Lec-22. 估计量的评价准则, 区间估计

主讲教师: 吴利苏 (**wulisu@sdust.edu.cn**)
主　　页: wulisu.cn

本次课内容

估计量的评价准则

- 无偏性准则
- 有效性准则
- 相合性准则

区间估计

枢轴量（补充）

对总体的未知参数可用不同方法求得不同的估计量，比如矩估计和极大似然估计. 如何评价不同估计量的好坏?

常用的评价准则有如下三条:

(1) 无偏性准则

(2) 有效性准则

(3) 相合性准则

定义 (无偏性准则)

设参数 θ 的估计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, ..., X_n)$, 若

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

则称 $\hat{\theta}$ 是 θ 的一个无偏估计量.

定义 (无偏性准则)

设参数 θ 的估计量 $\hat{\theta}=\hat{\theta}(X_1, X_2, \ldots, X_n)$, 若

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

则称 $\hat{\theta}$ 是 θ 的一个无偏估计量.

- 若 $E(\hat{\theta})\neq\theta$, 则 $E\left(\hat{\theta}\right)-\theta$ 称为估计量 $\hat{\theta}$ 的系统误差.
- 若

$$\lim_{n\to\infty} E\left(\hat{\theta}\right)=\theta,$$

则称 $\hat{\theta}$ 是 θ 的渐近无偏估计量.

无偏性: $E\left(\hat{\theta}\right) = \theta$

- 无偏性的统计意义是指在大量重复试验下，由$\hat{\theta}(X_1, \cdots, X_n)$ 给出的估计的平均恰是θ, 从而无偏性保证了 $\hat{\theta}$没有系统误差.

例

工厂长期为商家提供某种商品，假设生产过程相对稳定，产品合格率为 θ, 虽然一批货的合格率可能会高于 θ , 或低于 θ, 但无偏性能够保证在较长一段时间内合格率接近 θ，所以双方互不吃亏。

但作为顾客购买商品，只有二种可能，即买到的是合格品或不合格品，此时无偏性没有意义。

例

设总体 X 的一阶和二阶矩存在，

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

(1) 证明：样本均值 $\overline{X}$ 和样本方差 S^2 分别是 μ 和 σ^2 的无偏估计；

(2) 判断：B_2 是否为 σ^2 的无偏估计？是否为 σ^2 的渐近无偏估计？

(1) 证: 因 $X_1, X_2, \cdots, X_n$ 与 X 同分布, 故有:

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

故 $\overline{X}$ 是 μ 的无偏估计.

$$E\left(S^2\right) = \sigma^2$$

故 S^2 是 σ^2 的无偏估计.

(2) $B_2 = \frac{n-1}{n} S^2$

$$E(B_2) = \frac{n-1}{n} E\left(S^2\right) = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

故 B_2 不是 σ^2 的无偏估计.

$$\lim_{n\to\infty} E(B_2) = \lim_{n\to\infty} \frac{n-1}{n}\sigma^2 = \sigma^2$$

故 B_2 是 σ^2 的渐近无偏估计. □

例

设总体 X 服从均匀分布 $U(0,\theta),\theta$ 是未知参数，样本 $X_1,\cdots,X_n$.

(1) 求 θ 的矩估计，判断是否无偏；

(2) 求 θ 的极大似然估计，判断是否无偏.

解 (1)：矩估计:
由

$$\mu_1 = E(X) = \int_0^\theta \frac{1}{\theta} x dx = \frac{\theta}{2}.$$

$\Rightarrow \theta = 2\mu_1$
$\Rightarrow \theta$ 的矩估计 $\hat{\theta} = 2\overline{X}$.
因为

$$E(\hat{\theta}) = E\left(2\overline{X}\right) = 2E(\overline{X}) = 2E(X) = \theta,$$

所以 $\hat{\theta} = 2\overline{X}$ 是 θ 的无偏估计.

