Lec 2/20

Monday, February 20, 2017 14:54

Thm 12.1 (Neyman-Pearson Lemma)

type-1 error rate is α.

If C is a critical region of size α and K is a constant s.t.

$C = \{\vec{x} : \frac{L_0(\vec{x})}{L_1(\vec{x})} \le K\}$, Then C is a most powerful critical region

of size α for testing $\theta = \theta_0$ vs $\theta = \theta_1$. (simple-vs-simple).

Note: Compute size by $\alpha = \mathbb{P}(\text{reject } H_0 ; H_0 \text{ is true})$

Power $= 1 - \beta = \mathbb{P}(\text{reject } H_0 ; H_0 \text{ is false})$

Ex: normal population:

(Critical region determines a test).

$H_0 : \mu = \mu_0$ vs $H_1 : \mu = \mu_1$, $\mu_1 > \mu_0$

Data: a RS $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, 1)$.

Use NP lemma to find most powerful CR of size α.

Sol: $L_0(\vec{x}) = \prod_{i=1}^{n} f(x_i ; \mu_0) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \mu_0)^2}$

$L_1(\vec{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \mu_1)^2\right)$

So $L_0/L_1 = \exp\left(-\frac{1}{2}\sum(x_i - \mu_0)^2 + \frac{1}{2}\sum(x_i - \mu_1)^2\right)$.

$= \exp\left(\sum_{i=1}^{n} x_i(\mu_0 - \mu_1) - \frac{n}{2}(\mu_0^2 - \mu_1^2)\right)$

NP Lemma critical region: find constant K st.

$$C = \left\{\vec{x} : \exp\left(\sum_{i=1}^{n} x_i(\mu_0 - \mu_1) - \frac{n}{2}(\mu_0^2 - \mu_1^2)\right) \le K\right\}$$

$$= \left\{\vec{x} : \sum x_i(\mu_0 - \mu_1) \le \log(k) + \frac{n}{2}(\mu_0^2 - \mu_1^2)\right\}$$

$$= \left\{ \vec{x} : \sum x_i(\mu_0 - \mu_1) \le \log(k) + \frac{n}{2}(\mu_0^2 - \mu_1^2) \right\}$$

$$= \left\{ \vec{x} : \sum x_i \ge \frac{\log(k) + \frac{n}{2}(\mu_0^2 - \mu_1^2)}{\mu_0 - \mu_1} \right\}$$

$$= \left\{ \vec{x} : \bar{X} \ge \underbrace{\frac{\log(k) + \frac{n}{2}(\mu_0^2 - \mu_1^2)}{n(\mu_0 - \mu_1)}}_{} \right\} = \{ \vec{x} : \bar{X} \ge \tilde{k} \}.$$

$\tilde{K}$ a constant. finding $\tilde{K}$ equiv to finding K.

so find $\tilde{K}$ s.t. C has size α.

$\bar{X} \sim N(\mu_0, \frac{1}{n})$ under H_0 so

$(\bar{X} - \mu_0)\sqrt{n} \sim N(0,1)$ and $\bar{X} \ge \tilde{K}$ when $(\bar{X} - \mu_0)\sqrt{n} \ge (\tilde{k} - \mu_0)\sqrt{n}$.

so $(\tilde{K} - \mu_0)\sqrt{n} = Z_\alpha \Rightarrow \tilde{K} = \frac{Z_\alpha}{\sqrt{n}} + \mu_0$.

So the test should be:

Reject H_0 if $\bar{X} \ge \mu_0 + \frac{Z_\alpha}{\sqrt{n}}$

Accept H_0 if $\bar{X} < \mu_0 + \frac{Z_\alpha}{\sqrt{n}}$.

people don't like to say this instead: fail to reject H_0.

Note: the test doesn't depend on μ_1, but β does.

this test always minimizes β for any μ_1 though.

Ex: let $X \sim \text{Bin}(2, \theta)$. $\mathbb{P}(X = x; \theta) = \binom{2}{x}\theta^x(1-\theta)^{2-x}$ $\quad x = 0, 1, 2$

Test: $H_0: \theta = \frac{1}{2}$ vs $H_1: \theta = \frac{3}{4}$.

Sol. $L_0(x) = \binom{2}{x}\left(\frac{1}{2}\right)^2$

$L_1(x) = \binom{2}{x}\left(\frac{3}{4}\right)^x\left(\frac{1}{4}\right)^{2-x}$

$$L_0 = \begin{cases} 0.25 & x=0 \\ 0.5 & x=1 \\ 0.25 & x=2 \end{cases}$$

$$L_1 = \begin{cases} 0.0625 & x=0 \\ 0.375 & x=1 \\ 0.5625 & x=2 \end{cases}$$

$$\frac{L_0}{L_1} = \begin{cases} \frac{0.25}{0.0625} = 4 & x=0 \\ \frac{0.5}{0.375} = \frac{4}{3} & x=1 \\ \frac{0.25}{0.5625} = \frac{4}{9} & x=2 \end{cases}$$

NP lemma: find K s.t. $\frac{L_0}{L_1} \le K$ inside C w/ size $\le \alpha$.

HW: give K s.t. $\alpha \le 0.25$. and s.t. $\alpha \le 0.5$