# 第七章 参数估计

- § 7.1 点估计
- § 7.2 基于截尾样本的最大似然估计
- § 7.3 估计量的评选标准
- § 7.4 区间估计
- § 7.5 正态总体均值和方差的区间估计
- § 7.6 (0－1)分布参数的区间估计
- § 7.7 单侧置信区间

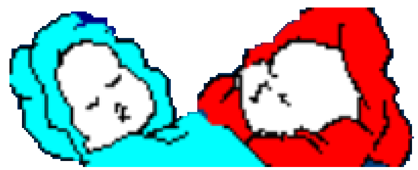
# 第七章 参数估计

- § 7.1 点估计
- § 7.2 基于截尾样本的最大似然估计
- § 7.3 估计量的评选标准
- § 7.4 区间估计
- § 7.5 正态总体均值和方差的区间估计
- § 7.6 (0－1)分布参数的区间估计
- § 7.7 单侧置信区间

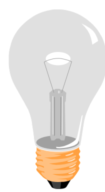
# §7.1 点估计

从本章开始讨论统计推断的两类基本问题：参数估计和假设检验问题，本章讨论总体参数的**点估计和区间估计**。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.

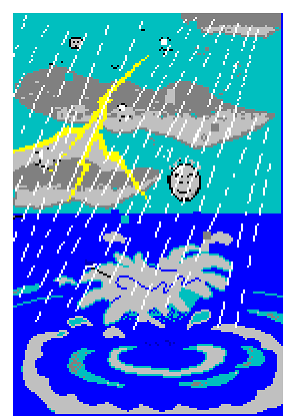
估计新生儿的平均体重



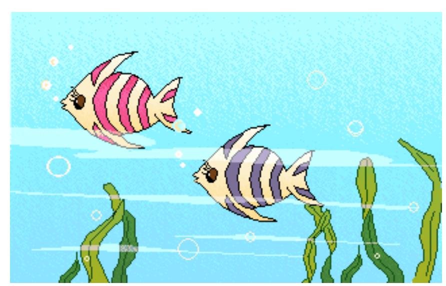
估计废品率



估计平均降雨量



估计湖中鱼数



…

…

# §7.1 点估计

- 问题：在总体形式已知时，对参数进行估计，有3个问题：

(1) 估计值?(点估计)

(2) 统计量选择的科学性（评选标准）

(3) 估计值的可信度? （区间估计）

- 点估计:

设总体X的分布函数的形式为已知，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参数的点估计问题

- 如：已知X～$N(\mu,\sigma^2)$，但参数$\mu,\sigma^2$为未知，需要估计

# § 7.1 点估计

- 点估计问题用数学模型来描述：
  - 设总体X的分布函数$F(x;\theta)$的形式为已知，$\theta$是待估参数，$X_1,X_2,\ldots,X_n$是X的一个样本，$x_1, x_2,\ldots, x_n$是相应的一个样本值
  - 点估计问题就是要**基于样本构造一个适当的统计量** $\hat{\theta}(X_1, X_2,\ldots,X_n)$，并用它的一个观察值 $\hat{\theta}(x_1, x_2,\ldots, x_n)$作为未知参数的近似值。
  - 称 $\hat{\theta}(X_1,X_2,\ldots,X_n)$为$\theta$的估计量，$\hat{\theta}(x_1, x_2,\ldots, x_n)$为$\theta$的估计值
- 不致混淆情况下，称估计量和估计值为估计，并都简记为 $\hat{\theta}$
  - 估计量是样本的函数，**样本值不同，$\theta$ 的估计值一般不同**，这体现了样本的个性

# §7.1 点估计

- 例1：某炸药厂，一天中发生着火现象的次数X～$\pi(\lambda)$泊松分布，参数$\lambda$未知，现有以下样本值，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 | 总计250天 |

- 解：首先找到一个对$\lambda$的估计量

已知$\lambda = E(X)$，而样本均值的数学期望$E(\overline{X}) = E(X)$，即$A_k \xrightarrow{P} \mu_k|_{k=1}$

所以可以用样本均值$\overline{X}$来估计总体均值$\lambda$

样本容量为250，均值相当于250天的着火次数相加/250

- 即$\lambda$的估计量 $\hat{\lambda} = \hat{E}(X) = \dfrac{1}{n}\sum_{k=1}^{n} X_k$，n＝250

