概统作业 (Week 14)

PB20000113 孔浩宇

June 25, 2023

1 (P306 T6)

(1) 对 $X_{(n)}$ 有分布函数和密度函数

$$F_n(x) = \begin{cases} 0, & (x \le 0) \\ \dfrac{x^n}{\theta^n}, & (x \in (0,\theta)) \\ 1. & (x \ge \theta) \end{cases} \qquad f_n(x) = \begin{cases} \dfrac{n \cdot x^{n-1}}{\theta^n}, & (x \in (0,\theta)) \\ 0. & (\text{其他}) \end{cases}$$

故功效函数为

$$\beta_\Psi(\theta) = \mathbb{P}_\theta\left(X_{(n)} \le 2.5 \middle| H_0\right) = F_n(2.5) = \left(\frac{2.5}{\theta}\right)^n \quad (\theta \ge 3)$$

对于 $\forall\ \theta \in H_0$，有

$$\theta \ge 3,\ \alpha_{1\Psi}(\theta) = \beta_\Psi(\theta) = \left(\frac{2.5}{\theta}\right)^n \le \left(\frac{2.5}{3}\right)^n$$

故有显著性水平

$$\alpha = \left(\frac{2.5}{3}\right)^n = \left(\frac{5}{6}\right)^n$$

(2)

$$\alpha \le 0.05 \ \Rightarrow\ \left(\frac{5}{6}\right)^n \le 0.05 \ \Rightarrow\ n \ge \log_{\frac{5}{6}} 0.05 \ \Rightarrow\ n \ge 16.43$$

即 n 至少为 17.

2 (P307 T9)

检验

$$H_0 : \mu = 105.02 \quad \leftrightarrow \quad H_1 : \mu \ne 105.02$$

8 个省的数据分别记为 $X_1, X_2, ..., X_8$，记 $n = 8$，有统计量

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = 105.0425$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 = 13.23$$

取检验统计量

$$T = \frac{\sqrt{n}(\overline{X} - \mu)}{S} \sim t_{n-1} \quad (\mu = \mu_0)$$

拒绝域

$$W = \left\{ |T| > t_{n-1}\left(\frac{\alpha}{2}\right) \right\}$$

代入数据，得

$$T = \frac{\sqrt{8} \times 0.0225}{\sqrt{13.23}} = 0.0175, \qquad t_{n-1}\left(\frac{\alpha}{2}\right) = t_7(0.025) = 2.3646$$

有

$$|T| = 0.0175 < 2.3646 = t_{n-1}\left(\frac{\alpha}{2}\right)$$

不在拒绝域内，故在显著性水平 5% 下，不能拒绝 H_0.

3

(1) 不妨记旧工艺下样本分布为 $X \sim N(\mu_1, \sigma_1^2)$，$m = 12$，有样本均值 $\overline{X}$ 与样本方差 S_1^2

$$\overline{X}_1 = \frac{6+4+5+5+6+5+5+6+4+6+7+4}{12} = 5.25$$

$$S_1^2 = \frac{1}{m-1} \sum_{i=1}^{m} \left(X_i - \overline{X}\right)^2 = 0.9318$$

记新工艺下样本分布为 $X \sim N(\mu_2, \sigma_2^2)$，$n = 12$，有样本均值 $\overline{Y}$ 与样本方差 S_2^2

$$\overline{Y} = \frac{2+1+2+2+1+0+3+2+1+0+1+3}{12} = 1.5$$

$$S_2^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2 = 1$$

取检验统计量

$$F = \frac{S_1^2}{S_2^2} \sim F(m-1, n-1) \quad (\sigma_1^2 = \sigma_2^2)$$

检验

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \leftrightarrow \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

拒绝域

$$W = \left\{ F < F_{11,11}\left(1 - \frac{\alpha}{2}\right) \text{ 或 } F > F_{11,11}\left(\frac{\alpha}{2}\right) \right\}$$

代入数据，有

$$F_{11,11}\left(1 - \frac{\alpha}{2}\right) = F_{11,11}(0.975) = 0.288 \qquad F_{11,11}\left(\frac{\alpha}{2}\right) = F_{11,11}(0.025) = 3.474$$

$$F = \frac{S_1^2}{S_2^2} = 0.9318 \notin W$$

故不能拒绝 H_0，认为两个总体的方差想等。

(2) 即检验

$$H_0 : \mu_1 - \mu_2 \geq 3 \quad \leftrightarrow \quad H_1 : \mu_1 - \mu_2 < 3$$

