UNIFORMLY MORE POWERFUL TESTS FOR HYPOTHESES CONCERNING LINEAR INEQUALITIES AND NORMAL MEANS

Roger L. Berger

Institute of Statistics Mimeo Series No. 1913

January 1988

Dr. Roger L. Berger
Statistics Dept.
Box 8203
N. C. State University
Raleigh, NC 27695-8203

This research was funded in part by Subgrant SG-652 from the U. S. Army Research Office through the Academy of Applied Science. Kristin Rahn and Soku Yi are thanked for computer programming assistance that produced tables and figures.

KEY WORDS: Likelihood ratio test; Uniformly more powerful; Linear inequality; Qualitative interaction; Majorization

ABSTRACT

In this paper we consider some hypothesis testing problems regarding normal means. In these problems, the hypotheses are defined by linear inequalities on the means. We show that in certain problems the likelihood ratio rest (LRT) is not very powerful. We describe a test that has the same size, α, as the LRT and is uniformly more powerful. The test is easily implemented since its critical values are standard normal percentiles. The increase in power with the new test can be substantial. For example the new test's power is $1/2\alpha$ times bigger (10 times bigger for $\alpha = .05$) than the LRT's power for some parameter points in a simple example.

Specifically let $\underset{\sim}{X} = (X_1,\ldots,X_p)'$ $(p \geq 2)$ be a multivariate normal random vector with unknown mean $\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_p)'$ and known, nonsingular covariance matrix $\underset{\sim}{\Sigma}$. We consider testing the null hypothesis

$$H_0: \quad \underset{\sim}{b}_i'\underset{\sim}{\mu} \leq 0 \text{ for some } i = 1,\ldots,k$$

versus the alternative hypothesis

$$H_1: \quad \underset{\sim}{b}_i'\underset{\sim}{\mu} > 0 \text{ for all } i = 1,\ldots,k \ .$$

Here $\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_k$ $(k \geq 2)$ are specified p-dimensional vectors that define the hypotheses. Many types of relationships among the means may be described with the linear inequalities in H_0 and H_1. Two interesting types of hypotheses are those that specify the signs of the means and those that describe an order relationship among the means. Some examples of alternative hypotheses that can be specified in this way are these:

$$H_1^S: \quad \mu_i > 0 \quad i = 1,\ldots,p \qquad \text{(sign testing)}$$

$$H_1^O: \quad \mu_1 < \mu_2 <\ldots< \mu_p \qquad \text{(simple order)}$$

H_1^L: $\mu_1 < \mu_i < \mu_p$ $\quad i = 2,\ldots,p-1$ $\quad$ (simple loop)

H_1^T: $\mu_1 < \mu_i$ $\quad i = 2,\ldots,p$ $\quad$ (simple tree).

If $\mu_i = \nu_{2i} - \nu_{1i}$ where ν_{ji} is the average response of the i^{th} patient subset to the j^{th} treatment, then H_1^S states that Treatment 2 is better than Treatment 1 for all subsets. If the μ_i are regression coefficients, then H_1^S states that the mean response increases with each independent variable. In any case, these relationships would be the alternative hypothesis. Rejection of H_0 by a test with small size would be taken as strong evidence confirming that the specified sign or order relationship is true.

Sasabuchi (1980) showed that the size-α LRT of H_0 versus H_1 is the test that rejects H_0 if

$$Z_i = \frac{\underset{\sim}{b}_i'\underset{\sim}{X}}{\sqrt{\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_i}} \geq z_\alpha \qquad \text{for all } i = 1,\ldots,k$$

where z_α is the upper 100α percentile of a standard normal distribution. This test is biased and has very low power if all the values $\underset{\sim}{b}_i'\underset{\sim}{\mu}$, $i = 1,\ldots,k$, are only slightly bigger than zero. We show that the following test is better than the LRT in that it has the same size, α, and has higher power at all parameter points. Let the size of the test, α, be less than .5. Define the integer J by the inequality $J-1 < 1/2\alpha \leq J$. Define the constants $c_0,\ldots,c_J$ by $c_0 = \infty$; $c_j = z_{j\alpha}$, $j=1,\ldots,J-1$; and $c_J = 0$. (Again, $z_{j\alpha}$ is the upper $100(j\alpha)$ percentile of a standard normal distribution.) In many cases, a size-α test that is uniformly more power than the LRT is the test that rejects H_0 if

$$\underset{\sim}{X} \in \bigcup_{j=1}^{J} R_j$$

where $R_j = \{\underset{\sim}{x}: \ c_j \leq z_i \leq c_{j-1},\ i = 1,\ldots,k\}$ and $z_i = \underset{\sim}{b}_i'\underset{\sim}{x}/\sqrt{\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_i}$ is the LRT statistic. The set R_1 is the rejection region of the LRT. So this test is obviously more powerful than the LRT. But we show that if, for each $i = 1,\dots,k$, there exists an $m \neq i$ such that $\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$, then this test is also a size-α test. It is easy to verify that this condition is satisfied, for example, for all the H_1 hypotheses mentioned in the preceding paragraph, except the simple tree, if $\underset{\sim}{\Sigma}$ is diagonal.

Tests that are even more powerful than those described in the previous paragraph might exist. We discuss an example of such a test. Its rejection region is the union of the above rejection region and more sets R_j, $j > J$. But despite this test's superior power properties, it is rather counterintuitive because its rejection region includes sample points $\underset{\sim}{x}$ for which $\underset{\sim}{b}_i'\underset{\sim}{x} < 0$ for *all* $i = 1,\dots,k$. Since $\underset{\sim}{X}$ is an estimate of $\underset{\sim}{\mu}$, it might seem strange to conclude that $\underset{\sim}{b}_i'\underset{\sim}{\mu} > 0$ for all $i = 1,\dots,k$ when the observed estimate of $\underset{\sim}{\mu}$ satisfies $\underset{\sim}{b}_i'\underset{\sim}{x} < 0$ for all $i = 1,\dots,k$. Thus tests such as in this example may be primarily of theoretical interest.

All of the previously mentioned results are derived in the $\underset{\sim}{\Sigma}$ known case. Sasabuchi (1980) showed that, if $\underset{\sim}{\Sigma}$ is unknown, the LRT is very similar to the LRT described above. The differences are that $\underset{\sim}{\Sigma}$ is replaced by an estimate in the denominator of Z_i and z_α is replaced by t_α, a t-distribution percentile. This suggests that if the constants $c_0,\dots,c_J$ are defined in terms of t percentiles rather than normal percentiles, a uniformly more powerful size-α test in the unknown $\underset{\sim}{\Sigma}$ case might be obtained. However we discuss an example that shows that this is not always true. The test constructed in this way has size greater than α. But in the example the size of the test converges to α quickly as the degrees of freedom for the estimate of $\underset{\sim}{\Sigma}$ becomes large. So even for moderate degrees of freedom (≥ 10), this test might be preferable to the LRT since its power is approximately α and it is much more powerful than the LRT.

We also briefly consider this two-sided hypothesis testing problem. Consider testing

$$H_0^2: \ \underset{\sim}{b}_i'\mu \le 0 \text{ for some } i = 1,\ldots,k, \text{ and } \underset{\sim}{b}_i'\mu \ge 0 \text{ for some } i = 1,\ldots,k$$

versus

$$H_1^2: \ \underset{\sim}{b}_i'\mu > 0 \text{ for all } i = 1,\ldots,k \text{ or } \underset{\sim}{b}_i'\mu < 0 \text{ for all } i = 1,\ldots,k.$$

Sasabuchi (1980) showed that the LRT rejects H_0^2 if $Z_i \ge c$ for all $i = 1,\ldots,k$ or $Z_i \le -c$ for all $i = 1,\ldots,k$. The determination of the constant c that yields a size-α test is more difficult than in the one-sided case. Sasabuchi gave some conditions under which $c = z_\alpha$. We consider only the special case of testing

$$H_0^2: \ \mu_i \le 0 \text{ for some } i = 1,\ldots,p \text{ and } \mu_i \ge 0 \text{ for some } i = 1,\ldots,p$$

versus

$$H_1^2: \ \mu_i > 0 \text{ for all } i = 1,\ldots,p \text{ or } \mu_i < 0 \text{ for all } i = 1,\ldots,p .$$

and $\underset{\sim}{\Sigma} = \text{diag}(\sigma_1^2,\ldots,\sigma_p^2)$, a diagonal matrix. Let the constants $c_0,\ldots,c_J$ be defined as above and define $c_{J+1},\ldots,c_{2J}$ by $c_j = -c_{2J-j}$. Then we show that the test that rejects H_0^2 if

$$\underset{\sim}{X} \in \bigcup_{j=1}^{2J} R_j$$

where $R_j = \{\underset{\sim}{x}: \ c_j \le x_i/\sigma_i \le c_{j-1}, \ i = 1,\ldots,p\}$ is a size-α test that is uniformly more powerful than the LRT. For the special case of $p = 2$, this provides a test that is uniformly more powerful than a test proposed by Gail and Simon (1985) for testing for a qualitative interaction.

1. TESTING PROBLEM AND LRT

Let $\underset{\sim}{X}' = (X_1,\ldots,X_p)$ be a p-variate ($p \geq 2$) normal random vector with unknown mean $\underset{\sim}{\mu}$ and nonsingular covariance matrix $\underset{\sim}{\Sigma}$. We will use the notation $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$. Throughout the paper, except in Section 5, $\underset{\sim}{\Sigma}$ will be assumed to be known. The results in this paper can be considered to be approximately true if $\underset{\sim}{\Sigma}$ is unknown but a large sample is available for estimating $\underset{\sim}{\Sigma}$. In many applications, $\underset{\sim}{\Sigma}$ will be a diagonal matrix, that is, the p populations with means $\mu_1,\ldots,\mu_p$ will be independent populations and X_i will be the sample mean of a random sample from the i^{th} population. But we will consider the more general setting.

Let $\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_k$ be k ($k \geq 2$) specified p-dimensional vectors. We consider testing the null hypothesis

$$H_0: \ \underset{\sim}{b}_i'\underset{\sim}{\mu} \leq 0 \text{ for some } i = 1,\ldots,k$$

(1.1) versus the alternative hypothesis

$$H_1: \ \underset{\sim}{b}_i'\underset{\sim}{\mu} > 0 \text{ for all } i = 1,\ldots,k \ .$$

For this to be meaningful, H_1 must be nonempty. (We use the symbol H_1 to denote the set of $\underset{\sim}{\mu}$ vectors specified by the hypothesis, as well as the statement of the hypothesis.) This would not be the case, for example, if $\underset{\sim}{b}_1 = -\underset{\sim}{b}_2$. To simplify the discussion we also assume that there are no redundant vectors in $\{\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_k\}$. That is there is no $\underset{\sim}{b}_j$ such that $\{\underset{\sim}{\mu}: \ \underset{\sim}{b}_i'\underset{\sim}{\mu} > 0, \ i = 1,\ldots,k\} = \{\underset{\sim}{\mu}: \ \underset{\sim}{b}_i'\underset{\sim}{\mu} > 0, \ i = 1,\ldots,k, \ i \neq j\}$. The requirement that there are no redundant vectors only simplifies notation and proofs and in no way restricts the hypothesis testing problems we are considering. If $p = 2$, then $k = 2$ because any larger set of $\underset{\sim}{b}$s would have a redundant vector. But if $p \geq 3$, then k might be any number bigger than one. Sasabuchi (1980) discusses other conditions that are equivalent to the requirement that H_1 is nonempty and $\{\underset{\sim}{b}_1,\ldots\underset{\sim}{b}_k\}$ has no redundant vectors.

Sasabuchi (1980) showed that the size-α LRT of H_0 versus H_1 is the test that rejects H_0 if

$$(1.2) \qquad Z_i = \frac{\underset{\sim}{b}_i'\underset{\sim}{X}}{\sqrt{\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_i}} \geq z_\alpha \quad \text{for all } i = 1,\ldots,k$$

where z_α is the upper 100α percentile of the standard normal distribution.

Actually, Sasabuchi (1980) considered a slightly different testing problem. The null hypothesis considered by Sasabuchi was

$$H_0: \ \underset{\sim}{b}_i'\underset{\sim}{\mu} \geq 0 \text{ for all } i = 1,\ldots,k \text{ with equality for at least one } i.$$

Sasabuchi's alternative hypothesis was the same as ours. So Sasabuchi was testing the null hypothesis that $\underset{\sim}{\mu}$ is on the boundary of the convex polyhedral cone defined by $\{\underset{\sim}{\mu}: \ \underset{\sim}{b}_i'\underset{\sim}{\mu} \geq 0 \text{ for all } i = 1,\ldots,k\}$ versus the alternative that $\underset{\sim}{\mu}$ is in the interior of the cone. We are testing the null hypothesis that $\underset{\sim}{\mu}$ is not in the interior of the cone versus the alternative that it is in the interior. In some cases our formulation may be more appropriate because the hypotheses do not artificially restrict the natural parameter space of $\underset{\sim}{\mu}$. But it is easy to modify Sasabuchi's argument to see that, in either case, the LRT has a rejection region of the form, reject H_0 if $Z_i \geq c$ for all $i = 1,\ldots,k$. To show that (1.2) is the size-α LRT for (1.1), it remains to show that $c = z_\alpha$ yields a size-α test. Our null hypothesis is a much larger set than Sasabuchi's. So when we take the supremum over H_0 of the rejection probability, we could get a larger value of the size of the test. But, in fact, the suprema over both sets are the same, as we will show in Section 3, and (1.2) does define the size-α LRT in our problem as well as Sasabuchi's.

