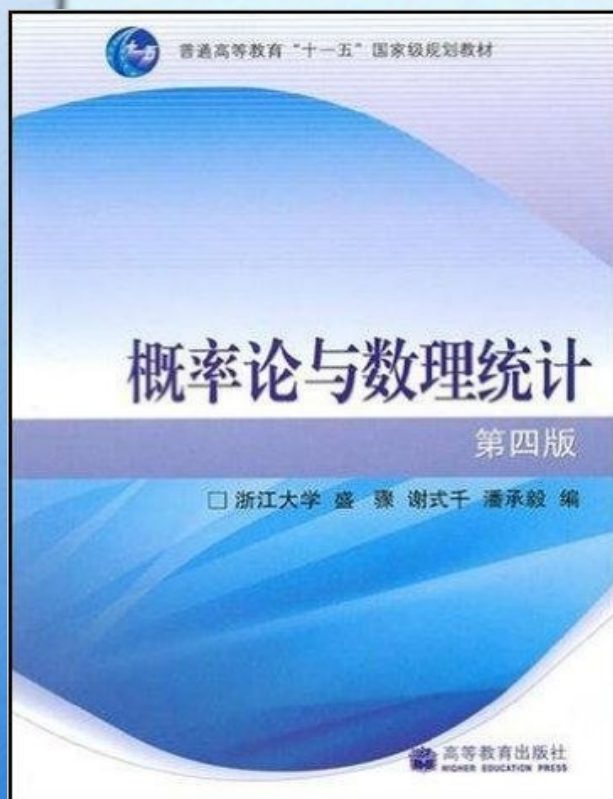


目 录



- 第一章 概率论的基本概念
- 第二章 随机变量及其分布
- 第三章 多维随机变量及其分布
- 第四章 随机变量的数字特征
- 第五章 大数定律及中心极限定理
- 第六章 样本及抽样分布
- 第七章 参数估计
- 第八章 假设检验
- 第九章 方差分析及回归分析

第七章 参数估计

第一节 点估计

第二节 基于截尾样本的最大似然估计

第三节 估计量的评选标准

第四节 区间估计

第五节 正态总体均值与方差的区间估计

第六节 $(0-1)$ 分布参数的区间估计

第七节 单侧置信区间

第一节 点估计

一、点估计问题的提法

二、估计量的求法

三、小结

矩估计思想:用样本矩估计总体矩；再借助待估计参数和总体矩的等式关系估计出待估计参数的点估计。

最大似然估计思想
一次试验就出现的事件
有较大的概率.

引例

某位同学与一位猎人一起外出打猎。

一只野兔从前方窜过。

只听一声枪响，野兔应声倒下。



如果要你推测，
是谁打中的呢？
你是如何想的呢？

因为只发一枪便打中，猎人射击的水平高，即一发子弹打中兔子概率大于这位同学。

看来这一枪是猎人射中的。

其概率模型为

令 X 为打一枪的中弹数，命中的概率为 p 则 $X \sim B(1, p)$, p 未知。

设想事先知道 p 只有两种可能：

$p=0.9$ (猎人命中率) 或 $p=0.1$ (同学命中率)

两人中有一人打枪，估计这一枪是谁打的，
即估计总体 X 的参数 p 的值

当兔子中弹，即 $\{X=1\}$ 发生了

若 $p=0.9$ ，则 $P\{X=1\}=0.9$

若 $p=0.1$ ，则 $P\{X=1\}=0.1$

现有样本观测值 $x=1$ ，什么样的参数使该样本值出现的可能性最大呢？

则选 $p=0.9$ ，即是猎人开的枪。

2. 最大似然估计法

似然函数的定义

(1) 总体 X 属离散型

若总体 X 属离散型, 设分布律

$P\{X = x\} = p(x; \theta), \theta \in \Theta$ 的形式为已知, θ 为待估参数, Θ 是 θ 可能的取值范围. 设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的样本, 则 $X_1, X_2, \dots, X_n$ 的联合分布律为 $\prod_{i=1}^n p(x_i; \theta)$.

又设 $x_1, x_2, \dots, x_n$ 为相应于样本 $X_1, X_2, \dots, X_n$ 的一个样本值 .

则样本 $X_1, X_2, \dots, X_n$ 取到观察值 $x_1, x_2, \dots, x_n$ 的概率,

即 $\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$ 发生的概率为

$$L(\theta) = L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n p(x_i; \theta), \quad \theta \in \Theta$$

$L(\theta)$ 称为样本的似然函数 .

最大似然估计法

得到样本值 $x_1, x_2, \dots, x_n$ 时, 选取使似然函数 $L(\theta)$

取得最大值的 $\hat{\theta}$ 作为未知参数 θ 的估计值,

即 $L(x_1, x_2, \dots, x_n; \hat{\theta}) = \max_{\theta \in \Theta} L(x_1, x_2, \dots, x_n; \theta)$.

(其中 Θ 是 θ 可能的取值范围)

这样得到的 $\hat{\theta}$ 与样本值 $x_1, x_2, \dots, x_n$ 有关, 记为

$\hat{\theta}(x_1, x_2, \dots, x_n)$, 参数 θ 的最大似然估计值,

$\hat{\theta}(X_1, X_2, \dots, X_n)$ 参数 θ 的最大似然估计量.

