A Test of a Multivariate Normal Mean with Composite Hypotheses Determined by Linear Inequalities

Author(s): S. Sasabuchi

Source: *Biometrika*, Aug., 1980, Vol. 67, No. 2 (Aug., 1980), pp. 429-439

Published by: Oxford University Press on behalf of Biometrika Trust

Stable URL: https://www.jstor.org/stable/2335486

Biometrika (1980), **67**, 2, *pp.* 429–39
Printed in Great Britain

A test of a multivariate normal mean with composite hypotheses determined by linear inequalities

BY S. SASABUCHI
Department of Mathematics, Kyushu University, Fukuoka, Japan

SUMMARY

In this paper we propose a new multivariate generalization of a one-sided test in a way different from that of Kudô (1963). Let X be a p-variate normal random variable with the mean vector μ and a known covariance matrix. We consider the null hypothesis that μ lies on the boundary of a convex polyhedral cone determined by linear inequalities; the alternative is that μ lies in its interior. A two-sided version is also discussed. This paper provides likelihood ratio tests and some applications along with some discussion of the geometry of convex polyhedral cones.

Some key words: Convex polyhedral cone; Likelihood ratio test; Linear inequalities; Multivariate normal distribution; Order restriction.

1. INTRODUCTION

Let $X = (X_1, \ldots, X_p)'$ be a p-variate normal random variable with mean vector $\mu = (\mu_1, \ldots, \mu_p)'$ and known covariance matrix Σ. We are concerned with a testing problem where both null and alternative hypotheses are determined by k linear inequalities $b_i' \mu \geqslant 0$ $(i = 1, \ldots, k)$, where b_i's are p-dimensional vectors and k may be different from p.

The pioneering work on this type of problem is by Bartholomew (1959a, b), where Σ is diagonal, the null hypothesis is determined by equalities and the alternatives by inequalities. Kudô (1963), whose work is closely related to Bartholomew's, considered the null hypothesis $H\colon \mu = 0$, or that the mean vector is at the origin, and the alternative hypothesis

$$K\colon \mu_i \geqslant 0 \quad (i = 1, \ldots, p), \quad \max_{1 \leqslant i \leqslant p} \mu_i > 0,$$

or that the mean vector is in the positive orthant. A generalization of this problem is discussed by Kudô & Choi (1975) where the alternative hypothesis is that μ lies in a convex cone determined by $k \geqslant p$ linear inequalities. All this work concerns a multivariate generalization of the one-sided test.

In this paper, we propose a new version of the multivariate generalization. A prototype is the testing problem with the null hypothesis $H_0\colon \mu_i \geqslant 0$ $(i = 1, \ldots, p)$, where equality holds for at least one value of i, and the alternative $K_0\colon \mu_i > 0$ $(i = 1, \ldots, p)$. The investigation of this was initiated by Inada (1978), where the prototype problem is treated in the bivariate case. We treat the general problem where the null hypothesis corresponds to the boundary of the convex polyhedral cone determined by the linear inequalities and the alternative corresponds to its interior.

Here we state the geometrical description of the test in the bivariate case, which has a strong intuitive appeal. Let X_1, X_2 be normally independently distributed with means μ_1 and μ_2 and common unit variance. Given a pair of half-lines l_1 and l_2 starting from the origin and forming an angle θ, the null hypothesis is that (μ_1, μ_2) lies on either l_1 or l_2, and the alternative is that it lies inside the angle θ. The likelihood ratio test can easily be derived and the critical region is the angle formed by shifting the original angle along its bisector as shown in Fig. 1a, where l_i' is parallel to l_i $(i = 1, 2)$ and the distance between l_1 and l_1' is equal to that between l_2 and l_2'. The least favourable case in the null hypothesis is the limiting one where (μ_1, μ_2) tends to infinity along either l_1 or l_2. The distance between l_1 and l_1' and also between l_2 and l_2' can easily be computed from the univariate normal distribution function.

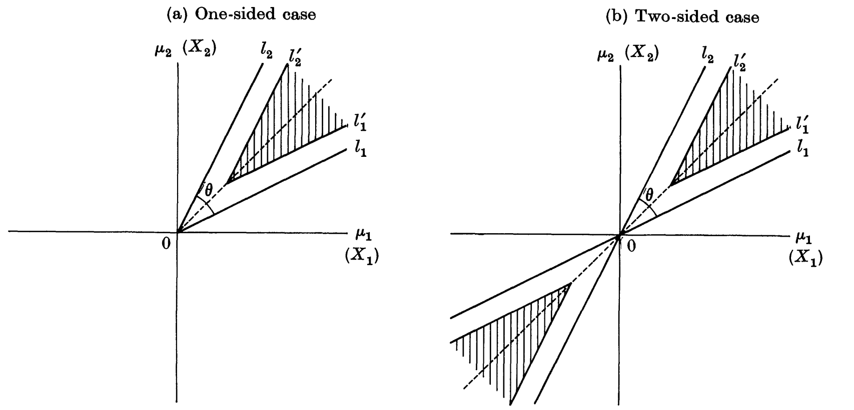

Fig. 1. Critical region superimposed on parameter space (a) for one-sided case, (b) two-sided case; H: (μ_1, μ_2) lies on either l_1 or l_2 and K: (μ_1, μ_2) lies inside the angle θ formed by l_1 and l_2; critical region is shaded angle; broken line is bisector of angle θ.