- $\lambda$的估计值 $\hat{\lambda} = \hat{E}(X) = \dfrac{1}{250}\sum_{k=1}^{250} X_k = \dfrac{1}{250}(0\times 75 + 1\times 90 + \cdots + 6\times 1) = 1.22$
- 可见对未知参数的估计量的构造方法是值得研究的

# §7.1 点估计

- **有两种典型的点估计方法：**
  - **矩估计法和最大似然估计法**
- **(一)矩估计法**
  - **它是基于一种简单的“替换”思想建立起来的一种估计方法**
  - **是英国统计学家*K.*皮尔逊最早提出的**
  - **其基本思想是用样本矩估计总体矩**
  - **理论依据是大数定律**

# §7.1 点估计

- 可以通过构造样本矩来构造估计量，因为**总体矩总可以表示为未知参数的函数**，如果**令总体矩等于样本矩的观察值，则构建了一个等式**
- 设X为连续型随机变量，概率密度为

$$f(x;\theta_1,\theta_2,\ldots,\theta_k)$$

- 或X为离散型随机变量，其分布律为

$$P\{X=x\}=p(x;\theta_1,\theta_2,\ldots,\theta_k)$$

  - 其中$\theta_1,\theta_2,\ldots,\theta_k$为待估的$k$个未知参数，
  - $X_1,X_2,\ldots,X_n$是来自总体X的样本

# §7.1 点估计

- 假设总体X的前$k$阶总体矩存在，记为$\mu_1, \mu_2,\ldots,\mu_k$，其中对任意的$\mu_l$有

$$X为连续型：\mu_l = E(X^l) = \int_{-\infty}^{\infty} x^l f(x;\theta_1,\theta_1,\cdots,\theta_k)dx$$

$$或离散型：\mu_l = E(X^l) = \sum_{i=1}^{\infty} x_i^l p(x_i;\theta_1,\theta_1,\cdots,\theta_k)$$

- 它们都是待估参数$\theta_1, \theta_2,\ldots, \theta_k$的函数，写成方程组得

$$\begin{cases} \mu_1 = \mu_1(\theta_1,\theta_1,\cdots,\theta_k) \\ \mu_2 = \mu_2(\theta_1,\theta_1,\cdots,\theta_k) \\ \qquad\cdots \\ \mu_k = \mu_k(\theta_1,\theta_1,\cdots,\theta_k) \end{cases}$$

# § 7.1 点估计

- 一般情况下，包含$k$个未知参数的方程组可以求出$k$个未知参数，即
$$\begin{cases} \theta_1 = \theta_1(\mu_1, \mu_1, \cdots, \mu_k) \\ \theta_2 = \theta_2(\mu_1, \mu_1, \cdots, \mu_k) \\ \cdots \\ \theta_k = \theta_k(\mu_1, \mu_1, \cdots, \mu_k) \end{cases}$$
- 但总体矩的形式容易求得，即待估参数的函数，但实际值未知，而我们有辛钦定理的推论，即

若总体X的k阶矩$E(X^k)=\mu_k$存在，则由辛钦定理，当$n\to\infty$时，样本k阶矩$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P} \mu_k$，$k=1,2,\ldots$，

由依概率收敛的性质知道

$$g(A_1, A_2,\ldots, A_k) \xrightarrow{P} g(\mu_1, \mu_2,\ldots, \mu_k)$$，g为连续函数

# §7.1 点估计

- 因此我们就以样本$l$阶矩$A_l$代替上式中的总体$l$阶矩$\mu_l$，$l=1,2,\ldots k$，并以

$$\hat{\theta}_l=\theta_l(A_1, A_2,\ldots, A_k)\text{，}l=1,2,\ldots k$$

作为未知参数$\theta_l$的估计量，$l=1,2,\ldots k$，这种估计量称为矩估计量，矩估计量的观察值称为矩估计值

- **矩估计法定义**：
- 象以上过程那样，基于样本矩依概率收敛于总体矩，并进而有样本矩的连续函数依概率收敛于总体矩的连续函数的性质，而用样本矩（或样本矩的连续函数）作为相应的总体矩（或总体矩的连续函数）的估计量，这种估计方法称为矩估计法。

