

第七章 点估计与估计量的评价

本章导读

上一章解决了“怎样从总体取得样本、统计量具有怎样的分布”. 本章进一步追问: 总体分布的形式已知, 但参数未知时, 能否用样本给参数一个具体的估计?

本章分成两条主线:

怎样估: 矩估计、最大似然估计

估得怎样: 无偏性、有效性、一致性.

矩估计利用“样本矩接近总体矩”, 最大似然估计利用“已经出现的数据在所选参数下应尽可能容易出现”. 二者可能给出相同结果, 也可能不同; 得到估计量后, 还必须评价其质量.

7.1 点估计的基本概念

7.1.1 从未知参数到样本函数

设总体 X 的分布形式已知, 但含未知参数 θ ; $X_1, \dots, X_n$ 是来自该总体的简单随机样本. 我们的目标是构造一个不含未知参数的统计量, 用它近似 θ .

定义 7.1 (点估计、估计量与估计值)

若统计量

$$\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$$

被用来估计未知参数 θ , 则称 $\hat{\theta}$ 为 θ 的估计量. 观测到样本值 $x_1, \dots, x_n$ 后, 所得数值

$$\hat{\theta}(x_1, \dots, x_n)$$

称为 θ 的**估计值**. 直接用一个数估计参数的方法称为**点估计**.

对象	典型记号	性质
未知参数	θ	固定但未知, 不是随机变量
估计量	$\hat{\theta}(X_1, \dots, X_n)$	随样本改变, 是统计量、随机变量
估计值	$\hat{\theta}(x_1, \dots, x_n)$	代入数据后得到的确定数值

帽子 “ $\widehat{}$ ” 表示估计, 不表示它一定等于真实参数. 点估计只给出一个数, 不直接说明误差可能有多大; 下一章的区间估计将补充不确定性范围.

小结

点估计的目标是把未知参数表示成样本的函数. 书写时必须区分固定的未知参数、随机的估计量和观测后的估计值; 后续评价的对象是估计量, 而不只是某次得到的数字.

7.2 矩估计法

7.2.1 思想来源: 样本矩逼近总体矩

设总体 k 阶原点矩存在, 记

$$\mu'_k = E(X^k), \quad A_k = \frac{1}{n} \sum_{i=1}^n X_i^k.$$

由于 $X_1^k, \dots, X_n^k$ 独立同分布, 辛钦大数定律给出

$$A_k \xrightarrow{P} \mu'_k.$$

因此样本量较大时, 样本矩通常接近对应总体矩. 矩估计法把这种“接近”直接写成方程.

定义 7.2 (矩估计法)

设总体分布含 k 个未知参数 $\theta_1, \dots, \theta_k$. 选取 k 个存在的总体矩, 并令它们等于对应样本矩, 得到关于未知参数的方程组. 解出

$$\hat{\theta}_j = \widehat{\theta}_j(X_1, \dots, X_n), \quad j = 1, \dots, k,$$

称为各参数的矩估计量.

最常用的原点矩方程组为

$$\frac{1}{n} \sum_{i=1}^n X_i^r = E(X^r), \quad r = 1, \dots, k.$$

若参数本身就是期望或方差, 直接用

$$\bar{X} \approx E(X), \quad \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \approx \text{Var}(X)$$

往往更简洁. 注意矩估计方差时使用分母为 n 的二阶样本中心矩, 不是分母为 $n-1$ 的无偏样本方差.

注: [矩估计的统一步骤]

1. 数清未知参数个数, 确定需要几个独立方程;
2. 计算相应总体矩, 并写成参数的函数;
3. 令总体矩等于样本矩;
4. 解方程或方程组, 并确认结果只含样本和已知量.

使用前还要检查所选总体矩是否存在. 矩估计的合理性来自大数定律, 但不意味着它天然无偏或最有效.

7.2.2 一个参数的矩估计

例题 7.3 (连续型总体的矩估计与一致性)

总体密度为

$$f(x; \theta) = \begin{cases} \frac{\theta^2}{x^3} e^{-\theta/x}, & x > 0, \\ 0, & \text{其他}, \end{cases} \quad \theta > 0.$$

求 θ 的矩估计量, 并证明它具有一致性.

