

Семинар - 10.

теория

Гипотезы. Критерий Неймана-Пирсона

Постановка задачи. Имеется выборка $(X_1, \dots, X_n)$ из распределения $\mathbb{P}_\theta, \theta = \{\theta_0, \theta_1\}$. Построить наиболее мощный критерий уровня значимости $\alpha \in (0, 1)$ для проверки простой гипотезы против простой альтернативы

$$\begin{aligned} H_0 : \theta &= \theta_0; \\ H_1 : \theta &= \theta_1 \quad (\theta_1 > \theta_0). \end{aligned} \quad (1)$$

Определение (Статистика отношения правдоподобий). Статистикой отношения правдоподобий называется статистика

$$\Lambda(\mathbf{x}) = \frac{p_{\theta_1}(\mathbf{x})}{p_{\theta_0}(\mathbf{x})} = \frac{\prod_{i=1}^n p_{\theta_1}(x_i)}{\prod_{i=1}^n p_{\theta_0}(x_i)}.$$

Введем функцию

$$\varphi(t) = \mathbb{P}_0(\Lambda(\mathbf{X}) \geq t) \quad (2)$$

и предположим, что она непрерывна.

Теорема 1 (Нейман, Пирсон). Пусть дана выборка $\mathbf{X} = (X_1, \dots, X_n)$ из распределения $\mathbb{P}_\theta, \theta = \{\theta_0, \theta_1\}$. Пусть $p_{\theta_0}, p_{\theta_1} > 0$. Допустим, что функция φ , определенная в (2) непрерывна. Зафиксируем произвольное $\alpha \in (0, 1)$ и найдем $t_\alpha \geq 0$, удовлетворяющее равенству $\varphi(t_\alpha) = \alpha$ для φ , определенной в (2). Тогда в задаче (1) статистический критерий δ^* с критической областью

$$\mathcal{X}_1^* = \{\mathbf{x} \in \mathbb{R}^n \mid \Lambda(\mathbf{x}) \geq t_\alpha\} \quad (3)$$

является наиболее мощным среди всех критериев уровня значимости α .

Теорема 2 (О монотонном отношении правдоподобий). Допустим, что проверяется простая гипотеза $H_0: \theta = \theta_0$ против односторонней альтернативы $H_1: \theta > \theta_0$. Пусть статистика отношения правдоподобия является монотонной функцией от достаточной статистики $T(\mathbf{X})$, то есть такой, что правдоподобие представляется в виде $L_{\mathbf{X}}(\theta) = g(T(\mathbf{X}), \theta)h(\mathbf{X})$, где g и h — некоторые функции. Тогда в задаче (H_0, H_1) существует равномерно наиболее мощный критерий, совпадающий с критерием Неймана-Пирсона для простой гипотезы и простой альтернативы.

Примечание. В условиях Теоремы (2) гипотезы со знаками $(=, >)$ могут быть заменены на гипотезы $(=, <)$, $(\geq, <)$ или $(\leq, >)$; критерий останется таким же. В случае монотонного убывания отношения правдоподобия знак в критерии меняется на противоположный.

еще теория

Задачи:

1. Пусть дана выборка $X_1, \dots, X_n$ из нормального распределения $\mathcal{N}(\theta, \sigma^2)$, σ^2 известна. Построить наиболее мощный критерий для проверки гипотезы $H_0: \theta = \theta_0$ против альтернативы $H_1: \theta = \theta_1$ ($\theta_1 > \theta_0$) на уровне значимости $\alpha \in (0, 1)$. Найти мощность этого критерия.

получение про
переход к достат
статистике:

допускаем

$$\begin{aligned}
 & \mathcal{N}(\theta, \sigma^2) \\
 & x_{1-\alpha} : \bar{X} \geq c_\alpha \\
 & c_\alpha : \alpha = P_{\theta_0} \left(\bar{X} \geq c_\alpha \right) \\
 & \quad \quad \quad \mathcal{N}(\theta_0, \frac{\sigma^2}{n}) \rightarrow \mathcal{N}(0,1) \\
 & = P_{\theta_0} \left(\frac{\sqrt{n}}{\sigma} (\bar{X} - \theta_0) \geq \frac{\sqrt{n}}{\sigma} (c_\alpha - \theta_0) \right) \\
 & = 1 - \Phi_{\mathcal{N}(0,1)} \left(\frac{\sqrt{n}}{\sigma} (c_\alpha - \theta_0) \right) = \alpha \\
 & \frac{\sqrt{n}}{\sigma} (c_\alpha - \theta_0) = x_{1-\alpha} \\
 & c_\alpha = \theta_0 + \frac{\sigma}{\sqrt{n}} x_{1-\alpha}
 \end{aligned}$$