The LRT may be thought of as an *intersection-union* test. Consider the individual testing problems of testing H_{0i}: $b_i'\mu \leq 0$ versus H_{1i}: $b_i'\mu > 0$. The uniformly most powerful size-α test of H_{0i} versus H_{1i} rejects H_{0i} if $Z_i \geq z_\alpha$. Our H_0 is the union of the H_{0i}, $i = 1,\ldots,k$. H_0 can be rejected only if every one of the H_{0i}s is rejected and that is what the LRT does. The LRT rejects H_0 if and only if every H_{0i} is rejected, that is, every $Z_i \geq z_\alpha$. A test constructed in this way was called an intersection-union test by Gleser (1973). More recent papers that use intersection-union tests include Berger (1982), Cohen, Gatsonis and Marden (1983), and Berger and Sinclair (1984).

The LRT also has two optimality properties. In some cases, it has been shown to be uniformly most powerful among all monotone, level-α tests and admissible. A test is said to be *monotone* if x is a sample point in the rejection region of the test and x^* is another sample point that satisfies $b_i'x^* \geq b_i'x$ for all $i = 1,\ldots,k$ then x^* is also in the rejection region of the test. Under various conditions it has been shown that the LRT is uniformly most powerful among all level-α tests with this monotonicity property. Lehmann (1952) first proved a result of this type. Cohen and Marden (1983) prove the result under various conditions. Cohen and Marden (1983) also show that, in a bivariate problem, the LRT is *admissible* in that no other test exists that has a uniformly smaller power function on H_0 and a uniformly bigger power function on H_1. This result has been generalized by Nomakuchi and Sakata (1987).

Despite these good properties the LRT has some deficiencies. It is a biased test in that the power function is less than α for some $\mu \in H_1$. In fact, this bias can be quite extreme. For example, suppose $\Sigma = \text{diag}(\sigma_1^2,\ldots,\sigma_p^2)$ is a diagonal matrix and consider the sign testing problem:

$$H_0^S: \quad \mu_i \le 0 \quad \text{for some } i = 1,\dots,p$$

(1.3) versus

$$H_1^S: \quad \mu_i > 0 \quad \text{for all } i = 1,\dots,p .$$

The LRT rejects H_0^S if $Z_i = X_i/\sigma_i \ge z_\alpha$ for all $i = 1,\dots,p$. If $\underset{\sim}{\mu} = \underset{\sim}{0}$, $Z_1,\dots,Z_p$ are independent normal (0,1) random variables. So the power function at $\underset{\sim}{\mu} = \underset{\sim}{0}$ is $\beta(\underset{\sim}{0}) = P_{\underset{\sim}{0}}(Z_1 \ge z_\alpha,\dots,Z_p \ge z_\alpha) = \alpha^p$ which is much less than α. Of course, $\underset{\sim}{\mu} = \underset{\sim}{0} \varepsilon H_0$ but the power function is continuous. So for $\underset{\sim}{\mu} \varepsilon H_1$ that are close to $\underset{\sim}{0}$, the power will be approximately α^p. To some extent this bias is unavoidable. A result of Lehmann (1952) shows that in some problems of the type we are considering, *no* unbiased, nonrandomized tests exists. Nomakuchi and Sakata (1987) also discuss this. But, in fact, tests do exist that have the same size as the LRT and are uniformly more powerful. Tests with this property are described in Sections 3 and 4. For example, for the above problem, the test in Section 3 has power equal to $\alpha^{p-1}/2$ at $\underset{\sim}{\mu} = \underset{\sim}{0}$. Thus this test's power is $(\alpha^{p-1}/2)/\alpha^p = 1/2\alpha$ times as big as the LRT's at some parameter points. This is a tenfold increase if $\alpha = .05$ and a fiftyfold increase if $\alpha = .01$.

Robertson and Wegman (1978) found the LRT for the testing problem in which H_1 is the null hypothesis and H_0 is the alternative hypothesis. That is the null hypothesis states that $\underset{\sim}{\mu}$ is in a cone and the alternative hypothesis states that $\underset{\sim}{\mu}$ is not in the cone. The test statistic is quite different in that case. The test statistic involves isotonic regression estimates of $\underset{\sim}{\mu}$ and the critical values for the test are percentiles for weighted sums of chi-squared or beta distributions. Our formulation of the problem is the appropriate one if one wishes to determine if the data confirm that the inequalities specified in H_1 are true.

For many computations, it is more convenient to consider this transformed version of the original problem that was used by Sasabuchi (1980). Let $\underset{\sim}{T}$ be a p x p nonsingular matrix such that $\underset{\sim}{T}\underset{\sim}{\Sigma}\underset{\sim}{T}' = \underset{\sim}{I}_p$, the p x p identity matrix. Thus $\underset{\sim}{T}^{-1}(\underset{\sim}{T}^{-1})' = \underset{\sim}{\Sigma}$. Make the transformation $\underset{\sim}{Y} = \underset{\sim}{T}\underset{\sim}{X}$. Then $\underset{\sim}{Y} \sim N_p(\underset{\sim}{\theta},\underset{\sim}{I}_p)$ where $\underset{\sim}{\theta} = \underset{\sim}{T}\underset{\sim}{\mu}$. Define the vectors $\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_k$ by $\underset{\sim}{a}_i' = \underset{\sim}{b}_i'\underset{\sim}{T}^{-1}$. Then $\underset{\sim}{b}_i'\underset{\sim}{\mu} = \underset{\sim}{a}_i'\underset{\sim}{\theta}$. So $\{\underset{\sim}{\theta}: \underset{\sim}{a}_i'\underset{\sim}{\theta} > 0, i = 1,\ldots,k\} \neq \Phi$ if and only if $\{\underset{\sim}{\mu}: \underset{\sim}{b}_i'\underset{\sim}{\mu} > 0, i = 1,\ldots,k\} \neq \Phi$ and $\{\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_k\}$ has no redundant vectors if and only if $\{\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_k\}$ has no redundant vectors. Thus our original testing problem is equivalent to observing $\underset{\sim}{Y}$ and testing

$$H_0: \quad \underset{\sim}{a}_i'\underset{\sim}{\theta} \leq 0 \quad \text{for some } i = 1,\ldots,k$$

versus

$$H_1: \quad \underset{\sim}{a}_i'\underset{\sim}{\theta} > 0 \quad \text{for all } i = 1,\ldots,k .$$

The LRT rejects H_0 if $Z_i = (\underset{\sim}{a}_i'\underset{\sim}{Y})/\sqrt{\underset{\sim}{a}_i'\underset{\sim}{a}_i} \geq z_\alpha$ for all $i = 1,\ldots,k$. We will consistently use the notation $\underset{\sim}{X}$, $\underset{\sim}{\mu}$, $\underset{\sim}{b}_i$ for quantities in the original problem and $\underset{\sim}{Y}$, $\underset{\sim}{\theta}$, $\underset{\sim}{a}_i$ for quantities in the transformed problem.

In Section 2, we prove some preliminary results that will later be used to show that various tests are size-α tests. The reader may only wish to read the theorems' statements on first reading. But Definitions 2.1 and 2.2 should be noted. In Section 3 we describe a size-α test that is uniformly more powerful than the LRT. We compare the powers of the two tests for the sign testing problem (1.3) when p = 2. In Section 4 we discuss an even more powerful test for the sign testing problem (1.3). In Section 5, the sign testing problem (1.3) with p = 2 is considered with an unknown variance. In Section 6, a two-sided version of the problem is considered and a size-α test that is uniformly more powerful than the LRT is described for a sign testing problem.

2. PRELIMINARY THEOREMS

The following results will be used in subsequent sections to prove that various tests are size-α tests. If $\underset{\sim}{g}$ is a vector, we use the notation $\|\underset{\sim}{g}\| = \sqrt{\underset{\sim}{g}'\underset{\sim}{g}}$.

<u>**LEMMA 2.1**</u>: Let $\underset{\sim}{g}$ and $\underset{\sim}{h}$ be two vectors such that $-\|\underset{\sim}{g}\|\cdot\|\underset{\sim}{h}\| < \underset{\sim}{g}'\underset{\sim}{h} \leq 0$. Define $\underset{\sim}{d} = \gamma\underset{\sim}{g} + \delta\underset{\sim}{h}$, where γ and δ are the numbers

$$\gamma = \frac{-\underset{\sim}{g}'\underset{\sim}{h}}{\|\underset{\sim}{g}\| \sqrt{\|\underset{\sim}{g}\|^2\|\underset{\sim}{h}\|^2 - (\underset{\sim}{g}'\underset{\sim}{h})^2}}$$

and

$$\delta = \frac{\|\underset{\sim}{g}\|}{\sqrt{\|\underset{\sim}{g}\|^2\|\underset{\sim}{h}\|^2 - (\underset{\sim}{g}'\underset{\sim}{h})^2}} \quad .$$

Then $\underset{\sim}{d}$ satisfies (i) $\|\underset{\sim}{d}\| = 1$, (ii) $\underset{\sim}{d}'\underset{\sim}{g} = 0$, (iii) $\underset{\sim}{d}'\underset{\sim}{h} > 0$, (iv) if $\underset{\sim}{y}$ is such that both $\underset{\sim}{g}'\underset{\sim}{y} \geq 0$ and $\underset{\sim}{h}'\underset{\sim}{y} \geq 0$ then $\underset{\sim}{d}'\underset{\sim}{y} \geq 0$, and (v) if $\underset{\sim}{y}$ is such that both $\underset{\sim}{g}'\underset{\sim}{y} \leq 0$ and $\underset{\sim}{h}'\underset{\sim}{y} \leq 0$ then $\underset{\sim}{d}'\underset{\sim}{y} \leq 0$.

<u>**PROOF**</u>: Note that the conditions on $\underset{\sim}{g}$ and $\underset{\sim}{h}$ imply that $\underset{\sim}{g} \neq 0$ and $\underset{\sim}{h} \neq 0$ and, hence, all ratios are well defined. (i), (ii), and (iii) are easily verified. Notice that $\delta > 0$ and $\gamma \geq 0$. So if $\underset{\sim}{g}'\underset{\sim}{y} \geq 0$ and $\underset{\sim}{h}'\underset{\sim}{y} \geq 0$, then $\underset{\sim}{d}'\underset{\sim}{y} = \gamma(\underset{\sim}{g}'\underset{\sim}{y}) + \delta(\underset{\sim}{h}'\underset{\sim}{y}) \geq 0$, verifying (iv). (v) is similar. ∎

<u>**LEMMA 2.2**</u>: Let $\underset{\sim}{g}$, $\underset{\sim}{h}$ and $\underset{\sim}{d}$ be as in Lemma 2.1. If $\underset{\sim}{y}$ is a vector and c is a number such that $\underset{\sim}{g}'\underset{\sim}{y}/\|\underset{\sim}{g}\| \geq c$ and $\underset{\sim}{h}'\underset{\sim}{y}/\|\underset{\sim}{h}\| \geq c$, then

$$\underset{\sim}{d}'\underset{\sim}{y} \geq c \sqrt{\frac{\|\underset{\sim}{g}\|\cdot\|\underset{\sim}{h}\| - \underset{\sim}{g}'\underset{\sim}{h}}{\|\underset{\sim}{g}\|\cdot\|\underset{\sim}{h}\| + \underset{\sim}{g}'\underset{\sim}{h}}} \quad .$$

If $\underset{\sim}{y}$ is a vector and c is a number such that $\underset{\sim}{g}'\underset{\sim}{y}/\|\underset{\sim}{g}\| \leq c$ and $\underset{\sim}{h}'\underset{\sim}{y}/\|\underset{\sim}{h}\| \leq c$, then

$$\mathbf{d}'\mathbf{y} \le c \sqrt{\frac{\|\mathbf{g}\|\cdot\|\mathbf{h}\| - \mathbf{g}'\mathbf{h}}{\|\mathbf{g}\|\cdot\|\mathbf{h}\| + \mathbf{g}'\mathbf{h}}} .$$

<u>**PROOF:**</u> We will prove the first result. Replace all the "$\ge$" with "$\le$" in the following to prove the second result. Define the vector

$$\mathbf{f} = \frac{c}{\|\mathbf{g}\|}\mathbf{g} + \frac{c(\|\mathbf{g}\|\cdot\|\mathbf{h}\| - \mathbf{g}'\mathbf{h})}{\|\mathbf{g}\|(\mathbf{h}'\mathbf{d})}\mathbf{d} .$$

Recalling that $\mathbf{g}'\mathbf{d} = 0$, notice that $\mathbf{g}'\mathbf{f}/\|\mathbf{g}\| = c$ and $\mathbf{h}'\mathbf{f}/\|\mathbf{h}\| = c$. Thus $\mathbf{g}'(\mathbf{y} - \mathbf{f})/\|\mathbf{g}\| \ge 0$ and $\mathbf{h}'(\mathbf{y} - \mathbf{f})/\|\mathbf{h}\| \ge 0$ and, hence, $\mathbf{g}'(\mathbf{y} - \mathbf{f}) \ge 0$ and $\mathbf{h}'(\mathbf{y} - \mathbf{f}) \ge 0$. By Lemma 2.1, this implies $\mathbf{d}'(\mathbf{y} - \mathbf{f}) \ge 0$, that is, $\mathbf{d}'\mathbf{y} \ge \mathbf{d}'\mathbf{f}$. It is easily verified that

$$\mathbf{d}'\mathbf{f} = c \sqrt{\frac{\|\mathbf{g}\|\cdot\|\mathbf{h}\| - \mathbf{g}'\mathbf{h}}{\|\mathbf{g}\|\cdot\|\mathbf{h}\| + \mathbf{g}'\mathbf{h}}} ,$$

proving the first result. ∎

The constants used to define the rejection regions for our tests were defined in the Abstract. For completeness we repeat the definition.

