第七章: 参数估计

引言

数理统计的任务是用样本去推断总体. 参数估计是统计推断的一种重要形式, 包括**点估计**和**区间估计**.

定义 0.1

设有参数分布族 $\{f(x,\theta),\theta\in\Theta\}$, 其中 Θ 是参数空间, $f(x,\theta)$ 的分布式已知, 但其分布与未知参数 θ 有关. $X_1,\cdots,X_n$ 是从总体中抽出的简单随机样本, 利用样本对未知参数 θ 或其函数 $g(\theta)$ 进行估计.

例 0.1

例如, $X_1,\cdots,X_n$ $i.i.d.\sim\mathcal{N}(\mu,\sigma^2)$, 记 $\theta=(\mu,\sigma^2)$. 通过样本对 μ 和 σ^2 或其函数 $g(\theta)=\mu/\sigma^2$ 的值作出估计, 这就是参数估计问题.

定义 0.2

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 为从某个总体中抽取的样本,

$$\hat{g}(\boldsymbol{X}) = \hat{g}(X_1, \cdots, X_n)$$

是样本的函数, 用 $\hat{g}(\boldsymbol{X})$ 作为 $g(\theta)$ 的估计, 称为点估计 *(point estimation)*.

注 0.1

有如下问题:

(1) 如何找到点估计量? 矩估计、极大似然估计

(2) 如何评价不同的点估计量? 点估计的评价准则

例 0.2

设 $X_1, \cdots, X_n$ 是取自某总体 F 一组简单样本. 对此总体的均值 $\theta = E_F(X)$ 可有哪些估计量?

解: 估计量可有如下选择:

$$\hat{\theta}_1 = \overline{X} = \frac{1}{n}(X_1 + \cdots + X_n),$$

$$\hat{\theta}_2 = \frac{1}{2}\left(X_{(1)} + X_{(n)}\right),$$

$$\hat{\theta}_3 = m_{1/2},$$

$$\hat{\theta}_4 = X_1.$$

其中 $X_{(1)}$ 和 $X_{(n)}$ 为样本最小和最大次序统计量, $m_{1/2}$ 为样本中位数. □

7.1 矩估计

矩估计: 定义

定义 1.1

设 $X_1, \cdots, X_n$ 是从总体 F 中抽取的简单随机样本. 这时, 样本矩可用来估计 F 的相应的总体矩. 即样本 k 阶原点矩

$$A_m = \frac{1}{n}\sum_{i=1}^{n} X_i^m, \quad k = 1, 2, \cdots$$

是总体 k 阶原点矩 $\alpha_m = E(X^m)$ 的自然矩估计量. 样本 k 阶中心矩

$$M_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k, \quad k = 2, 3, \cdots$$

是总体 k 阶中心矩 $\mu_k = E\big(X - E(X)\big)^k$ 的自然矩估计量.

矩估计: 定义

定义 1.2

一般 α_m 是总体参数 $\theta_1, \cdots, \theta_k$ 的函数, 记为 $\alpha_m(\theta_1, \cdots, \theta_k)$. 因此, 令总体的前 k 阶原点矩与同阶样本原点矩相等, 就得到关于 $\theta_1, \cdots, \theta_k$ 的一个方程组

$$\alpha_m(\theta_1, \cdots, \theta_k) = A_m, \quad m = 1, 2, \cdots, k.$$

解这个方程组, 其解记为

$$\hat{\theta}_i = \hat{\theta}_i(X_1, X_2, \cdots, X_n), \quad i = 1, \cdots, k.$$

称 $\hat{\theta}_i$ 为 $\theta_1, \cdots, \theta_k$ 的矩估计.

矩估计: 理论背景

注 1.1

因为样本 $X_1, X_2, \cdots, X_n$ 是独立同分布的, 于是, $X_1^m, \cdots, X_n^m$ 也是独立同分布的, 因而

$$E(X_1^m) = E(X_2^m) = \cdots = E(X_n^m) = \alpha_m.$$

按照大数定律, 样本原点矩 A_m 作为 $X_1^m, X_2^m, \cdots, X_n^m$ 的算术平均数依概率收敛到均值 $\alpha_m = E(X_i^m)$, 即

$$A_m = \frac{1}{n}\sum_{i=1}^{n} X_i^m \xrightarrow{p} \alpha_m.$$

于是, 对充分大的 n, 有 $\alpha_m(\theta_1, \cdots, \theta_k) \approx A_m$, 将“约等号”换成“等号”就得到矩估计.

矩估计: 步骤

矩估计的求法:

- 求分布的总体矩, 如一阶原点矩 (期望) 或二阶原点矩 (方差).
- 求对应阶数的样本矩.
- 令样本矩等于总体矩阵, 求解方程得到矩估计.

矩估计的原则:

- n 个参数寻找 n 个对应关系.
- **尽量使用低阶矩!**

矩估计: 均匀分布例子

例 1.1

设 $X_1, \cdots, X_n$ $i.i.d. \sim$ 均匀分布 $U[\theta_1, \theta_2]$, 参数 $\theta = (\theta_1, \theta_2)$, 其中 $-\infty < \theta_1 < \theta_2 < \infty$. 求 θ_1 和 θ_2 的矩估计量.

解: 设 $X \sim U[\theta_1, \theta_2]$, 由均匀分布的性质可知

$$E(X) = \alpha_1 = (\theta_1 + \theta_2)/2, Var(X) = \mu_2 = (\theta_2 - \theta_1)^2/12.$$

对应的样本矩分别为一阶样本原点矩 $\overline{X}$ 和二阶样本中心矩 S_n^2. 令样本矩等于总体矩阵可得,

$$\hat{\theta}_1 = \alpha_1 - \sqrt{3\mu_2} = \overline{X} - \sqrt{3}S_n,$$
$$\hat{\theta}_2 = \alpha_1 + \sqrt{3\mu_2} = \overline{X} + \sqrt{3}S_n.$$

注意, 此处 $S_n^2 = \sum_{i=1}^n (X_i - \overline{X})^2/n$, 不是样本方差 S^2. □

矩估计: 正态分布例子

例 1.2

设 $X_1, \cdots, X_n$ *i.i.d.* $\sim N(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty$, $\sigma > 0$. 求 μ 和 σ^2 的矩估计量.

解: 先求总体矩 $\alpha_1 = \mu, \mu_2 = \sigma^2$. 再求对应的样本矩

$$A_1 = \bar{X}, \quad M_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

令 $\alpha_1 = A_1$ 和 $\mu_2 = M_2$ 可得

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

注意, 此处 $S_n^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/n$, 也不是样本方差 S^2. $\square$

矩估计: 伽马分布例子

例 1.3

设 $X_1, \cdots, X_n$ *i.i.d.* $\sim \Gamma(\alpha, \lambda)$, 其中 $\alpha, \lambda > 0$. 求 α 和 λ 的矩估计量.

解: 先求总体矩 $\alpha_1 = \alpha/\lambda$, $\mu_2 = \alpha/\lambda^2$. 再求对应的样本矩

$$A_1 = \bar{X}, \quad M_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = S_n^2.$$

令 $\alpha_1 = A_1$ 和 $\mu_2 = M_2$ 可得

$$\hat{\alpha} = \frac{\bar{X}^2}{S_n^2}, \quad \hat{\lambda} = \frac{\bar{X}}{S_n^2}.$$

□

7.2 极大似然估计

极大似然估计: 似然函数

定义 2.1

设 $f(\boldsymbol{x},\theta)=f(x_1,\cdots,x_n,\theta)$ 为样本 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 的概率函数. 当 $\boldsymbol{x}$ 固定时, 把 $f(\boldsymbol{x},\theta)$ 看成 θ 的函数, 称为似然函数 *(likelihood function)*, 记为

$$L(\theta,\boldsymbol{x})=f(\boldsymbol{x},\theta),\quad \theta\in\Theta,\boldsymbol{x}\in\mathscr{X},$$

其中 Θ 为参数空间, $\mathscr{X}$ 为样本空间. 称 $\log L(\theta,\boldsymbol{x})$ 为对数似然函数, 记为 $l(\theta,\boldsymbol{x})$.

