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

一、矩估计

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

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

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

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

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．

二. 最大似然法估计

最大似然法的理论依据---极大似然原理

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。

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

$$L(\theta)=\boldsymbol{P}(X_1=x_1,\cdots,X_n=x_n)$$

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 $\hat{\theta}_L$

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

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

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5　设总体 $X \sim N(\mu,\sigma^2)$，μ,σ^2未知．$x_1,\dots,x_n$ 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 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$-\infty < x < \infty$

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

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

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

令 $t=\sigma^2$

$t=\sigma^2$

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

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

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

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

其中 $\theta>0$,求θ,μ的最大似然估计.

解：似然函数为

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

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

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

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

对 $\mu \le \min x_i, L(\theta,\mu)>0$, 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

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

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。

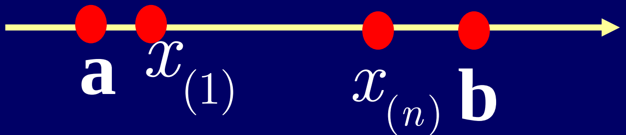
解：令$x_{(1)} = \min\{x_1, x_2, \cdots x_n\}$

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

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

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

§7.2 估计量的评选标准

例 设总体 $X \sim U(a, b)$, a, b 未知

$$\hat{a}_{矩} = \bar{X} - \sqrt{3B_2} \quad \hat{b}_{矩} = \bar{X} + \sqrt{3B_2}$$

$$\hat{a}_{MLE} = X_{(1)}, \quad \hat{b}_{MLE} = X_{(n)}$$

需要讨论以下几个问题:

(1) 我们希望一个"好的"估计量具有什么特性?

(2) 怎样决定一个估计量是否比另一个估计量"好"?

(3) 如何求得合理的估计量?

常用的几条标准是：

1．无偏性

2．有效性

3．均方误差

4．相合性

一、无偏性

定义1.设 $\hat{\theta}(X_1,\cdots,X_n)$ 是未知参数θ的估计量，若

$E(\hat{\theta})=\theta$ 则称$\hat{\theta}$为θ的无偏估计.

例1 设总体X的k阶矩 $\mu_k=E(X^k)$存在 $(X_1,X_2,\cdots,X_n)$是总体X的样本,

证明：$A_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^k$ 是μ_k的无偏估计.

证 由于 $E(X_i^k)=\mu_k \quad i=1,2,\cdots,n$ 因而

$$E(A_k)=E(\frac{1}{n}\sum_{i=1}^{n}X_i^k)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=\frac{1}{n}\cdot n\cdot\mu_k=\mu_k$$

特别地

(1) 样本均值 $\overline{X}$ 是总体期望 $E(X)$ 的无偏估计

(2) 样本二阶原点矩 $A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ 是总体二阶原点矩

$\mu_2 = E(X^2)$ 的无偏估计.

例2 设总体 X 的期望 与方差存在, X 的样本为

$(X_1, X_2, \cdots, X_n)$ $(n > 1)$. 证明

(1) $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ 不是 $D(X)$ 的无偏估计;

(2) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ 是 $D(X)$ 的无偏估计.

证 前已证 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}^2$

$$E(X_i)=E(X)=\mu,\ D(X_i)=D(X)=\sigma^2$$

$$E(\bar{X})=E(X)=\mu,\ D(\bar{X})=\frac{\sigma^2}{n}$$

$$E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)-E(\bar{X}^2)$$

$$=(\sigma^2+\mu^2)-(\frac{\sigma^2}{n}+\mu^2)=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

故 $E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=\sigma^2$

例3 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases} \qquad \theta>0 \text{ 为常数}$$

$(X_1,X_2,\cdots,X_n)$ 为 X 的一个样本

证明 $\overline{X}$ 与 $n\min\{X_1,X_2,\cdots,X_n\}$ 都是 θ 的无偏估计

证 $X\sim Exp\left(\dfrac{1}{\theta}\right)$ $\qquad E(X)=\theta$

故 $E(\overline{X})=E(X)=\theta$

$\overline{X}$ 是 θ 的无偏估计量.

令 $Z=\min\{X_1,X_2,\cdots,X_n\}$

$$F_Z(z)=1-P(X_1>z,X_2>z,\cdots,X_n>z)$$
$$=1-P(X_1>z)P(X_2>z)\cdots P(X_n>z)$$
$$=1-\prod_{i=1}^{n}(1-P(X_i\le z))=\begin{cases}0 & z<0\\ 1-e^{-\frac{nz}{\theta}} & z\ge 0\end{cases}$$

$$f_Z(z)=\begin{cases}0 & z<0\\ \dfrac{n}{\theta}e^{-\frac{nz}{\theta}} & z\ge 0\end{cases}$$

即 $Z\sim Exp\left(\dfrac{n}{\theta}\right)$ $\quad E(Z)=\dfrac{\theta}{n}$ $\quad E(nZ)=\theta$

故 nZ 是θ的无偏估计量.