DEFINITION 2.1: For $0 < \alpha < .5$, define the integer J by the inequality $J - 1 < 1/2\alpha \le J$. Define the constants $c_0,\dots,c_{2J}$ as follows: $c_0 = \infty$, $c_j = z_{j\alpha}$, $j = 1,\dots,J-1$; $c_J = 0$; $c_j = -c_{2J-j}$, $j = J+1,\dots,2J$

Notice that $c_0 > c_1 > \dots > c_{2J}$. If $1/2\alpha$ is an integer (as it is for $\alpha = .10$, $.05$ and $.01$), then $c_1, c_2, \dots, c_{2J-1}$ are the normal percentiles, z_α, $z_{2\alpha},\dots,z_{(2J-1)\alpha}$. For any α, if Z has a standard normal distribution, then $P(c_j \le Z \le c_{j-1}) = \alpha$ for $j = 1,\dots, J-1$, and $j = J+2,\dots,2J$. $P(c_J \le Z \le c_{J-1}) = P(c_{J+1} \le Z \le c_J) \le \alpha$ with equality if $1/2\alpha$ is an integer.

LEMMA 2.3: Let $\underset{\sim}{g}$ and $\underset{\sim}{h}$ satisfy the conditions in Lemma 2.1. Define the sets $S_1^*,\ldots,S_{2J}^*$ by

$$(2.1)\qquad S_j^* = \left\{\underset{\sim}{y}:\ c_j \le \frac{\underset{\sim}{g}'\underset{\sim}{y}}{\|\underset{\sim}{g}\|} \le c_{j-1},\quad c_j \le \frac{\underset{\sim}{h}'\underset{\sim}{y}}{\|\underset{\sim}{h}\|} \le c_{j-1}\right\} .$$

Let $\underset{\sim}{Y} \sim N_p(\underset{\sim}{\theta},\underset{\sim}{I}_p)$. If $\underset{\sim}{g}'\underset{\sim}{\theta} = 0$, then

$$P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in \bigcup_{j=1}^{2J} S_j^*) \le \alpha \quad .$$

PROOF: Let

$$r = \sqrt{\frac{\|\underset{\sim}{g}\|\cdot\|\underset{\sim}{h}\| - \underset{\sim}{g}'\underset{\sim}{h}}{\|\underset{\sim}{g}\|\cdot\|\underset{\sim}{h}\| + \underset{\sim}{g}'\underset{\sim}{h}}} \quad .$$

Let $\underset{\sim}{d}$ be the vector described in Lemma 2.1. Define the sets $S_1^\dagger,\ldots,S_{2J}^\dagger$ by

$$S_j^+ = \left\{\underset{\sim}{y}:\ c_j \le \frac{\underset{\sim}{g}'\underset{\sim}{y}}{\|\underset{\sim}{g}\|} \le c_{j-1},\quad c_j r \le \underset{\sim}{d}'\underset{\sim}{y} \le c_{j-1} r\right\} .$$

Lemma 2.2 implies that $S_j^* \subset S_j^\dagger$. Also

$$S_j^+ \cap S_{j+1}^+ \subset \left\{\underset{\sim}{y}:\ \frac{\underset{\sim}{g}'\underset{\sim}{y}}{\|\underset{\sim}{g}\|} = c_j\right\} ,$$

a set with probability zero, and $S_j^\dagger \cap S_i^\dagger = \phi$ if $|j-i| > 1$. Thus

$$(2.2)\qquad P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in \bigcup_{j=1}^{2J} S_j^*) \le P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in \bigcup_{j=1}^{2J} S_j^+) = \sum_{j=1}^{2J} P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in S_j^+) .$$

The random variables $\underset{\sim}{g}'\underset{\sim}{Y}/\|\underset{\sim}{g}\|$ and $\underset{\sim}{d}'\underset{\sim}{Y}$ are independent normal random variables since $\underset{\sim}{d}'\underset{\sim}{I}_p\underset{\sim}{g} = \underset{\sim}{d}'\underset{\sim}{g} = 0$. And $\underset{\sim}{g}'\underset{\sim}{Y}/\|\underset{\sim}{g}\|$ has a standard normal distribution if $\underset{\sim}{g}'\underset{\sim}{\theta} = 0$. Thus

$$\sum_{j=1}^{2J} P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in S_j^+) = \sum_{j=1}^{2J} P_{\underset{\sim}{\theta}}\left(c_j \le \frac{\underset{\sim}{g}'\underset{\sim}{Y}}{\|\underset{\sim}{g}\|} \le c_{j-1},\quad c_j r \le \underset{\sim}{d}'\underset{\sim}{Y} \le c_{j-1} r\right)$$

$$= \sum_{j=1}^{2J} P_{\underset{\sim}{\theta}}\left(c_j \le \frac{\underset{\sim}{g}'\underset{\sim}{Y}}{\|\underset{\sim}{g}\|} \le c_{j-1}\right) P_{\underset{\sim}{\theta}}\left(c_j r \le \underset{\sim}{d}'\underset{\sim}{Y} \le c_{j-1} r\right) \quad \text{(independence)}$$

$$\leq \sum_{j=1}^{2J} \alpha P_{\theta}(c_j r \leq d'Y \leq c_{j-1} r) \qquad \text{(property of } c_0,\ldots,c_{2J})$$

$$= \alpha \ .$$

This with (2.2) yields the desired result. ∎

The construction in the proof of Lemma 2.3 is illustrated in Figure 2.1 for the case, $p = k = 2$, $\alpha = .2$, $g' = (-1,2)$, $h' = (1,0)$. In this case $1/2\alpha = 2.5$ so $J = 3$ and $c_1 = -c_5 = .84$, $c_2 = -c_4 = .25$, $c_3 = 0$. The diamond shaped regions are the sets $S_1^*,\ldots,S_6^*$. The rectangular regions with dotted borders are the sets $S_1^\dagger,\ldots,S_6^\dagger$. Note that $S_j^* \subset S_j^\dagger$. It is because the angle $\eta \leq 90°$, i.e, $g'h \leq 0$, that we can construct the rectangles to contain the diamonds. The dotted lines with negative slope correspond to the ys for which $d'y = c_j r$, $j = 1,\ldots,5$.

The rejection regions for the tests we will consider are formed from the following sets.

DEFINITION 2.2. Let $z_i = z_i(x) = (b_i'x)/\sqrt{b_i'\Sigma b_i}$. For α, J, and $c_0,\ldots,c_{2J}$ as in Definition 2.1, define the following sets.

$$R_j = \{x: \ c_j \leq z_i \leq c_{j-1}, \quad i = 1,\ldots,k\} \qquad j = 1,\ldots,2J.$$

Under the transformation $y = Tx$ described in Section 1, the set R_j is mapped onto the set

$$S_j = \left\{y: \ c_j \leq \frac{a_i'y}{\sqrt{a_i'a_i}} = \frac{a_i'y}{\|a_i\|} \leq c_{j-1}, \quad i = 1,\ldots,k\right\} .$$

SETS FROM LEMMA 2.3

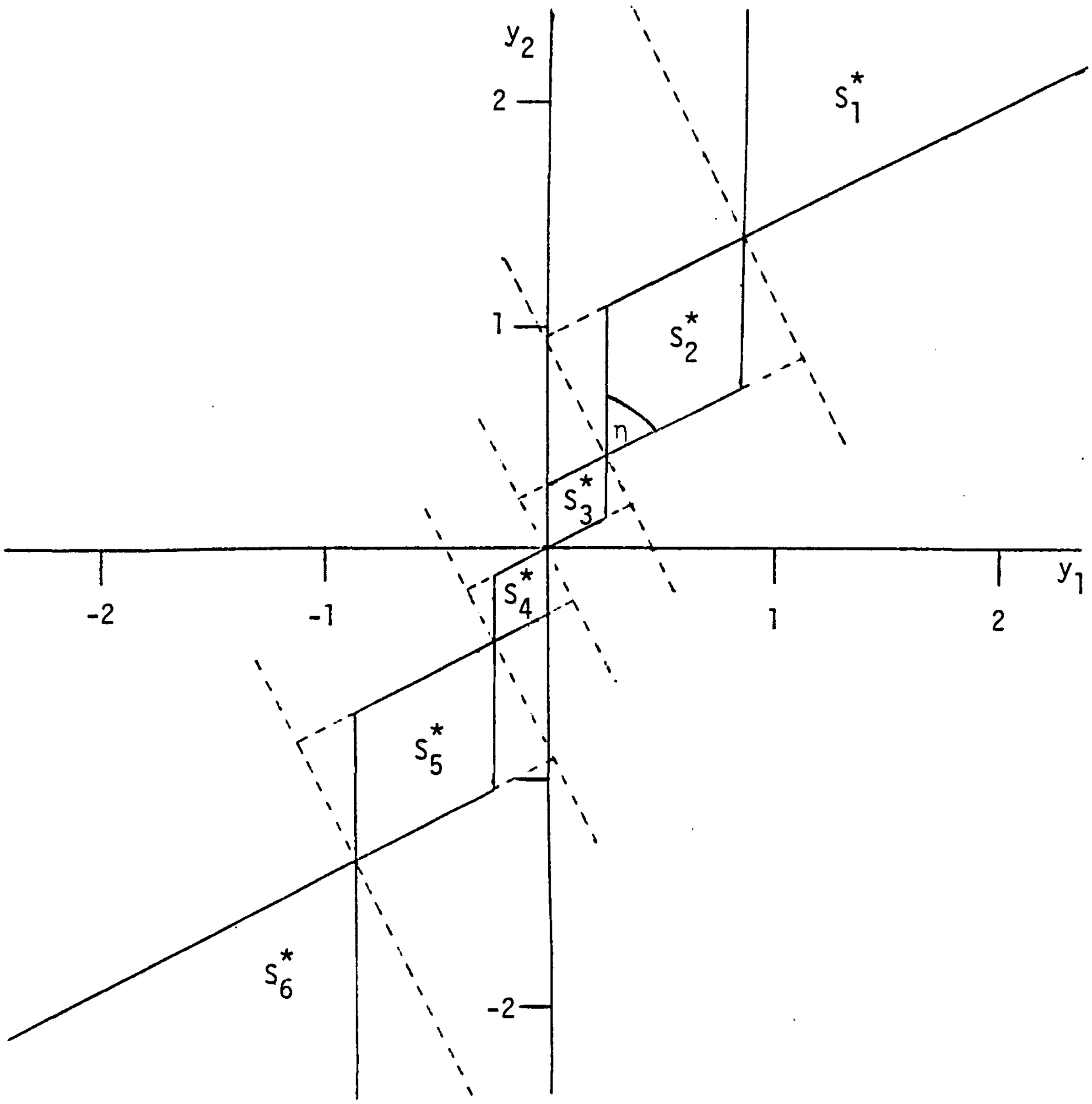

Figure 2.1

The following two theorems will be used to show that various tests are size-α. We state them in terms of the original quantities $\underset{\sim}{X}$, $\underset{\sim}{\mu}$, and $\underset{\sim}{b}_i$ as that is the context in which they will be used