只有一个未知参数, 先计算一阶总体矩. 令 $u = \theta/x$, 可得

$$E(X) = \int_0^\infty x \frac{\theta^2}{x^3} e^{-\theta/x} dx = \theta.$$

令 $\bar{X} = E(X)$, 得到

$$\hat{\theta}_M = \bar{X}.$$

由于 $X_1, \dots, X_n$ 独立同分布且 $E(X_i) = \theta$ 存在, 辛钦大数定律给出

$$\hat{\theta}_M = \bar{X} \xrightarrow{P} \theta.$$

故 $\hat{\theta}_M$ 是 θ 的一致估计量.

方法选择: 一个参数优先尝试一阶矩方程. **结果检查:** θ 与 X 具有相同量纲, $\bar{X}$ 的量纲也相同.

例题 7.4 (从矩估计量到矩估计值)

设 $0 < \theta < 1/2$, 总体分布为

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

样本值为 3, 1, 3, 0, 3, 1, 2, 3, 求 θ 的矩估计值.

总体期望为

$$E(X) = 2\theta(1-\theta) + 2\theta^2 + 3(1-2\theta) = 3 - 4\theta.$$

令 $\bar{X} = 3 - 4\theta$, 得矩估计量

$$\hat{\theta}_M = \frac{3 - \bar{X}}{4}.$$

本组数据的样本均值为 $\bar{x} = 16/8 = 2$, 所以矩估计值为

$$\hat{\theta}_M = \frac{3 - 2}{4} = \frac{1}{4}.$$

结果检查: $1/4$ 落在参数空间 $(0, 1/2)$ 内.

7.2.3 两个参数的矩估计

例题 7.5 (正态总体两个未知参数)

设 $X_1, \dots, X_n$ 来自 $N(\mu, \sigma^2)$, μ, σ^2 均未知. 求二者的矩估计量.

可以令前两个原点矩相等:

$$\bar{X} = E(X) = \mu, \quad \frac{1}{n} \sum_{i=1}^n X_i^2 = E(X^2) = \sigma^2 + \mu^2.$$

解得

$$\hat{\mu}_M = \bar{X}, \quad \hat{\sigma}_M^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

也可直接用样本一阶原点矩估计期望、样本二阶中心矩估计方差, 结果相同.

易错提醒: 这里 $\hat{\sigma}_M^2$ 的分母为 n , 它一般不是无偏估计; 矩估计法追求的是矩方程匹配, 不自动保证无偏性.

小结

矩估计把样本矩与总体矩对应相等, 再解参数方程. 一个参数通常先用 $\bar{X} = E(X)$; 多个参数需要同样多个独立方程. 使用时先检查矩存在, 再检查解是否只含样本量, 并区分二阶样本中心矩与无偏样本方差.

7.3 最大似然估计法

7.3.1 核心思想: 让已观测数据尽可能“像是会出现的”

固定一组已经观测到的样本值 $x_1, \dots, x_n$. 不同参数会给这组数据不同的可能程度. 最大似然法选择使该组数据的似然函数最大的参数.

定义 7.6 (似然函数)

若总体为离散型, 概率质量函数为 $p(x; \theta)$, 则

$$L(\theta) = \prod_{i=1}^n p(x_i; \theta).$$

若总体为连续型, 密度为 $f(x; \theta)$, 则

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta).$$

$L(\theta)$ 称为样本观测值的**似然函数**.

连续型情形中, 单点概率为零, 所以 $L(\theta)$ 不是“恰好取这些点的概率”; 密度乘积反映的是这些观测点附近出现的相对可能程度. 似然函数中数据已固定, 变量是参数 θ .

定义 7.7 (最大似然估计)

若参数值 $\hat{\theta}$ 使

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta),$$

则称 $\hat{\theta}$ 为 θ 的**最大似然估计值**. 把小写样本值换回随机样本, 所得统计量称为**最大似然估计量**, 常记为 $\hat{\theta}_L$.