Сила или мощность:

$$\begin{aligned}
 W(\delta, \theta_1) &= P_{\theta_1} \left(\bar{X} \geq c_\alpha \right) \\
 & \quad \quad \quad \mathcal{N}(\theta_1, \frac{\sigma^2}{n}) \rightarrow \mathcal{N}(0,1) \\
 & = P_{\theta_1} \left(\frac{\sqrt{n}}{\sigma} (\bar{X} - \theta_1) \geq \frac{\sqrt{n}}{\sigma} (c_\alpha - \theta_1) \right) \\
 & = 1 - \Phi_{\mathcal{N}(0,1)} \left(\frac{\sqrt{n}}{\sigma} (c_\alpha - \theta_1) \right) \\
 & = 1 - \Phi \left(\frac{\sqrt{n}}{\sigma} (\theta_0 - \theta_1) + x_{1-\alpha} \right), \text{ где } \Phi = F_{\mathcal{N}(0,1)} \\
 & \quad \quad \quad c_\alpha = \theta_0 + \frac{\sigma}{\sqrt{n}} x_{1-\alpha}
 \end{aligned}$$

2. Пусть дана выборка $X_1, \dots, X_n$ из нормального распределения $\mathcal{N}(\mu, \theta)$, μ известно. Построить наиболее мощный критерий для проверки гипотезы $H_0 : \theta = \theta_0$ против альтернативы $H_1 : \theta = \theta_1$ ($\theta_1 > \theta_0$) на уровне значимости $\alpha \in (0, 1)$. Найти мощность этого критерия.

$$\begin{aligned}
 & (2) \quad (\mu, \theta) \\
 & \mu_0 : \theta = \theta_0 \\
 & \mu_1 : \theta = \theta_1 \quad (\theta_1 > \theta_0) \\
 & \alpha \in (0, 1) \\
 & \Delta(x) = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_1}} \exp\left\{-\frac{(x_i - \mu)^2}{2\theta_1}\right\}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_0}} \exp\left\{-\frac{(x_i - \mu)^2}{2\theta_0}\right\}} = \\
 & \quad \quad \quad = \frac{\theta_0^n \exp\left\{-\frac{1}{2\theta_1} \sum_{i=1}^n (x_i - \mu)^2\right\}}{\theta_1^n \exp\left\{-\frac{1}{2\theta_0} \sum_{i=1}^n (x_i - \mu)^2\right\}} = \frac{\theta_0^n}{\theta_1^n} \exp\left\{T(x) \cdot \frac{(\theta_1^2 - \theta_0^2)}{2\theta_0^2\theta_1^2}\right\} \\
 & \quad \quad \quad T(x) = \sum_{i=1}^n (x_i - \mu)^2
 \end{aligned}$$

знак опять
состав!

$$\begin{aligned}
 & \chi^2_{1-\alpha} \left\{ \sum (x_i - \mu)^2 \geq C_\alpha \right\} \\
 & C_\alpha: \alpha = P_{\theta_0} \left(\sum (x_i - \mu)^2 \geq C_\alpha \right) = \\
 & = P_{\theta_0} \left(\underbrace{\frac{1}{\theta_0^2} \sum (x_i - \mu)^2}_{\chi^2(n)} \geq \frac{1}{\theta_0^2} C_\alpha \right) = 1 - F_{\chi^2(n)} \left(\frac{1}{\theta_0^2} C_\alpha \right) = \alpha \\
 & \frac{1}{\theta_0^2} C_\alpha = \chi^2_{1-\alpha} \\
 & C_\alpha = \theta_0^2 \chi^2_{1-\alpha} \\
 & W(\delta, \theta_1) = 1 - F_{\chi^2(n)} \left(\frac{\theta_0^2}{\theta_1^2} \chi^2_{1-\alpha} \right)
 \end{aligned}$$

