Ann. Inst. Statist. Math.
Vol. 46, No. 1, 137–145 (1994)

LIKELIHOOD RATIO TESTS FOR A CLASS OF NON-OBLIQUE HYPOTHESES*

XIAOMI HU** AND F. T. WRIGHT

Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211, U.S.A.

(Received October 21, 1992; revised May 6, 1993)

Abstract. In this paper, we study the likelihood ratio tests for situations in which the null and alternative hypotheses are determined by two polyhedral cones, C_1 and C_2, which are nested so that $C_1 \subset L \subset C_2$ and L is a linear space. The two cones are proved to be non-oblique. Members which satisfy this nesting condition are easily identified and include the cases in which $C_1 = L$ or $L = C_2$. When testing two non-oblique hypotheses with variances unknown, the least favorable point within the null hypothesis has not been determined in general. However, for the situation considered here, the zero vector is shown to be least favorable within the null hypothesis. Two sets of hypotheses are said to be equivalent if they lead to the same likelihood ratio test. For two non-oblique polyhedral cones, C_1 and C_2, four sets of equivalent hypotheses are identified. If $C_1 \subset L \subset C_2$, then the two cones in each of these four sets of hypotheses are similarly nested with a linear space in between.

Key words and phrases: Inferences subject to inequality constraints, iterated projection property, non-oblique cones, order restricted inferences.

1. Introduction

Bartholomew (1959, 1961) developed the likelihood ratio tests (LRTs) of homogeneity of normal means with the alternative restricted by an ordering on the means, such as $\mu_1 \le \mu_2 \le \cdots \le \mu_k$ or $\mu_1 \le \mu_2 \le \cdots \le \mu_h \ge \mu_{h+1} \ge \cdots \ge \mu_k$. Robertson and Wegman (1978) considered the LRTs with the order restriction as the null hypothesis and no restrictions on the alternative. The collection of mean vectors which satisfy a particular order restriction is a closed, convex cone. The cones determined by the most restrictive ordering, homogeneity of components, and the least restrictive ordering, no restriction, are linear spaces. Robertson *et*

* This research was supported in part by the National Institutes for Health under Grant 1 R01 GM42584-01A1.

** Part of this work is taken from this author's dissertation submitted in partial fulfillment for the Ph.D. Degree at the University of Missouri-Columbia.

al. (1988) summarize many of the details involved in making statistical inferences subject to order restrictions.

The collection of mean vectors which satisfy a set of linear inequality constraints, such as $\sum_{j=1}^{k} a_{ij}\mu_j \leq 0$ for $i = 1, 2, \ldots, m$, is also a closed, convex cone which is said to be polyhedral. With μ a normal mean vector, Raubertas *et al.* (1986) extend the results above to obtain the LRTs of H_0 versus $H_a - H_0$ where $H_0 : \mu \in C_1$ and $H_a : \mu \in C_2$ with C_1 and C_2 polyhedral cones, provided C_1 or C_2 is a linear space. We say such sets of hypotheses are of the LC type if C_1 is a linear space and C_2 is a polyhedral cone but not a linear space. Sets of hypotheses of the CL and CC type are defined similarly.

Warrack and Robertson (1984) studied an example of CC type hypotheses and found that the LRT is dominated by the LRT of another set of LC type hypotheses in that case. Menendez and Salvador (1991) examined the same issue in a general context. Using the concept of obliqueness of two cones, which was introduced by Warrack and Robertson (1984), Menendez *et al.* (1992*a*) explained this dominance. Menendez *et al.* (1992*b*) studied the LRTs for CC type hypotheses in which the two cones are not oblique. Non-obliqueness is characterized by an iterated projection property. In particular, two nested cones, $C_1 \subset C_2$, satisfy this property if the projection of any vector onto C_1 is the same as the projection onto C_1 of the projection of that vector onto C_2. Of course, this property depends on the distance function which determines the projection. The iterated projection property is difficult to verify in some situations.

