第三次习题课

2023 年 4 月 22 日

目录

- 基本概念回顾
  - 最小方差无偏估计
  - 区间估计
- 作业题讲解
- 补充例题

基本概念回顾

均方误差 (mean square error, MSE)

设 $\hat{g}(\boldsymbol{X})$ 为 $g(\theta)$ 的估计量, 则称 $\mathrm{E}_\theta(\hat{g}(\boldsymbol{X})-g(\theta))^2$ 为 $\hat{g}(\boldsymbol{X})$ 的均方误差.

- 若存在 $\hat{g}^*(\boldsymbol{X})$, 使得对 $g(\theta)$ 的任一估计量 $\hat{g}(\boldsymbol{X})$ 都有

$$\mathrm{E}_\theta(\hat{g}^*(\boldsymbol{X})-g(\theta))^2 \le \mathrm{E}_\theta(\hat{g}(\boldsymbol{X})-g(\theta))^2,\ \forall\theta\in\Theta,$$

  则称 $\hat{g}^*(\boldsymbol{X})$ 为 $g(\theta)$ 的一致最小均方误差估计.
- 一致最小均方误差估计常不存在.
- 当 $\hat{g}^*(\boldsymbol{X})$ 为 $g(\theta)$ 的无偏估计时,
  $\mathrm{MSE}(\hat{g}^*(\boldsymbol{X}))=\mathrm{Var}_\theta(\hat{g}^*(\boldsymbol{X}))$.
- 参数的无偏估计存在, 则称此参数为可估参数; 若参数函数的无偏估计存在, 则称此函数为可估函数.

基本概念回顾

一致最小方差无偏估计 (uniform minimum variance unbiased estimator, UMVUE)

设 $\mathcal{F} = \{f(x, \theta), \theta \in \Theta\}$ 是一个参数分布族, $g(\theta)$ 为定义在 Θ 上的可估函数. 设 $\hat{g}^*(\boldsymbol{X}) = \hat{g}^*(X_1, \cdots, X_n)$ 为 $g(\theta)$ 的一个无偏估计, 若对 $g(\theta)$ 的任一无偏估计 $\hat{g}(\boldsymbol{X})$ 都有

$$\mathrm{Var}_\theta(\hat{g}^*(\boldsymbol{X})) \le \mathrm{Var}_\theta(\hat{g}(\boldsymbol{X})), \ \forall \theta \in \Theta,$$

则称 $\hat{g}^*(\boldsymbol{X})$ 是 $g(\theta)$ 的一致最小方差无偏估计.

寻找 UMVUE 的方法:

- 充分完全统计量法
- 零无偏估计法
- Cramér-Rao 不等式法

基本概念回顾

Rao-Blackwell 定理

设 $T = T(\boldsymbol{X})$ 是一个充分统计量, 而 $\hat{g}(\boldsymbol{X})$ 是 $g(\theta)$ 的一个估计, 则

- $h(T) = \mathrm{E}(\hat{g}(\boldsymbol{X})|T)$ 是一个统计量.
- $h(T)$ 和 $\hat{g}(\boldsymbol{X})$ 有着相同的偏倚; 若 $\hat{g}(\boldsymbol{X})$ 是无偏的则 $h(T)$ 也是无偏的.
- $\mathrm{Var}_\theta(h(T)) \le \mathrm{Var}_\theta(\hat{g}(\boldsymbol{X}))$, $\theta \in \Theta$, 其中等号成立当且仅当 $\mathrm{P}_\theta(\hat{g}(\boldsymbol{X}) = h(T)) = 1$.

该定理表明一致最小方差无偏估计一定是充分统计量的函数.

基本概念回顾

充分完全统计量法:

Lehmann-Scheffé 定理 (L-S 定理)

设 $X \sim \{f(x,\theta), \theta \in \Theta\}$, Θ 为参数空间. 令 $X_1, \cdots, X_n$ 为从总体 X 中抽取的样本, $g(\theta)$ 为定义在 Θ 上的可估函数, $T(\boldsymbol{X})$ 为一个充分完全统计量. 若 $\hat{g}(T(\boldsymbol{X}))$ 为 $g(\theta)$ 的一个无偏估计, 则 $\hat{g}(T(\boldsymbol{X}))$ 是 $g(\theta)$唯一的 UMVUE.

