This is a reprint from

Advances in Statistical Decision Theory and Applications

edited by

S. Panchapakesan and N. Balakrishnan
Birkhäuser, Boston
1997

# 15

# *Likelihood Ratio Tests and Intersection-Union Tests*

**Roger L. Berger**

*North Carolina State University, Raleigh, NC*

**Abstract.** The likelihood ratio test (LRT) method is a commonly used method of hypothesis test construction. The intersection-union test (IUT) method is a less commonly used method. We will explore some relationships between these two methods. We show that, under some conditions, both methods yield the same test. But, we also describe conditions under which the size-$\alpha$ IUT is uniformly more powerful than the size-$\alpha$ LRT. We illustrate these relationships by considering the problem of testing $\mathrm{H}_0 : \min\{|\mu_1|, |\mu_2|\} = 0$ versus $\mathrm{H}_a : \min\{|\mu_1|, |\mu_2|\} > 0$, where $\mu_1$ and $\mu_2$ are means of two normal populations.

*Keywords and phrases*: likelihood ratio test, intersection-union test, size, power, normal mean, sample size.

## 15.1 Introduction and Notation

The likelihood ratio test (LRT) method is probably the most commonly used method of hypothesis test construction. Another method, which is appropriate when the null hypothesis is expressed as a union of sets, is the intersection-union test (IUT) method. We will explore some relationships between tests that result from these two methods. We will give conditions under which both methods yield the same test. But, we will also give conditions under which the size-$\alpha$ IUT is uniformly more powerful than the size-$\alpha$ LRT.

Let $\boldsymbol{X}$ denote the random vector of data values. Suppose the probability distribution of $\boldsymbol{X}$ depends on an unknown parameter $\theta$. The set of possible values for $\theta$ will be denoted by $\Theta$. $L(\theta|\boldsymbol{x})$ will denote the likelihood function for the observed value $\boldsymbol{X} = \boldsymbol{x}$. We will consider the problem of testing the null hypothesis $\mathrm{H}_0 : \theta \in \Theta_0$ versus the alternative hypothesis $\mathrm{H}_a : \theta \in \Theta_0^c$, where $\Theta_0$ is a specified subset of $\Theta$ and $\Theta_0^c$ is its complement.

The likelihood ratio test statistic for this problem is defined to be

$$\lambda(\boldsymbol{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta|\boldsymbol{x})}{\sup_{\theta \in \Theta} L(\theta|\boldsymbol{x})}.$$

A LRT rejects $\mathrm{H}_0$ for small values of $\lambda(\boldsymbol{x})$. That is, the rejection region of a LRT is a set of the form $\{\boldsymbol{x} : \lambda(\boldsymbol{x}) < c\}$, where $c$ is a chosen constant. Typically, $c$ is chosen so that the test is a size-$\alpha$ test. That is, $c = c_\alpha$ is chosen to satisfy

$$\sup_{\theta \in \Theta_0} P_\theta(\lambda(\boldsymbol{X}) < c_\alpha) = \alpha, \tag{15.1}$$

where $\alpha$ is the Type-I error probability chosen by the experimenter.

We will consider problems in which the null hypothesis set is conveniently expressed as a union of $k$ other sets, i.e., $\Theta_0 = \cup_{i=1}^k \Theta_i$. (We will consider only finite unions, although arbitrary unions can also be considered.) Then the hypotheses to be tested can be stated as

$$\mathrm{H}_0 : \theta \in \bigcup_{i=1}^k \Theta_i \quad \text{versus} \quad \mathrm{H}_a : \theta \in \bigcap_{i=1}^k \Theta_i^c. \tag{15.2}$$

The IUT method is a natural method for constructing a hypothesis test for this kind of problem. Let $R_i, i = 1, \ldots, k$ denote a rejection region for a test of $\mathrm{H}_{i0} : \theta \in \Theta_i$ versus $\mathrm{H}_{ia} : \theta \in \Theta_i^c$. Then the IUT of $\mathrm{H}_0$ versus $\mathrm{H}_a$, based on $R_1, \ldots, R_k$, is the test with rejection region $R = \cap_{i=1}^k R_i$. The rationale behind an IUT is simple. The overall null hypothesis, $\mathrm{H}_0 : \theta \in \cup_{i=1}^k \Theta_i$, can be rejected only if each of the individual hypotheses, $\mathrm{H}_{i0} : \theta \in \Theta_i$, can be rejected.

An IUT was described as early as 1952 by Lehmann. Gleser (1973) coined the term IUT. Berger (1982) proposed IUTs for acceptance sampling problems, and Cohen, Gatsonis and Marden (1983a) proposed IUTs for some contingency table problems. Since then many authors have proposed IUTs for a variety of problems. The IUT method is the reverse of Roy's (1953) well-known union-intersection method, which is useful when the null hypothesis is expressed as an intersection.

Berger (1982) proved the following two theorems about IUTs.