因对数函数严格递增, 最大化 $L(\theta)$ 与最大化对数似然

$$\ell(\theta) = \ln L(\theta)$$

等价. 取对数可把连乘变成求和, 便于求导.

注: [最大似然的标准流程]

1. 写清参数空间与样本必须满足的取值条件;
2. 构造似然函数 $L(\theta)$;
3. 取对数得到 $\ell(\theta)$;
4. 求驻点, 并结合参数空间、二阶导数或单调性判断最大值;
5. 把样本值 x_i 换成随机样本 X_i , 写出估计量.

若支持集依赖参数, 最大值可能出现在边界, 不能只会令导数为零.

7.3.2 连续型总体: 对数求导

例题 7.8 (连续型总体的最大似然估计)

仍设

$$f(x; \theta) = \frac{\theta^2}{x^3} e^{-\theta/x}, \quad x > 0, \text{quad} \theta > 0.$$

求 θ 的最大似然估计量.

对一组正的样本值 $x_1, \dots, x_n$,

$$L(\theta) = \prod_{i=1}^n \frac{\theta^2}{x_i^3} e^{-\theta/x_i} = \frac{\theta^{2n}}{\prod_{i=1}^n x_i^3} \exp\left(-\theta \sum_{i=1}^n \frac{1}{x_i}\right).$$

取对数:

$$\ell(\theta) = 2n \ln \theta - 3 \sum_{i=1}^n \ln x_i - \theta \sum_{i=1}^n \frac{1}{x_i}.$$

令一阶导数为零,

$$\ell'(\theta) = \frac{2n}{\theta} - \sum_{i=1}^n \frac{1}{x_i} = 0,$$

得到

$$\hat{\theta}_L = \frac{2n}{\sum_{i=1}^n 1/X_i}.$$

并且 $\ell''(\theta) = -2n/\theta^2 < 0$, 所以该驻点确为唯一最大值点.

关键转折: 矩估计计算 $E(X)$; 最大似然估计不积分, 而是把密度值连乘后对参数求极值.

7.3.3 离散型总体: 按取值频数归类

例题 7.9 (同一总体的矩估计与最大似然估计可以不同)

对前述分布

$$P(X=0) = \theta^2, \text{quad} P(X=1) = 2\theta(1-\theta),$$

$$P(X=2) = \theta^2, \text{quad} P(X=3) = 1-2\theta, \text{quad} 0 < \theta < \frac{1}{2},$$

给定样本值 3, 1, 3, 0, 3, 1, 2, 3, 求最大似然估计值.

样本中 0, 1, 2, 3 的频数依次为 1, 2, 1, 4, 故

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

于是

$$\ell'(\theta) = \frac{6}{\theta} - \frac{2}{1-\theta} - \frac{8}{1-2\theta} = 0.$$

解得

$$\theta = \frac{7 \pm \sqrt{13}}{12}.$$

其中带加号的根不在 $(0, 1/2)$ 内, 故

$$\hat{\theta}_L = \frac{7 - \sqrt{13}}{12}.$$

结果检查: 必须用参数空间筛根. 本题矩估计值为 $1/4$, 与最大似然估计值不同, 这不表示其中一个计算错误, 而是两种原则不同.

7.3.4 分段密度与隐藏约束

例题 7.10 (先用密度规范性减少参数)

总体密度为

$$f(x; \alpha, \beta) = \begin{cases} \alpha, & -1 < x < 0, \\ \beta, & 0 < x < 1, \\ 0, & \text{其他,} \end{cases}$$

样本值为 $-0.5, 0.3, -0.2, -0.6, -0.1, 0.4, 0.5, -0.8$. 求 α 的最大似然估计值.

由密度规范性, $\alpha + \beta = 1$, 故 $\beta = 1 - \alpha$, 参数空间为 $0 < \alpha < 1$. 样本中有 5 个负值、3 个正值, 因此

$$L(\alpha) = \alpha^5(1-\alpha)^3.$$

令

$$\ell'(\alpha) = \frac{5}{\alpha} - \frac{3}{1-\alpha} = 0,$$

得

$$\hat{\alpha}_L = \frac{5}{8}.$$

方法选择: 看到多个参数, 先检查规范性等条件能否把它们联系起来; 分段密度则按样本落入各段的频数写似然.

