第七章

参数估计

统计推断的基本问题

- 参数估计问题
  - 点估计
  - 区间估计
- 假设检验问题

数理统计问题：如何选取样本来对总体的种种统计特征作出判断。

参数估计问题：知道随机变量（总体）的分布类型，但它的一个或多个参数未知，根据样本来估计总体的参数，这类问题称为参数估计(paramentric estimation)

参数估计的类型——点估计、区间估计

例如，$X \sim N(\mu, \sigma^2)$,

若μ, σ^2未知，通过构造样本的函数，给出它们的估计值或取值范围就是参数估计的内容.

点估计

区间估计

雨课堂 Rain Classroom

参数估计的类型

点估计 —— 估计未知参数的值

区间估计——

估计未知参数的取值范围,

并使此范围包含未知参数

真值的概率为给定的值.

§7.1 点估计方法

点估计的思想方法

设总体X的分布函数的形式已知，但含有一个或多个未知参数：$\theta_1, \theta_2, \ldots, \theta_k$

设 $X_1, X_2, \cdots, X_n$为总体的一个样本

构造 k 个统计量：

$$\left.\begin{array}{l}\theta_1(X_1, X_2, \cdots, X_n) \\ \theta_2(X_1, X_2, \cdots, X_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \theta_k(X_1, X_2, \cdots, X_n)\end{array}\right\} \quad \text{随机变量}$$

当测得样本值$(x_1, x_2, \cdots, x_n)$时,代入上述统计量，即可得到 k 个数:

$$\left.\begin{array}{l} \hat{\theta}_1(x_1, x_2, \cdots, x_n) \\ \hat{\theta}_2(x_1, x_2, \cdots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \hat{\theta}_k(x_1, x_2, \cdots, x_n) \end{array}\right\} \text{数 值}$$

称数 $\hat{\theta}_1, \cdots, \hat{\theta}_k$ 为未知参数 $\theta_1, \cdots, \theta_k$ 的估计值

对应统计量 为未知参数 $\theta_1, \cdots, \theta_k$ 的估计量

问题
- 如何构造统计量？
- 如何评价估计量的好坏？

三种常用的点估计方法

数字特征法：以样本的数字特征作为相应总体数字特征的估计量。

矩法估计

最大似然估计法

数字特征法

数字特征法：以样本的数字特征作为相应总体数字特征的估计量。

以样本均值 $\bar{X}$ 作为总体均值 μ 的点估计量，即

$$\hat{\mu}=\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \qquad \text{点估计值} \qquad \hat{\mu}=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

以样本方差 S^2 作为总体方差 σ^2 的点估计量，即

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

点估计值 $$\widehat{\sigma^2}=S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

例1 一批钢件的**20**个样品的屈服点（**t/cm²**）为

4.98	**5.11**	**5.20**	**5.20**	**5.11**	**5.00**	**5.35**
5.61	**4.88**	**5.27**	**5.38**	**5.48**	**5.27**	**5.23**
4.96	**5.15**	**4.77**	**5.35**	**5.38**	**5.54**	

试估计该批钢件的平均屈服点及其方差。

解　由数字特征法，得屈服点及方差的估计值为

$$\hat{\mu}=\bar{x}=\frac{1}{20}\sum_{i=1}^{20}x_i=5.21$$

$$\widehat{\sigma^2}=S^2=\frac{1}{20-1}\sum_{i=1}^{20}(x_i-5.21)^2=0.049$$

k 阶矩的概念

定义 设 X 为随机变量，若 $E(|X|^k)$ 存在，则称 $E(X^k)$ 为 X 的 k 阶原点矩，记作 $\mu_k = E(X^k)$

样本的 k 阶原点矩，记作 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$

结论： $A_k \xrightarrow{P} \mu_k \quad (n \to \infty), k = 1, 2, \cdots$

原因： 辛钦大数定律

作用： 矩估计法的理论依据

参数的矩法估计

矩法估计：用样本的矩作为总体矩的估计量，具体做法如下

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

这里总体**X**的分布函数中含有**k**个参数。从上式中解出这**k**个参数可得

参数的矩法估计

从上式中解出这**k**个参数可得

$$\begin{cases} \theta_1 = \theta_1(\mu_1, \cdots \mu_k) \\ \theta_2 = \theta_2(\mu_1, \cdots \mu_k) \\ \vdots \\ \theta_k = \theta_k(\mu_1, \cdots \mu_k) \end{cases}$$