**Theorem 15.1.1** *If $R_i$ is a level-$\alpha$ test of* $\mathrm{H}_{0i}$, *for $i = 1, \ldots, k$, then the* IUT *with rejection region $R = \bigcap_{i=1}^k R_i$ is a level-$\alpha$ test of* $\mathrm{H}_0$ *versus* $\mathrm{H}_a$ *in* (15.2).

An important feature in Theorem 15.1.1 is that each of the individual tests is performed at level-$\alpha$. But the overall test also has the same level $\alpha$. There is no need for an adjustment, e.g., Bonferroni, for performing multiple tests. The reason there is no need for such a correction is the special way the individual tests are combined. $H_0$ is rejected only if every one of the individual hypotheses, $H_{0i}$, is rejected.

Theorem 15.1.1 asserts that the IUT is level-$\alpha$. That is, its size is at most $\alpha$. In fact, a test constructed by the IUT method can be quite conservative. Its size can be much less that the specified value $\alpha$. But, Theorem 15.1.2 (a generalization of Theorem 2 in Berger (1982)) provides conditions under which the IUT is not conservative; its size is exactly equal to the specified $\alpha$.

**Theorem 15.1.2** *For some $i = 1, \ldots, k$, suppose $R_i$ is a size-$\alpha$ rejection region for testing $H_{0i}$ versus $H_{ai}$. For every $j = 1, \ldots, k,\ j \neq i$, suppose $R_j$ is a level-$\alpha$ rejection region for testing $H_{0j}$ versus $H_{aj}$. Suppose there exists a sequence of parameter points $\theta_l, l = 1, 2, \ldots,$ in $\Theta_i$ such that*

$$\lim_{l\to\infty} P_{\theta_l}(\boldsymbol{X} \in R_i) = \alpha,$$

*and, for every $j = 1, \ldots, k,\ j \neq i$,*

$$\lim_{l\to\infty} P_{\theta_l}(\boldsymbol{X} \in R_j) = 1.$$

*Then the IUT with rejection region $R = \bigcap_{i=1}^{k} R_i$ is a size-$\alpha$ test of $H_0$ versus $H_a$.*

Note that in Theorem 15.1.2, the one test defined by $R_i$ has size exactly $\alpha$. The other tests defined by $R_j,\ j = 1, \ldots, k,\ j \neq i$, are level-$\alpha$ tests. That is, their sizes may be less than $\alpha$. The conclusion is the IUT has size $\alpha$. Thus, if rejection regions $R_1, \ldots, R_k$ with sizes $\alpha_1, \ldots, \alpha_k$, respectively, are combined in an IUT and Theorem 15.1.2 is applicable, then the IUT will have size equal to $\max_i\{\alpha_i\}$.

## 15.2 Relationships Between LRTs and IUTs

For a hypothesis testing problem of the form (15.2), the LRT statistic can be written as

$$\lambda(\boldsymbol{x}) = \frac{\sup_{\theta\in\Theta_0} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})} = \frac{\max_{1\le i\le k} \sup_{\theta\in\Theta_i} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})} = \max_{1\le i\le k} \frac{\sup_{\theta\in\Theta_i} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})}.$$

But,

$$\lambda_i(\boldsymbol{x}) = \frac{\sup_{\theta\in\Theta_i} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})}$$

is the LRT statistic for testing $H_{i0} : \theta \in \Theta_i$ versus $H_{ia} : \theta \in \Theta_i^c$. Thus, the LRT statistic for testing $H_0$ versus $H_a$ is

$$\lambda(\boldsymbol{x}) = \max_{1\le i\le k} \lambda_i(\boldsymbol{x}). \tag{15.3}$$

The LRT of $H_0$ is a combination of tests for the individual hypotheses, $H_{10}, \ldots, H_{k0}$. In the LRT, the individual LRT statistics are first combined via (15.3). Then, the critical value, $c_\alpha$ that yields a size-$\alpha$ test is determined by (15.1).

Another way to combine the individual LRTs is to use the IUT method. For each $i = 1, \ldots, k$, the critical value that defines a size-$\alpha$ LRT of $H_{i0}$ is the value $c_{i\alpha}$ that satisfies

$$\sup_{\theta \in \Theta_{i0}} P_\theta(\lambda_i(\boldsymbol{X}) < c_{i\alpha}) = \alpha. \tag{15.4}$$

Then, $R_i = \{\boldsymbol{x} : \lambda_i(\boldsymbol{x}) < c_{i\alpha}\}$ is the rejection region of the size-$\alpha$ LRT of $H_{i0}$, and, by Theorem 15.1.1, $R = \cap_{i=1}^k R_i$ is the rejection region of a level-$\alpha$ test of $H_0$. If the conditions of Theorem 15.1.2 are satisfied, this IUT has size-$\alpha$.

