TA section 8

JERRY C.

jerryc520.github.io/teach/MS.html

May 28, 2024

Review: Hypothesis Testing

- Given $\theta \in \Theta \subseteq \Omega = \Omega_0 \cup \Omega_1$ (parametric space). Define
  $H_o : \theta \in \Omega_0$ (null space) vs.
  $H_1 : \theta \in \Omega_1 := \Omega \setminus \Omega_0$ (alternative space), $\Omega_0 \cap \Omega_1 = \phi$.
- $C := \{\boldsymbol{X} : T(\boldsymbol{X}) \geq d | H_o\}$: rejection region. We say, we reject H_o iff $T(\boldsymbol{X}) \geq d$ for some $d > 0$ (落在拒絕域).
- testing rule (decision rule):

$$\phi(\boldsymbol{X}) = \boldsymbol{I}(T(\boldsymbol{X}) \in C),$$

  is a testing function.
- We want to control the probabilities of two errors (risks): for $\alpha, \beta \in [0, 1]$,

$$\alpha := \mathbb{P}(\text{type I error}) = \mathbb{P}(\text{reject } H_o \mid H_o \text{ is true}) = \mathbb{E}[\phi(\boldsymbol{X}) | H_o];$$

$$\beta := \mathbb{P}(\text{type II error}) = \mathbb{P}(\text{NOT reject } H_o \mid H_0 \text{ is false (or } H_1 \text{ is true)}).$$

- $\alpha \uparrow (\downarrow) \Rightarrow \beta \downarrow (\uparrow)$; $\alpha + \beta \neq 1$;
- 犯 type I error 較嚴重 (本身無罪卻被判有罪), 對犯 type II error 較容忍 (本身有罪卻沒被判有罪); 即: $\mathbb{P}$(type I error) 應優先控制!
- *Do not reject H_o* is preferred over *Accept H_o* (why ?);
- *Accept H_o* is at a risk of a type II error (沒控制犯 type II error 的機率).

Decision rule

- There are two possible decisions:
  Conclude that there is enough evidence to reject H_o (support H_1 is true);
  Conclude that there is not enough evidence to reject H_o.

Types of Hypotheses

- " $=$ " 放在 null;
- composite hypothesis: $H_o : \theta \in \Omega_o$ vs. $H_1 : \theta \in \Omega_1$;
- simple hypothesis: $H_o : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$, 其中 $\theta_0, \theta_1 \in \{$ singleton $\}$;
- two-sided (two-tailed): $H_o : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$;
- left-sided (left-tailed): $H_o : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$;
- right-sided (right-sided): $H_o : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$.

Size vs. Power

- power function $K(\theta) := \mathbb{P}(\text{reject } H_o \, |\theta) = \mathbb{P}_\theta(T(\boldsymbol{X}) \in C)$;
- $K(\theta)$ is an increasing function of θ, $\lim_{\theta \to -\infty} K(\theta) = 0$ and $\lim_{\theta \to \infty} K(\theta) = 1$;
- size α test: $\alpha = \sup_{\theta \in \Omega_0} K(\theta)$;
- level α test: $\alpha \geq \sup_{\theta \in \Omega_0} K(\theta)$; α is called the significant level;
- power of a test: $1 - \beta := K(\theta \in \Omega_1) = \mathbb{P}(\text{reject } H_o \, |H_1 \text{ is true})$;
- Consistent Test: If a test with the sequence of power functions $\{K_n(\theta)\}$, such that, for any fixed $\theta \in \Omega_1$, $\lim\limits_{n \to \infty} K_n(\theta) = 1$.
- Unbiased Test: If a test with power function $K(\theta)$, such that for every $\theta' \in \Omega_1$ and $\theta'' \in \Omega_0$, $K(\theta') \geq K(\theta'')$.

P-value

- p-value is a test statistic $p(\boldsymbol{X})$, such that $p(\boldsymbol{x}) \in [0,1]$ for any $\boldsymbol{X} = \boldsymbol{x}$, if for every $\theta \in \Omega_0$ (i.e., under H_o), $\alpha \in [0,1]$,

$$\mathbb{P}_\theta(p(\boldsymbol{X}) \leq \alpha) \leq \alpha,$$

(對應之 type I error 發生機率要小於 α), then we say $p(\boldsymbol{X})$ is valid.

- A test rejecting H_o is a level α test if and only if $p(\boldsymbol{X}) \leq \alpha$ (決定出要 reject H_o 之最小的顯著水準).

• Theorem

Let $T(\boldsymbol{X})$ be a testing statistic, with the rejection region $C = \{\boldsymbol{X} : T(\boldsymbol{X}) \geq d | H_o : \theta \in \Omega_0\}$. Then, for any $\boldsymbol{X} = \boldsymbol{x}$, define

$$p(\boldsymbol{x}) := \sup_{\theta \in \Omega_0} \mathbb{P}_\theta(T(\boldsymbol{X}) \geq T(\boldsymbol{x})),$$

then $p(\boldsymbol{X})$ is a valid p-value.