用$\mathbf{A_i}$分别代替上式中的$\boldsymbol{\mu}_\mathbf{i}$

$$\left.\begin{cases} \hat{\theta}_1 = \theta_1(A_1, \cdots A_k) \\ \hat{\theta}_2 = \theta_2(A_1, \cdots A_k) \\ \vdots \\ \hat{\theta}_k = \theta_k(A_1, \cdots A_k) \end{cases}\right\}$$

未知参数$\theta_1, \ldots, \theta_k$的矩估计量

代入一组样本值得 k 个数,可得矩估计量的观察值

矩估计值

例2 设某总体X的数学期望为E(X)=μ，方差D(X)=σ²，X_1，X_2，…，X_n为样本，试求μ和σ²的矩估计量。

解 总体的k阶原点矩为 $$\mu_1 = \mu$$

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

由上述两个式子可得 $$\mu = \mu_1 \qquad \sigma^2 = \mu_2 - \mu_1^2$$

样本的k阶原点矩为 $$A_1 = \bar{X} \qquad A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

由矩法估计，令A_i代替μ_i，可得μ和σ²的矩估计量为

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

结论：不管总体**X**服从何种分布，总体期望和方差的矩估计量分别为

$$\hat{\mu}=\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

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

估计值为 $\hat{\mu}=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ $\quad$ $\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$

例3 设X_1，X_2，…，X_n为总体X的样本，试求下列总体分布参数的矩估计量。

$$(1)\ X \sim N\left(\mu,\sigma^2\right)\quad (2)X \sim b(N,p)(N\text{已知})$$

解 （1）由于 $E(X)=\mu \quad D(X)=\sigma^2$

所以参数μ和σ²的矩估计量为

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

（2）由于 $E(X)=Np$

所以 $N\hat{p}=\bar{X}$

得参数p的矩估计量为 $\hat{p}=\frac{1}{N}\bar{X}=\frac{1}{N}\cdot\frac{1}{n}\sum_{i=1}^{n}X_i$

例4 设从某灯泡厂某天生产的灯泡中随机抽取**10**只灯泡，测得其寿命为**(**单位**:**小时**)**

1050， 1100， 1080， 1120， 1200

1250， 1040， 1130， 1300， 1200

试用矩法估计该厂这天生产的灯泡的平均寿命及寿命分布的方差.

解

$$\hat{E}(X) = \bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 1147(h)$$

$$\hat{D}(X) = \hat{\sigma}^2 = \frac{1}{10}\sum_{i=1}^{10} x_i^2 - \bar{x}^2 = 6821(h^2).$$

例5 设总体 $X \sim U[0,\theta]$, θ未知, 求参数 θ的 矩法估计量.

解　　由于 $\mu_1 = E(X) = \dfrac{\theta}{2}$,　　$\therefore \theta = 2\mu_1$

又样本的一阶矩为　　$A_1 = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$,

由矩法估计，令A_1代替θ_1，可得 θ的矩估计量为

$$\hat{\theta} = 2\overline{X}$$

参数的最大似然估计法

思想方法：一次试验就出现的

事件有较大的概率

例如：有两外形相同的箱子，各装100个球

一箱	99个白球	1 个红球
一箱	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：很可能第一箱.

例　设总体 X 服从0-1分布，且$P\{X=1\}=p$，用最大似然法求p的估计值.