THEOREM 2.1: Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$. Suppose the set $\{\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_k\}$ is such that H_1 in (1.1) is nonempty and has no redundant vectors. Suppose further that for each $i = 1,\ldots,k$ there is an $m \in \{1,\ldots,k\}$ (m will depend on i) such that $\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$. Let $0 < \alpha < .5$ and $c_0,\ldots,c_{2J}$ and $R_1,\ldots,R_{2J}$ be as in Definitions 2.1 and 2.2. If $\underset{\sim}{\mu}$ satisfies $\underset{\sim}{b}_i'\underset{\sim}{\mu} = 0$ for some $i \in \{1,\ldots,k\}$, then

$$P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in \bigcup_{j=1}^{2J} R_j) \leq \alpha.$$

PROOF: Using the transformation $\underset{\sim}{Y} = \underset{\sim}{T}\underset{\sim}{X}$, we have that

$$P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in \bigcup_{j=1}^{2J} R_j) = P_{\underset{\sim}{\theta}}(Y \in \bigcup_{j=1}^{2J} S_j)$$

where $\underset{\sim}{\theta} = \underset{\sim}{T}\underset{\sim}{\mu}$. Note that $\underset{\sim}{a}_i'\underset{\sim}{\theta} = \underset{\sim}{b}_i'\underset{\sim}{T}^{-1}\underset{\sim}{T}\underset{\sim}{\mu} = 0$. Let m be such that $\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$. Then $\underset{\sim}{a}_i'\underset{\sim}{a}_m = \underset{\sim}{b}_i'\underset{\sim}{T}^{-1}(\underset{\sim}{T}^{-1})'\underset{\sim}{b}_m = \underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$. Since H_1 is nonempty, $\underset{\sim}{a}_i'\underset{\sim}{a}_m > -\|\underset{\sim}{a}_i\|\cdot\|\underset{\sim}{a}_m\|$. ($\underset{\sim}{a}_i'\underset{\sim}{a}_m$ cannot be less than $-\|\underset{\sim}{a}_i\|\cdot\|\underset{\sim}{a}_m\|$ and $\underset{\sim}{a}_i'\underset{\sim}{a}_m = -\|\underset{\sim}{a}_i\|\cdot\|\underset{\sim}{a}_m\|$ implies $\underset{\sim}{a}_m = -r\underset{\sim}{a}_i$ for some positive constant r. But this would imply H_1 is empty.) Thus $\underset{\sim}{a}_i$ and $\underset{\sim}{a}_m$ satisfy the conditions on $\underset{\sim}{g}$ and $\underset{\sim}{h}$ in Lemmas 2.1, 2.2, and 2.3. Notice that with $\underset{\sim}{g} = \underset{\sim}{a}_i$ and $\underset{\sim}{h} = \underset{\sim}{a}_m$, S_j from Definition 2.1 is a subset of S_j^* from (2.1). Thus from Lemma 2.3 we have

$$P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in \bigcup_{j=1}^{2J} R_j) = P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in \bigcup_{j=1}^{2J} S_j) \leq P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in \bigcup_{j=1}^{2J} S_j^*) \leq \alpha . \quad \blacksquare$$

The second theorem we will use is quite general and unrelated to the special structure we have used up to now. But we have not found it stated in the literature in this generality.

THEOREM 2.2: Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$. Let R be a set and $\underset{\sim}{b}$ be a vector such that $\underset{\sim}{b}'\underset{\sim}{x} \geq 0$ for every $\underset{\sim}{x} \in R$. Let $\underset{\sim}{\mu}$ be a vector such that $\underset{\sim}{b}'\underset{\sim}{\mu} \leq 0$. Then there exists a vector $\underset{\sim}{\mu}^*$ such that $\underset{\sim}{b}'\underset{\sim}{\mu}^* = 0$ and $P_{\underset{\sim}{\mu}*}(\underset{\sim}{X} \in R) \geq P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in R)$.

PROOF: If $\underset{\sim}{b}'\underset{\sim}{\mu} = 0$ then $\underset{\sim}{\mu}^* = \underset{\sim}{\mu}$ satisfies the requirements. So assume $\underset{\sim}{b}'\underset{\sim}{\mu} < 0$. Let $\underset{\sim}{T}$ be the nonsingular matrix defined in Section 1. Make the transformation $\underset{\sim}{Y} = \underset{\sim}{T}\underset{\sim}{X}$. Let $Q = \underset{\sim}{T}R = \{\underset{\sim}{y}: \underset{\sim}{y} = \underset{\sim}{T}\underset{\sim}{x}, \underset{\sim}{x} \in R\}$. Then $\underset{\sim}{Y} \sim N_p(\underset{\sim}{\theta} = \underset{\sim}{T}\underset{\sim}{\mu},\underset{\sim}{I}_p)$ and $P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in R) = P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in Q)$. Let $\underset{\sim}{a}' = \underset{\sim}{b}'\underset{\sim}{T}^{-1}$. Then $\underset{\sim}{a}'\underset{\sim}{y} = \underset{\sim}{b}'\underset{\sim}{x} \geq 0$ for all $\underset{\sim}{y} \in Q$ and $\underset{\sim}{a}'\underset{\sim}{\theta} = \underset{\sim}{b}'\underset{\sim}{\mu} < 0$. Now let $\underset{\sim}{O} = (\underset{\sim}{o}_1,\ldots,\underset{\sim}{o}_p)'$ be an orthogonal matrix with $\underset{\sim}{o}_1 = \underset{\sim}{a}/\|\underset{\sim}{a}\|$. Make the transformation $\underset{\sim}{U} = \underset{\sim}{O}\underset{\sim}{Y}$. Let $P = \underset{\sim}{O}Q = \{\underset{\sim}{u}: \underset{\sim}{u} = \underset{\sim}{O}\underset{\sim}{y}, \underset{\sim}{y} \in Q\}$. Then $\underset{\sim}{U} \sim N_p(\underset{\sim}{\nu} = \underset{\sim}{O}\underset{\sim}{\theta},\underset{\sim}{I}_p)$ and $P_{\underset{\sim}{\nu}}(\underset{\sim}{U} \in P) = P_{\underset{\sim}{\theta}}(\underset{\sim}{Y} \in Q)$. For every $\underset{\sim}{u} \in P$, $u_1 = \underset{\sim}{a}'\underset{\sim}{y}/\|\underset{\sim}{a}\| \geq 0$. Also, $\nu_1 = \underset{\sim}{a}'\underset{\sim}{\theta}/\|\underset{\sim}{a}\| < 0$. Thus

$$P_{\underset{\sim}{\nu}}(\underset{\sim}{U} \in P) = \int\cdots\int_P (2\pi)^{-p/2} \exp -\left(\frac{1}{2}\sum_{i=1}^{p}(u_i-\nu_i)^2\right) du_1\ldots du_p$$

$$< \int\cdots\int_P (2\pi)^{-p/2} \exp -\left(\frac{1}{2}(u_1 - 0)^2 + \frac{1}{2}\sum_{i=2}^{p}(u_i-\nu_i)^2\right) du_1\ldots du_p$$

$$= P_{\underset{\sim}{\nu}*}(U \in P)$$

where $\underset{\sim}{\nu}^{*\prime} = (0,\nu_2,\ldots,\nu_p)$. Now making the two inverse transformations we have $\underset{\sim}{\mu}^* = \underset{\sim}{T}^{-1}\underset{\sim}{O}'\underset{\sim}{\nu}^*$ and $P_{\underset{\sim}{\mu}*}(\underset{\sim}{X} \in R) = P_{\underset{\sim}{\nu}*}(\underset{\sim}{U} \in P) > P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in R)$. Furthermore, since $\nu_1^* = 0$,

$$\underset{\sim}{O}'\underset{\sim}{\nu}^* = \sum_{i=2}^{p} \nu_i\underset{\sim}{o}_i \ .$$

Hence

$$\underset{\sim}{b}'\underset{\sim}{\mu}^* = \underset{\sim}{b}'\underset{\sim}{T}^{-1}\underset{\sim}{O}'\underset{\sim}{\nu}^* = \underset{\sim}{a}'\left(\sum_{i=2}^{p}\nu_i\underset{\sim}{o}_i\right) = \|\underset{\sim}{a}\|\underset{\sim}{o}_1'\left(\sum_{i=2}^{p}\nu_i\underset{\sim}{o}_i\right) \ .$$

But $\underset{\sim}{o}_1'\underset{\sim}{o}_i = 0$ for $i = 2,\ldots,p$ since $\underset{\sim}{O}$ was orthogonal. Hence $\underset{\sim}{b}'\underset{\sim}{\mu}^* = 0$ as was to be shown. ▮

Of course $\underset{\sim}{\mu}^*$ in Theorem 2.2 could be thought of as the projection of $\underset{\sim}{\mu}$ onto the $\{\underset{\sim}{\mu}\colon\ \underset{\sim}{b}'\underset{\sim}{\mu} = 0\}$ where the projection is in terms of a norm defined by $\underset{\sim}{\Sigma}$.

3. A TEST THAT IS MORE POWERFUL THAN THE LRT

Under certain conditions, the following test will be shown to be a size-α test that is uniformly more powerful than the LRT for the testing problem described in (1.1).

DEFINITION 3.1: For values of α that satisfy $0 < \alpha < .5$, define *Test I* to be the test that rejects H_0 if

$$\underset{\sim}{X} \in \bigcup_{j=1}^{J} R_j$$

where the sets R_j are defined in Definition 2.2. That is, if $Z_i = (\underset{\sim}{b}_i'\underset{\sim}{X})/\sqrt{\underset{\sim}{b}_i'\Sigma\underset{\sim}{b}_i}$ then Test I rejects H_0 if, for some $j \in \{1,\ldots,J\}$, $c_j \leq Z_i \leq c_{j-1}$ for all $i = 1,\ldots,k$.

Note that only half of the sets $R_1,\ldots,R_{2J}$ from Definition 2.2 are in the rejection region of Test I.

EXAMPLE 3.1: Let $p = k = 2$. Suppose X_1 and X_2 are independent and $X_i \sim N_1(\mu_i,\sigma_i^2)$. Let $\underset{\sim}{b}_1' = (1,0)$ and $\underset{\sim}{b}_2' = (0,1)$ so we are testing H_0: $\mu_1 \leq 0$ or $\mu_2 \leq 0$ versus H_1: $\mu_1 > 0$ and $\mu_2 > 0$. Then $Z_i = X_i/\sigma_i$ and Test I rejects H_0 if $c_j \leq Z_1,Z_2 \leq c_{j-1}$ for some $j \in \{1,\ldots,J\}$. For example, if $\alpha = .10$, then $J = 5$ and $c_1 = 1.28$, $c_2 = .84$, $c_3 = .52$, $c_4 = .25$ and $c_5 = 0$. So this rejection region consists of the five rectangles labelled $R_1,\ldots,R_5$ in Figure 3.1.

EXAMPLE 3.2: Let $p = k = 2$ and X_1 and X_2 be independent with $X_i \sim N_1(\mu_i,\sigma_i^2)$. Consider testing H_0: $2\mu_2 \leq \mu_1$ or $\mu_1 \leq 0$ versus H_1: $0 < \mu_1 < 2\mu_2$. Then $\underset{\sim}{b}_1' = (-1,2)$, $\underset{\sim}{b}_2' = (1,0)$, $Z_1 = (2X_2-X_1)/\sqrt{\sigma_1^2+4\sigma_2^2}$ and $Z_2 = X_1/\sigma_1$. For the case of $\sigma_1 = \sigma_2 = 1$ and $\alpha = .2$, these are the conditions used to construct Figure 2.1. The rejection region for Test I is $S_1^* \cup S_2^* \cup S_3^*$

