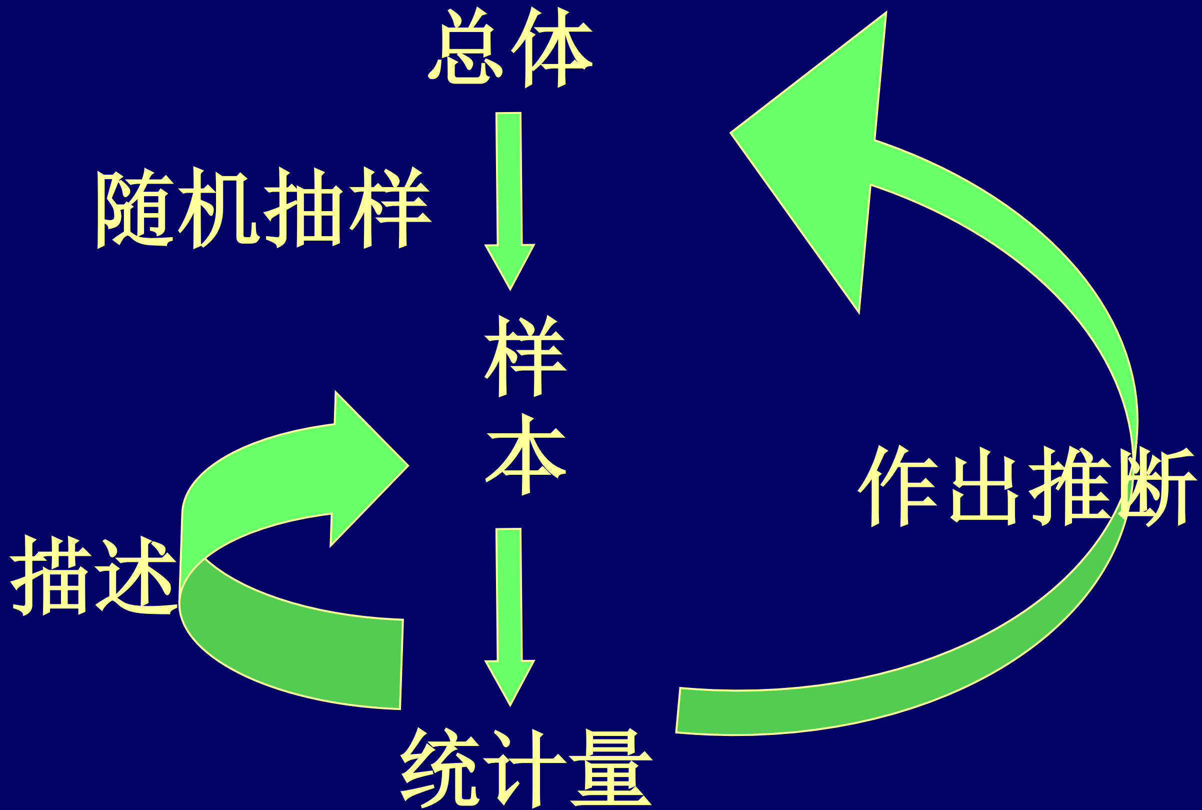
第一节 参数的点估计

- 点估计概念
- 求估计量的方法
- 估计量的评选标准

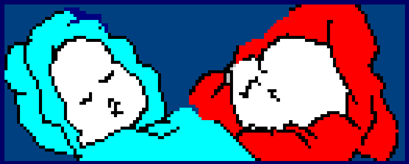


研究统计量的性质和评价一个统计推断的优良性，完全取决于其抽样分布的性质.

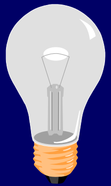
参数估计

参数估计问题是利用从总体抽样得到的信息来估计总体的某些参数或者参数的某些函数.

估计新生儿的体重

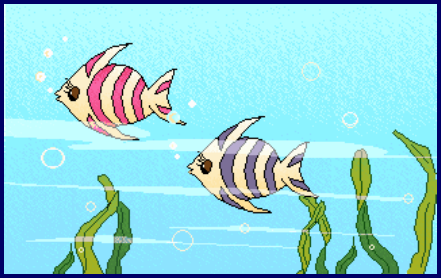


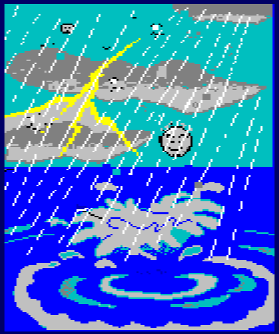
估计废品率



估计湖中鱼数

估计降雨量





…

…

在参数估计问题中，假定总体分布形式已知，未知的仅仅是一个或几个参数.

