11: HYPOTHESIS TESTING II

Larsen & Marx 6.4, 6.5

Prof Amanda Luby

Last time, we laid the groundwork for a more theoretical treatment of hypothesis testing. Today, we're going to continue that thread by talking about *power functions* and *testing errors* in a more theoretical framework. We'll end by introducing the *likelihood ratio test*, a method for deriving hypothesis test procedures.

1 Power Function

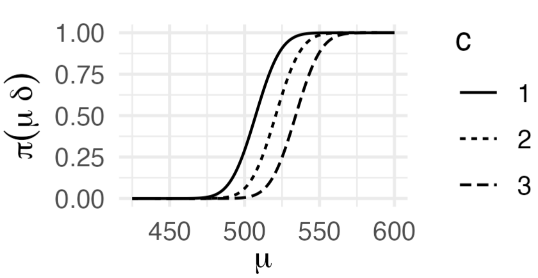
In order to generalize hypothesis testing procedures, it is useful to define the *power function* of a test (sometimes called a *power curve*).

Power Function

Let δ be a test procedure and denote $\pi(\theta|\delta)$ as the power function of the test. If δ is defined in terms of T and rejection region R, then:

Note:

Example: Math curriculum example from last time: In the year of the study, 86 sophomores were randomly selected to participate in a special set of classes that integrated geometry and algebra. Those students averaged 502 on the SAT-I math exam; the nationwide average was 494 with a standard deviation of 124. Find the power function $\pi(\theta|\delta)$ for the test δ defined last time.



2 Types of Errors

In any hypothesis test procedure, there are two ways we can be wrong: we can (1) conclude H_0 is true when H_1 is actually true, or we can (2) conclude H_0 is false when H_0 is actually true.

	H_0 True	H_1 True
Reject H_0	Type I Error	Correct
Fail to reject H_0	Correct	Type II Error

- If $\theta \in \Omega_0$:
- If $\theta \in \Omega_1$:

Solution:

Level-α_0 Test

A test that satisfies the above is called a *level* α_0 *test* and we say it has *significance level* α_0. In addition, the *size* $\alpha(\delta)$ of a test is defined as:

A test is a level α_0 test if and only if its size is at most α_0. If H_0 is simple, $\alpha(\delta) = \pi(\theta|\delta)$.

Example: Suppose that a random sample $X_1, ..., X_n$ is taken from the uniform distribution on the interval $[0, \theta]$, where θ is unknown but positive, and suppose we wish to test the following hypotheses. Find the power function and size of the test.

$H_0 : 3 \leq \theta \leq 4$

$H_1 : \theta < 3$ or $\theta > 4$

3 Likelihood Ratio Test

Many of the most popular hypothesis tests used in practice have the same conceptual heritage - a fundamental notion known as the *Generalized likelihood ratio* or GLR.

Example: Suppose $X_1, ..., X_n \sim Unif(0, \theta)$ and we wish to test $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$.

Generalized likelihood ratio

Let $y_1, ..., y_n$ be iid from $f_y(y; \theta)$. The generalized likelihood ratio is defined as:

Generalized likelihood ratio test

A generalized likelihood ratio test (GLRT) is one that rejects H_0 when

Let f_Λ denote the PDF of the GLR under H_0. If we knew what the pdf was, we could find λ^* and δ by solving:

Generally, however, we can't find f_Λ. Instead, we find a quantity W that we *do* know the distribution of, and show that Λ is a monotone function of W. Then, a test based on W is equivalent to one based on Λ.

Back to example: