Outline

10/10/2023

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

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

<u>Ex</u> $X_1, \ldots, X_n \sim P$ $\quad Y_1, \ldots, Y_m \sim Q$ $\quad H_0 : P = Q$ 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.)

A non-randomized test partitions $\mathcal{X}$ into

$R = \{x : \phi(x) = 1\}$ <u>rejection region</u>

$A = \{x : \phi(x) = 0\}$ <u>acceptance region</u>

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

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

fully summarizes test's behavior

ϕ is a <u>level-α test</u> $(\alpha \in [0,1])$ if $\sup_{\theta \in \Theta_0} \beta_\phi(\theta) \leq \alpha$

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

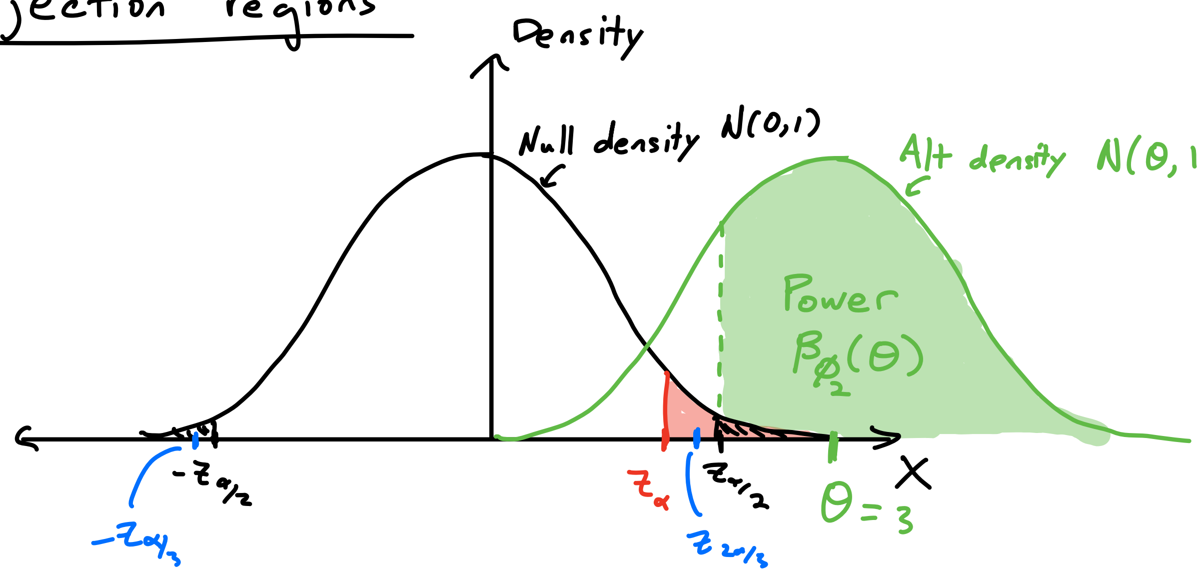
Let $\quad z_\alpha = \Phi^{-1}(1-\alpha)$, $\quad \Phi =$ normal cdf.

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

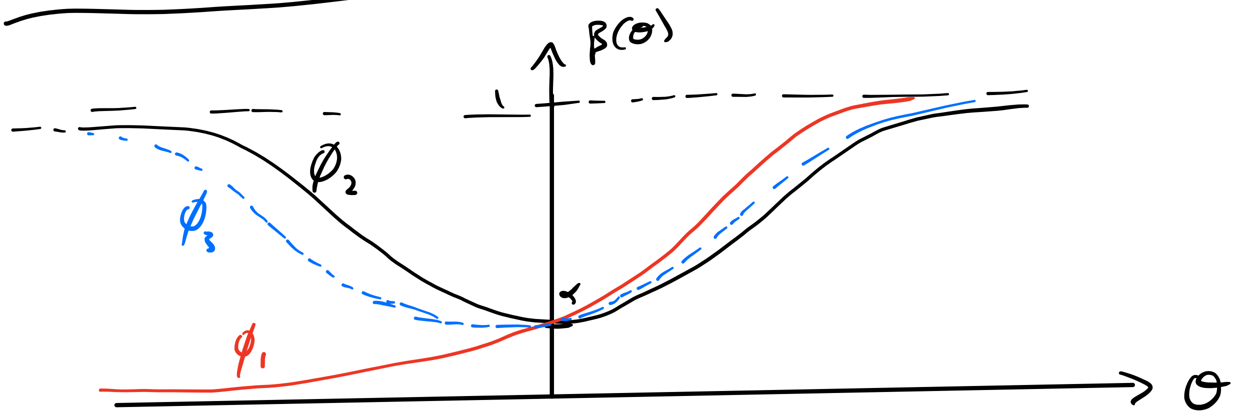
$\phi_1(x) = 1\{x > z_\alpha\}$ (1-sided test)