极大似然估计: 似然函数

注 2.1

似然函数和概率函数是同一表达式, 但表示两种不同含意.

- 当把 θ 固定, 将其看成定义在样本空间 $\mathscr{X}$ 上的函数时, 称为概率函数.
- 当把 $\boldsymbol{x}$ 固定, 将其看成定义在参数空间 Θ 上的函数时, 称为似然函数.

例 2.1

设罐子里有许多黑球和红球. 假定已知它们的比例是 *1:3*, 但不知道是黑球多还是红球多. 也就是说抽出一个黑球的概率是 *1/4* 或者 *3/4*. 如果有放回地从罐子中抽 n 个球, 根据抽样数据, 说明抽到黑球的概率是 *1/4* 还是 *3/4*.

极大似然估计: 似然函数

解: 令 X_i 表示第 i 次抽球的结果, 即

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 次抽出为黑球}, \\ 0, & \text{其他}. \end{cases}$$

记每次抽样中抽到黑球的概率为 θ, 此处 θ 只取可能的两个值 $\theta_1 = 1/4$ 和 $\theta_2 = 3/4$ 之一.

- 记 $X = \sum_{i=1}^{n} X_i$, 则 $X \sim b(n, \theta)$.
- 样本分布族为 $\mathscr{F} = \{f(x, \theta_1), f(x, \theta_2)\}$, 其中 $f(x, \theta_1)$ 为 $b(n, \theta_1)$, $f(x, \theta_2)$ 为 $b(n, \theta_2)$.

根据抽样结果, 判断样本来自总体 $f(x, \theta_1)$ 还是 $f(x, \theta_2)$?

极大似然估计: 似然函数

为简单记, 取 $n = 3$. 当 $x = 0, 1, 2, 3$ 时似然函数取值

x	0	1	2	3
$L(\theta_1, x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$
$L(\theta_2, x)$	$\frac{1}{64}$	$\frac{9}{64}$	$\frac{27}{64}$	$\frac{27}{64}$

- 当 $x = 0, 1$ 时, $L(\theta_1, x) > L(\theta_2, x) \Rightarrow$ 样本来自 $f(x, \theta_1)$, $\hat{\theta}_1 = 1/4$.
- 当 $x = 2, 3$ 时, $L(\theta_2, x) > L(\theta_1, x) \Rightarrow$ 样本来自 $f(x, \theta_2)$, $\hat{\theta}_2 = 3/4$.

更具体地, 从罐子中抽出 3 个球, 3 个球中没有黑球, 直观上黑球和红球的比例“看起来更像”是 1 : 3, 即抽出一个黑球的概率为 1/4. $\square$

极大似然估计: 似然函数

注 2.2

上例表明,

- 若
$$L(\theta_1, x) > L(\theta_2, x),$$
则倾向于认为样本 $\boldsymbol{X}$ 来自总体 $f(x, \theta_1)$, 即真实参数 θ 为 θ_1 的理由比认为样本 $\boldsymbol{X}$ 来自总体 $f(x, \theta_2)$, 即真实参数 θ 为 θ_1 的理由更充分.
- 或者说, 真实参数 θ 为 θ_1 的"似然性"更大些. 这样自然把"似然性"最大 (看起来最像) 的那个值作为真实参数 θ 的估计值, 即极大似然估计值.

定义 2.2

设 $\boldsymbol{X} = \{X_1, \cdots, X_n\}$ 是从参数分布族 $\{f(x, \theta), \theta \in \Theta\}$ 中抽取的简单随机样本, $L(\theta, \boldsymbol{x})$ 是似然函数, 若存在统计量 $\hat{\theta} = \hat{\theta}(\boldsymbol{X})$, 满足条件

$$L(\hat{\theta}, \boldsymbol{x}) = \sup_{\theta \in \Theta} L(\theta, \boldsymbol{x}), \quad \boldsymbol{x} \in \mathscr{X},$$

或等价地使得对数似然函数

$$l(\hat{\theta}^*, \boldsymbol{x}) = \sup_{\theta \in \Theta} l(\theta, \boldsymbol{x}), \quad \boldsymbol{x} \in \mathscr{X},$$

则称 $\hat{\theta}^*$ 为 θ 的极大似然估计 (*maximum likelihood estimation, MLE*). 若待估函数是 $g(\theta)$, 则定义 $g\big(\hat{\theta}^*(\boldsymbol{X})\big)$ 为 $g(\theta)$ 的 *MLE*.

极大似然估计: 极大似然估计求法 I

定理 2.1 (用微积分中求极值的方法)

设 $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$ 为参数向量, 若 $l(\theta, \boldsymbol{x})$ 的极大值在参数空间 Θ 的内点处 (而非边界点) 达到, 则此点必为似然方程组

$$\frac{\partial l(\boldsymbol{\theta}, \boldsymbol{x})}{\partial \theta_i} = 0, \quad i = 1, 2, \cdots, k$$

的解. 方程的解为参数极大似然估计 $\hat{\boldsymbol{\theta}}$.

注 2.3

若似然函数 $L(\theta, \boldsymbol{x})$ 是 θ 的连续可微函数, 则可用微积分中求极值的方法去求 θ 的 *MLE*, 即找使 $L(\theta, \boldsymbol{x})$ 达到最大时 θ 的值. 由于 $L(\theta, \boldsymbol{x})$ 与 $\log L(\theta, \boldsymbol{x}) = l(\theta, \boldsymbol{x})$ 拥有相同的极值点, 可用 $l(\theta, \boldsymbol{x})$ 代替 $L(\theta, \boldsymbol{x})$.

极大似然估计: 极大似然估计求法 I

但方程的解是否一定是 θ 的 MLE 呢?

- **$\hat{\theta}$ 满足似然方程只是 MLE 的必要条件, 而非充分条件.**
- 一般只有满足:
  (1) 似然函数的极大值在参数空间 Θ 内部达到.
  (2) 似然方程只有唯一解.

  则似然方程之解 $\hat{\theta}$ 必为 θ 的 MLE.

因此求出似然方程的解后, 要验证它为 θ 的 MLE, 有时并非易事.

极大似然估计: 二项分布例子

例 2.2

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 是从两点分布族 $\{b(1,p) : 0 < p < 1\}$ 中抽取的简单随机样本, 求 p 的 MLE.

解: 似然函数为

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

故有

$$l(p, \boldsymbol{x}) = \log p\left(\sum_{i=1}^n x_i\right) + \log(1-p)\left(n - \sum_{i=1}^n x_i\right).$$

极大似然估计: 二项分布例子

对数似然方程为

$$\frac{\partial l(p,\boldsymbol{x})}{\partial p} = \frac{1}{p}\sum_{i=1}^{n} x_i - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0,$$

解得

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}.$$

注意到 $\sum_{i=1}^{n} x_i \le n$, 很容易验证 $\ln L(p)$ 的二阶导数在 $\bar{x}$ 处取负值, 于是 $\bar{x}$ 是 $\ln L(p)$ 的最大值点. 因此, $\bar{X}$ 是 p 的极大似然估计. □

极大似然估计: 正态分布例子

例 2.3

设 $X_1, \cdots, X_n \sim \mathcal{N}(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty, \sigma^2 > 0$. 求 μ 和 σ^2 的 *MLE*.

解: 样本 $X_1, \cdots, X_n$ 的分布为

$$f(\boldsymbol{x}, \theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}.$$

对数似然函数为

$$l(\theta, \boldsymbol{x}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

极大似然估计: 正态分布例子

由对数似然方程组

$$\frac{\partial l(\theta, \boldsymbol{x})}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0,$$

$$\frac{\partial l(\theta, \boldsymbol{x})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0,$$

解得

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

易验证似然方程组有唯一解 $(\hat{\mu}, \hat{\sigma}^2)$, 且一定是最大值. 因此, $\hat{\mu}, \hat{\sigma}^2$ 为 μ 和 σ^2 的 MLE. $\square$

极大似然估计: 极大似然估计求法 II

例 2.4

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 是从均匀分布族 $\{U(0,\theta) : \theta > 0\}$ 中抽取的简单随机样本. 求 θ 的 *MLE*.

