mathcal{A}\|_T \geqq c] \geqq 1 - \epsilon.$$

PROOF. Choose $a = a(\epsilon) > 0$ such that

$$(5.6)\quad P[\|Y - \nu\|_T \leqq a] \geqq 1 - \epsilon/2.$$

Choose $b = b(\epsilon) > 0$ such that

$$P[T^{-1} - bI \text{ is positive definite}] \geqq 1 - \epsilon/2;$$

that is, such that $P[\lambda_{\min}(T^{-1}) > b] \geqq 1 - \epsilon/2$, where $\lambda_{\min}(T^{-1})$ is the smallest characteristic root of T^{-1}. It then follows that

$$P[b\,\|x - y\|_I < \|x - y\|_T \text{ for all } x, y \text{ in } \mathcal{E}_p] \geqq 1 - \epsilon/2,$$

so

$$P[b\,\|\nu - \mathcal{A}\|_I < \|\nu - \mathcal{A}\|_T \text{ for all } \nu \in \mathcal{E}_p,\ \mathcal{A} \subset \mathcal{E}_p] \geqq 1 - \epsilon/2.$$

Therefore $\|\nu - \mathcal{A}\|_I \geqq M$ implies $P[bM \leqq \|\nu - \mathcal{A}\|_T] \geqq 1 - \epsilon/2$. Hence by (5.6), $\|\nu - \mathcal{A}\|_I \geqq M$ implies $P[\|Y - \mathcal{A}\|_T \geqq bM - a] \geqq 1 - \epsilon$. Setting $M = M(\epsilon) = (a + c)/b$, we obtain (5.5). □

THEOREM 5.3. *The power of the* LRT (5.2) *approaches one uniformly in* μ *and* Σ *as* $\|\mu - \mathcal{P}_1\|_\Sigma \to \infty$ *with* $\mu \in \mathcal{P}_2$.

PROOF. Precisely, we show that for $\epsilon > 0$ there exists $M(\epsilon) > 0$ (not depending on μ, Σ, $\mathcal{P}_1$, $\mathcal{P}_2$) such that

$$\text{(5.7)} \quad \mu \in \mathcal{P}_2 \text{ and } \|\mu - \mathcal{P}_1\|_\Sigma \geqq M(\epsilon) \text{ imply } P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] \geqq 1 - \epsilon.$$

First, $\|X - \mathcal{P}_2\|_S \leqq \|X - \mu\|_S$ since $\mu \in \mathcal{P}_2$, so from (5.2) we have

$$U(\mathcal{P}_1, \mathcal{P}_2) \geqq [\|X - \mathcal{P}_1\|_S^2 - \|X - \mu\|_S^2][1 + \|X - \mu\|_S^2]^{-1}.$$

The distribution of $\|X - \mu\|_S^2 = (X - \mu)'S^{-1}(X - \mu)$ is independent of μ and Σ; thus it suffices to show that (5.7) holds with $U(\mathcal{P}_1, \mathcal{P}_2)$ replaced by $\|X - \mathcal{P}_1\|_S$. Letting $Y = \Sigma^{-\frac{1}{2}}X$, $\nu = \Sigma^{-\frac{1}{2}}\mu$, $\mathcal{A} = \Sigma^{-\frac{1}{2}}(\mathcal{P}_1)$, and $T = \Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}}$, we see that $V \equiv Y - \nu \sim \mathfrak{N}(0, I)$ and $T \sim \mathcal{W}(N - 1, I)$. By Lemma 4.2, $\|X - \mathcal{P}_1\|_S = \|Y - \mathcal{A}\|_T$ and $\|\mu - \mathcal{P}_1\|_\Sigma = \|\nu - \mathcal{A}\|_I$. Thus, Lemma 5.2 provides the desired result. □

6. One-sided alternatives. We now study the null distribution of the LRT statistic in the case where $\mathcal{P}_1 = \{0\}$, the set consisting of the origin alone, and $\mathcal{P}_2 = \mathcal{C}$, a cone as defined in section 2. We write $U(\mathcal{C})$ for $U(\{0\}, \mathcal{C})$. From (5.2) and (5.3) we have

$$\text{(6.1)} \quad \begin{aligned} U(\mathcal{C}) &= [\|X\|_S^2 - \|X - \pi_S(X; \mathcal{C})\|_S^2][1 + \|X - \pi_S(X; \mathcal{C})\|_S^2]^{-1} \\ &= \|\pi_S(X; \mathcal{C})\|_S^2[1 + \|X - \pi_S(X; \mathcal{C})\|_S^2]^{-1}. \end{aligned}$$

In order to carry out the LRT at size α we must determine that positive number c_α which satisfies

$$\sup_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{C}) \geqq c_\alpha] = \alpha.$$

In general, there is no explicit formula for $P_{0,\Sigma}[U(\mathcal{C}) \geqq c]$. In Theorem 6.2 however, we obtain sharp upper and lower bounds (as Σ varies) on the null distribution of $U(\mathcal{C})$, and these bounds (and therefore also c_α) do not depend on the particular cone $\mathcal{C}$ (provided $\mathcal{C}$ contains an open set). Using the upper bound an explicit formula for c_α is obtained which can be readily applied.

LEMMA 6.1. *For any cone* $\mathcal{C}$ *and any* $\Sigma > 0$,

$$\text{(6.2)} \quad P_{0,\Sigma}[U(\mathcal{C}) \geqq c] \geqq \tfrac{1}{2}P[\chi_1^2/\chi_{N-p}^2 \geqq c],$$

$$\text{(6.3)} \quad P_{0,\Sigma}[U(\mathcal{C}) \geqq c] \leqq \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c],$$

where χ_1^2, χ_{p-1}^2, χ_p^2, *and* χ_{N-p}^2 *denote independent chi-square variates.*

PROOF. By (5.4) we have $U(\mathcal{L}) \leqq U(\mathcal{C}) \leqq U(\mathcal{H}^+)$ where the halfline $\mathcal{L}$ and the halfspace $\mathcal{H}^+$ are as in 2.1°. Hence for any $\Sigma > 0$,

$$P_{0,\Sigma}[U(\mathcal{L}) \geqq c] \leqq P_{0,\Sigma}[U(\mathcal{C}) \geqq c] \leqq P_{0,\Sigma}[U(\mathcal{H}^+) \geqq c].$$

By Corollary 7.6 (with $q = 0, r = 1, s = p - 1$),

$$P_{0,\Sigma}[U(\mathcal{L}) \geqq c] = \tfrac{1}{2}P[\chi_1^2/\chi_{N-p}^2 \geqq c]. \tag{6.4}$$

Also, by an argument similar to that leading to (8.4), one finds that

$$P_{0,\Sigma}[U(\mathcal{H}^+) \geqq c] = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c]. \tag{6.5}$$ □