参数估计问题的一般提法

设有一个统计总体，总体的分布函数为 $F(x,\theta)$，其中θ为未知参数 (θ 可以是向量) .

现从该总体抽样，得样本

$$X_1,X_2,\ldots,X_n$$

要依据该样本对参数 θ 作出估计, 或估计 θ 的某个已知函数 $g(\theta)$.

这类问题称为参数估计.

参数估计 $\begin{cases} \text{点估计} \\ \text{区间估计} \end{cases}$

例如估计某队男生的平均身高. 身高X~ $N(\mu, 0.1^2)$

从该总体选取容量为5的样本，样本观测值为

1.65 1.67 1.68 1.78 1.69

我们的任务是根据这5个数，估计总体均值 μ 的近似值.

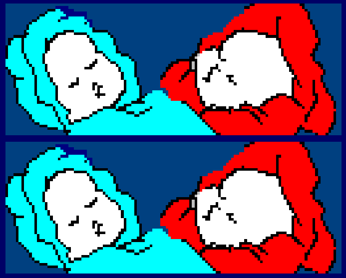
估计 μ 为1.68， 这是点估计.

估计 μ 在区间 [1.57, 1.84] 内，这是区间估计.

一、点估计概念

例1 已知某地区新生婴儿的体重 $X \sim N(\mu, \sigma^2)$，

$(\mu, \sigma$ 未知 $)$

 … 

随机抽查100个婴儿，得100个体重数据

10,7,6,6.5,5,5.2, …

而全部信息就由这100个数组成.

据此, 我们应如何估计 μ 和 σ 呢 ?

为估计μ:

我们需要构造出适当的样本的函数 $T(X_1,X_2,\ldots X_n)$，每当有了样本，就代入该函数中算出一个值，用来作为μ的估计值.

$T(X_1,X_2,\ldots X_n)$ 称为参数μ的点估计量,

把样本值代入$T(X_1,X_2,\ldots X_n)$ 中，得到μ的一个点估计值.

我们知道, 若 $X \sim N\left(\mu,\sigma^2\right)$, 则 $E(X)=\mu$.

由辛钦大数定律,

样本体重的平均值

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < \varepsilon\} = 1$$

用样本体重的均值 $\overline{X}$ 估计 μ.

类似地，用样本体重的方差 S^2 估计 σ^2.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

二、寻求估计量的方法

1. 矩估计法

2. 极大似然法

3. 最小二乘法

4. 贝叶斯方法　　　……

1. 矩估计法

K.皮尔逊

由辛钦大数定律，

若总体 X 的数学期望 $E(X)=\mu$ 有限，则有

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

$$\Downarrow$$

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k \xrightarrow{P} E(X^k)=\mu_k \quad (k=1,2,\cdots)$$

$$g(A_1,A_2,\cdots,A_k) \xrightarrow{P} g(\mu_1,\mu_2,\cdots,\mu_k)$$

其中 g 为连续函数.

例1　设总体 X 在 $[a, b]$ 上服从均匀分布，a, b 未知．$X_1, X_2, \dots, X_n$ 是来自 X 的样本，试求 a, b 的估计量．

解：$\mu_1 = E(X) = \dfrac{a+b}{2}$

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

即 $\begin{cases} a+b = 2\mu_1 \\ b-a = \sqrt{12(\mu_2 - \mu_1^2)} \end{cases}$　解得　$\begin{aligned} a &= \mu_1 - \sqrt{3(\mu_2 - \mu_1^2)} \\ b &= \mu_1 + \sqrt{3(\mu_2 - \mu_1^2)} \end{aligned}$

于是 a, b 的估计量为

$$a = \bar{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}, \quad b = \bar{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

定义 用样本原点矩估计相应的总体原点矩，又用样本原点矩的连续函数估计相应的总体原点矩的连续函数，这种参数点估计法称为矩估计法.

理论依据： 辛钦大数定律

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

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k \xrightarrow{P} E(X^k)=\mu_k \quad (k=1,2,\cdots)$$

$$g(A_1,A_2,\cdots,A_k) \xrightarrow{P} g(\mu_1,\mu_2,\cdots,\mu_k)$$

其中 g 为连续函数.