REJECTION REGIONS FOR LRT, TEST I AND TEST II

IN BIVARIATE SIGN TESTING PROBLEM, α = .10

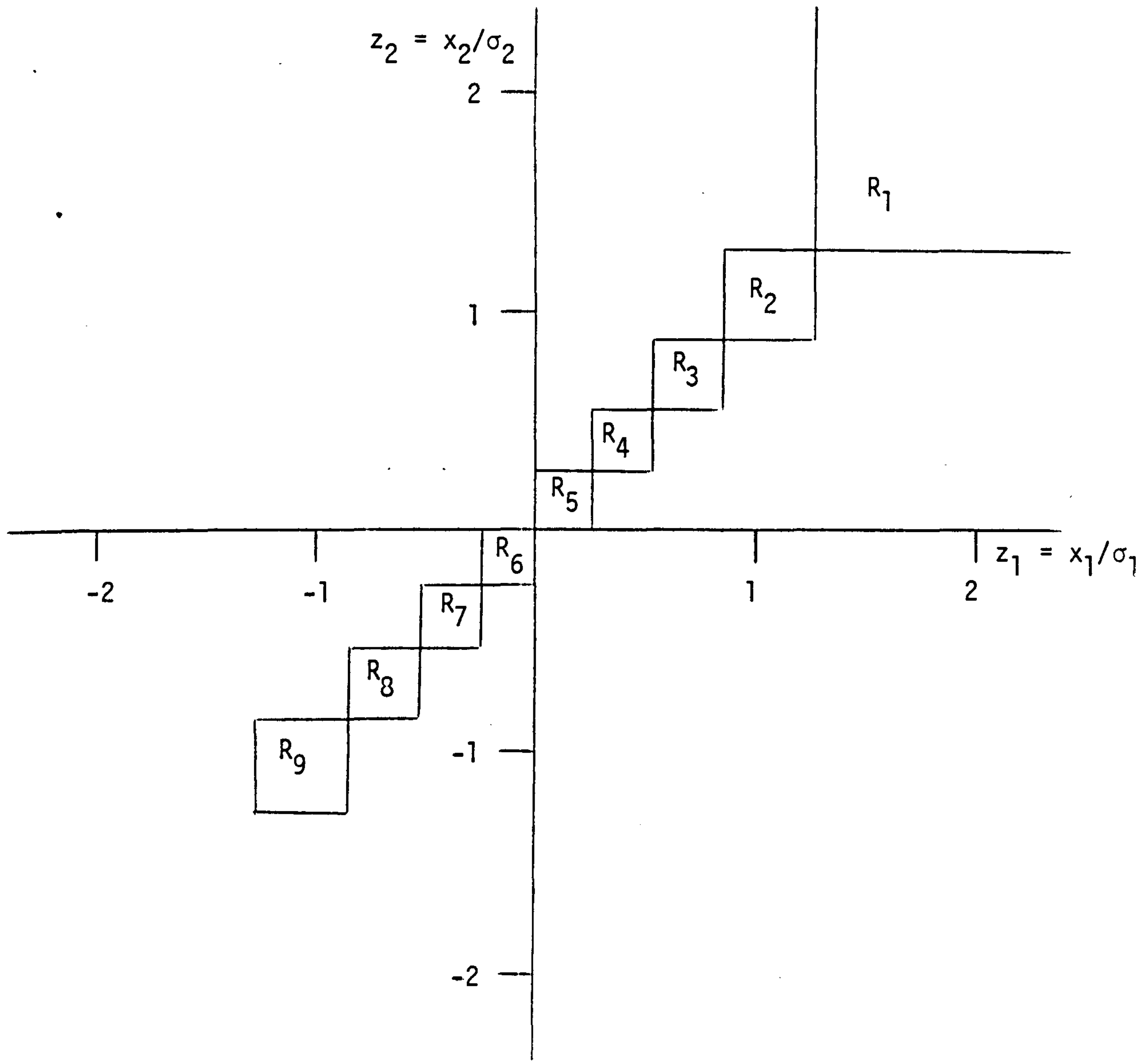

Figure 3.1

in Figure 2.1 where the axes are now the $x_1 - x_2$ axes. For a smaller, more common value of α, the picture would be similar but with more, but smaller, diamond shaped regions in the rejection region.

We now prove that Test I has the properties we desire.

THEOREM 3.1: For the testing problem described in (1.1), suppose that for each $i = 1,\ldots,k$ there exists and $m \in \{1,\ldots,k\}$ (m will depend on i) such that $\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$. If $0 < \alpha < .5$ then Test I is a size-α test and Test I is uniformly more powerful than the size-α LRT.

PROOF: The size-α LRT, as found by Sasabuchi (1980), rejects H_0 if $Z_i \geq z_\alpha$ for all $i = 1,\ldots,k$. But $c_0 = \infty$ and $c_1 = z_\alpha$. So the set R_1 is the rejection region of the size-α LRT. Since R_1 is a subset of the rejection region Test I, Test I is uniformly more powerful than the size-α LRT.

Let $H_s = \{\underset{\sim}{\mu}\colon\ \underset{\sim}{b}_i'\underset{\sim}{\mu} \geq 0$ for all $i = 1,\ldots,k$ and $\underset{\sim}{b}_i'\underset{\sim}{\mu} = 0$ for some $i\}$. Sasabuchi showed that $\sup_{\underset{\sim}{\mu}\in H_s} P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in R_1) = \alpha$, that is, the LRT is a size-α test for Sasabuchi's null hypothesis, H_s. But $H_s \subset H_0$ and

$$R_1 \subset \bigcup_{j=1}^{J} R_j ,$$

so

$$\begin{aligned} \alpha &= \sup_{\mu\in H_s} P_{\mu}(\underset{\sim}{X} \in R_1) \\ &\leq \sup_{\mu\in H_0} P_{\mu}(\underset{\sim}{X} \in \bigcup_{j=1}^{J} R_j) \qquad (3.1) \\ &= \text{size of Test I.} \end{aligned}$$

Now, let $\underset{\sim}{\mu} \in H_0$. Then there exists an i such that $\underset{\sim}{b}_i'\underset{\sim}{\mu} \leq 0$. For all

$$\underset{\sim}{X} \in \bigcup_{j=1}^{J} R_j ,$$

we have

$$\frac{\underset{\sim}{b}_i'\underset{\sim}{x}}{\sqrt{\underset{\sim}{b}_i'\Sigma\underset{\sim}{b}_i}} = z_i \geq c_J = 0$$

and, hence, $\underset{\sim}{b}_i'\underset{\sim}{x} \geq 0$. Thus by Theorem 2.2, there is a $\underset{\sim}{\mu}^*$ with $\underset{\sim}{b}_i'\underset{\sim}{\mu}^* = 0$ such that

$$(3.2) \qquad P_{\mu^*}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{J} R_j) \geq P_{\mu}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{J} R_j) .$$

By Theorem 2.1, the conditions on $\{\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_k\}$ imply that

$$(3.3) \qquad \alpha \geq P_{\mu^*}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{2J} R_j) > P_{\mu^*}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{J} R_j) .$$

Since $\underset{\sim}{\mu} \varepsilon H_0$ was arbitrary, (3.2) and (3.3) imply

$$(3.4) \qquad \alpha \geq \sup_{\mu \varepsilon H_0} P_{\mu}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{J} R_j) = \text{size of Test I.}$$

(3.1) and (3.4) show that Test I is a size-α test. ∎

It may seem curious that one can take a size-α test (the LRT), add sets of positive probability to the rejection region, and still have a size-α test. This is possible because, although $\sup_{\underset{\sim}{\mu}\varepsilon H_s} P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \varepsilon R_1) = \alpha$, $P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \varepsilon R_1) < \alpha$ for every $\underset{\sim}{\mu} \varepsilon H_0$. Sasabuchi (1980) showed that the supremum was only attained in a limit as one $\underset{\sim}{b}_i'\underset{\sim}{\mu} = 0$ and all other $\underset{\sim}{b}_j'\underset{\sim}{\mu} \to \infty$. Test I's power function satisfies

$$P_{\mu}(\underset{\sim}{X} \varepsilon R_1) < P_{\mu}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{J} R_j) < \alpha$$

for all $\underset{\sim}{\mu} \varepsilon H_0$.

To illustrate quantitatively the improvement in power that is provided by Test I, consider again the bivariate sign testing problem from Example 3.1. We use $\sigma_1 = \sigma_2 = 1$. We use $\alpha = .10$ so the rejection region for Test I is

$R_1 \cup \ldots \cup R_5$ in Figure 3.1 and the rejection region for the LRT is just R_1. Let

$$\beta_I(\underset{\sim}{\mu}) = P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{J} R_j) \quad \text{and} \quad \beta_L(\underset{\sim}{\mu}) = P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \varepsilon R_1)$$

be the power functions of Test I and the LRT, respectively. These two functions are graphed for certain $\underset{\sim}{\mu}$ values in Figures 3.2a, b, and c. In each figure the lowest graph is $\beta_L(\underset{\sim}{\mu})$ and the middle graph is $\beta_I(\underset{\sim}{\mu})$. (The top graph is the graph of the power function of a test that will be described in Section 4.) In Figure 3.2a, the graphs are for values of $\underset{\sim}{\mu}' = (0,\mu)$, $\mu \geq 0$. These values are on the boundary of H_0 so the graphs are everywhere less than $\alpha = .10$. An unbiased test would have a power function equal to $\alpha = .10$ for all these $\underset{\sim}{\mu}$ values. Test I and the LRT are biased but Test I is less so, having a minimum value of $\beta_I(\underset{\sim}{0}) = .05$ rather than $\beta_L(\underset{\sim}{0}) = .01$. Figure 3.2b provides power function graphs for mean vectors on the diagonal, $\underset{\sim}{\mu}' = (\mu,\mu)$, $\mu \geq 0$. $\beta_I(\underset{\sim}{\mu})$ is noticably above $\beta_L(\underset{\sim}{\mu})$ for $\mu \leq 2$ with the largest difference, $\beta_I(\underset{\sim}{\mu}) - \beta_L(\underset{\sim}{\mu}) \approx .07$ occurring is the range $.5 < \mu < 1$. Figure 3.2c has power function graphs for mean vectors of the form $\underset{\sim}{\mu}' = (.5\mu,\mu)$, $\mu \geq 0$. $\beta_I(\underset{\sim}{\mu})$ is noticably larger than $\beta_L(\underset{\sim}{\mu})$ for $\mu \leq 3$ with the maximum difference, $\beta_I(\underset{\sim}{\mu}) - \beta_L(\underset{\sim}{\mu}) \approx .06$ occurring the the range $.5 < \mu < 1.1$. Another comparison of β_I to β_L is given in Figure 3.3 where contours of the ratio $\beta_I(\underset{\sim}{\mu})/\beta_L(\underset{\sim}{\mu})$ are graphed. The ratio is always bigger than one, of course, since $\beta_I(\underset{\sim}{\mu}) > \beta_L(\underset{\sim}{\mu})$. The graph shows that the ratio is 5 at $\underset{\sim}{\mu} = \underset{\sim}{0}$, as is easily calculated, $.05/.01 = 5$. The ratio converges to one as $\mu_1 \to \infty$ and $\mu_2 \to \infty$, because, in this case both $\beta_I(\underset{\sim}{\mu}) \to 1$ and $\beta_L(\underset{\sim}{\mu}) \to 1$. But the ratio remains large for a reasonable range of $\underset{\sim}{\mu}$ values. For example, Test I is 40% better than the LRT, that is, $\beta_I(\underset{\sim}{\mu})/\beta_L(\underset{\sim}{\mu}) = 1.4$, for $\underset{\sim}{\mu}$ with $\mu_1 + \mu_2 \approx 1.8$. (The contours in Figure 3.3 are not straight lines. We are