THEOREM 6.2. *Let $\mathcal{C}$ be a cone. Then*

$$\inf_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{C}) \geqq c] = \tfrac{1}{2}P[\chi_1^2/\chi_{N-p}^2 \geqq c]. \tag{6.6}$$

Furthermore, if $\mathcal{C}$ contains a p-dimensional open set,

$$\sup_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{C}) \geqq c] = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c]. \tag{6.7}$$

In this latter case the level α cutoff point c_α for the LRT *is the unique solution of the equation*

$$\alpha = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c_\alpha] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c_\alpha]. \tag{6.8}$$

PROOF. Let $\beta = \tfrac{1}{2}P[\chi_1^2/\chi_{N-p}^2 \geqq c]$ and $\gamma = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c]$. By (6.2) we have $\inf_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{C}) \geqq c] \geqq \beta$. To obtain the opposite inequality we exhibit a sequence $\{\Sigma_n\}$, $\Sigma_n > 0$, such that as $n \to \infty$

$$P_{0,\Sigma_n}[U(\mathcal{C}) \geqq c] \to \beta. \tag{6.9}$$

Let $\mathcal{C}_\lambda$, $\mathcal{L}$, and $\{A_n\}$ be as given in 2.2° and 2.4°, and let $\Sigma_n = (A_n'A_n)^{-1}$. Then

$$\beta \leqq P_{0,\Sigma_n}[U(\mathcal{C}) \geqq c] \leqq P_{0,\Sigma_n}[U(\mathcal{C}_\lambda) \geqq c] = P_{0,I}[U(A_n(\mathcal{C}_\lambda)) \geqq c], \tag{6.10}$$

where the equality is obtained by applying Lemma 4.2 in (6.1) and noting that $A_nX \sim \mathfrak{N}(0, I)$ and $A_nSA_n' \sim \mathcal{W}(N-1, I)$. Now for fixed X and S, $U(A_n(\mathcal{C}_\lambda)) \to U(\mathcal{L})$ as $n \to \infty$ by Lemma 4.3 (i). Therefore

$$P_{0,I}[U(A_n(\mathcal{C}_\lambda)) \geqq c] \to P_{0,I}[U(\mathcal{L}) \geqq c] = \beta,$$

the equality coming from (6.4). This, together with (6.10), yields (6.9).

Next, assuming that $\mathcal{C}$ contains an open set, we obtain (6.7) by exhibiting a sequence $\{\Sigma_n\}$ such that as $n \to \infty$

$$P_{0,\Sigma_n}[U(\mathcal{C}) \geqq c] \to \gamma. \tag{6.11}$$

Let $\mathcal{C}_\nu$, $\mathcal{H}^+$, and $\{B_n\}$ be as given in 2.3° and 2.4°, and let $\Sigma_n = (B_n'B_n)^{-1}$. Then as in (6.10),

$$\gamma \geqq P_{0,\Sigma_n}[U(\mathcal{C}) \geqq c] \geqq P_{0,\Sigma_n}[U(\mathcal{C}_\nu) \geqq c] = P_{0,I}[U(B_n(\mathcal{C}_\nu)) \geqq c].$$

Applying Lemma 4.3 (ii) and using (6.5), as $n \to \infty$

$$P_{0,I}[U(B_n(\mathcal{C}_\nu)) \geqq c] \to P_{0,I}[U(\mathcal{H}^+) \geqq c] = \gamma.$$ □

REMARK. For the problem of testing $H:\mu = 0$ vs. $K:\mu \varepsilon \mathcal{C}$ under assumptions (a) or (b) (see Introduction), power studies have indicated that the LRT has greater power than the best invariant test for testing H against the unrestricted alternative $\mu \neq 0$. In the case Σ unknown, the best invariant test for testing H vs. $\mu \neq 0$ is Hotelling's T^2-test, which rejects H if $\|X\|_S^2 = X'S^{-1}X \geqq c$ and which is unbiased. It follows from Theorem 6.2 and the continuity of the power function of the LRT, that the LRT is not unbiased for testing H vs. K (assuming $\mathcal{C}$ contains an open set). This implies that there exist alternative points near H at which the power of the level α LRT is less than that of the level α T^2-test. Another possibly unfavorable (for the LRT) observation is the following. Since $\|X\|_S^2 \geqq U(\mathcal{C})$, for any μ, Σ we have that $P_{\mu,\Sigma}[U(\mathcal{C}) \geqq c_\alpha] \leqq P_{\mu,\Sigma}[\|X\|_S^2 \geqq c_\alpha]$. Therefore, the level α LRT is uniformly less powerful than the level α^* T^2-test, where, however,

$$\alpha^* = P[\chi_p^2/\chi_{N-p}^2 \geqq c_\alpha] = \alpha + \tfrac{1}{2}\{P[\chi_p^2/\chi_{N-p}^2 \geqq c_\alpha] - P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c_\alpha]\} > \alpha.$$

These considerations indicate that there is a need for a careful comparison of the power of the LRT with that of the T^2-test before the LRT can be wholeheartedly recommended for use in practice.

The above results can be applied to the problem of testing a subset of the components of μ against one-sided alternatives. Partition μ in the form $\mu' = (\mu_1', \mu_2', \mu_3')$, with $\mu_1:q \times 1$, $\mu_2:r \times 1$, $\mu_3:s \times 1$, $q + r + s = p$, and partition Σ, X, and S accordingly.

We are concerned with the testing problem (5.1) when $\mathcal{P}_1$ and $\mathcal{P}_2$ have the form

$$\mathcal{P}_1 = \{\mu:\mu_1 \varepsilon \mathcal{E}_q,\ \mu_2 = 0,\ \mu_3 = 0\},$$
$$\mathcal{P}_2 = \{\mu:\mu_1 \varepsilon \mathcal{E}_q,\ \mu_2 \varepsilon \mathcal{C},\ \mu_3 = 0\}. \tag{6.12}$$

The LRT statistic $U(\mathcal{P}_1, \mathcal{P}_2)$ is given in (5.2). Using (2.1) repeatedly and setting $Y = X_2 - S_{23}S_{33}^{-1}X_3$, $T = S_{22\cdot3}$ we have

$$\begin{aligned} U(\mathcal{P}_1, \mathcal{P}_2) &= [\|Y\|_T^2 - \|Y - \pi_T(Y;\mathcal{C})\|_T^2] \\ &\qquad [1 + X_3'S_{33}^{-1}X_3 + \|Y - \pi_T(Y;\mathcal{C})\|_T^2]^{-1} \\ &= [\|Z\|_T^2 - \|Z - \pi_T(Z;\mathcal{C})\|_T^2]\,[1 + \|Z - \pi_T(Z;\mathcal{C})\|_T^2]^{-1} \end{aligned} \tag{6.13}$$

where $Z = Y/(1 + X_3'S_{33}^{-1}X_3)^{\frac{1}{2}}$. By applying 3.7° it can be shown that when $\mu_3 = 0$ the distribution depends only on μ_2 and $\Sigma_{22\cdot3}$ (see the proof of the following theorem and also of Theorem 8.5.)