In general, the two methods of combining $\lambda_1(\boldsymbol{x}), \ldots, \lambda_k(\boldsymbol{x})$ need not yield the same test. But, the following theorem gives a common situation in which the two methods do yield the same test. Theorems 15.2.1 and 15.2.2 are similar to Theorems 5 an 6 in Davis (1989).

**Theorem 15.2.1** *If the constants $c_{1\alpha}, \ldots, c_{k\alpha}$ defined in* (15.4) *are all equal and the conditions of Theorem 15.1.2 are satisfied, then the size-$\alpha$* LRT *of* $H_0$ *is the same as the* IUT *formed from the individual size-$\alpha$* LRT*s of* $H_{10}, \ldots, H_{k0}$.

**Proof:** Let $c = c_{1\alpha} = \cdots = c_{k\alpha}$. The rejection region of the IUT is given by

$$\begin{aligned} R &= \bigcap_{i=1}^k \{\boldsymbol{x} : \lambda_i(\boldsymbol{x}) < c_{i\alpha}\} = \bigcap_{i=1}^k \{\boldsymbol{x} : \lambda_i(\boldsymbol{x}) < c\} \\ &= \{\boldsymbol{x} : \max_{1 \le i \le k} \lambda_i(\boldsymbol{x}) < c\} = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) < c\}. \end{aligned}$$

Therefore, $R$ has the form of an LRT rejection region. Because each of the individual LRTs has size-$\alpha$ and the conditions of Theorem 15.1.2 are satisfied, $R$ is the size-$\alpha$ LRT. $\square$

Theorem 15.2.1 is particularly useful in situations in which the individual LRT statistics (or a transformation of them) have simple known distributions. In this case, the determination of the critical values, $c_{1\alpha}, \ldots, c_{k\alpha}$, is easy. But the distribution of $\lambda(\boldsymbol{X}) = \max_{1 \le i \le k} \lambda_i(\boldsymbol{X})$ may be difficult, and the determination of its critical value, $c_\alpha$, from (15.1) may be difficult. Examples of this kind of analysis may be found in Sasabuchi (1980), Sasabuchi (1988a), and Sasabuchi (1988b). In these papers about normal mean vectors, the alternative hypothesis is a polyhedral cone. The individual LRTs are expressed in terms of $t$-tests, each one representing the LRT corresponding to one face of the cone. All of the $t$-tests are based on the same degrees of freedom, so all the critical values are equal. Assumptions are made that ensure that the conditions of Theorem 15.1.2 are satisfied, and, in this way, the LRT is expressed as an intersection of $t$-tests. Sasabuchi does not use the IUT terminology, but it is clear that this is the argument that is used.

Theorem 15.2.1 gives conditions under which, if $c_{1\alpha} = \cdots = c_{k\alpha}$, the size-$\alpha$ LRT and size-$\alpha$ IUT are the same test. But, if the $c_{i\alpha}$s are not all equal, these two tests are not the same, and, often, the IUT is the uniformly more powerful test. Theorem 15.2.2 gives conditions under which this is true.

**Theorem 15.2.2** *Let $c_{1\alpha}, \ldots, c_{k\alpha}$ denote the critical values defined in* (15.4). *Suppose that for some $i$ with $c_{i\alpha} = \min_{1\le j\le k}\{c_{j\alpha}\}$, there exists a sequence of parameter points $\theta_l, l = 1, 2, \ldots,$ in $\Theta_i$ such that the following two conditions are true:*

(i) $\lim_{l\to\infty} P_{\theta_l}(\lambda_i(\boldsymbol{X}) < c_{i\alpha}) = \alpha$,

(ii) *For any $j \neq i$,* $\lim_{l\to\infty} P_{\theta_l}(\lambda_j(\boldsymbol{X}) < c_{i\alpha}) = 1$.

*Then, the following are true:*

(a) *The critical value for the size-$\alpha$* LRT *is $c_\alpha = c_{i\alpha}$.*

(b) *The* IUT *with rejection region $R = \cap_{j=1}^{k}\{\boldsymbol{x} : \lambda_j(\boldsymbol{x}) < c_{j\alpha}\}$ is a size-$\alpha$ test.*

(c) *The* IUT *in* (b) *is uniformly more powerful than the size-$\alpha$* LRT.

