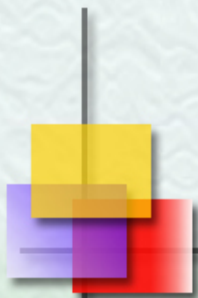


# 第一节 点估计

一、点估计问题的提法

二、估计量的求法

三、小结



# 一、点估计问题的提法

设总体  $X$  的分布函数形式已知, 但它的一个或多个参数为未知, 借助于总体  $X$  的一个样本来估计总体未知参数的值的问题称为**点估计问题**.

**例1** 在某炸药制造厂, 一天中发生着火现象的次数  $X$  是一个随机变量, 假设它服从以  $\lambda > 0$  为参数的泊松分布, 参数  $\lambda$  为未知, 设有以下的样本值, 试估计参数  $\lambda$ .



着火次数 $k$	0	1	2	3	4	5	6	
发生 $k$ 次着 火的天数 $n_k$	75	90	54	22	6	2	1	$\Sigma = 250$

解 因为  $X \sim \pi(\lambda)$ , 所以  $\lambda = E(X)$ .

用样本均值来估计总体的均值  $E(X)$ .

$$\bar{x} = \frac{\sum_{k=0}^6 kn_k}{\sum_{k=0}^6 n_k} = \frac{1}{250} (0 \times 75 + 1 \times 90 + 2 \times 54 + 3 \times 22 + 4 \times 6 + 5 \times 2 + 6 \times 1) = 1.22.$$

故  $E(X) = \lambda$  的估计为 1.22.





## 点估计问题的一般提法

设总体  $X$  的分布函数  $F(x; \theta)$  的形式为已知,  $\theta$  是待估参数.  $X_1, X_2, \dots, X_n$  是  $X$  的一个样本,  $x_1, x_2, \dots, x_n$  为相应的一个样本值.

点估计问题就是要构造一个适当的统计量  $\hat{\theta}(X_1, X_2, \dots, X_n)$ , 用它的观察值  $\hat{\theta}(x_1, x_2, \dots, x_n)$  来估计未知参数  $\theta$ .

$\hat{\theta}(X_1, X_2, \dots, X_n)$  称为  $\theta$  的估计量. } 通称估计,  
 $\hat{\theta}(x_1, x_2, \dots, x_n)$  称为  $\theta$  的估计值. } 简记为  $\hat{\theta}$ .

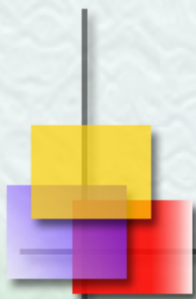


**例2** 在某纺织厂细纱机上的断头次数  $X$  是一个随机变量, 假设它服从以  $\lambda > 0$  为参数的泊松分布, 参数  $\lambda$  为未知, 现检查了150只纱锭在某一时间段内断头的次数, 数据如下, 试估计参数  $\lambda$ .

断头次数 $k$	0	1	2	3	4	5	6	
断头 $k$ 次的纱锭数 $n_k$	45	60	32	9	2	1	1	150

**解** 先确定一个统计量  $\bar{X}$ , 再计算出  $\bar{X}$  的观察值  $\bar{x}$ , 把  $\bar{x}$  作为参数  $\lambda$  的估计值.

$$\bar{x} = 1.133. \quad \lambda \text{ 的估计值为 } 1.133.$$



## 二、估计量的求法

由于估计量是样本的函数, 是随机变量, 故对不同的样本值, 得到的参数值往往不同, 如何求估计量是关键问题.