THEOREM 6.3. *Let $\mathcal{P}_1$, $\mathcal{P}_2$ be given by* (6.12), *with $\mathcal{C}$ a cone in $\mathcal{E}_r$. Under the null hypothesis $H:\mu \varepsilon \mathcal{P}_1$,*

$$\inf_{\{\Sigma_{22\cdot3}>0\}} P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = \tfrac{1}{2}P[\chi_1^2/\chi_{N-s-r}^2 \geqq c].$$

Furthermore, if $\mathcal{C}$ contains an r-dimensional open set,

$$\sup_{\{\Sigma_{22\cdot3}>0\}} P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = \tfrac{1}{2}P[\chi_{r-1}^2/\chi_{N-s-r}^2 \geqq c] + \tfrac{1}{2}P[\chi_r^2/\chi_{N-s-r}^2 \geqq c],$$

from which the level α cutoff point for the LRT *may be determined.*

PROOF. If $\mu \in \mathcal{P}_1$, it follows from 3.2° and 3.7° that $Z \sim \mathfrak{N}(0, \Sigma_{22\cdot3})$, $T \sim \mathfrak{W}(N - 1 - s, \Sigma_{22\cdot3})$, and Z and T are independent. Comparing (6.13) with (6.1) we see that the null distribution of $U(\mathcal{P}_1, \mathcal{P}_2)$ is identical to that of $U(\mathcal{C})$, with $\Sigma_{22\cdot3}$ substituted for Σ, r for p, and $N - s$ for N. The result then follows from Theorem 6.2. □

7. Testing $H:\mu = 0$ against $K:\mu \geqq 0$. In this section the LRT for the testing problem considered in Section 6 is studied in further detail for the special case $\mathcal{C} = \mathcal{O}$, the positive orthant. Nüesch attempted to treat the case Σ unknown but his work is invalidated by an error (see Shorack (1967), p. 1751).

For purposes of comparison we state the results of Bartholomew, Kudô, and Nüesch for the case Σ known: the LRT for testing $H:\mu = 0$ vs. $K:\mu \in \mathcal{O}$ with Σ known rejects H if $\|\pi_\Sigma(X; \mathcal{O})\|_\Sigma^2 \geqq c$; the null distribution of the LRT statistic is given by

$$P_{0,\Sigma}[\|\pi_\Sigma(X; \mathcal{O})\|_\Sigma^2 \geqq c] = \sum_{k=1}^{p} P[\chi_k^2 \geqq c]w(p, k, \Sigma) \tag{7.1}$$

where the weight $w(p, k, \Sigma)$ is the probability that exactly k of the p components of $\pi_\Sigma(X; \mathcal{O})$ are strictly positive. Discussions concerning these weights and relevant tables are given by Bartholomew, Miles, Chacko, Nüesch (1964), and Shorack (see also Theorem 7.4). Notice that $\pi_\Sigma(X; \mathcal{O})$ is uniquely determined since $\mathcal{O}$ is convex. We abbreviate $\pi_\Sigma(X; \mathcal{O})$ by π_Σ.)

From (6.1), the LRT statistic for testing $H:\mu = 0$ vs. $K:\mu \in \mathcal{O}$ with Σ unknown is $U(\mathcal{O})$. Theorem 6.2 can be applied to provide the value of the level α cutoff point c_α for the LRT. The fact that $\mathcal{O}$ is a convex polyhedral cone whose faces are sections of subspaces, however, enables us to obtain explicitly in Theorem 7.4 the distribution of $U(\mathcal{O})$ under the null hypothesis $H:\mu = 0$. This is achieved by conditioning on the event that $\pi_S(X; \mathcal{O})$ lies in a given face of $\mathcal{O}$. We abbreviate $U(\mathcal{O})$ by U, $\pi_S(X; \mathcal{O})$ by π_S, and write the ith component of π_S as $\pi_{S,i}$.

To begin the evaluation of $P_{0,\Sigma}[U \geqq c]$, set up a 1-1 correspondence between the 2^p faces of $\mathcal{O}$ (including the "faces" $\{x: x > 0\}$ and $\{x: x = 0\}$) and the 2^p subsets of the set $\{1, \cdots, p\}$ of indices as follows: to any subset M of $\{1, \cdots, p\}$ let correspond the face $\mathfrak{X}_M = \{x: x_i > 0 \text{ if } i \in M, x_i = 0 \text{ if } i \notin M\}$. Then the event that π_S lies in the face $\mathfrak{X}_M$ is given by

$$\{(X, S): \pi_S \in \mathfrak{X}_M\} = \{(X, S): \pi_{S,i} > 0 \quad \text{if} \quad i \in M, \pi_{S,i} = 0 \quad \text{if} \quad i \notin M\}. \tag{7.2}$$

Suppose without loss of generality that M consists of the last k members of $\{1, \cdots, p\}$. Partition X, S, and Σ as follows:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \qquad S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{7.3}$$

with $X_2: k \times 1$, $S_{22}: k \times k$, and $\Sigma_{22}: k \times k$. (Note: for a general subset M with k elements, permute the components of X so that those whose indices belong to M occupy the final k positions, and permute the rows and columns of S and Σ accordingly.) A proof of the next lemma is given by Nüesch (1966).

LEMMA 7.1. (i) $\{(X, S): \pi_S \in \mathfrak{X}_M\} = \{(X, S): X_2 - S_{21}S_{11}^{-1}X_1 > 0, S_{11}^{-1}X_1 \leqq 0\}$.

(ii) *Given that* $\pi_S \varepsilon \mathfrak{X}_M$, *then*

$$\pi_S = \begin{pmatrix} 0 \\ X_2 - S_{21} S_{11}^{-1} X_1 \end{pmatrix}.$$

THEOREM 7.2. *Under the hypothesis* $H:\mu = 0$, *the conditional distribution of* U *given the event* $\{\pi_S \varepsilon \mathfrak{X}_M\}$ *is*

$$P_{0,\Sigma}[U \geqq c \mid \pi_S \varepsilon \mathfrak{X}_M] = P[\chi_k^2/\chi_{N-p}^2 \geqq c], \tag{7.4}$$

where $k = \#(M)$, *the number of elements in* M (i.e., *the dimension of the face* $\mathfrak{X}_M$).

