第六章、统计学基本概念, 第七章、估计

- 第十三次课

  §6.1 引言

  §6.2 若干基本概念

  §7.1 最大似然估计

- 第十四次课

  §7.2 矩估计

  §7.3 估计的无偏性

  §7.4 无偏估计的优良性

- 第十五次课

  §7.5 估计的相合性

  §7.6 估计的渐近分布

  §7.7 置信区间和置信限

§6.1 引言, §6.2 若干基本概念

- 概率vs 统计

  理论vs 应用, 极限定理: LLN, CLT.
- 数据(data), 收集、分析...; 思想、方法!
- 总体模型$X \sim F_\theta$, $\theta \in \Theta$.
- 样本$\vec{X} = (X_1, \cdots, X_n)$, 其中$X_1, \cdots, X_n \sim$ i.i.d. F_θ
- 样本量n,

  样本值$\vec{x} = (x_1, \cdots, x_n)$,
- 研究对象: θ 或$g(\theta)$.

例: 有N 个产品, 其中M 个次品. $N, M >> 1$, 未知.

目标: 调查次品率$p = \frac{M}{N}$.

方法: 任取n 个.

- 总体$\Omega_0 = \{1, \cdots, M, M+1, \cdots N\}$. 古典概型.
- 根据目标, 对样本i 赋值, 产生随机变量(总体!):

$$X(i) = \begin{cases} 1, & i \leqslant M, \\ 0, & i \geqslant M+1. \end{cases}$$

  总体分布: $F_\theta =$ 两点分布, 参数$\theta = p = \frac{M}{N}$. $\Theta = [0, 1]$.
- 任取n 个. 产生新模型:

  $\Omega = \{\omega = (i_1, \cdots, i_n), i_1, \cdots, i_n \in \Omega_0\}$.

  得到n 个样本: $X_1(\omega) = X(i_1), \cdots, X_n(\omega) = X(i_n)$.
- 近似地, $X_1, \cdots, X_n$ 是定义在Ω 上的i.i.d随机变量, 都$\sim F_\theta$.

- 统计量: 样本的函数$T(\vec{X})$, 输入数据$\vec{x}$ 输出值$T(\vec{x})$.
  样本均值: $\frac{1}{n}(X_1+\cdots+X_n)=\bar{X}$,
  样本方差: $\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^2$.
- 样本值$\{x_i, i=1,\cdots,n\}\to$ **新模型(随机变量ξ)**:
  $P(\xi=x_i)=\frac{1}{n}$, $i=1,\cdots,n$.
  样本均值: $E\xi=\frac{1}{n}(x_1+\cdots+x_n)=\bar{x}$.
  样本方差: $\mathrm{var}(\xi)=\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2$.
- 正交分解$E(\xi-a)^2=\mathrm{var}(\xi)+(a-\bar{x})^2$:
  $\frac{1}{n}\sum_{i=1}^n(x_i-a)^2=\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2+(\bar{x}-a)^2$.
- 两类主要问题:
  参数估计: 给出θ 的估计值$\hat{\theta}$ 或区间$[\underline{\theta},\bar{\theta}]$.
  假设检验: 对某假设回答“是”或“否”.
- 其他: 回归、贝叶斯统计.

§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: 大概率事件发生.

似然函数: 参数(与数据)的函数.

- 离散型:

  $L(\theta) = \prod_{i=1}^n P_\theta(X_i = x_i)(= P_\theta(X_i = x_i, i = 1, \cdots, n))$.
- 连续型: $L(\theta) = \prod_{i=1}^n p_\theta(x_i)(= p_{\theta,\vec{X}}(x_1, \cdots, x_n))$.
- 定义1.1. θ 的最大似然估计指$L(\theta)$ 的最大值点, 记为$\hat{\theta}(x_1, \cdots, x_n)$.
- $L(\theta) = L(\theta, \vec{x})$, 当数据变化时, 似然函数会变.
- 统计量: 理论$\hat{\theta}(\vec{X})$ vs 应用$\hat{\theta}(\vec{x})$.