In this paper, we consider the LRTs for a special class of CC type hypotheses which are non-oblique. We say that the hypotheses H_0 and H_a are of CLC type if $C_1 \subset L \subset C_2$ with L a linear space. We will show that in such settings the cones C_1 and C_2 are non-oblique. The members of this class are easy to identify and include all of the problems discussed in Raubertas *et al.* (1986) and Robertson *et al.* (1988). Menendez *et al.* (1992*b*) point out that when testing non-oblique hypotheses with unknown variances, the least favorable point in the null hypothesis has not been determined in general. For the types of hypotheses studied here the zero vector is least favorable and hence the significance level can be computed by taking the mean equal to zero.

In Section 2, we prepare the necessary tools for the development of the tests. In Section 3, the LRTs are developed for CLC type hypotheses with polyhedral cones in the cases of independent random samples and a random sample from a multivariate distribution. In Section 4, equivalent sets of hypotheses are discussed and four sets of equivalent hypotheses are identified for the variance known case.

2. Preliminaries

Assume that the Hilbert space R^k has inner product $\langle \cdot, \cdot \rangle$. With $\|\cdot\|$ the norm induced by the inner product, the projection of the vector x onto set D, $P(x \mid D)$, is a vector that minimizes $\|x - y\|$ over $y \in D$. For a closed, convex set D and any vector x, $P(x \mid D)$ exists and is unique. Furthermore,

$$P(x \mid D) = x^* \quad \text{if and only if } x^* \in D \text{ and } \langle x - x^*, x^* - y \rangle \geq 0 \text{ for each } y \in D. \tag{2.1}$$

The characterization given in (2.1) will be used repeatedly to verify that a candidate is the desired projection. For a set A, A^P denotes the set $\{y : \langle x, y\rangle \le 0,$ for each $x \in A\}$ which is a closed, convex cone. For a closed, convex cone C, C^P is called the polar cone associated with C. It is well known that for any vector x, $x = P(x \mid C) + P(x \mid C^P)$, and $\langle P(x \mid C), P(x \mid C^P)\rangle = 0$. Define $A + B$ as the set $\{x + y : x \in A, y \in B\}$. If both A and B are convex cones, so is $A + B$. Hestenes (1975) gives an example where both A and B are closed, convex cones, but $A + B$ is not closed. When A and B are orthogonal, the notation $A \oplus B$ is used for $A + B$. These definitions and related results are given in many references, see for example Robertson *et al.* (1988).

LEMMA 2.1. *If $C_1 \subset L \subset C_2$, C_1 and C_2 are closed, convex cones in R^k and L is a linear space, then the following conclusions hold:*

(1) *For any set D in L, D and C_2^P are orthogonal;*

(2) *For any closed, convex set D in L, D and C_2 are non-oblique, i.e. $P(P(x \mid C_2) \mid D) = P(x \mid D)$ for each $x \in R^k$;*

(3) *$P(x \pm y \mid C_2) = P(x \mid C_2) \pm y$ and $P(x \pm y \mid C_2^P) = P(x \mid C_2^P)$ for each $x \in R^k$ and each $y \in L$; and*

(4) *$\|P(x \mid C_2) - P(x \mid C_1)\| \le \|P(x - y \mid C_2) - P(x - y \mid C_1)\|$ for each $x \in R^k$ and each $y \in C_1 \oplus C_2^P$.*

PROOF. (1) is true since $C_2^P \subset L^\perp$. (2) and (3) can be proved by checking the conditions in (2.1). We now consider (4). For each $x \in R^k$ and each $y \in C_1$ by conclusion (2) and the definition of a projection,

$$\|P(x \mid C_2) - P(x \mid C_1)\| \le \|P(x \mid C_2) - (y + P(x - y \mid C_1))\|.$$

