

## □ 点估计的极大似然估计法

思想方法：一次试验就出现的事件有较大的概率

例如：有两个外形相同的箱子，都装有100个球

第一箱 99个白球， 1个红球

第二箱 1个白球， 99个红球

现从两箱中任取一箱，并从箱中任取一球，  
结果所取得的球是白球。

**问** 所取的球来自哪一箱？

**答** 第一箱。

**例6** 设总体  $X$  服从0-1分布, 且  $P(X = 1) = p$ ,  
用极大似然法求  $p$  的估计值。

**解**  $X$  的概率分布为

$$P(X = 1) = p,$$

$$P(X = 0) = 1 - p,$$

可以写成

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

设  $X_1, X_2, \dots, X_n$  为总体  $X$  的样本,

设  $x_1, x_2, \dots, x_n$  为总体  $X$  的样本值,

则

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$$= P(X_1 = x_1) \cdot P(X_2 = x_2) \cdots P(X_n = x_n)$$

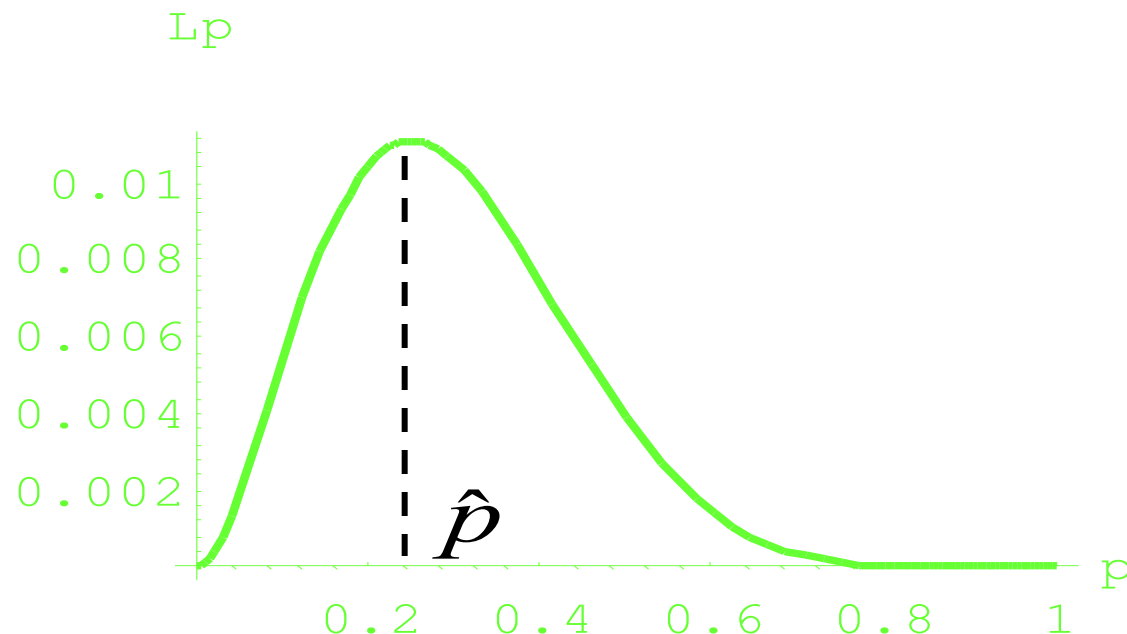
$$= p^{x_1} (1-p)^{1-x_1} \cdot p^{x_2} (1-p)^{1-x_2} \cdots p^{x_n} (1-p)^{1-x_n}$$

$$= p^{\sum_{i=1}^n x_i} (1-p)^{n - \sum_{i=1}^n x_i}$$

$$\equiv L(p)$$

$$x_i = 0, 1, i = 1, 2, \dots, n$$

对于不同的  $p$ ,  $L(p)$  不同, 见右下图



现经过一次试验, 事件

$$(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

发生了, 则  $p$  的取值应使这个事件发生的概率  $L(p)$  最大。

在容许的范围内选择  $p$  , 使  $L(p)$  最大

$$L(P) = p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i}$$

注意到,  $\ln L(p)$  是  $L$  的单调增函数, 故若某个  $p$  使  $\ln L(p)$  最大, 则这个  $p$  必使  $L(p)$  最大。

$$\frac{d\ln L}{dp} = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} \stackrel{\text{令}}{=} 0 \Rightarrow \hat{p} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

$$\left( \frac{d^2 \ln L}{dp^2} = -\frac{\sum_{i=1}^n x_i}{p^2} - \frac{n - \sum_{i=1}^n x_i}{(1-p)^2} < 0 \right)$$

所以  $\hat{p} = \bar{x}$  为所求  $p$  的估计值.