- 唯一性的意义: 设 $\hat{g}$ 和 $\hat{g}_1$ 是 $g(\theta)$ 的两个估计量, 若 $\mathrm{P}_\theta(\hat{g} = \hat{g}_1) = 1$ 对一切 $\theta \in \Theta$, 则视 $\hat{g}$ 和 $\hat{g}_1$ 是同一个估计量.
- 指数族的应用 (本质上还是寻找充分完全统计量).

基本概念回顾

零无偏估计法:

定理 3

设 $\boldsymbol{X}$ 是来自 $\{F_\theta, \theta \in \Theta\}$ 的样本, $\mathcal{G} = \{\hat{g} : \mathrm{E}_\theta \hat{g}^2(\boldsymbol{X}) < \infty, \forall \theta \in \Theta\}$, $\hat{g}(\boldsymbol{X})$ 是属于 $\mathcal{G}$ 中的一个无偏估计量, 记 $\mathcal{U}$ 为 $\mathcal{G}$ 中所有零无偏估计量的集合. 则 $\hat{g}(\boldsymbol{X})$ 是 $g(\theta)$ 的 UMVUE 的充分必要条件为

$$\mathrm{E}_\theta(\hat{g}(\boldsymbol{X})U) = 0,\ \forall U \in \mathcal{U},\ \forall \theta \in \Theta.$$

- 一个推论: $T(\boldsymbol{X})$ 为 θ 的充分统计量, $h(T) \in \mathcal{G}$ 是 $g(\theta)$ 的一个无偏估计, 对 $\forall \theta \in \Theta$, 和 $\delta(T)$ 满足 $\mathrm{E}_\theta \delta(T) = 0$, 都有

  $$\mathrm{Cov}_\theta(h(T), \delta(T)) = \mathrm{E}_\theta[h(T)\delta(T)] = 0,\ \forall \theta \in \Theta,$$

  则 $h(T)$ 是 $g(\theta)$ 的 UMVUE.
- 该定理可用于证明 UMVUE 不存在.

基本概念回顾

Cramer-Rao 不等式法:

C-R 正则条件

- 参数空间 Θ 是直线上的某个开区间;
- 分布支撑 $A = \{x|f(x,\theta) > 0\}$ 不依赖于 θ;
- 对 $\forall x \in A$, $f(x,\theta)$ 关于 θ 可导, 且积分与微分运算可交换, 即
$$\frac{\partial}{\partial\theta}\int f(x,\theta)dx = \int \frac{\partial}{\partial\theta} f(x,\theta)dx;$$
- 存在
$$0 < I(\theta) = \mathrm{E}_\theta\left[\frac{\partial \log f(x,\theta)}{\partial\theta}\right]^2 < \infty$$

基本概念回顾

Cramer-Rao 下界

设 $\mathcal{F} = \{f(x,\theta), \theta \in \Theta\}$ 是 C-R 正则分布族，$g(\theta)$ 是一个可微函数，设 $X_1, \cdots, X_n$ 是来自总体 $\{f(x,\theta)\}$ 的一个简单随机样本，$\hat{g}(\boldsymbol{X})$ 是 $g(\theta)$ 的一个无偏估计量且满足积分 $\int \cdots \int \hat{g}(\boldsymbol{x}) f(\boldsymbol{x},\theta) d\boldsymbol{x}$ 可在积分号下对 θ 求导，则有

$$\mathrm{Var}_\theta(\hat{g}) \geq \frac{[g'(\theta)]^2}{nI(\theta)}, \forall \theta \in \Theta.$$

利用 C-R 不等式判断无偏估计 $\hat{g}(\boldsymbol{X})$ 是否为 $g(\theta)$ 的 UMVUE:

1. 验证分布族是否为 C-R 正则分布族
2. 计算 Fisher 信息量 $I(\theta)$ 和无偏估计 $\hat{g}(\boldsymbol{X})$ 的方差，看其是否达到 C-R 下界

基本概念回顾

Theorem 5

在 C-R 正则条件下，可估函数 $g(\theta)$ 不恒等于一个常数，则其无偏估计量 $\hat{g}(\boldsymbol{X})$ 达到 C-R 下界当且仅当存在 $N \in \mathcal{B}$, $\mu(N) = 0$, 当 $x \notin N$, $\theta \in \Theta$ 时有

$$f(x, \theta) = C(\theta) h(x) \exp\{Q(\theta) \hat{g}(x)\},$$

其中 $C(\theta) > 0$, Q 在 Θ 上可微、严格单调增，且 $h(x)$ 恒正.