This problem can be generalized to the two-sided case by extending the pair of half-lines to the pair of lines crossing at the origin. The likelihood ratio test and the critical region can be found quite analogously; see Fig. 1b. Unfortunately the least favourable case can be determined similarly only when $\theta \leqslant \frac{1}{2}\pi$.

In § 2, notation and preparatory discussion of convex cones are given, which are necessary because we may have the number, k, of inequalities determining the convex cone larger than p. An example is the convex cone corresponding to the simple loop restriction, that is, $\mu_1 \leqslant \mu_i \leqslant \mu_p$ $(i = 2, \ldots, p-1)$. Section 3 is devoted to the determination of the likelihood ratio test. The critical region can be expressed in a strikingly simple form and the percentage points can be tabulated from readily available tables of the standard normal distribution. The two-sided version is discussed in § 4, where we show that under certain conditions on the linear inequalities the percentage point can be computed similarly. In § 5, some testing problems determined by the order restrictions of Barlow *et al.* (1972, Chapter 3) are discussed as applications. Further generalizations are given in § 6, where the assumption of known variance matrix is relaxed.

2. Notation and preliminaries

Let x be a vector in R^p, p-dimensional Euclidean space, where the Euclidean norm $\|x\| = \surd(x'x)$ is defined, and, for a set S in R^p, let $-S$ be the set $\{x\colon -x \in S\}$.

Let $N_p(\mu, \Sigma)$ be a p-variate normal distribution with mean vector μ and covariance matrix Σ. A testing problem with null hypothesis H and alternative K is denoted by $[H, K]$. We deal with testing problems where H and K are determined by inequalities, and for simplicity we use the notation that $\geqslant_*$ means that at least one equality holds and $\geqslant^*$ means that at least one strict inequality holds.

Let $\{a_1, \ldots, a_k\}$ be a set of vectors in R^p, and let $\mathscr{J}$ be a convex polyhedral cone defined by $\{a_1, \ldots, a_k\}$, that is,

$$\mathscr{J} = \{y;\ a_i' y \geqslant 0\ (i = 1, \ldots, k)\}. \tag{2·1}$$

The boundary and the interior of $\mathscr{J}$ are

$$\partial \mathscr{J} = \{y;\ a_i' y \geqslant_* 0\ (i = 1, \ldots, k)\}, \quad \mathscr{J}^0 = \{y;\ a_i' y > 0\ (i = 1, \ldots, k)\},$$

respectively.

The boundary has k faces F_j;

$$F_j = \{y;\ a_j' y = 0, a_i' y \geqslant 0, i \in \{1, \ldots, k\} - \{j\}\} \quad (j = 1, \ldots, k), \tag{2·2}$$

some of which may be of dimension less than $p-1$, or equivalently the relative interior of F_j;

$$\mathscr{J}_j = \{y;\ a_j' y = 0, a_i' y > 0, i \in \{1, \ldots, k\} - \{j\}\} \quad (j = 1, \ldots, k) \tag{2·3}$$

may be empty for some j.

We do not impose restrictions on k and p. Indeed, if $k < p$, $\mathscr{J}$ may be a half-plane or a half-space.

Generally there may be redundant vectors in $\{a_1, \ldots, a_k\}$, and this will create technical difficulties. Also $\mathscr{J}^0$ may be empty, and in this case the alternative hypotheses become vacuous. In order to exclude these cases, we prepare ourselves with some definitions and theorems.

Definition 2·1. A vector a_j is said to be redundant in $\{a_1, \ldots, a_k\}$ if

$$\{y;\ a_i' y \geqslant 0\ (i = 1, \ldots, k)\} = \{y;\ a_i' y \geqslant 0,\ i \in \{1, \ldots, k\} - \{j\}\},$$

or equivalently, there does not exist a vector y such that $a_i' y \geqslant 0$ $(i \neq j)$, $a_j' y < 0$.

The following is adopted from the definition of Davis (1954). The trivial case $\lambda_i = 0$ $(i = 1, \ldots, k)$ is excluded from the definition.

Definition 2·2. A set of vectors $\{a_1, \ldots, a_k\}$ is said to be with positive relations if there exist nonnegative numbers $\lambda_1, \ldots, \lambda_k$, not all simultaneously zero, such that $\Sigma_i \lambda_i a_i = 0$; otherwise the set is said to be without positive relations.

Remark 2·1. If $\{a_1, \ldots, a_k\}$ is with positive relations, it is linearly dependent; equivalently, if $\{a_1, \ldots, a_k\}$ is linearly independent, it is without positive relations.

Now we state a theorem which will play an important role in §§ 3 and 4.

THEOREM 2·1. *Suppose that $\{a_1, \ldots, a_k\}$ is without positive relations and has no redundant vector in it. Then $\mathscr{J}^0 \neq \phi$, and all the faces F_j $(j = 1, \ldots, k)$ are exactly $p-1$ dimensional and their relative interiors are nonempty, that is $\mathscr{J}_j \neq \phi$ $(j = 1, \ldots, k)$.*

This theorem is a direct consequence of the three theorems in the Appendix.