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

Rejection regions



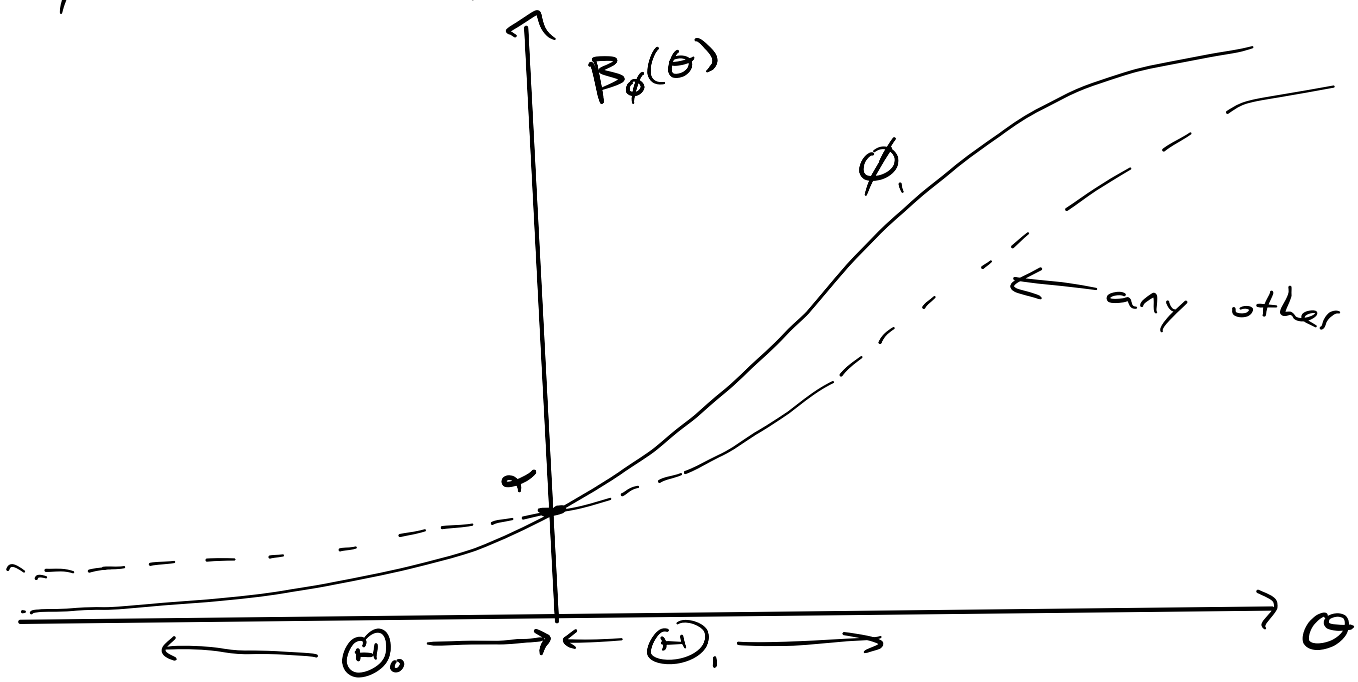
Power functions:



Sometimes a unique best test exists:

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

ϕ_1 is best possible level-α test.



Likelihood Ratio Test

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

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

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

Let $LR(x) = p_1(x)/p_0(x)$, where p_1, p_0 are densities

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

Likelihood ratio test (LRT):

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

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

Intuition: (discrete case)

Power under H_1: $\int_R p_1(x)\,d\mu(x)$ BANG

Sig. level: $\int_R p_0(x)\,d\mu(x)$ BUCK

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

Neyman-Pearson

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 We are interested in maximization problem

$$\underset{\phi : \mathcal{X} \to [0,1]}{\text{maximize}} \quad \mathbb{E}_1[\phi(X)] \quad \text{s.t.} \quad \mathbb{E}_0[\phi(X)] \le \alpha$$

Lagrange form:

$$\text{maximize} \quad \mathbb{E}_1[\phi(x)] - \lambda \mathbb{E}_0[\phi(X)]$$

$$= \int \phi(x)\left(p_1(x) - \lambda p_0(x)\right) d\mu(x)$$

$$= \int \phi(x)\left(\frac{p_1(x)}{p_0(x)} - \lambda\right) dP_0(x)$$

Solution(s):

$$\phi(x) = \begin{cases} 1 & \text{if } LR > \lambda \\ 0 & \text{if } LR < \lambda \\ \text{arbitrary} & \text{if } LR = \lambda \end{cases}$$

$\Rightarrow \phi^*$ maximizes Lagrangian for $\lambda = c$

Consider any other test $\tilde{\phi}(x)$, $\mathbb{E}_0 \tilde{\phi}(x) \le \alpha$

$$\begin{aligned} \mathbb{E}_1 \tilde{\phi} &\le \mathbb{E}_1 \tilde{\phi} - c\mathbb{E}_0 \tilde{\phi} + c\alpha && c(\alpha - \mathbb{E}_0 \tilde{\phi}) \ge 0 \\ &\le \mathbb{E}_1 \phi^* + c\mathbb{E}_0 \phi^* + c\alpha && \phi^* \text{ maxes Lagrangian} \\ &\le \mathbb{E}_1 \phi^* && c(\alpha - \mathbb{E}_0 \phi^*) \ge 0 \end{aligned}$$