PROOF. Applying Lemma 7.1 (ii) in (6.1) and using (2.1), we find that $\pi_S \varepsilon \mathfrak{X}_M$ implies

$$U = (X_2 - S_{21}S_{11}^{-1}X_1)'S_{22\cdot1}^{-1}(X_2 - S_{21}S_{11}^{-1}X_1)/(1 + X_1'S_{11}^{-1}X_1).$$

Then by Lemma 7.1 (i), 3.2°, and 3.7°, we have (omitting subscripts from $P_{0,\Sigma}$)

$$\begin{aligned} P[U \geqq c, \pi_S \varepsilon \mathfrak{X}_M] &= P[\gamma'S_{22\cdot1}^{-1}\gamma \geqq c, \gamma > 0, S_{11}^{-1}X_1 \leqq 0] \\ &= P[\gamma'S_{22\cdot1}^{-1}\gamma \geqq c, \gamma > 0]P[S_{11}^{-1}X_1 \leqq 0], \end{aligned} \tag{7.5}$$

$$P[\pi_S \varepsilon \mathfrak{X}_M] = P[\gamma > 0, S_{11}^{-1}X_1 \leqq 0] = P[\gamma > 0]P[S_{11}^{-1}X_1 \leqq 0], \tag{7.6}$$

where $\gamma = \gamma(X, S) = (X_2 - S_{21}S_{11}^{-1}X_1)/(1 + X_1'S_{11}^{-1}X_1)^{\frac{1}{2}}$. By 3.7°, however, $\gamma \sim \mathfrak{N}(0, \Sigma_{22\cdot1})$ and hence by Lemma 3.2 (i)

$$\begin{aligned} P[\gamma'S_{22\cdot1}^{-1}\gamma \geqq c, \gamma > 0] &= P[\gamma'S_{22\cdot1}^{-1}\gamma \geqq c]P[\gamma > 0] \\ &= P[\chi_k^2/\chi_{N-p}^2 \geqq c]P[\gamma > 0], \end{aligned} \tag{7.7}$$

the last equality a consequence of 3.5°. □

Notice that the conditional distribution of U does not depend on Σ, and depends on the face $\mathfrak{X}_M$ only through its dimension k. For any $x \varepsilon \mathfrak{O}$, let $K(x) =$ the number of positive components of $x =$ the dimension of the face of $\mathfrak{O}$ in which x lies. Then by the last remark

$$P_{0,\Sigma}[U \geqq c \mid K(\pi_S) = k] = P[\chi_k^2/\chi_{N-p}^2 \geqq c]. \tag{7.8}$$

From (7.4) and (7.8) we obtain the following expressions for the null distribution of the LRT statistic U:

$$\begin{aligned} P_{0,\Sigma}[U \geqq c] &= \textstyle\sum_M P[\chi_k^2/\chi_{N-p}^2 \geqq c]P_{0,\Sigma}[\pi_S \varepsilon \mathfrak{X}_M] \\ &= \textstyle\sum_{k=1}^{p} P[\chi_k^2/\chi_{N-p}^2 \geqq c]P_{0,\Sigma}[K(\pi_S) = k], \end{aligned} \tag{7.9}$$

where in the first summation M ranges over all subsets of $\{1, \cdots, p\}$ and $k = \#(M)$.

Again suppose that M consists of the last k elements of $\{1, \cdots, p\}$.

LEMMA 7.3. *Let* $\psi_p(A)$ *be as defined in Lemma* 3.1. *Then*

$$P_{0,\Sigma}[\pi_S \varepsilon \mathfrak{X}_M] = P_{0,\Sigma}[\pi_\Sigma \varepsilon \mathfrak{X}_M] = \psi_k(\Sigma_{22\cdot1}^{-\frac{1}{2}})\psi_{p-k}(\Sigma_{11}^{\frac{1}{2}}). \tag{7.10}$$

PROOF. From Lemma 7.1, $\{\pi_\Sigma \varepsilon \mathfrak{X}_M\} = \{X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1 > 0, \Sigma_{11}^{-1}X_1 \leqq 0\}$. Since X_1 and $X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1$ are independent (omitting subscripts from $P_{0,\Sigma}$)

$$P[\pi_\Sigma \varepsilon \mathfrak{X}_M] = P[X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1 > 0]P[\Sigma_{11}^{-1}X_1 \leqq 0]. \tag{7.11}$$

By Lemma 3.1, $P[\gamma > 0] = P[X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1 > 0] = \psi_k(\Sigma_{22.1}^{-\frac{1}{2}})$ since γ and $X_2 - \Sigma_{21}\Sigma_{11}^{-\frac{1}{2}}X_1$ are both distributed as $\mathfrak{N}(0, \Sigma_{22\cdot1})$. Next let

$$Y = \Sigma_{11}^{-\frac{1}{2}}X_1 \sim \mathfrak{N}(0, I),\ T = \Sigma_{11}^{-\frac{1}{2}}S_{11}\Sigma_{11}^{-1} \sim \mathfrak{W}(N - 1, I),$$

and $Z = T^{-1}Y$. Then $P[S_{11}^{-1}X_1 \leqq 0] = P[\Sigma_{11}^{-\frac{1}{2}}Z \leqq 0]$ and we again apply Lemma 3.1 to conclude $P[S_{11}^{-1}X_1 \leqq 0] = P[\Sigma_{11}^{-\frac{1}{2}}X_1 \leqq 0] = \psi_{p-k}(\Sigma_{11}^{\frac{1}{2}})$. Comparing (7.6) and (7.11) completes the proof. □

Summing (7.10) over all M having exactly k elements,

$$P_{0,\Sigma}[K(\pi_S) = k] = P_{0,\Sigma}[K(\pi_\Sigma) = k] = \textstyle\sum_{\#(M)=k} \psi_k(\Sigma_{22.1}^{-\frac{1}{2}})\psi_{p-k}(\Sigma_{11}^{\frac{1}{2}}). \tag{7.12}$$

Notice that $P_{0,\Sigma}[K(\pi_\Sigma) = k] = w(p, k, \Sigma)$, the weights appearing in (7.1).

THEOREM 7.4. *Under the hypothesis $H:\mu = 0$, the distribution of the* LRT *statistic U is*

$$\begin{aligned} P_{0,\Sigma}[U \geqq c] &= \textstyle\sum_M P[\chi_k^2/\chi_{N-p}^2 \geqq c]\psi_k(\Sigma_{22.1}^{-\frac{1}{2}})\psi_{p-k}(\Sigma_{11}^{\frac{1}{2}}) \\ &= \textstyle\sum_{k=1}^p P[\chi_k^2/\chi_{N-p}^2 \geqq c]w(p, k, \Sigma), \end{aligned}$$

where $w(p, k, \Sigma) = \sum_{\#(M)=k} \psi_k(\Sigma_{22.1}^{-\frac{1}{2}})\psi_{p-k}(\Sigma_{11}^{\frac{1}{2}})$. In the first summation M ranges over all subsets of $\{1,\cdots, p\}$ and k is the number of elements in M. The partitioning of Σ is in accordance with the note after (7.3).

REMARK. Notice that $P_{0,\Sigma}[U = 0] = P_{0,\Sigma}[S^{-1}X \leqq 0] = P_{0,\Sigma}[\Sigma^{-1}X \leqq 0] = w(p, 0, \Sigma) = \psi_p(\Sigma^{\frac{1}{2}}) > 0$.