常用构造估计量的方法: (两种)

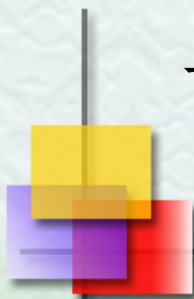
**矩估计法和最大似然估计法.**





# 1. 矩估计法

设  $X$  为连续型随机变量, 其概率密度为  $f(x; \theta_1, \theta_2, \dots, \theta_k)$ , 或  $X$  为离散型随机变量, 其分布律为  $P\{X = x\} = p(x; \theta_1, \theta_2, \dots, \theta_k)$ , 其中  $\theta_1, \theta_2, \dots, \theta_k$  为待估参数, 若  $X_1, X_2, \dots, X_n$  为来自  $X$  的样本, 假设总体  $X$  的前  $k$  阶矩存在, 且均为  $\theta_1, \theta_2, \dots, \theta_k$  的函数, 即



$$\mu_l = E(X^l) = \int_{-\infty}^{+\infty} x^l f(x; \theta_1, \theta_2, \dots, \theta_k) dx \quad (X \text{ 为连续型})$$

$$\text{或 } \mu_l = E(X^l) = \sum_{x \in R_X} x^l p(x; \theta_1, \theta_2, \dots, \theta_k), \quad (X \text{ 为离散型})$$

其中  $R_X$  是  $x$  可能取值的范围,  $l = 1, 2, \dots, k$

因为样本矩  $A_l = \frac{1}{n} \sum_{i=1}^n X_i^l$  依概率收敛于相应的

总体矩  $\mu_l$  ( $l = 1, 2, \dots, k$ ),

样本矩的连续函数依概率收敛于相应的总体矩的连续函数.





## 矩估计法的定义

用样本矩来估计总体矩,用样本矩的连续函数来估计总体矩的连续函数,这种估计法称为**矩估计法**.

矩估计法的具体做法: 令  $\mu_l = A_l$ ,  $l = 1, 2, \dots, k$ .

这是一个包含  $k$  个未知参数  $\theta_1, \theta_2, \dots, \theta_k$  的方程组, 解出其中  $\theta_1, \theta_2, \dots, \theta_k$ .

用方程组的解  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$  分别作为  $\theta_1, \theta_2, \dots, \theta_k$  的估计量, 这个估计量称为矩估计量.

矩估计量的观察值称为矩估计值.



**例3** 设总体  $X$  在  $[0, \theta]$  上服从均匀分布, 其中  $\theta$  ( $\theta > 0$ ) 未知,  $(X_1, X_2, \dots, X_n)$  是来自总体  $X$  的样本, 求  $\theta$  的估计量.

**解** 因为  $\mu_1 = E(X) = \frac{\theta}{2}$ ,

根据矩估计法, 令  $\frac{\hat{\theta}}{2} = A_1 = \bar{X}$ ,

所以  $\hat{\theta} = 2\bar{X}$  为所求  $\theta$  的估计量.



**例4** 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  $X$  的样本, 求  $a$ ,  $b$  的估计量.

解 
$$\mu_1 = E(X) = \frac{a+b}{2},$$

$$\mu_2 = E(X^2) = D(X) + [E(X)]^2 = \frac{(a-b)^2}{12} + \frac{(a+b)^2}{4},$$

$$\text{令 } \frac{a+b}{2} = A_1 = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$\frac{(a-b)^2}{12} + \frac{(a+b)^2}{4} = A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2,$$





$$\text{即} \begin{cases} a + b = 2A_1, \\ b - a = \sqrt{12(A_2 - A_1^2)}. \end{cases}$$

解方程组得到 $a, b$ 的矩估计量分别为

$$\hat{a} = A_1 - \sqrt{3(A_2 - A_1^2)} = \bar{X} - \sqrt{\frac{3}{n} \sum_{i=1}^n (X_i - \bar{X})^2},$$

$$\hat{b} = A_1 + \sqrt{3(A_2 - A_1^2)} = \bar{X} + \sqrt{\frac{3}{n} \sum_{i=1}^n (X_i - \bar{X})^2}.$$



例5 设总体  $X$  服从几何分布, 即有分布律

$$P\{X = k\} = p(1-p)^{k-1} \quad (k = 1, 2, \cdots),$$

其中  $p$  ( $0 < p < 1$ ) 未知,  $(X_1, X_2, \cdots, X_n)$  是来自总体  $X$  的样本, 求  $p$  的估计量.