By conclusion (3) above, the right hand side is $\|P(x - y \mid C_2) - P(x - y \mid C_1)\|$. Thus

$$\|P(x \mid C_2) - P(x \mid C_1)\| \le \|P(x - y \mid C_2) - P(x - y \mid C_1)\|.$$

Applying this result to the nested structure $C_2^P \subset L^\perp \subset C_1^P$, we have that

$$\|P(x - y \mid C_1^P) - P(x - y \mid C_2^P)\| \le \|P(x - y - z \mid C_1^P) - P(x - y - z \mid C_2^P)\|$$

for each z in C^P. Thus,

$$\|P(x \mid C_2) - P(x \mid C_1)\| \le \|P(x - (y + z) \mid C_2) - P(x - (y + z) \mid C_1)\|. \quad \square$$

A polyhedral cone is a set of points which satisfy a finite set of linear inequalities. With A an m by k matrix, the polyhedral cone $\{x \in R^k : Ax \le 0\}$ is denoted by $C[A]$, where an inequality between a vector and a real number is to be interpreted component wise. Let A_i be a matrix consisting of some rows in A, and $\bar{A}_i$ be the matrix consisting of the rows in A but not in A_i. The matrix without any rows is denoted by $\emptyset$. We assume that $C[A]$ has no redundant rows, i.e. $C[A_i] \ne C[A]$ if $A_i \ne A$. The face of $C[A]$ associated with A_i, $F[A_i]$, is defined as $\{x : A_i x = 0, \bar{A}_i x \le 0\}$. The face $F[A_i]$ is in the linear space $\{x : A_i x = 0\}$ which is called the null space associated with A_i and denoted by $N[A_i]$. The interior of the face $F[A_i]$ with respect to the topology in $N[A_i]$ is the set $\{x : A_i x = 0, \bar{A}_i x < 0\}$ and is denoted by $F^0[A_i]$. It is known that $C[A]$ can be partitioned into a finite number of interiors of faces.

LEMMA 2.2. *If $L_1 \subset C \subset L_2$ where L_1, L_2 are linear subspaces of R^k, $C = C[A]$ is a polyhedral cone, then for each x in R^k the following are equivalent:*
(1) $P(x \mid C) \in F^o[A_i]$;
(2) $P(x \mid C) = P(x \mid N[A_i]) \in F^o[A_i]$; *and*
(3) $P(x \mid N[A_i]) - P(x \mid L_1) \in F^o[A_i]$ *and* $P(x \mid L_2) - P(x \mid N[A_i]) \in C^P$.

PROOF. The implications are proved in the order indicated below.

(1) $\Rightarrow$ (2): For each $y \in N[A_i]$ there exists $\delta > 0$ such that $P(x \mid C) \pm \delta y \in F^o[A_i] \subset C$ since $P(x \mid C) \in F^o[A_i]$ and $F^o[A_i]$ is an open set with respect to the topology in $N[A_i]$. By (2.1) we have $\langle x - P(x \mid C), y\rangle = 0$ which establishes $P(x \mid C) = P(x \mid N[A_i])$.

(2) $\Rightarrow$ (3): $L_1 \subset C$ implies that $P(x \mid N[A_i]) - P(x \mid L_1)$ and $P(x \mid N[A_i])$ are in the same interior of a face. Notice that $P(x \mid L_2) - P(x \mid N[A_i]) = P(x \mid C^P) - P(x \mid L_2^\perp)$ and $L_2^\perp \subset C^P$. So, $P(x \mid L_2) - P(x \mid N[A_i]) \in C^P$.