设总体的分布函数中含有k个未知参数$\theta_1,\theta_2,\cdots,\theta_k$,

Step 1 计算总体的矩

$$EX^i = g_i(\theta_1,\theta_2,\cdots,\theta_k) = \mu_i \qquad i=1,2,\ldots,k$$

Step 2 解方程组

$$\theta_j = h_j(\mu_1,\mu_2,\cdots,\mu_k) \qquad j=1,2,\ldots,k$$

Step 3 用样本矩A_i估计总体矩μ_i

即可得诸 θ_j 的矩估计量：

$$\hat{\theta}_j = h_j(A_1,A_2,\cdots,A_k) \qquad j=1,2,\ldots,k$$

矩估计量的观察值称为矩估计值 .

例2　设总体 X 的均值 μ 和方差 $\sigma^2(>0)$都存在，μ,σ^2 未知． $X_1,\ldots,X_n$ 是来自 X 的样本，试求 μ,σ^2 的矩估计量．

解: $\mu_1 = E\left(X\right) = \mu$

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

解得　$\mu = \mu_1$

$$\sigma^2 = \mu_2 - \mu_1^2$$

于是 μ,σ^2 的矩估计量为 $\mu = A_1 = \bar{X}$

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

例3 设总体X的概率密度为

$$f(x)=\begin{cases}(\alpha+1)x^{\alpha}, & 0<x<1\\ 0, & \text{其它}\end{cases}$$

其中 $\alpha>1$ 是未知参数 $,X_1,X_2,\dots,X_n$ 是取自 X 的样本,

求参数 α 的矩估计.

解: $$\mu_1=E(X)=\int_0^1 x(\alpha+1)x^{\alpha}dx$$

$$=(\alpha+1)\int_0^1 x^{\alpha+1}dx=\frac{\alpha+1}{\alpha+2}$$

解得 $$\alpha=\frac{2\mu_1-1}{1-\mu_1}$$

故 α 的矩估计量为 $$\hat{\alpha}=\frac{2\overline{X}-1}{1-\overline{X}}$$

例 4 设$X_1,X_2,\ldots X_n$是取自总体 X 的一个样本

$$X \sim f(x)=\begin{cases}\dfrac{1}{\theta}e^{-(x-\mu)/\theta}, & x\ge\mu\\ 0, & \text{其它}\end{cases}$$ θ,μ为未知参数

其中$\theta>0$，求 θ,μ 的矩估计.

解 由密度函数知 $X-\mu$ 具有均值为 θ 的指数分布

故 $\begin{cases}E(X-\mu)=\theta\\ D(X-\mu)=\theta^2\end{cases}$ 即 $\begin{cases}E(X)=\mu+\theta\\ D(X)=\theta^2\end{cases}$

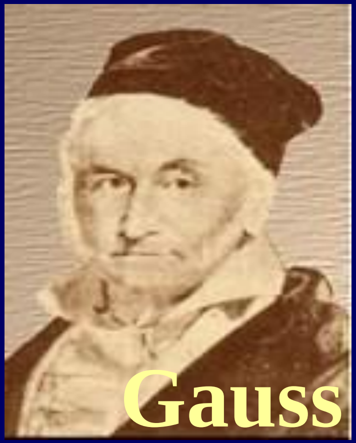
解得 $\theta=\sqrt{D(X)}$ $\mu=E(X)-\sqrt{D(X)}$

故 $\hat{\theta}=\sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$ $\hat{\mu}=\overline{X}-\sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$

矩法的优点是简单易行,并不需要事先知道总体是什么分布 .

缺点是，当总体类型已知时，没有充分利用分布提供的信息 . 一般场合下,矩估计量不具有唯一性 .

2. 最大似然法



1821

1922

最大似然法的基本思想

极大似然原理的直观想法是：一个随机试验如有若干个可能的结果 A、B、C，….。若在一次试验中，结果 A 出现，则一般认为试验条件对 A 出现有利，也即 A 出现的概率最大。

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？ 你会如何想呢?

例1、设有外形完全相同的两个箱子，甲箱有 99 个白球，1个黑球；乙箱有 1 个白球 99 个黑球。今随机地抽取一箱，再从取出的一箱中抽取一球，结果为白球。问这球从哪一个箱子中取出的？

解：甲箱中抽得白球的概率 $P(\text{白} \mid \text{甲}) = \dfrac{99}{100}$

乙箱中抽得白球的概率 $P(\text{白} \mid \text{乙}) = \dfrac{1}{100}$

由此看到，这一白球从甲箱中抽出的概率比从乙箱中抽出的概率大的多。

例2、要估计某车间生产的一批产品的不合格率 p.

