概率论与数理统计

主讲教师：荣腾中

重庆大学

1929

CHONGQING UNIVERSITY

Chapter 1

参数估计

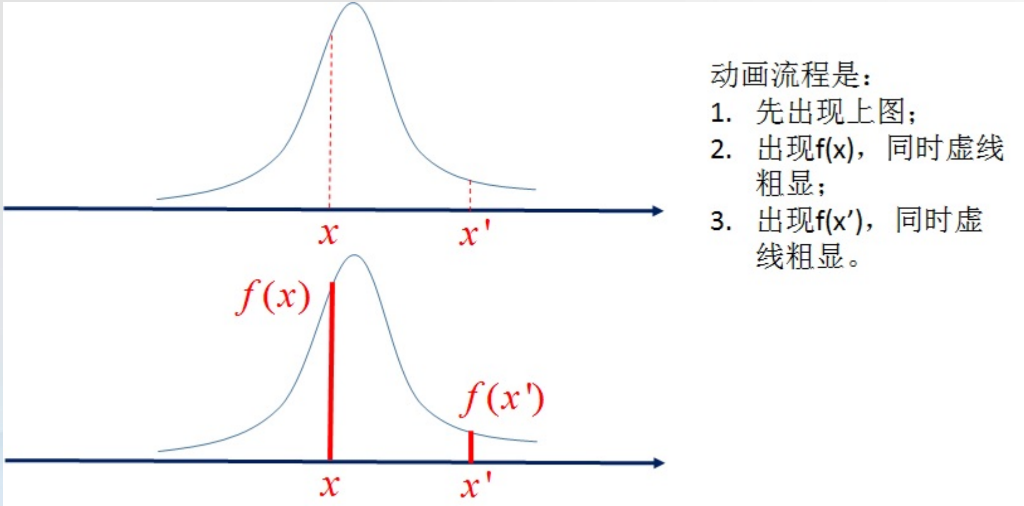
1.1 连续型分布的似然估计

在离散型分布中，似然函数为样本发生的概率为

$$L(\theta) = P\{X_1 = x_1, \cdots, X_n = x_n\} = \prod_{i=1}^{n} P\{X_i = x_i\}$$

连续型随机分布，是不是同样可行呢？

$$P\{X_1 = x_1, \cdots, X_n = x_n\} = 0$$



在 $(X_1, X_2, \cdots, X_n)$ 的联合分布中，联合密度函数为

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

由似然法思想，在联合密度函数点 $(x_1, x_2, \cdots, x_n) \in R^n$ 上

的取值应最大。记

$$L(\theta)=\prod_{i=1}^{n} f_X(x_i;\theta),\theta\in\Theta$$

为似然函数。则参数的极大似然估计 $\hat{\theta}$ 应满足

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

例 1.1.1. 设总体 X 有如下密度函数

$$f_X(x) = \begin{cases} (\theta+1)x^\theta & 0 < x < 1 \\ 0 & \text{其他} \end{cases}, \quad \theta > -1$$

求参数 θ 的极大似然估计。

解：建立似然函数

$$L(\theta) = \prod_{i=1}^{n} (\theta+1){x_i}^\theta = (\theta+1)^n \prod_{i=1}^{n} {x_i}^\theta,\ \theta > -1$$

取对数

$$\ln L(\theta) = n\ln(\theta+1) + \theta\sum_{i=1}^{n}\ln x_i$$

求导数令为 0，建立似然方程

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0 \Longrightarrow \frac{n}{\theta+1} + \sum_{i=1}^{n} \ln x_i = 0$$

$$\boldsymbol{\theta} = \frac{-n}{\sum \ln x_i} - 1$$

由于 $\frac{\partial^2 \ln L(\theta)}{\partial \theta^2} = -\frac{n}{(\theta+1)^2} < 0$，所以 θ 的极大似然估计量为

$$\hat{\theta} = \frac{-n}{\sum \ln X_i} - 1$$

它与矩估计是 $\frac{1}{1-\bar{X}} - 2$ 不同的。

例 1.1.2. 设总体 $X \sim U[0,\theta]$，求参数 θ 的极大似然估计量。

解：因为总体 X 的密度函数为

$$f_X(x) = \frac{1}{\theta},\ 0 \le x \le \theta$$

似然函数为

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n},\ 0 \le x_i \le \theta$$

取对数

$$\ln L(\theta) = -n\ln(\theta)$$

求导数，建立似然方程

$$\frac{\partial \ln L(\theta)}{\partial \theta} = -\frac{n}{\theta} = 0$$

似然方程无解，说明求偏导数不能求得极值点，则参数可能在边界处取得似然函数最大值。

回到似然函数来分析

$$L(\theta) = \frac{1}{\theta^n},\ 0 \le x_i \le \theta$$

要使似然函数极大，则 θ 要极小，而

$$0 \le x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} \le \theta$$

则当 $\theta = x_{(n)}$ 时，$L(\theta)$ 最大。所以参数的极大似然估计量为

$$\hat{\theta} = X_{(n)}$$

1.1.1 极大似然估计的优缺点

极大似然估计克服了矩法估计的一些缺点：

优点
- ① 利用了总体的分布信息
- ② 不要求总体矩一定存在
- ③ 对样本容量没有要求

缺点
- ① 似然方程可能无解，需要讨论
- ② 似然方程可能非常复杂，只能求数值解获得估计值