一般地, 设 $X$ 为离散型随机变量, 其分布律为

$$P(X = x) = f(x, \theta), \quad x = u_1, u_2, \dots, \theta \in \Theta$$

$X_1, X_2, \dots, X_n$ 为总体 $X$ 的样本,

$x_1, x_2, \dots, x_n$ 为总体 $X$ 的样本值,

则 $X_1, X_2, \dots, X_n$ 的概率分布为

$$\begin{aligned} & P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ &= f(x_1, \theta) f(x_2, \theta) \cdots f(x_n, \theta) \end{aligned}$$

记为

$$= L(x_1, x_2, \dots, x_n, \theta) \stackrel{\text{或}}{=} L(\theta) \quad x_i = u_1, u_2, \dots,$$

$$i = 1, 2, \dots, n, \theta \in \Theta$$

称 $L(\theta)$ 为样本的似然函数

当给定一组样本值时,  $L(\theta)$  就是参数  $\theta$  的函数, 极大似然估计法的思想就是:

选择适当的  $\theta = \hat{\theta}$ , 使  $L(\theta)$  取最大值, 即

$$\begin{aligned} & L(x_1, x_2, \dots, x_n, \hat{\theta}) \\ &= \max_{\theta \in \Theta} \{ f(x_1, \theta) f(x_2, \theta) \cdots f(x_n, \theta) \} \end{aligned}$$

则称这样得到的  $\hat{\theta}$

$$\hat{\theta} = g(x_1, x_2, \dots, x_n)$$

为参数  $\theta$  的极大似然估计值

称统计量

$$\hat{\theta} = g(X_1, X_2, \dots, X_n)$$

为参数  $\theta$  的极大似然估计量



**注1** 若随机变量 $X$ 连续, 取 $f(x_i, \theta)$ 为 $X_i$ 的密度函数

似然函数为 
$$L(\theta) = \prod_{i=1}^n f(x_i, \theta)$$

**注2** 未知参数个数可以不止一个, 如 $\theta_1, \theta_2, \dots, \theta_k$

设 $X$ 的密度函数(或分布率)为  $f(x, \theta_1, \theta_2, \dots, \theta_k)$

则定义似然函数为

$$\begin{aligned} & L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k) \\ &= L(\theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^n f(x_i, \theta_1, \theta_2, \dots, \theta_k) \end{aligned}$$

$$-\infty < x_i < +\infty, i = 1, 2, \dots, n \quad (\theta_1, \theta_2, \dots, \theta_k) \in \Theta$$

若  $L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k)$  关于  $\theta_1, \theta_2, \dots, \theta_k$  可微, 则称

$$\frac{\partial}{\partial \theta_r} L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k) = 0 \quad r = 1, 2, \dots, k$$

为似然方程组

若对于某组给定的样本值  $x_1, x_2, \dots, x_n$ , 参数  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$  使得似然函数取得最大值, 即

$$\begin{aligned} & L(x_1, x_2, \dots, x_n; \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k) \\ &= \max_{(\theta_1, \theta_2, \dots, \theta_k) \in \Theta} \{L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k)\} \end{aligned}$$

则称  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$  为  $\theta_1, \theta_2, \dots, \theta_k$  的极大似然估计值

显然,

$$\hat{\theta}_r = g_r(x_1, x_2, \dots, x_n) \quad r = 1, 2, \dots, k$$

称统计量

$$\hat{\theta}_r = g_r(X_1, X_2, \dots, X_n) \quad r = 1, 2, \dots, k$$

为  $\theta_1, \theta_2, \dots, \theta_k$  的极大似然估计量

**例7** 设总体  $X \sim N(\mu, \sigma^2)$ ,  $x_1, x_2, \dots, x_n$  是  $X$  的一组样本值, 求  $\mu, \sigma^2$  的极大似然估计.

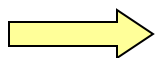
**解**  $L(x_1, x_2, \dots, x_n; \mu, \sigma^2)$

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

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

似然  
方程  
组为

$$\begin{cases} \left( \frac{\partial}{\partial \mu} \ln L \right) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \\ \left( \frac{\partial}{\partial (\sigma^2)} \ln L \right) = \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \mu)^2 - \frac{n}{2(\sigma^2)} = 0 \end{cases}$$



$$\begin{cases} \hat{\mu}_{mle} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x} \\ \hat{\sigma}_{mle}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}$$

$\mu, \sigma^2$  的极大似然估计量分别为

$$\frac{1}{n} \sum_{i=1}^n X_i = \bar{X} \qquad \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = S_n^2$$

# 求未知参数的极大似然估计值(量)的方法

1) 写出似然函数 $L$

2) 求出  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$  , 使得

$$L(x_1, x_2, \dots, x_n; \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)$$
$$= \max_{(\theta_1, \theta_2, \dots, \theta_k) \in \Theta} \{L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k)\}$$

若  $L$  是  $\theta_1, \theta_2, \dots, \theta_k$  的可微函数, 解似然方程组

$$\frac{\partial}{\partial \theta_r} L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k) = 0$$

$$r = 1, 2, \dots, k$$

可求得未知参数的极大似然估计值  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$

然后, 再求得极大似然估计量.

