HW 8, Lab 5 due tonight
Friday: little bit of R

Snapshot of grades on moodle (through HW 6/quiz 2)

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.

$H_0: \theta \in \Omega_0$
$H_1: \theta \in \Omega_1$

- T: test statistic – function of our data (X)
- θ: unknown parameter
- R: rejection region: $\{T : \text{reject } H_0\}$
- Ω_0, Ω_1: null and alternative parameter spaces
- δ: decision rule/hypothesis test → "reject H_0 when T ..."

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:
>
> $\pi(\theta|\delta) = P(T \in R \mid \theta)$ for $\theta \in \Omega$
>
> π is a function of θ and gives $P(\text{reject } H_0)$

Note: Ideal power function?

(1) $\theta \in \Omega_0 \rightarrow \pi(\theta|\delta) = 0$

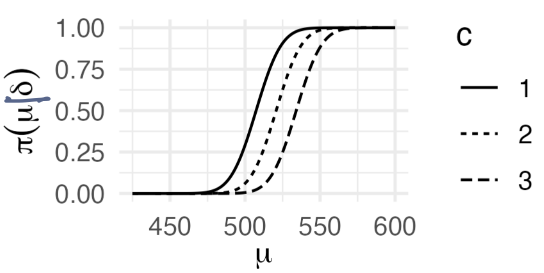
(2) $\theta \in \Omega_1 \rightarrow \pi(\theta|\delta) = 1$

correct decision w/ probability 1

in practice, want $\pi(\theta|\delta)$ close to 0 when $\theta \in \Omega_0$ and close to 1 when $\theta \in \Omega_1$

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.

P(reject H_0)



$H_0: \mu = 494$
$H_1: \mu > 494$

Decision rule δ: reject H_0 when

$$\underbrace{Z = \frac{\bar{Y} - 494}{124/\sqrt{86}} > 1.64}_{\text{rejection region}}$$

T

$R = [c, \infty)$, $c = 1.64$

$\mu_0 = 494$
$\sigma = 124$
$n = 86$

$\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$

true pop. param

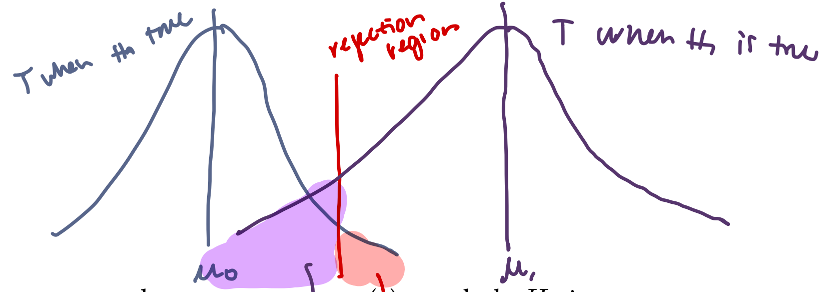
$$\pi(\theta|\delta) = P(T \in R \mid \mu) = P\left(\frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}} > c \mid \mu\right) = P(\bar{Y} > c \cdot \sigma/\sqrt{n} + \mu_0 \mid \mu)$$

$$= P\left(\frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} > \frac{c\,\sigma/\sqrt{n}}{\sigma/\sqrt{n}} + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(Z > c + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)$$

$$= 1 - \Phi\left(c + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)$$

$P(Z \le z^*)$
pnorm()



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.

P(Type I) = α

P(Type II)

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

- If $\theta \in \Omega_0$: $\pi(\theta|\delta) = P(\text{Type I error})$
- If $\theta \in \Omega_1$: $1 - \pi(\theta|\delta) = P(\text{type II error})$

$\rightarrow$ only 1 possible error for any θ, but we never know whether $\theta \in \Omega_0$ or $\theta \in \Omega_1$

Solution:

Choose $\alpha_0 \in (0,1)$ such that $\pi(\theta|\delta) \le \alpha_0$ for all $\theta \in \Omega_0$

If simple H_0: Set α_0 such that $\pi(\theta|\delta) \le \alpha_0$ for $\theta = \theta_0$

$\pi_\delta(\theta)$

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:

$$\max_{\theta \in \Omega_0} \pi(\theta|\delta) \leftarrow \max(\text{Type I error}) \text{ over all possible } \theta \in \Omega_0$$

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

Hw a posted tonight due Wed

Example: Suppose that a random sample $X_1, \ldots, X_n$ 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.

Quiz 3 - Mon before Thank.

