Lecture 20: p-values and Neyman-Pearson test

Ciaran Evans

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.

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

threshold

$$\overbrace{\sup_{\theta \in \Theta_0} \beta(\theta)}^{\text{worst case type I error}} = \sup_{\theta \in \Theta_0} P_\theta(T(X_1, ..., X_n) > c_\alpha) = \alpha$$

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

$$p = \underbrace{\sup_{\theta \in \Theta_0} P_\theta}_{\text{probability under } H_0}(T(X_1, ..., X_n) > \underbrace{T(x_1, ..., x_n)}_{\text{observed test statistic}})$$

Proof of theorem

$$p = \inf\{\alpha : \text{reject } H_0\} = \inf\{\alpha : T(x_1, \dots, x_n) > c_\alpha\}$$

As $\alpha \downarrow$, $c_\alpha \uparrow$

cut off corresponding to p-value

$$c_p = \sup\{c_\alpha : T(x_1, \dots, x_n) > c_\alpha\} = T(x_1, \dots, x_n)$$

and by definition of c_α

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

$$\Rightarrow \sup_{\theta \in \Theta_0} P_\theta(T(X_1, \dots, X_n) > c_p) = p$$

$T(x_1, \dots, x_n)$

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

//

(probability of "our data or more extreme")

Recap: hypothesis testing and power function

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

Given observed data $X_1, ..., X_n$:

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

The **power function** $\beta(\theta)$ is

$$\beta(\theta) = P_\theta((X_1, ..., X_n) \in \mathcal{R})$$

Goal: maximize power for $\theta \in \Theta_1$, subject to control of power for $\theta \in \Theta_0$

Wald test for normal mean

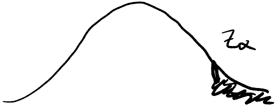
Let $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ with σ^2 known. We wish to test

$$H_0 : \mu = \mu_0 \qquad H_A : \mu = \mu_1$$

where $\mu_1 > \mu_0$. The Wald test rejects if

look in the right tail

$$\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} > z_\alpha$$



$$\Rightarrow \quad \bar{X}_n > \underbrace{\mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha}_{c_0}$$

Wald test for normal mean

Let $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ with σ^2 known. We wish to test

$$H_0 : \mu = \mu_0 \qquad H_A : \mu = \mu_1$$

where $\mu_1 > \mu_0$. The Wald test rejects if

$$\overline{X}_n > \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha$$

We know that $\beta(\mu_0) = \alpha$ for this test.

Our question: Is there a *better* test for these hypotheses?

- To answer this question, we will need to introduce the *Neyman-Pearson* test

Rearranging

$H_0: \mu = \mu_0 \qquad H_A: \mu = \mu_1 \qquad , \mu_1 > \mu_0$

reject H_0 when $\bar{X}_n > c_0$. Equivalent to

$$\frac{L(\mu_1)}{L(\mu_0)} = \frac{f(X_1, \ldots, X_n \mid \mu_1)}{f(X_1, \ldots, X_n \mid \mu_0)} = \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_i (X_i - \mu_1)^2\right\}}{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_i (X_i - \mu_0)^2\right\}} > K_0$$

(for some K_0)

$$\bar{X}_n > c_0 \iff 2n\bar{X}(\mu_0 - \mu_1) < 2nc_0(\mu_0 - \mu_1)$$

$$\iff \sum_i X_i^2 - \sum_i X_i^2 + 2n\bar{X}(\mu_0 - \mu_1) - n\mu_0^2 + n\mu_1^2 < 2nc_0(\mu_0 - \mu_1) - n\mu_0^2 + n\mu_1^2$$

$$\iff -\frac{1}{2\sigma^2}\sum_i (X_i - \mu_1)^2 + \frac{1}{2\sigma^2}\sum_i (X_i - \mu_0)^2 > -\frac{1}{2\sigma^2}\left(2nc_0(\mu_0 - \mu_1) - n\mu_0^2 + n\mu_1^2\right)$$

$$\iff \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_i (X_i - \mu_1)^2\right\}}{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_i (X_i - \mu_0)^2\right\}} > \underbrace{\exp\left\{\quad \text{"} \quad \text{"}\quad\right\}}_{K_0}$$