解　　总体 X 的概率分布为

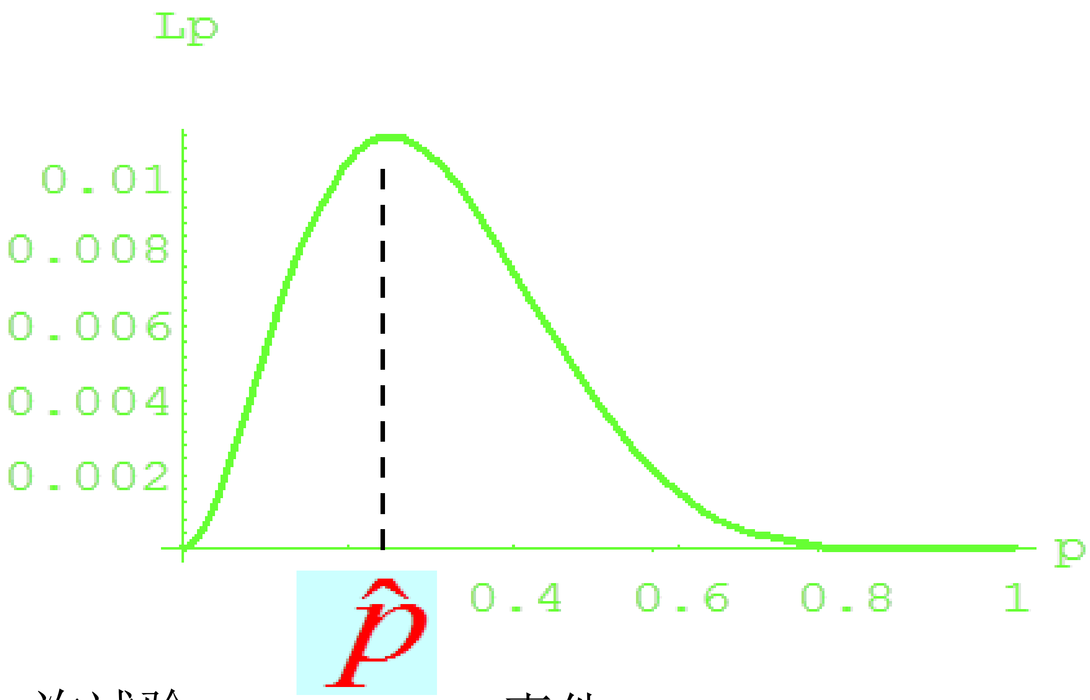
$$P\{X=x\}=p^x(1-p)^{1-x},\quad x=0,1$$

设 $x_1, x_2, \cdots, x_n$为总体样本$X_1, X_2, \cdots, X_n$的样本值，

则 $P\{X_1=x_1, X_2=x_2, \cdots, X_n=x_n\}$

$$=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}=L(p)\quad x_i=0,1, i=1,2,\cdots,n$$

对于不同的 p ， L (p)不同，见右下图



现经过一次试验， 事件

$$\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}$$

发生了， 则 p 的取值应使这个事件发生

的概率最大.

在容许范围内选择 **p**，使**L**(**p**)最大

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。

$$\frac{\mathrm{d}\ln L}{\mathrm{d}p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n-\sum_{i=1}^{n} x_i}{1-p} \overset{令}{=} 0 \quad \Rightarrow \hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

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

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

一般，设 X 为离散型随机变量，其分布律为

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

则样本 $X_1,X_2,\cdots,X_n$的概率分布为

$$P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}$$

$$=f(x_1,\theta)f(x_2,\theta)\cdots f(x_n,\theta)$$

记为 或

$$=L(x_1,x_2,\cdots,x_n,\theta)=L(\theta)\qquad x_i=u_1,u_2,\cdots,$$

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

称 $L(\theta)$ 为样本的似然函数

最大似然法的思想

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

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

称这样得到的

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

为参数 θ 的最大似然估计值

称统计量

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

为参数 θ 的最大似然估计量

注1　　若 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　　未知参数可以不止一个，如 $\theta_1,\ldots,\theta_k$

设 X 的密度(或分布)为 $$f(x,\theta_1,\cdots,\theta_k)$$

则定义似然函数为

$$L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)$$

$$=L(\theta_1,\cdots,\theta_k)=\prod_{i=1}^{n} f(x_i,\theta_1,\cdots,\theta_k)$$

$$-\infty<x_i<+\infty,\ i=1,2,\cdots,n \qquad (\theta_1,\cdots,\theta_k)\in\Theta$$

若 $L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)$ 关于 $\theta_1,\ldots,\theta_k$可微,则称

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

为似然方程组

若对于某组给定的样本值 $x_1,x_2,\cdots,x_n$,

参数 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ 使似然函数取得最大值,即

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

则称 $\hat{\theta}_1,\cdots,\hat{\theta}_k$ 为 $\theta_1,\ldots,\theta_k$ 的**最大似然估计值**

显然，

$$\hat{\theta}_r = g(x_1, x_2, \cdots, x_n) \qquad r = 1,2,\cdots,k$$

称统计量

$$\hat{\theta}_r = g(X_1, X_2, \cdots, X_n) \qquad r = 1,2,\cdots,k$$

为 $\theta_1, \theta_2, \ldots, \theta_k$ 的**最大似然估计量**

参数的最大似然估计法

求解方法：

（1）构造似然函数

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

（2）取自然对数

$$\ln L(x_1,x_2,\cdots,x_n,\theta)=\sum_{i=1}^{n}\ln f(x_i,\theta)$$

（3）令

$$\frac{d\ln L}{d\theta}=0$$

其解$\hat{\theta}$ 即为参数θ的最大似然估计值。

若总体的密度函数中有多个参数θ_1，θ_2，…，θ_n，则将第（3）步改为 $\frac{\partial \ln L}{\partial \theta_i}=0,(i=1,2,\cdots,n)$