(3) $\Rightarrow$ (1): Clearly, $P(x \mid N[A_i]) \in F^o[A_i] \subset C$. Applying conclusion (1) of Lemma 2.1 and $P(x \mid L_2) - P(x \mid N[A_i]) \in C^P$, checking the conditions in (2.1), we have that $P(P(x \mid L_2) \mid C) = P(x \mid N[A_i])$. But by conclusion (2) of Lemma 2.1, $P(P(x \mid L_2) \mid C) = P(x \mid C)$. Therefore $P(x \mid C) \in F^o[A_i]$. □

Let W be a k by k positive definite matrix. For any vectors x and y in R^k, define $\langle x, y\rangle_W$ as $x^T W y$. Then $\langle \cdot, \cdot \rangle_W$ is a well defined inner product. The norm induced by this inner product is denoted by $\| \cdot \|_W$. The projection of a vector x onto set D is denoted by $P_W(x \mid D)$. The following lemma is essentially Lemma 3.1 of Shapiro (1985). We present its proof because it may be of interest from a methodological point of view.

LEMMA 2.3. *If X is a k-dimensional normal random vector with zero mean and covariance matrix W^{-1}, C is a cone in R^k and L is a linear space, then $\|P_W(X \mid L)\|_W^2$ conditioned on $P_W(X \mid L) \in C$ has a central chi-squared distribution with degrees of freedom* $\dim(L)$.

PROOF. If $Y \sim N(0, V)$ and $AVA^T = \mathrm{diag}(\sigma^2, \ldots, \sigma^2, 0, \ldots, 0)$, then $(AY)^T(AY)$ and $AY \in C$ are independent by Basu's Theorem, cf. Hogg and Craig (1978) since $(AY)^T(AY)$ is sufficient for σ^2, but AY/σ is ancillary. Let PX be $P_W(X \mid L)$, then P is idempotent and self-adjoint, cf. Bachman and Narici (1972). Thus, $\sqrt{W} P W^{-1} P^T \sqrt{W}$ is symmetric and idempotent. Therefore, there exists an orthogonal matrix Q such that $Q\sqrt{W} P W^{-1} P^T \sqrt{W} Q^T = \mathrm{diag}(1, \ldots, 1, 0, \ldots, 0)$. Applying the claim at the beginning of the proof with $Y = P_W(X \mid L)$, $V = PW^{-1}P^T$ and $A = Q\sqrt{W}$, we see that $\|P_W(X \mid L)\|_W^2$ and $P_W(X \mid L) \in C$ are independent. But $\|P_W(X \mid L)\|_W^2$ has chi-squared distribution with degrees of freedom equal to $r = \mathrm{rank}(P) = \dim(L)$. The conclusion follows. □

In the nested structure $C_1 \subset L \subset C_2$, if C_2 is the polyhedral cone $C[A]$, then C_1 is in $N[A]$ since $L \subset N[A]$. So, without loss of generality, we assume that $C_1 = N[A] \cap C[B]$.

LEMMA 2.4. *If $X \sim N(0, W^{-1})$ and $C_1 = N[A] \cap C[B]$ and $C_2 = C[A]$ are polyhedral cones in R^k, then $\|X - P_W(X \mid C_2)\|_W^2$, $\|P_W(X \mid C_2) - P_W(X \mid C_1)\|_W^2$ and $\|P_W(X \mid C_1)\|_W^2$, conditioned on $P_W(X \mid C_1) \in N[A] \cap F^o[B_j]$ and $P_W(X \mid C_2) \in F^o[A_i]$, are independent central chi-squared variables with the degrees of freedom $k - \dim(N[A_i])$, $\dim(N[A_i]) - \dim(N[A] \cap N[B_j])$ and $\dim(N[A] \cap N[B_j])$.*