Choosing threshold

Note c, γ are not really two "free parameters"

We need both to solve one equation:

$$\alpha = \mathbb{E}_0 \phi^*(x)$$
$$= \mathbb{P}_0(LR > c) + \gamma \mathbb{P}_0(LR = c)$$

Case 1: $LR(x)$ continuous

γ irrelevant, set $c =$ upper α quantile of LR

Case 2: $LR(x)$ discrete

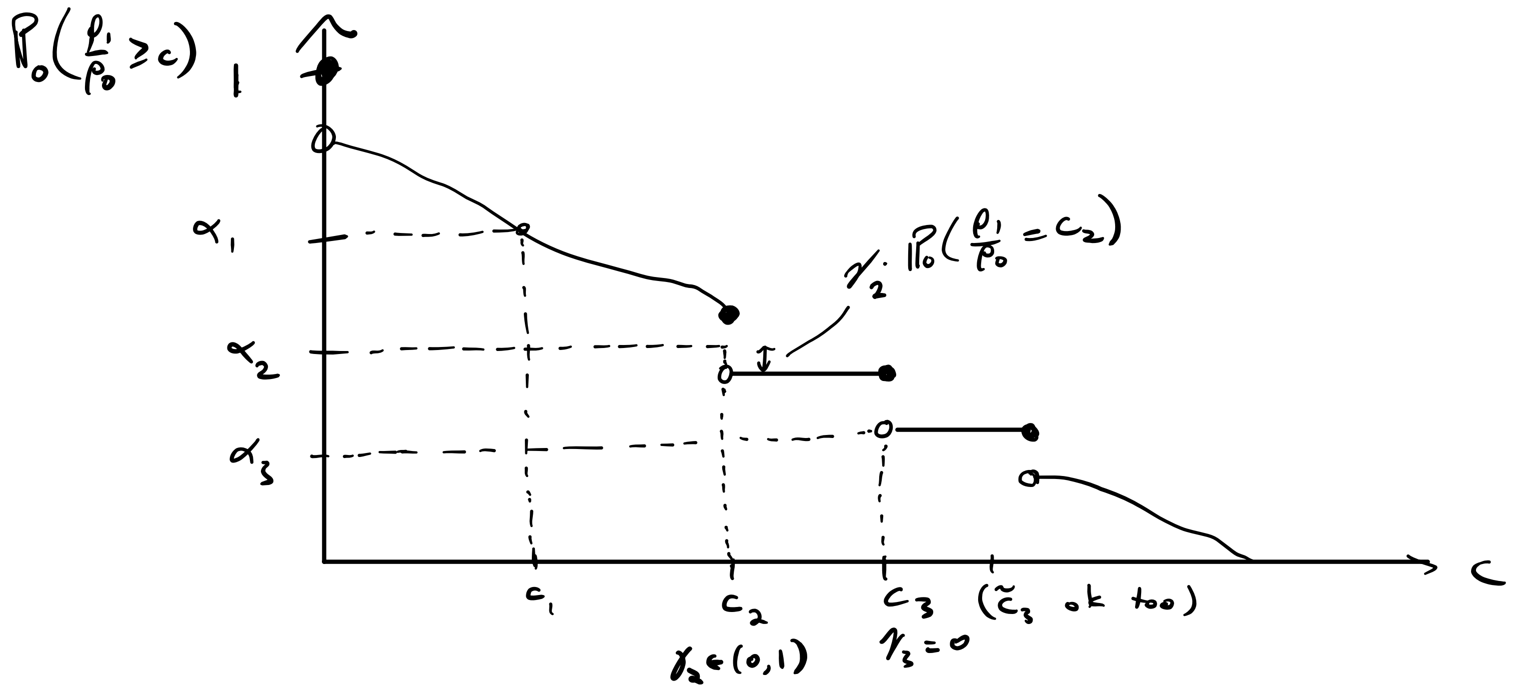
$\mathbb{P}_0(LR > c)$ jumps down at discrete values $LR(x)$

Set $c = \max\{t : \mathbb{P}_0(LR \geq t) \geq \alpha\}$

Then $\mathbb{P}_0(LR > c) \leq \alpha$

$$\gamma = (\alpha - \mathbb{P}_0(LR > c)) / \mathbb{P}_0(LR = c)$$

General case:



Ex 1-param. exp. fam.

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

$H_0: \eta = \eta_0$ vs. $H_1: \eta = \eta_1 > \eta_0$

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

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

$\phi^*(x)$ 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 $X_1, \ldots, X_n \overset{iid}{\sim} p_\eta(x)$, same H_0, H_1:

Reject for large $\sum_{i=1}^n T(X_i)$

Important: $\phi^*(X)$ depends only on η_0 and $\text{sgn}(\eta_1 - \eta_0)$, not on η_1.

Next topic: Uniformly most powerful (UMP) tests