Chap 8 参数估计

统计推断的基本问题

- 参数估计 (Chap8)
  - 点估计
  - 区间估计
- 假设检验 (Chap9)

§ 8.1　参数的点估计

什么是参数估计？

例如，$X \sim N(\mu, \sigma^2)$,

若μ, σ^2未知，在抽取样本、获取样本观测值之后，

通过构造样本的函数(统计量), 给出它们的估计值或取值区间，就是参数估计的内容.

点估计　区间估计

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，使得这个范围包含未知参数真值的概率为给定的值.

§8.1 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，但它含有一个或多个未知参数：$\theta_1,\theta_2,\ldots,\theta_k$

设 $X_1, X_2,\ldots, X_n$为总体的一个样本

构造 k 个统计量：

$$\left.\begin{array}{l}
\theta_1(X_1,X_2,\cdots,X_n) \\[2ex]
\theta_2(X_1,X_2,\cdots,X_n) \\[2ex]
\cdots\cdots\cdots\cdots\cdots\cdots \\[2ex]
\theta_k(X_1,X_2,\cdots,X_n)
\end{array}\right\}\text{随机变量}$$

当测得一组样本值$(x_1, x_2, \ldots, x_n)$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$\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k$ 为未知参数$\theta_1, \theta_2, \cdots, \theta_k$ 的估计值

称对应的统计量 $\theta_1(X_1, X_2, \cdots, X_n)$, $\theta_2(X_1, X_2, \cdots, X_n)$,

$\cdots$, $\theta_k(X_1, X_2, \cdots, X_n)$

为未知参数 $\theta_1, \theta_2, \cdots, \theta_k$ 的估计量

问题

如何构造统计量？

如何评价估计量的好坏？

常用的点估计方法

□ 矩法

方法： 用样本的 k 阶矩作为总体的 k 阶矩的估计量，建立含有待估计参数的方程，从而可解出待估计参数

即：

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to E(X)$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \to E(X^2)$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^3 \to E(X^3)$$

记作：

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \hat{E}(X)$$

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

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

由
$$\begin{cases} \dfrac{1}{n}\sum_{i=1}^{n} X_i = \hat{E}(X) \\ \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 = \hat{E}(X^2) \end{cases}$$

$$\Longrightarrow \quad \hat{\mu} = \hat{E}(X) = \bar{X}$$

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

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

即：以样本均值估计总体均值，

以样本的2阶中心距估计方差

例2 设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \ldots, X_n$ 为总体的样本，求 μ, σ^2 的矩法估计量。

解 $$\hat{\mu}_{矩} = \hat{E}(X) = \bar{X}$$

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

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

例3 设总体 $X \sim E(\lambda), X_1, X_2, \ldots, X_n$为总体的样本，求$\lambda$的矩法估计量。

解 $E(X) = \dfrac{1}{\lambda}$ 则 $\hat{E}(X) = \dfrac{1}{\hat{\lambda}} = \overline{X}$

故 $\hat{\lambda} = \dfrac{1}{\overline{X}}$

❑ 点估计的极大似然估计法

思想方法：一次试验就出现的事件有较大的概率

例如: 有两个外形相同的箱子,都装有100个球

第一箱 99个白球， 1个红球

第二箱 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 第一箱.

例6 设总体 X 服从0-1分布，且$P(X=1)=p$, 用极大似然法求 p 的估计值。

解 X 的概率分布为

$$P(X=1)=p,$$

$$P(X=0)=1-p,$$

可以写成

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

设 $X_1, X_2, \ldots, X_n$ 为总体 X 的样本,

设 $x_1, x_2, \ldots, x_n$ 为总体 X 的样本值,

则
$$P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)$$

$$= P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \cdots \cdot P(X_n = x_n)$$

$$= p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \cdots \cdot p^{x_n}(1-p)^{1-x_n}$$

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

$$\equiv L(p)$$

$$x_i = 0,1, i = 1,2,\cdots,n$$

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

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

所以 $\hat{p}=\bar{x}$ 为所求 p 的估计值.

