Lecture 20: Hypothesis testing framework

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:

Ex: $H_0: p = 0.5$
$H_A: p = 0.7$
(don't need to cover parameter space)

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

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

otherwise, H_0 is a composite hypothesis
(likewise for H_A)

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

simple null $\quad$ composite alternative

Outcomes

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

(Assume that Θ_0 and Θ_1 partition the parameter space)

The outcome of the test is a decision to either **reject** H_0 or **fail to reject** H_0.

Four possibilities

Decision		Truth: H_0 is true	Truth: H_A is true (H_0 is false)
	Fail to reject H_0	Yay!	type II error (false negative)
	Reject H_0	type I error (false positive)	Yay!

Goal:

1) Control type I error at some specified level (α)

2) Minimize type II error among tests that control type I error

Constructing a test

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

observe data $X_1, \ldots, X_n$

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

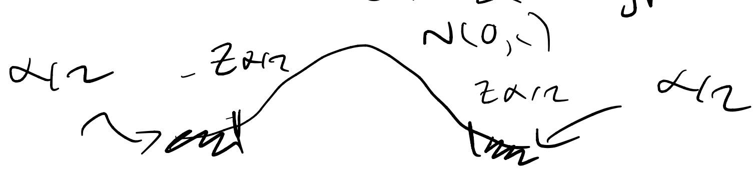
Example: $X_1, \ldots, X_n$ iid from pop. w/ mean μ, variance σ^2

$H_0 : \mu = \mu_0$ $\quad$ $H_A : \mu \neq \mu_0$ $\quad$ (assume σ^2 known)

1. $T(X_1, \ldots, X_n) = \dfrac{\bar{X}_n - \mu_0}{\sigma / \sqrt{n}}$
2. $R = \left\{ (x_1, \ldots, x_n) : \left| \dfrac{\bar{x}_n - \mu_0}{\sigma / \sqrt{n}} \right| > z_{\frac{\alpha}{2}} \right\}$

rejection region is determined by our desired type I error rate (α)

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



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

Power function

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

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

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 $\beta(\theta) \le \alpha$ for $\theta \in \Theta_0$

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

$\Phi = N(0,1)$ cdf

Example

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

$\bar{X}_n \approx N(\mu, \frac{\sigma^2}{n})$

$\iff \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$

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

reject H_0 when $\left| \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}}$

Power function:

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

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

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

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

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

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

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

Fix n, $|\mu-\mu_0| \to \infty$ $\Rightarrow$ $\beta(\mu) \to 1$

Fix μ, $n \to \infty$ $\Rightarrow$ $\beta(\mu) \to 1$

$$\Phi\Bigg(\underbrace{-z_{\alpha/2} - \frac{\sqrt{n}(\mu-\mu_0)}{\sigma}}_{\substack{\to -\infty \\ \text{or } \to \infty}}\Bigg)$$

(sign of $\mu-\mu_0$)

For any n, $\mu=\mu_0$: $\beta(\mu_0) = \alpha$

Rejecting H_0

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

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

p-values

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

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

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

Example

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

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