(2) X 的概率密度

$$f_X(x;\theta)=\begin{cases}\frac{1}{\theta}, & 0\le x\le\theta\\ 0, & \text{其他}\end{cases}$$

极大似然估计:

$$L(\theta)=\begin{cases}\dfrac{1}{\theta^n}, & 0\le x_1,\cdots,x_n\le\theta,\\ 0, & \text{其它}.\end{cases}$$

$L(\theta)$ 关于 $\theta>0$ 递减,
而 θ 的范围为 $\theta\ge x_{(n)}=\max\{x_1,...,x_n\}$,
所以, θ 的极大似然估计量

$$\hat{\theta}=X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}.$$

$X_{(n)} = \max\{X_1, X_2, \cdots, X_n\}$ 的分布函数为

$$F_{X_{(n)}}(x) = [F(x)]^n = \begin{cases} 0, & x < 0, \\ \dfrac{x^n}{\theta^n}, & 0 \le x \le \theta, \\ 1, & x > \theta. \end{cases}$$

求导数得密度函数为

$$f_{X_{(n)}}(x) = \begin{cases} \frac{nx^{n-1}}{\theta^n}, & 0 \le x \le \theta, \\ 0, & \text{其它}. \end{cases}$$

因此有

$$
\begin{aligned}
E\left(\hat{\theta}\right) = E\left(X_{(n)}\right) &= \int_0^{\theta} \frac{x \cdot n x^{n-1}}{\theta^n} dx \\
&= \frac{n}{n+1}\theta \neq \theta
\end{aligned}
$$

所以$\hat{\theta} = X_{(n)}$作为参数θ的估计是有偏的. $\square$

纠偏方法

- 如果 $E\left(\hat{\theta}\right) = a\theta + b, \theta \in \Theta$, 其中 a,b 是常数，且 $a \neq 0$, 则 $\frac{1}{a}(\hat{\theta} - b)$是 θ 的无偏估计.
- 在上例中,

$$E(X_{(n)}) = \frac{n}{n+1}\theta,$$

取

$$X_{(n)}^{*} = \frac{n+1}{n}X_{(n)},$$

则 $X_{(n)}^{*}$ 是 θ 的无偏估计.

定义 (有效性准则)

设 $\hat{\theta}_1, \hat{\theta}_2$ 是 θ 的两个无偏估计，如果

$$D(\hat{\theta}_1) \leq D(\hat{\theta}_2), \quad \forall \theta \in \Theta$$

不等号至少对某一个 $\theta \in \Theta$ 成立，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

定义 (有效性准则)

设 $\hat{\theta}_1, \hat{\theta}_2$ 是 θ 的两个无偏估计，如果

$$D(\hat{\theta}_1) \leq D(\hat{\theta}_2), \quad \forall \theta \in \Theta$$

不等号至少对某一个 $\theta \in \Theta$ 成立，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

- 方差较小的无偏估计量是一个更有效的估计量.

例

设总体为 $X, E(X) = \mu, D(X) = \sigma^2 > 0$, $X_1, \cdots, X_n$ 是一样本. 对 $1 \le k \le n$, 令

$$\hat{\theta}_k = \frac{1}{k}(X_1 + \ldots + X_k)$$

即 $\hat{\theta}_k$ 为前 k 个样本平均值. 显然, $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$ 均是参数 μ 的无偏估计.
问: 在估计量 $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$ 中, 哪个 $\hat{\theta}_k$ 作为参数 μ 的估计最有效?

解：

$$D(\hat{\theta}_k) = \frac{1}{k^2}\sum_{i=1}^{k} D(X_i) = \frac{\sigma^2}{k},$$

即估计量的方差随着 k 的增加而减少，

$\therefore \hat{\theta}_n$ 最有效. $\square$

定义

设 $\hat{\theta}(X_1, \cdots, X_n)$ 为参数 θ 的估计量，
若对于任意 $\theta \in \Theta$，当 $n \to +\infty$ 时，

$$\hat{\theta}_n \xrightarrow{\quad P \quad} \theta$$

则称 $\hat{\theta}_n$ 为 θ 的相合估计量.

注

- 相合性是对一个估计量的基本要求. 不具备相合性的估计量不予考虑.
- 相合性只有在样本容量很大时, 才显现其优越性, 实际应用中很难做到. 在实际工程中往往使用无偏性和有效性进行评价.
- 无偏性、有效性、相合性是评价估计量的一些基本标准, 还有其他侧重点的评价标准.