Uniformly Most Powerful (UMP) test

- simple hypothesis: Neyman-Pearson lemma $\Rightarrow$ MP test
- composite hypothesis: monotone likelihood ratio (MLR) family $\Rightarrow$ UMP test
- UMP test $\Rightarrow$ MP test

Example 1

- $X_1, X_2 \sim$ i.i.d. $U[\theta, \theta+1]$. Under $H_o : \theta = 0$ vs. $H_1 : \theta = 0.5$, consider two testing rules $\phi_1(X_1) = \boldsymbol{I}(X_1 > 0.95)$ and $\phi_2(X_1, X_2) = \boldsymbol{I}(X_1 + X_2 > k)$, for $k \in [1, 2]$.
- size: $\alpha_1 = \mathbb{E}[\phi_1(X_1)|H_o] = \mathbb{P}(X_1 > 0.95|\theta = 0) = 0.05$, and

$$\alpha_2 = \mathbb{P}(X_1 + X_2 > k|\theta = 0) = \int_{1-k}^{1}\int_{k-x_1}^{1} 1 dx_2 dx_1 = (2-k)^2/2,$$

  so, $k^* = 2 - \sqrt{2\alpha_2}$.
- If $\alpha_1 = \alpha_2$, $k^* = 2 - \sqrt{(0.1)} \approx 1.68$.
- power of ϕ_1:

$$K_1(\theta) = \mathbb{P}_\theta(X_1 > 0.95) = \begin{cases} 0, & [\theta \le -0.05] \\ \theta + 0.05, & [-0.05 < \theta \le 0.95] \\ 1, & [0.95 < \theta]. \end{cases}$$

  So, power of ϕ_1: $K_1(\theta = 0.5) = 0.55$.

- let $Y = X_1 + X_2$, where $X_i \sim U[0.5, 1.5]$ under H_1,
  let $Z = X_1 \Rightarrow X_2 = Y - Z$, Jacobian $|J| = |\partial(X_1, X_2)/\partial(Z, Y)| = 1$,

$$f_{X_1,X_2}(x_1, x_2) = 1,\ 0.5 \leq x_1, x_2 \leq 1.5$$

$$\Rightarrow f_{Y,Z}(y, z) = 1,\ 0.5 \leq z \leq 1.5, 0.5 \leq y - z \leq 1.5$$

$$\Rightarrow f_Y(y) = \int_{\max(0.5,y-1.5)}^{\min(1.5,y-0.5)} 1\, dz = \begin{cases} y-1, & y \in [1,2] \\ 3-y, & y \in (2,3] \\ 0, & o.w. \end{cases}$$

  So, power of ϕ_2: for $k \in [1, 2]$,
  $K_2(\theta = 0.5) = \mathbb{P}(Y > k|\theta = 0.5) = \int_k^2 (y-1)dy + \int_2^3 (3-y)dy =$
  $0.5 - k^2/2 + k$.
- Take $k^* = 1.68$, $K_2(\theta = 0.5) = 0.7688$.
- 即, 當考慮兩個檢定函數 ϕ_1, ϕ_2, 控制相同的 size 之下,ϕ_2 的 power 較大。

Example 2

- $X \sim P(\lambda)$, for testing $H_o : \lambda \leq 1$ vs. $H_1 : \lambda > 1$. Consider one sample observed $X = 3$,
- the p-value of X is:
  $p(\boldsymbol{x}) = \mathbb{P}(X \geq 3|\lambda = 1) = 1 - \mathbb{P}(X < 3|\lambda = 1) =$
  $1 - e^{-1}1^2/2! - e^{-1}1^1/1! - e^{-1}1^0/0! \approx 0.0803.$

Example 3

- 假設 $X_1, \cdots, X_n$, 為服從 $\mathcal{N}(\mu, \sigma^2)$ 分佈之隨機樣本。考慮檢定虛無假設 $H_o : \mu = \mu_0$, 且檢定統計量為 $\bar{X}_n$。令 $z_0 = \sqrt{n}(\bar{x}_0 - \mu_0)/\sigma$, 此時假設 $Z \sim \mathcal{N}(0, 1)$,
- 則可以求得 p-value 為:

$$p(\boldsymbol{x}) = \mathbb{P}(\bar{X}_n \geq \bar{x}_0 | \mu_0) = \begin{cases} 2(1 - \Phi(|z_0|)), & H_1 : \mu \neq \mu_0 \\ 1 - \Phi(z_0), & H_1 : \mu > \mu_0 \\ \Phi(z_0), & H_1 : \mu < \mu_0. \end{cases}$$