解 
$$\mu_1 = E(X) = \sum_{k=1}^{\infty} k p (1-p)^{k-1} = \frac{1}{p},$$

$$\text{令 } \frac{1}{\hat{p}} = A_1 = \bar{X},$$

所以  $\hat{p} = \frac{1}{\bar{X}}$  为所求  $p$  的估计量.



例6 设总体  $X$  的均值  $\mu$  和方差  $\sigma^2$  都存在, 且有  $\sigma^2 > 0$ , 但  $\mu$  和  $\sigma^2$  均为未知, 又设  $X_1, X_2, \dots, X_n$  是一个样本, 求  $\mu$  和  $\sigma^2$  的矩估计量.

解  $\mu_1 = E(X) = \mu,$

$$\mu_2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2,$$

$$\text{令} \begin{cases} \mu = A_1, \\ \sigma^2 + \mu^2 = A_2. \end{cases}$$

解方程组得到矩估计量分别为  $\hat{\mu} = A_1 = \bar{X},$

$$\hat{\sigma}^2 = A_2 - A_1^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$





上例表明:

**总体均值与方差的矩估计量的表达式不因不同的总体分布而异.**

例  $X \sim N(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  未知, 即得  $\mu, \sigma^2$  的矩估计量

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

一般地,

用样本均值  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  作为总体  $X$  的均值的矩估计,

用样本二阶中心矩  $B_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  作为总体

$X$  的方差的矩估计.



## 2. 最大似然估计法

(1) 设总体  $X$  属离散型

### 似然函数的定义

设分布律  $P\{X = k\} = p(x; \theta)$ ,  $\theta$  为待估参数,  $\theta \in \Theta$ ,

(其中  $\Theta$  是  $\theta$  可能的取值范围)

$X_1, X_2, \dots, X_n$  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, 即事件  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$  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  $\theta$  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  $\Theta$  是  $\theta$  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  $\theta$  的最大似然估计值,

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



## (2) 设总体 $X$ 属连续型

### 似然函数的定义

设概率密度为  $f(x; \theta)$ ,  $\theta$  为待估参数,  $\theta \in \Theta$ ,

(其中  $\Theta$  是  $\theta$  可能的取值范围)

$X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本,

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

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  $\theta$  的最大似然估计值,

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





最大似然估计法是由费舍尔引进的.

费舍尔

求最大似然估计量的步骤:

(一) 写出似然函数

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

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

(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  $\theta$  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  $\theta$  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  $\theta_i$  ( $i = 1, 2, \dots, k$ ) 的最大似然估计值  $\hat{\theta}_i$ .



例7 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  $P\{X = x\} = p^x (1-p)^{1-x}$ ,  $x = 0, 1$ ,

$$\begin{aligned} \text{似然函数 } L(p) &= \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} \\ &= p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i}, \end{aligned}$$





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

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

解得  $p$  的最大似然估计值  $p = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$

$p$  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.



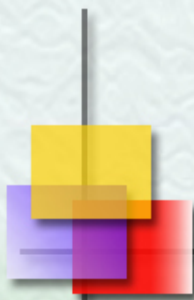
**例8** 设  $X$  服从参数为  $\lambda$  ( $\lambda > 0$ ) 的泊松分布,  $X_1, X_2, \dots, X_n$  是来自  $X$  的一个样本, 求  $\lambda$  的最大似然估计量.

**解** 因为  $X$  的分布律为

$$P\{X = x\} = \frac{\lambda^x}{x!} e^{-\lambda}, \quad (x = 0, 1, 2, \dots, n)$$

所以  $\lambda$  的似然函数为

$$L(\lambda) = \prod_{i=1}^n \left( \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \right) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n (x_i!)},$$



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

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

$$\text{解得 } \lambda \text{ 的最大似然估计值 } \lambda = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x},$$

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

这一估计量与矩估计量是相同的。





例9 设总体  $X \sim N(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  为未知参数,  $x_1, x_2, \dots, x_n$  是来自  $X$  的一个样本值, 求  $\mu$  和  $\sigma^2$  的最大似然估计量.