解: 样本 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 的联合密度为

$$f(\boldsymbol{x}, \theta) = \begin{cases} \frac{1}{\theta^n}, & 0 < x_1, \cdots, x_n < \theta, \\ 0, & \text{其他}. \end{cases}$$

因为均匀分布 $U(0,\theta)$ 的支撑集依赖 θ, 似然函数 $L(\theta, \boldsymbol{x})$ 不是 θ 的连续函数. **不能用对数似然函数求微商的方法求解 θ 的 MLE.**

极大似然估计: 极大似然估计求法 II

只能从定义的角度出发, 为使 $L(\theta, \boldsymbol{x})$ 达到极大, 应使分母上 θ 尽可能地小, 但 θ 不能太小以致 L 为 0.

$$L(\theta, \boldsymbol{x}) = \frac{1}{\theta^n} \prod_{i=1}^{n} I(x_i \leq \theta) = \frac{1}{\theta^n} I(\theta \geq x_{(n)}),$$

其中 $I(\cdot)$ 为示性函数, $x_{(n)}$ 表示最大次序统计量. 要使得似然函数达到最大, 需满足 (1) θ 尽可能地小, (2) $I(\theta \geq x_{(n)}) = 1$, 即

$$\hat{\theta} = X_{(n)}.$$

$\square$

例 2.5

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 是从均匀分布族 $\{U(a,b) : a < b\}$ 中抽取的简单随机样本. 求 a 和 b 的 *MLE*.

解: 给定样本 $\boldsymbol{x}$ 时, θ 的似然函数为

$$L(a,b,\boldsymbol{x}) = \begin{cases} \frac{1}{(b-a)^n}, & a \le x_i \le b, i = 1,2,\cdots,n \\ 0, & \text{其他} \end{cases}$$

在支撑集 $a \le x_i \le b, i = 1,2,\cdots,n$ 上

$$L(a,b,\boldsymbol{x}) = \frac{1}{(b-a)^n} \cdot I(a \le x_{(1)} \le x_{(n)} \le b).$$

为使 $L(a,b,\boldsymbol{x})$ 达到最大, 需使得 $b-a$ 尽量的小, 即

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

7.3 点估计量的优良准则

点估计量的优良准则: 均方误差准则

定义 3.1

设 $\hat{g}(\boldsymbol{X})$ 为 $g(\theta)$ 的估计量, 则称 $E_\theta\left[\left(\hat{g}(\boldsymbol{X})-g(\theta)\right)^2\right]$ 为均方误差 (*Mean Square error, MSE*). 特别地, 当 $g(\theta)=\theta$ 时, 设其估计量为 $\hat{\theta}$, 则均方误差为 $E_\theta[(\hat{\theta}-\theta)^2]$.

性质 3.1

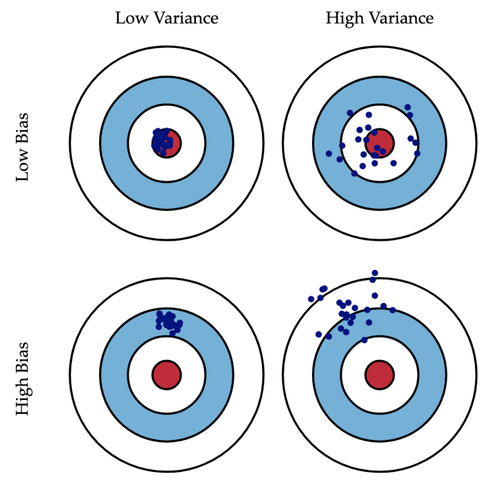
均方误差可进行如下分解:

$$\begin{aligned} MSE = E_\theta[(\hat{\theta}-\theta)^2] &= E_\theta\left[(\hat{\theta}-E[\hat{\theta}])+(E[\hat{\theta}]-\theta)\right]^2 \\ &= E_\theta(\hat{\theta}-E[\hat{\theta}])^2+(E[\hat{\theta}]-\theta)^2+2E\left[(\hat{\theta}-E[\hat{\theta}])\cdot(E[\hat{\theta}]-\theta)\right] \\ &= Var(\hat{\theta})+\left(E[\hat{\theta}]-\theta\right)^2. \end{aligned}$$

其中 $E(\hat{\theta})-\theta$ 是估计量 $\hat{\theta}$ 的偏差 (*Bias*).

点估计量的优良准则: 均方误差准则

方差刻画了估计量的波动程度, 偏差刻画了估计量偏离真实值 θ 的程度.



例 3.1

令 $X_1, \cdots, X_n$ *i.i.d.* $\sim N(\mu, \sigma^2)$. 利用样本均值 $\overline{X}$ 和样本方差 S^2 来估计 μ 和 σ^2. 试求两个估计量的 *MSE*.

解: 因为

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

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

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

其中

$$E(X_i^2) = Var(X_i) + [E(X_i)]^2 = \sigma^2 + \mu^2,$$
$$E(\bar{X}^2) = Var(\bar{X}) + [E(\bar{X})]^2 = \frac{\sigma^2}{n} + \mu^2.$$

将其代入上式可知,

$$\begin{aligned} E(S^2) &= \frac{n}{n-1}(\sigma^2 + \mu^2) - \frac{n}{n-1}\left(\frac{\sigma^2}{n} + \mu^2\right) \\ &= \sigma^2. \end{aligned}$$

所以 $\overline{X}$ 和 S^2 是无偏的.

点估计量的优良准则: 均方误差准则

求 MSE 需求参数 μ 和 σ^2 的方差, 显然 $Var(\overline{X}) = \sigma^2/n$. 对样本方差 S^2, 因为 $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, 所以

$$Var\left(\frac{(n-1)S^2}{\sigma^2}\right) = \frac{(n-1)^2}{\sigma^4} Var(S^2) = 2(n-1),$$

即 $Var(S^2) = 2\sigma^4/(n-1)$. 因此

$$MSE_{\overline{X}} = Var(\overline{X}) = \frac{\sigma^2}{n},$$
$$MSE_{S^2} = Var(S^2) = \frac{2\sigma^4}{n-1}.$$

$\square$

点估计量的优良准则: 均方误差准则

例 3.2

令 $X_1, \cdots, X_n$ $i.i.d. \sim \mathcal{N}(\mu, \sigma^2)$. 利用

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

来估计 σ^2, 试求估计量 S_n^2 的 MSE.

解: 先求该估计的偏差,

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

其偏差为

$$|E(\hat{\sigma}^2) - \sigma^2| = \frac{1}{n} > 0.$$

再求估计的方差

$$Var(\hat{\sigma}^2) = Var\left(\frac{n-1}{n}S^2\right) = \frac{(n-1)^2}{n^2}\frac{2\sigma^4}{n-1} = \frac{2(n-1)\sigma^4}{n^2}.$$

该估计量的 MSE 为

$$MSE_{\hat{\sigma}^2} = \left(E(\hat{\sigma}^2) - \sigma^2\right)^2 + Var(\hat{\sigma}^2) = \frac{(2n-1)\sigma^4}{n^2}.$$

由于

$$\frac{(2n-1)\sigma^4}{n^2}\bigg/\frac{2\sigma^4}{n-1} = \frac{(2n-1)(n-1)}{2n^2} = \frac{2n^2-3n+1}{2n^2} < 1.$$

因此,

$$MSE_{S_n^2} = \frac{(2n-1)\sigma^4}{n^2} < \frac{2\sigma^4}{n-1} = MSE_{S^2}.$$

可见, 虽然估计量 S_n^2 有偏差, 但 MSE 较小. $\square$

点估计量的优良准则: 均方误差准则

例 3.3

设 $X \sim b(100, \theta)$, $0 < \theta < 1$, 分别利用

$$\delta_1 = \frac{X}{100}, \quad \delta_2 = \frac{X+3}{100}, \quad \delta_3 = \frac{X+3}{106}$$

来估计 θ, 比较三个估计量的 *MSE*.

