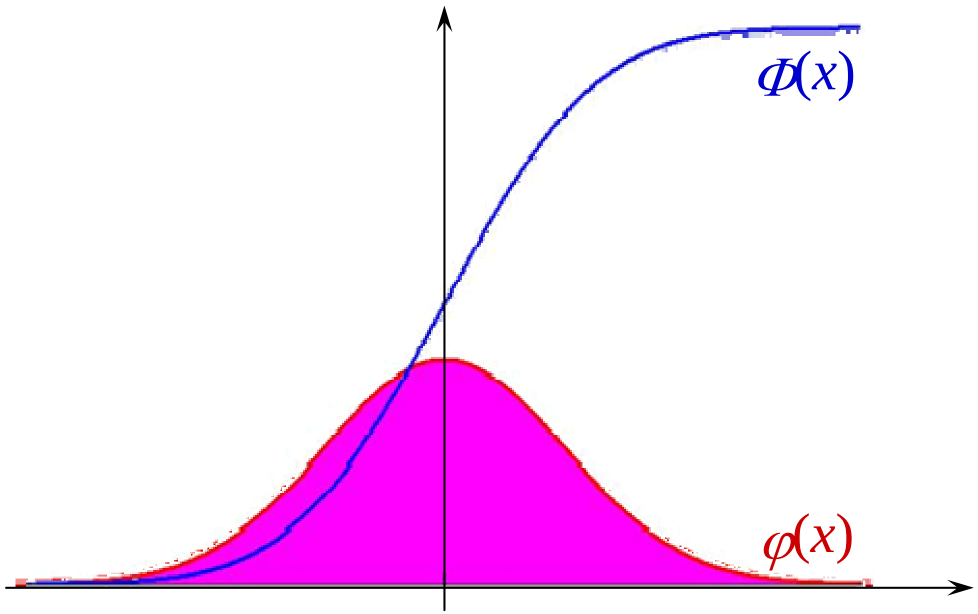
湖北省精品课程

概率论与数理统计



华中科技大学 概率统计系

叶 鹰 副教授

7.2.2 极大似然估计法

引例（P_{118}例7.5）袋中放有黑球和白球共4个，今有放回地抽球三次，得到2次白球，1次黑球，试问如何估计袋中白球个数？

解 按矩估计：总体 $X \sim B(1, m/4)$，m为未知参数，

x_1, x_2, x_3为样本观察值，$\alpha_1 = E(X) = m/4 \Longrightarrow m = 4\alpha_1$

$$\hat{m} = 4\bar{x} = 4 \times \frac{1}{3}(1+1+0) = \frac{8}{3}$$

7.2.2 极大似然估计法

引例（P_{118}例7.5）袋中放有黑球和白球共4个，今有放回地抽球三次，得到2次白球，1次黑球，试问如何估计袋中白球个数?

解 设袋中的白球数为m，记$p=m/4$，X为抽到的白球数。

则 $X \sim B(3, p)$

当$\boldsymbol{m = 3}$时，

$\boldsymbol{P(X = 2)}$**最大，**

故 $\hat{m}=3$ 。

m	p \ X	0	1	2	3
0	0	1	0	0	0
1	1/4	27/64	27/64	9/64	1/64
2	2/4	8/64	24/64	24/64	8/64
3	3/4	1/64	9/64	27/64	27/64
4	1	0	0	0	1

极大似然原则： 已发生的事件，其概率应该最大。故未知参数的选择应有利于该事件的发生。

似然函数：

$$L(x_1,x_2,\cdots,x_n;\theta)=\begin{cases}\prod_{i=1}^{n} f(x_i;\theta) & \text{其中} f(x;\theta)\text{为}X\text{的密度函数}\\ \prod_{i=1}^{n} P_\theta(X=x_i) & \text{其中} P_\theta(X=x)\text{为}X\text{的分布律}\end{cases}$$

θ 的**极大似然估计值** $\hat{\theta}(x_1,x_2,\cdots,x_n)$：

$$L(x_1,x_2,\cdots,x_n;\hat{\theta})=\max_{\theta} L(x_1,x_2,\cdots,x_n;\theta)$$

θ 的**极大似然估计量** $\hat{\theta}(X_1,X_2,\cdots,X_n)$

通常由对数似然方程 $\dfrac{\partial \ln L(\theta)}{\partial \theta_i}=0 \quad i=1,2,\cdots,k$ 解出 θ_i