POWER FUNCTIONS FOR $\mu_1 = 0, \quad 0 < \mu_2 < 4$

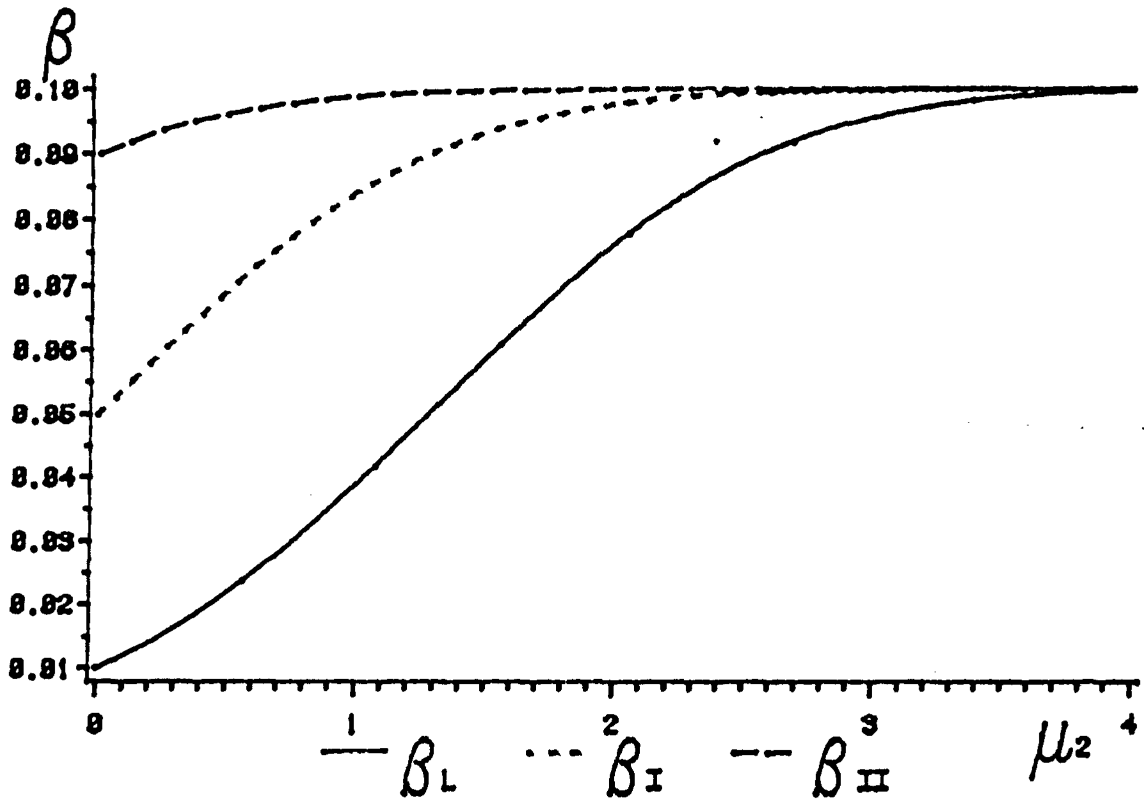

Figure 3.2a

POWER FUNCTIONS FOR $\mu_1 = \mu_2$, $0 < \mu_2 < 4$

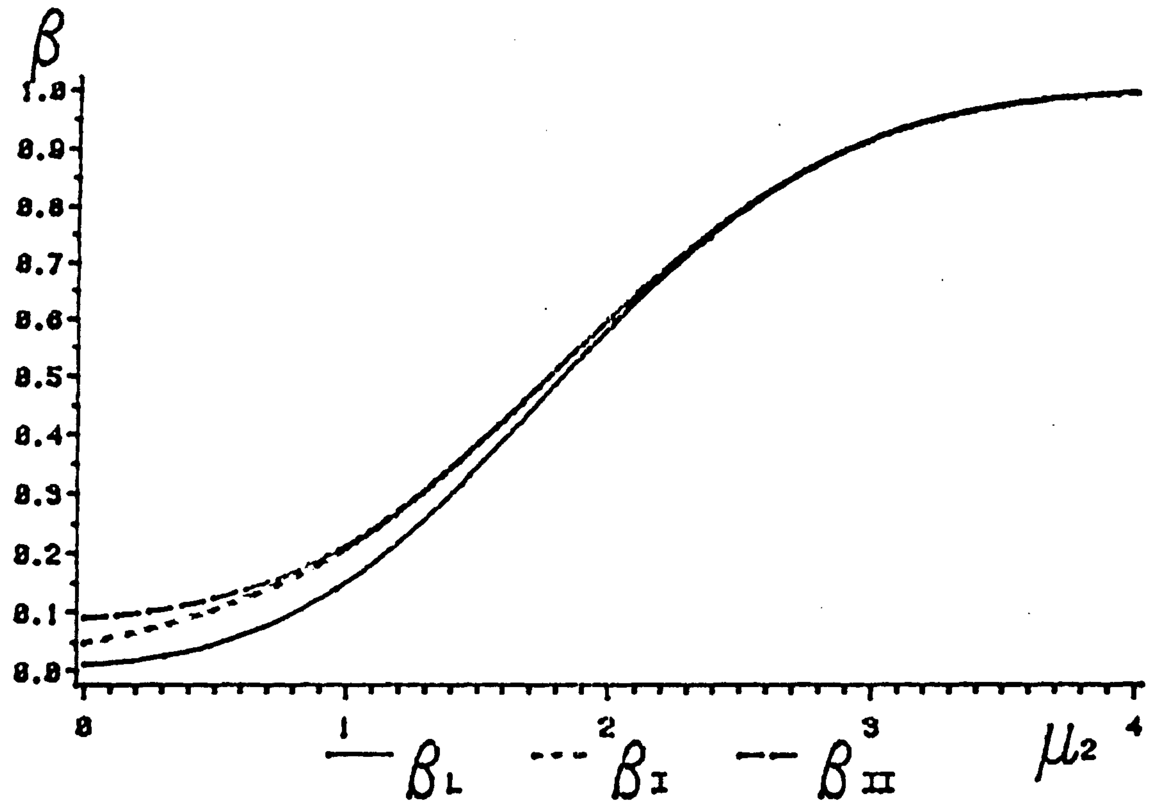

Figure 3.2b

POWER FUNCTIONS FOR $\mu_1 = .5*\mu_2, \quad 0 < \mu_2 < 4$

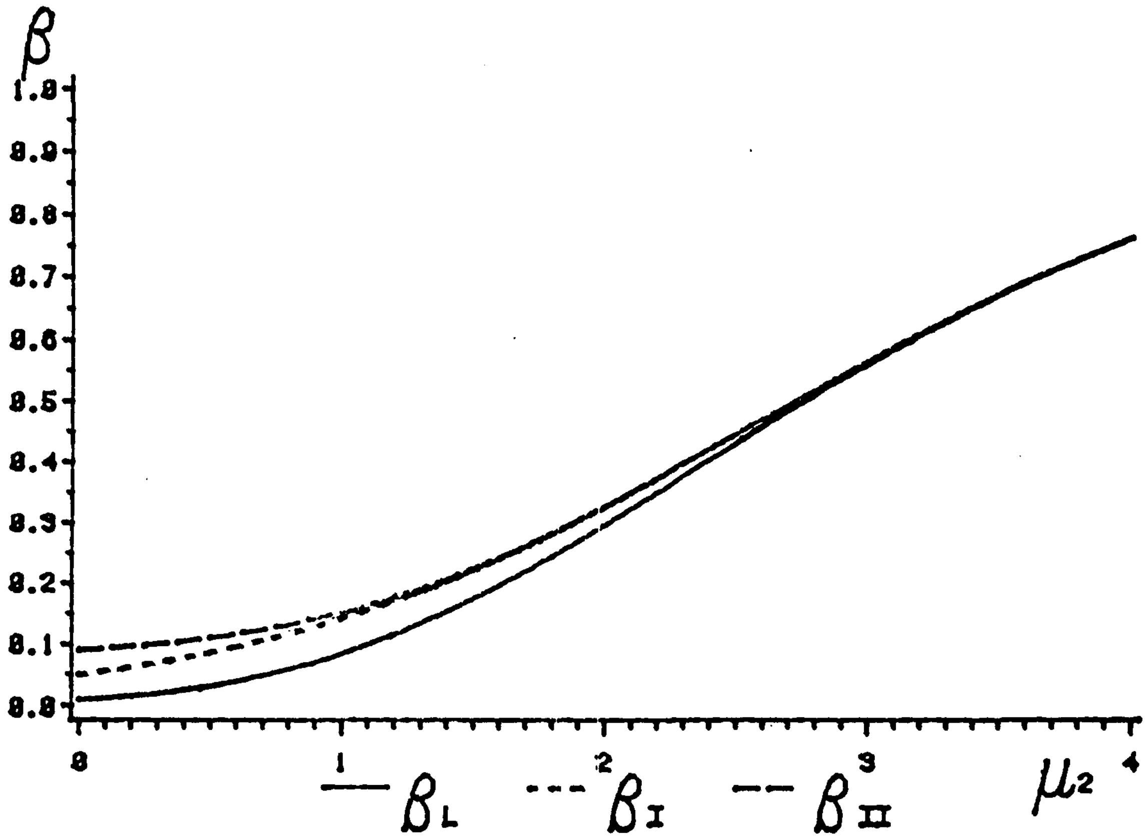

Figure 3.2c

CONTOUR PLOT OF RATIO = β_I/β_L

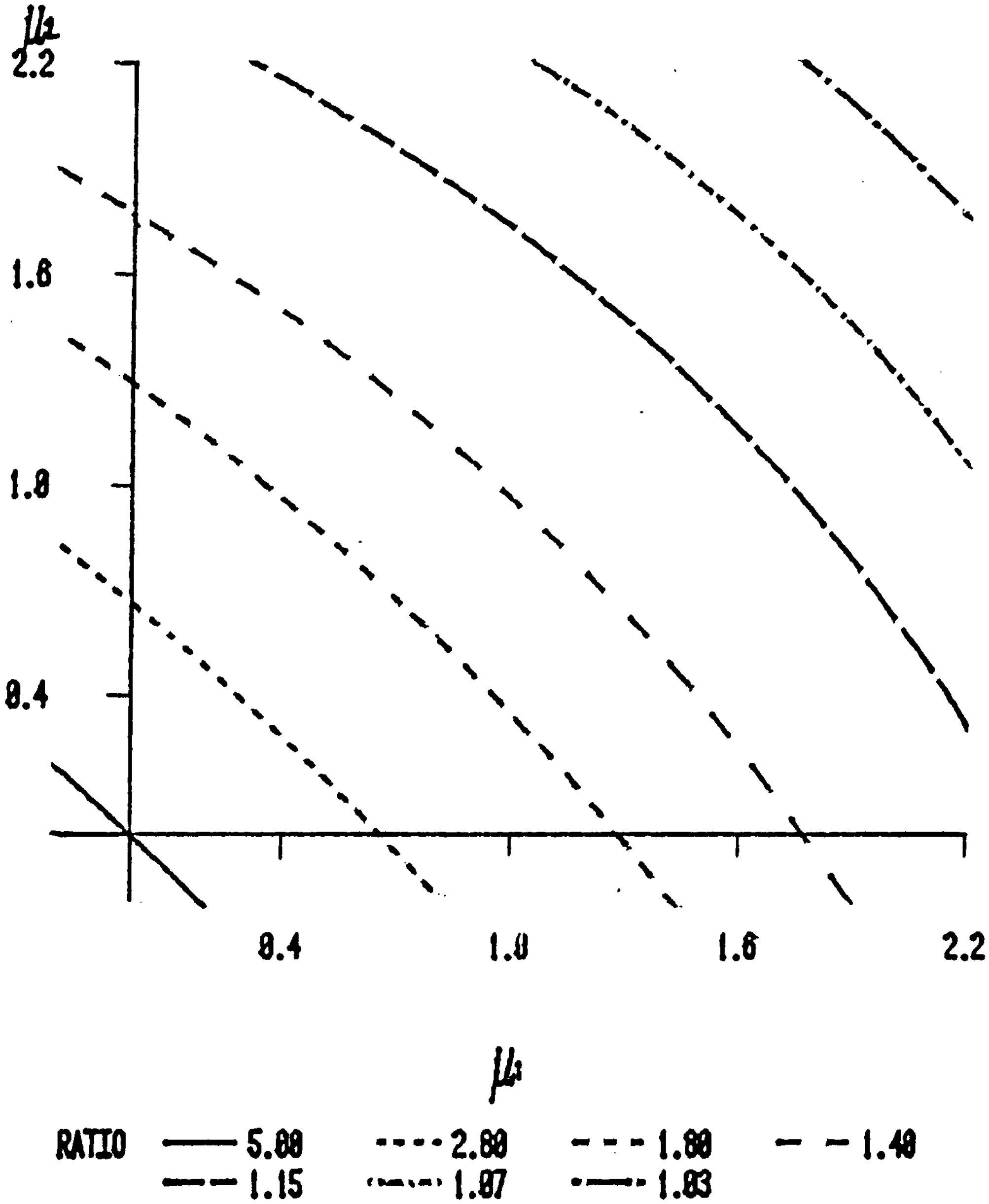

Figure 3.3

only approximating the $\beta_I(\underset{\sim}{\mu})/\beta_L(\underset{\sim}{\mu}) = 1.4$ contour with the line $\mu_1 + \mu_2 = 1.8$ in this statement.)

The power comparison mentioned after (1.3) can be easily made. Let $X_1,\ldots,X_p$ be independent and let $X_i \sim N_1(\mu_i,\sigma_i^2)$. Consider testing H_0: $\mu_i \leq 0$ for some $i = 1,\ldots,p$ versus H_1: $\mu_i > 0$ for all $i = 1,\ldots,p$. Then $Z_i = X_i/\sigma_i$ and the LRT rejects H_0 if $Z_i \geq z_\alpha$ for all $i = 1,\ldots,p$. If $\underset{\sim}{\mu} = \underset{\sim}{0}$, $Z_i \sim N_1(0,1)$ for all i and the Z_is are independent. Hence $\beta_L(\underset{\sim}{0}) = P(Z_1 \geq z_\alpha,\ldots,Z_p \geq z_\alpha) = \alpha^p$. For test I,

$$\beta_I(\underset{\sim}{0}) = \sum_{j=1}^{J} P_{\underset{\sim}{0}}(R_j) = \sum_{j=1}^{J} P_{\underset{\sim}{0}}(c_j \leq Z_i \leq c_{j-1},\ i = 1,\ldots,p) .$$

The first J-1 terms in the sum are all α^p since $c_j = z_{j\alpha}$, $j = 0,\ldots,J-1$. The last term may be less than α^p but, if $1/2\alpha$ is an integer, the last term is also α^p and $J = 1/2\alpha$. Thus $\beta_I(\underset{\sim}{0}) \approx J\alpha^p \approx \alpha^p/2\alpha = \alpha^{p-1}/2$. So, for this sign testing problem, $\beta_I(\underset{\sim}{0})/\beta_L(\underset{\sim}{0}) \approx (\alpha^{p-1}/2)/\alpha^p = 1/2\alpha$. $1/2(.1) = 5$ is the value of the contour at $\underset{\sim}{\mu} = 0$ in Figure 3.3.