一般地，设X为离散型随机变量，其分布律为

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

$X_1, X_2,\ldots, X_n$为总体 X 的样本,

$x_1, x_2,\ldots, x_n$为总体 X 的样本值,

则$X_1, X_2,\ldots, X_n$的概率分布为

$$P(X_1=x_1,X_2=x_2,\cdots,X_n=x_n)$$

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

$$\overset{\text{记为}}{=} L(x_1,x_2,\cdots,x_n,\theta)\overset{\text{或}}{=}L(\theta)\quad x_i=u_1,u_2,\cdots,$$

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

称$L(\theta)$为样本的似然函数

当给定一组样本值时，$L(\theta)$就是参数θ的函数，极大似然估计法的思想就是：

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

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

则称这样得到的$\hat{\theta}$

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

为参数 θ 的极大似然估计值

称统计量

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

为参数 θ的极大似然估计量

注1 若随机变量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 未知参数个数可以不止一个, 如θ_1, θ_2,…, θ_k

设X的密度函数(或分布率)为 $f(x,\theta_1,\theta_2,\cdots,\theta_k)$

则定义似然函数为

$$L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)$$

$$=L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} f(x_i,\theta_1,\theta_2,\cdots,\theta_k)$$

$$-\infty<x_i<+\infty, i=1,2,\cdots,n \quad (\theta_1,\theta_2,\cdots,\theta_k)\in\Theta$$

若 $L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)$ 关于 $\theta_1,\ \theta_2,\ldots,\ \theta_k$ 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 $x_1, x_2,\ldots, x_n$，参数 $\hat\theta_1,\hat\theta_2,\cdots,\hat\theta_k$ 使得似然函数取得最大值，即

$$L(x_1,x_2,\cdots,x_n;\hat\theta_1,\hat\theta_2,\cdots,\hat\theta_k)$$

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

则称 $\hat\theta_1,\hat\theta_2,\cdots,\hat\theta_k$ 为 $\theta_1,\ \theta_2,\ldots,\ \theta_k$ 的极大似然估计值

显然，

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

称统计量

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

为$\theta_1, \theta_2, \ldots, \theta_k$ 的极大似然估计量

例7 设总体 $X \sim N(\mu, \sigma^2)$，$x_1, x_2, \ldots, x_n$ 是 X 的一组样本值，求 μ, σ^2 的极大似然估计.

解 $L(x_1, x_2, \cdots, x_n; \mu, \sigma^2)$

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

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

似然方程组为

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

$$\Rightarrow \begin{cases} \hat{\mu}_{mle} = \dfrac{1}{n}\sum_{i=1}^{n}x_i = \bar{x} \\ \hat{\sigma^2}_{mle} = \dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \end{cases}$$

μ, σ^2 的极大似然估计量分别为

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

求未知参数的极大似然估计值(量)的方法

1）写出似然函数L

2）求出 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ ，使得

$$L(x_1,x_2,\cdots,x_n;\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k)$$

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

若 L 是 $\theta_1,\theta_2,\cdots,\theta_k$ 的可微函数, 解似然方程组

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

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

可求得未知参数的极大似然估计值 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$

然后, 再求得极大似然估计量.

若 L 不是$\theta_1,\theta_2,\cdots,\theta_k$ 的可微函数, 需用其它方法求极大似然估计值. 请看下例：

例8 设 $X \sim U(a,b)$, $x_1, x_2, \ldots, x_n$ 是 X 的一个样本, 求 a, b 的极大似然估计值与极大似然估计量.

解 X 的密度函数为

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

似然函数为

$$L(x_1,x_2,\cdots,x_n;a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a<x_i<b,\ i=1,2,\cdots,n\\ 0, & \text{其它}\end{cases}$$

似然函数只有当 $a < x_i < b,\ i = 1,2,\ldots,n$ 时才能获得最大值, 且 a 越大, b 越小, L 越大.

