M4056 Hypothesis Testing IV. **November 3,5, 2010**

A. Review

I'll summarize the testing situation in a nutshell. Let θ parametrize a a family of random variables X_θ. Let $\vec{X}_\theta$ be a sample consisting of n random variables that are i.i.d. X_θ.[1] Assume that a value of θ is fixed, and sample data $\vec{x}$ is produced. The statistician makes judgments about θ based upon $\vec{x}$.

In hypothesis testing,

- we assume that the set of all possible values of θ is the disjoint union of two subsets Θ_0 and Θ_1; we take H_0 be the "hypothesis" that θ belongs to Θ_0 rather than Θ_1;
- we divide the possible values of $\vec{X}$ into two disjoint sets: A, the acceptance region, and R, the rejection region, thus creating a "test" of H_0.

The *power function* of a test is $\beta(\theta) := P_\theta(\vec{X} \in R)$. This is the probability of rejection, as a function of the parameter. The *size* of a test is $\sup\{\, \beta(\theta) \mid \theta \in \Theta_0 \,\}$. This is the maximum probability of rejection if H_0 is true. If the size of a test is less than a given number, we say the *level of significance* is (better than) that number.

In practice, a test constructed to meet two criteria. First, a level of significance is given. Second, among tests attaining the desired level of significance, the more powerful tests are sought. Of course, the power varies with θ, so in general comparing power means comparing two functions. If T and T' are two tests determined by rejection regions R and R' and having power functions β and β', we say T *is uniformly more powerful than* T' if $\beta \geq \beta'$ on the set Θ_1. In other words, for all θ, if $H_0(\theta)$ is false, then $P_\theta(\vec{X} \in R) \geq P_\theta(\vec{X} \in R')$, i.e., T has greater probability of rejecting H_0 than T'.

B. Hypotheses with Simple Alternatives

This concerns the situation in which θ has only two possible values. In this case, we say that we have *simple alternative hypotheses*:

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta = \theta_1. \tag{$*$}$$

Example. Tony's Seafood is having a special on 20-pound packages of frozen shrimp. The shrimp come in two sizes: jumbo (12 to the pound) and large (16 to the pound). You figure the jumbo are a good deal, but not the large. Unfortunately, the packages are not marked. Fortunately, however, the sales clerk will weigh four shrimp from the package you select, at your request. How can you use this information to decide whether or not to buy? (Assume both kinds of shrimp are normally distributed by weight, with a standard deviation of 1/4 ounce.)

---

[1] Note that most authors omit the subscript θ, or attach it to other symbols. For example, if R is a subset of sample space many authors will write $P(\vec{X} \in R \mid \theta)$ or $P_\theta(\vec{X} \in R)$ in place of $P(\vec{X}_\theta \in R)$.

C. The Likelihood Ratio Test for Simple Alternatives

Likelihood ratio tests for simple alternatives are quite simple in from. Let $f(\vec{x} \mid \theta_0)$ and $f(\vec{x} \mid \theta_1)$ be the *pdf*s (or *pmf*s) of the sample under the two hypotheses. According to 8.2.1, the likelihood ratio for H_0 is:

$$\lambda(\vec{x}) = \frac{f(\vec{x}|\theta_0)}{\max\{f(\vec{x}|\theta_0), f(\vec{x}|\theta_1)\}}.$$

The likelihood ratio test determined by a number c (with $0 \leq c \leq 1$) consists in rejecting H_0 if $\lambda(\vec{x}) < c$. You should think of c as a number chosen by the statistician. In order to increase significance, a smaller c is chosen.

If in a likelihood ratio test of simple alternatives $\lambda(\vec{x}) < 1$, then $\lambda(\vec{x}) = \lambda^*(\vec{x})$, where

$$\lambda^*(\vec{x}) = \frac{f(\vec{x}|\theta_0)}{f(\vec{x}|\theta_1)}.$$

Thus, for $c < 1$, the likelihood ratio test is equivalent to rejecting H_0 if:

$$f(\vec{x}|\theta_0) < cf(\vec{x}|\theta_1).$$

Setting $k = 1/c$, we get a test T_k with rejection region

$$R_k = \{\, \vec{x} \mid f(\vec{x}|\theta_1) > kf(\vec{x}|\theta_0) \,\}. \tag{T_k}$$

(This is the test referred to in (8.3.1) of the textbook.) Let $\beta_k(\theta)$ be the power function of this test and let α_k be its size. For the rest of these notes, we reserve the symbols T_k, R_k, α_k and β_k for likelihood ratio tests of the form we have just discussed.