Let $A : \mathcal{E}_p \rightarrow \mathcal{E}_p$ be a non singular linear transformation. Applying Lemma 4.2, the LRT statistic $U(A(\mathcal{O}))$ for testing $H:\mu = 0$ vs. $K:\mu \varepsilon A(\mathcal{O})$ can be put into the form

$$U(A(\mathcal{O})) = \|\pi_T(Z; \mathcal{O})\|_T^2 [1 + \|Z - \pi_T(Z; \mathcal{O})\|_T^2]^{-1}, \tag{7.13}$$

where $Z = A^{-1}X \sim \mathfrak{N}(A^{-1}\mu, A^{-1}\Sigma A'^{-1})$ and

$$T = A^{-1}SA'^{-1} \sim \mathfrak{W}(N - 1, A^{-1}\Sigma A'^{-1}).$$

COROLLARY 7.5. *Under the hypothesis $H:\mu = 0$, the distribution of the* LRT *statistic $U(A(\mathcal{O}))$ is*

$$\begin{aligned} P_{0,\Sigma}[U(A(\mathcal{O})) \geqq c] &= \textstyle\sum_M P[\chi_k^2/\chi_{N-p}^2 \geqq c]\psi_k(\Lambda_{22.1}^{-\frac{1}{2}})\psi_{p-k}(\Lambda_{11}^{\frac{1}{2}}) \\ &= \textstyle\sum_{k=1}^p P[\chi_k^2/\chi_{N-p}^2 \geqq c]w(p, k, \Lambda) \end{aligned}$$

where $\Lambda = A^{-1}\Sigma A'^{-1}$. The partitioning of Λ is in accordance with the note after (7.3).

Finally, the preceding results can be applied to the problem of testing a subset of the components of the mean against one-sided alternatives of the form $A(\mathcal{O})$.

COROLLARY 7.6. *Let $\mathcal{P}_1$, $\mathcal{P}_2$ be defined by* (6.12) *with $\mathcal{C}$ replaced by $A(\mathcal{O})$.*

Under the hypothesis $H:\mu \varepsilon \mathcal{P}_1$, *the distribution of the* LRT *statistic* $U(\mathcal{P}_1, \mathcal{P}_2)$ *is*

$$P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = \sum_M P[\chi_k^2/\chi_{N-s-r}^2 \geqq c]\psi_k(\Lambda_{22.1}^{-\frac{1}{2}})\psi_{r-k}(\Lambda_{11}^{\frac{1}{2}})$$
$$= \sum_{k=1}^r P[\chi_k^2/\chi_{N-s-,}^2 \geqq c]w(r, k, \Lambda)$$

where $\Lambda = A^{-1}\Sigma_{22\cdot3}A'^{-1}$. *In the first summation* M *ranges over all subsets of* $\{1, \cdots, r\}$ *and* Λ *is partitioned in accordance with the note after* (7.3).

8. One-sided hypotheses. The geometrical techniques of Section 6 can be applied in other special cases of the general problem (5.1). For example, an easy modification of the argument leading to Theorem 6.2 provides the following sharp bounds for the null distribution of the LRT statistic for testing $H:\mu = 0$ vs. $K:\mu \varepsilon \mathcal{E}_p - \mathcal{C}$ (the complement of $\mathcal{C}$, where $\mathcal{C}$ is a cone in $\mathcal{E}_p$). The LRT statistic $U(\mathcal{E}_p - \mathcal{C})$ is given by (6.1) with $\mathcal{C}$ replaced by $\mathcal{E}_p - \mathcal{C}$. Note that the bounds do not depend on $\mathcal{C}$.

THEOREM 8.1. *For any cone* $\mathcal{C}$,

$$\sup_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{E}_p - \mathcal{C}) \geqq c] = P[\chi_p^2/\chi_{N-p}^2 \geqq c].$$

If $\mathcal{C}$ *contains a p-dimensional open set then*

$$\inf_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{E}_p - \mathcal{C}) \geqq c] = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c].$$

A more interesting special case of (5.1) is that involving one-sided hypotheses. Specifically, we treat the problem of testing $H:\mu \varepsilon \mathcal{C}$ vs. $K:\mu \varepsilon \mathcal{E}_p$ where $\mathcal{C}$ is a cone. From (5.2), the LRT statistic for this problem has the simple form

$$U(\mathcal{C}, \mathcal{E}_p) = \|X - \pi_S(X; \mathcal{C})\|_S^2 = \|X - \mathcal{C}\|_S^2,$$

which is simply the distance (determined by S) from X to $\mathcal{C}$. We derive sharp upper and lower bounds on the null distribution of $U(\mathcal{C}, \mathcal{E}_p)$ (Theorems 8.3 and 8.4), thereby providing an explicit formula for the level α cutoff point c^α, i.e., that positive number satisfying

$$\sup_{\{\mu\varepsilon\mathcal{C},\Sigma>0\}} P_{\mu,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c^\alpha] = \alpha.$$

These bounds (and also c^α) do not depend on the particular cone $\mathcal{C}$.

The next lemma, which is needed in the derivation of the upper bound, states an interesting "monotonicity" property of the power function $P_{\mu,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c]$, considered as a function of μ with Σ fixed.

LEMMA 8.2. *If the cone* $\mathcal{C}$ *is convex, then for any* $\mu_1, \mu_2 \varepsilon \mathcal{E}_p$ *satisfying* $\mu_2 - \mu_1 \varepsilon \mathcal{C}$ *we have*

$$P_{\mu_2,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c] \leqq P_{\mu_1,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c].$$

PROOF. For any $\gamma \varepsilon \mathcal{E}_p$ let $\mathcal{C}_\gamma = \mathcal{C} + \gamma \equiv \{x: x = z + \gamma, z \varepsilon \mathcal{C}\}$. If $\gamma \varepsilon \mathcal{C}$, $\mathcal{C}_\gamma \subset \mathcal{C}$ by the convexity of $\mathcal{C}$, and thus $\|X - \mathcal{C}\|_S^2 \leqq \|X - \mathcal{C}_\gamma\|_S^2$. Hence, setting

$\mu_2 - \mu_1 = \gamma \varepsilon \mathcal{C}$, we have

$$\begin{aligned} P_{\mu_2,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c] &\leqq P_{\mu_2,\Sigma}[\|X - \mathcal{C}_\gamma\|_S^2 \geqq c] \\ &= P_{\mu_2,\Sigma}[\|(X - \gamma) - \mathcal{C}\|_S^2 \geqq c] \\ &= P_{\mu_1,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c]. \qquad \square \end{aligned}$$

THEEOREM 8.3. *For any cone* $\mathcal{C}$

$$\begin{aligned} \sup_{\{\mu\varepsilon\mathcal{C},\Sigma>0\}} P_{\mu,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c] &= \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p+1}^2 \geqq c] \\ &\quad + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c] \\ &= \sup_{\{\Sigma>0\}} P_{0,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c]. \end{aligned} \tag{8.1}$$