例1（P_{120}例7.6）设寿命$X \sim E(1/\theta)$，试未知参数θ的极大似然估计量。

解 $L(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n}(\frac{1}{\theta}e^{-\frac{x_i}{\theta}})\qquad x_i>0,\quad i=1,2,\cdots,n$

$$=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i}$$

$$\frac{\partial}{\partial\theta}\ln L=\frac{\partial}{\partial\theta}[-n\ln\theta-\frac{1}{\theta}n\overline{x}]\ =-\frac{n}{\theta}+\frac{n}{\theta^2}\overline{x}=0$$

解得 估计值 $\hat{\theta}=\overline{x}$

估计量 $\hat{\theta}=\overline{X}$ 与矩估计相同。

例2（P$_{121}$例7.8）设总体X~$N(\mu,\sigma^2)$，求μ和σ^2的极大似然估计。

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

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

$$\begin{cases}\dfrac{\partial}{\partial\mu}\ln L=\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=\dfrac{1}{\sigma^2}(n\bar{x}-n\mu)=0\\[2ex]\dfrac{\partial}{\partial\sigma^2}\ln L=-\dfrac{n}{2}\cdot\dfrac{1}{\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2=0\end{cases}$$

$\Rightarrow$ $\hat{\mu}=\bar{x}$， $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\tilde{S}^2$ 与矩估计相同。

例3（P_{174}例7.8）设总体 $X \sim U(a,b)$，试求a 和b 的极大似然估计。

解

$$L(a,b)=\prod_{i=1}^{n} f(x_i;a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a \le x_1,x_2,\cdots,x_n \le b \\ 0, & \text{其它}\end{cases}$$

$\because$ 在 $a \leqslant x_1, x_2, \cdots, x_n \leqslant b$ 下

$$a\uparrow，\ b\downarrow \ \Longrightarrow \ (b-a)\downarrow \ \Longrightarrow \ L(b,a)\uparrow$$

$\because$ 当 $a=\min\limits_{1\leqslant i\leqslant n}\{x_i\}$，$b=\max\limits_{1\leqslant i\leqslant n}\{x_i\}$ 时， $L(b,a)=\max$

即a,b的极大似然估计为：$\hat{a}=X_1^*$, $\hat{b}=X_n^*$。

性质： 若 $\hat{\theta}$ 为 θ 的极大似然估计，$\boldsymbol{u=u(\theta)}$ 有反函数 $\boldsymbol{\theta=\theta(u)}$，则 $\boldsymbol{\hat{u}=u(\hat{\theta})}$ 为 $\boldsymbol{u=u(\theta)}$ 的极大似然估计。

事实上由 $$L(x_1,x_2,\ldots,x_n;\hat{\theta})=\max_{\theta\in\Theta}L(x_1,x_2,\ldots,x_n;\theta)$$

记 $\hat{u}=u(\hat{\theta})$，则 $\hat{\theta}=\theta(\hat{u})$，因此

$$L(x_1,x_2,\ldots,x_n;\theta(\hat{u}))=\max_{u\in U}L(x_1,x_2,\ldots,x_n;\theta(u))$$

例3（P_{122}例7.10） 设总体 $X\sim N(\mu,\sigma^2)$，求 σ 的极大似然估计。

解 由 $\sigma>0,\sigma=\sqrt{\sigma^2}$ 有反函数，再由例2知 $\hat{\sigma}^2=\tilde{S}^2$

故
$$\hat{\sigma}=\tilde{S}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

（练习17.4）

§7.3 估计量的评选原则

7.3.1 无偏性

若$E(\hat{\theta})=\theta$， 则称$\hat{\theta}$为θ的**无偏估计量**。

例1 证明样本均值$\overline{X}$为总体期望$\mu=E(X)$的无偏估计。

解 $$E(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=E(X)=\mu$$

一般(P_{123}例7.11) $$E(A_K)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=E(X^k)=\alpha_k$$

例2 (P$_{123}$例7.12）证明样本方差 $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 是总体方差 $D(X)=\sigma^2$ 的无偏估计。

