Outline

10/12/2021

1) Hypothesis testing
2) Neyman - Pearson Lemma

Hypothesis Testing

In <u>hypothesis testing</u>, we use data X to infer which of two submodels generated X

Model $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$

Null hypothesis $H_0 : \theta \in \Theta_0$

Alternative hyp. $H_1 : \theta \in \Theta_1$

(Whenever H_1 unspecified, assume $\Theta_1 = \Theta \setminus \Theta_0$)

H_0 is "default choice": we either

1. accept H_0 (fail to reject, no definite concl.)
2. reject H_0 (conclude Θ_0 false, Θ_1 true)

<u>Ex</u> $X \sim N(\theta, 1)$ $H_0 : \theta \leq 0$ vs $H_1 : \theta > 0$

or $H_0 : \theta = 0$ vs $H_1 : \theta \neq 0$

<u>Ex</u> $X_1, \dots, X_n \sim P$ $Y_1, \dots, Y_m \sim Q$ $H_0 : P = Q$ vs $H_1 : P \neq Q$

[Common conceptual objection: we "know" $\theta \neq 0$ or $P \neq Q$ already, why bother?
We will return to this.]

Power Function

Can describe a test formally by its <u>critical function</u> (a.k.a. <u>test function</u>)

$$\phi(x) = \begin{cases} 0 & \text{accept } H_0 \\ \pi \in (0,1) & \text{reject w.p. } \pi \\ 1 & \text{reject } H_0 \end{cases}$$

In practice, randomization rarely used $(\phi(x) = \{0,1\})$

(In theory, simplifies discussions.)

For non-randomized ϕ, the <u>rejection region</u> is

$$R = \{x : \phi(x) = 1\}$$

$$A = \mathcal{X} \setminus R \quad \text{called } \underline{\text{acceptance region}}$$

<u>Power function</u>: $\beta_\phi(\theta) = \mathbb{E}_\theta[\phi(X)]$ rejection prob.

$$= \mathbb{P}_\theta[\text{Reject } H_0]$$

fully summarizes behavior of test.

The <u>significance level</u> of ϕ is $\sup_{\theta \in \Theta_0} \beta_\phi(\theta)$

We say ϕ is a <u>level-α test</u> if its significance level is $\leq \alpha \in [0,1]$ (Sorry)

Ubiquitous choice is $\alpha = 0.05$

["Most influential offhand remark in history of science"]

<u>Goal</u>: Maximize $\beta_\phi(\theta)$ on Θ_1, subject to level-α constraint

Examples

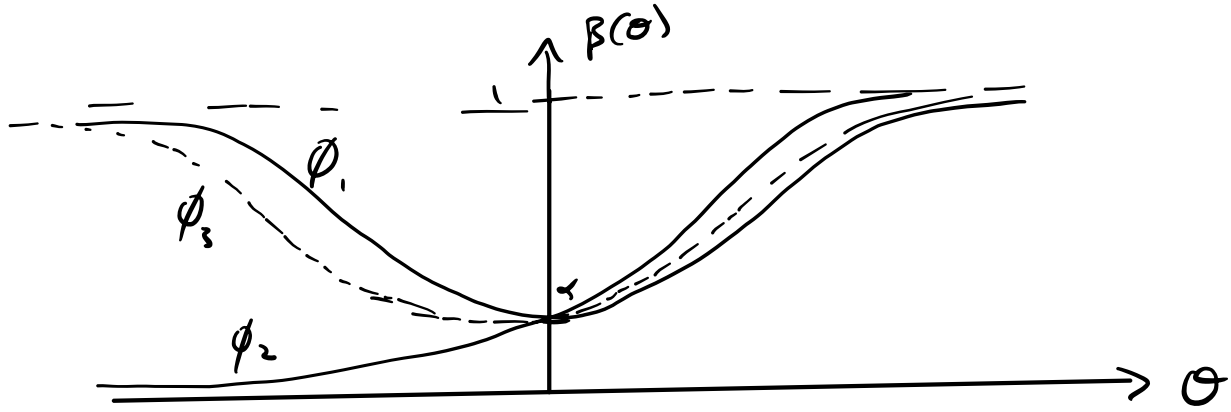
Ex $X \sim N(\theta, 1)$ $\quad H_0: \theta = 0 \quad H_1: \theta \neq 0$

Let $z_\alpha = \Phi^{-1}(1-\alpha)$, Φ = normal cdf.