若  $L$  不是  $\theta_1, \theta_2, \dots, \theta_k$  的可微函数, 需用其它方法求极大似然估计值. 请看下例:

**例8** 设  $X \sim U(a, b)$ ,  $x_1, x_2, \dots, x_n$  是  $X$  的一个样本, 求  $a, b$  的极大似然估计值与极大似然估计量.

**解**  $X$  的密度函数为

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其它} \end{cases}$$

似然函数为

$$L(x_1, x_2, \dots, x_n; a, b) = \begin{cases} \frac{1}{(b-a)^n}, & a < x_i < b, \\ 0, & i = 1, 2, \dots, n \\ & \text{其它} \end{cases}$$



似然函数只有当  $a < x_i < b, i = 1, 2, \dots, n$  时才能获得最大值, 且  $a$  越大,  $b$  越小,  $L$  越大.

令 
$$x_{\min} = \min \{x_1, x_2, \dots, x_n\}$$
$$x_{\max} = \max \{x_1, x_2, \dots, x_n\}$$

取 
$$\hat{a} = x_{\min}, \quad \hat{b} = x_{\max}$$

则对满足  $a \leq x_{\min} \leq x_{\max} \leq b$  的一切  $a < b$ , 都有

$$\frac{1}{(b-a)^n} \leq \frac{1}{(x_{\max} - x_{\min})^n}$$

故  $\hat{a} = x_{\min}, \hat{b} = x_{\max}$

是  $a, b$  的极大似然估计值.

$$X_{\min} = \min\{X_1, X_2, \dots, X_n\}$$

$$X_{\max} = \max\{X_1, X_2, \dots, X_n\}$$

分别是  $a, b$  的极大似然估计量.

## 极大似然估计值的不变性原理

设  $\hat{\theta}$  是  $\theta$  的极大似然估计值,  $u(\theta)$  ( $\theta \in \Theta$ ) 是  $\theta$  的函数, 且具有单值的反函数  $\theta = \theta(u)$ ,  $u \in U$  则  $\hat{u} = u(\hat{\theta})$  是  $u(\theta)$  的极大似然估计值.

如：在正态分布总体 $N(\mu, \sigma^2)$ 中， $\sigma^2$ 的极大似然估计值为

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

$\sigma = \sqrt{\sigma^2}$  是 $\sigma^2$ 的单值函数，且具有单值的反函数，故 $\sigma$ 的极大似然估计值为

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

$\lg \sigma$ 的极大似然估计值为

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

$$P(X > t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi} \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}} e^{-\frac{(x - \bar{x})^2}{2 \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}} dx$$

矩估计就不具有这个性质.

例如 设  $X$  的密度函数为

$$f(x) = \begin{cases} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} e^{-\frac{x^2}{2\sigma^2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$E(X) = \sqrt{\frac{2}{\pi}} \sigma, \quad D(X) = \left(1 - \frac{2}{\pi}\right) \sigma^2$$

$X_1, X_2, \dots, X_n$  为总体的样本

由矩法, 令

$$E(X) = \sqrt{\frac{2}{\pi}} \sigma = \bar{X}$$

$$E(X^2) = D(X) + E^2(X) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

得 $\sigma$ 与 $\sigma^2$ 的矩法估计量为

$$\hat{\sigma} = \sqrt{\frac{\pi}{2}} \bar{X}$$

——不具有不变性

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \neq (\hat{\sigma})^2$$

**例4** 设  $X$  的密度函数为:

$$f(x; \theta) = \begin{cases} e^{-(x-\theta)}, & \theta \leq x < +\infty \\ 0, & \text{其它} \end{cases}$$

**解**

似然函数为

$$L(x_1, x_2, \dots, x_n; \theta) = \begin{cases} \sum_{i=1}^n e^{-(x_i - \theta)}, & \theta \leq x_i < +\infty, \\ 0, & \text{其它} \end{cases} \quad i = 1, 2, \dots, n$$

当  $\theta \leq x_i < +\infty$  时,  $L(x_1, x_2, \dots, x_n; \theta) = \sum_{i=1}^n e^{-(x_i - \theta)}$

是 $\theta$ 的增函数, 故 $\theta = \min\{x_1, x_2, \dots, x_n\}$ 时,



$L(x_1, x_2, \dots, x_n; \theta)$  取最大值

当  $\theta \leq x_i < +\infty$  不成立时,  $L(x_1, x_2, \dots, x_n; \theta) = 0$

故 $\theta$ 的极大似然估计为 $\min\{x_1, x_2, \dots, x_n\}$