

第一章 参数估计

1.1 求矩估计与最大似然估计

Remark. 矩估计与最大似然估计

矩估计

令 $EX^k = \frac{1}{n} \sum_{i=1}^n X_i^k$ 或者 $E(X - EX)^k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, k = 1, 2, \dots$ 得到 $\theta_1, \theta_2 \dots$ 的矩估计量

$$\begin{cases} EX = \bar{X}, & \text{一个参数} \\ EX^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 & \text{两个参数} \end{cases}$$

最大似然估计

(1) 对样本点 x_1, x_2, \dots, x_n , 似然函数为 $L(\theta) \begin{cases} \prod_{i=1}^n p(x_i; \theta) \\ \prod_{i=1}^n f(x_i; \theta) \end{cases}$

(2) 似然函数两端取对数求导

(3) 令 $\frac{d \ln L(\theta)}{d\theta} = 0$ 就可以得到 θ 的最大似然估计值

一个关于规范的小提示, 如果问估计值用小写字母 (样本值), 问估计量用大写字母 (随机变量)

1. (2002, 数一) 设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $0 < \theta < \frac{1}{2}$ 为未知参数, 利用总体 X 的如下样本值 3, 1, 3, 0, 3, 1, 2, 3, 求 θ 的矩估计值与最大似然估计值。

Solution.

(矩估计) 这道题只有一个参数, 只需要用一阶矩估计 $EX = 2\theta(1-\theta) + 2\theta^2 + 3 - 6\theta = \bar{X}$, 其中 $\bar{X} = \frac{16}{8} = 2$, 故 θ 的矩估计值 $\hat{\theta} = \frac{1}{4}$

(最大似然估计) 对于样本 3, 1, 3, 0, 3, 1, 2, 3, 似然估计函数为

$$L(\theta) = 4\theta^6(1-\theta)^2(1-2\theta)^4$$

令 $\frac{d\ln\theta}{d\theta} = 0$ 有 $\theta = \frac{7+\sqrt{13}}{12}$ 又 $0 < \theta < \frac{1}{2}$, 故最终 $\theta = \frac{7-\sqrt{13}}{12}$

□

2. (2011, 数一) 设 X_1, X_2, \dots, X_n 为来自正态总体 $N(\mu, \sigma^2)$ 的简单随机样本, 其中 μ 已知, $\sigma^2 > 0$ 未知, 样本均值为 \bar{X} , 样本方差为 S^2 。

(1) 求 σ^2 的最大似然估计量 $\hat{\sigma}^2$;

(2) 求 $E(\hat{\sigma}^2)$ 与 $D(\hat{\sigma}^2)$ 。

Solution.

(1) 对于样本 X_1, \dots, X_n 其最大似然函数为

$$L(\sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

注意参数为 σ^2 , 令 $\frac{d\ln\sigma^2}{d\sigma^2} = 0$, 有 $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$

(2) 这种题优先考虑 χ^2 分布的期望与方差结论, 有题 (1) 有

$$\frac{X_i - \mu}{\sigma} \sim N(0, 1) \implies \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$$

故 $E(\hat{\sigma}^2) = \sigma^2, D(\hat{\sigma}^2) = \frac{2\sigma^4}{n}$

□

3. (2022, 数一、三) 设 X_1, X_2, \dots, X_n 为来自期望为 θ 的指数分布总体的简单随机样本, Y_1, Y_2, \dots, Y_m 为来自期望为 2θ 的指数分布总体的简单随机样本, 两个样本相互独立。利用 X_1, X_2, \dots, X_n 与 Y_1, Y_2, \dots, Y_m ,

(1) 求 θ 的最大似然估计量 $\hat{\theta}$;

(2) 求 $D(\hat{\theta})$ 。

Solution. 这是双总体, 但基本上和单总体一致, 不要被唬住了哦!

(1) 由题有 $X \sim E(\frac{1}{\theta})$, $Y \sim E(\frac{1}{2\theta})$, 故其概率密度分别为

$$f_X(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad f_Y(y) = \begin{cases} \frac{1}{2\theta} e^{-\frac{y}{2\theta}}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

则对于样本 X_1, X_2, \dots, X_n 与 Y_1, Y_2, \dots, Y_n , 最大似然估计函数为

$$L(\theta) = \left(\frac{1}{2}\right)^m \theta^{-(m+n)} e^{-\frac{1}{\theta}(\sum_{i=1}^n X_i + \frac{1}{2} \sum_{j=1}^m Y_j)}$$

则令 $\frac{d \ln \theta}{d \theta} = 0$, 有 $\hat{\theta} = \frac{1}{n+m}(\sum_{i=1}^n X_i + \frac{1}{2} \sum_{j=1}^m Y_j)$

(2)

$$\begin{aligned} D(\hat{\theta}) &= \left(\frac{1}{m+n}\right)^2 D\left(\sum_{i=1}^n X_i + \frac{1}{2} \sum_{j=1}^m Y_j\right) \\ &= \frac{\theta^2}{m+n} \end{aligned}$$

□

1.2 估计量的评价标准

Remark. 估计量的评价标准

- (1) (无偏性) 设 $\hat{\theta}$ 为 θ 的估计量, 若 $E\hat{\theta} = \theta$ 则称其为 θ 无偏估计量
- (2) (有效性) 设 $\hat{\theta}_1, \hat{\theta}_2$ 为 θ 的无偏估计, 若 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$ 则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效
- (3) 设 $\hat{\theta}$ 为 θ 的估计量, 若 $\hat{\theta}$ 依概率收敛于 θ , 则称 $\hat{\theta}$ 为 θ 一致 (相合) 估计量
一致性的考点在于 $\frac{1}{n} \sum_{i=1}^n \square \xrightarrow{P} E\square$