Remark 2·2. The second statement of the theorem ensures that we can make all $a_i' y$ $(i \neq j)$ simultaneously infinite, keeping $a_j' y = 0$.

3. ONE-SIDED TEST

Let $X = (X_1, \ldots, X_p)'$ be a random vector distributed as $N_p(\mu, \Sigma)$, where Σ is a known nonsingular matrix. Although Σ is either a unit or a diagonal matrix in most of the applications, it is convenient to work in the fairly general framework.

Let $B = (b_1, \ldots, b_k)'$ be a given $k \times p$ matrix which satisfies the following assumption.

Assumption 1. $\{b_1, \ldots, b_k\}$ is without positive relations and has no redundant vector in it.

We consider the following testing problem:

$$\text{null hypothesis, } H_1;\; b_i'\mu \geqslant_* 0 \quad (i = 1, \ldots, k),$$
$$\text{alternative, } K_1;\; b_i'\mu > 0 \quad (i = 1, \ldots, k).$$

Remark 3·1. Any set of k linearly independent vectors satisfies Assumption 1, as guaranteed by Remark 2·1 and Theorem A·2, and $[H_0, K_0]$ in § 1 is a special case of $[H_1, K_1]$.

Let T be a $p \times p$ nonsingular matrix such that $T\Sigma T' = I_p$, that is $\Sigma = T^{-1}(T^{-1})'$, and make the transformation

$$Y = TX, \quad \theta = T\mu. \tag{3·1}$$

Then Y is a random vector distributed as $N_p(\theta, I_p)$.

Define a $k \times p$ matrix A as

$$A = \begin{pmatrix} a_1' \\ \vdots \\ a_k' \end{pmatrix} = \begin{pmatrix} b_1'T^{-1} \\ \vdots \\ b_k'T^{-1} \end{pmatrix} = BT^{-1}. \tag{3·2}$$

The set of vectors $\{a_1, \ldots, a_k\}$ satisfies Assumption 1, and Theorem 2·1 is valid for the convex polyhedral cone $\mathscr{J}$ defined by (2·1).

We have $b_i'\mu = a_i'\theta$ $(i = 1, \ldots, k)$, and hence the problem $[H_1, K_1]$ is transformed to the following problem $[H_1', K_1']$;

$$H_1'\colon a_i'\theta \geqslant_* 0 \quad (i = 1, \ldots, k), \quad K_1'\colon a_i'\theta > 0 \quad (i = 1, \ldots, k).$$

Now H_1' and K_1' can be expressed as $\theta \in \partial\mathscr{J}$ and $\theta \in \mathscr{J}^0$, respectively. In this paper we identify a hypothesis with the corresponding set of parameters. For example, we write $H_1' = \partial\mathscr{J}$, $K_1' = \mathscr{J}^0$ and $H_1' \cup K_1' = \mathscr{J}$.

The likelihood function is $L = C_0 \exp(-\frac{1}{2}\| Y - \theta \|^2)$, where C_0 is a positive constant independent of θ. The likelihood ratio for the problem $[H_1, K_1]$ or $[H_1', K_1']$ is given by

$$\lambda = \frac{\sup\limits_{H_1'} L}{\sup\limits_{H_1' \cup K_1'} L} = \frac{\sup\limits_{\theta \in \partial\mathscr{J}} \exp(-\frac{1}{2}\| Y - \theta \|^2)}{\sup\limits_{\theta \in \mathscr{J}} \exp(-\frac{1}{2}\| Y - \theta \|^2)} = \frac{\exp(-\frac{1}{2}\inf\limits_{\theta \in \partial\mathscr{J}} \| Y - \theta \|^2)}{\exp(-\frac{1}{2}\inf\limits_{\theta \in \mathscr{J}} \| Y - \theta \|^2)}.$$

Because

$$\inf_{\theta \in \mathscr{J}} \| Y - \theta \|^2 = 0 \quad (Y \in \mathscr{J}^0), \quad \inf_{\theta \in \mathscr{J}} \| Y - \theta \|^2 = \inf_{\theta \in \partial\mathscr{J}} \| Y - \theta \|^2 \quad (Y \notin \mathscr{J}^0),$$

we have that

$$\lambda = \begin{cases} \exp(-\frac{1}{2}\inf\limits_{\theta \in \partial\mathscr{J}} \| Y - \theta \|^2) & (Y \in \mathscr{J}^0), \\ 1 & (Y \notin \mathscr{J}^0), \end{cases}$$

and the critical region can be written as

$$Y \in \mathscr{J}^0, \quad \inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2 \geqslant s_1^2, \tag{3·3}$$

where s_1 is a positive constant. We note that

$$\partial \mathscr{J} = \bigcup_{i=1}^{k} F_i, \tag{3·4}$$

where $F_i = \{y;\, a_i' y = 0, a_j' y \geqslant 0 \ (j = 1, \ldots, k)\}$. Let B_i be the smallest subspace containing F_i, and, for a given point Y, let $\hat{Y}_i$ denote the projection of Y onto B_i. As F_i is assured to be exactly $p-1$ dimensional faced by Theorem 2·1, B_i and Y_i are expressed as

$$B_i = \{y;\, a_i' y = 0\}, \quad \hat{Y}_i = Y - \frac{a_i' Y}{a_i' a_i} a_i, \tag{3·5}$$

respectively.