**Proof:** To prove (a), recall that the LRT rejection region using critical value $c_{i\alpha}$ is

$$\{\boldsymbol{x} : \lambda(\boldsymbol{x}) < c_{i\alpha}\} = \bigcap_{j=1}^{k}\{\boldsymbol{x} : \lambda_j(\boldsymbol{x}) < c_{i\alpha}\}. \tag{15.5}$$

For each $j = 1, \ldots, k$, because $c_{i\alpha} = \min_{1\le j\le k}\{c_{j\alpha}\}$ and $\{\boldsymbol{x} : \lambda_j(\boldsymbol{x}) < c_{j\alpha}\}$ is a size-$\alpha$ rejection region for testing $\mathrm{H}_{j0}$ versus $\mathrm{H}_{ja}$, $\{\boldsymbol{x} : \lambda_j(\boldsymbol{x}) < c_{i\alpha}\}$ is a level-$\alpha$ rejection region for testing $\mathrm{H}_{j0}$ versus $\mathrm{H}_{ja}$. Thus, by Theorem 15.1.1, the LRT rejection region in (15.5) is level-$\alpha$. But, in fact, this LRT rejection region is size-$\alpha$ because

$$\begin{aligned}
\sup_{\theta\in\Theta_0} P_\theta(\lambda(\boldsymbol{X}) < c_{i\alpha}) &\geq \lim_{l\to\infty} P_{\theta_l}(\lambda(\boldsymbol{X}) < c_{i\alpha}) \\
&= \lim_{l\to\infty} P_{\theta_l}\left(\bigcap_{j=1}^{k}\{\lambda_j(\boldsymbol{X}) < c_{i\alpha}\}\right) \\
&= 1 - \lim_{l\to\infty} P_{\theta_l}\left(\bigcup_{j=1}^{k}\{\lambda_j(\boldsymbol{X}) < c_{i\alpha}\}^c\right) \\
&\geq 1 - \lim_{l\to\infty}\sum_{j=1}^{k} P_{\theta_l}\left(\{\lambda_j(\boldsymbol{X}) < c_{i\alpha}\}^c\right) \\
&= 1 - (1-\alpha) = \alpha.
\end{aligned}$$

The last inequality follows from (i) and (ii).

For each $j = 1, \ldots, k$, $\{\boldsymbol{x} : \lambda_j(\boldsymbol{x}) < c_{j\alpha}\}$ is a level-$\alpha$ rejection region for testing $\mathrm{H}_{j0}$ versus $\mathrm{H}_{ja}$. Thus, Theorem 15.1.2, (i), and (ii) allow us to conclude part (b) is true.

Because $c_{i\alpha} = \min_{1 \le j \le k}\{c_{j\alpha}\}$, for any $\theta \in \Theta$,

$$P_\theta(\lambda(\boldsymbol{X}) < c_{i\alpha}) = P_\theta\left(\bigcap_{j=1}^{k}\{\lambda_j(\boldsymbol{X}) < c_{i\alpha}\}\right) \le P_\theta\left(\bigcap_{j=1}^{k}\{\lambda_j(\boldsymbol{X}) < c_{j\alpha}\}\right). \tag{15.6}$$

The first probability in (15.6) is the power of the size-$\alpha$ LRT, and the last probability in (15.6) is the power of the IUT. Thus, the IUT is uniformly more powerful. □

In part (c) of Theorem 15.2.2, all that is proved is that the power of the IUT is no less than the power of the LRT. However, if all the $c_{j\alpha}$s are not equal, the rejection region of the LRT is a proper subset of the rejection region of the IUT, and, typically, the IUT is strictly more powerful than the LRT. An example in which the critical values are unequal and the IUT is more powerful than the LRT is discussed in Berger and Sinclair (1984). They consider the problem of testing a null hypothesis that is the union of linear subspaces in a linear model. If the dimensions of the subspaces are unequal, then the critical values from an $F$-distribution have different degrees of freedom and are unequal.

## 15.3 Testing $\mathrm{H}_0 : \min\{|\mu_1|, |\mu_2|\} = 0$

In this section, we consider an example that illustrates the previous results. We find that the size-$\alpha$ IUT is uniformly more powerful than the size-$\alpha$ LRT. We then describe a different IUT that is much more powerful than both of the preceding tests. This kind of improved power, that can be obtained by judicious use of the IUT method, has been described for other problems by Berger (1989) and Liu and Berger (1995). Saikali (1996) found tests more powerful than the LRT for a one-sided version of the problem we consider in this section.

Let $X_{11}, \ldots, X_{1n_1}$ denote a random sample from a normal population with mean $\mu_1$ and variance $\sigma_1^2$. Let $X_{21}, \ldots, X_{2n_2}$ denote an independent random sample from a normal population with mean $\mu_2$ and variance $\sigma_2^2$. All four parameters, $\mu_1$, $\mu_2$, $\sigma_1^2$, and $\sigma_2^2$, are unknown. We will consider the problem of testing the hypotheses

$$\mathrm{H}_0 : \mu_1 = 0 \text{ or } \mu_2 = 0 \quad \text{versus} \quad \mathrm{H}_a : \mu_1 \ne 0 \text{ and } \mu_2 \ne 0. \tag{15.7}$$

Another way to express these hypotheses is

$$\mathrm{H}_0 : \min\{|\mu_1|, |\mu_2|\} = 0 \quad \text{versus} \quad \mathrm{H}_a : \min\{|\mu_1|, |\mu_2|\} > 0.$$

The parameters $\mu_1$ and $\mu_2$ could represent the effects of two different treatments. Then, $\mathrm{H}_0$ states that at least one treatment has no effect, and $\mathrm{H}_a$ states that both treatments have an effect.