例

设总体 X 的 k 阶矩 $E(X^k)=\mu_k(k\geq 2)$ 存在，$X_1,\cdots,X_n$ 是一样本，证明：

(1) $A_l=\frac{1}{n}\sum_{i=1}^{n}X_i^l$ 是 μ_l 的相合估计；

(2) B_2, S^2 是 $D(X)=\sigma^2$ 的相合估计；

(3) S 是 σ 的相合估计.

证明：(1) 由辛钦大数定律知，对 $l=1,...,k$,

$$A_l=\frac{1}{n}\sum_{i=1}^{n}X_i^l\xrightarrow{\quad P\quad}\mu_l=E(X^l),$$

因此 A_l 是 $E(X^l)$ 的相合估计. 特别地，$\overline{X}$ 是 $\mu_1=E(X)$ 的相合估计，A_2 是 $\mu_2=E(X^2)$ 相合估计.
(2) 因为 $D(X)=\sigma^2=\mu_2-\mu_1^2$,

$$B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=A_2-\overline{X}^2,$$

根据依概率收敛性质，$B_2=A_2-\overline{X}^2$ 是 σ^2 的相合估计.
而 $S^2=\frac{n}{n-1}B_2$ 也是 σ^2 的相合估计.
(3) $S=\sqrt{S^2}$ 是 σ 的相合估计. □

区间估计

- 根据具体样本观测值, 点估计提供一个明确的数值.
- 但这种判断的把握有多大, 点估计本身并没有告诉人们. 为弥补这种不足, 提出区间估计的概念.

区间估计

设 X 是总体，$X_1, ..., X_n$ 是一样本. 区间估计的目的是找到两个统计量：

$$\hat{\theta}_1 = \hat{\theta}_1(X_1, \ldots, X_n), \quad \hat{\theta}_2 = \hat{\theta}_2(X_1, \ldots, X_n),$$

使随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 以一定可靠程度盖住 θ.

定义

设总体 X 的分布函数 $F(x;\theta)$, θ 未知. 对给定值 $\alpha(0<\alpha<1)$, 有两个统计量

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

使得

$$P\Big\{\underline{\theta}(X_1,\cdots,X_n)<\theta<\overline{\theta}(X_1,\cdots,X_n)\Big\}\geq 1-\alpha$$

则

- $(\underline{\theta},\overline{\theta})$ 称为 θ 的置信水平为 $1-\alpha$ 的(双侧) 置信区间;
- $\underline{\theta}$ 和 $\overline{\theta}$ 分别称为置信下限和置信上限.

注

- 参数 θ 虽然未知，但是确定的值.
- $\underline{\theta}, \overline{\theta}$ 是统计量，随机的，依赖于样本。
- 置信区间 $(\underline{\theta}, \overline{\theta})$ 不唯一，依赖于样本.
- 对于有些样本观察值，区间覆盖 θ, 但对于另一些 样本观察值, 区间则不能覆盖θ.

置信区间的含义

设总体 $X \sim N(\mu, 4)$, μ 未知， $X_1, ..., X_4$ 是一样本. 则 $\overline{X} \sim N(\mu, 1)$.

$$\begin{aligned} P(\overline{X} - 2 < \mu < \overline{X} + 2) &= P(|\overline{X} - \mu| < 2) \\ &= 2\Phi(2) - 1 = 0.9544 \end{aligned}$$

$\Rightarrow (\overline{X} - 2, \overline{X} + 2)$ 是 μ 的置信水平为 0.95 的置信区间.

置信区间的含义

若 $\mu = 0.5$, 当 $\bar{x}$ 分别为 3, 2, 1 时，对应置信区间为：

$$(-1, 3) \quad (1, 5) \quad (0, 4)$$

对于一个具体的区间而言，或者包含真值，或者不包含真值，无概率可言.

$(\overline{X}-2, \overline{X}+2)$ 是 μ 的置信水平为 0.95 的置信区间中“置信水平为 0.95”的意义是什么？

置信区间的含义

一般地,

$$P\{\underline{\theta}(X_1,\ldots,X_n) < \theta < \overline{\theta}(X_1,\ldots,X_n)\} = 1-\alpha,$$

则置信区间 $(\underline{\theta},\overline{\theta})$ 的含义为：

- 反复抽样多次 (各次样本容量都为 n). 每个样本值确定一个区间 $(\underline{\theta},\overline{\theta})$, 每个这样的区间或包含 θ 的真值，或不包含 θ 的真值. 按伯努利大数定律，在这些区间中，包含 θ 真值的比例约为 $1-\alpha$.