用随机变量 X 来描述一件产品是合格品或不合格品。

$$X=\begin{cases}1 & \text{这件产品是不合格品}\\ 0 & \text{这件产品是合格品}\end{cases}$$

X 的分布列为 $P(X=x)=p^{x}(1-p)^{1-x}, x=0,1$

随机地从中抽取一个容量为 n 子样 $X_1,X_2,\cdots,X_n$，子样取到观测值 $x_1,x_2,\cdots,x_n$ 的概率为

$$P(X_1=x_1,X_2=x_2,\cdots,X_n=x_n)=\prod_{i=1}^{n}P(X_i=x_i)$$

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

其中 $x_i=0$ 或 1, $i=1,2,\cdots,n.$

这个概率是未知参数 p 的函数, 用 $L(p)$ 表示, 称作似然函数, 即

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

----似然函数 likelyhood

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

---对数似然函数

$$L(\hat{p}_L) = \max_{p\in(0,1)} L(p) = \max_{p\in(0,1)} p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

$$\ln L(\hat{p}_L) = \max_{p\in(0,1)} \ln(L(p)) = \max_{p\in(0,1)}\left(\sum_{i=1}^{n} x_i \ln p + (n-\sum_{i=1}^{n} x_i)\ln(1-p)\right)$$

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

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

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

解方程得 $$\hat{p}_L = \hat{p}(x_1, \cdots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

它使 **L(p)** 达到极大, 称为参数 p 的**最大似然估计值**

其相应的统计量 $$\hat{p}_L = \frac{1}{n}\sum_{i=1}^{n} X_i$$

称作参数 p 的**最大似然估计量**.

若 X 是 离散型总体, 分布律为 $P(X=x)=f(x;\theta),\ \theta\in\Theta$

求极大似然估计的一般步骤为

Step1. 计算似然函数

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

Step2. 求似然函数或对数似然函数的极大值点 $\hat{\theta}_L$

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

$$\ln L(\hat{\theta}_L;x_1,\cdots,x_n)=\max_{\theta\in\Theta}\ln L(\theta;x_1,\cdots,x_n)$$

这样得到的 $\hat{\theta}_{\mathrm{L}}$ 与样本值 $x_1, x_2, \cdots, x_n$ 有关,

常记为 $\hat{\theta}_L(x_1, \cdots, x_n)$,称 $\hat{\theta}_L(x_1, \cdots, x_n)$ 为参数 θ

的最大似然估计值, 而相应的统计量 $\hat{\theta}_L(X_1, \cdots, X_n)$

称为参数 θ 的最大似然估计量.

求函数极大值的常用方法

1. 驻点法：求似然函数或对数似然函数的驻点

一阶导数为**0**，二阶导数小于**0**的点为函数的极大值点

$$\frac{d\mathrm{L}(\theta)}{\mathrm{d}\theta}\Big|_{\hat{\theta}}=0,\ \frac{d^2\mathrm{L}(\theta)}{\mathrm{d}\theta^2}\Big|_{\hat{\theta}}<0 \qquad \frac{d\ln\mathrm{L}(\theta)}{\mathrm{d}\theta}\Big|_{\hat{\theta}}=0,\ \frac{d^2\ln\mathrm{L}(\theta)}{\mathrm{d}\theta^2}\Big|_{\hat{\theta}}<0$$

若 θ是向量，上述方程必须用方程组代替 .

2. 分析似然函数的单独性

如果似然函数是单调函数，则在区间端点处取到极值

例3、设随机变量 X服从泊松分布，$p(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$，

$k=0,1,2,\cdots$，其中 $\lambda>0$ 是未知参数。

求 λ 的最大似然估计。

解：设 $x_1,x_2,\cdots,x_n$ 为样本 $X_1,X_2,\cdots,X_n$ 的一组观测值，

似然函数为 $L(\lambda)=L(\lambda;x_1,\cdots,x_n)=\frac{\lambda^{x_1}}{x_1!}e^{-\lambda}\cdots\frac{\lambda^{x_n}}{x_n!}e^{-\lambda}=\frac{\lambda^{\sum_{i=1}^n x_i}}{x_1!x_2!\cdots x_n!}e^{-n\lambda}$