4. AN EVEN MORE POWERFUL TEST

Test I from Section 3 is not necessarily the most powerful size-α test. In some cases there exists size-α tests that are uniformly more powerful than Test I. In this section we will give an example of such a test.

We shall call the more powerful test, *Test II.* Test II will reject H_0 if

$$\underset{\sim}{X} \in \bigcup_{j=1}^{M} R_j$$

where $J < M < 2J$. The rejection region for Test II consists of the rejection region for Test I plus some more of the sets R_j. Hence, Test II is obviously more powerful than Test I or the LRT. But the verification that Test II is a size-α test is more difficult. Theorem 2.2 cannot be used because the rejection region does not lie on one side of a plane defined by $\underset{\sim}{b}'\underset{\sim}{x} = 0$. In fact, it is not obvious that we will still have a size-α test if we add even one set, R_{J+1}, to the rejection region of Test I.

Test II may be primarily of theoretical interest because it has a rather counterintuitive property. For any $\underset{\sim}{x} \in R_j$, where $j > J$, $\underset{\sim}{b}_i'\underset{\sim}{x} \le 0$ for all $i = 1,\ldots k$. Thus, if we reject H_0 for such an $\underset{\sim}{x}$, we are deciding that $\underset{\sim}{b}_i'\underset{\sim}{\mu} > 0$ for all $i = 1,\ldots,k$ even though the estimate of $\underset{\sim}{\mu}$, $\underset{\sim}{x}$, satisfies $\underset{\sim}{b}_i'\underset{\sim}{x} \le 0$ for all $i = 1,\ldots,k$. Although we will demonstrate in an example that M can be chosen so that Test II is a size-α test that is uniformly more powerful than Test I, the important question might be this. Is there a size-α test with power comparable to Test II that only rejects for $\underset{\sim}{x}$ such that $\underset{\sim}{b}_i'\underset{\sim}{x} \ge 0$ for all $i = 1,\ldots,k$?

The example we will consider is again the bivariate sign testing problem. $\underset{\sim}{X} \sim N_2(\underset{\sim}{\mu},\underset{\sim}{I}_2)$ (for simplicity of notation we consider both variances equal to one). We are testing H_0: $\mu_1 \le 0$ or $\mu_2 \le 0$ versus H_1: $\mu_1 > 0$ and $\mu_2 > 0$.

For $\alpha = .10$, Test I has the rejection region

$$\underset{\sim}{X} \in \bigcup_{j=1}^{5} R_j,$$

where the R_j are depicted in Figure 3.1. We will show that Test II, with rejection region

$$\underset{\sim}{X} \in \bigcup_{j=1}^{9} R_j ,$$

is also size-α = .10. To compute the size of Test II we will use techniques involving majorization and Schur convexity. The reader is referred to Marshall and Olkin (1974) for all definitions regarding these concepts. Each of the sets, $R_j = \{(x_1,x_2): c_j \leq x_1,x_2 \leq c_{j-1}\}$ is a Schur convex set and any union of Schur convex sets is a Schur convex set. Thus the rejection region for Test II, for any M, is a Schur convex set. The density of $\underset{\sim}{X}$ is Schur concave. By Theorem 2.1 of Marshall and Olkin (1974), the power function of Test II,

$$\beta_{II}(\underset{\sim}{\mu}) = P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in \bigcup_{j=1}^{M} R_j) ,$$

is a Schur concave function. That is, if $\underset{\sim}{\mu}$ majorizes $\underset{\sim}{\mu}^*$, i.e., $\mu_1 + \mu_2 = \mu_1^* + \mu_2^*$ and $\max\{\mu_1,\mu_2\} \geq \max\{\mu_1^*,\mu_2^*\}$, then $\beta_{II}(\underset{\sim}{\mu}^*) \geq \beta_{II}(\underset{\sim}{\mu})$.

The size of Test II is $\sup_{\underset{\sim}{\mu} \in H_0} \beta_{II}(\underset{\sim}{\mu})$. We wish to determine the largest M, $J < M < 2J$, (if any exists) for which the size is α. Let $\underset{\sim}{\mu} \in H_0$ with $\mu_1 + \mu_2 \geq 0$. Since $\min\{\mu_1,\mu_2\} \leq 0$, $\underset{\sim}{\mu}$ majorizes $\underset{\sim}{\mu}^* = (\mu_1 + \mu_2, 0)'$. The rejection region of Test II, for any M, is subset of

$$\bigcup_{j=1}^{2J} R_j .$$

So by Theorem 2.1, $\beta_{II}(\underset{\sim}{\mu}) \leq \beta_{II}(\underset{\sim}{\mu}^*) \leq \alpha$. Now let $\underset{\sim}{\mu} \in H_0$ with $\mu_1 + \mu_2 \leq 0$. Then $\underset{\sim}{\mu}$ majorizes $\underset{\sim}{\mu}^* = (\bar{\mu},\bar{\mu})'$ where $\bar{\mu} = (\mu_1 + \mu_2)/2$. Thus $\beta_{II}(\underset{\sim}{\mu}) \leq \beta_{II}(\underset{\sim}{\mu}^*)$.

If we can show that $\beta_{II}(\bar{\mu},\bar{\mu}) \le \alpha$ for all $\bar{\mu} \le 0$, then we will have verified that Test II is a size-α test. Furthermore, we actually need only verify $\beta_{II}(\bar{\mu},\bar{\mu}) \le \alpha$ for $c_M \le \bar{\mu} \le 0$. The reason for this is that for every

$$\underset{\sim}{x} \in \bigcup_{j=1}^{M} R_j ,$$

$x_1 + x_2 \ge 2c_M$. Thus by translating the problem so that (c_M,c_M) is the origin, we can use Theorem 2.2 to show that $\beta_{II}(\bar{\mu},\bar{\mu}) \le \beta_{II}(c_M,c_M)$ for all $\bar{\mu} < c_M$. For α = .10, .05 and .01, we calculated $\beta_{II}(\bar{\mu},\bar{\mu})$ for $c_M \le \bar{\mu} \le 0$ on a grid with spacing .001 to find the maximum M for which $\beta_{II}(\bar{\mu},\bar{\mu}) \le \alpha$ for all such $\bar{\mu}$. The results are shown in Table 4.1. So Test II with M equal to the value in the table is a size-α test. In the table we also list the value of $\bar{\mu}$, $c_M \le \bar{\mu} \le 0$, at which $\beta_{II}(\bar{\mu},\bar{\mu})$ is maximized and the maximum value of $\beta_{II}(\bar{\mu},\bar{\mu})$. But note that the size of Test II is α, not the value listed as $\beta_{II}(\bar{\mu},\bar{\mu})$. The $\sup_{\underset{\sim}{\mu} \in H_0} \beta_{II}(\underset{\sim}{\mu})$ occurs, as with Test I and the LRT, in the limit of parameter points $(0,\mu)$ as $\mu \to \infty$.

The power function for Test II for α = .10 is the top graph in Figures 3.2a, b and c. For $\underset{\sim}{\mu}$ near $\underset{\sim}{0}$, $\beta_{II}(\underset{\sim}{\mu})$ is 1.8 times bigger than $\beta_{I}(\underset{\sim}{\mu})$ and 9 times bigger than $\beta_{L}(\underset{\sim}{\mu})$. In Figure 3.2a, one can see that Test II is much more nearly an unbiased test than either of the other two. But, despite these superior power properties, Test II is probably only of theoretical interest, for the reasons mentioned earlier.

TABLE 4.1 VALUE OF M THAT GIVES SIZE α FOR TEST II

α	M	$\bar{\mu}$ at which maximum occurs	$\beta_{II}(\bar{\mu},\bar{\mu})$
.10	9	.000	.01000
.05	19	-.884	.04906
.01	95	-.901	.00985

5. UNKNOWN VARIANCE EXAMPLE

The previous sections all dealt with models in which $\underset{\sim}{\Sigma}$, the covariance matrix, is known. Sasabuchi (1980) also considered two models in which $\underset{\sim}{\Sigma}$ was unknown. (See Sasabuchi (1987a and 1987b) for much more detail.) He considered the model in which $\underset{\sim}{\Sigma}$ was completely unknown and the model in which $\underset{\sim}{\Sigma} = \sigma^2\underset{\sim}{\Lambda}$, σ^2 unknown and $\underset{\sim}{\Lambda}$ known. He showed that the LRTs for these models were very similar to the LRT in the known $\underset{\sim}{\Sigma}$ case. Namely the test statistics Z_i were the same, except $\underset{\sim}{\Sigma}$ was replaced by an estimate, and the critical value z_α was replaced with a t-distribution percentile, t_α.

Because of the similarities it is natural to ask whether making the same changes in Test I will yield a test that is size-α and uniformly more powerful than the LRT. The answer, unfortunately, is that, in general, this method of test construction does not yield a size-α test. We describe an example that illustrates this fact. But the example also illustrates that even for moderate sample sizes, the test has size close to α and may be preferable to the LRT because of its higher power.

Consider again the bivariate sign testing problem. We are testing H_0: $\mu_1 \leq 0$ or $\mu_2 \leq 0$ versus H_1: $\mu_1 > 0$ and $\mu_2 > 0$. Suppose X_1 and X_2 are independent with $X_i \sim N_1(\mu_i, \sigma^2)$. Let S^2 an independent estimate of σ^2 such that $\nu S^2/\sigma^2$ has a chi-squared distribution with ν degrees of freedom (df). Typically S^2 will be a pooled estimate of σ^2 based on samples from the two populations that gave rise to X_1 and X_2. The LRT rejects H_0 if $X_1/S > t_\alpha$ and $X_2/S > t_\alpha$ where t_α is the upper 100α percentile of a t-distribution with ν df. Define $c_0, \ldots, c_J$ as in Definition 2.1 except with $t_{j\alpha}$, t-distribution percentiles, replacing the normal percentiles, $z_{j\alpha}$. Letting $R_j^* = \{(x_1, x_2, s)\colon\ c_j \leq x_1/s \leq c_{j-1},\ c_j \leq x_2/s \leq c_{j-1}\}$, we consider the analogue of Test I, the test that rejects H_0 if

$$(X_1, X_2, S) \in \bigcup_{j=1}^{J} R_j^* \ .$$

If we let $h_\sigma(s)$ be the density of S, the power function of this test can be expressed as

$$(5.1) \qquad \beta_I(\mu,\sigma) = \int_0^\infty \sum_{j=1}^{J} P_{\mu,\sigma}(c_j s \le X_1 \le c_{j-1}s,\ c_j s \le X_2 \le c_{j-1}s) h_\sigma(s) ds \ .$$

Using Theorem 2.2 on the integrand in (5.1), it can easily be shown that

$$(5.2) \qquad \sup_{(\mu_1,\mu_2,\sigma)\varepsilon H_0} \beta_I(\mu,\sigma) = \sup_{0\le\mu<\infty} \beta_I((\mu,0),1) \ .$$

We calculated $\beta_I((\mu,0),1)$ for values of μ between 0 and 20 by increments of .1 using numeric integration. We did the calculations for α = .10 and .05 and various df. In every case we found that $\beta_I((\mu,0),1)$ increased to a maximum that was greater than α and then decreased to α. The results of these calculations are summarized in Table 5.1. The table gives $\beta((0,0),1)$, the value of μ that gives the largest value of $\beta((\mu,0),1)$ called μ_{max}, $\beta_I((\mu_{max},0),1)$ which by (5.2) is the size of the test, and $\beta_I((14,0),1)$. The calculations always gave the same value of $\beta_I((\mu,0),1)$ for all $\mu \ge 14$. It is clear that $\lim_{\mu\to\infty} \beta_I((\mu,0),1) = \alpha$. So the nearness of $\beta_I((14,0),1)$ to α gives an indication of the accuracy of the computations. In all cases, $|\beta_I((14,0),1)-\alpha| \le .0002$.

It can be seen from Table 5.1 that in every case the size of the test, $\beta((\mu_{max},0),1)$, is greater than α. So in this example, this construction does not yield a size-α test. But the size of the test does approach α as the df becomes large. Even for df as small as ten (which could correspond to sample sizes of six from each population if S^2 is the pooled estimate of σ^2) the size of this test is close to α. So for moderate or large df, this test might be preferable to the LRT since its size is approximately α and it has higher power.