Now the following lemma is important.

Lemma 3·1. *For $Y \in \mathscr{J}^0$, the following equality holds*:

$$\inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2 = \min_{1 \leqslant i \leqslant k} \| Y - \hat{Y}_i \|^2. \tag{3·6}$$

Remark 3·2. Now $\hat{Y}_i$ is the closest point of Y to B_i, but it is not necessarily the closest point of Y to F_i, and hence the lemma is not trivial. For $Y \notin \mathscr{J}$, (3·6) does not necessarily hold.

Proof of Lemma 3·1. First, from (3·4) and the fact that B_i contains F_i, we have that

$$\begin{aligned} \inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2 &= \min_{1 \leqslant i \leqslant k} \inf_{\theta \in F_i} \| Y - \theta \|^2 \\ &\geqslant \min_{1 \leqslant i \leqslant k} \inf_{\theta \in B_i} \| Y - \theta \|^2 \\ &= \min_{1 \leqslant i \leqslant k} \| Y - \hat{Y}_i \|^2. \end{aligned}$$

Secondly we prove the inverse inequality. We consider the following two cases. If $\hat{Y}_i \in F_i$, then $\hat{Y}_i \in \partial \mathscr{J}$, and hence we have that

$$\inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2 \leqslant \| Y - \hat{Y}_i \|^2.$$

If $\hat{Y}_i \notin F_i$, then $\hat{Y}_i \notin \mathscr{J}$, and hence, from the assumption that $Y \in \mathscr{J}^0$, there exists a vector $Y_0 \in \partial \mathscr{J}$ such that $\| Y - Y_0 \|^2 \leqslant \| Y - \hat{Y}_i \|^2$. Therefore

$$\begin{aligned} \inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2 \leqslant \| Y - Y_0 \|^2 &\leqslant \| Y - \hat{Y}_i \|^2 \\ &\leqslant \min_{1 \leqslant i \leqslant k} \| Y - \hat{Y}_i \|^2. \end{aligned}$$

From (3·3), (3·5) and Lemma 3·1, the critical region is given by

$$a_i' Y > 0 \quad (i = 1, \ldots, k), \quad \min_{1 \leqslant i \leqslant k} \left\| \frac{a_i' Y}{a_i' a_i} a_i \right\|^2 \geqslant s_1^2,$$

which is equivalent to

$$a_i' Y / \| a_i \| \geqslant s_1 \quad (i = 1, \ldots, k). \tag{3·7}$$

Finally, from (3·1) and (3·2), we have $a_i' Y = b_i' X$, $\| a_i \|^2 = b_i' \Sigma b_i$ for $i = 1, \ldots, k$, and hence (3·7) is equivalent to

$$b_i' X / \sqrt{(b_i' \Sigma b_i)} \geqslant s_1 \quad (i = 1, \ldots, k). \tag{3·8}$$

The next problem is to determine the constant s_1 to achieve significance level α.

Let us put

$$U_i = b_i'(X-\mu)/\sqrt{(b_i' \Sigma b_i)} = a_i'(Y-\theta)/\| a_i \| \quad (i = 1, \ldots, k). \tag{3·9}$$

Then $U = (U_1, \ldots, U_k)'$ is distributed as $N_k(0, R)$, where R is the correlation matrix of $BX = AY$, independent of μ, that is θ.

The constant s_1 is determined by the relation

$$\begin{aligned} \alpha &= \sup_{\theta \in \partial \mathscr{I}} \mathrm{pr}\{a_i' Y / \| a_i \| \geqslant s_1 \quad (i = 1, \ldots, k)\} \\ &= \max_{1 \leqslant j \leqslant k} \sup_{\theta \in F_j} \mathrm{pr}\{U_i \geqslant s_1 - a_i' \theta / \| a_i \| \quad (i = 1, \ldots, k)\}. \end{aligned} \tag{3·10}$$

Now $\mathrm{pr}\{U_i \geqslant s_1 - a_i'\theta/\| a_i \| \ (i = 1, \ldots, k)\}$ is monotone increasing in each of $a_1'\theta, \ldots, a_k'\theta$. Because $\{a_1, \ldots, a_k\}$ satisfies Assumption 1, Theorem 2·1 ensures that, given j, we can make all $a_i'\theta$ $(i \neq j)$ arbitrarily large under the constraint $a_j'\theta = 0$; see Remark 2·2. Therefore we have that

$$\alpha = \max_{1 \leqslant j \leqslant k} \mathrm{pr}(U_j \geqslant s_1). \tag{3·11}$$

Because U_j is distributed as $N_1(0,1)$, we get that

$$\alpha = \max_{1 \leqslant j \leqslant k} Q(s_1) = Q(s_1), \tag{3·12}$$

where

$$Q(s_1) = 1 - \Phi(s_1) = \int_{s_1}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2} dt.$$

Summarizing the above, we have the following theorem.