The level α *cutoff point* c^α *for the* LRT *is the unique solution to the equation*

$$\alpha = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p+1}^2 \geqq c^\alpha] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c^\alpha].$$

PROOF. Trivially,

$$\sup_{\{\Sigma>0\}} P_{0,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c] \leqq \sup_{\{\mu\varepsilon\mathcal{C},\Sigma>0\}} P_{\mu,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c].$$

In order to show that

$$\sup_{\{\mu\varepsilon\mathcal{C},\Sigma>0\}} P_{\mu,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c] \leqq \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p+1}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c], \tag{8.2}$$

fix $\mu \varepsilon \mathcal{C}$ and $\Sigma > 0$ and let $\mathcal{L}$ be any halfline contained in $\mathcal{C}$. Then $\|X - \mathcal{C}\|_S^2 \leqq \|X - \mathcal{L}\|_S^2$, so

$$P_{\mu,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c] \leqq P_{\mu,\Sigma}[\|X - \mathcal{L}\|_S^2 \geqq c] \leqq P_{0,\Sigma}[\|X - \mathcal{L}\|_S^2 \geqq c], \tag{8.3}$$

the second inequality following from Lemma 8.2 ($\mathcal{L}$ is convex). Applying Lemma 4.2 with $A = \Sigma^{-\frac{1}{2}}$, $P_{0,\Sigma}[\|X - \mathcal{L}\|_S^2 \geqq c] = P[\|Z - \mathcal{L}^*\|_T^2 \geqq c]$, where $Z = \Sigma^{-\frac{1}{2}}X \sim \mathfrak{N}(0, I)$, $T = \Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}} \sim \mathcal{W}(N - 1, I)$, and $\mathcal{L}^* = \Sigma^{-\frac{1}{2}}(\mathcal{L})$ is again a halfline. By means of an orthogonal transformation we may take $\mathcal{L}^* = \{x{:}x_1 \geqq 0, x_2 = \cdots = x_p = 0\}$. Partition Z and T as follows: $Z' = (Z_1, Z_2')$, $T = (T_{ij})$ $(i, j = 1, 2)$, where Z_1 and T_{11} are both scalars. Then using (2.1) we easily find that $\|Z - \mathcal{L}^*\|_T^2 = Z_2'T_{22}^{-1}Z_2$ if $Z_1 - T_{12}T_{22}^{-1}Z_2 \geqq 0$ and $\|Z - \mathcal{L}^*\|_T^2 = Z'T^{-1}Z$ if $Z_1 - T_{12}T_{22}^{-1}Z_2 < 0$. Therefore, since $1 + Z_2'T_{22}^{-1}Z_2$ and $T_{11\cdot2}$ are positive scalars,

$$\begin{aligned} &P[\|Z - \mathcal{L}^*\|_T^2 \geqq c] \\ &\quad = P[Z_2'T_{22}^{-1}Z_2 \geqq c, (Z_1 - T_{12}T_{22}^{-1}Z_2)/(1 + Z_2'T_{22}^{-1}Z_2)^{\frac{1}{2}} \geqq 0] \\ &\qquad + P[Z'T^{-1}Z \geqq c, (Z_1 - T_{12}T_{22}^{-1}Z_2)T_{11.2}^{-1} < 0]. \end{aligned}$$

By 3.7° and 3.5°, the first term on the right is equal to

$$P[Z_2'T_{22}^{-1}Z_2 \geqq c]P[Z_1 - T_{12}T_{22}^{-1}Z_2 \geqq 0] = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p+1}^2 \geqq c].$$

Since $(Z_1 - T_{12}T_{22}^{-1}Z_2)T_{11.2}^{-1}$ is the first component of $T^{-1}Z$, Lemma 3.2(ii) implies that the second term on the right is equal to

$$P[Z'T^{-1}Z \geqq c]P[Z_1 - T_{12}T_{22}^{-1}Z_2 < 0] = \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c].$$

Thus we conclude that

$$P_{0,\Sigma}[\|X - \mathcal{L}\|_S^2 \geqq c] = \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p+1}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c]. \tag{8.4}$$

To complete the proof of the theorem it suffices to show that

$$\sup_{\{\Sigma>0\}} P_{0,\Sigma}[\|X - \mathcal{C}\|_S^2 \geqq c] \geqq \tfrac{1}{2}P[\chi_{p-1}^2/\chi_{N-p+1}^2 \geqq c] + \tfrac{1}{2}P[\chi_p^2/\chi_{N-p}^2 \geqq c]. \tag{8.5}$$

As in the proof of Theorem 6.2, there exist a halfline $\mathcal{L}$ and a sequence $\{A_n\}$ of nonsingular linear transformations such that $\bigcap_{n=1}^{\infty} A_n(\mathcal{C}) = \mathcal{L}$. By Lemma 4.3(i) this implies that $\|X - A_n(\mathcal{C})\|_S^2 \to \|X - \mathcal{L}\|_S^2$ as $n \to \infty$. Therefore, with $\Sigma_n = (A_n'A_n)^{-1}$ we apply Lemma 4.2 to obtain

$$P_{0,\Sigma_n}[\|X - \mathcal{C}\|_S^2 \geqq c] = P_{0,I}[\|X - A_n(\mathcal{C})\|_S^2 \geqq c] \to P_{0,I}[\|X - \mathcal{L}\|_S^2 \geqq c]. \quad \square$$

REMARK. It should be noted that it is not necessarily true that for Σ fixed,

$$\sup_{\{\mu\varepsilon\mathcal{C}\}} P_{\mu,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c] = P_{0,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c].$$

By Lemma 8.2, however, this is true if $\mathcal{C}$ is convex.

THEOREM 8.4. *If the cone $\mathcal{C}$ contains a p-dimensional open set, then for any fixed* $\Sigma > 0$

$$\inf_{\{\mu\varepsilon\mathcal{C}\}} P_{\mu,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) \geqq c] = 0.$$

PROOF. We shall prove the stronger result that for Σ fixed,

$$\sup_{\{\mu\varepsilon\mathcal{C}\}} P_{\mu,\Sigma}[U(\mathcal{C}, \mathcal{E}_p) = 0] = 1. \tag{8.6}$$