解方程组即可。

例 假设（X_1，X_2，⋯，X_n）是取自总体X的样本，X服从参数θ为的指数分布，其概率密度函数为 $f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}},\theta>0,x>0$ 求参数θ的矩估计和最大似然估计量。

解 设$x_1,x_2,\cdots,x_n$是相应于$X_1,X_2,\cdots,X_n$的样本观察值

则似然函数为

$$L(\theta)=\prod_{i=1}^{n}(\frac{1}{\theta}e^{-\frac{x_i}{\theta}})=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i}$$

取对数 $$\ln L(\theta)=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}x_i$$

似然方程为 $$\frac{d\ln L(\theta)}{d\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}x_i=0$$

解　设$x_1,x_2,\cdots,x_n$是相应于$X_1,X_2,\cdots,X_n$的样本观察值

则似然函数为

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

取对数　$$\ln L(\theta)=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}x_i$$

似然方程为　$$\frac{d\ln L(\theta)}{d\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}x_i=0$$

所以θ的最大似然估计量为　$$\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$$

例 假设（X_1，X_2，…，X_n）是取自正态总体$N(\mu,\sigma^2)$的样本，求μ和σ^2的最大似然估计量。

解 设$x_1,x_2,\cdots,x_n$是相应于$X_1,X_2,\cdots,X_n$的样本观察值

构造似然函数

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

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

取对数

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

续解　　求偏导数，并令其为**0**

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

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

解得

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

所以μ，σ²的最大似然估计量为

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

与矩估计量

相同

例：设总体X的概率分布率为：$\begin{pmatrix} 1 & 2 & 3 \\ \theta & \theta/2 & 1-3\theta/2 \end{pmatrix}$，其中$\theta>0$未知，

现得到样本观测值2，3，2，1，3，求θ的矩估计与极大似然估计。

解：(1) 矩估计

$$\mu_1 = E(X) = \sum x_k p_k = \theta + 2\times\theta/2 + 3\times(1-3\theta/2) = 3 - 5\theta/2$$

$\therefore \theta = 2/5(3-\mu_1)$　　用$A_1=\bar{x}=2.2$代替μ_1，可得θ的矩估计为

$$\hat{\theta} = 2/5(3-2.2) = 0.32$$

(2) 极大似然估计

$$L(\theta) = P\{X_1=2, X_2=3, X_3=2, X_4=1, X_5=3\}$$

$$= (\theta/2)(1-3\theta/2)(\theta/2)\theta(1-3\theta/2)$$

$$= \tfrac{1}{16}\theta^3(2-3\theta)^2$$

$$\ln L(\theta) = -\ln 16 + 3\ln\theta + 2\ln(2-3\theta)$$

$$\frac{d\ln L(\theta)}{d\theta} = \frac{3}{\theta} - \frac{6}{2-3\theta} = 0 \qquad \Rightarrow \hat{\theta} = 0.4$$

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

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

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

可得未知参数的最大似然估计值 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$

然后,再求得最大似然估计量.

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

例：设总体X服从$[0,\theta]$上的均匀分布，$\theta>0$未知，

试由样本$x_1,x_2,\cdots,x_n$求出θ的极大似然估计和矩估计。

解：(1) 极大似然估计

因X的概率密度为：$f(x;\theta)=\begin{cases}\frac{1}{\theta} & 0\le x\le\theta \\ 0 & 其它\end{cases}$

故参数θ的似然函数为：$L(\theta)=\begin{cases}\frac{1}{\theta^n} & 0\le x_1,x_2,\cdots,x_n\le\theta \\ 0 & 其它\end{cases}$

由于$\frac{d\ln(\theta)}{d\theta}=-\frac{n}{\theta}\neq 0$,不能用微分法求$\hat{\theta}_L$

以下从定义出发求$\hat{\theta}_L$：

因为$0\le x_i\le\theta$,故θ的取值范围最小为$x_{(n)}=max\{x_1,x_2,\cdots,x_n\}$

又$L(\theta)=\frac{1}{\theta^n}$对$\theta>x_{(n)}$的$\theta$是减函数，$\theta$越小，$L$越大，故$\hat{\theta}_L=x_{(n)}$时，$L$最大；

所以θ的极大似然估计量为$\hat{\theta}_L=X_{(n)}=max\{x_1,x_2,\cdots,x_n\}$

(2) 矩估计

由$E(X)=\int_0^\theta\frac{1}{\theta}xdx=\frac{\theta}{2}=\bar{X}\ \Rightarrow\hat{\theta}=2\bar{X}$