PROOF. By Lemma 2.2, the variables in this lemma conditioned on the events in this lemma are identical in distribution to the variables $\|P_W(X \mid N^\perp[A_i])\|_W^2$, $\|P_W(X \mid N[A_i] \cap (N[A] \cap N[B_j])^\perp)\|_W^2$ and $\|P_W(X \mid N[A] \cap N[B_j])\|_W^2$ conditioned on the three events $P_W(X \mid N^\perp[A_i]) \in C_2^P$, $P_W(X \mid N[A_i] \cap (N[A] \cap N[B_j])^\perp) \in C_1^P \cap F^o[A_i]$ and $P_W(X \mid N[A] \cap N[B_j]) \in N[A] \cap F^o[B_j]$. But $P_W(X \mid N[A] \cap N[B_j])$, $P_W(X \mid N^\perp[A_i])$ and $P_W(X \mid N[A_i] \cap (N[A] \cap N[B_j])^\perp)$ are independent and C_2^P, $C_1^P \cap F^o[A_i]$, $N[A] \cap F^o[B_j]$ are cones. The desired conclusion follows from Lemma 2.3. □

3. Likelihood ratio tests

In this section, we consider the LRTs for CLC type hypotheses with polyhedral cones. As noted earlier, we may suppose that $C_1 = N[A] \cap C[B]$, $C_2 = C[A]$ and $L = N[A]$. We consider the two cases of independent random samples and a random sample from a multivariate population.

Case 1 (Independent random samples). Suppose Y_{ij} for $j = 1, 2, \ldots, n_i$ and $i = 1, 2, \ldots, k$ are independent and $Y_{ij} \sim N(\mu_i, \sigma_i^2)$. Let $\mu = (\mu_1, \mu_2, \ldots, \mu_k)^T$, $\bar{Y} = (\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_k)^T$ with $\bar{Y}_i$ the mean of the i-th random sample and $\Sigma = \operatorname{diag}(\sigma_1^2/n_1, \sigma_2^2/n_2, \ldots, \sigma_k^2/n_k)$.

Case 2 (A multivariate random sample). Suppose $Y_j = (Y_{1j}, Y_{2j}, \ldots, Y_{kj})^T$ for $j = 1, 2, \ldots, n$ are independent and identically distributed k-dimensional normal random vectors with mean $\mu = (\mu_1, \mu_2, \ldots, \mu_k)^T$ and covariance matrix V. Then $\bar{Y} = (\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_k)^T$, with $\bar{Y}_i$ the mean of Y_{ij} for $j = 1, 2, \ldots, n$, has covariance $\Sigma = V/n$.

Let $M(A)$ be a collection of A_i, submatrices of A, for which $F^o[A_i]$ partitions $C[A]$ and let $M(B)$ be defined similarly. With $W = \Sigma^{-1}$, let P_{ij} be the probability, under $\mu = 0$, that $P_W(\bar{Y} \mid C_1)$ is in $N[A] \cap F^o[B_j]$ and $P_W(\bar{Y} \mid C_2)$ is in $F^o[A_i]$.

Variances known. In this section, we consider the two sampling plans described above and suppose that Σ is known. Let χ_ν^2 denote a chi-squared random variable with ν degrees of freedom and $\chi_0^2 \equiv 0$. The test statistic in (3.1) below and its null distribution are given in Menendez *et al.* (1992*b*), but (3.2) below gives a more explicit relationship between the degrees of freedom for the chi-squared random variables and the P_{ij}.

THEOREM 3.1. *Let C_1, C_2, W and P_{ij} be defined as above and let $H_0 : \mu \in C_1$ and $H_a : \mu \in C_2$. If W is known, then the LRT of H_0 versus $H_a - H_0$ rejects H_0 for large values of*

$$T = \|P_W(\bar{Y} \mid C_2) - P_W(\bar{Y} \mid C_1)\|_W^2; \tag{3.1}$$

$\mu = 0$ is least favorable within H_0, i.e. the supremum of $P[T \geq t]$ over $\mu \in C_1$ occurs at $\mu = 0$; and under $\mu = 0$,

$$P[T \geq t] = \sum_{A_i \in M(A), B_j \in M(B)} P_{ij} P(\chi^2_{\dim(N[A_i]) - \dim(N[A] \cap N[B_j])} \geq t). \tag{3.2}$$