令
$$x_{\min} = \min\ \{x_1, x_2,\ldots, x_n\}$$
$$x_{\max} = \max\ \{x_1, x_2,\ldots, x_n\}$$

取
$$\hat{a} = x_{\min}, \qquad \hat{b} = x_{\max}$$

则对满足$a \le x_{\min} \le x_{\max} \le b$的一切 $a < b$, 都有

$$\frac{1}{(b-a)^n} \le \frac{1}{(x_{\max} - x_{\min})^n}$$

故 $\hat{a} = x_{\min}, \ \hat{b} = x_{\max}$

是 a , b 的极大似然估计值.

$$X_{\min} = \min\{X_1, X_2, \cdots, X_n\}$$

$$X_{\max} = \max\{X_1, X_2, \cdots, X_n\}$$

分别是 a , b 的极大似然估计量.

极大似然估计值的不变性原理

设 $\hat{\theta}$ 是 θ 的极大似然估计值, $u(\theta)$ $(\theta \in \Theta)$ 是 θ 的函数，且具有单值的反函数 $\theta = \theta(u)$, $u \in U$ 则 $\hat{u} = u(\hat{\theta})$ 是 $u(\theta)$ 的极大似然估计值.

如：在正态分布总体$N(\mu,\sigma^2)$中，σ^2的极大似然估计值为

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

$\sigma = \sqrt{\sigma^2}$ 是σ^2的单值函数，且具有单值的反函数，故σ的极大似然估计值为

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

$\lg\sigma$的极大似然估计值为

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

$$P(X > t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}} e^{-\frac{(x-\overline{x})^2}{2\cdot\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}} dx$$

矩估计就不具有这个性质.

例如 设 X 的密度函数为

$$f(x)=\begin{cases}\sqrt{\dfrac{2}{\pi}}\dfrac{1}{\sigma}e^{-\frac{x^2}{2\sigma^2}} & x>0\\ 0 & x\le 0\end{cases}$$

$$E(X)=\sqrt{\frac{2}{\pi}}\sigma,\ D(X)=\left(1-\frac{2}{\pi}\right)\sigma^2$$

$X_1, X_2, \ldots, X_n$ 为总体的样本

由矩法，令

$$E(X)=\sqrt{\frac{2}{\pi}}\,\sigma=\overline{X}$$

$$E(X^2)=D(X)+E^2(X)=\sigma^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$$

得σ与σ^2的矩法估计量为

$$\hat{\sigma}=\sqrt{\frac{\pi}{2}}\,\overline{X}$$

$$\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}X_i^2\neq(\hat{\sigma})^2$$

——不具有不变性

§ 8.2 点估计的评价标准

对于同一个未知参数, 不同的方法得到的估计量可能不同,于是提出问题

1、应该选用哪一种估计量?

2、用什么标准来评价一个估计量的好坏?

常用标准

(1) 无偏性

(2) 有效性

(3) 一致性

无偏性

定义 设 $(X_1, X_2, \cdots, X_n)$ 是总体X的样本，

$\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ 是总体参数θ的估计量，

$E(\hat{\theta})$存在, 如果：

$$E(\hat{\theta}) = \theta$$

则称 $\hat{\theta}$ 是θ的无偏估计量.

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，

求证：不论 X 服从什么分布,

$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 μ_k的无偏估计量

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

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

$$= \frac{1}{n}\cdot n \cdot \mu_k = \mu_k$$

特别地,

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

样本二阶原点矩 $A_2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2$

是总体二阶原点矩 $\mu_2 = E(X^2)$ 的无偏估计量

例2 设总体 X 的期望 $E(X)$与方差 $D(X)$存在, $(X_1, X_2, \cdots, X_n)$ 是 X 的一个样本, $n > 1$． 求证：

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

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

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

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

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

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

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

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

$$E(\overline{X}^2)=E^2(\overline{X}_i)+D(\overline{X})$$
$$=\mu^2+\frac{\sigma^2}{n}$$

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

$$=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

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

通常，在估计时：
确实以样本均值估计总体均值

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

但却不以2阶中心距 S_n^2 估计总体方差，
而是以样本方差 S^2 估计总体方差

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

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 设总体期望为 $E(X)=\mu$, 方差 $D(X)=\sigma^2$,