$H_0 : 3 \le \theta \le 4$

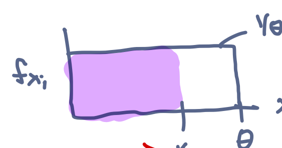
$H_1 : \theta < 3 \text{ or } \theta > 4$

Recall that the MLE is $\hat{\theta} = X_{max} \rightarrow \hat{\theta}$ will be close to θ for large n, but always $< \theta$

Define $\delta = \begin{cases} \text{Do not reject} & 2.9 < X_{max} < 4 \\ \text{Reject } H_0 & X_{max} \le 2.9 \text{ or } X_{max} \ge 4 \end{cases}$

$R = (-\infty, 2.9] \cup [4, \infty)$

$\pi(\theta|\delta) = P(X_{max} \in R) = P(X_{max} \le 2.9 | \theta) + P(X_{max} \ge 4 | \theta)$



$P(X_i \le x) = \frac{x}{\theta}$

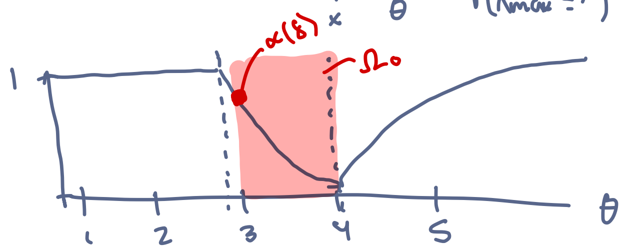
$\Downarrow$

$P(X_{max} \le x) = \left(\frac{x}{\theta}\right)^n$

$\theta < 2.9$: $P(X_{max} \le 2.9) = 1$, $P(X_{max} \ge 4) = 0$ $\}$ $\pi(\theta|\delta) = 1$

$2.9 \le \theta \le 4$: $P(X_{max} \le 2.9) = \left(\frac{2.9}{\theta}\right)^n$, $P(X_{max} \ge 4) = 0$ $\}$ $\pi(\theta|\delta) = \left(\frac{2.9}{\theta}\right)^n$

$\theta > 4$: $P(X_{max} \le 2.9) = \left(\frac{2.9}{\theta}\right)^n$, $P(X_{max} \ge 4) = 1 - \left(\frac{4}{\theta}\right)^n$ $\}$ $\pi(\theta|\delta) = \left(\frac{2.9}{\theta}\right)^n + 1 - \left(\frac{4}{\theta}\right)^n$



$\alpha(\delta) = \left(\frac{2.9}{3}\right)^n$

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 $H_1 : \theta < \theta_0$.

Basic idea: How much more likely is H_0 compared to H_1 for given data?

Recall the likelihood function: $L(\theta) = \prod_{i=1}^{n} f_x(x_i; \theta) = \begin{cases} (1/\theta)^n & 0 \le x_i \le \theta \\ 0 & \text{otherwise} \end{cases}$

We'll maximize $L(\theta)$ twice: once under H_0 and once under H_1

$$\theta \in \Omega_0 : \max_{\theta \in \Omega_0} L(\theta) = L(\theta_0) = \begin{cases} \left(\frac{1}{\theta_0}\right)^n & 0 \le x_{max} \le \theta \\ 0 & \text{otherwise} \end{cases}$$

$$\theta \in \Omega_1 : \max_{\theta \in \Omega_1} L(\theta) = \begin{cases} \left(\frac{1}{x_{max}}\right)^n & 0 \le x_{max} \le \theta \\ 0 & \text{otherwise} \end{cases}$$

How much more likely is H_0 compared to H_1?

$$\frac{(1/\theta_0)^n}{(1/x_{max})^n} = \left(\frac{x_{max}}{\theta_0}\right)^n \rightarrow$$

close to 1: not much evidence against H_0

close to 0: lots of evidence against H_0

Generalized likelihood ratio

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

$$\lambda = \frac{\max_{\Omega_0} L(\theta)}{\max_{\Omega_1} L(\theta)} \rightarrow \frac{L(\Omega_0)}{L(\Omega_1)} \text{ (shorthand)}$$

Generalized likelihood ratio test

A generalized likelihood ratio test (GLRT) is one that rejects H_0 when $0 < \lambda < \lambda^*$

where λ^* is chosen so that $P(0 < \Lambda \le \lambda^* \mid \theta \in \Omega_0) = \alpha$

random variable version of λ

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:

$$\alpha = \int_0^{\lambda^*} f_\Lambda(\lambda)\, d\lambda$$

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:

$$P(\Lambda \le \lambda^* \mid \theta \in \Omega_0) = \alpha$$

$$\downarrow$$

$$P\left(\left(\frac{X_{max}}{\theta_0}\right)^n \le \lambda^* \mid \theta = \theta_0\right)$$

$$= P\left(\frac{X_{max}}{\theta_0} \le \sqrt[n]{\lambda^*} \mid \theta = \theta_0\right)$$

$$= P(X_{max} \le \theta_0 \sqrt[n]{\lambda^*} \mid \theta = \theta_0)$$

$$P(X_{max} \le x) = \left(\frac{x}{\theta}\right)^n$$

Let $W = X_{max}$ $\quad w^* = \theta_0 \sqrt[n]{\lambda^*}$

$$\alpha = \left(\frac{\theta_0 \sqrt[n]{\lambda^*}}{\theta_0}\right)^n$$

$$\begin{cases} \sqrt[n]{\alpha} = \sqrt[n]{\lambda^*} \\ \alpha = \lambda^* \end{cases}$$

$\to P(W \le w^* \mid \theta = \theta_0) = \alpha$

$\to w^* = \theta_0 \sqrt[n]{\lambda^*} = \theta_0 \sqrt[n]{\alpha}$

δ = reject H_0 if $W = X_{max} \le \theta_0 \sqrt[n]{\alpha}$

= reject H_0 if $\Lambda = \left(\frac{X_{max}}{\theta_0}\right)^n \le \lambda^* = \alpha$