解: 对 $0 < \theta < 1$, 分别求三个估计量的 MSE 为

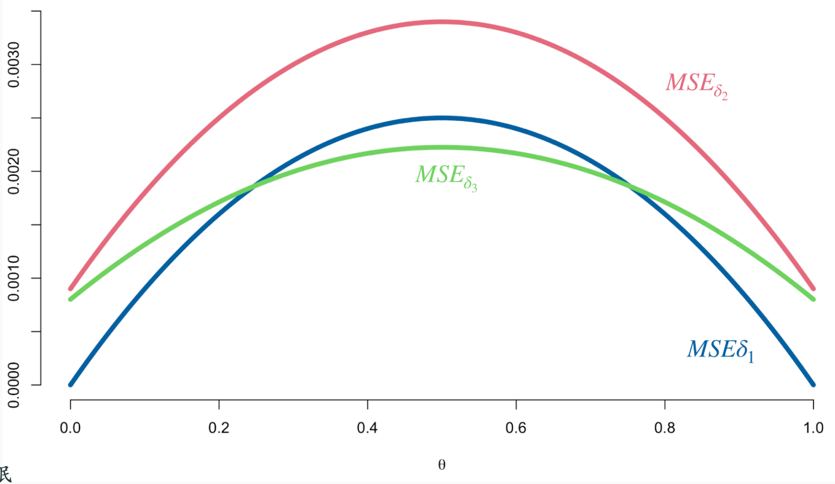
$$MSE_{\delta_1} = \left(E\left[\frac{X}{100}\right] - \theta\right)^2 + Var\left(\frac{X}{100}\right) = \frac{\theta(1-\theta)}{100};$$

$$MSE_{\delta_2} = \left(E\left[\frac{X+3}{100}\right] - \theta\right)^2 + Var\left(\frac{X+3}{100}\right) = \frac{9+\theta(1-\theta)}{100^2};$$

$$MSE_{\delta_3} = \left(E\left[\frac{X+3}{106}\right] - \theta\right)^2 + Var\left(\frac{X+3}{106}\right) = \frac{(9-8\theta)(1+8\theta)}{106^2}.$$

点估计量的优良准则: 均方误差准则

对 $0 < \theta < 1$, 作出三个估计量 MSE 的函数图,



点估计量的优良准则: 均方误差准则

在上例的讨论中, 只要知道 θ 的值, 可以选择最优的估计量.

- 但实际是 θ **是未知的待估参数, 所以才会产生参数估计问题**.
- 该如何选择在 MSE 准则下最优的估计?
- 此时, 常常制定新的, 更合理的评价准则, 求出新准则下最优的估计. 一条常用的准则是:
- **在偏差为 0 (无偏, Unbiased) 的一簇估计中选取方差较小的 (有效), 方差最小的估计量称为最小方差无偏估计.**

点估计量的优良准则: 无偏性

定义 3.2 (无偏性)

设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 为从总体 $\{f(x,\theta):\theta\in\Theta\}$ 中抽取的样本 $g(\theta)$ 是定义在 Θ 上的已知函数. $\hat{g}(\boldsymbol{X})=\hat{g}(X_1,\cdots,X_n)$ 是 $g(\theta)$ 的一个估计量, 如果

$$E_\theta[\hat{g}(\boldsymbol{X})]=g(\theta),\quad \theta\in\Theta,$$

则称 $\hat{g}(\boldsymbol{X})$ 为 $g(\theta)$ 的一个无偏估计 (*unbiased estimation*).

注 3.1

无偏性有两层含义: (1) 无系统偏差, 即由于样本随机性产生的正负偏差平均起来值为 0, 不存在总是高估或者低估某一参数的情况; (2) 要求估计量大量重复使用, 在多次重复使用下给出接近真实值 $g(\theta)$ 的估计.

例 3.4

设 $X_1, \cdots, X_n$ 是来自均值为 μ, 方差为 σ^2 总体的一个样本 (没有正态假设). 验证 X_1 和 $\overline{X}$ 是 μ 的无偏估计; 验证 S^2 是 σ^2 的无偏估计.

证明: 由样本均值和方差的性质,

$$E(X_1) = E(\overline{X}) = \mu, \quad E(S^2) = \sigma^2.$$

所以无偏性成立. 需注意, 无偏估计可以不只有一个. 此外, 这里也解释了样本方差 S^2 要除以 $n-1$ 是出于无偏性的考虑. □

点估计量的优良准则: 无偏性

例 3.5

设 $X_1, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的简单随机样本, 样本方差 S^2 是 σ^2 的无偏估计. 验证 S 不是 σ 的无偏估计.

证明: 因为 $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$,

$$Y \equiv \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \Rightarrow S = \frac{\sigma}{\sqrt{n-1}} Y^{1/2}.$$

求 S 的期望转换为求 $Y^{1/2}$ 的期望. 先从 Y 的密度函数出发,

$$f(y) = \frac{1}{2^{\frac{n-1}{2}} \Gamma(\frac{n-1}{2})} y^{\frac{n-1}{2}-1} e^{-\frac{y}{2}}, \quad y > 0.$$

点估计量的优良准则: 无偏性

从而

$$
\begin{aligned}
E(Y^{1/2}) &= \int_0^\infty y^{1/2} \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})} y^{\frac{n-1}{2}-1} e^{-\frac{y}{2}} dy \\
&= \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})} \int_0^\infty y^{\frac{n}{2}-1} e^{-\frac{y}{2}} dy \\
&= \frac{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})} = \sqrt{2}\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}.
\end{aligned}
$$

所以,

$$
E(S) = E\left[\frac{\sigma}{\sqrt{n-1}} Y^{1/2}\right] = \sqrt{\frac{2}{n-1}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \sigma \neq \sigma.
$$

因此, S 不是 σ 的无偏估计. $\square$

点估计量的优良准则: 有效性

定义 3.3

设 $\hat{g}_1(\boldsymbol{X}) = \hat{g}_1(X_1, \cdots, X_n)$ 和 $\hat{g}_2(\boldsymbol{X}) = \hat{g}_2(X_1, \cdots, X_n)$ 为 $g(\theta)$ 的两个不同无偏估计量, 若

$$D_\theta(\hat{g}_1(\boldsymbol{X})) \leq D_\theta(\hat{g}_2(\boldsymbol{X})), \quad \forall \theta \in \Theta,$$

且至少存在一个 $\theta \in \Theta$, 使得严格不等号成立, 则称估计量 $\hat{g}_1(\boldsymbol{X})$ 比 $\hat{g}_2(\boldsymbol{X})$ 有效.

例 3.6

在无偏估计类中, 方差越小的估计量越有效. 利用 X_1 和 $\overline{X}$ 估计 μ, 方差为 $D_{X_1} = \sigma^2 > \frac{\sigma^2}{n} = D(\overline{X})$. 可见 $\overline{X}$ 比 X_1 更有效, 且 n 越大, $\overline{X}$ 对 μ 的估计越有效.

点估计量的优良准则: 有效性

例 3.7

设 $X_1, \cdots, X_n$ 是抽自均值为 μ 的总体, 考虑 μ 的两个估计

$$\hat{\mu} = \bar{X}, \quad \hat{\mu}_{(-i)} = \frac{1}{n-1}\sum_{j\neq i} X_j,$$

这里 $\hat{\mu}_{(-i)}$ 表示去掉第 i 个样本 X_i 后, 对其余 $n-1$ 个样本所求的样本均值. 显然, $\hat{\mu}$ 和 $\hat{\mu}_{(-i)}$ 都是无偏估计. 二者哪个更有效? 因为 $X_1, \cdots, X_n$ 独立同分布,

$$Var(\hat{\mu}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{\sigma^2}{n},$$

$$Var(\hat{\mu}_{(-i)}) = \frac{\sigma^2}{n-1} > \frac{\sigma^2}{n} = Var(\hat{\mu}).$$

因此, $\hat{\mu} = \bar{X}$ 更有效.