$(X_1, X_2, \cdots, X_n)$为总体X的一个样本，

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

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

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

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

$$=\sum_{i=1}^{n}c_i\mu=\mu\sum_{i=1}^{n}c_i=\mu$$

(2) $$D(\hat{\mu})=D(\bar{X})=\frac{1}{n}\sigma^2$$

$$D(\hat{\mu}_1)=D(\sum_{i=1}^{n}c_iX_i)=\sum_{i=1}^{n}D(c_iX_i)$$

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

证明 $\hat{\mu}=\bar{X}$ 比 $\hat{\mu}_1=\sum_{i=1}^{n}c_iX_i$ 更有效

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

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

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

$$\Rightarrow \quad \frac{1}{n}\le \sum_{i=1}^{n} c_i^2$$

"=" 成立当且仅当 $c_1 = c_2 = \ldots = c_n$

$$\Rightarrow \quad D(\hat{\mu}) < D(\hat{\mu}_1)$$

结论 算术均值比加权均值更有效.

算术均值是所有线性无偏估计中方差最小的，称为最小方差无偏估计.

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

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

例7(2) 知 $\hat{\mu}_3$ 最有效.

一致性

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

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

则称 $\hat{\theta}$ 是总体参数θ的一致(或相合)估计量.

一致性估计量仅在样本容量 n 足够大时,才显示其优越性.

例6 样本k阶矩 $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ 是总体k阶矩 $E(X^k)$ 的一致估计

证明：$$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^k) = E(X^k)$$

故，$$1 \geq P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i^k - E(X^k)\right| < \varepsilon\right\} \geq 1 - \frac{D\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right)}{\varepsilon^2}$$

$$= 1 - \frac{D\left(\sum_{i=1}^{n} X_i^k\right)}{n^2\varepsilon^2} = 1 - \frac{\sum_{i=1}^{n} D(X_i^k)}{n^2\varepsilon^2} = 1 - \frac{nD(X^k)}{n^2\varepsilon^2} \to 1$$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i^k - E(X^k)\right| < \varepsilon\right\} = 1$$

例7 正态分布总体 $N(\mu,\sigma^2)$ 的样本方差 S^2 是总体方差 σ^2 的一致估计

证明： $E(S^2)=\sigma^2$ 故，

$$1\geq P\left\{\left|S^2-\sigma^2\right|<\varepsilon\right\}=P\left\{\left|S^2-E(S^2)\right|<\varepsilon\right\}\geq 1-\frac{D(S^2)}{\varepsilon^2}$$

其中， $D(S^2)=D\left[\frac{(n-1)S^2}{\sigma^2}\cdot\frac{\sigma^2}{n-1}\right]$

$$=(\frac{\sigma^2}{n-1})^2\cdot D\left[\frac{(n-1)S^2}{\sigma^2}\right]=(\frac{\sigma^2}{n-1})^2\cdot 2(n-1)=\frac{2\sigma^4}{n-1}$$

故， $1\geq P\left\{\left|S^2-\sigma^2\right|<\varepsilon\right\}\geq 1-\frac{2\sigma^4}{(n-1)\varepsilon^2}$

$$\lim_{n\to\infty}P\left\{\left|S^2-\sigma^2\right|<\varepsilon\right\}=1$$

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、一致估计量.

证 由例3 知 $\overline{X}$ 是θ的无偏估计量.

$$1\ge P\left\{\left|\overline{X}-\theta\right|<\varepsilon\right\}\ge 1-\frac{D(\overline{X})}{\varepsilon^2}=1-\frac{\theta^2}{n\varepsilon^2}$$

$$\lim_{n\to\infty}P\left\{\left|\overline{X}-\theta\right|<\varepsilon\right\}=1$$

所以 $\overline{X}$ 是 θ 的一致估计量, 证毕.