# § 7.1 点估计

- **矩估计法的一般方法:**
- 1° 由总体X的概率分布的形式计算总体矩的形式，即

$$\mu_l = E(X^l) = \mu_l(\theta_1, \theta_2, \ldots, \theta_k),\ l = 1,2,\ldots k,$$

它们是未知参数$\theta_1, \theta_2, \ldots, \theta_k$的函数，$k$的取值与待估的未知参数个数相同

- 2° 将它们联立得到方程组，并解出未知参数用总体矩表示的函数形式
- 3° 直接以样本矩代替总体矩，$A_l$代替上式中的总体$l$阶矩$\mu_l$，$l=1,2,\ldots k$，得到$k$个估计量 $\hat{\theta}_l = \theta_l(A_1, A_2, \ldots, A_k)$，$l=1,2,\ldots k$，即得到$k$个未知参数的矩估计量
- 4 ° 以样本值计算出样本矩的观察值并分别代入矩估计量，得到各个未知参数的估计值

# § 7.1 点估计

**例2** 设总体 $X$ 在 $[a,b]$ 上服从均匀分布其中 $a$, $b$ 未知, $(X_1, X_2, \cdots, X_n)$ 是来自总体 $X$ 的样本, 求 $a$, $b$ 的矩估计量

解 1° 有两个参数，先求前二阶矩

$$\mu_1 = E(X) = \frac{a+b}{2},$$

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

2° 求解未知参数的表达式，用总体矩表示

易知 $a+b=2\mu_1$

$$b-a=\sqrt{12(\mu_2-\mu_1^2)}$$

联立方程，解得
$$a=\mu_1-\sqrt{3(\mu_2-\mu_1^2)}$$
$$b=\mu_1+\sqrt{3(\mu_2-\mu_1^2)}$$

# § 7.1 点估计

**3° 以样本$k$阶矩$A_1, A_2$，代替总体$k$阶矩$\mu_1$，$\mu_2$，**

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

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

其中$A_1 = \bar{X},\ A_2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^{\,2}$

# §7.1 点估计

➲ 例3　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$X_1,X_2,\ldots,X_n$是来自X的样本，试求$\mu$和$\sigma^2$的矩估计量

解：
$$\begin{cases}\mu_1 = E(X) = \mu \\ \mu_2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2\end{cases}$$

解得 $\mu = \mu_1$,
$$\sigma^2 = \mu_2 - \mu_1^2$$

分别以$A_1$，$A_2$，代替$\mu_1$，$\mu_2$，得到$\mu$和$\sigma^2$的矩估计量为

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

总体均值与方差的矩估计量的表达式不因总体分布而异，其主要原因是方差与一阶矩和二阶矩满足 $D(X) = E(X^2) - [E(X)]^2$

# § 7.1 点估计

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

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

一般地:

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

用样本二阶中心矩$B_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$作为总体$X$的方差的矩估计

矩估计法的优点是简单易行

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 .

# § 7.1 点估计

**例　设总体$X$的分布密度为**

$$f(x;\theta)=\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}\quad(-\infty<x<+\infty,\theta>0)$$

$X_1,X_2,\ldots,X_n$为总体$X$的样本,求参数$\theta$的矩估计量

**解：** 由于$f(x;\theta)$只含有一个未知参数$\theta$

一般只需求出$E(X)$便能得到$\theta$的矩估计量，但是

$$E(X)=\int_{-\infty}^{+\infty}xf(x;\theta)dx=\int_{-\infty}^{+\infty}x\cdot\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}dx=0$$

即$E(X)$不含有$\theta$, 故不能由此得到$\theta$的矩估计量.

