

第七章、估计

§7.1 最大似然估计

- 总体 $X \sim F_\theta$. 目标: 给出 θ 的估计值 $\hat{\theta}$.
- 思想: 大概 事件发生. 支撑: 概 的主观置信度含义.
- 似然函数:

$$L(\theta) = p(x_1, \theta) \cdots p(x_n, \theta),$$

其中, $p(\theta, x)$ 为总体的分布列/密度.

- 定义 1.1. θ 的最大似然估计 $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ 指 $L(\theta)$ 的最大值. 即, $\hat{\theta} \in \Theta$, 且

$$L(\hat{\theta}) = \max_{\theta \in \Theta} p(x_1, \theta) \cdots p(x_n, \theta).$$

- 在似然函数 $L(\theta) = p(x_1, \theta) \cdots p(x_n, \theta)$ 中, 视数据 \vec{x} 为已知, 视参数 θ 为未知.
- 用 $\hat{\theta}(x_1, \cdots, x_n)$ 则强调计算, 用 $\hat{\theta}(X_1, \cdots, X_n)$ 则强调理论.



例1.1(正态模型/测量模型). 测试飞机最大飞行速度, 得到 $n = 15$ 个数据: x_1, \dots, x_n . 试估计飞机的最大飞行速度的均值.

- 假设飞机最大飞行速度 X 服从正态分布 $N(\mu, \sigma^2)$.
- $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, \infty)$. μ 为待估参数, σ^2 为讨厌参数.
- 似然函数:

$$L(\theta) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\}.$$

- 进一步, 还可求 $\tau = \sigma^2$ 的最大似然估计.
- 将 $\hat{\mu} = \bar{x}$ 代入, 记 $a = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$.

$$L(\bar{x}, \tau) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\}$$

$$= \frac{1}{(\sqrt{2\pi})^n} \tau^{-\frac{n}{2}} e^{-\frac{n}{2} \cdot \frac{a}{\tau}} = \frac{1}{(\sqrt{2\pi})^n} \exp \left\{ -\frac{n}{2} \left(\ln \tau + \frac{a}{\tau} \right) \right\}.$$

$\hat{\tau}$ 即为 $\ln \tau + \frac{a}{\tau}$ 的最小值

$$\frac{d}{d\tau} \left(\ln \tau + \frac{a}{\tau} \right) = \frac{1}{\tau} - \frac{a}{\tau^2} = \frac{1}{\tau^2} (a - \tau) \Rightarrow \hat{\tau} = a.$$

- 称 $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ 为**样本方差**.

例1.2(次品 的估计). 某工人生产20 件产品, 检查出恰有一件为次品. 估计该工人生产的次品 .

• 总体 $X \sim B(1, p)$, $p \in [0, 1]$. 样本量: $n = 20$.

• 似然函数:

$$L(p) = p^s(1-p)^{n-s}, \quad \text{其中 } s = x_1 + \dots + x_n.$$

• \hat{p} 也为 $\ln L(p) = s \ln p + (n-s) \ln(1-p)$ 的最大值:

$$\frac{d}{dp} \ln L(p) = \frac{s}{p} - \frac{n-s}{1-p} \Rightarrow \hat{p} = \frac{s}{n}.$$

• $n = 20$, $s = 1$, 因此, $\hat{p} = \frac{1}{20}$.

例1.4. 总体: $X \sim U(0, \theta)$, 数据: X_1, \dots, X_n , (样本量: n).

求: θ 的最大似然估计.

- 似然函数:

$$L(\theta) = \prod_{i=1}^n \frac{1}{\theta} \mathbf{1}_{\{0 \leq x_i \leq \theta\}} = \frac{1}{\theta^n} \mathbf{1}_{\{0 \leq x_1, \dots, x_n \leq \theta\}}$$

• 仅当 $\theta \geq \max_{1 \leq i \leq n} x_i$ 时, $L(\theta) > 0$.

• 当 $\theta \geq \max_{1 \leq i \leq n} x_i$ 时, $L(\theta) = \frac{1}{\theta^n}$, 关于 θ 单调下降.

- 从而, $\hat{\theta} = \max_{1 \leq i \leq n} x_i$, 即 $\max_{1 \leq i \leq n} X_i$.



例1.6. 总体: $X \sim \mathcal{P}(\lambda)$, 样本量: n . 求: λ 的最大似然估计.

• 似然函数:

$$L(\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \frac{1}{\prod_{i=1}^n x_i!} \lambda^{n\bar{x}} e^{-\lambda n} = \frac{1}{\prod_{i=1}^n x_i!} e^{n(\bar{x} \ln \lambda - \lambda)}.$$

• λ 是 $\bar{x} \ln \lambda - \lambda$ 的最大值:

$$\frac{d}{d\lambda} (\bar{x} \ln \lambda - \lambda) = \frac{\bar{x}}{\lambda} - 1 \Rightarrow \lambda = \bar{x}, \text{ 或 } X.$$