证明 $E(S^2)=E[\dfrac{1}{n-1}(\sum_{i=1}^{n}X_i^2-n\overline{X}^2)]$

$$=\frac{1}{n-1}[nE(X_1^2)-nE(\overline{X}^2)]$$

$$\xrightarrow[E(X^2)=D(X)+(EX)^2]{D(X)=E(X^2)-(EX)^2}=\frac{1}{n-1}[n(\sigma^2+\mu^2)-n(D(\overline{X})+\mu^2)]$$

$$=\frac{1}{n-1}[n\sigma^2+n\mu^2-\sigma^2-n\mu^2)]$$

$$=\sigma^2$$

注1 例2说明$\tilde{S}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$不是$\sigma^2$的无偏估计。

$$\because \quad E(\tilde{S}^2)=E(\frac{n-1}{n}S^2)=\frac{n-1}{n}\sigma^2 \xrightarrow[n\to\infty]{} \sigma^2$$

~ 渐近无偏估计

注2 $\hat{\theta}$为θ的无偏估计，但$u(\hat{\theta})$不一定是$u(\theta)$的无偏估计。

例如：若$D(X)>0$，则

$$E[(\hat{\mu})^2]=E(\overline{X}^2)=D(\overline{X})+[E(\overline{X})]^2>[E(\overline{X})]^2=\mu^2 \text{ (习题7.13)}$$

注3 无偏估计不唯一，如X_1和$\overline{X}$均为$\mu=E(X)$的无偏估计。

事实上对任何$c_1, c_2, \cdots, c_n$，当$c_1+c_2+\cdots+c_n=1$时，

$$E(\sum_{i=1}^{n}c_iX_i)=\mu$$

7.3.2 有效性

若 $E(\hat{\theta}_1)=E(\hat{\theta}_2)=\theta$，且 $D(\hat{\theta}_1)\le D(\hat{\theta}_2)$，则称$\hat{\theta}_1$ 比 $\hat{\theta}_2$ **有效**

比如 $E(X_1)=E(\overline{X})=\mu$ ，但$D(X_1)=\sigma^2$, $D(\overline{X})=\dfrac{1}{n}\sigma^2$

故 $\hat{\mu}=\overline{X}$比 $\hat{\mu}_1=X_1$ 有效。

$\hat{\theta}_0$为θ的**最小方差无偏估计量**～

$E(\hat{\theta}_0)=\theta$ ，且对θ的一切无偏估计$\hat{\theta}$ 有:

$$D(\hat{\theta}_0)\le D(\hat{\theta})$$

例3 在总体期望$\mu=E(X)$的线性无偏估计类

$\overline{U}=\left\{\hat{\mu}=\sum_{i=1}^{n}c_iX_i \;\middle|\; \sum_{i=1}^{n}c_i=1\right\}$ 中求μ最小方差无偏估计。

解 $E(\hat{\mu})=\sum_{i=1}^{n}c_iE(X_i)=\sum_{i=1}^{n}c_i\mu=\mu$， $\hat{\mu}\in\overline{U}$

由**Cauchy-Schwarz**不等式 $(\sum_{i=1}^{n}a_ib_i)^2\le(\sum_{i=1}^{n}a_i^2)(\sum_{i=1}^{n}b_i^2)$

$$D(\hat{\mu})=\sum_{i=1}^{n}c_i^2D(X)=\frac{1}{n}(\sum_{i=1}^{n}1^2)(\sum_{i=1}^{n}c_i^2)D(X)$$

$$\ge\frac{1}{n}(\sum_{i=1}^{n}1\times c_i)^2D(X)=\frac{1}{n}D(X)$$

而$D(\overline{X})=\frac{1}{n}D(X)$，故$\hat{\mu}_0=\overline{X}$是$\mu$的最小方差线性无偏估计。

7.3.3 一致性

若 $\forall\ \varepsilon > 0$， 有 $\lim\limits_{n\to\infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1$

则称 $\hat{\theta}$ 为 θ 的一致估计。

例4 由大数定律 $\lim\limits_{n\to\infty} P(|\overline{X} - \mu| < \varepsilon) = 1$

知样本均值 $\overline{X}$ 是总体均值 μ 的一致估计。

例5 证明正态总体的样本方差 S^2 是 σ^2 的一致估计。

证 $\because \frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1)$，$E(S^2) = \sigma^2$

$$D(S^2) = \frac{\sigma^4}{(n-1)^2} D(\frac{(n-1)}{\sigma^2} S^2) = \frac{2\sigma^4}{(n-1)}$$