3. Пусть $X_1, \dots, X_n$ – выборка из экспоненциального распределения $\text{Exp}(\theta)$. Постройте равномерно наиболее мощный критерий уровня значимости α для проверки гипотезы $H_0: \theta = \theta_0$ против альтернативы $H_1: \theta < \theta_0$. Вычислите его мощность.

$$\begin{aligned}
 & H_0: \theta = \theta_0 \quad H_1: \theta = \theta_1 \quad (\theta_1 < \theta_0) \\
 & L(x) = \frac{\prod_{i=1}^n \theta_1 e^{-\theta_1 x_i}}{\prod_{i=1}^n \theta_0 e^{-\theta_0 x_i}} = \frac{\theta_1^n}{\theta_0^n} \exp \left\{ \sum_{i=1}^n x_i (\theta_0 - \theta_1) \right\} \\
 & = \frac{\theta_1^n}{\theta_0^n} \exp \left\{ (\theta_0 - \theta_1) n \bar{X} \right\} \geq t_\alpha \\
 & \left\{ \frac{n \bar{X}}{\theta_0} \geq C_\alpha \right\} \rightarrow ?
 \end{aligned}$$

$$\begin{aligned}
 & P_{\theta_0} (n \bar{X} \geq C_\alpha) = \alpha \\
 & F_{\sum X_i} (C_\alpha) = 1 - \alpha \\
 & C_\alpha - 1 - \alpha \text{-квантиль в } F_{\sum X_i} \\
 & \text{почему важно знать распред.} \\
 & X_i \sim \text{Exp}(\theta), x \geq 0. \\
 & \sum X_i \sim \text{Gamma}(n, \theta)
 \end{aligned}$$

Короткое обсуждение:

4. Пусть $X_1, \dots, X_n$ – выборка из распределения Бернулли $\text{Bin}(1, \theta)$. Проверьте гипотезу $H_0: \theta = \theta_0$ против альтернативы $H_1: \theta = \theta_1$ ($\theta_1 < \theta_0$) на уровне значимости α .

дост. стат. $\Rightarrow \left\{ \bar{X} \leq C_\alpha \right\}$ - Binom

так как дискретное,
то получим ил. не найдется,
берем ближайшее до заданного есть.

$$\begin{aligned}
 & P_{\theta_0} (n \bar{X} \geq C_\alpha) = \alpha_0 \leq \alpha. \\
 & F_{\sum X_i} (C_\alpha) = \alpha_0 \leq \alpha \\
 & C_\alpha - \alpha_0 \text{-квантиль в } \text{Binom}(n, \theta_0). \\
 & \sum C_\alpha.
 \end{aligned}$$

5. Имеется X_1 – выборка объема 1. Проверьте гипотезу

$$H_0 : X_1 \sim \mathcal{U}(0, 1)$$

против альтернативы

$$H_1 : X_1 \sim \text{Exp}(1).$$

Постройте наиболее мощный критерий уровня значимости α для различения этих гипотез и вычислите его мощность.

$$\begin{aligned} \mathcal{X}_1 &= \mathcal{X}_1^* \cup (1, +\infty) = \{x \in [0, 1] \cup (1, +\infty)\} \\ H_0: X_1 &\sim \mathcal{U}(0, 1) & \leftarrow F_0(x) \\ H_1: X_1 &\sim \text{Exp}(1) & \leftarrow F_1(x) \\ x \in [0, 1]: & \Lambda(x) = \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} = \frac{1 \cdot e^{-x}}{1} = e^{-x} \\ x > 1: & \Lambda(x) = \frac{0}{0} = 1 \\ \Lambda(x) &= e^{-x} \Rightarrow \mathcal{X}_1^* = \{x \in [0, 1]: e^{-x} \geq t_\alpha\} \\ P_{\theta_0}(e^{-X_1} \geq t_\alpha) &= P(X_1 \leq -\ln t_\alpha) = \alpha \end{aligned}$$

$$\begin{aligned} \Lambda(X) &\uparrow \text{от } T(X) \in \mathbb{R} \\ \alpha &= F(C_\alpha) \\ &= C_\alpha \\ \Rightarrow C_\alpha &= \alpha \\ \downarrow \\ \mathcal{X}_1^* &= \{x \in [0, 1]: x \leq \alpha\} \\ &= \{x \in [1-\alpha, 1]\} \end{aligned}$$