习题一、11. 在未名湖中捕鱼80条, 标记后放回, 再捕鱼100条, 其中4条有标记. 猜湖中的总鱼数N.

- $n=1$, X_1 =有标记的鱼数$\sim$ 超几何分布, $x_1=4$.
- $\theta=N$. $L(\theta,x_1)=P(X_1=x_1)=\dfrac{C_{80}^{x_1}C_{N-80}^{100-x_1}}{C_N^{100}}$.
- $L(\theta,4)=\dfrac{C_{80}^{4}C_{N-80}^{96}}{C_N^{100}}=:p_N$.
- p_N 的最大值点为$\hat{N}=1999$ 或2000 (皆可).

例1.1. 测试飞机最大飞行速度, 得到n 个数据: $x_1, \cdots, x_n$. 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$N(\mu, \sigma^2)$.
  假设的合理性vs 不合理性.
- 确定参数$\theta = (\mu, \sigma^2)$. θ: 待估参数, σ^2: 讨厌参数.
  写出似然函数.
  $L(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}$.
- 求$L(\theta)$ 的最大值点.
  固定σ^2, 即求$\sum_{i=1}^n (x_i - \mu)^2$ 的最小值点, 注意:
  $\frac{1}{n}\sum_{i=1}^n (x_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2 + (\bar{x} - \mu)^2$.
  $\hat{\mu} = \bar{x}$ (样本均值).
- 一般情况, $\hat{\mu}$ 可能会依赖于讨厌参数σ^2, 还需求出$\hat{\sigma}^2$ 再代入.

进一步, 还可求$\tau=\sigma^2$ 的最大似然估计.

- 似然函数.

  已知$\hat{\mu}=\bar{x}$. 此时, $\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2=: a$,

  $L(\bar{x},\sigma^2)=\frac{1}{(\sqrt{2\pi\sigma^2})^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2}=\frac{1}{(\sqrt{2\pi})^n}\tau^{-\frac{n}{2}}e^{-\frac{na}{2\tau}}$.
- 求$L(\bar{x},\tau)$ 即$\log L(\bar{x},\tau)=C-\frac{n}{2}\log\tau-\frac{na}{2\tau}$ 的最大值点.

  $\frac{d}{d\tau}\log L(\bar{x},\tau)=-\frac{n}{2\tau}+\frac{na}{2\tau^2}=\frac{n}{2\tau^2}(a-\tau)$.

  因此, $\hat{\tau}=a=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ (样本方差).

习题七、9. 假设$X_1, \cdots, X_n$ i.i.d.$\sim N(\mu, \sigma^2)$, $\mu \geqslant 0$, $\sigma^2 = 1$. 求μ 的最大似然估计.

- 确定参数$\theta := \mu$, 及其范围$\Theta = [0, \infty)$.

  写出似然函数,

  $L(\mu) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$.
- 求$L(\theta)$ 的最大值点.

  $\sum_{i=1}^n (x_i - \mu)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2$.
- 注意参数范围!

  若$\bar{x} \in \Theta$ (即$\bar{x} \geqslant 0$), 则$\hat{\mu} = \bar{x}$.

  否则, $\bar{x} < 0$, 在$\mu \in \Theta$ 中取$(\bar{x} - \mu)^2$ 的最小值点, 即$\hat{\mu} = 0$.

例1.5. 假设$X_1, \cdots, X_n$ i.i.d.$\sim$ Exp(λ). 求λ 的最大似然估计.

- $L(\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda n \bar{x}}$.
- 求$\log L(\lambda) = n \log \lambda - n\lambda\bar{x}$ 的最大值点.
  $\frac{d}{d\lambda} \log L(\lambda) = n(\frac{1}{\lambda} - \bar{x})$
- $\hat{\lambda} = \frac{1}{\bar{x}}$为$\lambda$的最大似然估计.
- $\hat{\lambda}(\vec{X}) = \frac{1}{\bar{X}}$为$\lambda$的最大似然估计量.

- 注: 有时用$\theta := \frac{1}{\lambda}$做参数, $p_\theta(x) = \frac{1}{\theta} e^{-x/\theta} 1_{x \geqslant 0}$, $EX = \theta$.