7.3.5 支持集依赖参数: 最大值在边界

例题 7.11 (不能机械地令导数为零)

总体密度为

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

求 θ 的最大似然估计量.

记 $x_{(1)} = \min(x_1, \dots, x_n)$. 只有当 $\theta < x_{(1)}$ 时似然为正, 并且

$$L(\theta) = 2^n \exp \left[-2 \sum_{i=1}^n x_i + 2n\theta \right].$$

在允许范围内, $L(\theta)$ 随 θ 严格递增, 所以最优参数应尽量靠近上边界 $x_{(1)}$. 按本课程的边界约定, 最大似然估计量写为

$$\hat{\theta}_L = X_{(1)} = \min(X_1, \dots, X_n).$$

严谨说明: 题面使用严格不等式 $x > \theta$ 时, $\theta = x_{(1)}$ 处似然为零, 严格地说只有当 $\theta \uparrow x_{(1)}$ 时取得上确界; 若支持集写成 $x \geq \theta$, 最大值就在边界取得. 考试按课程结论写 $X_{(1)}$, 同时理解这个边界细节.

7.3.6 两个未知参数与不变性

定理 7.12 (正态总体的最大似然估计)

设 $X_1, \dots, X_n$ 来自 $N(\mu, \sigma^2)$, 且 μ, σ^2 均未知, 则

$$\hat{\mu}_L = \bar{X}, \quad \hat{\sigma}_L^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

证明: 证明目标: 对正态似然的两个未知参数分别求偏导并解方程组.

出发点: 固定样本值后,

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

关键转换: 取对数, 把 σ^2 作为一个整体求偏导.

$$\ell(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

似然方程为

$$\begin{aligned} \frac{\partial \ell}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0, \\ \frac{\partial \ell}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \mu)^2 = 0. \end{aligned}$$

先由第一式得 $\hat{\mu} = \bar{x}$, 再代入第二式得

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

换回随机样本即得结论.

证明小结: 两个参数对应两个偏导方程, 通常先解较简单的一个再回代. 方差的最大似然估计分母为 n , 与矩估计相同, 但不是无偏样本方差 S^2 . ■

定理 7.13 (最大似然估计的不变性)

若 $\hat{\theta}_L$ 是 θ 的最大似然估计, 且 g 为一函数, 则 $g(\hat{\theta}_L)$ 是 $g(\theta)$ 的最大似然估计.

例如, 若 $\hat{\theta}_L = \bar{X}$, 则 $3\bar{X} + 1$ 是 $3\theta + 1$ 的最大似然估计. 矩估计一般没有这样的不变性, 不能不经重新列矩方程就直接套用.

小结

最大似然估计固定观测数据, 把似然看成参数的函数并求最大值. 常规题用“写似然—取对数—求导—筛选参数空间—换回大写样本”; 分段密度要按频数归类, 支持集依赖参数时要检查单调性和边界, 不能只寻找驻点.

7.4 估计量的评价标准

得到估计量并不代表估计质量已经足够好. 本课程用无偏性、有效性和一致性从三个角度评价.

7.4.1 无偏性: 长期平均不系统偏高或偏低

定义 7.14 (无偏估计)

若对任意 $\theta \in \Theta$, 估计量 $\hat{\theta}$ 的期望存在且

$$E_{\theta}(\hat{\theta}) = \theta,$$

则称 $\hat{\theta}$ 是 θ 的无偏估计量.

无偏性描述重复抽样后估计量的平均位置, 不保证某一次估计一定接近真值, 也不保证方差小.

例题 7.15 (证明矩估计量无偏)

总体密度为

$$f(x; \theta) = \begin{cases} \frac{6x}{\theta^3}(\theta - x), & 0 < x < \theta, \\ 0, & \text{其他,} \end{cases} \quad \theta > 0.$$

求 θ 的矩估计量, 并判断其是否无偏.