PROOF. Clearly, the LRT rejects H_0 for large values of $\|\bar{Y} - P_W(\bar{Y} \mid C_1)\|_W^2 - \|\bar{Y} - P_W(\bar{Y} \mid C_2)\|_W^2$ and by conclusion (1) of Lemma 2.1 this equals T which is given in (3.1). Furthermore, it follows from conclusion (4) of Lemma 2.1 that the distribution of T with $\mu \in C_1$ is stochastically less than or equal to the distribution of T with $\mu = 0$. Thus, zero is least favorable within the null hypothesis. Because $\{F^o[A_i] : A_i \in M(A)\}$ partitions C_2 and $\{N[A] \cap F^o[B_j] : B_j \in M(B)\}$ partitions C_1, (3.2) follows from the law of total probability and Lemma 2.4. □

Variances unknown. For independent random samples, i.e. Case 1, one commonly assumes that $\sigma_i^2 = a_i\sigma^2$ with a_i known and σ^2 unknown. Thus with $u_i = n_i/a_i$ for $i = 1, 2, \ldots, k$ and $U = \mathrm{diag}(u_1, u_2, \ldots, u_k)$, $\Sigma = \sigma^2 U^{-1}$ and $W = U/\sigma^2$. In Case 2, it is commonly assumed that $V = \sigma^2\Sigma_0$ where Σ_0 is known and σ^2 is unknown. Then $W = U/\sigma^2$ where $U = n\Sigma_0^{-1}$. For Cases 1 and 2 respectively, define

$$R = \sum_{i=1}^{k} \frac{u_i}{n_i} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 \quad \text{and} \quad R = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^T U (Y_i - \bar{Y}) \tag{3.3}$$

and note that the degrees of freedom associated with R are $\nu = n_1 + n_2 + \cdots + n_k - k$ and $\nu = nk - k$, respectively.

Since σ^2 is unknown, one cannot compute distances with respect to the metric determined by W, however a projection with respect to W is the same as the projection with respect to U. Thus, P_{ij} is unchanged if the two projections in its definition are taken with respect to U rather than W.

Let $F(a, b)$ denote a random variable with an F distribution with parameters a and b, let $\nu(A_i) = \dim(N[A_i])$ and let $\nu(B_j) = \dim(N[A] \cap N[B_j])$.

THEOREM 3.2. *Let C_1, C_2, U, P_{ij}, R and ν be defined as above and let $H_0 : \mu \in C_1$ and $H_a : \mu \in C_2$. If U is known and σ^2 is unknown, then the LRT of H_0 versus $H_a - H_0$ rejects H_0 for large values of*

$$T_1 = \frac{\|P_U(\bar{Y} \mid C_2) - P_U(\bar{Y} \mid C_1)\|_U^2}{R + \|\bar{Y} - P_U(\bar{Y} \mid C_2)\|_U^2}; \tag{3.4}$$

$\mu = 0$ is least favorable within H_0; and for $\mu = 0$ and $t > 0$,

$$(3.5)\quad P[T_1 \geq t] = \sum_{A_i \in M(A), B_j \in M(B)} P_{ij} P\left[F(\nu(A_i) - \nu(B_j), \nu + k - \nu(A_i)) \geq \frac{\nu + k - \nu(A_i)}{\nu(A_i) - \nu(B_j)} t\right]$$

where the summation is taken over A_i and B_j with $\nu(A_i) > \nu(B_j)$.