D. Significance and Power for Tests of Simple Alternatives

When testing simple alternatives, the size and the power function are also particularly simple in form (regardless of whether the test is a likelihood ratio test). Let T be a test of simple alternatives with rejection region R. The size of T is

$$\alpha = P_{\theta_0}(\vec{X} \in R) = \int_R f(\vec{x}|\theta_0)d\vec{x} = \int \mathbf{1}_R f(\vec{x}|\theta_0)d\vec{x},$$

where $\mathbf{1}_R$ is the indicator function of R. If T' is another test with rejection region R', then its size is $\alpha' = \int \mathbf{1}_{R'} f(\vec{x}|\theta_0)d\vec{x}$. Then,

$$\begin{aligned} T \text{ has greater size than } T' \;&\Leftrightarrow\; \int \mathbf{1}_R f(\vec{x}|\theta_0)d\vec{x} \geq \int \mathbf{1}_{R'} f(\vec{x}|\theta_0)d\vec{x}. \\ &\Leftrightarrow\; \int (\mathbf{1}_R - \mathbf{1}_{R'}) f(\vec{x}|\theta_0)d\vec{x} \geq 0. \end{aligned} \tag{sig}$$

T is uniformly more powerful than T' if and only if $\beta(\theta_1) \geq \beta'(\theta_1)$. Since

$$\beta(\theta_1) = P_{\theta_1}(\vec{X} \in R) = \int \mathbf{1}_R f(\vec{x}|\theta_1)d\vec{x},$$

we deduce:

$$T \text{ is uniformly more powerful than } T' \;\Leftrightarrow\; \int (\mathbf{1}_R - \mathbf{1}_{R'}) f(\vec{x}|\theta_1)d\vec{x} \geq 0. \tag{pow}$$

E. Size and Power of T_k Compared to Other Tests

Let $k > 0$ and let T_k be a likelihood ratio test of simple alternatives, as in section **C**. The relationships of the size and power of T_k to any other test T of the alternatives have elegant mathematical representations.

Let T be another test. (T need not be a likelihood ratio test.) Let R be its rejection region, let $\beta(\theta)$ be its power function and let α be its size. Consider the product

$$(\mathbf{1}_{R_k} - \mathbf{1}_R) \cdot \big(f(\vec{x}|\theta_1) - kf(\vec{x}|\theta_0)\big).$$

If $\vec{x} \in R_k$ then both factors are ≥ 0. When $x \notin R_k$, both factors are ≤ 0. This shows that this product is ≥ 0. It follows that the integral over sample space is also non-negative:

$$0 \leq \int (\mathbf{1}_{R_k} - \mathbf{1}_R)\big(f(\vec{x}|\theta_1) - kf(\vec{x}|\theta_0)\big)\, d\vec{x}$$

But the integral can be expressed in terms of the power and size:

$$0 \leq \big(\beta_k(\theta_1) - \beta(\theta_1)\big) - k(\alpha_k - \alpha).$$

We can rewrite this:

$$k(\alpha_k - \alpha) \leq \big(\beta_k(\theta_1) - \beta(\theta_1)\big). \tag{$**$}$$

The entire argument also works with the roles of T_k and T reversed, giving $k(\alpha - \alpha_k) \leq \big(\beta(\theta_1) - \beta_k(\theta_1)\big)$.

E. The Neyman-Pearson Lemma.

Let k be a positive number. Let T_k be the test of H_0 described at the end of section **C**.

Part (a). *If T is any test of H_0 with size less than or equal to α_k, then T_k is uniformly more powerful than T. In other words, the likelihood ratio test is uniformly most powerful among all tests of H_0 with size at most α_k.*

Proof. This is a direct consequence of ($**$), for if $\alpha_k \geq \alpha$, then $\beta_k(\theta_1) \geq \beta(\theta_1)$. /////

Part (b). *If T is uniformly most powerful among all tests of level α_k, then T has size α_k and T is defined by a rejection region R that differs from R_k by a set that has measure 0 with respect to both $f(\vec{x}|\theta_0)$ and $f(\vec{x}|\theta_1)$.*

Proof. Since T has leveI α_k, $\alpha \leq \alpha_k$. If T is (uniformly) most powerful among all tests of level α_k then by Part (a), $\beta(\theta_1) = \beta_k(\theta_1)$. From this and ($**$), it follows that that $\alpha_k \leq \alpha$. Hence $\alpha_k = \alpha$. The last statement follows from these two equalities by the formulae for α and β given in secton D. /////