$\phi_2(x) = 1\{|x| > z_{\alpha/2}\}$ (2-sided test)

$\phi_1(x) = 1\{x > z_\alpha\}$ (1-sided test)

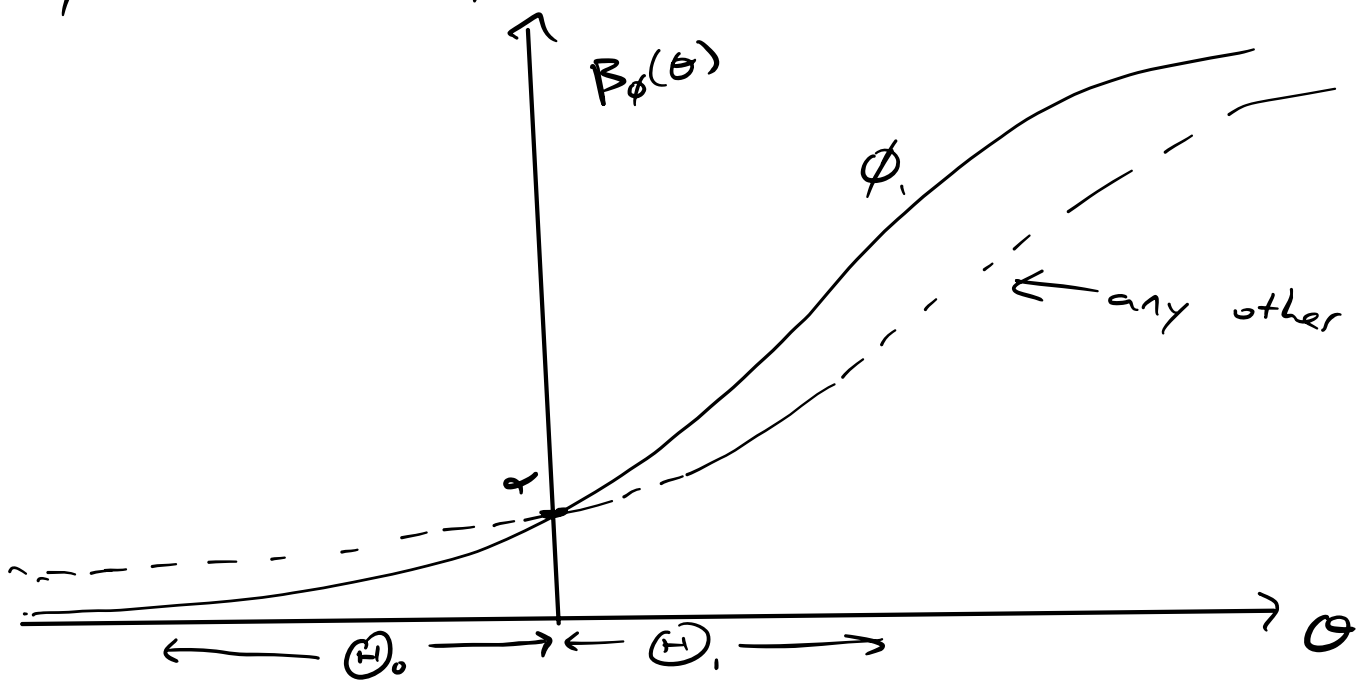
$\phi_3(x) = 1\{x < z_{\alpha/3} \text{ or } x > z_{2\alpha/3}\}$



Sometimes a unique best test exists:

Ex: $X \sim N(\theta, 1)$ $\quad H_0: \theta \leq 0 \quad H_1: \theta > 0$

ϕ_1 is best possible level-α test.



Likelihood Ratio Test

A <u>simple</u> hypothesis is a single distribution:

$$\Theta_0 = \{\theta_0\} \quad \text{or} \quad \Theta_1 = \{\theta_1\}$$

When null/alt. both simple, there exists a unique best test which <u>rejects for large values of</u> the likelihood ratio:

$$\phi^*(x) = \begin{cases} 1 & \dfrac{p_1(x)}{p_0(x)} > c \\ \gamma & \dfrac{p_1(x)}{p_0(x)} = c \\ 0 & \dfrac{p_1(x)}{p_0(x)} < c \end{cases}$$

where p_1, p_0 are null/alt. densities.