7.4 正态总体的区间估计 (一)

正态总体的区间估计

定义 4.1

设有一个参数分布族 $\{f(x,\theta),\theta\in\Theta\}$, $g(\theta)$ 是定义在参数空间 Θ 上的已知函数, $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从分布族某总体 $f(x,\theta)$ 中抽取的样本, 令 $\hat{g}_1(\boldsymbol{X})$ 和 $\hat{g}_2(\boldsymbol{X})$ 为定义在样本空间 $\mathscr{X}$ 上, 取值在 Θ 上的两个统计量, 且 $\hat{g}_1(\boldsymbol{X})\le\hat{g}_2(\boldsymbol{X})$, 则称随机区间 $[\hat{g}_1(\boldsymbol{X}),\hat{g}_2(\boldsymbol{X})]$ 为 $g(\theta)$ 的一个区间估计 (*interval estimation*).

注 4.1

注意定义中"随机区间"的概念, $\hat{g}_1(\boldsymbol{X})$ 和 $\hat{g}_2(\boldsymbol{X})$ 是统计量, 随样本变动. 而参数 $g(\theta)$ 是固定的. 由于样本是随机的, 不能保证在任何情况下, 随机区间 $[\hat{g}_1(\boldsymbol{X}),\hat{g}_2(\boldsymbol{X})]$ 必定包含 $g(\theta)$, 只能以一定概率保证. 希望随机区间 $[\hat{g}_1(\boldsymbol{X}),\hat{g}_2(\boldsymbol{X})]$ 包含 $g(\theta)$ 的概率越大越好.

正态总体的区间估计

任何一个满足条件 $\hat{g}_1 \le \hat{g}_2$ 的统计量 $\hat{g}_1 \le \hat{g}_2$ 都可构成 $g(\theta)$ 的一个区间估计 $[\hat{g}_1, \hat{g}_2]$. 如何挑选一个好的呢? 有两个要素: **可靠性**和**精度**.

- 可靠性是待估参数 $g(\theta)$ 包含住 $[\hat{g}_1, \hat{g}_2]$ 内的可能性大小. **可能性越大, 可靠性越强**.
- 精度可由随机区间的平均长度度量. **长度越短, 精度越高**.

二者常常是矛盾的, 不可能同时高. **在保证一定可靠性的前提下, 选择精度尽可能高的区间估计**. 如果应用中要求可靠性和精度同时提高, 只能增加样本量.

正态总体的区间估计

定义 4.2 (置信系数 (可靠性指标))

设随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 为参数 θ 的一个区间估计, 则 $[\hat{\theta}_1, \hat{\theta}_2]$ 包含 θ 的概率 $P_\theta(\hat{\theta}_1 \le \theta \le \hat{\theta}_2)$ 称为此区间估计的置信水平 *(confidence level)*. 置信水平在参数空间 Θ 上的下确界

$$\inf_{\theta \in \Theta} P_\theta(\hat{\theta}_1 \le \theta \le \hat{\theta}_2)$$

称为该区间估计的置信系数 *(coefidence coefficient)*.

定义 4.3 (精度指标)

精度的标准不止一个, 最常见的是随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 的平均长度 $E_\theta(\hat{\theta}_2 - \hat{\theta}_1)$. 平均长度越短, 精度越高.

正态总体的区间估计

例 4.1

设样本 $(X_1, \cdots, X_n)$ 来自正态总体 $\mathcal{N}(\mu, \sigma^2)$, 其中 $\mu \in \mathbb{R}$, $\sigma^2 > 0$. μ 和 σ^2 的估计量分别是样本均值 $\overline{X}$ 和样本方差 $S^2 = \sum_{i=1}^n (X_i - \overline{X})^2/(n-1)$, 用 $[\overline{X} - kS/\sqrt{n}, \overline{X} + kS/\sqrt{n}]$ 作总体均值 μ 的区间估计. 考虑其置信度和精度.

解: 记 $\theta = (\mu, \sigma^2)$, 上述区间估计的置信度为

$$\begin{aligned} P_\theta\left(\overline{X} - kS/\sqrt{n} \le \mu \le \overline{X} + kS/\sqrt{n}\right) &= P_\theta\left(|\sqrt{n}(\overline{X} - \mu)/S| \le k\right) \\ &= P(|T| \le k), \end{aligned}$$

其中 $T = \sqrt{n}(\overline{X} - \mu)/S \sim t_{n-1}$, 其分布与 θ 无关, 因而区间估计的置信系数为 $P(|T| \le k)$. 显然, **k 越大, 区间的置信系数越大, 区间就越可靠**.

正态总体的区间估计

由于 $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, 所以区间的平均长度为

$$l_k = E\left[\frac{2k}{\sqrt{n}}S\right] = \frac{2k}{\sqrt{n}}E[S] = \frac{2\sqrt{2}k\sigma\Gamma(n/2)}{\sqrt{n(n-1)}\Gamma\big((n-1)/2\big)}.$$

显然, k **越大, 区间也越长, 精度越差**.

- 在样本量 n 固定时, 为了提高置信度, 需增加 k 值, 从而放大了区间, 降低了精度.
- 反过来, 为了提高精度, 需要减少 k 值, 缩短区间, 降低了置信度.

置信度和精度相互制约着. 要想同时提高可靠性和精度, 只能增大样本量 n.

正态总体的区间估计: 置信区间

定义 4.4

设 $[\hat{\theta}_1(\boldsymbol{X}), \hat{\theta}_2(\boldsymbol{X})]$ 是参数 θ 的一个区间估计, 若对给定的 $0 < \alpha < 1$, 有

$$P_\theta\big(\hat{\theta}_1(\boldsymbol{X}) \leq \theta \leq \hat{\theta}_2(\boldsymbol{X})\big) \geq 1 - \alpha, \quad \theta \in \Theta,$$

则称 $[\hat{\theta}_1(\boldsymbol{X}), \hat{\theta}_2(\boldsymbol{X})]$ 是 θ 的置信水平为 $1 - \alpha$ 的置信区间 *(confidence interval)*.

注 4.2

设 $\alpha = 0.05$, 则 $1 - \alpha = 0.95$, 若把置信区间 $[\hat{\theta}_1(\boldsymbol{X}), \hat{\theta}_2(\boldsymbol{X})]$ 重复 *100* 次, 平均约有 *95* 次随机区间 $[\hat{\theta}_1(\boldsymbol{X}), \hat{\theta}_2(\boldsymbol{X})]$ 包含真参数 θ, 平均约有 *5* 次随机区间 $[\hat{\theta}_1(\boldsymbol{X}), \hat{\theta}_2(\boldsymbol{X})]$ 不包含 θ.

例 4.2

设 $\boldsymbol{X} = (X_1, \cdots, X_{10}) \sim \mathcal{N}(\mu, \sigma^2)$, 其中 $\mu \in \mathbb{R}, \sigma^2 > 0$. 参数 μ 的置信水平为 $1-\alpha$ 的置信区间为

$$[\overline{X} - kS/\sqrt{n}, \overline{X} + kS/\sqrt{n}].$$

若分别 $\alpha = 0.1$ 和 $\alpha = 0.5$, 利用模拟方法比较二者的置信区间.

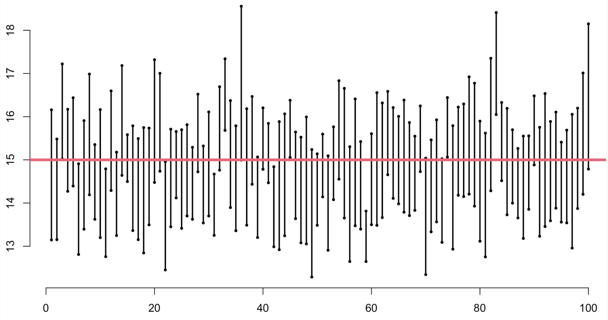
解: 设 $\mu = 15, \sigma^2 = 4$. (1) 若取 $\alpha = 0.1$, 由 α 确定 $k = 1.833$. 利用随机模拟的方法从 $N(15, 4)$ 中产生一个容量为 10 的样本, 例如 14.85, 13.01,13.50, 14.93, 6.97, 13.80, 17.95, 13.37, 16.29, 2.38.