§ 8.3　区间估计

引例：已知 $X \sim N(\mu, \sigma^2)$，σ^2已知，μ未知

$x_1, x_2, \ldots, x_n$ 是一组样本值。

估计未知参数μ

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{\sqrt{\sigma^2 / n}} \sim N(0, 1)$$

则 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right|>z_{1-\alpha/2}\right)=\alpha$$

故 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right|\le z_{1-\alpha/2}\right)=1-\alpha$$

即 $$P\left(\bar{X}-z_{1-\alpha/2}\sqrt{\sigma^2/n}\le\mu\le\bar{X}+z_{1-\alpha/2}\sqrt{\sigma^2/n}\right)=1-\alpha$$

称随机区间 $$\left(\bar{X}-z_{1-\alpha/2}\sqrt{\sigma^2/n},\ \bar{X}+z_{1-\alpha/2}\sqrt{\sigma^2/n}\right)$$

为末知参数 μ 的置信度为1-α的置信区间.

置信区间的意义

$$\left(\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n},\ \bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n}\right)$$

反复抽取容量为n的样本， 都可得到一个区间,这个区间可能包含未知参数 μ 的真值,也可能不包含未知参数的真值, 包含真值的区间占1-α.

$\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n}$ —— μ 的置信下限

$\bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n}$ —— μ 的置信上限

$1-\alpha$ —— 置信度

置信区间的定义

设 θ 是一个待估计的参数, α 是一给定的数, $(0 < \alpha < 1)$. 若能找到两个统计量

$$\hat{\theta}_1(X_1, X_2, \cdots, X_n),\quad \hat{\theta}_2(X_1, X_2, \cdots, X_n)$$

使得 $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$

则称随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 为参数 θ 的置信度为 $1 - \alpha$ 的置信区间, 分别称 $\hat{\theta}_1, \hat{\theta}_2$ 为置信下限与置信上限, $1 - \alpha$ 称为置信水平或置信度.

几点说明 $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1-\alpha$

- 置信区间的长度 $\hat{\theta}_2 - \hat{\theta}_1$ 反映了估计的精度

  $\hat{\theta}_2 - \hat{\theta}_1$ 越小，估计的精度越高.

例：$(\bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n}) - (\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n}) = 2 \cdot z_{1-\alpha/2}\sqrt{\sigma^2/n}$

- α 反映了估计的可靠程度，α 越小，越可靠.

  α 越小，$1-\alpha$ 越大，估计的可靠程度越高，但

  这时，$\hat{\theta}_2 - \hat{\theta}_1$ 往往增大，因而估计的精度降低.

- α 确定后，置信区间的选取方法不唯一，常选长度最小的一个.

求置信区间的步骤

- 寻找一个样本的函数

$$g(X_1, X_2, \cdots, X_n, \theta)$$ —称为**枢轴量**

例如
$$\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} = g(X_1, X_2, \cdots, X_n, \mu) \sim N(0,\ 1)$$

它含有待估参数，不含其它未知参数，它的分布已知，且分布不依赖于待估参数(常由θ的点估计出发考虑).

❑ 给定置信度 $1-\alpha$，定出两个常数 a , b ,使得

$$P(a < g(X_1, X_2, X_n, \theta) < b) = 1-\alpha$$

例如

$$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right| \le z_{1-\alpha/2}\right) = 1-\alpha$$

❑ 由 $a < g(X_1, X_2, X_n, \theta) < b$ 解出

$$\underline{\theta}(X_1, X_2, \cdots, X_n) \quad \bar{\theta}(X_1, X_2, \cdots, X_n)$$

得置信区间 $(\underline{\theta}, \bar{\theta})$

例如 $\left(\bar{X}-z_{1-\alpha/2}\sqrt{\sigma^2/n}, \bar{X}+z_{1-\alpha/2}\sqrt{\sigma^2/n}\right)$

置信区间常用公式

(一) 一个正态总体$X \sim N(\mu, \sigma^2)$的情形

(1) 方差σ^2已知, μ的置信区间

$$\left(\overline{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) \quad \cdots\cdots\cdots\cdots \quad (1)$$

推导 由 $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ 选取枢轴量

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

由 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right|>z_{1-\alpha/2}\right)=\alpha$$

得 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right|\le z_{1-\alpha/2}\right)=1-\alpha$$

即 $$P\left(\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

故 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\quad \overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