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

$$\frac{\partial\ln L(\lambda)}{\partial\lambda}=-n+\frac{1}{\lambda}\sum_{i=1}^n x_i=0\Rightarrow\hat{\lambda}_L=\bar{x}$$

故 λ 的最大似然估计量为 $\hat{\lambda}_L(X_1,\cdots X_n)=\bar{X}$

最大似然估计法也适用于分布中含多个未知参数

$\theta_1,\theta_2,\cdots,\theta_k$ 的情况。这时似然函数 $L=L(\theta_1,\theta_2,\cdots,\theta_k)$

分别令 $\dfrac{\partial L}{\partial \theta_i}=0,\ i=1,2,\cdots,k.$

或令 $\dfrac{\partial \ln L}{\partial \theta_i}=0,\ i=1,2,\cdots,k.$

解上述由 ***k*** 个方程组成的方程组，即可得到未知参数

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

若 总体X 是连续型R.V, 其密度函数为 $f(x;\theta),\theta\in\Theta$

子样 $X_1,X_2,\cdots,X_n$ ，其观测值 $x_1,x_2,\cdots,x_n$

样本 $(X_1,X_2,\cdots,X_n)$ 落入点 $(x_1,x_2,\cdots,x_n)$ 邻域内的概率为

$$\boldsymbol{P}(\mathrm{x}_1<\mathrm{X}_1\le \mathrm{x}_1+\Delta \mathrm{x}_1,\cdots,\mathrm{x}_n<\mathrm{X}_n\le \mathrm{x}_n+\Delta \mathrm{x}_n)=\prod_{i=1}^{n}\mathrm{f}(\mathrm{x}_i;\theta)\Delta \mathrm{x}_i$$

同离散型一样,取 θ 的估计值 $\hat{\theta}_L$,使概率 $\prod_{i=1}^{n} f(x_i;\theta)\Delta x_i$

取到最大值。

由于 Δx_i 是不依赖于 θ 的增量，故只需考虑函数

$$L(\theta)=L(\theta;x_1,\cdots,x_n)=\prod_{i=1}^{n} f(x_i;\theta)$$ 的最大值。

所以连续总体样本的似然函数为

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

为样本 $(X_1,X_2,\cdots,X_n)$ 的联合密度函数。

若 $L(\hat{\theta}_L;x_1,\cdots,x_n)=\max\limits_{\theta\in\Theta}L(\theta;x_1,\cdots,x_n)$

则称 $\hat{\theta}_L(x_1,\cdots,x_n)$ 为 θ 的最大似然估计值。

求参数的最大似然估计的一般步骤

Step1. 计算样本的似然函数

1）若总体是离散型R.V，似然函数为样本的联合概率

$$\mathrm{L}(\theta) = \boldsymbol{P}(\mathrm{X}_1 = \mathrm{x}_1, \cdots, \mathrm{X}_\mathrm{n} = \mathrm{x}_\mathrm{n})$$

2）若总体是连续型R.V，似然函数为样本的联合密度

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

Step2. 求似然函数或对数似然函数的极大值点 $\hat{\theta}_L$

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

$$\ln \mathrm{L}(\hat{\theta}_\mathrm{L}; \mathrm{x}_1, \cdots, \mathrm{x}_\mathrm{n}) = \max_{\theta\in\Theta} \ln \mathrm{L}(\theta; \mathrm{x}_1, \cdots, \mathrm{x}_\mathrm{n})$$

例5 设总体 $X \sim N(\mu,\sigma^2)$，μ,σ^2未知．$x_1,\ldots,x_n$ 是来自 X 的样本值，试求 μ,σ^2的最大似然估计量．

解： X 的概率密度为 $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$， $-\infty<x<\infty$

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

$$=(2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2]$$

于是 $LnL=-\dfrac{n}{2}\ln(2\pi)-\dfrac{n}{2}\ln\sigma^2-\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

令 $t=\sigma^2$

$t=\sigma^2$

$$\frac{\partial}{\partial\sigma^2}LnL=\frac{\partial}{\partial t}LnL=-\frac{n}{2t}+\frac{1}{2t^2}\sum_{i=1}^{n}(x_i-\mu)^2=0$$

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

令 $$\frac{\partial}{\partial\mu}LnL = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}x_i - n\mu\right) = 0$$

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

解得 $$\mu = \frac{1}{n}\sum_{i=1}^{n}x_i = \overline{x} \qquad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$$

μ,σ^2 的最大似然估计量为

$$\mu = \overline{X}, \qquad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

例6 设$X_1,X_2,\ldots X_n$是取自总体X的一个样本

$$X \sim f(x)=\begin{cases}\theta x^{\theta-1}, & 0<x<1\\ 0, & \text{其它}\end{cases}$$ 其中$\theta>0$,

求 θ 的极大似然估计.