4. 设总体 X 的概率密度为

$$f(x) = \begin{cases} 2e^{-2(x-\theta)}, & x > \theta \\ 0, & x \leq \theta \end{cases}$$

其中 $\theta > 0$ 为未知参数, X_1, X_2, \dots, X_n 为来自总体 X 的简单随机样本。

- (1) 求 θ 的最大似然估计量 $\hat{\theta}$;
- (2) 问 $\hat{\theta}$ 是否为 θ 的无偏估计量? 并说明理由。

Solution.(1) 对于样本 X_1, X_2, \dots, X_n 的最大似然估计函数为

$$L(\theta) = \prod_{i=1}^n 2e^{-2(x_i - \theta)} = 2^n e^{-\sum_{i=1}^n (x_i - \theta)}$$

显然 $L(\theta)$ 关于 θ 是单调递增的, 则根据最大似然的定义, 应该取使得 $L(\theta)$ 最大的值, 而由题目有 $X_1 > \theta, X_2 > \theta, \dots$, 故 $\hat{\theta} = \min \{X_1, X_2, \dots, X_n\}$

(2) 由概率密度函数有 $F_X(x) = \int_{-\infty}^x f(t)dt$, 故

$$F_X(x) = \int_{-\infty}^x f(t)dt = \begin{cases} 1 - e^{-2(x-\theta)}, & x > \theta \\ 0, & x \leq \theta \end{cases}$$

故 $F_{min} = 1 - [1 - F_X(x)]^n$ 即

$$F_{min} = \begin{cases} 1 - e^{-2n(x-\theta)}, & x > \theta \\ 0, & x \leq \theta \end{cases}$$

故

$$f_{min} = \begin{cases} 2ne^{-2n(x-\theta)}, & x > \theta \\ 0, & x \leq \theta \end{cases}$$

由期望的定义有

$$E\hat{\theta} = \int_{\theta}^{+\infty} 2nxe^{-2n(x-\theta)} = \theta + \frac{1}{2n}$$

□

5. (2010, 数一) 设总体 X 的概率分布为

X	1	2	3
P	$1 - \theta$	$\theta - \theta^2$	θ^2

其中参数 $\theta \in (0, 1)$ 未知, N_i 表示来自总体 X 的简单随机样本 (样本容量为 n) 中等于 i 的个数 ($i = 1, 2, 3$) 求常数 a_1, a_2, a_3 使得 $T = \sum_{i=1}^3 a_i N_i$ 为 θ 的无偏估计量, 并求 T 的方差.

Solution. 由题可知 $N_i \sim B(n, p)$, 具体来说有

$$\begin{cases} N_1 \sim B(n, 1 - \theta) \\ N_2 \sim B(n, \theta - \theta^2) \\ N_3 \sim B(n, \theta^2) \end{cases}$$

且有 $N_1 + N_2 + N_3 = n$

故 $ET = \sum_{i=1}^3 a_i EN_i = n[a_1 + (a_2 - a_1)\theta + (a_3 - a_2)\theta^2] = \theta$, 只需要令

$$\begin{cases} a_1 = 0 \\ a_2 = \frac{1}{n} \\ a_3 = \frac{1}{n} \end{cases}$$

求 $DT = \frac{1}{n^2}D(n - N_1) = \frac{1}{n^2}DN_1 = \frac{\theta(1-\theta)}{n}$

□

1.3 区间估计与假设检验

Remark (区间估计与假设检验). 这一节内容很少, 只需要掌握置信度的概念, 假设检验的基本过程与第一类错误/第二类错误的概念即可

1. 置信度与置信区间

设总体 X 的分布函数 $F(x, \theta)$ 含有一个**未知参数** $\theta, \theta \in \Theta$ 其中 Θ 是其所有可能取值的集合, 对于给定值 $0 < \alpha < 1$, 若由来自总体 X 的样本 X_1, X_2, \dots, X_n 确定了两个统计量 $\theta_1, \theta_2, \theta_1 \leq \theta_2$ 对于 $\forall \theta \in \Theta$ 都有

$$P\{\hat{\theta}_1 < \theta < \hat{\theta}_2\} \geq 1 - \alpha$$

则称区间 (θ_1, θ_2) 为 θ 置信水平为 $1 - \alpha$ 的置信区间, $\hat{\theta}_1, \hat{\theta}_2$ 分别称置信水平为 $1 - \alpha$ 的双侧置信区间的置信下限和置信上限, $1 - \alpha$ 称为置信水平或置信度

2. 原假设 H_0 与备择假设 H_1

类型	H_0	H_1
双边检验	$\theta = \theta_0$	$\theta \neq \theta_0$
单边检验-左边	$\theta \geq \theta_0$	$\theta < \theta_0$
单边检验-右边	$\theta \leq \theta_0$	$\theta > \theta_0$

3. 假设检验的过程

- (1) 根据题意写出原假设 H_0 和备择假设 H_1
- (2) 选择检验方式, 写出检验统计量及其分布
- (3) 根据给定的显著性水平确定拒绝域
- (4) 统计检验统计量的值, 做出推断

4. 第一类错误/第二类错误

类型	含义	犯错的概率
第一类错误	原假设 H_0 为真, 但却拒绝 H_0 , 即弃真概率	$\alpha = P\{\text{拒绝}H_0 \mid H_0\text{为真}\}$
第二类错误	原假设 H_0 为假, 但却接受 H_0 , 即取伪概率	$\beta = p\{\text{接受}H_0 \mid H_0\text{不真}\}$

- (1) 仅控制犯第一类错误的检验称为显著检验, α 为显著性水平
- (2) 当样本容量固定时, α 和 β 中任意一个减少, 另一个必然增大; 如果要使 α 和 β 同时减少, 只能增大样本容量