例4 设总体 $X \sim U(0,\theta)$,其密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta} & 0<x<\theta, \\ 0 & \text{其他}\end{cases}$$

$\theta>0$ 为常数

$(X_1, X_2, \cdots, X_n)$ 为 X 的一个样本

θ 的矩估计量为 $\hat{\theta}_1=2\bar{X}$

θ 的极大似然估计量为 $\hat{\theta}_2=X_{(n)}=\max\{X_1, X_2, \cdots, X_n\}$

证明 $\hat{\theta}_1=2\bar{X}$ 与 $\hat{\theta}=\dfrac{n+1}{n}X_{(n)}$ 都是θ 的无偏估计

证明：$E(\hat{\theta}_1)=2E(\bar{X})=2\times\dfrac{\theta}{2}=\theta$

$\therefore \hat{\theta}_1=2\bar{X}$ 是θ 的无偏估计量.

$$\hat{\theta}_2 = X_{(n)} = \max\{X_1, X_2, \cdots, X_n\} \qquad F_{x_{(n)}}(x) = [F(x)]^n$$

$$\therefore f_{x_{(n)}}(x) = F'_{x_{(n)}}(x) = nf(x)[F(x)]^{n-1} = \begin{cases} \dfrac{n}{\theta}\left(\dfrac{x}{\theta}\right)^{n-1} & 0 < x < \theta \\ 0 & \text{其他} \end{cases}$$

$$E\hat{\theta}_2 = EX_{(n)} = \int_{-\infty}^{\infty} x f_{x_{(n)}}(x)dx = \int_0^{\theta} n\left(\frac{x}{\theta}\right)^n dx = \frac{n}{n+1}\theta$$

$$\therefore E\hat{\theta} = \frac{n+1}{n} EX_{(n)} = \theta$$

$$\therefore \hat{\theta} = \frac{n+1}{n} X_{(n)}$$ 是θ 的无偏估计量.

2、有效性

定义 设 $\hat{\theta}_1 = \theta_1(X_1, X_2, \cdots, X_n)$

$$\hat{\theta}_2 = \theta_2(X_1, X_2, \cdots, X_n)$$

都是总体参数θ的无偏估计量, 且

$$D(\hat{\theta}_1) < D(\hat{\theta}_2)$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.

例5 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases}\qquad \theta>0 \text{ 为常数}$$

由例3可知, $\overline{X}$ 与 $n\min\{X_1,X_2,\cdots,X_n\}$ 都是θ的无偏估计量,问哪个估计量更有效?

解 $\mathrm{Var}(\overline{X})=\dfrac{\theta^2}{n}$, $\mathrm{Var}(n\min\{X_1,X_2,\cdots,X_n\})=\theta^2$

所以, $\overline{X}$ 比 $n\min\{X_1,X_2,\cdots,X_n\}$ 更有效.

例6 设总体 $X \sim U(0,\theta)$,其密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta} & 0<x<\theta, \\ 0 & \text{其他}\end{cases}$$

$\theta>0$ 为常数

$(X_1, X_2, \cdots, X_n)$ 为 X 的一个样本

$\hat{\theta}_1 = 2\bar{X}$ 与 $\hat{\theta} = \dfrac{n+1}{n} X_{(n)}$ 都是 θ 的无偏估计

问哪一个更有效?

解: $D(\hat{\theta}_1) = 4D(\bar{X}) = 4 \times \dfrac{\theta^2}{12n} = \dfrac{\theta^2}{3n}$

$$D(\hat{\theta}) = D\left(\frac{n+1}{n} X_{(n)}\right) = \frac{\theta^2}{n(n+2)}$$

当 $n>2$ 时 $D(\hat{\theta}) < D(\hat{\theta}_1)$

$\therefore \hat{\theta}$ 较 $\hat{\theta}_1$ 有效。

例7 设总体 X，且 $E(X)=\mu$，$D(X)=\sigma^2$

$(X_1, X_2, \cdots, X_n)$为总体 X 的一个样本

(1) 设常数 $c_i \neq \frac{1}{n}$ $i=1,2,\cdots,n.$ $\sum_{i=1}^{n} c_i = 1.$

证明 $\hat{\mu}_1 = \sum_{i=1}^{n} c_i X_i$ 是 μ 的无偏估计

(2) 证明 $\hat{\mu} = \overline{X}$ 比 $\hat{\mu}_1 = \sum_{i=1}^{n} c_i X_i$ 更有效

证 (1) $E(\hat{\mu}_1) = \sum_{i=1}^{n} c_i E(X_i) = \sum_{i=1}^{n} c_i \mu = \mu$

(2) $$D(\hat{\mu}_1)=\sum_{i=1}^{n}c_i^2 D(X_i)=\sigma^2\sum_{i=1}^{n}c_i^2$$

而 $$1=\left(\sum_{i=1}^{n}c_i\right)^2=\sum_{i=1}^{n}c_i^2+2\sum_{1\le i<j\le n}c_ic_j$$

$$<\sum_{i=1}^{n}c_i^2+\sum_{1\le i<j\le n}(c_i^2+c_j^2)\le n\sum_{i=1}^{n}c_i^2$$

$$\Rightarrow \sum_{i=1}^{n}c_i^2>\frac{1}{n} \Rightarrow D(\hat{\mu})=\frac{1}{n}\sigma^2<D(\hat{\mu}_1)$$

结论 算术均值比加权均值更有效.