正态总体的区间估计: 置信区间

计算其样本均值 $\bar{x} = 14.705$, 样本标准差 $s = 1.843$. 从而得到参数 μ 的区间估计为

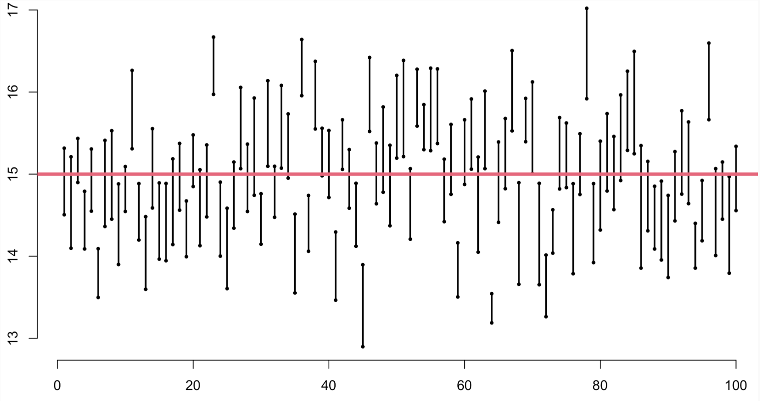
$$\left[14.705 - \frac{1.833 \times 1.843}{\sqrt{10}}, 14.705 + \frac{1.833 \times 1.843}{\sqrt{10}}\right] \approx [13.637, 15.773].$$

该区间包含真实参数 $\mu = 15$. 现重复这样的方法 100 次, 可得到 100 个样本, 得到 100 个区间.



正态总体的区间估计: 置信区间

(2) 若取 $\alpha = 0.5$, 由 α 可确定 $k = 0.7027$. 同样利用随机模拟的方法生成随机样本, 重复 100 次, 得到 100 个区间, 作出下图.



例 4.3

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 是从总体 $N(\mu, \sigma^2)$ 中抽取的简单随机样本, 其中 σ^2 已知, 求 μ 的置信系数为 $1-\alpha$ 的置信区间.

解: 显然, μ 的点估计是 $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$, 将其标准化

$$U = \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} \sim \mathcal{N}(0, 1),$$

其分布与 μ 无关. 由正态分布的对称性可知,

$$P_\mu\left(\left|\frac{\sqrt{n}(\overline{X} - \mu)}{\sigma}\right| \leq Z_{\alpha/2}\right) = 1 - \alpha,$$

此处 $Z_{\alpha/2}$ 为标准正态分布的上侧 $\alpha/2$ 分位数.

正态总体的区间估计: 枢轴量法

经不等式等价变形, 可知

$$P_\mu\left(\overline{X}-\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}<\mu<\overline{X}+\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\right)=1-\alpha.$$

因此

$$\left[\overline{X}-\frac{\sigma}{\sqrt{n}}Z_{\alpha/2},\overline{X}+\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\right]$$

为 μ 的置信系数 $1-\alpha$ 的置信区间. $\square$

正态总体的区间估计: 枢轴量法

利用枢轴量法构造置信区间的步骤如下:

(1) 找待估参数 μ 的一个良好估计, 如点估计 $T(\boldsymbol{X}) = \overline{X}$.

(2) 构造一个 $T(\boldsymbol{X})$ 与 μ 的函数 $\varphi(T, \mu)$, 使其满足 **其表达式与待估参数 μ 有关, 其分布与待估参数 μ 无关**, 则称随机变量 $\varphi(T, \mu)$ 为**枢轴变量**. 如变量 U.

(3) 对给定的 $0 < \alpha < 1$, 决定两个常数 a 和 b, 使得

$$P_\mu(a \leq \varphi(T, \mu) \leq b) = 1 - \alpha.$$

求解括号中的不等式得到 $\hat{\mu}_1(\boldsymbol{X}) \leq \mu \leq \hat{\mu}_2(\boldsymbol{X})$, 则有

$$P_\mu(\hat{\mu}_1(\boldsymbol{X}) \leq \mu \leq \hat{\mu}_2(\boldsymbol{X})) = 1 - \alpha.$$

这表明 $[\hat{\mu}_1(\boldsymbol{X}), \hat{\mu}_2(\boldsymbol{X})]$ 是 μ 置信水平为 $1 - \alpha$ 的置信区间.

正态总体 μ 的区间估计: 枢轴量法 (均值未知, 方差已知)

例 4.4 (均值未知, 方差已知)

设某车间生产零件的长度 $X \sim \mathcal{N}(\mu, 0.09)$, 若得到一组样本观察值为

$$12.6,\ 13.4,\ 12.8,\ 13.2.$$

求零件平均长度 μ 的 95% 的置信区间.

解: 由样本观测值计算 $\overline{X} = 13, n = 4, \sigma = 0.3$. 查表求得 $Z_{0.025} = 1.96$, 所以

$$\left[\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}, \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}\right] = \left[13 - \frac{0.3}{\sqrt{4}} \times 1.96, 13 + \frac{0.3}{\sqrt{4}} \times 1.96\right]$$
$$\approx [12.71, 13.29]$$

为参数 μ 的置信水平为 $1 - \alpha$ 的置信区间. □

μ 的区间估计: 枢轴量法 (均值未知, 方差未知)

应用上, σ^2 常常是未知的. 考虑

$$T = \frac{\sqrt{n}(\overline{X} - \mu)}{S} \sim t_{n-1}.$$

其中 T 的表达式与 μ 有关, 但其分布与 μ 无关, 可取为枢轴量. 由于 t 分布关于原点对称, 令

$$P(|T| \le c) = P\left(-c \le \frac{\sqrt{n}(\overline{X} - \mu)}{S} \le c \right) = 1 - \alpha,$$

则 $c = t_{n-1}(\alpha/2)$. 变形得到 μ 的置信系数为 $1 - \alpha$ 的置信区间为

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

μ 的区间估计: 枢轴量法 (均值未知, 方差未知)

例 4.5 (均值未知, 方差未知)

为测得某种溶液中的甲醛浓度, 取得 4 个独立测定的平均值 $\overline{X} = 8.34\%$, 样本标准差 $S = 0.03\%$, 设测量近似服从正态分布, 求总体均值 μ 的 95% 的置信区间.

解: 因为 $1 - \alpha = 0.95, n = 4, t_{n-1}(\alpha/2) = t_3(0.025) = 3.182$, 故有 $St_{n-1}(\alpha/2)/\sqrt{n} = 0.03 \times 3.182/2 = 0.0477, \overline{X} = 8.34$. 由此,

$$\left[\overline{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2), \overline{X} + \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2)\right] = [8.292\%, 8.388\%]$$

为 μ 的置信系数为 95% 的置信区间. $\square$

σ^2 的区间估计: 枢轴量法 (均值已知)

当 μ 已知时, σ^2 的一个良好估计为

$$S_\mu^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2,$$

且 $nS_\mu^2/\sigma^2 \sim \chi_n^2$, 则取 $T = nS_\mu^2/\sigma^2$ 为枢轴量, 其表达式与 σ^2 有关, 但其分布与 σ^2 无关, 找到 c_1 和 c_2 使得

$$P\left(c_1 \le \frac{nS_\mu^2}{\sigma^2} \le c_2\right) = 1-\alpha.$$

满足上式要求的 c_1 和 c_2 有无穷多对, 其中有一对 c_1 和 c_2, 使得区间长度最短. 但这样一对 c_1 和 c_2 不易求且表达式复杂, 应用不方便.

σ^2 的区间估计: 枢轴量法 (均值已知)

一般令 c_1 和 c_2 满足

$$P\left(\frac{nS_\mu^2}{\sigma^2} < c_1\right) = \frac{\alpha}{2}, \quad P\left(\frac{nS_\mu^2}{\sigma^2} > c_2\right) = \frac{\alpha}{2}.$$

由 χ^2 分布的上侧分位数表可知, $c_1 = \chi_n^2(1-\alpha/2)$, $c_2 = \chi_n^2(\alpha/2)$, 即有

$$P_{\sigma^2}\left(\chi_n^2\big(1-\frac{\alpha}{2}\big) \le \frac{nS_\mu^2}{\sigma^2} \le \chi_n^2\big(\frac{\alpha}{2}\big)\right) = 1-\alpha.$$

利用不等式变形得到 σ^2 的置信系数为 $1-\alpha$ 的置信区间为

$$\left[\frac{nS_\mu^2}{\chi_n^2(\alpha/2)}, \frac{nS_\mu^2}{\chi_n^2(1-\alpha/2)}\right].$$

σ^2 的区间估计: 枢轴量法 (均值已知)

例 4.6

为了解一台测量长度的仪器的精度, 对一根长 *30mm* 的标准金属棒进行了 *6* 次测量, 结果 *(*单位*: mm)* 是 *30.1, 29.9, 29.8, 30.3, 29.6.* 假如测量值服从正态分布 $N(30, \sigma^2)$, 求 σ^2 置信水平为 *0.95* 的置信区间.