由切比雪夫不等式 $P(|S^2 - \sigma^2| < \varepsilon) \geq 1 - \dfrac{2\sigma^4}{(n-1)\varepsilon^2} \xrightarrow[n\to\infty]{} 1$

习题选讲

练习12.5 海船上雷达显示器是半径为R的一个圆，由方位物（灯塔）反射回来的信号以光点的形式均匀分布呈现在这个圆的任意一点处，求光点到圆心距离的数学期望。

解 设光点到圆心的坐标为(X,Y)，由题意知(X,Y)的密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi R^2}, & x^2+y^2\le R^2\\ 0, & \text{其他}\end{cases}$$

$$E(\sqrt{X^2+Y^2})=\iint_{x^2+y^2\le R^2}\frac{\sqrt{x^2+y^2}}{\pi R^2}dxdy$$

$$=\int_0^{2\pi}\int_0^{R}\frac{r}{\pi R^2}rdrd\theta \quad =\frac{2}{3}R$$

习题选讲

练习13.3 在长为l 的线段上任取两点，求两点间距离的期望及方差。

解I 设此两点为X和Y，则(X,Y)的联合密度函数为

$$f(x,y)=f(x)f(y)=\begin{cases}1/l^2, & 0\le x,y\le l\\ 0, & \text{其他}\end{cases}$$

$$E(|X-Y|)=\int_0^l\int_0^l |x-y|\tfrac{1}{l^2}dxdy$$

$$=\int_0^l[\int_0^x \tfrac{x-y}{l^2}dy]dx+\int_0^l[\int_0^y \tfrac{y-x}{l^2}dx]dy=\frac{l}{3}$$

$$E(|X-Y|^2)=\int_0^l\int_0^l (x-y)^2\tfrac{1}{l^2}dxdy=\frac{l^2}{6}$$

$$D(|X-Y|)=\frac{l^2}{6}-(\frac{l}{3})^2=\frac{l^2}{18}$$

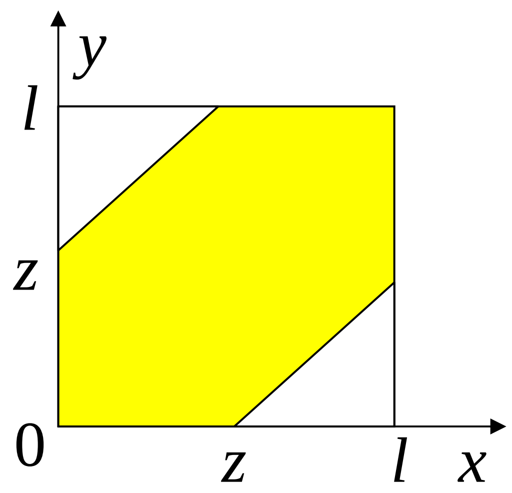
习题选讲

练习13.3 在长为l 的线段上任取两点，求两点间距离的期望及方差。

解II 设此两点为X和Y，则 $Z=|X-Y|$ 的分布函数为

$$F_Z(z)=P(|X-Y|\le z)=\iint_{|x-y|\le z} f(x,y)dxdy$$

$$=\begin{cases} 0, & z<0 \\ [l^2-(l-z)^2]/l^2, & 0\le z\le l \\ 1, & z>1 \end{cases}$$



$$f_Z(z)=F_Z'(z)=2(l-z)/l^2 \qquad 0\le z\le 1$$

$$E(Z)=\int_0^1 z\times 2(l-z)/l^2dz=l/3$$

$$E(Z^2)=\int_0^1 z^2\times 2(l-z)/l^2dz=l^2/6$$

$$D(Z)=l^2/18$$

习题选讲

习题4.7 设一个试验有m个等可能的结果，求至少一个结果连接发生k次的独立试验的期望次数。

解 设X_k为有一个结果接连发生k次的实验次数，

记$E_k=E(X_k)$。则

$$X_k = X_{k-1} + [1\times\frac{1}{m} + (1-\frac{1}{m})X_k]$$

$$E_k = E_{k-1} + 1\times\frac{1}{m} + (1-\frac{1}{m})E_k \Rightarrow E_k = mE_{k-1} + 1$$

$$E_1 = 1 \Rightarrow E_k = 1 + m + m^2 + \cdots + m^{k-1} = \frac{m^k - 1}{m-1} \qquad (m>1)$$