(2) 方差σ^2未知，μ的置信区间

$$\left(\bar{X}-t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\ \bar{X}+t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right) \quad \cdots\cdots\cdots \quad (2)$$

推导　选取枢轴量 $T=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim T(n-1)$

由 $P\left(\left|\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\right|\geq t_{1-\frac{\alpha}{2}}(n-1)\right)=\alpha$

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

知 $$P\left(\left|\frac{\bar{X}-\mu}{S/\sqrt{n}}\right|<t_{1-\frac{\alpha}{2}}(n-1)\right)=1-\alpha$$

$$P\left(\bar{X}-t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}<\mu<\bar{X}+t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)=1-\alpha$$

故 μ 的置信区间为

$$\left(\bar{X}-t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\ \bar{X}+t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)$$

用到：设总体$X \sim N(\mu，\sigma^2)$，$(X_1，X_2，\ldots\ldots X_n)$是总体 X 的样本，S^2为样本方差，证明：

$$T=\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

证明：

$$X_i \sim N(\mu,\sigma^2)$$

$$\bar{X} \sim N(\mu,\frac{\sigma^2}{n})$$

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/ n-1}} \sim t(n-1)$$

(3) 当 μ 已知时，方差 σ^2 的 置信区间

取枢轴量 $Q=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$

由概率

$$P\left(\chi^2_{\frac{\alpha}{2}}(n)<\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}<\chi^2_{1-\frac{\alpha}{2}}(n)\right)=1-\alpha$$

得 σ^2 的置信度为 $1-\alpha$ 置信区间为

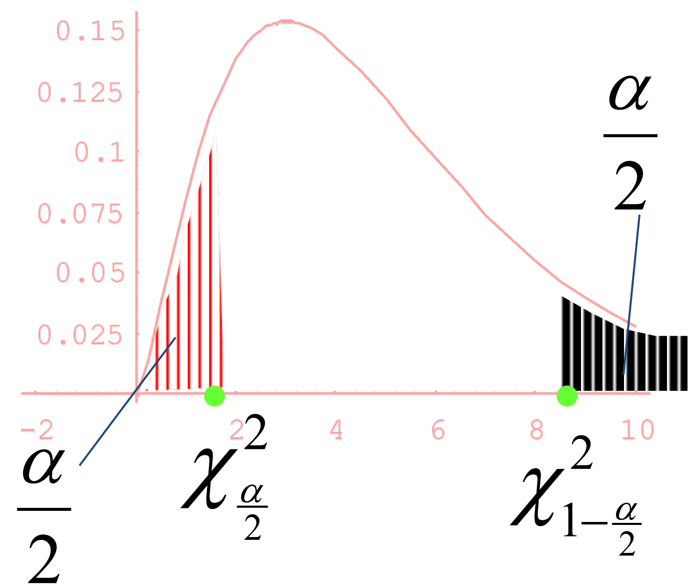
$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)},\ \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right) \quad \cdots\cdots\cdots \quad (3)$$

(4) 当 μ 未知时，方差 σ^2 的置信区间

P177 Th3

选取 $K=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ 则由

$$P\left(\chi^2_{\frac{\alpha}{2}}<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{1-\frac{\alpha}{2}}\right)=1-\alpha$$



得 σ^2 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\ \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right) \quad \cdots\cdots\cdots \quad (4)$$

例1 某工厂生产一批滚珠, 其直径 X 服从正态分布$N(\mu,\sigma^2)$, 现从某天的产品中随机抽取6件,测得直径为

$$15.1,\ 14.8,\ 15.2,\ 14.9,\ 14.6,\ 15.1$$

(1) 若$\sigma^2=0.06$, 求 μ 的置信度为95%的 置信区间;

(2) 若σ^2未知,求 μ 的置信度为95%的置信区间;

(3) 求方差σ^2的置信度为95%的置信区间.