# §7.1 点估计

为此，求

$$E(X^2)=\int_{-\infty}^{+\infty}x^2 f(x;\theta)dx=\int_{-\infty}^{+\infty}x^2\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}dx$$

$$=\frac{1}{\theta}\int_0^{+\infty}x^2e^{-\frac{x}{\theta}}dx=2\theta^2$$

故令 $\frac{1}{n}\sum_{i=1}^{n}X_i^2=2\hat{\theta}^2$

于是解得$\theta$的矩估计量为 $\hat{\theta}=\sqrt{\frac{1}{2n}\sum_{i=1}^{n}X_i^2}$

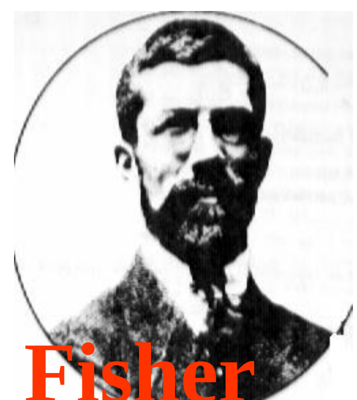
# §7.1 点估计

## ➲ (二)最大似然估计法

- **最大似然法**是在总体类型已知条件下使用的一种参数估计方法

- 它首先是由德国数学家高斯在1821年提出的，然而这个方法常归功于英国统计学家Fisher， Fisher在1922年重新发现了这一方法，并首先研究了这种方法的一些性质

Gauss



Fisher

# §7.1 点估计

**似然函数**（当X为离散型随机变量时）

- 设总体X属离散型，其分布律$P\{X=x\}=p(x;\theta)$，$\theta\in\Theta$的形式为已知，$\theta$为待估参数，$\Theta$是$\theta$可能取值的范围。
- 设$X_1,X_2,\ldots,X_n$是来自X的样本，则$X_1,X_2,\ldots,X_n$的联合分布律为$\prod_{i=1}^{n} p(x_i;\theta)$。
- 又设$x_1, x_2,\ldots, x_n$是相应于样本$X_1,X_2,\ldots,X_n$的一个样本值。易知样本$X_1,X_2,\ldots,X_n$取到观察值$x_1, x_2,\ldots, x_n$的概率，亦即事件$\{X_1=x_1,X_2=x_2,\ldots, X_n=x_n\}$发生的概率为

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

- 该概率随$\theta$取值而变化，是$\theta$的函数，**$L(\theta)$称为样本的似然函数**，其中$x_1, x_2,\ldots, x_n$是已知的样本值，它们都是常数

# §7.1 点估计

R.A.Fisher引进了最大似然估计法

主要思路：

- 现在$X_1,X_2,\ldots,X_n$取到观察值$x_1, x_2,\ldots, x_n$这一情况已经发生，说明取到这一样本值的概率$L(\theta)$比较大
- 显然如果$\theta$的取值使得出现$x_1, x_2,\ldots, x_n$的概率为小概率事件是不合理的。
- 这样我们自然认为使得样本值$x_1, x_2,\ldots, x_n$出现的概率$L(\theta)$ (它是$\theta$的函数) 取值很大的$\theta$值作为待估参数$\theta$的估计量更为合理。

# §7.1 点估计

- 最大似然估计法就是固定样本观察值$x_1, x_2,\ldots, x_n$，在$\theta$取值的可能范围$\Theta$内，挑选使似然函数$L(\theta)=L(x_1, x_2,\ldots, x_n; \theta)$达到最大值的参数值$\hat{\theta}$，作为$\theta$的估计值。即取$\hat{\theta}$使

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

- 注意"在$\theta$取值的可能范围$\Theta$内"这一条件，当极大值点不在$\Theta$范围内时，应在$\Theta$内选取似然函数最大的点作为估计值。
- 这样得到的估计值与样本值$x_1, x_2,\ldots, x_n$有关，是它们的函数，常记为$\hat{\theta}(x_1, x_2,\ldots, x_n)$,称为参数$\theta$的最大似然估计值，而相应的统计量$\hat{\theta}(X_1,X_2,\ldots,X_n)$称为参数$\theta$的最大似然估计量
- 注意：最大似然估计是先求估计值表达形式,再给出估计量

# §7.1 点估计

当X为连续型随机变量时

- 若总体X属连续型，其概率密度$f(x;\theta)$，$\theta\in\Theta$的形式为已知，$\theta$为待估参数，$\Theta$是$\theta$可能取值的范围。

- 设$X_1,X_2,\ldots,X_n$是来自X的样本，则$X_1,X_2,\ldots,X_n$的联合概率密度为$\prod_{i=1}^{n} f(x_i;\theta)$