Cohen, Gatsonis and Marden (1983b) considered tests of (15.7) in the variance known case. They proved an optimality property of the LRT in a class of monotone, symmetric tests.

### 15.3.1 Comparison of LRT and IUT

Standard computations yield that, for $i = 1$ and 2, the LRT statistic for testing $\mathrm{H}_{i0} : \mu_i = 0$ is

$$\lambda_i(x_{11}, \ldots, x_{1n_1}, x_{21}, \ldots, x_{2n_2}) = \left(1 + \frac{t_i^2}{n_i - 1}\right)^{-n_i/2},$$

where $\overline{x}_i$ and $s_i^2$ are the sample mean and variance from the $i$th sample and

$$t_i = \frac{\overline{x}_i}{s_i/\sqrt{n_i}} \tag{15.8}$$

is the usual $t$-statistic for testing $\mathrm{H}_{i0}$. Note that $\lambda_i$ is computed from both samples. But, because the likelihood factors into two parts, one depending only on $\mu_1$, $\sigma_1^2$, $\overline{x}_1$ and $s_1^2$ and the other depending only on $\mu_2$, $\sigma_2^2$, $\overline{x}_2$ and $s_2^2$, the part of the likelihood for the sample not associated with the mean in $\mathrm{H}_{i0}$ drops out of the LRT statistic.

Under $\mathrm{H}_{i0}$, $t_i$ has a Student's $t$ distribution. Therefore, the critical value that yields a size-$\alpha$ LRT of $\mathrm{H}_{i0}$ is

$$c_{i\alpha} = \left(1 + \frac{t^2_{\alpha/2, n_i - 1}}{n_i - 1}\right)^{-n_i/2},$$

where $t_{\alpha/2,n_i-1}$ is the upper $100\alpha/2$ percentile of a $t$ distribution with $n_i - 1$ degrees of freedom. The rejection region of the IUT is the set of sample points for which $\lambda_1(\boldsymbol{x}) < c_{1\alpha}$ and $\lambda_2(\boldsymbol{x}) < c_{2\alpha}$. This is more simply stated as reject $\mathrm{H}_0$ if and only if

$$|t_1| > t_{\alpha/2,n_1-1} \quad \text{and} \quad |t_2| > t_{\alpha/2,n_2-1}. \tag{15.9}$$

Theorem 15.1.2 can be used to verify that the IUT formed from these individual size-$\alpha$ LRTs is a size-$\alpha$ test of $\mathrm{H}_0$. To verify the conditions of Theorem 15.1.2, consider a sequence of parameter points with $\sigma_1^2$ and $\sigma_2^2$ fixed at any positive values, $\mu_1 = 0$, and let $\mu_2 \to \infty$. Then, $P(\lambda_1(\boldsymbol{x}) < c_{1\alpha}) = P(|t_1| > t_{\alpha/2,n_1-1}) = \alpha$, for any such parameter point. However, $P(\lambda_2(\boldsymbol{x}) < c_{2\alpha}) = P(|t_2| > t_{\alpha/2,n_2-1}) \to 1$ for such a sequence because the power of the $t$-test converges to 1 as the noncentrality parameter goes to infinity.

If $n_1 = n_2$, then $c_{1\alpha} = c_{2\alpha}$, and, by Theorem 15.2.1, this IUT formed from the individual LRTs is the LRT of $\mathrm{H}_0$.

If the sample sizes are unequal, the constants $c_{1\alpha}$ and $c_{2\alpha}$ will be unequal, and the IUT will not be the LRT. In this case, let $c = \min\{c_{1\alpha}, c_{2\alpha}\}$. By Theorem 15.2.2, $c$ is the critical value that defines a size-$\alpha$ LRT of $\mathrm{H}_0$. The same sequence as in the preceding paragraph can be used to verify the conditions of Theorem 15.2.2, if $c_{1\alpha} < c_{2\alpha}$. If $c_{1\alpha} > c_{2\alpha}$, a sequence with $\mu_2 = 0$ and $\mu_1 \to \infty$ can be used.

If $c = c_{1\alpha} < c_{2\alpha}$, then the LRT rejection region, $\lambda(\boldsymbol{x}) < c$, can be expressed as

$$|t_1| > t_{\alpha/2,n_1-1} \quad \text{and} \quad |t_2| > \left\{ \left[ \left( 1 + \frac{t^2_{\alpha/2,n_1-1}}{n_1 - 1} \right)^{n_1/n_2} - 1 \right] (n_2 - 1) \right\}^{1/2} . \tag{15.10}$$

The cutoff value for $|t_2|$ is larger than $t_{\alpha/2,n_2-1}$, because this rejection region is a subset of the IUT rejection region.