解: 此处 $n = 6, \mu = 30$, 得 $\sum_{i=1}^{6}(X_i - \mu)^2 = 0.35, \alpha = 0.05$, 查表得 $\chi_6^2(0.025) = 14.4494, \chi_6^2(0.975) = 1.2375$. 计算

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\chi_n^2(\alpha/2)} = \frac{0.35}{14.4494} \approx 0.0242,$$

$$\hat{\sigma}_2^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\chi_n^2(1-\alpha/2)} = \frac{0.35}{1.2375} \approx 0.2828.$$

因此, σ^2 置信水平为 0.95 的置信区间为 $[0.0242, 0.2828]$. □

σ^2 的区间估计: 枢轴量法 (均值未知)

记 $\theta = (\mu, \sigma^2)$. 此时

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

是 σ^2 的良好无偏估计. 因为 $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. 取

$$T = (n-1)S^2/\sigma^2$$

为枢轴量, 其表达式与 σ^2 有关, 而其分布与 σ^2 无关. 找到 d_1 和 d_2 使得

$$P_\theta\left(d_1 \le \frac{(n-1)S^2}{\sigma^2} \le d_2\right) = 1 - \alpha.$$

σ^2 的区间估计: 枢轴量法 (均值未知)

取 $d_1 = \chi^2_{n-1}(1-\alpha/2), d_2 = \chi^2_{n-1}(\alpha/2)$, 故有

$$P_{\sigma^2}\left(\chi^2_n(1-\frac{\alpha}{2}) \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_n(\frac{\alpha}{2})\right) = 1-\alpha.$$

利用不等式变形得到 σ^2 的置信系数为 $1-\alpha$ 的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_n(\alpha/2)}, \frac{(n-1)S^2}{\chi^2_n(1-\alpha/2)}\right].$$

σ^2 的区间估计: 枢轴量法 (均值未知)

例 4.7

为测得某种溶液中的甲醛浓度, 取得 *4* 个独立测定的平均值 $\overline{X} = 8.34\%$, 样本标准差 $S = 0.03\%$, 设测量近似服从正态分布, 求总体方差 σ^2 的 *95%* 的置信区间.

解: 同上例, 由 $n-1=3, \alpha/2=0.025$, 查表得

$$\chi_3^2(0.025) = 9.348, \chi_3^2(0.975) = 0.216, S^2 = 0.0009$$

可知

$$\left[\frac{(n-1)S^2}{\chi_n^2(\alpha/2)}, \frac{(n-1)S^2}{\chi_n^2(1-\alpha/2)}\right] = [0.00029, 0.0125]$$

为 σ^2 的置信系数为 95% 的置信区间. □

7.5 正态总体的区间估计 (二)

正态总体的区间估计

设

- $X_1, \cdots, X_m$ 是自正态总体 $\mathcal{N}(\mu_1, \sigma_1^2)$ 抽取的简单随机样本.
- $Y_1, \cdots, Y_n$ 是自正态总体 $\mathcal{N}(\mu_2, \sigma_2^2)$ 抽取的简单随机样本.
- 且 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_n$ 相互独立.

设 $\overline{X}, \overline{Y}$ 和 S_1^2, S_2^2 分别为这两组样本的样本均值和样本方差,其中

$$S_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \overline{X})^2, \quad S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2.$$

正态总体的区间估计: $\mu_1 - \mu_2$ 的区间估计

(1) 当 $m = n, \sigma_1^2, \sigma_2^2$ 未知时. 令 $Z_i = Y_i - X_i, i = 1, 2, \cdots, n$, 且记 $\tilde{\mu} = \mu_1 - \mu_2, \tilde{\sigma}^2 = \sigma_1^2 + \sigma_2^2$, 则有

$$Z_i \sim N(\tilde{\mu}, \tilde{\sigma}^2), \quad i = 1, 2, \cdots, n.$$

$\overline{Z} = \overline{Y} - \overline{X}$ 是一个良好的无偏估计, 枢轴量为

$$T_Z = \frac{\sqrt{n}(\overline{Z} - \tilde{\mu})}{S_Z} \sim t_{n-1},$$

其中 $S_Z^2 = \sum_{i=1}^n (Z_i - \overline{Z})^2/(n-1)$, T_Z 的表达式与 $\tilde{\mu}$ 有关, 但其分布与 $\tilde{\mu}$ 无关. 因此, $\tilde{\mu}$ 的 $1-\alpha$ 的置信区间为

$$\left[\overline{Z} - \frac{S_Z}{\sqrt{n}} t_{n-1}(\alpha/2), \overline{Z} + \frac{S_Z}{\sqrt{n}} t_{n-1}(\alpha/2)\right],$$

(2) 当 σ_1^2 和 σ_2^2 已知时. 易知 $\overline{Y}-\overline{X}$ 是 $\mu_1-\mu_2$ 的一个良好无偏估计, 且 $\overline{Y}-\overline{X}\sim N(b-a,\sigma_1^2/m+\sigma_2^2/n)$, 由此可知

$$U=\frac{\overline{Y}-\overline{X}-(b-a)}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}\sim N(0,1).$$

U 的表达式与 $\mu_1-\mu_2$ 有关, 但其分布与 $\mu_1-\mu_2$ 无关, 取 U 为枢轴量, 故有

$$P_{\mu_1,\mu_2}=\left(\left|\frac{\overline{Y}-\overline{X}-(\mu_2-\mu_1)}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}\right|\le u_{\alpha/2}\right)=1-\alpha.$$

括号中不等式的等价变形得到 $b-a$ 的置信系数为 $1-\alpha$ 的置信区间, 故有

$$\left[\overline{Y}-\overline{X}-u_{\alpha/2}\sqrt{\sigma_1^2/m+\sigma_2^2/n},\quad \overline{Y}-\overline{X}+u_{\alpha/2}\sqrt{\sigma_1^2/m+\sigma_2^2/n}\right].$$

正态总体的区间估计: $\mu_1 - \mu_2$ 的区间估计

(3) 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时. 令

$$
\begin{aligned}
S_\omega^2 &= \frac{1}{m+n-2}[(m-1)S_1^2 + (n-1)S_2^2] \\
&= \frac{1}{m+n-2}\left[\sum_{i=1}^{m}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(Y_i - \overline{Y})^2\right],
\end{aligned}
$$

显然 $\overline{Y} - \overline{X}$ 是 $b - a$ 的无偏估计, 可知

$$
T_\omega = \frac{\overline{Y} - \overline{X} - (b-a)}{S_\omega}\sqrt{\frac{mn}{m+n}} \sim t_{n+m-2},
$$

其中 T_ω 的表达式与 $b - a$ 有关, 但其分布与 $b - a$ 无关.

