Lecture 19: Hypothesis testing framework

Ciaran Evans

General framework for hypothesis tests

Definition: Let $\theta \in \mathbb{R}^p$ be some parameter of interest. A **hypothesis** is a statement about θ. The goal of a hypothesis test is to compare two competing hypotheses:

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

- $\Theta_0 \cap \Theta_1 = \emptyset$
- If $\Theta_0 = \{\theta_0\}$, H_0 is a simple hypothesis

otherwise, Θ_0 is a composite hypothesis

(likewise for Θ_1)

Ex: $H_0 : \begin{bmatrix} \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\qquad$ $H_A : \begin{bmatrix} \beta_3 \\ \beta_4 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

simple null $\qquad$ composite alternative

Outcomes

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

The test either **rejects** H_0 or **fails to reject** H_0. Possible outcomes:

Truth

Decision	H_0 is true	H_0 is false
Fail to reject	✓	type II error (false negative)
reject	type I error (false positive)	✓

Goal

	H_0 is true	H_0 is false
fail to reject	correct decision	type II error
reject	type I error	correct decision

Goal: Minimize type II error rate, subject to control of type I error rate

Constructing a test

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

Observe data $x_1, \dots, x_n$

① Calculate a test statistic $T_n = T(X_1, \dots, X_n)$

② Choose a rejection region

$R = \{ (x_1, \dots, x_n) : \text{reject } H_0 \}$

③ Reject H_0 if $(X_1, \dots, X_n) \in R$

Constructing a test

Given observed data $X_1, \dots, X_n$:

1. Calculate a test statistic $T_n = T(X_1, \dots, X_n)$
2. Choose a rejection region $\mathcal{R} = \{(x_1, \dots, x_n) : \text{ reject } H_0\}$
3. Reject H_0 if $(X_1, \dots, X_n) \in \mathcal{R}$

Example: $X_1, \dots, X_n$ iid with mean μ and variance σ^2

① $T(X_1, \dots, X_n) = \dfrac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$

② $\mathcal{R} = \left\{ (x_1, \dots, x_n) : \left| \dfrac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}} \right\}$

rejection region is determined by desired type I error rate α

Prob. of a type I error $= P(\text{reject } H_0 \mid H_0 \text{ is true})$

$= P\left(\left| \dfrac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}} \,\Big|\, \mu = \mu_0 \right)$

Power function

Suppose we reject H_0 when $(X_1, ..., X_n) \in R$. The **power function** $\beta(\theta)$ is

$$\beta(\theta) = \underbrace{P_\theta((X_1, ..., X_n) \in R)}_{\text{probability of rejecting } H_0 \text{ if true parameter is } \theta}$$

function of θ

Goal: Want $\beta(\theta)$ small when $\theta \in \Theta_0$
and $\beta(\theta)$ large when $\theta \in \Theta_1$

Formally: ① Fix $\alpha \in [0,1]$
② Try to maximize $\beta(\theta)$ for $\theta \in \Theta_1$
subject to $\beta(\theta) \leq \alpha$ for $\theta \in \Theta_0$

- If $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$, our test is size α
- If $\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$, our test is level α

Example

$X_1, \ldots, X_n$ iid from a population with mean μ and variance σ^2.

$$H_0 : \mu = \mu_0 \qquad H_A : \mu \neq \mu_0$$

$$\beta(\mu) = P_\mu\left(\left| \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}} \right)$$

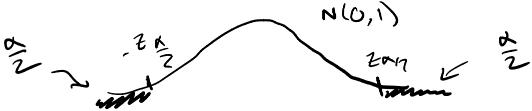
$$= \underbrace{P_\mu\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}} \right)} + P_\mu\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} < -z_{\frac{\alpha}{2}} \right)$$

$$P_\mu\left(\frac{\bar{X}_n - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}} \right)$$

$$\hookrightarrow P_\mu\left(\underbrace{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}_{\approx N(0,1) \quad (CLT)} > z_{\alpha/2} - \left(\frac{\mu - \mu_0}{\sigma/\sqrt{n}} \right) \right)$$

$$\Rightarrow \beta(\mu) \approx 1 - \Phi\left(z_{\frac{\alpha}{2}} - \left(\frac{\mu - \mu_0}{\sigma/\sqrt{n}} \right) \right) + \Phi\left(-z_{\frac{\alpha}{2}} - \left(\frac{\mu - \mu_0}{\sigma/\sqrt{n}} \right) \right)$$

Example



$X_1, \ldots, X_n$ iid from a population with mean μ and variance σ^2.

$$H_0 : \mu = \mu_0 \qquad H_A : \mu \neq \mu_0$$

$$\beta(\mu) \approx 1 - \Phi\left(z_{\frac{\alpha}{2}} - \left(\frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right)\right) + \Phi\left(-z_{\frac{\alpha}{2}} - \left(\frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right)\right)$$

For any n: if $\mu = \mu_0$ (H_0 is true)

$$\beta(\mu) \approx 1 - \Phi(z_{\frac{\alpha}{2}}) + \Phi(-z_{\frac{\alpha}{2}}) = \alpha$$

<u>Fix n</u>: as $|\mu - \mu_0| \to \infty$, $\beta(\mu) \to 1$

<u>Fix $\mu \neq \mu_0$</u>: as $n \to \infty$, $\beta(\mu) \to 1$

Rejecting H_0

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

Question: A hypothesis test rejects H_0 if $(X_1, ..., X_n)$ is in the rejection region $\mathcal{R}$. Are there any issues if we only use a rejection region to test hypotheses?

rejection region: only gives binary info
(in region or not)

p-values

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

Given α, we construct a rejection region $\mathcal{R}$ and reject H_0 when $(X_1, ..., X_n) \in \mathcal{R}$. Let $(x_1, ..., x_n)$ be an observed set of data.

Definition: The **p-value** for the observed data $(x_1, ..., x_n)$ is the smallest α for which we reject H_0.

Intuition:

reject H_0 when p-value $< \alpha$

$\Rightarrow \inf\{\alpha : \text{reject } H_0\} = \inf\{\alpha : \text{p-value} < \alpha\}$

$= \text{p-value}$

p-values

Suppose we have a test which rejects H_0 when $T(X_1, ..., X_n) > c_\alpha$, where c_α is chosen so that

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \sup_{\theta \in \Theta_0} P_\theta(T(X_1, ..., X_n) > c_\alpha) = \alpha$$

Let $x_1, ..., x_n$ be a set of observed data.

Theorem: The p-value for the set of observed data $x_1, ..., x_n$ is

$$p = \sup_{\theta \in \Theta_0} P_\theta(T(X_1, ..., X_n) > \underbrace{T(x_1, ..., x_n)}_{\text{observed test statistic}})$$

"probability of observed test statistic or 'more extreme', under H_0"

Proof of theorem