The critical values $c_{i\alpha}$ were computed for the three common choices of $\alpha =$ .10, .05, and .01, and for all sample sizes $n_i = 2, \ldots, 100$. On this range it was found that $c_{i\alpha}$ is increasing in $n_i$. So, at least on this range, $c = \min\{c_{1\alpha}, c_{2\alpha}\}$ is the critical value corresponding to the smaller sample size. This same property was observed by Saikali (1996).

### 15.3.2 More powerful test

In this section we describe a test that is uniformly more powerful than both the LRT and the IUT. This test is similar and may be unbiased. The description of this test is similar to tests described by Wang and McDermott (1996).

The more powerful test will be defined in terms of a set, $S$, a subset of the unit square. $S$ is the union of three sets, $S_1$, $S_2$, and $S_3$, where

$$\begin{aligned} S_1 &= \{(u_1,u_2) : 1-\alpha/2 < u_1 \le 1, 1-\alpha/2 < u_2 \le 1\} \\ &\quad \bigcup \{(u_1,u_2) : 0 \le u_1 < \alpha/2, 1-\alpha/2 < u_2 \le 1\} \\ &\quad \bigcup \{(u_1,u_2) : 1-\alpha/2 < u_1 \le 1, 0 \le u_2 < \alpha/2\} \\ &\quad \bigcup \{(u_1,u_2) : 0 \le u_1 < \alpha/2, 0 \le u_2 < \alpha/2\}, \end{aligned}$$

$$\begin{aligned} S_2 &= \{(u_1,u_2) : \alpha/2 \le u_1 \le 1-\alpha/2, \alpha/2 \le u_2 \le 1-\alpha/2\} \\ &\quad \bigcap \Big( \{(u_1,u_2) : u_1 - \alpha/4 \le u_2 \le u_1 + \alpha/4\} \\ &\qquad \bigcup \{(u_1,u_2) : 1 - u_1 - \alpha/4 \le u_2 \le 1 - u_1 + \alpha/4\} \Big), \end{aligned}$$

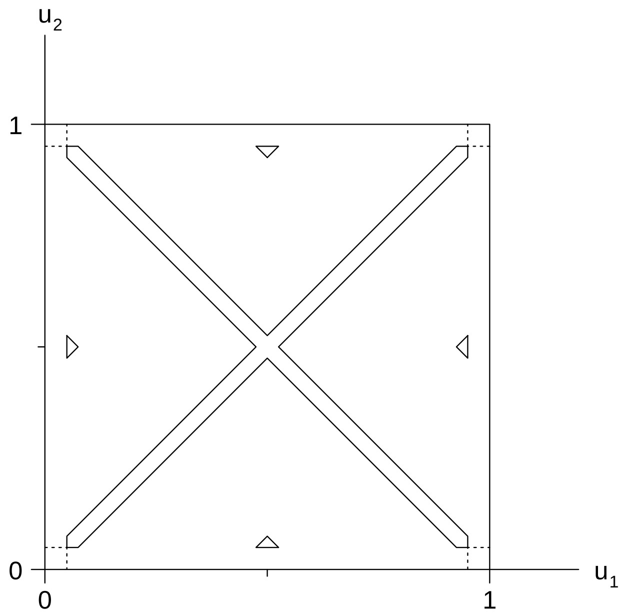


Figure 15.1: The set $S$ for $\alpha = .10$. Solid lines are in $S$, dotted lines are not.

and

$$
\begin{aligned}
S_3 \;=\; & \{(u_1,u_2) : \alpha/2 \le u_1 \le 1-\alpha/2, \alpha/2 \le u_2 \le 1-\alpha/2 \\
& \bigcap \Big( \{(u_1,u_2) : |u_1 - 1/2| + 1 - 3\alpha/4 \le u_2\} \\
& \quad \bigcup \{(u_1,u_2) : -|u_1 - 1/2| + 3\alpha/4 \ge u_2\} \\
& \quad \bigcup \{(u_1,u_2) : |u_2 - 1/2| + 1 - 3\alpha/4 \le u_1\} \\
& \quad \bigcup \{(u_1,u_2) : -|u_2 - 1/2| + 3\alpha/4 \ge u_1\} \Big) .
\end{aligned}
$$

The set $S$ for $\alpha = .10$ is shown in Figure 15.1. $S_1$ consists of the four squares in the corners. $S_2$ is the middle, X-shaped region. $S_3$ consists of the four small triangles.

The set $S$ has this property. Consider any horizontal or vertical line in the unit square. Then the total length of all the segments of this line that intersect with $S$ is $\alpha$. This property implies the following theorem.

