§8.2 N-P引理和似然比检验

- 简单假设检验问题: $\Theta = \{\theta_0, \theta_1\}$.

$$H_0 : \theta = \theta_0 \leftrightarrow H_1 : \theta = \theta_1.$$

- 似然函数: $L(\vec{x}, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$. (以连续型为例)
- 似然比否定域/似然比检验:

$$\mathcal{W}_\lambda = \{\vec{x} : L(\vec{x}, \theta_1) > \lambda L(\vec{x}, \theta_0)\}.$$

- 定理2.1. (Neyman-Pearson 引理) 若λ_0 使得

$$P_{\theta_0}\left(\vec{X} \in \mathcal{W}_{\lambda_0}\right) = \alpha,$$

  则$\mathcal{W}_{\lambda_0}$ 是水平为α 的UMP 否定域.
- 定理2.2. 上述$\mathcal{W}_{\lambda_0}$ 是无偏否定域, 因而也是UMPU 否定域.

- N-P 引理的证明: 假设$\tilde{\mathcal{W}}$ 也是水平为α 的否定域, 往验证

$$P_{\theta_1}(\vec{X} \in \mathcal{W}) \geqslant P_{\theta_1}(\vec{X} \in \tilde{\mathcal{W}}), \quad \text{其中 } W_{\lambda_0} \triangleq \mathcal{W}.$$

- 由假设知, $\mathcal{W}$ 与$\tilde{\mathcal{W}}$ 都是水平为α 的否定域, 即

$$P_{\theta_0}(\vec{X} \in \mathcal{W}) = P_{\theta_0}(\vec{X} \in \tilde{\mathcal{W}}) = \alpha.$$

- 左右两边同时扣除交事件$\{\vec{X} \in \mathcal{W} \cap \tilde{\mathcal{W}}\}$ 的概率. 即, 需验证

$$P_{\theta_0}(\vec{X} \in \mathcal{W} \backslash \tilde{\mathcal{W}}) = P_{\theta_0}(\vec{X} \in \tilde{\mathcal{W}} \backslash \mathcal{W})$$

$$\Rightarrow L \triangleq P_{\theta_1}(\vec{X} \in \mathcal{W} \backslash \tilde{\mathcal{W}}) \geqslant P_{\theta_1}(\vec{X} \in \tilde{\mathcal{W}} \backslash \mathcal{W}) \triangleq R.$$

- 由$\mathcal{W} = \mathcal{W}_{\lambda_0} = \{\vec{x} : L(\vec{x}, \theta_1) > \lambda_0 L(\vec{x}, \theta_0)\}$ 知 "$\Rightarrow$" 成立:

$$L = \int_{\mathcal{W} \backslash \tilde{\mathcal{W}}} L(\vec{x}, \theta_1) d\vec{x} \geqslant \lambda_0 \int_{\mathcal{W} \backslash \tilde{\mathcal{W}}} L(\vec{x}, \theta_0) d\vec{x},$$

$$R = \int_{\tilde{\mathcal{W}} \backslash \mathcal{W}} L(\vec{x}, \theta_1) d\vec{x} \leqslant \lambda_0 \int_{\tilde{\mathcal{W}} \backslash \mathcal{W}} L(\vec{x}, \theta_0) d\vec{x}.$$

例2.1. $X \sim N(\mu, 1)$. 求假设检验问题$H_0: \mu = 0 \leftrightarrow H_1: \mu = 2$ 的水平为$\alpha = 0.05$ 的UMP 否定域.

- 似然函数与似然比:

$$\frac{L(\vec{x}, \theta_1)}{L(\vec{x}, \theta_0)} = \frac{\frac{1}{\sqrt{2\pi}^n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-2)^2}}{\frac{1}{\sqrt{2\pi}^n} e^{-\frac{1}{2}\sum_{i=1}^n x_i^2}} = e^{\frac{1}{2}\sum_{i=1}^n 4(x_i-1)}.$$

- 似然比否定域:

$$\mathcal{W}_\lambda = \left\{\vec{x} : \frac{L(\vec{x}, \theta_1)}{L(\vec{x}, \theta_0)} > \lambda\right\} = \{\vec{x} : \bar{x} > c\}.$$

- $T(x_1, \cdots, x_n) = \bar{x}$ 称为检验统计量.
- 根据α 选择λ (等价地, 选择c):

$$\alpha = P_{\theta_0}\left(\bar{X} > c\right) = P(Z > c\sqrt{n}) \Rightarrow c = z_{1-\alpha}/\sqrt{n}.$$

- 查表获得$z_{1-0.05} = 1.65$. 从而所求为

$$\mathcal{W} = \{\vec{x} : \bar{x} > 1.65/\sqrt{n}\}.$$

例2.2. $\Theta = \{0,1\}$. $f(x,0) = 1_{\{0<x<1\}}$, $f(x,1) = 2x1_{\{0<x<1\}}$. 求假设检验问题$H_0: \theta = 0 \leftrightarrow H_1: \theta = 1$ 的UMP 否定域.

- 似然函数与似然比:

$$\frac{L(\vec{x},\theta_1)}{L(\vec{x},\theta_0)} = \frac{2^n x_1 \cdots x_n 1_{\{0<x_1,\cdots,x_n<1\}}}{1_{\{0<x_1,\cdots,x_n<1\}}} = 2^n x_1 \cdots x_n 1_{\{0<x_1,\cdots,x_n<1\}}.$$

- 似然比否定域与检验统计量$T = T(x_1, \cdots, x_n)$:

$$\mathcal{W}_\lambda = \left\{\vec{x} : \frac{L(\vec{x},\theta_1)}{L(\vec{x},\theta_0)} > \lambda\right\} = \left\{\vec{x} : -2\sum_{i=1}^n \ln x_i < c\right\}.$$

- 根据α 选择c: 在H_0 下, $-2\ln X \sim \chi^2(2)$, 于是, $T \sim \chi^2(2n)$.

$$\alpha = P_{\theta_0}\left(-2\sum_{i=1}^n \ln X_i < c\right) \Rightarrow c = \chi^2_\alpha(2n).$$