DECIDING BETWEEN TWO SIMPLE HYPOTHESES

1. Definitions of size and power of tests

Let H_0 be any simple hypothesis, specifying a likelihood function $f_0 = f_0(x)$. For any test of the hypothesis, the *size* of the test is the probability of rejecting H_0 when it's true. The convention we have been using has been to set sizes α of tests at .05, but in testing against a simple alternative, other sizes with $0 \leq \alpha < 1$ are possible. If a test statistic has a discrete set of values, there may be no test based on it of size equal to 0.05.

Suppose H_1 is another simple hypothesis such that exactly one of H_0 and H_1 is true. Then the *power* of a test of H_0 against the alternative H_1 is the probability of (correctly) rejecting H_0 (and deciding in favor of H_1) if H_1 is true. Both small size and large power are good properties, but there is a tradeoff between the two.

Suppose we have two hypotheses, H_0 and H_1, such that if H_i holds, the likelihood function for one observation is f_i for $i = 0, 1$. In one sense, what we want to do is test the simple hypothesis H_0 against the simple alternative H_1. Another view is that we want to decide on one of H_0 or H_1.

2. The likelihood ratio and the Neyman–Pearson lemma

Let's define the *likelihood ratio* $LR(x)$ as $f_1(x)/f_0(x)$ if this is defined and finite; as $+\infty$ if $f_0(x) = 0 < f_1(x)$; and as 0 if $f_1(x) = f_0(x) = 0$. If we have $X_1, ..., X_n$ i.i.d. f_i where i is the same for all X_j, $i = 0$ or 1, and i is unknown, we get a likelihood ratio $\prod_{j=1}^{n}(f_1/f_0)(X_j)$. An undefined product $0 \cdot \infty$ is not possible because if, say, $f_1(X_j) = 0$, then with probability 1, $i = 0$, and then $f_0(X_k) > 0$ for all $k = 1, \ldots, n$. For each C with $0 < C < +\infty$ we can define two *likelihood ratio tests* of H_0 vs. H_1: one is to reject H_0 and decide in favor of H_1 if $LR(x) \geq C$, and the other is similar but with $> C$ instead of $\geq C$. If $LR(x)$ has a continuous distribution under each of H_0 and H_1, then the probability that $LR(x) = C$ is 0 under either hypothesis and the two tests are essentially equivalent.

For deciding between H_0 and H_1, a basic fact is the following:

Date: 18.650, Oct. 14, 2015.

Theorem 1 (Neyman–Pearson Lemma). *For any simple hypothesis* H_0 *and simple alternative* H_1*, for any* $0 < C < +\infty$*, and either likelihood ratio test* T *of* H_0 *vs.* H_1 *for that* C*, the power of* T *against* H_1 *is as large, or larger, than that of any other test* U *of* H_0 *vs.* H_1 *whose size is less or equal to that of* T*.*

This fact follows from the formulation of the Neyman–Pearson Lemma given in Rice, p. 332, and proved there. A correction to Rice's statement: "and significance level α" should be "has significance level α." In the terminology here one could say "has size α."

3. Costs of errors

The question then is, how to choose C. One consideration is the costs of errors. Let c_i be the cost (or loss) if H_i is true but not chosen, for $i = 0, 1$. Then c_0 and c_1 may be very different. For example, let H_0 be the hypothesis that an individual being tested does not have a disease D, and H_1 the hypothesis that the individual does have D. If the physician or tester decides in favor of H_1 while H_0 is true ("false positive"), then c_0 would include the cost (in money and time) of further tests until it was eventually realized that H_0 is true. Whereas, if a decision in favor of H_0 is made when H_1 is true, and if D is serious, and there are good treatments for it, but it goes untreated for a while in the given patient, the disease may get worse and lead to quite a high cost c_1. Such a situation is asymmetric: there is no reason to think that $c_0 = c_1$.

If a test has size α at H_0 and power β against H_1, then if H_0 is true, the *risk* (expected cost) is αc_0. If H_1 is true, the risk is $(1-\beta)c_1$. One might perhaps want to choose C in a likelihood ratio test so as to minimize the maximum of these two risks.

4. Prior probabilities

But there is yet another consideration. There may be information available, based on which one may be able to assign a *prior* probability π_0 that H_0 is true, and so $\pi_1 = 1 - \pi_0$ that H_1 is true, before doing the test. In the example, the disease D may have a known prevalence (relative frequency of occurring) in a population including the person being tested of π_1, and then $\pi_0 = 1 - \pi_1$. Including the prior probabilities, the overall risk of the test would be $\pi_0 \alpha c_0 + \pi_1(1-\beta)c_1$ and we'd like to choose C in the likelihood ratio test to minimize this overall risk.

The overall likelihood function is $f(x) = \pi_0 f_0(x) + \pi_1 f_1(x)$. After doing the test, the *posterior* probability of H_0, or the conditional probability that H_0 is true given that x is observed, is

$$P(H_0|x) = \frac{\pi_0 f_0(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)}, \tag{1}$$

which follows from Bayes' theorem (or formula) in case of discrete distributions. For continuous distributions of x we can consider the bivariate distribution of (i, x) where $i = 0$ or 1. The marginal density of x is f and its conditional density given i is f_i. The conditional probability that H_0 is true given x is as shown in (1). Likewise, the conditional, or posterior, probability that H_1 is true given that x is observed is

$$P(H_1|x) = \frac{\pi_1 f_1(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)}. \tag{2}$$

If x is observed, the conditional = posterior risk (expected cost) of choosing H_1 is

$$\frac{\pi_0 f_0(x) c_0}{\pi_0 f_0(x) + \pi_1 f_1(x)}, \tag{3}$$

and the conditional risk (expected cost) of choosing H_0 is

$$\frac{\pi_1 f_1(x) c_1}{\pi_0 f_0(x) + \pi_1 f_1(x)}. \tag{4}$$

We want to decide in favor of the hypothesis having smaller posterior risk. Since the denominators of (3) and (4) are the same, this means we want to choose H_1 if $\pi_1 f_1(x) c_1 > \pi_0 f_0(x) c_0$, choose H_0 if $\pi_1 f_1(x) c_1 < \pi_0 f_0(x) c_0$, and expect equal costs for either choice if $\pi_1 f_1(x) c_1 = \pi_0 f_0(x) c_0$. All factors make intuitive sense: for any of π_i, $f_i(x)$, or c_i to be larger than the corresponding numbers with $1 - i$ in place of i will incline us to choose H_i. We will choose H_1 if

$$LR(x) = \frac{f_1(x)}{f_0(x)} > \frac{\pi_0 c_0}{\pi_1 c_1},$$

which is a likelihood ratio test with $C = \pi_0 c_0 / (\pi_1 c_1)$.