概率统计 (A) 课程作业: 假设检验

周书予

2000013060@stu.pku.edu.cn

May 31, 2022

1

1.

$$\alpha = \mathbb{P}(\text{reject } H_0|H_0) = \mathbb{P}\left(\overline{X} \geqslant 0.5|\theta = 0.4\right) = 0.4^3 + 3 \times 0.4^2 \times (1-0.4) = 0.352$$
$$\beta = \mathbb{P}(\text{accept } H_0|H_1) = \mathbb{P}\left(\overline{X} < 0.5|\theta = 0.5\right) = (1-0.5)^3 + 3 \times (1-0.5)^2 \times 0.5 = 0.5$$

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

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

2

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

   对于原假设 $H_0 : \mu \leqslant \mu_0$, 根据 Neyman-Pearson 原则, 检验的拒绝域为
$$W = \left\{Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \geqslant z_\alpha\right\}$$

   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 $H_0 : \mu \geqslant \mu_0$, 根据 Neyman-Pearson 原则, 检验的拒绝域为
$$W = \left\{Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \leqslant -z_\alpha\right\}$$

   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 $T = \frac{\overline{X}-\mu_0}{\sqrt{S^2/n}}$, 其样本取值为 $t_0 = 0.713$.

   对于原假设 $H_0 : \mu \leqslant \mu_0$, 根据 Neyman-Pearson 原则, 检验的拒绝域为
$$W = \left\{T = \frac{\overline{X}-\mu_0}{\sqrt{S^2/n}} \geqslant t_\alpha(n-1)\right\}$$

   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

$$\alpha \geqslant \mathbb{P}\left(\text{reject } H_0|H_0\right) = \mathbb{P}\left(T \leqslant C|H_0\right) = F_{n-1}(C)$$

   因此 $C \leqslant F_{n-1}^{-1}(\alpha)$.

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

$$\alpha \geqslant \mathbb{P}\left(\text{reject } H_0|H_0\right) = \mathbb{P}\left(T \leqslant C|H_0\right) = F_n(C)$$

   因此 $C \leqslant F_n^{-1}(\alpha)$.

4

Y 的分布函数为

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

于是

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

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

5

$T = \frac{\overline{X} - \overline{Y}}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ 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 $G_k(\cdot)$)

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

从而 $C \geqslant G_k^{-1}(1-\alpha)$.

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.

$$L(x_1, \cdots, x_n, \mu = 0) = \mathbb{P}\left(x_1, \cdots, x_n|\mu = 0\right) = \prod_{i=1}^n \mathbb{P}\left(x_i|\mu = 0\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\sum_{i=1}^n x_i^2/2\right)$$

$$L(x_1, \cdots, x_n, \mu = 2) = \mathbb{P}\left(x_1, \cdots, x_n|\mu = 2\right) = \prod_{i=1}^n \mathbb{P}\left(x_i|\mu = 2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\sum_{i=1}^n (x_i-2)^2/2\right)$$

$$\lambda(\mathbf{X}) = \frac{L(X_1, \cdots, X_n, \mu = 2)}{L(X_1, \cdots, X_n, \mu = 0)} = \exp\left(2\sum_{i=1}^n X_i - 2n\right)$$

2.

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

3.

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

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

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

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

类似的也有

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

考虑拒绝域 W 的显著水平为 α, 即

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

而通过取最小的 λ_{LRT} 可以实现

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

因故, 有

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