Rearranging

Let $\mathbf{X} = X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ with σ^2 known. We wish to test

$$H_0 : \mu = \mu_0 \qquad H_A : \mu = \mu_1$$

where $\mu_1 > \mu_0$.

chose so that $\beta(\mu_0) = \alpha$

The Wald test rejects if $\overline{X}_n > c_0$, which is equivalent to rejecting when

$$\frac{L(\mu_1|\mathbf{X})}{L(\mu_0|\mathbf{X})} = \frac{f(X_1, ..., X_n|\mu_1)}{f(X_1, ..., X_n|\mu_0)} > k_0$$

chosen so that $\beta(\mu_0) = \alpha$

Neyman-Pearson test

Let $X_1, ..., X_n$ be a sample from some distribution with probability function f and parameter θ. To test

$$H_0 : \theta = \theta_0 \qquad H_A : \theta = \theta_1,$$

the **Neyman-Pearson** test rejects H_0 when

$$\frac{L(\theta_1|X)}{L(\theta_0|X)} = \frac{f(X_1, ..., X_n|\theta_1)}{f(X_1, ..., X_n|\theta_0)} > k$$

where k is chosen so that $\beta(\theta_0) = \alpha$.

Example

Let $X_1, ..., X_n \overset{iid}{\sim} Exponential(\theta)$, with pdf $f(x|\theta) = \theta e^{-\theta x}$. We want to test

$$H_0 : \theta = \theta_0 \qquad H_A : \theta = \theta_1,$$

where $\theta_1 < \theta_0$. The Neyman-Pearson test rejects when

$$\frac{L(\theta_1|\mathbf{X})}{L(\theta_0|\mathbf{X})} > k.$$

1. Calculate $\dfrac{L(\theta_1|\mathbf{X})}{L(\theta_0|\mathbf{X})}$
2. Rearrange the ratio to show that $\dfrac{L(\theta_1|\mathbf{X})}{L(\theta_0|\mathbf{X})} > k$ if and only if $\sum_i X_i > c$ for some c
3. Using the fact that $\sum_i X_i \sim Gamma(n, \theta)$, find c such that $\beta(\theta_0) = \alpha$

$$\frac{L(\theta_1 \mid X)}{L(\theta_0 \mid X)} = \frac{\theta_1^n e^{-\theta_1 \sum_i x_i}}{\theta_0^n e^{-\theta_0 \sum_i x_i}} > k \qquad f(x) = \theta e^{-\theta x}$$

$$L(\theta_1 \mid X) = f(x \mid \theta_1) = \prod_i \theta_1 e^{-\theta_1 x_i} = \theta_1^n e^{-\theta_1 \sum_i X_i}$$

$$\left(\frac{\theta_1}{\theta_0}\right)^n \exp\left\{ (\theta_0 - \theta_1) \sum_i X_i \right\} > k$$

$\sum_i X_i \sim$ Gamma$(n, \theta_0) \Rightarrow$

$$\exp\left\{ (\theta_0 - \theta_1) \underbrace{\sum_i X_i} \right\} > \left(\frac{\theta_0}{\theta_1}\right)^n k$$

$$X_i \sim \text{Exp}(\theta) = \text{Gamma}(1, \theta)$$

$$\sum_i X_i \sim \text{Gamma}(n, \theta)$$

$$\Rightarrow \quad (\theta_0 - \theta_1) \sum_i X_i > \log k + n \log\left(\frac{\theta_0}{\theta_1}\right)$$

$$\Rightarrow \quad \sum_i X_i > \frac{\log k + n \log\left(\frac{\theta_0}{\theta_1}\right)}{\theta_0 - \theta_1} = c$$

reject H_0 when $\sum_i X_i > c$

Under H_0, $\sum_i X_i \sim \text{Gamma}(n, \theta_0)$ $\qquad \beta(\theta_0) = \alpha$ if $c =$ upper α quantile of Gamma(n, θ_0)