

概率统计 (A) 课程作业: 假设检验

周书予

2000013060@stu.pku.edu.cn

June 1, 2022

1

1.

$$\begin{aligned}\alpha &= \mathbb{P}(\text{reject } H_0 | H_0) = \mathbb{P}(\bar{X} \geq 0.5 | \theta = 0.4) = 0.4^3 + 3 \times 0.4^2 \times (1 - 0.4) = 0.352 \\ \beta &= \mathbb{P}(\text{accept } H_0 | H_1) = \mathbb{P}(\bar{X} < 0.5 | \theta = 0.5) = (1 - 0.5)^3 + 3 \times (1 - 0.5)^2 \times 0.5 = 0.5\end{aligned}$$

显著性水平即为第 I 类错误发生的概率, 即 $\alpha = 0.352$.

2.

$$p = \mathbb{P}\left(Z \geq \frac{2}{3} | H_0\right) = \mathbb{P}\left(\bar{X} \geq \frac{2}{3} | \theta = 0.4\right) = 0.4^3 + 3 \times 0.4^2 \times (1 - 0.4) = 0.352$$

2

1. 记 $\mu_0 = 30$. 当 $\sigma = 1.1$ 已知时, 取统计量 $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$, 计算其样本取值为 $z_0 = 0.727$.

对于原假设 $H_0: \mu \leq \mu_0$, 根据 Neyman-Pearson 原则, 检验的拒绝域为

$$W = \left\{ Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \right\}$$

由于 $z_{0.05} = 1.645$, 故样本取值不在拒绝域内, 接受原假设.

2. 对于原假设 $H_0: \mu \geq \mu_0$, 根据 Neyman-Pearson 原则, 检验的拒绝域为

$$W = \left\{ Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha \right\}$$

同样, 样本取值不在拒绝域内, 接受原假设.

3. 当 σ 未知时, 取统计量 $T = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}}$, 其样本取值为 $t_0 = 0.713$.

对于原假设 $H_0: \mu \leq \mu_0$, 根据 Neyman-Pearson 原则, 检验的拒绝域为

$$W = \left\{ T = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} \geq t_\alpha(n-1) \right\}$$

由于 $t_{0.05}(5) = 2.015$, 故样本取值不在拒绝域内, 接受原假设.

3

1. 注意到当原假设 H_0 成立时, 统计量 $T = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$, 故

$$\alpha \geq \mathbb{P}(\text{reject } H_0 | H_0) = \mathbb{P}(T \leq C | H_0) = F_{n-1}(C)$$

因此 $C \leq F_{n-1}^{-1}(\alpha)$.

2. 当原假设 H_0 成立时, 统计量 $T = \sum_i \frac{(X_i - \mu)^2}{\sigma_0^2} \sim \chi^2(n)$, 故

$$\alpha \geq \mathbb{P}(\text{reject } H_0 | H_0) = \mathbb{P}(T \leq C | H_0) = F_n(C)$$

因此 $C \leq F_n^{-1}(\alpha)$.

4

Y 的分布函数为

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \left(\frac{y}{\theta}\right)^n, & 0 \leq y < \theta \\ 1, & y \geq \theta \end{cases}$$

于是

$$\alpha \geq \mathbb{P}(\text{reject } H_0 | H_0) = \mathbb{P}\left(Y > \frac{\theta_0}{a} \vee Y < \frac{\theta_0}{b} | \theta = \theta_0\right) = 1 - F_Y\left(\frac{\theta}{a}\right) + F_Y\left(\frac{\theta}{b}\right) = 1 - \frac{1}{a^n} + \frac{1}{b^n}$$

当 $\alpha > 1 - \frac{1}{a^n}$ 时, 有 $b \geq \sqrt[n]{\frac{1}{\alpha - 1 + \frac{1}{a^n}}}$, 即 b 的最小值为 $\sqrt[n]{\frac{1}{\alpha - 1 + \frac{1}{a^n}}}$. 当 $\alpha \leq 1 - \frac{1}{a^n}$ 时, 不存在满足条件的 b .

5

$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ 的样本取值为 1.639728.

当原假设 H_0 成立时, 该统计量服从近似 t 分布, 近似自由度为 $k = \frac{(t_1 + t_2)^2}{t_1^2/(n_1 - 1) + t_2^2/(n_2 - 1)}$, 其样本取值为 17.371796. 拒绝域为 $W = \{T | T > C\}$, 则有 (记自由度为 k 的 t 分布的分布函数为 $G_k(\cdot)$)

$$\alpha \geq \mathbb{P}(T > C | T \sim t(k)) = 1 - G_k(C)$$

从而 $C \geq G_k^{-1}(1 - \alpha)$.

当 α 分别取 0.05 和 0.01 时, 可分别计算 $G_k^{-1}(0.95) = 1.737467$, $G_k^{-1}(0.99) = 2.561309$, T 的样本取值均不落在拒绝域内, 因此结论均为接受原假设.