解：似然函数为

$$L(\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta-1}\qquad (0<x_i<1)\quad 1\le i\le n$$

对数似然函数为 $$\ln L(\theta)=n\ln\theta+(\theta-1)\sum_{i=1}^{n}\ln x_i$$

求导并令其为0 $$\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\ln x_i=0$$

得 $\theta^*=-n\Big/\sum_{i=1}^{n}\ln x_i$ 即为θ的MLE .

例7 设$X_1,X_2,\dots X_n$是取自总体X的一个样本

$$X \sim f(x)=\begin{cases}\dfrac{1}{\theta}e^{-(x-\mu)/\theta}, & x\ge\mu\\ 0, & \text{其它}\end{cases}\quad \theta,\mu\text{为未知参数}$$

其中 θ >0,求θ,μ的最大似然估计.

解：似然函数为

$$L(\theta,\mu)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-(x_i-\mu)/\theta},\quad x_i\ge\mu\quad i=1,2,\dots,n$$

$$=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-\mu)},\quad \min x_i\ge\mu$$

对数似然函数为 $\ln L(\theta,\mu)=-n\ln\theta-\dfrac{1}{\theta}\sum_{i=1}^{n}(x_i-\mu)$

对数似

ln

用求导方法无法最终确定θ、μ,
利用函数的单调性来求.

对 θ, μ 分别求偏导并令其为0,

$$\frac{\partial \ln L(\theta,\mu)}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}(x_i - \mu)=0 \qquad (1)$$

$$\frac{\partial \ln L(\theta,\mu)}{\partial \mu} = \frac{n}{\theta} > 0 \qquad (2)$$

由(1)得 $\theta = \frac{1}{n}\sum_{i=1}^{n} x_i - \mu$

$$L(\theta,\mu)=\begin{cases}\dfrac{1}{\theta^n}e^{-\frac{1}{\theta}\sum\limits_{i=1}^{n}(x_i-\mu)}, & \min x_i \geq \mu \\ 0, & \text{其它}\end{cases}$$

对 $\mu \leq \min x_i, L(\theta,\mu) > 0$, 且是$\mu$的增函数

μ 取其它值时，$L(\theta,\mu)=0.$

故使$L(\theta,\mu)$达到最大的μ, 即 μ的MLE是

$$\mu^* = \min_{1\leq i\leq n} x_i$$

于是

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

即 θ^*, μ^*为 θ, μ的MLE .

例8 设总体X在[a，b]上服从均匀分布，a，b是未知参数，$x_1, x_2, \cdots x_n$为样本值，求a，b的极大似然估计量。

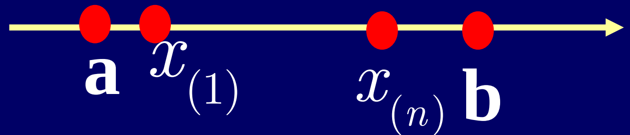
解：令$x_{(1)} = \min\{x_1, x_2, \cdots x_n\}$

$$x_{(n)} = \max\{x_1, x_2, \cdots x_n\}$$

则似然函数为 $L(a,b) = \dfrac{1}{(b-a)^n} \quad a \le x_{(1)}, b \ge x_{(n)}$

对于满足条件$a \le x_{(1)}, x_{(n)} \le b$ 的任意a，b有

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



$$b - a \ge x_{(n)} - x_{(1)} > 0$$

即 $L(a,b)$ 在 $a=x_{(1)}, b=x_{(n)}$ 时达到最大值，

故a，b的极大似然估计值为

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

a, b的极大似然估计量为

$$\hat{a}=\mathrm{X}_{(1)}=\min_{1\le i\le n} X_i,\quad \hat{b}=\mathrm{X}_{(n)}=\max_{1\le i\le n} X_i$$

四、小结

这一讲，我们介绍了参数点估计, 给出了寻求估计量最常用的矩法和极大似然法 .

参数点估计是用一个确定的值去估计未知的参数 . 看来似乎精确 ， 实际上把握不大 .