THEOREM 3·1. *The critical region of the likelihood ratio test for the problem* $[H_1, K_1]$ *is given by*

$$b_i' X / \sqrt{(b_i' \Sigma b_i)} \geqslant u(\alpha) \quad (i = 1, \ldots, k),$$

where α *is the significance level and* $u(\alpha)$ *is the upper* α *point of the standard normal distribution.*

Example 3·1: positive orthant. For a random sample $X_1, \ldots, X_n$ from $N_p(\mu, \Sigma)$, with Σ known and nonsingular, consider the testing problem, with $k \leqslant p$,

$$H_{11}: \mu_i \geqslant_* 0 \quad (i = 1, \ldots, k), \quad K_{11}: \mu_i > 0 \quad (i = 1, \ldots, k).$$

As the distribution of the sufficient statistic, $\bar{X} = n^{-1} \Sigma_j X_j$, is $N_p(\mu, n^{-1}\Sigma)$, Theorem 3·1 provides the critical region $\bar{X}_i \sqrt{n} / \sqrt{\sigma_{ii}} \geqslant u(\alpha)$ $(i = 1, \ldots, k)$, where $\bar{X} = (\bar{X}_1, \ldots, \bar{X}_p)'$, $\Sigma = ((\sigma_{ij}))$. When $p = k = 1$, this reduces to the ordinary one-sided test in the univariate case.

4. Two-sided test

In this section, we adopt the same notation and assumptions as in § 3, and identify a hypothesis with the corresponding set of parameters. The hypotheses considered in this section are

$$H_2 = H_1 \cup (-H_1), \quad K_2 = K_1 \cup (-K_1). \tag{4·1}$$

By the transformation (3·1) and (3·2), the problem $[H_2, K_2]$ is transformed to $[H_2', K_2']$, where H_2' and K_2' can be expressed as $\theta \in \partial \mathscr{J} \cup (-\partial \mathscr{J})$ and $\theta \in \mathscr{J}^0 \cup (-\mathscr{J}^0)$, respectively.

The likelihood ratio for the problem $[H_2, K_2]$, or $[H_2', K_2']$, is

$$\lambda = \frac{\sup_{H_2'} L}{\sup_{H_2' \cup K_2'} L} = \frac{\exp\left(-\frac{1}{2} \inf_{\theta \in \partial \mathscr{J} \cup (-\partial \mathscr{J})} \| Y - \theta \|^2\right)}{\exp\left(-\frac{1}{2} \inf_{\theta \in \mathscr{J} \cup (-\mathscr{J})} \| Y - \theta \|^2\right)},$$

$$\lambda = \begin{cases} \exp\left(-\frac{1}{2} \inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2\right) & (Y \in \mathscr{J}^0), \\ \exp\left(-\frac{1}{2} \inf_{\theta \in -\partial \mathscr{J}} \| Y - \theta \|^2\right) & (Y \in -\mathscr{J}^0), \\ 1 & \text{otherwise.} \end{cases}$$

The critical region is given by

$$\{Y \in \mathscr{J}^0, \inf_{\theta \in \partial \mathscr{J}} \| Y - \theta \|^2 \geqslant s_2^2\} \quad \text{or} \quad \{Y \in -\mathscr{J}^0, \inf_{\theta \in -\partial \mathscr{J}} \| Y - \theta \|^2 \geqslant s_2^2\}, \tag{4·2}$$

where s_2 is a positive constant. This is equivalent to

$$\{a_i' Y / \| a_i \| \geqslant s_2 \ (i = 1, \ldots, k)\} \quad \text{or} \quad \{a_i' Y / \| a_i \| \leqslant -s_2 \ (i = 1, \ldots, k)\}. \tag{4·3}$$

Furthermore, this is equivalent to

$$\{b_i' X / \surd(b_i' \Sigma b_i) \geqslant s_2 \ (i = 1, \ldots, k)\} \quad \text{or} \quad \{b_i' X / \surd b_i' \Sigma b_i \leqslant -s_2 \ (i = 1, \ldots, k)\}. \tag{4·4}$$

Remark 4·1. Let $R(s_1)$ denote the critical region for $[H_1, K_1]$, or $[H_1', K_1']$. Then, the critical region for $[H_2, K_2]$, or $[H_2', K_2']$, is the disjoint union, $R(s_2) \cup \{-R(s_2)\}$.

The constant s_2 is determined by the relation

$$\alpha = \sup_{\theta \in \partial \mathscr{J} \cup (-\partial \mathscr{J})} [\operatorname{pr}\{U_i \geqslant s_2 - a_i' \theta / \| a_i \| \ (i = 1, \ldots, k)\} + \operatorname{pr}\{-U_i \geqslant s_2 + a_i' \theta / \| a_i \| \ (i = 1, \ldots, k)\}], \tag{4·5}$$

where U_i $(i = 1, \ldots, k)$ are defined by (3·9). Let

$$f(\tau_1, \ldots, \tau_k;\, s_2) = \operatorname{pr}\{U_i \geqslant s_2 - \tau_i \ (i = 1, \ldots, k)\} + \operatorname{pr}\{-U_i \geqslant s_2 + \tau_i \ (i = 1, \ldots, k)\}. \tag{4·6}$$