(note dominating measure always exists, e.g. $P_0 + P_1$)

and c, γ chosen to make $\mathbb{E}_0 \phi^*(x) = \alpha$

ϕ^* is called the <u>likelihood ratio test</u> (LRT)

<u>Intuition:</u> (discrete case)

Power under H_1: $\int_{\mathbb{R}} p_1(x)\,d\mu(x)$ **BANG**

Sig. level: $\int_{\mathbb{R}} p_0(x)\,d\mu(x)$ **BUCK**

[Analogy: \$100 to buy as much flour as possible]

Neyman-Pearson

Proposition (Keener 12.1)

Suppose $c \geq 0$, ϕ^* maximizes Lagrangian:

$$\mathbb{E}_1[\phi(X)] - c\,\mathbb{E}_0[\phi(X)]$$

among all critical functions.

If $\mathbb{E}_0[\phi^*(X)] = \alpha$ then ϕ^* maximizes

$\mathbb{E}_1[\phi(X)]$ among all level-α tests ϕ

(Essentially, Lagrange multipliers work even though $\mathcal{X}$ infinite/cts./etc.)

Proof Suppose $\mathbb{E}_0[\phi(X)] \leq \alpha$. Then

$$\mathbb{E}_1[\phi(X)] \leq \mathbb{E}_1[\phi(X)] + c(\alpha - \mathbb{E}_0[\phi(X)])$$

$$\leq \mathbb{E}_1[\phi^*(X)] - c\,\mathbb{E}_1[\phi^*(X)] + c\alpha$$

$$= \mathbb{E}_1[\phi^*(X)]$$

$\boxtimes$

Theorem (Neyman-Pearson Lemma)

LRT with significance level α is optimal for testing $H_0: X \sim p_0$ vs. $H_1: X \sim p_1$

Proof Maximize the Lagrangian:

$$\mathbb{E}_1[\phi(X)] - c\,\mathbb{E}_0[\phi(X)]$$

$$= \int_{\mathcal{X}} (p_1(x) - c p_0(x))\,\phi(x)\,d\mu(x)$$

$$= \int_{p_1 > c p_0} |p_1 - c p_0|\,\phi\,d\mu - \int_{p_1 < c p_0} |p_1 - c p_0|\,\phi\,d\mu$$

Max. first term $\Rightarrow$ $\phi(X) = 1$ when $p_1(x) > c p_0(x)$

Min. second term $\Rightarrow$ $\phi(X) = 0$ when $p_1(x) < c p_0(x)$

Choose minimum c s.t.

$$\mathbb{P}_0\left[\frac{p_1(X)}{p_0(X)} > c\right] \le \alpha \le \mathbb{P}_0\left[\frac{p_1(X)}{p_0(X)} \ge c\right]$$

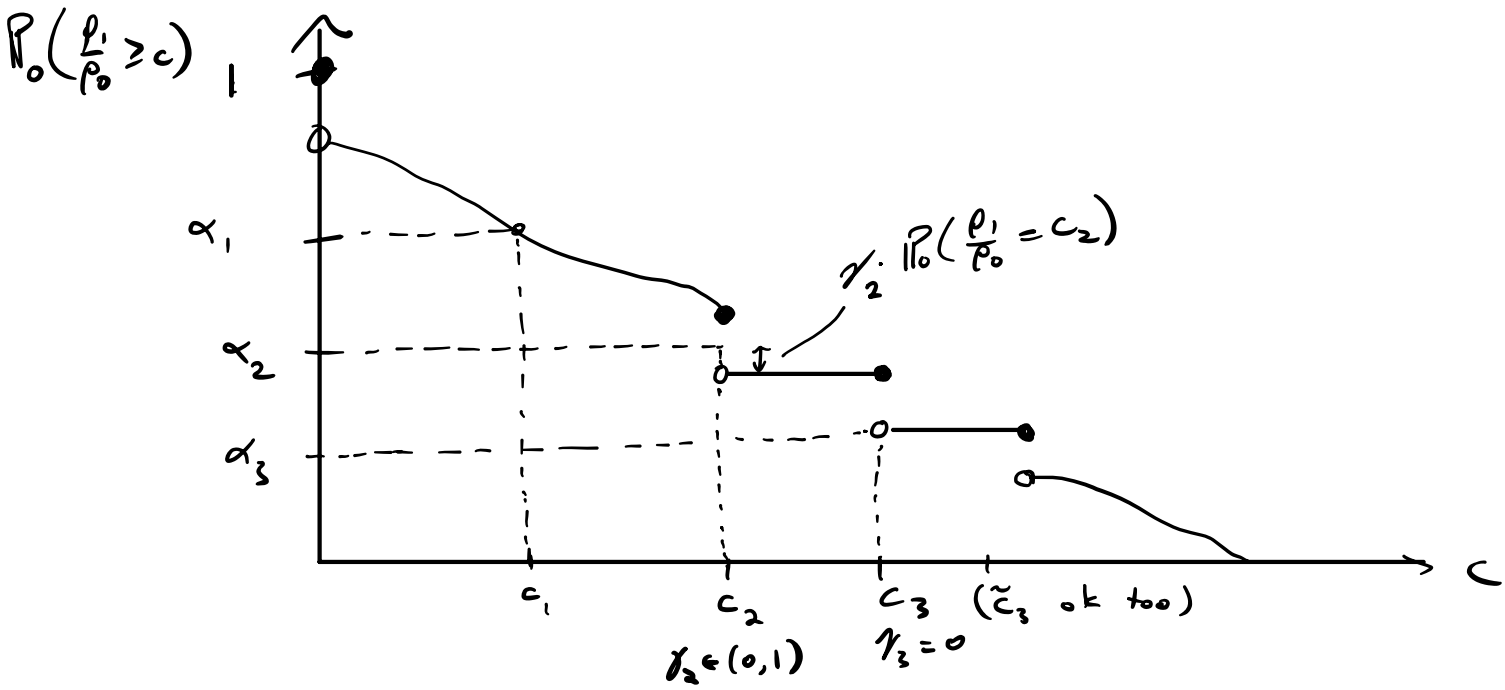
And choose γ to "top up" sig. level:

$$\mathbb{P}_0\left[\frac{p_1}{p_0} > c\right] + \gamma\,\mathbb{P}_0\left[\frac{p_1}{p_0} = c\right] = \alpha$$

(Apply Prop.) $\boxtimes$

Keener gives converse up to wiggle room if $\frac{p_1(X)}{p_0(X)} = c$ for more than 1 value of X.

Choosing c_α, γ_α for ϕ^*



<u>Keener Cor 12.4</u>: If $p_0 \overset{a.s.}{\neq} p_1$ and ϕ is LRT with level $\alpha \in (0,1)$ then $\mathbb{E}_1 \phi(X) > \alpha$

<u>Proof</u> $\mu(\{p_1 > p_0\})$, $\mu(\{p_1 < p_0\})$ both > 0

<u>$c \geq 1$:</u>

$$\underbrace{\mathbb{E}_1 \phi^* - \mathbb{E}_0 \phi^*}_{\text{Power} - \alpha} = \int_{\frac{p_1}{p_0} > 1} |p_1 - p_0| \phi^* d\mu - \cancelto{0}{\int_{\frac{p_1}{p_0} \leq 1} |p_1 - p_0| \phi^* d\mu}$$

$$> 0$$

<u>$c \leq 1$:</u>

$$\underbrace{\mathbb{E}_1(1-\phi^*) - \mathbb{E}_0(1-\phi^*)}_{1 - \text{Power} - (1-\alpha)} = -\int_{\frac{p_1}{p_0} < 1} |p_1 - p_0| (1-\phi^*) d\mu$$

$$< 0$$

Ex 1-param. exp. fam.

$$X \sim p_\eta(x) = e^{\eta T(x) - A(\eta)} h(x)$$

$$H_0: \eta = \eta_0 \quad \text{vs.} \quad H_1: \eta = \eta_1 > \eta_0$$

$$\frac{p_1(x)}{p_0(x)} = \frac{e^{\eta_1 T(x) - A(\eta_1)}}{e^{\eta_0 T(x) - A(\eta_0)}}$$

$$= e^{(\eta_1 - \eta_0) T(x) - (A(\eta_1) - A(\eta_0))}$$

$\phi^*(x)$ rejects for large $T(x)$:

$$\phi^*(x) = \begin{cases} 0 & T(x) < c \\ \gamma & T(x) = c \\ 1 & T(x) > c \end{cases}$$

choose c, γ to make

$$\mathbb{P}_{\eta_0}(T(x) > c) + \gamma \, \mathbb{P}_{\eta_0}(T(x) = c) = \alpha$$

Ex $X_1, \dots, X_n \overset{iid}{\sim} p_\eta(x)$, same H_0, H_1:
Reject for large $\sum_{i=1}^n T(X_i)$

Surprise: $\phi^*(x)$ depends only on η_0 and $\operatorname{sgn}(\eta_1 - \eta_0)$, <u>not</u> on η_1.

Next topic: Uniformly most powerful (UMP) tests