- 又设$x_1, x_2,\ldots, x_n$是相应于样本$X_1,X_2,\ldots,X_n$的一个样本值。则随机点$(X_1,X_2,\ldots,X_n)$落在点$(x_1, x_2,\ldots, x_n)$的邻域(边长为$dx_1, dx_2,\ldots, dx_n$的n维立方体)内的概率近似为$\prod_{i=1}^{n} f(x_i;\theta)dx_i$ 该值随$\theta$的变化而变化。

- 与离散型的情况一样，取$\theta$的估计值使得以上概率取到最大值，但因子 $\prod_{i=1}^{n} dx_i$ 不随$\theta$而变，故只需考虑函数

$$L(\theta)=L(x_1, x_2,\ldots, x_n; \theta)=\prod_{i=1}^{n} f(x_i;\theta)，\theta\in\Theta,$$

的最大值，这里$L(\theta)$称为样本的似然函数。

# §7.1 点估计

- 若估计值 $\hat{\theta}$ 使得

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

则称 $\hat{\theta}(x_1, x_2, \ldots, x_n)$为参数$\theta$的最大似然估计值，

称 $\hat{\theta}(X_1, X_2, \ldots, X_n)$为参数$\theta$的最大似然估计量。

- 在以上定义下，确定最大似然估计量的问题就归结为微分学中的求最大值问题了

- 在很多情况下，$p(x;\theta)$和$f(x;\theta)$关于$\theta$可微，$\hat{\theta}$常可从方程 $\frac{d}{d\theta}L(\theta) = 0$ 中解得。

- 又$L(\theta)$与$\ln L(\theta)$在同一$\theta$处取到极值，因此$\theta$的最大似然估计值$\hat{\theta}$ 也可从方程 $\frac{d}{d\theta}\ln L(\theta) = 0$中解得。而后者往往比较方便，并称为**对数似然方程**

# §7.1 点估计

➲ **最大似然估计的一般方法：**

1° 给出总体X的分布律或分布函数$p(x;\theta)$和$f(x;\theta)$

2° 写出似然函数$L(\theta)=\prod_{i=1}^{n} p(x_i;\theta)$ 或

$L(\theta)=\prod_{i=1}^{n} f(x_i;\theta)$，或相应的对数似然函数$\ln L(\theta)$

3° 由 $\frac{d}{d\theta}L(\theta)=0$ 或 $\frac{d}{d\theta}\ln L(\theta)=0$ 结合$\theta$的取值范围，求得参数$\theta$的估计值，进而给出估计量

就似然函数而言，离散型是样本分布律函数值，连续性是样本概率密度值

# § 7.1 点估计

**例4** 设 $X \sim b(1,p)$, $X_1, X_2, \cdots, X_n$是来自$X$的一个样本,求$p$的最大似然估计量

**解** 设$x_1, x_2, \cdots, x_n$为相应于样本$X_1, X_2, \cdots, X_n$的一个样本值,

$X$的分布律为 $P\{X = x\} = p^x (1-p)^{1-x}, x = 0,1$

似然函数 $L(p) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$

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

# § 7.1 点估计

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

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

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

$p$ 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，进一步验证了极大似然估计的合理性**

# § 7.1 点估计

## ➲ 多个未知参数的极大似然估计

**如果随机变量X的分布中含有多个未知参数$\theta_1$, $\theta_2,\ldots,\theta_k$。似然函数L是这些未知参数的函数，即$L(\theta_1,\theta_2,\ldots,\theta_k)$。**

**令** $$\frac{\partial}{\partial\theta_i}L(\theta_1,\theta_2,\cdots\theta_k)=0,\quad i=1,2,\ldots,k$$

**或** $$\frac{\partial}{\partial\theta_i}\ln L(\theta_1,\theta_2,\cdots\theta_k)=0,\quad i=1,2,\ldots,k \qquad (*)$$

**解上述的$k$个方程组成的方程组，即可得到各未知参数$\theta_i(i=1,2,\ldots,k)$的最大似然估计值$\hat{\theta}_i$，上述对数形式方程组称为对数似然方程组**

# § 7.1 点估计

**例5** 设总体 $X \sim N(\mu,\sigma^2)$, $\mu,\sigma^2$为未知参数, $x_1,x_2,\cdots,x_n$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

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

# § 7.1 点估计

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

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

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

# § 7.1 点估计

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

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

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

它们与相应的矩估计量相同.