Note that f is determined by Σ and B. Then (4·5) can be written as

$$\alpha = \sup_{\theta \in \partial \mathscr{J} \cup (-\partial \mathscr{J})} f(a_1' \theta / \| a_1 \|, \ldots, a_k' \theta / \| a_k \|;\, s_2). \tag{4·7}$$

But both U and $-U$ are distributed as $N_k(0, R)$, and hence, for all $(\tau_1, \ldots, \tau_k;\, s_2)$, we have that $f(-\tau_1, \ldots, -\tau_k;\, s_2) = f(\tau_1, \ldots, \tau_k;\, s_2)$. Therefore (4·7) reduces to

$$\begin{aligned} \alpha &= \sup_{\theta \in \partial \mathscr{J}} f(a_1' \theta / \| a_1 \|, \ldots, a_k' \theta / \| a_k \|; s_2) \\ &= \max_{1 \leqslant j \leqslant k} \sup_{\theta \in F_j} f(a_1' \theta / \| a_1 \|, \ldots, a_k' \theta / \| a_k \|;\, s_2). \end{aligned} \tag{4·8}$$

Especially, if $k = 1$,

$$\alpha = \sup_{a_1'\theta=0} f(a_1'\theta/\|a_1\|;\, s_2) = \mathrm{pr}\,(U_1 \geqslant s_2) + \mathrm{pr}\,(-U_1 \geqslant s_2) = 2Q(s_2).$$

But if $k \geqslant 2$, we cannot find a simple relation between α and s_2.

In order to assure that the least favourable case is the limiting case, we impose the following condition.

Condition 1. For all $s_2 > 0$, $f(\tau_1, \ldots, \tau_k;\, s_2)$ is monotone increasing in each of $\tau_1, \ldots, \tau_k$ on the set $\{(\tau_1, \ldots, \tau_k) \mid \tau_1 \geqslant 0, \ldots, \tau_k \geqslant 0\}$.

Under this condition, we can apply the same argument as in § 3 to get

$$\alpha = \max_{1 \leqslant j \leqslant k} f(\infty, \ldots, \infty, 0, \infty, \ldots, \infty) = \max_{1 \leqslant j \leqslant k} \mathrm{pr}\,(U_j \geqslant s_2) = Q(s_2).$$

Summarizing the above results, we have the following.

THEOREM 4·1. *The critical region of the likelihood ratio test for the problem* $[H_2, K_2]$ *is given by*

$$\{b_i' X/\surd(b_i' \Sigma b_i) \geqslant s_2 \ (i = 1, \ldots, k)\} \quad or \quad \{b_i' X/\surd(b_i' \Sigma b_i) \leqslant -s_2 \ (i = 1, \ldots, k)\},$$

where s_2 *is a positive constant depending on the significance level* α. *When* $k = 1$, $s_2 = u(\tfrac{1}{2}\alpha)$. *When* $k \geqslant 2$, $s_2 = u(\alpha)$, *if the function f defined by* (4·6) *satisfies Condition* 1.

Now we state a theorem which provides a sufficient condition for Condition 1. This condition is, when $p = k = 2$, reduced to that U_1 and U_2 are negatively correlated, or equivalently the angle of $\mathscr{J}$ is less than $\tfrac{1}{2}\pi$, as stated in § 1.

THEOREM 4·2. *Suppose that* $2 \leqslant k \leqslant p$ *and* $\mathrm{rank}\,(B) = k$, *that is* $B\Sigma B'$ *and* R *are nonsingular. Then if all the elements of* $(B\Sigma B')^{-1}$ *are nonnegative, or, equivalently, all the elements of* $R^{-1} = ((\rho^{ij}))$ *are nonnegative, then the function f defined by* (4·6) *satisfies* Condition 1.

Proof. Now

$$f(\tau_1, \ldots, \tau_k;\, s_2) = \int_{s_2-\tau_k}^{\infty} \cdots \int_{s_2-\tau_1}^{\infty} n(u_1, \ldots, u_k)\, du_1 \ldots du_k + \int_{s_2+\tau_k}^{\infty} \cdots \int_{s_2+\tau_1}^{\infty} n(u_1, \ldots, u_k)\, du_1 \ldots du_k, \tag{4·9}$$

where $n(\,.\,)$ is the density function of $N_k(0, R)$.