Note that $\{U(\mathcal{C}, \mathcal{E}_p) = 0\} = \{\|X - \mathcal{C}\|_S^2 = 0\} = \{X \varepsilon \mathcal{C}\}$. Let $\mathcal{P}_1 = \mathcal{E}_p - \mathcal{C}$ (the complement of $\mathcal{C}$) and $\mathcal{P}_2 = \mathcal{E}_p$. Since $\mathcal{C}$ contains an open set, there exists $\mu_0 \varepsilon \mathcal{C}$ such that $\|\mu_0 - \mathcal{P}_1\|_\Sigma \equiv \delta > 0$. Then $\mu_n = n\mu_0 \varepsilon \mathcal{C}$ satisfies $\|\mu_n - \mathcal{P}_1\|_\Sigma = n\delta \to \infty$. It follows from the proof of Theorem 5.3 (in particular, from (5.7) with $U(\mathcal{P}_1, \mathcal{P}_2)$ replaced by $\|X - \mathcal{P}_1\|_S$) that $\|\mu_n - \mathcal{P}_1\|_\Sigma \to \infty$ implies $P_{\mu_n,\Sigma}[\|X - \mathcal{P}_1\|_S \geqq d] \to 1$, for any positive number d. Since $\|X - \mathcal{P}_1\|_S \geqq d$ implies $X \varepsilon \mathcal{C}$, this implies (8.6). $\square$

To conclude this section we consider the problem of testing a one-sided hypothesis involving a subset of the components of the mean μ. Precisely, using the notation introduced in the paragraph preceding (6.12), we consider the problem of testing $H:\mu \varepsilon \mathcal{P}_1$ vs. $K:\mu \varepsilon \mathcal{P}_2$ where now

$$\mathcal{P}_1 = \{\mu:\mu_1 \varepsilon \mathcal{E}_q, \mu_2 \varepsilon \mathcal{C}, \mu_3 = 0\}, \tag{8.7}$$
$$\mathcal{P}_2 = \{\mu:\mu_1 \varepsilon \mathcal{E}_q, \mu_2 \varepsilon \mathcal{E}_r, \mu_3 = 0\}.$$

The LRT statistic $U(\mathcal{P}_1, \mathcal{P}_2)$ can be put into the form

$$U(\mathcal{P}_1, \mathcal{P}_2) = \|Z - \pi_T(Z; \mathcal{C})\|_T^2 = \|Z - \mathcal{C}\|_T^2$$

where $Z = (X_2 - S_{23}S_{33}^{-1}X_3)/(1 + X_3'S_{33}^{-1}X_3)^{\frac{1}{2}}$ and $T = S_{22\cdot 3} \sim \mathcal{W}(N - 1 - s, \Sigma_{22\cdot 3})$. By 3.2°, Z and T are independent. It is shown below that the distribution of $U(\mathcal{P}_1, \mathcal{P}_2)$ depends only on μ_2 and $\Sigma_{22\cdot 3}$.

THEOREM 8.5. *Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be given by* (8.7) *with $\mathcal{C}$ a cone in $\mathcal{E}_r$, and suppose that the hypothesis $H:\mu \varepsilon \mathcal{P}_1$ is true. Then*

$$(8.8) \quad \sup_{\{\mu_2 \varepsilon \mathcal{C}, \Sigma_{22\cdot 3}>0\}} P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = \tfrac{1}{2}P[\chi_{r-1}^2/\chi_{N-s-r-1}^2 \geqq c] + \tfrac{1}{2}P[\chi_r^2/\chi_{N-s-r}^2 \geqq c] = \sup_{\{\Sigma_{22\cdot 3}>0\}} \mathcal{P}_{\mu^*,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c],$$

where μ^* is any point in $\mathcal{P}_1$ with $\mu_2^* = 0$. If in addition $\mathcal{C}$ contains an r-dimensional open set, then for any fixed $\Sigma > 0$

$$(8.9) \qquad \inf_{\{\mu_2 \varepsilon \mathcal{C}\}} P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = 0.$$

PROOF. Trivially

$$(8.10) \quad \sup_{\{\Sigma_{22\cdot 3}>0\}} P_{\mu^*,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] \leqq \sup_{\{\mu_2 \varepsilon \mathcal{C}, \Sigma_{22\cdot 3}>0\}} P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c].$$

Next, $\mu_3 = 0$ for any $\mu \in \mathcal{P}_1$, so from 3.7° the conditional distribution of Z given (X_3, S_{33}) is $\mathfrak{N}(\nu, \Sigma_{22\cdot 3})$, where $\nu \equiv \mu_2/(1 + X_3'S_{33}^{-1}X_3)^{\frac{1}{2}}$ is in $\mathcal{C}$ if and only if $\mu_2 \varepsilon \mathcal{C}$. Therefore, conditional on (X_3, S_{33}), the distribution of $U(\mathcal{P}_1, \mathcal{P}_2)$ is identical with that of $U(\mathcal{C}, \mathcal{E}_p)$, with $\Sigma_{22\cdot 3}$ substituted for Σ, r for p, $N - s$ for N, and ν for μ. Hence, by Theorem 8.3, $P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c \mid X_3, S_{33}] \leqq \beta$ for any $\mu \varepsilon \mathcal{P}_1$, where β is the middle expression in (8.8), and so

$$(8.11) \qquad \sup_{\{\mu_2 \varepsilon \mathcal{C}, \Sigma_{22\cdot 3}>0\}} P_{\mu,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] \leqq \beta.$$

Finally, if $\mu^* \varepsilon \mathcal{P}_1$ is such that $\mu_2^* = 0$ then the conditional distribution of Z is $\mathfrak{N}(0, \Sigma_{22\cdot 3})$ so the unconditional distribution of $U(\mathcal{P}_1, \mathcal{P}_2)$ is identical with that of $U(\mathcal{C}, \mathcal{E}_p)$ (with the proper substitution of parameters). Therefore, by Theorem 8.3

$$\sup_{\{\Sigma_{22\cdot 3}>0\}} P_{\mu^*,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = \beta.$$

Now assume $\mathcal{C}$ contains an open set and fix $\Sigma > 0$. It follows by an argument similar to the proof of Theorem 8.4 and by the above remarks concerning the conditional distribution of $U(\mathcal{P}_1, \mathcal{P}_2)$ that there exists a sequence $\{\mu_n\}$, $\mu_n \varepsilon \mathcal{P}_1$, such that $P_{\mu_n,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c \mid X_3, S_{33}] \to 0$ as $n \to \infty$. Hence, by the bounded convergence theorem (noting that the distribution of (X_3, S_{33}) does not depend on μ_n)

$$P_{\mu_n,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c] = E\{P_{\mu_n,\Sigma}[U(\mathcal{P}_1, \mathcal{P}_2) \geqq c \mid X_3, S_{33}]\} \to 0. \qquad \square$$

9. The Schaafsma-Smid approach. In this section familiarity with the terminology of Schaafsma and Smid (1966) is assumed.

A straightforward application of the methods in Chapters 4 and 5 of Lehmann (1959) leads to the following characterization of the family D of all S.M.P.

similar level α tests for the problem of testing $H:\mu = 0$ vs. $K:\mu \varepsilon \mathcal{C}$ with Σ unknown.