计算总体期望:

$$E(X) = \int_0^{\theta} x \frac{6x}{\theta^3} (\theta - x) dx = \frac{\theta}{2}.$$

令 $\bar{X} = \theta/2$, 得

$$\hat{\theta} = 2\bar{X}.$$

于是

$$E(\hat{\theta}) = 2E(\bar{X}) = 2E(X) = \theta.$$

所以 $2\bar{X}$ 是 θ 的无偏估计量.

证明模式: 先明确待证等式 $E(\hat{\theta}) = \theta$, 再利用样本均值期望等于总体期望; 不要把“由矩估计得到”直接当成无偏的理由.

7.4.2 有效性: 同为无偏时比较波动

定义 7.16 (有效性比较)

设 $\hat{\theta}_1, \hat{\theta}_2$ 都是 θ 的无偏估计量, 方差均存在. 若

$$\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2),$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.

例题 7.17 (三个无偏估计量的有效性比较)

设 X_1, X_2, X_3 是来自期望为 μ 、方差为 σ^2 的总体的简单随机样本. 比较

$$\hat{\mu}_1 = \bar{X}, \quad \hat{\mu}_2 = \frac{1}{2}X_1 + \frac{1}{3}X_2 + \frac{1}{6}X_3, \quad \hat{\mu}_3 = X_1$$

的有效性.

三组系数之和均为 1, 故三者期望均为 μ , 都是无偏估计. 利用样本独立性,

$$\begin{aligned} \text{Var}(\hat{\mu}_1) &= \frac{\sigma^2}{3}, \\ \text{Var}(\hat{\mu}_2) &= \left(\frac{1}{4} + \frac{1}{9} + \frac{1}{36}\right)\sigma^2 = \frac{7}{18}\sigma^2, \\ \text{Var}(\hat{\mu}_3) &= \sigma^2. \end{aligned}$$

因为

$$\frac{1}{3} < \frac{7}{18} < 1,$$

所以 $\hat{\mu}_1$ 最有效, $\hat{\mu}_3$ 最无效.

关键条件: 先确认无偏, 再比较方差; 线性组合的方差能写成系数平方和, 是因为样本相互独立.

7.4.3 一致性: 样本增多时逼近真值

定义 7.18 (一致估计量)

若对任意 $\theta \in \Theta$, 当 $n \rightarrow \infty$ 时

$$\hat{\theta}_n \xrightarrow{P} \theta,$$

则称 $\hat{\theta}_n$ 是 θ 的一致估计量, 也称相合估计量.

证明: 证明目标: 说明只要总体期望 $E(X) = \mu$ 存在, 样本均值就是 μ 的一致估计量.

出发点: 简单随机样本 $X_1, X_2, \dots$ 独立同分布, 且共同期望为 μ .

关键转换: 把 $\bar{X}_n$ 识别为独立同分布随机变量的算术平均.

由辛钦大数定律,

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} E(X) = \mu.$$

因此 $\bar{X}_n$ 是 μ 的一致估计量.

证明小结: 证明一致性不是计算某个固定 n 下的期望, 而是研究 $n \rightarrow \infty$ 的概率极限. 遇到样本均值或样本矩, 应优先检查大数定律条件. ■

标准	数学条件	回答的问题
无偏性	$E(\hat{\theta}) = \theta$	重复抽样的平均位置是否正确
有效性	无偏前提下方差更小	哪个无偏估计量波动更小
一致性	$\hat{\theta}_n \xrightarrow{P} \theta$	样本量增大后是否逼近真值

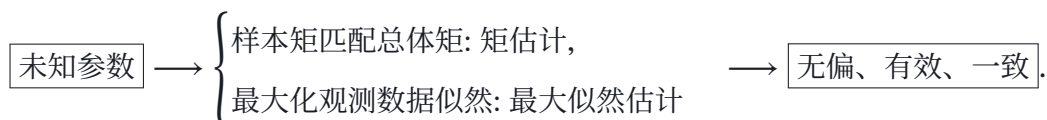
这三种标准不能互相替代: 无偏估计量可能方差很大; 有偏估计量也可能一致; 有效性比较在本章定义下以双方无偏为前提.