6

- 1.

$$\begin{aligned} L(x_1, \dots, x_n, \mu = 0) &= \mathbb{P}(x_1, \dots, x_n | \mu = 0) = \prod_{i=1}^n \mathbb{P}(x_i | \mu = 0) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\sum_{i=1}^n x_i^2/2\right) \\ L(x_1, \dots, x_n, \mu = 2) &= \mathbb{P}(x_1, \dots, x_n | \mu = 2) = \prod_{i=1}^n \mathbb{P}(x_i | \mu = 2) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\sum_{i=1}^n (x_i - 2)^2/2\right) \\ \lambda(\mathbf{X}) &= \frac{L(X_1, \dots, X_n, \mu = 2)}{L(X_1, \dots, X_n, \mu = 0)} = \exp\left(2\sum_{i=1}^n X_i - 2n\right) \end{aligned}$$

2.

$$\begin{aligned}
 \alpha &\geq \mathbb{P}(\text{reject } H_0 | H_0) = \mathbb{P}\left(\exp\left(2\sum_{i=1}^n X_i - 2n\right) \geq \lambda_{\text{LRT}} \middle| X_i \sim \text{i.i.d. } \mathcal{N}(0, 1)\right) \\
 &= \mathbb{P}\left(S \geq \frac{\ln \lambda_{\text{LRT}}}{2} + n \middle| S \sim \mathcal{N}(0, n)\right) \\
 &= \mathbb{P}\left(T \geq \frac{\ln \lambda_{\text{LRT}}}{2\sqrt{n}} + \sqrt{n} \middle| T \sim \mathcal{N}(0, 1)\right) \\
 &= 1 - \Phi\left(\frac{\ln \lambda_{\text{LRT}}}{2\sqrt{n}} + \sqrt{n}\right) \\
 \lambda_{\text{LRT}} &\geq \exp(2\sqrt{n}\Phi^{-1}(1 - \alpha) - 2n)
 \end{aligned}$$

即当显著水平为 $\alpha = 0.05$ 时, λ_{LRT} 的最小值为 $\exp(2\sqrt{n}\Phi^{-1}(0.95) - 2n)$.

3.

$$L(\mathbf{x}, 2)[\phi_{\text{LRT}}(\mathbf{x}) - \phi(\mathbf{x})] \geq \lambda_{\text{LRT}} L(\mathbf{x}, 0)[\phi_{\text{LRT}}(\mathbf{x}) - \phi(\mathbf{x})] \quad (1)$$

当 $\lambda(\mathbf{x}) = \frac{L(\mathbf{x}, 2)}{L(\mathbf{x}, 0)} \geq \lambda_{\text{LRT}}$ 时, $\mathbf{x} \in W_{\text{LRT}}$, 故 $\phi_{\text{LRT}}(\mathbf{x}) = 1$, 则 $\phi_{\text{LRT}}(\mathbf{x}) - \phi(\mathbf{x}) \geq 0$, 从而 eq. (1) 成立.

当 $\lambda(\mathbf{x}) < \lambda_{\text{LRT}}$ 时, $\mathbf{x} \notin W_{\text{LRT}}$, 故 $\phi_{\text{LRT}}(\mathbf{x}) = 0$, 则 $\phi_{\text{LRT}}(\mathbf{x}) - \phi(\mathbf{x}) \leq 0$, 从而 eq. (1) 也成立.

4. 注意到 $\beta_{\text{LRT}} = \mathbb{P}(\mathbf{X} \in W_{\text{LRT}} | H_1)$ 为备选假设 H_1 成立时, 原假设被拒绝的概率, 可以写为

$$\beta_{\text{LRT}} = \int_{\mathbf{x}} L(\mathbf{x}, 2) \phi_{\text{LRT}}(\mathbf{x}) d\mathbf{x}$$

类似的也有

$$\beta = \int_{\mathbf{x}} L(\mathbf{x}, 2) \phi(\mathbf{x}) d\mathbf{x}$$

考虑拒绝域 W 的显著水平为 α , 即

$$\int_{\mathbf{x}} L(\mathbf{x}, 0) \phi(\mathbf{x}) d\mathbf{x} \leq \alpha$$

而通过取最小的 λ_{LRT} 可以实现

$$\int_{\mathbf{x}} L(\mathbf{x}, 0) \phi_{\text{LRT}}(\mathbf{x}) d\mathbf{x} = \alpha$$

因故, 有

$$\begin{aligned}
 \beta_{\text{LRT}} - \beta &= \int_{\mathbf{x}} L(\mathbf{x}, 2)[\phi_{\text{LRT}}(\mathbf{x}) - \phi(\mathbf{x})] d\mathbf{x} \\
 &\geq \lambda_{\text{LRT}} \int_{\mathbf{x}} L(\mathbf{x}, 0)[\phi_{\text{LRT}}(\mathbf{x}) - \phi(\mathbf{x})] d\mathbf{x} \\
 &\geq \lambda_{\text{LRT}}(\alpha - \alpha) \\
 &= 0
 \end{aligned}$$