解 (1)若$\sigma^2=0.06$，求μ的置信度为95%的置信区间;

$$\left(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\quad \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

$\sigma^2=0.06$　　$n=6$　　$1\text{-}\alpha=0.95$

$$z_{1-\frac{\alpha}{2}}=z_{1-0.025}=1.96$$

由给定数据算得 $\bar{x}=\dfrac{1}{6}\sum_{i=1}^{6}x_i=14.95$

由公式(1) 得 μ的置信区间为

$$\left(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\quad \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

$$=(14.75,\quad 15.15)$$

(2) 若σ^2未知,求 μ 的置信度为95%的置信区间

$$\left(\bar{x}-\frac{s}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1),\quad \bar{x}+\frac{s}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)\right)$$

查表得 $t_{1-0.025}(5)=2.5706$

由给定数据算得 $\bar{x}=14.95$

$$s^2=\frac{1}{5}\left(\sum_{i=1}^{6}x_i^2-6\bar{x}^2\right)=0.051. \quad s=0.226$$

由公式 (2) 得 μ 的置信区间为

$$\left(\bar{x}-\frac{s}{\sqrt{6}}t_{0.025}(5),\quad \bar{x}+\frac{s}{\sqrt{6}}t_{0.025}(5)\right)=(14.71,\quad 15.187)$$

(3) 求方差σ^2的置信度为95%的置信区间

$$\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right)$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = 0.051.$$

查表得 $\chi^2_{0.025}(5) = 0.831,\quad \chi^2_{0.975}(5) = 12.833$

由公式 (4) 得 μ 的置信区间为

$$\left(\frac{5s^2}{\chi^2_{0.975}(5)}, \quad \frac{5s^2}{\chi^2_{0.025}(5)} \right) = (0.0199, \quad 0.3069)$$

(二) 单侧置信区间

定义 对于给定的 $\alpha\,(0<\alpha<1)$，θ 是待估参数

$(X_1, X_2, \cdots, X_n)$ 是总体 X 的样本,

若能确定一个统计量

$$\underline{\theta}=\underline{\theta}\,(X_1, X_2, \cdots, X_n)\quad(\text{或 } \overline{\theta}=\overline{\theta}\,(X_1, X_2, \cdots, X_n))$$

使得 $P(\theta>\underline{\theta})=1-\alpha$ （ 或 $P(\theta<\overline{\theta})=1-\alpha$ ）

则称 $(\underline{\theta}, +\infty)$ （或 $(-\infty, \overline{\theta})$）

为置信度为 $1-\alpha$ 的单侧置信区间.

$\underline{\theta}$ —— 单侧置信下限 $\overline{\theta}$ —— 单侧置信下限

例3 已知灯泡的寿命X服从正态分布，从灯泡中随机地抽取 5 只作寿命试验，测得寿命(以小时记) 为

1050， 1100， 1120， 1250， 1280

求灯泡寿命均值的置信度为0.95的单侧置信下限与灯泡寿命方差的置信度为0.95的单侧置信上限

解

$X \sim N(\mu, \sigma^2)$， μ, σ^2 未知

(1) 选取枢轴量 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

$$故\, P\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{1-\alpha}(n-1)\right)=1-\alpha$$

$$即\, P\left(\mu \ge \bar{X}-t_{1-\alpha}(n-1)\cdot S/\sqrt{n}\right)=1-\alpha$$

$n=5$, $\bar{x}=1160$, $s^2=\frac{1}{4}(\sum_{i=1}^{5} x_i^2-5\bar{x})=9950$

$$t_{1-\alpha}(n-1)=t_{1-0.05}(4)=2.1318$$

$$\underline{\mu}=\bar{x}-t_{1-0.05}(4)\times\frac{s}{\sqrt{5}}=1064.9$$

(2) 选取枢轴量 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$$\text{故}\quad P\left(\frac{(n-1)S^2}{\sigma^2} > \chi^2_{\alpha}(n-1)\right) = 1-\alpha$$

$$\text{即}\quad P\left(\sigma^2 < \frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}\right) = 1-\alpha$$

$$n = 5, \quad s^2 = \frac{1}{4}\left(\sum_{i=1}^{5} x_i^2 - 5\overline{x}\right) = 9950$$

$$\chi^2_{0.05}(4) = 0.711$$

$$\overline{\sigma^2} = \frac{4s^2}{\chi^2_{0.05}(4)} = 55977$$