正态总体的区间估计: $\mu_1 - \mu_2$ 的区间估计

取 T_ω 为枢轴量, 故有

$$P\left(\left|\frac{\overline{Y}-\overline{X}-(b-a)}{S_\omega}\right|\sqrt{\frac{mn}{m+n}} \le t_{m+n-2}(\alpha/2)\right) = 1-\alpha,$$

括号中不等式的等价变形得到 $b-a$ 的置信系数为 $1-\alpha$ 的置信区间为

$$\left[\overline{Y}-\overline{X}-S_\omega t_{m+n-2}\left(\frac{\alpha}{2}\right)\sqrt{\frac{1}{m}+\frac{1}{n}},\right.$$
$$\left.\overline{Y}-\overline{X}+S_\omega t_{m+n-2}\left(\frac{\alpha}{2}\right)\sqrt{\frac{1}{m}+\frac{1}{n}}\right],$$

例 5.1

某公司利用两条自动化流水线罐装矿泉水. 现从生产线上抽取样本 $X_1,\cdots,X_{12}$ 和 $Y_1,\cdots,Y_{17}$, 它们是每瓶矿泉水的体积 (单位: ml). 计算样本均值 $\overline{X}=501.1$ 和 $\overline{Y}=499.7$; 样本方差 $S_1^2=2.4$ 和 $S_2^2=4.7$. 假设这两条流水线所装的矿泉水的体积分别服从正态分布 $N(\mu_1,\sigma^2)$ 和 $N(\mu_2,\sigma^2)$, 试求 $\mu_2-\mu_1$ 置信系数为 0.95 的置信区间.

解: 由题意, $\overline{Y}-\overline{X}=-1.4$,

$$S_\omega^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{n+m-2}=\frac{11\times 2.4+16\times 4.7}{12+17-2}\approx 3.763.$$

查表求得 $t_{m+n-2}(0.025)=2.05$.

正态总体的区间估计: $\mu_1 - \mu_2$ 的区间估计例子

算得

$$
\begin{aligned}
&\left[\overline{Y} - \overline{X} - S_\omega t_{m+n-2}\left(\frac{\alpha}{2}\right)\sqrt{\frac{1}{m} + \frac{1}{n}},\right.\\
&\qquad \left.\overline{Y} - \overline{X} + S_\omega t_{m+n-2}\left(\frac{\alpha}{2}\right)\sqrt{\frac{1}{m} + \frac{1}{n}}\right]\\
=&\left[-1.4 - \sqrt{3.763} \times 2.05 \times \sqrt{\frac{1}{12} + \frac{1}{17}},\right.\\
&\qquad \left.-1.4 + \sqrt{3.763} \times 2.05 \times \sqrt{\frac{1}{12} + \frac{1}{17}}\right]\\
\approx&[-2.9, 0.1]
\end{aligned}
$$

为 $\mu_2 - \mu_1$ 的 95% 的置信区间. $\square$

7.6 非正态总体的区间估计

二项分布参数的置信区间

设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是两点分布 $B(1,p)$ 的简单随机样本, 求 p 的置信区间. 令 $S_n=\sum_{i=1}^n X_i$, 可知 $S_n\sim B(n,p)$. 利用中心极限定理, 有

$$\frac{S_n-np}{\sqrt{np(1-p)}}=\frac{\sqrt{n}(\overline{X}-p)}{\sqrt{p(1-p)}}\xrightarrow{d}N(0,1),\quad n\to\infty.$$

这表明当 n 充分大时, 随机变量 $T=\sqrt{n}(\overline{X}-p)/\sqrt{p(1-p)}$ 的极限分布是 $N(0,1)$, 与未知参数 p 无关, 可取 T 为枢轴量. 当 n 充分大时有

$$P(|T|\leq u_{\alpha/2})=P\left(\left|\frac{\sqrt{n}(\overline{X}-p)}{\sqrt{p(1-p)}}\right|\leq u_{\alpha/2}\right)\approx 1-\alpha.$$

二项分布参数的置信区间

由 $\hat{p} = S_n/n \xrightarrow{P} p$, 则当 $n \to \infty$, 有

$$\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1), \quad \sqrt{\frac{p(1-p)}{\hat{p}(1-\hat{p})}} \xrightarrow{P} 1.$$

故由 Slutsky 定理可知,

$$T = \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}} = \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}} \times \sqrt{\frac{p(1-p)}{\hat{p}(1-\hat{p})}} \xrightarrow{d} \mathcal{N}(0,1),$$

故可取 T 为枢轴量, 其极限分布与 p 无关. 令

$$P\left(\left|\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}}\right| \le u_{\alpha/2}\right) \approx 1-\alpha.$$

利用不等式变形得

$$\left[\hat{p} - u_{\alpha/2}\cdot\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad \hat{p} + u_{\alpha/2}\cdot\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right].$$

二项分布参数的置信区间例子

例 6.1

对某事件 A 作 120 次观察, A 发生 36 次. 试给出事件 A 发生概率 p 的 0.95 置信区间.

解: 由题意 $n = 120, \hat{p} = \bar{x} = 36/120 = 0.3$, $u_{0.025} = 1.96$, 于是

$$\left[0.3 - 1.96 \times \sqrt{\frac{0.3 \times 0.7}{120}}, 0.3 + 1.96 \times \sqrt{\frac{0.3 \times 0.7}{120}}\right] = [0.218, 0.382].$$

为比例 p 置信系数为 0.95 的置信区间. □

泊松分布参数的置信区间

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 为抽自 Poisson 总体 $\mathcal{P}(\lambda)$ 的简单随机样本, 求 λ 的置信区间. 记 $S_n = \sum_{i=1}^n X_i$, 则 S_n 为服从参数为 $n\lambda$ 的 Poisson 分布, 即

$$P(S_n = k) = \frac{e^{-n\lambda}(n\lambda)^k}{k!}, \quad k = 0, 1, 2, \cdots$$

当 n 充分大时, 由中心极限定理可知:

$$\frac{S_n - n\lambda}{\sqrt{n\lambda}} = \frac{\sqrt{n}(\overline{X} - \lambda)}{\sqrt{\lambda}} \xrightarrow{d} N(0, 1), \quad n \to \infty.$$

将随机变量 $T = \sqrt{n}(\overline{X} - \lambda)/\sqrt{\lambda}$ 作枢轴量, 其极限分布与未知参数 λ 无关.

泊松分布参数的置信区间

由 $\hat{\lambda} = \overline{X} \xrightarrow{P} \lambda$, 当 $n \to \infty$ 时有

$$\frac{\sqrt{n}(\overline{X} - \lambda)}{\sqrt{\lambda}} \xrightarrow{d} \mathcal{N}(0,1), \quad \frac{\sqrt{\lambda}}{\hat{\lambda}} \xrightarrow{P} 1.$$

由 Slutsky 定理可知,

$$\frac{\sqrt{n}(\overline{X} - \lambda)}{\sqrt{\hat{\lambda}}} = \frac{\sqrt{n}(\overline{X} - \lambda)}{\sqrt{\lambda}} \cdot \frac{\sqrt{\lambda}}{\hat{\lambda}} \xrightarrow{d} N(0,1).$$

令 $T = \sqrt{n}(\hat{\lambda} - \lambda)/\sqrt{\hat{\lambda}}$ 为枢轴量, 其极限分布与未知参数 λ 无关. 给定置信系数 $1 - \alpha$, 则有

$$P\left(\left|\frac{\sqrt{n}(\hat{\lambda} - \lambda)}{\sqrt{\hat{\lambda}}}\right| \leq u_{\alpha/2}\right) \approx 1 - \alpha.$$

不等式变形得到, λ 的置信系数近似为 $1 - \alpha$ 的置信区间为

$$\left[\hat{\lambda} - u_{\alpha/2}\sqrt{\hat{\lambda}/n}, \hat{\lambda} + u_{\alpha/2}\sqrt{\hat{\lambda}/n}\right],$$

泊松分布参数的置信区间

例 6.2

公共汽车站在一单位时间内（如 20 分钟）到达的乘客数服从泊松分布 $P(\lambda)$, 对不同的车站, 所不同的仅仅是参数 λ 的取值不同. 现对一城市某一公共汽车进行了 100 个单位时间的调查. 计算得到每 20 分钟内来到该车站的乘客数平均值 $\bar{X} = 15.2$ 人. 求参数 λ 的置信系数为 0.95 的置信区间.

解: $n = 100, \alpha = 0.05, Z_{\alpha/2} = 1.96, \bar{X} = 15.2$. 可得

$$\left[15.2 - 19.6 \times \sqrt{\frac{15.2}{100}}, 15.2 + 19.6 \times \sqrt{\frac{15.2}{100}}\right] = [14.44, 15.96]$$

为参数 λ 的置信系数为 0.95 的置信区间. $\square$