取检验统计量

$$T = \frac{\overline{X} - \overline{Y} - 3}{S_T \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t_{m+n-2} \quad (\mu_1 - \mu_2 = 3)$$

其中有优良点估计

$$S_T = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}}$$

拒绝域

$$W = \{T < -t_{m+n-2}(\alpha)\}$$

代入数据，得

$$S_T = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}} = \sqrt{\frac{11 \times 0.9318 + 11 \times 1}{12+12-2}} = 0.9828$$

$$T = \frac{\overline{X} - \overline{Y} - 3}{S_T\sqrt{\frac{1}{m} + \frac{1}{n}}} = \frac{5.25 - 1 - 3}{0.9828 \times \sqrt{\frac{1}{12} + \frac{1}{12}}} = 1.8693$$

$$t_{m+n-2}(\alpha) = t_{22}(0.05) = 1.7171$$

有

$$T = 1.8693 > -1.7171 \ \Rightarrow \ T \notin W$$

故不能拒绝 H_0，可以认为旧工艺下 NDMA 平均含量比新工艺下显著地大 3.

4 (P312 T38)

记马克 · 吐温文中 3 个字母组成的单字的比例为 $X \sim N(\mu_1, \sigma_1^2)$，$m = 8$，有统计量

$$\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i = 0.2319$$

$$S_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(X_i - \overline{X}\right)^2 = 2.12 \times 10^{-4}$$

记斯诺德格拉斯文中 3 个字母组成的单字的比例为 $Y \sim N(\mu_2, \sigma_2^2)$，$n = 10$，有统计量

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = 0.2097$$

$$S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 = 9.33 \times 10^{-5}$$

有

$$\sigma_1^2 = \sigma_2^2 = \sigma^2 \quad (\sigma^2 \text{ 未知})$$

检验

$$H_0 : \mu_1 = \mu_2 \quad \leftrightarrow \quad H_1 : \mu_1 \neq \mu_2$$

取检验统计量

$$T = \frac{\overline{X} - \overline{Y}}{S_T\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t_{m+n-2} \quad (\mu_1 - \mu_2 = 3)$$

其中有优良点估计

$$S_T = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}}$$

拒绝域

$$W = \{|T| > t_{m+n-2}(\alpha/2)\}$$

代入数据，得

$$S_T = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}} = \sqrt{\frac{7 \times 2.12 \times 10^{-4} + 9 \times 9.33 \times 10^{-5}}{8+10-2}} = 0.01205$$

$$T=\frac{\overline{X}-\overline{Y}}{S_T\sqrt{\frac{1}{m}+\frac{1}{n}}}=\frac{0.2319-0.2097}{0.01205\times\sqrt{\frac{1}{8}+\frac{1}{10}}}=3.884$$

$$t_{m+n-2}(\alpha/2)=t_{16}(0.025)=2.1199$$

有

$$|T|=3.884>2.1199\ \Rightarrow\ T\in W$$

故拒绝 H_0，即在显著性水平 $\alpha=0.05$ 下有差异，不能认为是一个人。

5 A

Proof.

对参数 μ 进行检验

$$H_0:\mu=\mu_0\quad\leftrightarrow\quad H_1:\mu\neq\mu_0$$

不妨记检验统计量为 T_n.

(1) σ^2 已知，此时拒绝域

$$\left\{|Z|>u_{\alpha/2}\right\}$$

由 $\alpha=0.05$ 时接受 H_0，有

$$|T_n|\leq u_{0.025}$$

又

$$u_{0.005}>u_{0.025}\ \Rightarrow\ |T_n|<u_{0.005}$$

即此时 $\alpha=0.01$ 时仍接受 H_0。

(2) σ^2 未知，此时拒绝域

$$\{|T|>t_{n-1}(\alpha/2)\}$$

由 $\alpha=0.05$ 时接受 H_0，有

$$|T_n|\leq t_{n-1}(0.025)$$

又

$$t_{n-1}(0.005)>t_{n-1}(0.025)\ \Rightarrow\ |T_n|<t_{n-1}(0.005)$$

即此时 $\alpha=0.01$ 时仍接受 H_0。

综上，当 $\alpha=0.01$ 时仍接受 H_0，故选择 A. □