**Theorem 15.3.1** *Let $U_1$ and $U_2$ be independent random variables. Suppose the supports of $U_1$ and $U_2$ are both contained in the interval $[0,1]$. If either $U_1$ or $U_2$ has a uniform$(0,1)$ distribution, then $P((U_1,U_2) \in S) = \alpha$.*

**Proof:** Suppose $U_1 \sim$ uniform$(0,1)$. Let $G_2$ denote the cumulative distribution function of $U_2$. Let $S(u_2) = \{u_1 : (u_1, u_2) \in S\}$, for each $0 \le u_2 \le 1$. Then,

$$P((U_1, U_2) \in S) = \int_0^1 \int_{S(u_2)} 1\, du_1\, dG_2(u_2) = \int_0^1 \alpha\, dG_2(u_2) = \alpha.$$

The second equality follows from the property of $S$ mentioned before the theorem.

If $U_2 \sim \text{uniform}(0,1)$, the result is proved similarly. □

Our new test, which we will call the $S$-test, of the hypotheses (15.7) can be described as follows. Let $F_i$, $i = 1, 2$, denote the cdf of a central $t$ distribution with $n_i - 1$ degrees of freedom. Let $U_i = F_i(t_i)$, $i = 1, 2$, where $t_i$ is the $t$ statistic defined in (15.8). Then the $S$-test rejects $\mathrm{H}_0$ if and only if $(U_1, U_2) \in S$.

$U_1$ and $U_2$ are independent because $t_1$ and $t_2$ are independent. If $\mu_1$ $(\mu_2) = 0$, then $F_1(t_1)$ $(F_2(t_2)) \sim \text{uniform}(0,1)$, and, by Theorem 15.3.1, $P((U_1, U_2) \in S) = \alpha$. That is, the $S$-test is a size-$\alpha$ test of $\mathrm{H}_0$. The event $(U_1, U_2) \in S_1$ is the same as the event in (15.9). So, the rejection region of the $S$-test contains the rejection region of the IUT from the the previous section, and the $S$-test is a size-$\alpha$ test that is uniformly more powerful than the size-$\alpha$ IUT.

We have seen that the IUT is uniformly more powerful than the LRT, and the $S$-test is uniformly more powerful than the IUT. Table 15.1 gives an example of the differences in power for these three tests. This example is for $n_1 = 5$, $n_2 = 30$ and $\alpha = .05$. The table gives the rejection probabilities for some parameter points of the form $(\mu_1, \mu_2) = (r\cos(\theta), r\sin(\theta))$, where $r = 0(.25)2$ and $\theta = 0(\pi/8)\pi/2$. These are equally spaced points on five lines emanating from the origin in the first quadrant. In Table 15.1, $\sigma_1^2 = \sigma_2^2 = 1$.

The $\theta = 0$ and $\theta = \pi/2$ entries in Table 15.1 are on the $\mu_1$ and $\mu_2$ axes, respectively. For the $S$-test, the rejection probability is equal to $\alpha$ for all such points. But, the other two tests are biased and their rejection probabilities are much smaller than $\alpha$ for $(\mu_1, \mu_2)$ close to $(0,0)$. For the IUT, the power converges to $\alpha$ as the parameter goes to infinity along either axis. For the LRT, this is also true along the $\mu_2$ axis. But, as is suggested by the table, for the LRT

$$\lim_{\mu \to \pm\infty} P(\text{reject } \mathrm{H}_0 | \mu_1 = \mu, \mu_2 = 0) = P(|T_{29}| > 2.384) = .024,$$

where $T_{29}$ has a central $t$ distribution with 29 degrees of freedom and 2.384 is the critical value for $t_2$ from (15.10). The power of the IUT along the $\mu_i$ axis is proportional to the power of a univariate, two-sided, size-$\alpha$ $t$-test of $\mathrm{H}_{0i} : \mu_i = 0$. Because the test of $\mathrm{H}_{01}$ is based on 4 degrees of freedom while the test of $\mathrm{H}_{02}$ is based on 29 degrees of freedom, the power increases more rapidly along the $\mu_2$ axis.

The sections of Table 15.1 for $\theta = \pi/8$, $\pi/4$ and $3\pi/8$ (except for $r = 0$) correspond to points in the alternative hypothesis. There it can be seen that the