例如 $X \sim N(\mu,\sigma^2)$，(X_1,X_2) 是一样本.

$$\left.\begin{aligned}\hat{\mu}_1 &= \frac{2}{3}X_1+\frac{1}{3}X_2 \\ \hat{\mu}_2 &= \frac{1}{4}X_1+\frac{3}{4}X_2 \\ \hat{\mu}_3 &= \frac{1}{2}X_1+\frac{1}{2}X_2\end{aligned}\right\}\text{都是}\mu\text{的无偏估计量}$$

由例7(2) 知 $\hat{\mu}_3$ 最有效.

3、均方误差

目 的： 对 有 偏 估 计 进 行评 价

均 方 误 差： $\mathrm{MSE}(\hat{\theta}) = E(\hat{\theta} - \theta)^2$

$$\mathrm{MSE}(\hat{\theta}) = E[(\hat{\theta} - E\hat{\theta}) + (E\hat{\theta} - \theta)]^2$$

$$= E(\hat{\theta} - E\hat{\theta})^2 + (E\hat{\theta} - \theta)^2$$

$$+ 2E[(\hat{\theta} - E\hat{\theta})(E\hat{\theta} - \theta)]$$

$$= Var(\hat{\theta}) + (E\hat{\theta} - \theta)^2.$$

4、相合性

定义 设 $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \cdots, X_n)$是总体参数$\theta$的估计量. 若对于任意的$\theta \in \Theta$, 当$n \to \infty$时, $\hat{\theta}_n$依概率收敛于θ, 即 $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} P(|\hat{\theta}_n - \theta| \geq \varepsilon) = 0$$

则称$\hat{\theta}_n$是总体参数θ的一致(或相合)估计量.

相合估计量仅在样本容量 n 足够大时, 才显示其优越性.

由辛钦定理

若总体 X 的数学期望 $E(X)=\mu$ 有限, 则有

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

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

故

$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$ 为 $E(X^k)=\mu_k \ (k=1,2,\cdots)$ 的相合估计量.

$g(A_1,A_2,\cdots,A_k)$ 为 $g(\mu_1,\mu_2,\cdots,\mu_k)$ 的相合估计量.

关于相合性的两个常用结论

1\. 样本 k 阶原点矩是总体 k 阶原点矩的相合估计. 由大数定律证明

2\. 设 $\hat{\theta}_n$ 是 θ 的无偏估计量, 且 $\lim\limits_{n\to\infty} \mathrm{Var}(\hat{\theta}_n) = 0$ 则 $\hat{\theta}_n$ 是 θ 的相合估计. 用切比雪夫不等式证明

例8 $X \sim f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases}$ $\theta>0$ 为常数

则 $\overline{X}$ 是θ的无偏、相合估计.

证 $E\overline{X}=\theta$

$$\lim_{n\to\infty} D(\overline{X}) = \lim_{n\to\infty}\frac{\theta^2}{n}=0$$

$$\therefore \overline{X} \xrightarrow{P} \theta$$

所以 $\overline{X}$是 θ 的相合估计, 证毕.

设总体 $X \sim N(0, \sigma^2)$，它的概率密度函数为

$$f(x;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$，其中 σ^2 为未知参数.

设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本

（1）求 σ^2 的极大似然估计量 $\widehat{\sigma_1}^2$；　　（2）求 σ^2 的矩估计量 $\widehat{\sigma_2}^2$.

（3）证明 $\widehat{\sigma_1}^2$ 是 σ^2 的无偏、相合估计；

(1) 求 σ^2 的极大似然估计量 $\hat{\sigma}_1^2$；　　(2) 求 σ^2 的矩估计量 $\hat{\sigma}_2^2$.

(3) 证明 $\hat{\sigma}_1^2$ 是 σ^2 的无偏、相合估计；

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

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

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

$\frac{d\ln L(\sigma^2)}{d\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}x_i^2=0$

$\therefore\ \hat{\sigma}_1^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2$

(3) $E(\hat{\sigma}_1^2)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)=\frac{1}{n}\sum_{i=1}^{n}\sigma^2=\sigma^2$

$\therefore\ \hat{\sigma}_1^2$ 是 σ^2 的无偏估计

由辛钦大数定律 $\hat{\sigma}_1^2=A_2=\frac{1}{n}\sum_{i=1}^{n}x_i^2\xrightarrow{P}E(X^2)=\sigma^2$　$\therefore\ \hat{\sigma}_1^2$ 是 σ^2 的相合估计

(2) $E(X)=0$

$E(X^2)=\int_{-\infty}^{+\infty}x^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx$

$=\sigma^2=\mu_2$

$\therefore\ \sigma^2$ 的矩估计量为

$\hat{\sigma}_2^2=A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$.