PROOF. It follows from standard arguments, see Robertson *et al.* ((1988), p. 63), and conclusion (1) of Lemma 2.1 that the LRT rejects H_0 for large values of T_1 which is given by (3.4). To show that $\mu = 0$ is least favorable, let $\mu \in C_1$ and set

$$S(\mu) = \frac{\|P_W(X + \mu \mid C_2) - P_W(X + \mu \mid C_1)\|_W^2}{Q(\nu) + \|X + \mu - P_W(X + \mu \mid C_2)\|_W^2}$$

where $Q(\nu) \sim \chi_\nu^2$, $X \sim N(0, W^{-1})$ and $Q(\nu)$ and X are independent. Then, by conclusions (3) and (4) of Lemma 2.1, $S(\mu)$ is stochastically less than or equal to $S(0)$. However,

$$(3.6)\qquad T_1 = \frac{\|P_W(\bar{Y} \mid C_2) - P_W(\bar{Y} \mid C_1)\|_W^2}{R/\sigma^2 + \|\bar{Y} - P_W(\bar{Y} \mid C_2)\|_W^2},$$

$R/\sigma^2 \sim \chi_\nu^2$, $\bar{Y} \sim N(\mu, W^{-1})$ and R/σ^2 and $\bar{Y}$ are independent. Hence, $\mu = 0$ is least favorable within the null hypothesis.

Let $\mu = 0$. Conditioning on $P_U(\bar{Y} \mid C_1) \in N[A] \cap F^o[B_j]$ and $P_U(\bar{Y} \mid C_2) \in F^o[A_i]$, applying Lemma 2.4 to (3.6) and appealing to the law of total probability, (3.5) is established. □

Remark. Let $B(a, b)$ denote a random variable with a beta distribution and $B(0, b) \equiv 0$. It is straightforward to show that the LRT of H_0 versus $H_a - H_0$ rejects H_0 for large values of

$$(3.7)\qquad T_2 = \frac{\|P_U(\bar{Y} \mid C_2) - P_U(\bar{Y} \mid C_1)\|_U^2}{R + \|\bar{Y} - P_U(\bar{Y} \mid C_1)\|_U^2}.$$

If T_2 is used as a test statistic, zero is still least favorable in H_0 and its null distribution is a mixture of beta distributions known as an E-bar-squared distribution. In particular, for $\mu = 0$ and $t > 0$,

$$(3.8)\qquad P[T_2 \geq t] = \sum_{A_i \in M(A), B_j \in M(B)} P_{ij} P\left[B\left(\frac{1}{2}\nu(A_i) - \frac{1}{2}\nu(B_j), \frac{1}{2}\nu + \frac{1}{2}k - \frac{1}{2}\nu(A_i)\right) \geq t\right]$$

where $M(A)$, $M(B)$, P_{ij}, $\nu(A_i)$, $\nu(B_j)$ and ν are defined as in Theorem 3.2.

Example. We consider a pesticide problem in which the effectiveness of different dosage levels of a pesticide is under investigation. A previous study has confirmed that $\mu_1 \le \mu_2 \le \mu_3$ where μ_1, μ_2 and μ_3 are the mean percentages of pests killed at three increasing dosage levels. It is suspected that the pests have developed an immunity to the pesticide, at least at the dosage levels currently being used. If this is the case, then we now believe that $\mu_1 = \mu_2 = \mu_3$. We wish to investigate the response of these pests to the pesticide at the three dosage levels studied earlier as well as at two larger dosage levels. If an immunity has been developed, we expect the response at the two new levels to be at least as good as at the other three, but because we have no information concerning the two new levels, it is not clear how they are related to each other. Hence, our null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3 \le \min(\mu_4, \mu_5)$. If the suspected immunity has not been developed, then we believe that $\mu_1 \le \mu_2 \le \mu_3$, but we do not know how μ_4 or μ_5 are related to each other or to the other means. Thus, $H_a : \mu_1 \le \mu_2 \le \mu_3$. The two hypotheses are of the CLC type and the linear space between them is $L = \{\mu \in R^5 : \mu_1 = \mu_2 = \mu_3\}$.

This example illustrates a type of CLC hypotheses which are a combination of the type of problems studied by Bartholomew (1961) and Robertson and Wegman (1978). The null imposes both equality and inequality constraints and the alternative is obtained by changing the equality constraints to inequality constraints in the null hypothesis.