解  $X$  的概率密度为

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

$X$  的似然函数为

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



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

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

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



$$\text{由 } \frac{1}{\sigma^2} \left[ \sum_{i=1}^n x_i - n\mu \right] = 0 \text{ 解得 } \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x},$$

$$\text{由 } -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \mu)^2 = 0 \text{ 解得}$$

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

故  $\mu$  和  $\sigma^2$  的最大似然估计量分别为

$$\hat{\mu} = \bar{X}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2. \quad \text{它们与相应的矩估计量相同.}$$





**例10** 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.

**解** 记  $x_{(l)} = \min(x_1, x_2, \dots, x_n)$ ,  
 $x_{(h)} = \max(x_1, x_2, \dots, x_n)$ ,

$X$  的概率密度为

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



因为  $a \leq x_1, x_2, \dots, x_n \leq b$  等价于  $a \leq x_{(l)}, x_{(h)} \leq b$ ,  
 作为  $a, b$  的函数的似然函数为

$$L(a, b) = \begin{cases} \frac{1}{(b-a)^n}, & a \leq x_{(l)}, b \geq x_{(h)}, \\ 0, & \text{其他} \end{cases}$$

于是对于满足条件  $a \leq x_{(l)}, b \geq x_{(h)}$  的任意  $a, b$  有

$$L(a, b) = \frac{1}{(b-a)^n} \leq \frac{1}{(x_{(h)} - x_{(l)})^n},$$



即似然函数  $L(a, b)$  在  $a = x_{(l)}$ ,  $b = x_{(h)}$  时  
取到最大值  $(x_{(h)} - x_{(l)})^{-n}$ ,

$a, b$  的最大似然估计值

$$\hat{a} = x_{(l)} = \min_{1 \leq i \leq n} x_i, \quad \hat{b} = x_{(h)} = \max_{1 \leq i \leq n} x_i,$$

$a, b$  的最大似然估计量

$$\hat{a} = \min_{1 \leq i \leq n} X_i, \quad \hat{b} = \max_{1 \leq i \leq n} X_i.$$





## 最大似然估计的性质

设  $\theta$  的函数  $u = u(\theta)$ ,  $\theta \in \Theta$  具有单值反函数  $\theta = \theta(u)$ ,  $u \in \quad$ . 又设  $\hat{\theta}$  是  $X$  的概率密度函数  $f(x; \theta)$  ( $f$  形式已知) 中的参数  $\theta$  的最大似然估计, 则  $\hat{u} = u(\hat{\theta})$  是  $u(\theta)$  的最大似然估计.

**证明** 因为  $\hat{\theta}$  是  $\theta$  的最大似然估计值,  
所以 
$$L(x_1, x_2, \dots, x_n; \hat{\theta}) = \max_{\theta \in \Theta} L(x_1, x_2, \dots, x_n; \theta),$$

其中  $x_1, x_2, \dots, x_n$  是来自总体  $X$  的一个样本值,

由于  $\hat{u} = u(\hat{\theta}), \quad \hat{\theta} = \theta(\hat{u}),$



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

于是  $\hat{u} = u(\hat{\theta})$  是  $u(\theta)$  的最大似然估计.

此性质可以推广到总体分布中含有多个未知参数的情况.

如例9中  $\sigma^2$  的最大似然估计值为  $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ ,

函数  $u = u(\sigma^2) = \sqrt{\sigma^2}$  有单值反函数  $\sigma^2 = u^2 (u \geq 0)$ ,  
故标准差  $\sigma$  的最大似然估计值为

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



### 三、小结

两种求点估计的方法:  $\left\{ \begin{array}{l} \text{矩估计法} \\ \text{最大似然估计法} \end{array} \right.$

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.

$$\text{似然函数 } L(\theta) = L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n p(x_i; \theta)$$

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