# §7.1 点估计

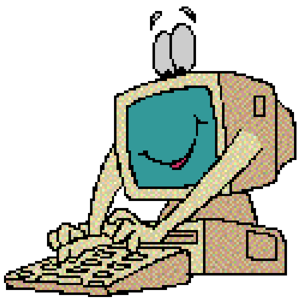
**例6** 设总体$X$在$[a,b]$上服从均匀分布,其中$a$, $b$未知, $x_1,x_2,\cdots,x_n$是来自总体$X$的一个样本值,求$a,b$的最大似然估计量.

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

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

$X$的概率密度为

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



# §7.1 点估计

因为$a \le x_1, x_2, \cdots, x_n \le b$等价于$a \le x_{(1)}, x_{(n)} \le b$,

作为$a, b$的函数的似然函数为

$$L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a \le x_{(1)},\ b \ge x_{(n)} \\ 0, & \text{其它}\end{cases}$$

无极值点

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

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

# §7.1 点估计

即似然函数 $L(a,b)$ 在 $a=x_{(1)},\ b=x_{(n)}$ 时

取到最大值 $(x_{(n)}-x_{(1)})^{-n}$,

$a,b$ 的最大似然估计值

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

$a,b$ 的最大似然估计量

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

# §7.1 点估计

- 用上述求导方法求参数的最大似然估计有时行不通，这时要用**极大似然原则**来求

- 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

- 最大似然估计函数的性质:

  - 待估参数的函数的估计量

设$\theta$的函数$u=u(\theta)$，$\theta\in\Theta$，具有单值反函数$\theta=\theta(u)$，$u\in U$，又设 $\hat{\theta}$是X的概率分布中参数$\theta$的最大似然估计，则 $\hat{u}=u(\hat{\theta})$是$u(\theta)$的最大似然估计。

# § 7.1 点估计

- 证：由于$\hat{\theta}$是$\theta$ 的最大似然估计，于是有

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

其中$x_1, x_2,\ldots, x_n$是X的一个样本值，而由题设有$\hat{u} = u(\hat{\theta})$，且有$\hat{\theta} = \theta(\hat{u})$

这样上式可写成

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

这样选取的估计量$\hat{u}$的函数$\theta(\hat{u}) = \hat{\theta}$使得$L(x_1, x_2,\ldots, x_n; \theta(\hat{u}))$最大，因此 $\hat{u} = u(\hat{\theta})$是$u(\theta)$的最大似然估计

# §7.1 点估计

- **当总体分布中含有多个未知参数时**，也具有上述性质
- 例如例5中的两个参数$\mu,\sigma^2$，得到$\hat{\sigma^2}=B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$
- 函数$u=u(\sigma^2)=\sqrt{\sigma^2}$有单值反函数$\sigma^2=u^2(u\geq 0)$，按上述性质得到标准差$\sigma$的最大似然估计
- $$\hat{\sigma}=\sqrt{\hat{\sigma^2}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$
- 很多情况下对数似然方程或方程组没有有限函数形式的解，（初等不可解）用数值计算方法求近似解。著名的有newton－raphson迭代算法，拟牛顿算法，切线法等

# 第七章 参数估计

- §7.1 点估计
- §7.2 基于截尾样本的最大似然估计
- §7.3 估计量的评选标准
- §7.4 区间估计
- §7.5 正态总体均值和方差的区间估计
- §7.6 (0－1)分布参数的区间估计
- §7.7 单侧置信区间

# §7.2基于截尾样本的最大似然估计

- 如果我们能够获得一个完全的样本，就可以用数理统计知识去分析总体的性质。

- 但有时只能获得部分样本，比如
  - 灯泡寿命的试验，可能在测试时间段内，仍然有灯泡没有点爆，这样就获得了部分样本，如何根据这种情况给出最大似然估计。
  - 可以考察一定的时间或者得到一定个数的样本点爆时间。这就是定时结尾样本和定数结尾样本
  - 本小节内容略去

# 本章作业

- $P_{173}$：1，2(1,3)，3(1,3)，4(1,3)