注: $\hat{g}(\boldsymbol{X})$ 的方差达不到 C-R 下界，不能说明 $g(\theta)$ 的 UMVUE 不存在. 但若 $\hat{g}(\boldsymbol{X})$ 的方差达到此下界，则 $\hat{g}(\boldsymbol{X})$ 一定是 $g(\theta)$ 的 UMVUE.

基本概念回顾

无偏估计的不足:

- UMVUE 在平方误差损失下可能是不可容许的.
- 无偏估计可能不存在.
- σ 的 UMVUE 不是 σ^2 的 UMVUE 的平方根. (MLE 可以, 但 UMVUE 通常不可以!)
- 无偏性无关紧要.

基本概念回顾

效率

$g(\theta)$ 无偏估计 $\hat{g}(\boldsymbol{X})$ 的效率定义为

$$e_{\hat{g}}(\theta) = \frac{[g'(\theta)]^2/(nI(\theta))}{\mathrm{Var}_\theta(\hat{g}(\boldsymbol{X}))}.$$

- $0 < e_{\hat{g}}(\theta) \leq 1$.
- $e_{\hat{g}}(\theta) = 1$, 有效估计.
- $\lim_{n\to\infty} e_{\hat{g}}(\theta) = 1$, 渐近有效估计.

基本概念回顾

区间估计

实参数 θ 的区间估计是任意一对统计量 $L(\boldsymbol{X})$ 和 $U(\boldsymbol{X})$, 其中对任意的 $\boldsymbol{X} \in \mathcal{X}$ 都满足 $L(\boldsymbol{X}) \leq U(\boldsymbol{X})$ 组成的区间 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$.

- 可靠度: $\mathrm{P}_\theta\left(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})]\right)$.
- 精确度: 由随机区间的平均长度度量.

基本概念回顾

置信区间 (CI)

设 $X_1, \cdots, X_n \sim F_\theta$, $\theta \in \Theta$. 设 $L(\boldsymbol{X}) \leq U(\boldsymbol{X})$ a.s. 为两个统计量, 则称区间 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ 为 θ 的 $100(1-\alpha)\%$ 置信区间, 若 $\mathrm{P}_\theta\left(L(\boldsymbol{X}) \leq \theta \leq U(\boldsymbol{X})\right) \geq 1-\alpha$, $\forall \theta \in \Theta$ 且"$=$" 至少对某个 θ 成立. $1-\alpha$ 称为置信区间的置信水平 (置信系数). 称区间 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ 为 θ 的大样本 (渐近)$100(1-\alpha)\%$ 置信区间, 若 $\mathrm{P}_\theta\left(L(\boldsymbol{X}) \leq \theta \leq U(\boldsymbol{X})\right) \to 1-\alpha$, $\forall \theta \in \Theta$ 当 $n \to \infty$, 且"$=$" 至少对某个 θ 成立.

- 置信下界: $\mathrm{P}_\theta\left(\theta \geq L(\boldsymbol{X})\right)$.
- 置信上界: $\mathrm{P}_\theta\left(\theta \leq U(\boldsymbol{X})\right)$.
- 若 $L(\boldsymbol{X})$ 为 θ 的 $1-\alpha_1$ 水平置信下界, $U(\boldsymbol{X})$ 为 θ 的 $1-\alpha_2$ 水平置信上界, 则 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ 为 θ 的 $1-\alpha_1-\alpha_2$ 的置信区间.

基本概念回顾

枢轴变量

称随机变量 $g(\boldsymbol{X}, \theta)$ 是枢轴变量，若它的分布与 θ 无关.

求解步骤:

- 找到一个枢轴变量 $g(\boldsymbol{X}, \theta)$ 和常数 $a < b$ 使得

$$\mathrm{P}_\theta\left(a \le g(\boldsymbol{X}, \theta) \le b\right) \ge 1 - \alpha$$

- 反推得到

$$C(\boldsymbol{X}) = \{\theta : a \le g(\boldsymbol{X}, \theta) \le b\} = \{\theta : L(\boldsymbol{X}, a) \le \theta \le U(\boldsymbol{X}, b)\}$$