First we observe that

$$\frac{\partial f}{\partial \tau_k} = \int_{s_2+\tau_{k-1}}^{\infty} \cdots \int_{s_2+\tau_1}^{\infty} g(u_1, \ldots, u_{k-1};\, s_2, \tau_k)\, du_1 \ldots du_{k-1} + \int \underset{\mathscr{D}}{\cdots} \int n(u_1, \ldots, u_{k-1}, s_2 - \tau_k)\, du_1 \ldots du_{k-1}, \tag{4·10}$$

where

$$g(u_1, \ldots, u_{k-1};\, s_2, \tau_k) = n(u_1, \ldots, u_{k-1}, s_2 - \tau_k) - n(u_1, \ldots, u_{k-1}, s_2 + \tau_k), \tag{4·11}$$

$$\mathscr{D} = \{(s_2 - \tau_1, \infty) \times \ldots \times (s_2 - \tau_{k-1}, \infty)\} \cap \{(s_2 + \tau_1, \infty) \times \ldots \times (s_2 + \tau_{k-1}, \infty)\}^C.$$

The second term of (4·10) is nonnegative. For the first term, we have that

$$g(u_1, \ldots, u_{k-1};\, s_2, \tau_k) = C_0 \exp\left[-\frac{1}{2}\left\{(s_2-\tau_k)^2 + 2\sum_{i=1}^{k-1} \rho^{ik} u_i (s_2-\tau_k)\right\}\right]$$

$$- C_0 \exp\left[-\frac{1}{2}\left\{(s_2+\tau_k)^2 + 2\sum_{i=1}^{k-1} \rho^{ik} u_i (s_2+\tau_k)\right\}\right],$$

where $C_0 > 0$, $\tau_i \geqslant 0$ $(i = 1, \ldots, k)$ and $\rho^{ik} \geqslant 0$ $(i = 1, \ldots, k-1)$, and hence $g \geqslant 0$ on the domain of integration, $s_2 + \tau_1 < u_1 < \infty, \ldots, s_2 + \tau_{k-1} < u_{k-1} < \infty$. Thus the first term of (4·10) is nonnegative.

As $\partial f/\partial \tau_k$ is nonnegative, f is monotone increasing in τ_k. The monotonicity of f in the other τ_i $(1 \leqslant i \leqslant k-1)$ is established in the same way.

It is interesting to note the following theorem which seems not to have been clearly stated in the recent literature.

THEOREM 4·3. *Let Λ be a covariance matrix with all covariances nonpositive; then all the elements of Λ^{-1} are nonnegative.*

This is readily proved by induction.

Example 4·1: positive or negative orthant. We adopt the same notation as in Example 3·1, and consider the problem

$$H_{21}\colon \{\mu_i \geqslant_* 0\ (i = 1, \ldots, k)\} \quad \text{or} \quad \{\mu_i \leqslant_* 0\ (i = 1, \ldots, k)\},$$

$$K_{21}\colon \{\mu_i > 0\ (i = 1, \ldots, k)\} \quad \text{or} \quad \{\mu_i < 0\ (i = 1, \ldots, k)\}.$$

By Theorem 4·1, we get the critical region

$$\{\bar{X}_i \sqrt{n}/\sqrt{\sigma_{ii}} \geqslant s_2\ (i = 1, \ldots, k)\} \quad \text{or} \quad \{\bar{X}_i \sqrt{n}/\sqrt{\sigma_{ii}} \leqslant -s_2\ (i = 1, \ldots, k)\}.$$

When $k = 1$, $s_2 = u(\frac{1}{2}\alpha)$. When $k \geqslant 2$, Theorem 4·2 and Theorem 4·3 provide the percentage points only in a special case, namely, if $\sigma_{ij} \leqslant 0$ $(1 \leqslant i, j \leqslant k, i \neq j)$, $s_2 = u(\alpha)$. Especially when $p = k = 1$, this reduces to the ordinary two-sided test in the univariate case.

5. EXAMPLES

Given p normal populations $N_1(\mu_i, \sigma_i^2)$ with known variances, and samples of sizes n_i from these populations, their means are denoted by $\bar{X}_i$ $(i = 1, \ldots, p)$.

Suppose that we have prior information that the μ_i''s are nondecreasing, or that the simple order restriction is valid. These p populations need not be distinct, or some means may be equal, and this possibility can be tested against the hypothesis that all are distinct, i.e. with at least one equality in H_{12},

$$H_{12}\colon \mu_1 \leqslant \ldots \leqslant \mu_p, \quad K_{12}\colon \mu_1 < \ldots < \mu_p.$$

Instead of the simple order, we can consider simple tree order, $\mu_1 \leqslant \mu_i$ $(i = 2, \ldots, p)$; or simple loop order, $\mu_1 \leqslant \mu_i \leqslant \mu_p$ $(i = 2, \ldots, p-1)$. Likelihood ratio tests can be derived from Theorem 3·1 in all the cases. Note that for simple loop order the number of inequalities exceeds p for $p > 4$.

In their two-sided versions, the form of the critical regions can be determined easily. For instance, the critical region of the two-sided version for simple order is $Z_i \geqslant s_2$ $(i = 1, \ldots, p-1)$

or $Z_i \leqslant -s_2$ $(i = 1, \ldots, p-1)$, where

$$Z_i = \frac{\bar{X}_{i+1} - \bar{X}_i}{\sqrt{(\sigma_i^2/n_i + \sigma_{i+1}^2/n_{i+1})}} \quad (i = 1, \ldots, p-1).$$

Because the Z_i's are negatively correlated, we can apply Theorems 4·1–4·3 to arrive at a simple relation, $s_2 = u(\alpha)$ for $p \geqslant 3$, and when $p = 2$, to get the ordinary two-sided test, $s_2 = u(\frac{1}{2}\alpha)$. For simple tree order, the Z_i's are replaced by

$$Z_i = \frac{\bar{X}_{i+1} - \bar{X}_1}{\sqrt{(\sigma_1^2/n_1 + \sigma_{i+1}^2/n_{i+1})}} \quad (i = 1, \ldots, p-1).$$

In this case, the covariance matrix of Z_i's unfortunately does not satisfy the condition of Theorem 4·2, and a simple relation between s_2 and α cannot be derived except for the usual case when $p = 2$ and $s_2 = u(\frac{1}{2}\alpha)$. The form of the critical region of the simple loop order hypothesis can be derived similarly but we have the same difficulty about determining the relation between s_2 and α.