TABLE 5.1 SIZE OF TEST I IN UNKNOWN VARIANCE EXAMPLE

Degrees of Freedom	$\beta_I((0,0),1)$	μ_{max}	size of test = $\beta_I((\mu_{max},0),1)$	$\beta_I((14,0),1)$
		$\alpha = .10$		
2	.0681	1.7	.1235	.0999
4	.0591	1.8	.1102	.1000
6	.0560	2.0	.1059	.1000
8	.0545	2.1	.1039	.1000
10	.0536	2.2	.1028	.1000
12	.0530	2.2	.1021	.1000
16	.0522	2.4	.1012	.1000
20	.0518	2.5	.1009	.1001
30	.0512	2.7	.1005	.1001
50	.0509	2.9	.1003	.1002
120	.0503	3.4	.1000	.1000
∞	.0500	∞	.1000	.1000
		$\alpha = .05$		
2	.0371	1.7	.0702	.0500
4	.0309	1.8	.0600	.0500
6	.0288	1.9	.0564	.0500
8	.0278	2.0	.0546	.0500
10	.0272	2.1	.0535	.0500
12	.0269	2.1	.0528	.0500
16	.0264	2.2	.0519	.0500
20	.0261	2.3	.0514	.0500
30	.0257	2.4	.0508	.0500
50	.0255	2.6	.0505	.0501
120	.0252	3.0	.0501	.0500
∞	.0250	∞	.0500	.0500

6. A UNIFORMLY MORE POWERFUL TEST IN A TWO-SIDED PROBLEM

In this section we return to the known covariance model and consider a two-sided problem involving linear inequalities. A two-sided version of the testing problem (1.1) is obtained if the alternative hypothesis is $H_1 \cup (-H_1)$ where H_1 is the set defined in (1.1). That is, consider testing

$$H_0^2: \quad \underset{\sim}{b}_i'\underset{\sim}{\mu} \leq 0 \text{ for some } i = 1,\ldots,k \text{ and } \underset{\sim}{b}_i'\underset{\sim}{\mu} \geq 0 \text{ for some } i = 1,\ldots,k$$

(6.1) versus

$$H_1^2: \quad \underset{\sim}{b}_i'\underset{\sim}{\mu} > 0 \text{ for all } i = 1,\ldots,k \text{ or } \underset{\sim}{b}_i'\underset{\sim}{\mu} < 0 \text{ for all } i = 1,\ldots,k .$$

Sasabuchi (1980) showed that the LRT rejects H_0^2 if $Z_i = \underset{\sim}{b}_i'\underset{\sim}{X}/\sqrt{\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_i} \geq c$ for all $i = 1,\ldots,k$ or $Z_i \leq -c$ for all $i = 1,\ldots,k$. Sasabuchi showed that in some cases the constant c that yields a size-α test is $c = z_\alpha$. We will show for the sign testing problem that Test III, the test that rejects H_0^2 if

$$\underset{\sim}{X} \in \bigcup_{j=1}^{2J} R_j ,$$

is a size-α test that is uniformly more powerful than the LRT. (The sets R_j are still the sets in Definition 2.2.) Note that if $c = z_\alpha$, then $R_1 \cup R_{2J}$ is the rejection region for the LRT. Thus Test III is obviously a more powerful test than the LRT. The difficulty is in showing that it is a size-α test.

As mentioned in Section 1, Sasabuchi (1980) actually considered a different null hypothesis. His null hypothesis was that $\underset{\sim}{\mu}$ was on the boundary of H_1^2. His alternative was H_1^2. For Sasabuchi's problem, Theorem 2.1 provides two insights. It shows that $c = z_\alpha$ is the constant that yields a size-α LRT in a broader class of problems than found by Sasabuchi. It also shows that for this broader class, Test III is a size-α test that is uniformly more powerful than the LRT. Sasabuchi's Theorems 4.1, 4.2, and 4.3 imply that if $2 \leq k \leq p$

and $\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$ for all $i \neq m$, $i, m = 1,\ldots,k$, then $c = z_\alpha$. Theorem 2.1 does not require $k \leq p$ and only requires $\underset{\sim}{b}_i'\underset{\sim}{\Sigma}\underset{\sim}{b}_m \leq 0$ for one m for each i. Since any $\underset{\sim}{\mu}$ on the boundary of H_1^a satisfies $\underset{\sim}{b}_i'\underset{\sim}{\mu} = 0$ for some i, Theorem 2.1 implies that $c = z_\alpha$ yields a size-α LRT for Sasabuchi's null hypothesis and that Test III is also a size-α test, under the conditions stated in the theorem.

For the rest of this section, except the final paragraph where extensions will be discussed, we consider the sign testing problem. Let $X_1,\ldots,X_p$ be independent, $X_i \sim N_1(\mu_i,\sigma_i^2)$ and consider testing

$$H_0^S: \quad \mu_i \leq 0 \text{ for some } i = 1,\ldots,p \text{ and } \mu_i \geq 0 \text{ for some } i = 1,\ldots,p$$

(6.2) versus

$$H_1^S: \quad \mu_i > 0 \text{ for all } i = 1,\ldots,p \text{ or } \mu_i < 0 \text{ for all } i = 1,\ldots,p \quad .$$

The LRT rejects H_0^S if $X_i/\sigma_i \geq z_\alpha$ for all $i = 1,\ldots,p$ or $X_i/\sigma_i \leq -z_\alpha$ for all $i = 1,\ldots,p$. Test III rejects H_0^S if for some $j = 1,\ldots,2J$, $c_j \leq X_i/\sigma_i \leq c_{j-1}$ for all $i = 1,\ldots,p$.

To see that Test III is a size-α test in this problem, let $Z_i = X_i/\sigma_i$. Then $\underset{\sim}{Z} = (Z_1,\ldots,Z_p)' \sim N_p(\underset{\sim}{\theta},\underset{\sim}{I}_p)$ where $\theta_i = \mu_i/\sigma_i$. The rejection region for Test III is a Schur convex subset of the $\underset{\sim}{Z}$ space and the density of $\underset{\sim}{Z}$ is Schur concave. We will use Theorem 2.1 of Marshall and Olkin (1974) to show that Test III is a size-α test. Let $\underset{\sim}{\mu} \in H_0^S$. If $\mu_i \geq 0$ for all $i = 1,\ldots,p$ or $\mu_i \leq 0$ for all $i = 1,\ldots,p$, then $\mu_i = 0$ for some i (since $\underset{\sim}{\mu} \in H_0^S$). By Theorem 2.1,

$$P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \in \bigcup_{j=1}^{2J} R_j) \leq \alpha \ .$$

So suppose $\mu_i > 0$ for some i and $\mu_i < 0$ for some i. Suppose

$$\sum_{i=1}^{p} \theta_i \geq 0.$$

Without loss of generality we can assume $\theta_1 \geq \ldots \geq \theta_n \geq 0 > \theta_{n+1} \geq \ldots \geq \theta_p$ where $1 \leq n < p$. Let

$$a = \left(\sum_{i=1}^{p} \theta_i \Big/ \sum_{i=1}^{n} \theta_n \right)$$

and define $\underset{\sim}{\theta}^*$ by $\theta_i^* = a\theta_i$, $i = 1,\ldots,n$, $\theta_i^* = 0$, $i = n + 1,\ldots,p$ and $\mu_i^* = \sigma_i \theta_i^*$, $i = 1,\ldots,p$. Since $0 \leq a < 1$, $\underset{\sim}{\theta}$ majorizes $\underset{\sim}{\theta}^*$. So by Theorem 2.1 of Marshall and Olkin (1974)

$$\begin{aligned} P_{\underset{\sim}{\mu}}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{2J} R_j) &= P_{\underset{\sim}{\theta}}(\bigcup_{j=1}^{2J} \{c_j \leq Z_i \leq c_{j-1}, \text{ for all } i = 1,\ldots,p\}) \\ &\leq P_{\underset{\sim}{\theta}^*}(\bigcup_{j=1}^{2J} \{c_j \leq Z_i \leq c_{j-1}, \text{ for all } i = 1,\ldots,p\}) \\ &= P_{\underset{\sim}{\mu}^*}(\underset{\sim}{X} \varepsilon \bigcup_{j=1}^{2J} R_j) \\ &\leq \alpha . \qquad \text{(Theorem 2.1)} . \end{aligned}$$

The case of $\underset{\sim}{\mu} \varepsilon H_0^S$,

$$\sum_{i=1}^{p} \theta_i < 0 ,$$

can be similarly handled. Thus Test III is a size-α test for testing (6.2).

An application in which the two-sided hypotheses might be of interest is suggested by Gail and Simon (1985). As mentioned in the Abstract, let $\mu_i = \nu_{2i} - \nu_{1i}$ where ν_{ji} is the average response of the i^{th} patient subset ($i = 1,\ldots,p$) to the j^{th} treatment ($j = 1,2$). If $\mu_i > 0$ for all $i = 1,\ldots,p$, then Treatment 2 is better in all subsets. If $\mu_i < 0$ for all $i = 1,\ldots,p$, then Treatment 1 is better in all subsets. Thus H_1^S states that the same treatment is better for all subsets. In the terminology of Gail and Simon, there is *no qualitative interaction* between treatment effects and patient subsets. Gail and Simon had H_1^0: "no qualitative interaction" as the null hypothesis. So the LRT they study is different than the one we have considered and our results are not generally applicable in their problem. But in one case, p = 2 patient subsets, our Test III provides a uniformly more powerful size-α test in the Gail and Simon problem. To see this, let $\mu_1 = \nu_{21} - \nu_{11}$, as before, but let $\mu_2 = \nu_{12} - \nu_{22}$. Now, H_1^0: $\mu_1 > 0, \mu_2 > 0$ or $\mu_1 < 0, \mu_2 < 0$ states that there is a qualitative interaction, as in the Gail and Simon formulation. For this special case, Zelterman (1987) has constructed an approximate test that is uniformly more powerful than the Gail and Simon LRT and locally most powerful at $\underset{\sim}{\mu} = \underset{\sim}{0}$.

We have only shown that Test III is a size-α test for the special sign testing problem (6.2). For the more general problem (6.1), Theorem 2.1 would still be useful. To use majorization and Schur convexity as we have, one would need to show that

$$\bigcup_{j=1}^{2J} R_j$$

(possibly after some transformation) is a Schur convex set. If p = k = 2, this is always possible. For example, in Figure 2.1, the plane can be rotated so that the line $y_1 = y_2$ bisects the sets $S_1^*,\ldots,S_6^*$. But, if $p \geq 3$, this may not be possible. For example, it does not seem that

$$\bigcup_{j=1}^{2J} R_j$$

resulting from the two-sided simple order hypothesis,

$$H_1^0: \quad \mu_1 < \ldots < \mu_p \text{ or } \mu_1 > \ldots > \mu_p ,$$

can be rotated to be a Schur convex set if $p \geq 3$. Thus other techniques may be needed to find uniformly more powerful tests in the general two-sided problem.

REFERENCES

Berger, Roger L. (1982), "Multiparameter Hypothesis Testing and Acceptance Sampling," *Technometrics,* 24, 295-300.

Berger, Roger L. and Sinclair, Dennis (1984), "Testing Hypotheses Concerning Unions of Linear Subspaces," *Journal of the American Statistical Association,* 79, 158-163.

Cohen, Arthur, Gatsonis, Constantine, and Marden, John I. (1983), "Hypothesis Testing for Marginal Probabilities in a 2 × 2 × 2 Contingency Table with Conditional Independence," *Journal of the American Statistical Association,* 78, 920-929.

Cohen, Arthur and Marden, John (1983), "Hypothesis Tests and Optimality Properties in Discrete Multivariate Analysis," *Studies in Econometrics, Time Series, and Multivariate Statistics,* Academic Press. 379-405

Gail, M. and Simon, R. (1985), "Testing for Qualitative Interactions between Treatment Effects and Patient Subsets," *Biometrics,* 41, 361-372.

Gleser, Leon J. (1973), "On a Theory of Intersection-Union Tests," *Institute of Mathematical Statistics Bulletin* (Abstract), 2, 233.

Lehmann, E. L. (1952), "Testing Multiparameter Hypotheses," *Annals of Mathematical Statistics,* 23, 541-552.

Marshall, Albert W. and Olkin, Ingram (1973), "Majorization in Multivariate Distributions," *Annals of Statistics,* 2, 1189-1200.

Nomakuchi, Kentaro and Sakata, Toshio (1987), "A Note on Testing Two Dimensional Normal Mean," to appear in *Annals of Statistical Mathematics.*

Robertson, Tim and Wegman, Edward J. (1978), "Likelihood Ratio Tests for Order Restrictions in Exponential Families," *Annals of Statistics,* 6, 485-505.

Sasabuchi, S. (1980), "A Test of a Multivariate Normal Mean with Composite Hypotheses Determined by Linear Inequalities," *Biometrika,* 67, 429-439.

Sasabuchi, Syoichi (1987a), "A Multivariate Test with Composite Hypotheses Determined by Linear Inequalities in the Case when the Covariance Matrix has an Unknown Scale Factor," to appear in *Memoirs of Faculty of Science, Kyushu University, Series A (Mathematics).*

Sasabuchi, Syoichi (1987b), "A Multivariate One-Sided Test with Composite Hypotheses in the Case when the Covariance Matrix Is Completely Unknown," to appear in *Memoirs of Faculty of Science, Kyushu University, Series A (Mathematics).*

Zelterman, Daniel (1987), "Test for Qualitative Interactions," Department of Mathematics and Statistics Technical Report, State University of New York, Albany.