置信区间的含义

如反复抽样 10000 次,

- 当 $\alpha = 0.05$, 即置信水平为 95% 时，10000 个区间中包含 θ 真值的约为 9500 个;
- 当 $\alpha = 0.01$, 即置信水平为 99% 时，10000 个区间中包含θ 的真值的约为 9900 个.

求置信区间步骤

设 θ 是总体的未知参数，$X_1, \cdots, X_n$ 为样本，给定置信水平 $1-\alpha$，

1. 构造枢轴量(不依赖 θ 及未知参数的函数)

$$W = W(X_1, \cdots, X_n; \theta).$$

2. 确定常数 a, b 使得

$$P\{a < W(X_1, \cdots, X_n; \theta) < b\} = 1 - \alpha$$

3. 解得 θ 的取值范围即为置信区间.

例

设总体 $X \sim N(\mu, \sigma^2)$, σ^2 已知，μ 未知，$X_1, \cdots, X_n$ 为样本，求 μ 的置信水平为 $1-\alpha$ 的置信区间.

枢轴量（补充）

枢轴量和统计量的区别：

(1) 枢轴量是样本和待估参数的函数，其分布不依赖于任何未知参数；

(2) 统计量只是样本的函数，其分布常依赖于未知参数.

- 枢轴量通常可由未知参数的点估计得到. 比如正态总体的区间估计.

单个正态总体 $N(\mu, \sigma^2)$ 的枢轴量

- μ 的枢轴量：

$$\begin{cases} \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1), & (\sigma^2 \text{ 已知}) \\ \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1), & (\sigma^2 \text{ 未知}) \end{cases}$$

- σ^2 的枢轴量：$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, μ 未知.

两个正态总体 $N(\mu_1,\sigma_1^2), N(\mu_2,\sigma_2^2)$ 的枢轴量

- $\mu_1-\mu_2$ 的枢轴量：

$$\begin{cases} \dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1), & (\sigma_1^2,\sigma_2^2 \text{ 已知}) \\ \dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2), & (\sigma_1^2=\sigma_2^2 \text{未知}) \end{cases}$$

其中 $S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$, $S_w=\sqrt{S_w^2}$.

- $\frac{\sigma_1^2}{\sigma_2^2}$ 的枢轴量: $\frac{S_1^2}{S_2^2}/\frac{\sigma_1^2}{\sigma_2^2} \sim F(n_1-1,n_2-1)$.
(μ_1,μ_2 未知)

单侧置信区间

定义

若

$$P\{\theta > \underline{\theta}(X_1, ..., X_n)\} \geq 1 - \alpha,$$

则 $(\underline{\theta}, \infty)$ 称为参数 θ 的置信水平为 $1 - \alpha$ 的单侧置信区间, $\underline{\theta}$ 称为单侧置信下限.

若

$$P\left\{\theta < \overline{\theta}(X_1, \ldots, X_n)\right\} \geq 1 - \alpha,$$

则 $(-\infty, \overline{\theta})$ 称为参数 θ 的置信水平为 $1 - \alpha$ 的单侧置信区间, $\overline{\theta}$ 称为单侧置信上限.

单侧置信区间和双侧置信区间的关系

$\underline{\theta}$ 是 θ 的置信水平为 $1-\alpha_1$ 的单侧置信下限，
$\overline{\theta}$ 是 θ 的置信水平为 $1-\alpha_2$ 的单侧置信上限，
$\Longrightarrow (\underline{\theta}, \overline{\theta})$是 θ 的置信度为 $1-\alpha_1-\alpha_2$ 的双侧置信区间.

证明: $P\{\underline{\theta} \geq \theta\} \leq \alpha_1, \quad P\{\theta \geq \overline{\theta}\} \leq \alpha_2$

$$P\{\underline{\theta} < \theta < \overline{\theta}\} = 1 - P\{\underline{\theta} \geq \theta\} - P\{\overline{\theta} \leq \theta\}$$
$$\geq 1-\alpha_1-\alpha_2. \qquad \square$$