| | $r$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | .00 | .25 | .50 | .75 | 1.00 | 1.25 | 1.50 | 1.75 | 2.00 |
| | | | | | $\theta = 0$ | | | | |
| $S$-test | .050 | .050 | .050 | .050 | .050 | .050 | .050 | .050 | .050 |
| IUT | .002 | .004 | .007 | .013 | .020 | .028 | .036 | .041 | .045 |
| LRT | .001 | .002 | .003 | .006 | .010 | .013 | .017 | .020 | .022 |
| | | | | | $\theta = \pi/8$ | | | | |
| $S$-test | .050 | .051 | .066 | .117 | .224 | .384 | .567 | .731 | .849 |
| IUT | .002 | .006 | .022 | .074 | .186 | .359 | .555 | .727 | .848 |
| LRT | .001 | .003 | .013 | .050 | .141 | .299 | .499 | .688 | .829 |
| | | | | | $\theta = \pi/4$ | | | | |
| $S$-test | .050 | .052 | .076 | .137 | .227 | .329 | .440 | .554 | .663 |
| IUT | .002 | .009 | .044 | .122 | .223 | .329 | .440 | .554 | .663 |
| LRT | .001 | .006 | .032 | .106 | .214 | .327 | .440 | .554 | .663 |
| | | | | | $\theta = 3\pi/8$ | | | | |
| $S$-test | .050 | .051 | .060 | .079 | .103 | .133 | .170 | .213 | .262 |
| IUT | .002 | .012 | .043 | .076 | .103 | .133 | .170 | .213 | .262 |
| LRT | .001 | .008 | .036 | .073 | .102 | .133 | .170 | .213 | .262 |
| | | | | | $\theta = \pi/2$ | | | | |
| $S$-test | .050 | .050 | .050 | .050 | .050 | .050 | .050 | .050 | .050 |
| IUT | .002 | .013 | .038 | .049 | .050 | .050 | .050 | .050 | .050 |
| LRT | .001 | .009 | .032 | .048 | .050 | .050 | .050 | .050 | .050 |

Table 15.1: Powers of $S$-test, IUT and LRT for $n_1 = 5$, $n_2 = 30$ and $\alpha = .05$. Power at parameters of form $(\mu_1, \mu_2) = (r\cos(\theta), r\sin(\theta))$ with $\sigma_1 = \sigma_2 = 1$.

$S$-test has much higher power than the other two tests, especially for parameters close to $(0, 0)$. The IUT, which is very intuitive and easy to describe, offers some power improvement over the LRT.

## 15.4 Conclusion

For a null hypothesis expressed as a union, as in (15.2), the IUT method is a simple, intuitive method of constructing a level-$\alpha$ test. We have described situations in which the IUT defined by size-$\alpha$ LRTs of the individual hypotheses is a uniformly more powerful test than the size-$\alpha$ LRT of the overall hypothesis. And, we have illustrated in an example how even more powerful tests might be found by careful consideration of the specific problem at hand.

**Acknowledgments.** I thank Yining Wang for helpful suggestions about Theorem 15.2.2.

---

## References

Berger, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling, *Technometrics*, **24**, 295–300.

Berger, R. L. (1989). Uniformly more powerful tests for hypotheses concerning linear inequalities and normal means, *Journal of the American Statistical Association*, **84**, 192–199.

Berger, R. L. and Sinclair, D. F. (1984). Testing hypotheses concerning unions of linear subspaces, *Journal of the American Statistical Association*, **79**, 158–163.

Cohen, A., Gatsonis, C. and Marden, J. I. (1983a). Hypothesis testing for marginal probabilities in a $2 \times 2 \times 2$ contingency table with conditional independence, *Journal of the American Statistical Association*, **78**, 920–929.

Cohen, A., Gatsonis, C. and Marden, J. I. (1983b). Hypothesis tests and optimality properties in discrete multivariate analysis, in *Studies in Econometrics, Time Series, and Multivariate Statistics* (Eds., S. Karlin, T. Amemiya and L. A. Goodman), pp. 379–405, New York:Academic Press.

Davis, L. J. (1989). Intersection union tests for strict collapsibility in three-dimensional contingency tables, *Annals of Statistics*, **17**, 1693–1708.

Gleser, L. J. (1973). On a theory of intersection-union tests, *Institute of Mathematical Statistics Bulletin*, **2**, 233. Abstract.

Lehmann, E. L. (1952). Testing multiparameter hypotheses, *Annals of Mathematical Statistics*, **23**, 541–552.

Liu, H. and Berger, R. L. (1995). Uniformly more powerful, one-sided tests for hypotheses about linear inequalities, *Annals of Statistics*, **23**, 55–72.

Roy, S. N. (1953). On a heuristic method of test construction and its use in multivariate analysis, *Annals of Mathematical Statistics*, **24**, 220–238.

Saikali, K. G. (1996). Uniformly more powerful tests for linear inequalities, Ph.D. dissertation, North Carolina State University.

Sasabuchi, S. (1980). A test of a multivariate normal mean with composite hypotheses determined by linear inequalities, *Biometrika*, **67**, 429–439.

Sasabuchi, S. (1988a). A multivariate test with composite hypotheses determined by linear inequalities when the covariance matrix has an unknown scale factor, *Memoirs of the Faculty of Science, Kyushu University, Series A*, **42**, 9–19.

Sasabuchi, S. (1988b). A multivariate one-sided test with composite hypotheses when the covariance matrix is completely unknown, *Memoirs of the Faculty of Science, Kyushu University, Series A*, **42**, 37–46.

Wang, Y. and McDermott, M. P. (1996). Construction of uniformly more powerful tests for hypotheses about linear inequalities, *Technical Report 96/05*, Departments of Statistics and Biostatistics, University of Rochester.