LEMMA 9.1. *For any non-zero* $\mu_0 \varepsilon \mathcal{C}$ *and any* $\Sigma_0 > 0$, *the most powerful similar level* α *test for testing* $H:\mu = 0, \Sigma > 0$ vs. *the simple alternative* $\mu = \mu_0$, $\Sigma = \Sigma_0$ *is the one-sided Student's-t test, denoted by* $\phi(\xi_0)$, *which rejects H if* $(N-1)^{\frac{1}{2}}\xi_0'X/(\xi_0'S\xi_0)^{\frac{1}{2}} \geqq t_{N-1;\alpha}$, *where* $\xi_0 = \Sigma_0^{-1}\mu_0$ *and* $t_{N-1;\alpha}$ *is the upper* α*-percentage point of the Student's-t distribution with* $N-1$ *degrees of freedom.* $\phi(\xi_0)$ *depends on* (μ_0, Σ_0) *only through* $\xi_0/\|\xi_0\|$.

When the true parameter values are (μ, Σ),

$$(N-1)^{\frac{1}{2}}\xi_0'X/(\xi_0'S\xi_0)^{\frac{1}{2}} \sim t'_{N-1}[\xi_0'\mu/(\xi_0'\Sigma\xi_0)^{\frac{1}{2}}],$$

where $t'_{N-1}(\lambda)$ denotes a non-central Student's-t variate with non-centrality parameter λ. Define $M(\lambda)$ for λ real by $M(\lambda) = P[t'_{N-1}(\lambda) \geqq t_{N-1;\alpha}]$. $M(\lambda)$ is a continuous, strictly increasing function of λ with $M(-\infty) = 0$, $M(0) = \alpha$, $M(\infty) = 1$. The power $P_{\mu,\Sigma}[\phi(\xi_0)$ rejects $H]$ of the test $\phi(\xi_0)$ at the alternative point (μ, Σ) is given by $M[\xi_0'\mu/(\xi_0'\Sigma\xi_0)^{\frac{1}{2}}]$. Since the test $\phi(\xi)$ is most powerful similar at the alternative point (μ, Σ), where $\xi = \Sigma^{-1}\mu$, the envelope power function is $M[(\mu'\Sigma^{-1}\mu)^{\frac{1}{2}}]$. Note that this approaches one if and only if $\mu'\Sigma^{-1}\mu \to \infty$.

THEOREM 9.2. *Suppose the cone* $\mathcal{C}$ *contains two or more distinct halflines. For any* S.M.P. *similar level* α *test* $\phi(\xi_0)$, *there exists a sequence* $\{(\mu_n, \Sigma_n)\}$ *of alternative points* $(\mu_n \varepsilon \mathcal{C}, \Sigma_n > 0)$ *such that as* $n \to \infty$

(i) $\mu_n'\Sigma_n^{-1}\mu_n \to \infty$,

(ii) $P_{\mu_n,\Sigma_n}[\phi(\xi_0)$ *rejects* $H] \to 0$ *or* α.

PROOF. Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be distinct halflines contained in $\mathcal{C}$ and let ν_i be any non-zero element in $\mathcal{L}_i$. Either $\xi_0 \not\varepsilon \mathcal{L}_1$ or $\xi_0 \not\varepsilon \mathcal{L}_2$ (or both). Suppose $\xi_0 \not\varepsilon \mathcal{L}_1$. If $\xi_0'\nu_1 < 0$, then with $\mu_n = n\nu_1$ and $\Sigma_n = \Sigma$ (any fixed $\Sigma > 0$) we have $\mu_n'\Sigma_n^{-1}\mu_n = n^2\nu_1'\Sigma^{-1}\nu_1 \to \infty$ and $P_{\mu_n,\Sigma_n}[\phi(\xi_0)$ rejects $H] = M[n\xi_0'\nu_1/(\xi_0'\Sigma\xi_0)^{\frac{1}{2}}] \to 0$. If $\xi_0'\nu_1 \geqq 0$, it is easy to construct a sequence $\{\Sigma_n\}$ such that $\nu_1'\Sigma_n^{-1}\nu_1 \to \infty$ and $\xi_0'\Sigma_n\xi_0 \to \infty$. Then with $\mu_n = \nu_1$ for all n, $\mu_n'\Sigma_n^{-1}\mu_n \to \infty$ and $P_{\mu_n,\Sigma_n}[\phi(\xi_0)$ rejects $H] = M[\xi_0'\nu_1/(\xi_0'\Sigma_n\xi_0)^{\frac{1}{2}}] \to \alpha$. In either case, (i) and (ii) hold. If $\xi_0 \varepsilon \mathcal{L}_2$, replace ν_1 by ν_2 in the preceding argument. □

Since $\mu'\Sigma^{-1}\mu$ is the natural measure of distance from the alternative point (μ, Σ) to the hypothesis H, this result states that for any S.M.P. similar test $\phi(\xi_0)$ there exist alternative points located arbitrarily far from H, which cannot be distinguished from H by $\phi(\xi_0)$. By contrast, Theorem 5.3 implies that the power of the LRT for this problem approaches one uniformly as $\mu'\Sigma^{-1}\mu \to \infty$. We conclude that in this testing situation the LRT is preferable to any S.M.P. similar test.

Acknowledgment. I wish to thank Professor Ingram Olkin for calling my attention to this problem area and for his guidance, criticism, and encouragement.

REFERENCES

ABELSON, R. P. and TUKEY, J. W. (1963). Efficient utilization of non-numerical information in quantitative analysis: general theory and the case of simple order. *Ann. Math. Statist.* **34** 1347–1369.

ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

BARTHOLOMEW, D. J. (1959a, b). A test of homogeneity for ordered alternatives, I, II. *Biometrika* **46** 36–48, 328–335.

BARTHOLOMEW, D. J. (1961a). A test of homogeneity of means under restricted alternatives. *J. Roy. Statist. Soc. Ser. B* **23** 239–281.

BARTHOLOMEW, D. J. (1961b). Ordered tests in the analysis of variance. *Biometrika* **48** 325–332.

CHACKO, V. J. (1963). Testing homogeneity against ordered alternatives. *Ann. Math. Statist.* **34** 945–956.

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

KUDÔ, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* **50** 403–418.

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

MILES, R. E. (1959). The complete amalgamation into blocks, by weighted means, of a finite set of real numbers. *Biometrika* **46** 317–327.

NÜESCH, P. E. (1964). Multivariate tests of location for restricted alternatives. Doctoral dissertation, Swiss Federal Institute of Technology, Zurich.

NÜESCH, P. E. (1966). On the problem of testing location in multivariate problems for restricted alternatives. *Ann. Math. Statist.* **37** 113–119.

SCHAAFSMA, W. and SMID, L. J. (1966). Most stringent somewhere most powerful tests against alternatives restricted by a number of linear inequalities. *Ann. Math. Statist.* **37** 1161–1172.

SHORACK, G. R. (1967). Testing against ordered alternatives in model I analysis of variance; normal theory and nonparametric. *Ann. Math. Statist.* **38** 1740–1753.

STEIN, C. M. (1966). Multivariate analysis (mimeographed notes, recorded by M. L. Eaton). Department of Statistics, Stanford University.