Part III-B: Probability Theory and Mathematical Statistics

Lecture by 李漫漫

Note by THF

2024 年 11 月 30 日

目录

Lecture 21

11.28

两个总体的参数估计

假设种类：

- $H_0 : \mu_1 = \mu_2, H_1 : \mu_1 \neq \mu_2$
- $H_0 : \mu_1 \geq \mu_2, H_1 : \mu_1 < \mu_2$
- $H_0 : \mu_1 \leq \mu_2, H_2 : \mu_1 > \mu_2$
- $^\star H_0 : \mu_1 - \mu_2 \geq c, H_1 : \mu_1 - \mu_2 < c$
- $^\star H_0 : \mu_1 - \mu_2 \leq c, H_1 : \mu_1 - \mu_2 > c$

一般使用点估计：$\overline{X} = \hat{\mu_1}, \overline{Y} = \hat{\mu_2}$，将假设转为：$\mu_1 - \mu_2 \geq c \Rightarrow \overline{X} - \overline{Y} \geq c$

原因：$\overline{X} - \overline{Y}$ 是 $\mu_1 - \mu_2$ 的最小方差无偏估计

Notation. 两个总体匹配/不独立

Example. 一种马达正常工作的平均电流不超过 0.8A，抽取 16 台马达，测得 $\overline{X} = 0.92, S^2 = 0.32$，假设电流符合正态分布 $X \sim N\left(\mu, \sigma^2\right)$，取 $\alpha = 0.05$，求厂家的话是否可信

Solve. 确定假设：有两种可能的假设：

a. $H_0 : \mu \leq 0.8, H_1 : \mu > 0.8$

b. $H_0 : \mu \geq 0.8, H_1 : \mu < 0.8$

确定假设统计量：由于 σ 未知，因此使用 t 统计量：$\frac{\overline{X} - \mu}{S}\sqrt{n} > t_{1-\frac{\alpha}{2}}(n-1)$

分别带入数据后发现：对于 $H_0 : \mu < 0.8$ 和 $H_0 : \mu \geq 0.8$，都不拒绝原假设

Notation. 对于假设检验，任何假设都有犯错误的可能，拒绝原假设的可能是充分的（α 一般较小），不拒绝原假设有较大的可能犯错误（β 可能更大），因此不拒绝原假设的结论需要加大样本量继续验证

正态总体的方差的检验

检验统计量不使用 μ ： $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$
继续化简：

$$\chi^2 = \frac{(n-1)\,S^2}{\sigma_0^2} = \frac{(n-1)\cdot\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma_0^2} = \frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma_0^2}.$$

此时使用 $S^2=\hat{\sigma}^2$ 来估计 σ ，如果 μ 已知则可以使用 $\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu\right)^2=\hat{\sigma}^2$ 来估计 σ

总体分布的卡方拟合优度检验

根据样本预测总体的分布种类（假设）

Example. 某建筑工地发生事故的记录：求 $\alpha = 0.05$ 下，数据是否符合泊松分布 $P(\lambda)$

表 1: 工地事故

事故数	天数
0	102
1	59
2	30
3	8
4	0
5	1
≥ 6	0
合计	200

Notation. 泊松分布：

$$X\sim P(\lambda)\quad p=P(X\leq x)=\sum_{k=1}^{n}\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}.$$

Solve. 设每天发生事故 i 次为事件 A_i，确定假设：

- 原假设 $H_0:\forall i, P(A_i)=p_i$
- 被则假设 $H_1:\exists i, P(A_i)\neq p_i$

λ 可以使用 $\overline{X}=\hat{\lambda}$ 估计，即 $\hat{\lambda}=\overline{x}=0.74$ ，使用 $P(0.74)$ 可以计算 $\hat{p}_i$

确定假设统计量:

$$\chi^2 = \sum_{i=1}^{m} \frac{n_i}{np_i} - n$$
$$\hat{\chi}^2 = \sum_{i=1}^{m} \frac{n_i}{n\hat{p}_i} - n$$

.

Lecture 22

11.30