(2) 总体 X 属连续型

似然函数的定义

设概率密度为 $f(x; \theta)$, θ 为待估参数, $\theta \in \Theta$, Θ 是 θ 可能的取值范围.

设 $X_1, X_2, \dots, X_n$ 是来自 X 的样本, 则

$X_1, X_2, \dots, X_n$ 的联合密度为 $\prod_{i=1}^n f(x_i; \theta)$.

设 $x_1, x_2, \dots, x_n$ 为相应于样本 $X_1, X_2, \dots, X_n$ 的一个样本值. 则随机点 $(X_1, X_2, \dots, X_n)$ 落在点

$(x_1, x_2, \dots, x_n)$ 的邻域 (边长分别为 $dx_1, dx_2, \dots, dx_n$ 的 n 维立方体) 内的概率近似地为 $\prod_{i=1}^n f(x_i; \theta) dx_i$,

$$L(\theta) = L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta),$$

$L(\theta)$ 称为样本的似然函数 .

若 $L(x_1, x_2, \dots, x_n; \hat{\theta}) = \max_{\theta \in \Theta} L(x_1, x_2, \dots, x_n; \theta)$.

$\hat{\theta}(x_1, x_2, \dots, x_n)$ 参数 θ 的最大似然估计值,

$\hat{\theta}(X_1, X_2, \dots, X_n)$ 参数 θ 的最大似然估计量.

求最大似然估计的步骤:

(一)写出似然函数

$$L(\theta) = L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^n p(x_i; \theta)$$

$$\text{或 } L(\theta) = L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta);$$

(二)取对数

$$\ln L(\theta) = \sum_{i=1}^n \ln p(x_i; \theta) \quad \text{或} \quad \ln L(\theta) = \sum_{i=1}^n \ln f(x_i; \theta);$$

(三) 对 θ 求导 $\frac{d \ln L(\theta)}{d \theta}$, 并令 $\frac{d \ln L(\theta)}{d \theta} = 0$, 对数似然方程

解方程即得未知参数 θ 的最大似然估计值 $\hat{\theta}$.

最大似然估计法也适用于分布中含有多个未知参数的情况. 此时只需令

$$\frac{\partial}{\partial \theta_i} \ln L = 0, \quad i = 1, 2, \dots, k. \quad \text{对数似然方程组}$$

解出由 k 个方程组成的方程组, 即可得各未知参数 θ_i ($i = 1, 2, \dots, k$) 的最大似然估计值 $\hat{\theta}_i$.

例1 设 $X \sim b(1, p)$, $X_1, X_2, \dots, X_n$ 是来自 X 的一个样本, 求 p 的最大似然估计量.

解 设 $x_1, x_2, \dots, x_n$ 为相应于样本 $X_1, X_2, \dots, X_n$ 的一个样本值, X 的分布律为

$$P\{X = x\} = p^x (1-p)^{1-x}, \quad x = 0, 1.$$

故似然函数 $L(p) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i}$

$$= p^{x_1} (1-p)^{1-x_1} p^{x_2} (1-p)^{1-x_2} \cdots p^{x_n} (1-p)^{1-x_n}$$
$$= p^{\sum_{i=1}^n x_i} (1-p)^{n - \sum_{i=1}^n x_i},$$

对数似然函数

而 $\ln L(p) = \left(\sum_{i=1}^n x_i \right) \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1 - p),$

令 $\frac{d}{dp} \ln L(p) = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1 - p} = 0,$

对数似然方程

解得 p 的最大似然估计值

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$

p 的最大似然估计量为 $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}.$

这一估计量与矩估计量是相同的.

例2 设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 为未知参数, $x_1, x_2, \dots, x_n$ 是来自 X 的一个样本值, 求 μ 和 σ^2 的最大似然估计量.

解 X 的概率密度为

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

似然函数为

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

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

$$\text{令} \begin{cases} \frac{\partial}{\partial \mu} \ln L(\mu, \sigma^2) = 0, \\ \frac{\partial}{\partial \sigma^2} \ln L(\mu, \sigma^2) = 0, \end{cases}$$

对数似然方程

$$\begin{cases} \frac{1}{\sigma^2} \left[\sum_{i=1}^n x_i - n\mu \right] = 0, \\ -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \mu)^2 = 0, \end{cases}$$

由前一式解得

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

代入后一式得

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

因此 μ, σ^2 的最大似然估计量分别为

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

它们与相应的矩估计量相同.

三、小结

求最大似然估计(*MLE*)的一般步骤是:

(1) 根据样本联合分布列(或联合密度)求出 θ 的似然函数 $L(\theta)$

(2) 求对数似然函数 $\ln L(\theta)$

(3) 求对数方程（组）、下结论

在统计问题中往往先使用最大似然估计法, 在最大似然估计法使用不方便时,再用矩估计法.

问题与思考

1、问题 设总体 X 在 $[a, b]$ 上服从均匀分布, 其中 a, b 未知, $x_1, x_2, \dots, x_n$ 是来自总体 X 的一个样本值, 求 a, b 的最大似然估计量.



2、思考

参数的点估计主要有两种, 估计量不一定唯一, 如何估计量的比较的优劣?

谢 谢！