基本概念回顾

单个正态总体参数的置信区间

$X_1, \cdots, X_n$ i.i.d. $N(\mu, \sigma^2)$. 记

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

σ^2 已知, 求 μ 的置信区间

枢轴变量: $U = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$

置信区间: $\left[\bar{X} - \frac{\sigma}{\sqrt{n}} u_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha/2}\right]$

σ^2 未知, 求 μ 的置信区间

枢轴变量: $T = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t_{n-1}$

置信区间: $\left[\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2), \bar{X} + \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2)\right]$

基本概念回顾

单个正态总体参数的置信区间

μ 已知, 求 σ^2 的置信区间

枢轴变量: $T = nS_\mu^2/\sigma^2 \sim \chi_n^2$, 其中 $S_\mu^2 = \sum_{i=1}^n (X_i - \mu)^2/n$

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

μ 未知, 求 σ^2 的置信区间

枢轴变量: $T = (n-1)S^2/\sigma^2 \sim \chi_{n-1}^2$

置信区间: $\left[\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha/2)}, \frac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha/2)}\right]$

基本概念回顾

两个正态总体参数的置信区间

$X_1, \cdots, X_m$ i.i.d. $N(a, \sigma_1^2)$, $Y_1, \cdots, Y_n$ i.i.d. $N(b, \sigma_2^2)$, 且 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_n$ 独立. 样本均值分别为 $\bar{X}$ 和 $\bar{Y}$, 样本方差分别为 S_1^2 和 S_2^2.

均值差 $b-a$ 的置信区间

- $m=n$ 时, 令 $\bar{Z}=\bar{Y}-\bar{X}$, 记 $\tilde{\mu}=b-a$, $\tilde{\sigma}^2=\sigma_1^2+\sigma_2^2$ 未知, 枢轴变量 $T_z=\frac{\sqrt{n}(\bar{Z}-\tilde{\mu})}{S_z} \sim t_{n-1}$.
- σ_1^2 和 σ_2^2 已知, 枢轴变量 $U=\frac{\bar{Y}-\bar{X}-(b-a)}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}} \sim N(0,1)$.
- $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知,
  令 $S_\omega^2=\frac{1}{m+n-2}\left[(m-1)S_1^2+(n-1)S_2^2\right]$,
  枢轴变量 $T_\omega=\frac{\bar{Y}-\bar{X}-(b-a)}{S_\omega}\sqrt{\frac{mn}{m+n}} \sim t_{n+m-2}$.
- $\sigma_1^2 \neq \sigma_2^2$ 未知, Behrens-Fisher 问题.

基本概念回顾

两个正态总体参数的置信区间

方差比 σ_1^2/σ_2^2 的置信区间

- 若 a 和 b 已知，记 $S_a^2 = \sum_{i=1}^m (X_i - a)^2/m$, $S_b^2 = \sum_{i=1}^n (Y_i - b)^2/m$, 枢轴变量 $F = \frac{S_a^2/\sigma_1^2}{S_b^2/\sigma_2^2} \sim F_{m,n}$.
- 若 a 和 b 未知，枢轴变量 $F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{m-1,n-1}$.

基本概念回顾

自助法置信区间

$$
\begin{array}{ccccccccc}
f & \longrightarrow & \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} & \Longrightarrow & \hat{f}_n & \longrightarrow & \begin{bmatrix} X^*_{1(\mathbf{1})} \\ X^*_{2(\mathbf{1})} \\ \vdots \\ X^*_{n(\mathbf{1})} \end{bmatrix}, & \cdots, & \begin{bmatrix} X^*_{1(M)} \\ X^*_{2(M)} \\ \vdots \\ X^*_{n(\mathbf{M})} \end{bmatrix} \\
 & & \downarrow & & & & \downarrow & & \downarrow \\
 & & \hat{\theta} & & & & \hat{\theta}^*_{\mathbf{1}}, & \cdots, & \hat{\theta}^*_M
\end{array}
$$

习题三 32. 设 $X_1, \ldots, X_n$ i.i.d. $\sim N(\mu, \sigma^2)$, $-\infty < \mu < +\infty$, $\sigma^2 > 0$ 都是未知参数, 试求下列参数函数的 UMVUE.

(1) $3\mu + 4\sigma^2$; (2) $\mu^2/(4\sigma^2)$.

习题三 33. 设 $X_1, \ldots, X_n$ i.i.d. 服从两点分布 $b(1, p)$, $0 < p < 1$ 是未知参数, 试求下列参数函数的 UMVUE.