The restrictions of simple order, simple tree and simple loop were employed in testing problems by Bartholomew (1961), and we have formulated different problems employing the same restrictions. Our arguments can be applied to other types of restrictions described by Barlow *et al.* (1972, Chapter 3).

6. Generalizations

Our results can be generalized in the following two ways. First, Σ may have an unknown scale factor, that is $\Sigma = \sigma^2 \Lambda$, where Λ is known and σ^2 is unknown, but an independent estimate $\hat{\sigma}^2$ distributed as a χ^2 distribution is available. Secondly Σ may be completely unknown but an independent estimate $\hat{\Sigma}$ distributed as a Wishart distribution may be available.

In both cases, generalizations parallel to those in the previous sections can be established. Namely, the critical regions are of the exactly similar forms, and the percentage points can be computed from tables of t distributions, except for the two-sided test in the second case, where the Condition 1 cannot be employed. A detailed discussion will be published elsewhere.

The author is grateful to Professor A. Kudô for guidance and encouragement. He also thanks Mr K. Inada, Kagoshima University, and the referee for valuable suggestions. He also wishes to thank Miss T. Sasada for typewriting.

Appendix

Theorems related to linear inequalities and the convex cone

First we state two important theorems related to linear inequalities. These are essential not only for constructing the hypotheses but also for deriving the critical regions. Their original statements and proofs are given, for example, by Mangasarian (1969, Chapter 2).

Theorem A·1: *Gordan's theorem. The interior $\mathscr{J}^0$ is nonempty if and only if $\{a_1, \ldots, a_k\}$ is without positive relations.*

Theorem A·2: *Farkas's lemma. A vector a_j is redundant if and only if it can be written as*

$$a_j = \sum_{i \neq j} \lambda_i a_i, \quad \lambda_i \geqslant 0, \quad i \in \{1, \ldots, k\} - \{j\}. \tag{A·1}$$

The following theorem provides a geometrical interpretation of 'redundancy'. The theorem may well be known, but the present author has not been able to find it stated.

THEOREM A·3. *Suppose that* $\mathscr{J}^0 \neq \phi$. *Then* a_j *is redundant if and only if* $\mathscr{J}_j = \phi$.

Proof. Note that the assumption is, by Theorem A·1, equivalent to that $\{a_1, \ldots, a_k\}$ is without positive relations.

If a_j is redundant, then, from Theorem A·2, a_j can be written as (A·1), and λ_i, $i \in \{1, \ldots, k\} - \{j\}$, are assured to be not all simultaneously zero, by the assumption. Therefore, $a_j' y > 0$ for any y with $a_i' y > 0$, $i \in \{1, \ldots, k\} - \{j\}$, and hence $\mathscr{J}_j = \phi$.

Conversely, if $\mathscr{J}_j = \phi$, then, Motzkin's theorem (Mangasarian, 1969, Chapter 2) implies that there exist real numbers $\lambda_1, \ldots, \lambda_k$ such that

$$\lambda_i \geqslant^* 0, \quad i \in \{1, \ldots, k\} - \{j\}, \quad \sum_{i=1}^{k} \lambda_i a_i = 0.$$

By the assumption, $\lambda_j < 0$, and a_j can be written as

$$a_j = \sum_{i \neq j} \frac{\lambda_i}{-\lambda_j} a_i, \quad \frac{\lambda_i}{-\lambda_j} \geqslant 0, \quad i \in \{1, \ldots, k\} - \{j\},$$

and hence a_j is redundant by Theorem A·2.

REFERENCES

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. & BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions*. New York: Wiley.

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

BARTHOLOMEW, D. J. (1959b). A test of homogeneity for ordered alternatives. II. *Biometrika* **46**, 328–35.

BARTHOLOMEW, D. J. (1961). A test of homogeneity of means under restricted alternatives (with discussion). *J. R. Statist. Soc.* B **23**, 239–81.

DAVIS, C. (1954). Theory of positive linear dependence. *Am. J. Math.* **76**, 733–46.

INADA, K. (1978). Some bivariate tests of composite hypotheses with restricted alternatives. *Rep. Fac. Sci., Kagoshima Univ.* (*Math. Phys. Chem.*) No. 11, 25–31.

KUDÔ, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* **50**, 403–18.

KUDÔ, A. & CHOI, J. R. (1975). A generalized multivariate analogue of the one sided test. *Mem. Fac. Sci., Kyushu Univ.* A **29**, 303–28.

MANGASARIAN, O. L. (1969). *Nonlinear Programming*. New York: McGraw-Hill.

[*Received April* 1979. *Revised November* 1979]