小结

无偏性看期望, 有效性在无偏前提下看方差, 一致性看样本量趋于无穷时的依概率收敛. 解题时先写对应定义式, 再选择期望运算、方差比较或大数定律, 不能只凭“看起来接近”下结论.

本章总结

一、知识结构



二、两种点估计方法对比

比较项	矩估计	最大似然估计
核心依据	样本矩逼近总体矩	已观测样本的似然应尽可能大
主要计算	求总体矩、列方程组	写似然、取对数、求极值
前提关注	所用总体矩必须存在	参数空间、支持集和独立性
多参数处理	建立同样多个矩方程	建立同样多个似然偏导方程
常见陷阱	把分母 n 写成 $n - 1$	忽略边界、参数范围或分段取值
性质	一般无不变性	一一变换下有不变性

三、方法选择与检查顺序

- 矩估计:** 数参数、选矩、算总体矩、列等式、解参数、查参数范围.
- 最大似然:** 定数据与参数角色、写支持条件、构造似然、取对数、求极值、筛参数空间、换回大写样本.
- 无偏性:** 直接计算 $E(\hat{\theta})$, 与 θ 比较.
- 有效性:** 先验证所有候选量无偏, 再比较方差.
- 一致性:** 写成依概率收敛目标, 优先寻找样本均值或样本矩并调用大数定律.

四、重点辨析与易错点

- 估计量是随机变量, 估计值是一次观测得到的数, 真实参数则固定但未知.
- 似然 $L(\theta; x_1, \dots, x_n)$ 与联合密度形式相同, 但用途不同: 数据固定, 参数变化.
- 最大似然的驻点只是候选点; 还要检查参数空间、二阶导数、端点和单调性.
- 连续型样本的似然是密度乘积, 不是若干单点概率的乘积.
- 正态方差的矩估计和最大似然估计均除以 n ; 无偏样本方差除以 $n - 1$.
- 无偏不等于误差小, 一致也不等于每个有限样本下都无偏.

五、证明掌握清单

- **A 级, 应能独立复现:** 根据矩估计思想建立方程并求解; 为离散型和连续型总体构造似然并处理驻点或单调边界; 用期望证明无偏; 在无偏前提下用方差比较有效性; 用大数定律证明一致性.
- **B 级, 应理解证明思路:** 样本矩为何能逼近总体矩; 最大似然为什么选择使已观察样本最可能出现的参数.
- **C 级, 知道结论即可:** 最大似然估计的一般相合性和渐近正态性; 充分性、完备性与克拉默-拉奥下界的证明. 这些内容本阶段不要求复现.

六、自测题

练习 7.19

用自己的话区分参数 θ 、估计量 $\hat{\theta}$ 和估计值, 并说明为什么评价无偏性时研究的是估计量.

练习 7.20

设总体 $X \sim P(\lambda)$, 求 λ 的矩估计量, 并证明它无偏且一致.

练习 7.21

设总体在区间 $(0, \theta)$ 上服从均匀分布. 分别求 θ 的矩估计量和最大似然估计量, 并说明最大似然为何来自样本最大值这一边界.

练习 7.22

设 $X_1, \dots, X_n$ 来自 $N(\mu, \sigma^2)$, μ, σ^2 均未知. 写出二者的矩估计和最大似然估计, 并判断 $\hat{\sigma}^2$ 是否无偏.

练习 7.23

若 X_1, X_2 独立同分布, $E(X_i) = \mu$, $\text{Var}(X_i) = \sigma^2$, 比较

$$\hat{\mu}_1 = \frac{X_1 + X_2}{2}, \quad \hat{\mu}_2 = \frac{3X_1 + X_2}{4}$$

的无偏性与有效性.

练习 7.24

某似然函数在参数空间内没有驻点, 但对参数严格递增. 说明下一步应检查什么, 为什么不能直接断言“最大似然估计不存在”.

完成自测后, 应能概括本章的统一逻辑: 先用样本构造参数估计量, 再分别从平均位置、波动大小和大样本极限评价它.