(1) p^s; (2) $p^s + (1-p)^{n-s}$ ($0 < s < n$ 为整数).

作业题讲解

习题三 **35.** 设 $\boldsymbol{X} = (X_1, \ldots, X_n)$ 是从下列几何分布中抽取的简单样本:

$$P(X_1 = i) = \theta(1-\theta)^{i-1}, \quad i = 1, 2, \ldots, \ 0 < \theta < 1.$$

(1) 试证明 $T = T(\boldsymbol{X}) = \sum_{i=1}^{n} X_i$ 是 θ 的充分完全统计量, 且服从帕斯卡 (负二项) 分布

$$P_\theta(T = t) = \binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}, \quad t = n, n+1, n+2, \ldots.$$

(2) 计算 $E_\theta(T)$, 并由此求 θ^{-1} 的 UMVUE.
(3) 试证明

$$\psi(X_1) = \begin{cases} 1, & X_1 = 1, \\ 0, & X_1 = 2, 3, \ldots \end{cases}$$

是 θ 的无偏估计, 计算 $E_\theta(\psi(X_1) \mid T = t)$, 并由此求得 θ 的 UMVUE.

作业题讲解

补充作业 1. 设 $X_1, \ldots, X_n$ 来自如下离散分布

$$\mathsf{P}(X_1 = \theta - 1) = \mathsf{P}(X_1 = \theta) = \mathsf{P}(X_1 = \theta + 1) = \frac{1}{3}, \quad \theta \in \mathbb{Z}.$$

作业题讲解

补充作业 2. 设 $X_1, \ldots, X_n$ 来自均匀分布 $U(\theta - 1/2, \theta + 1/2)$, $\theta \in \mathbb{R}$. 证明: 对任意 θ 的非常值函数, 不存在 UMVUE.

作业题讲解

习题三 42. 设 $X_1, \ldots, X_n$ i.i.d. $\sim N(0, \sigma^2)$, $\sigma^2 > 0$ 为未知参数.
(1) 求 σ^2 的无偏估计方差的 Cramér-Rao 下界.
(2) 求 σ^2 的一致最小方差无偏估计及它的效率.

作业题讲解

习题三 **43.** 设 $X_1, \ldots, X_n$ i.i.d. $\sim N(a, \sigma^2)$, a 已知, 证明:
$\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i - a|$ 为 σ 的无偏估计量, 且有效率为 $1/(\pi - 2)$.

补充例题

补充例题 1. 设 $X_1, \ldots, X_n$ i.i.d. $\sim U(0, \theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

补充例题

补充例题 1. 设 $X_1, \ldots, X_n$ i.i.d. $\sim U(0, \theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

注: 最大次序统计量 $X_{(n)}$ 对参数 θ 是充分的, 但不是完全的, 因为 $\theta \geq 1$!

补充例题

补充例题 1. 设 $X_1, \ldots, X_n$ i.i.d. $\sim U(0, \theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

注: 最大次序统计量 $X_{(n)}$ 对参数 θ 是充分的, 但不是完全的, 因为 $\theta \geq 1$!

解: $X_1, \ldots, X_n$ 的联合密度函数为 $\theta^{-n} I_{(0,\theta)}(x_{(n)})$. 由因子分解定理知 $X_{(n)}$ 是参数 θ 的充分统计量.

$X_{(n)}$ 的概率密度函数为

$$f(x, a) = n\theta^{-n} x^{n-1} I_{(0,\theta)}(x).$$

对于任意一个 θ 的零无偏估计 $U(X_{(n)})$ 使得 $\mathrm{E}_\theta U(X_{(n)}) = 0$ 有

$$\mathrm{E}_\theta U(X_{(n)}) = n\theta^{-n} \int_0^1 U(x) x^{n-1} dx + n\theta^{-n} \int_1^\theta U(x) x^{n-1} dx,$$

这推出 $U(x) = 0$ $a.s.$ P_θ 在 $[1, \infty)$ 上成立, 且 $\int_0^1 U(x) x^{n-1} dx = 0$.