4. Equivalent sets of hypotheses

Two sets of null and alternative hypotheses are said to be (likelihood ratio) equivalent if the LRTs for these two sets of hypotheses have the same rejection regions for each significance level $\alpha \in (0, 1)$. The next result is motivated by the observation in Menendez *et al.* (1992*b*) that the LRTs for situations (A), (C) and (D) below have the same critical regions.

THEOREM 4.1. *If $\bar{Y}$ is a k-dimensional normal random vector with unknown mean μ and known covariance matrix W^{-1} and C_1 and C_2 are two closed, convex cones in R^k which are non-oblique, then the following sets of null and alternative hypotheses are equivalent.*

(A) $H_0 : \mu \in C_1$, $H_a : \mu \in C_2$,

(B) $H_0 : \mu \in C_2^P$, $H_a : \mu \in C_1^P$,

(C) $H_0 : \mu = 0$, $H_a : \mu \in C_2 \cap C_1^P$,

(D) $H_0 : \mu \in (C_2 \cap C_1^P)^P$, $H_a : \mu \in R^k$.

PROOF. By Theorem 5.2 of Zarantonello (1971), C_1^P and C_2^P are non-oblique because C_1 and C_2 are. Thus, from the work of Menendez *et al.* (1992*b*), we see that $\mu = 0$ is least favorable within H_0 for each of the testing situations; for the testing situations (A), (C) and (D) the test statistics are the same; and for the hypotheses in (B) the test statistic is $\|P_W(\bar{Y} \mid C_1^P) - P_W(\bar{Y} \mid C_2^P)\|_W^2$ which clearly is the same as for (A). $\square$

It is instructive to notice that (C) is of LC type, (D) is of CL type. Both are special cases of CLC type of hypotheses. Furthermore, if $C_1 \subset L \subset C_2$, then (A), (B), (C) and (D) are all of CLC type.

Acknowledgements

The authors want to thank the referees for their comments which improved this paper.

References

Bachman, G. and Narici, L. (1972). *Functional Analysis*, Academic Press, New York.

Bartholomew, D. J. (1959). A test of homogeneity for ordered alternatives, *Biometrika*, **46**, 36–48 and 328–335.

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

Hestenes, M. R. (1975). *Optimization Theory*, Wiley, New York.

Hogg, R. V. and Craig, A. F. (1978). *Introduction to Mathematical Statistics*, Macmillan, New York.

Menendez, J. A. and Salvador, B. (1991). Anomalies of the likelihood ratio test for testing restricted hypotheses, *Ann. Statist.*, **19**, 889–898.

Menendez, J. A., Rueda, C. and Salvador, B. (1992*a*). Dominance of likelihood ratio tests under cone constraints, *Ann. Statist.*, **20**, 2087–2099.

Menendez, J. A., Rueda, C. and Salvador, B. (1992*b*). Testing non-oblique hypotheses, *Comm. Statist. Theory Methods*, **21**, 471–484.

Raubertas, R. F., Lee, C. I. C. and Nordheim, E. V. (1986). Hypothesis tests for normal means constrained by linear inequalities, *Comm. Statist. Theory Methods*, **15**, 2809–2833.

Robertson, T. and Wegman, E. J. (1978). Likelihood ratio tests for order restrictions in exponential families, *Ann. Statist.*, **6**, 485–505.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inferences*, Wiley, New York.

Shapiro, A. (1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints, *Biometrika*, **72**, 133–144.

Warrack, G. and Robertson, T. (1984). A likelihood ratio test regarding two nested but oblique order restricted hypotheses, *J. Amer. Statist. Assoc.*, **79**, 881–889.

Zarantonello, E. H. (1971). Projections on convex sets in Hilbert space and spectral theory, *Contributions to Nonlinear Functional Analysis* (ed. E. H. Zarantonello), 237–424, Academic Press, New York.