补充例题

补充例题 1. 设 $X_1, \ldots, X_n$ i.i.d. $\sim U(0, \theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

解:

定理 3 的推论

$T(\boldsymbol{X})$ 为 θ 的充分统计量, $h(T) \in \mathcal{G}$ 是 $g(\theta)$ 的一个无偏估计, 对 $\forall \theta \in \Theta$, 和 $\delta(T)$ 满足 $\mathrm{E}_\theta \delta(T) = 0$, 都有

$$\mathrm{Cov}_\theta(h(T), \delta(T)) = \mathrm{E}_\theta[h(T)\delta(T)] = 0, \ \ \theta \in \Theta,$$

则 $h(T)$ 是 $g(\theta)$ 的 UMVUE.

补充例题

补充例题 1. 设 $X_1, \dots, X_n$ i.i.d. $\sim U(0,\theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

解:

定理 3 的推论

$T(\boldsymbol{X})$ 为 θ 的充分统计量, $h(T) \in \mathcal{G}$ 是 $g(\theta)$ 的一个无偏估计, 对 $\forall \theta \in \Theta$, 和 $\delta(T)$ 满足 $\mathrm{E}_\theta \delta(T) = 0$, 都有

$$\mathrm{Cov}_\theta(h(T), \delta(T)) = \mathrm{E}_\theta[h(T)\delta(T)] = 0, \ \theta \in \Theta,$$

则 $h(T)$ 是 $g(\theta)$ 的 UMVUE.

考虑 $h(X_{(n)})$. 为满足 $\mathrm{E}_\theta[h(X_{(n)})U(X_{(n)})] = 0$, 必须有

$$\int_0^1 h(x)U(x)x^{n-1}dx = 0$$

补充例题

补充例题 1. 设 $X_1, \ldots, X_n$ i.i.d. $\sim U(0, \theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

解: 因此我们可以考虑以下函数

$$h(x) = \begin{cases} c, & 0 \leq x \leq 1, \\ bx, & x > 1, \end{cases}$$

其中 b 和 c 是两个常数. 由之前的讨论可知,

$$\mathrm{E}_\theta[h(X_{(n)})U(X_{(n)})] = 0, \forall \theta \geq 1.$$

又由 $\mathrm{E}_\theta[h(X_{(n)})] = \theta$, 我们得到

$$\begin{aligned} \theta =& c\mathrm{P}_\theta(X_{(n)} \leq 1) + b\mathrm{E}_\theta[X_{(n)} I_{(1,\infty)}(X_{(n)})] \\ =& c\theta^{-n} + \frac{bn}{n+1}(\theta - \theta^{-n}). \end{aligned}$$

补充例题

补充例题 1. 设 $X_1, \ldots, X_n$ i.i.d. $\sim U(0, \theta)$, 其中 $\theta \geq 1$. 寻找 θ 的一个 UMVUE.

解: 取 $c = 1$, $b = \frac{n+1}{n}$. 因此 θ 的一个 UMVUE 为

$$h(X_{(1)}) = \begin{cases} 1, & 0 \leq X_{(n)} \leq 1, \\ \frac{n+1}{n} X_{(n)}, & X_{(n)} > 1. \end{cases}$$

补充例题

补充例题 2. 设 $X_1, \ldots, X_n$ i.i.d. $\sim N(\theta, \theta)$, $\theta > 0$ 未知, 寻找一个枢轴变量并构造 θ 的置信区间.

补充例题

补充例题 **2.** 设 $X_1, \ldots, X_n$ i.i.d. $\sim N(\theta, \theta)$, $\theta > 0$ 未知, 寻找一个枢轴变量并构造 θ 的置信区间.

解: 定义变量 $R = \frac{n(\bar{X}-\theta)^2}{\theta}$. 注意到 $\bar{X} \sim N(\theta, \theta/n)$, 则 $R \sim \chi_1^2$, R 的分布与 θ 无关, 所以 R 是一个枢轴变量.

令 $\chi_1^2(\alpha)$ 为 χ_1^2 的 $(1-\alpha)$ 分位数, 且 $c = \chi_1^2(\alpha)/(2n)$. 从而 $(1-\alpha)$ 置信区间为

$$C(\boldsymbol{X}) = \{\theta : R \leq \chi_1^2(\alpha)\} = \{\theta : \theta^2 - 2(\bar{X} + c)\theta + \bar{X}^2 \leq 0\}.$$

若 $\bar{X} \geq -c/2$, 则

$$C(\boldsymbol{X}) = \left[\bar{X} + c - \sqrt{2c\bar{X} + c^2}, \bar{X} + c + \sqrt{2c\bar{X} + c^2}\right].$$

否则, $C(\boldsymbol{X})$ 是空集.

感谢观看！