概统 Assignment 13 & 14 答案

李屹

2022 年 12 月 23 日

A13Q1. Y_2 为 $X_{7,8,9}$ 的均值，故 $S^2 \sim \dfrac{\sigma^2}{2}\chi^2(2)$，且 S^2 与 Y_2 相互独立，自然也与 $Y_1 - Y_2$ 相互独立，又有 $Y_1 \sim N(\mu, \dfrac{\sigma^2}{6})$，$Y_2 \sim N(\mu, \dfrac{\sigma^2}{3})$，$Y_1 - Y_2 \sim \dfrac{\sigma}{\sqrt{2}}N(0,1)$，由定义得 $Z \sim t(2)$。■

A13Q2.

1. $\left|\dfrac{\partial(X_1, X_2)}{\partial(X_1+X_2, X_1-X_2)}\right| = \dfrac{1}{2}$，且 (X_1, X_2) 服从二维正态分布，因此 (X_1+X_2, X_1-X_2) 也服从二维正态分布，$Cov(X_1+X_2, X_1-X_2) = E((X_1+X_2)(X_1-X_2)) - E(X_1+X_2)E(X_1-X_2) = E(X_1^2) - E(X_2^2) = 0$，可得 X_1+X_2 与 X_1-X_2 相互独立。

2. (X_1+X_2)，$(X_1-X_2) \sim N(0, 2\sigma^2)$，两者独立同分布，由定义得 $(X_1+X_2)^2$，$(X_1-X_2)^2 \sim \sqrt{2}\sigma\chi^2(1)$，仍然独立同分布，由定义得 $\dfrac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$。

■

A13Q3.

1. $E(X) = \displaystyle\int_0^1 (\theta+1)x^{\theta+1}\,\mathrm{d}x = \frac{\theta+1}{\theta+2} = \bar{X}$，故矩估计为 $\hat{\theta} = \dfrac{1}{1-\bar{X}} - 2$。$\ln L(\theta) = n\ln(\theta+1) + \theta(\displaystyle\sum_{i=1}^{n}\ln X_i)$，有唯一最大值，解 $\dfrac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta) = 0$ 得极大似然估计 $\hat{\theta} = -\dfrac{n}{\displaystyle\sum_{i=1}^{n}\ln X_i} - 1$。

2. $E(X) = \displaystyle\int_\theta^\infty 2x\mathrm{e}^{-2(x-\theta)}\,\mathrm{d}x = \theta + \frac{1}{2} = \bar{X}$，矩估计为 $\hat{\theta} = \bar{X} - \dfrac{1}{2}$。$\ln L(\theta) = n(\ln 2 + 2\theta) - 2(\displaystyle\sum_{i=1}^{n} X_i)$，随 θ 增加单调递增，由于 $x \geq \theta$，有 $\theta \leq \min(X_i)$，极大似然估计为 $\hat{\theta} = \min(X_i)$。

3. 为满足归一性条件，只可能是 $x > \theta_1 \geq \theta_2 > 0$。
   $E(X) = \int_{\theta_1}^{\infty} \frac{x}{\theta_2}\mathrm{e}^{-\frac{x-\theta_1}{\theta_2}}\,\mathrm{d}x = \theta_1 + \theta_2$，$E(X^2) = \int_{\theta_1}^{\infty} \frac{x^2}{\theta_2}\mathrm{e}^{-\frac{x-\theta_1}{\theta_2}}\,\mathrm{d}x = \theta_1^2 + 2\theta_1\theta_2 + 2\theta_2^2$，解得 $\hat{\theta}_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$，$\hat{\theta}_1 = \bar{X} - \hat{\theta}_2$。上面解仅在 $\hat{\theta}_2 \leq \frac{1}{2}\bar{X}$ 时满足归一性条件，在条件不满足时，严格来说应当属于不存在矩估计值，实践中可以考虑基于不等式 $\theta_1 \geq \theta_2$ 适当放弃高阶矩得近似解 $\hat{\theta}_1 = \hat{\theta}_2 = \frac{1}{2}\bar{X}$。$\ln L(\theta_1, \theta_2) = n(\frac{\theta_1 - \bar{X}}{\theta_2} - \ln\theta_2)$，随 θ_1 单调递增，而 $x > \theta_1$ 给出 $\theta_1 < \min(X_i)$，是开区间，因此极大似然估计值不存在。

4. $E(X) = \frac{a+b}{2}$，$E(X^2) = \frac{a^2+b^2+ab}{3}$，解得矩估计为 $\hat{a} = \bar{X} - \sqrt{3}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$，$\hat{b} = \bar{X} + \sqrt{3}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$。在 $\{X_i\}$ 只有一个值时，可以认为矩估计不存在。$\ln L(a, b) = -n\ln(b-a)$，随 b 增大单调增加，随 a 增大单调减小，由 $a \leq x \leq b$ 可得极大似然估计 $\hat{a} = \min(X_i)$，$\hat{b} = \max(X_i)$。在 $\{X_i\}$ 只有一个值时，可以认为极大似然估计不存在。

■

A13Q4. 由正态分布的极大似然估计得 $\hat{\mu} = \bar{X}$，$P(X \leq 3)$ 的极大似然估计为 $\Phi(3 - \bar{X})$。■

A13Q5. $E(\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2) = \sum_{i=1}^{n-1}E((X_{i+1} - X_i)^2) = 2(n-1)\sigma^2$，因此 $c = \frac{1}{2(n-1)}$。■

A14Q1. 无偏性要求 $E(a\hat{\theta}_1 + b\hat{\theta}_2) = (a+b)\theta = \theta$，故 $a + b = 1$。设 $D(\hat{\theta}_2) = S$，$D(a\hat{\theta}_1 + b\hat{\theta}_2) = (2a^2 + b^2)S$，取 $a = \frac{1}{3}$，$b = \frac{2}{3}$ 时有最小值 $\frac{2}{3}S$。■

A14Q2.

由均匀分布得 $E(X_i) = \theta + \frac{1}{2}$，$E(X_{(n)}) = \theta + \frac{n}{n+1}$，故 $E(\hat{\theta}_1) = \theta$，$E(\hat{\theta}_2) = \theta$，即都是无偏估计。

注意到均匀分布中 $X_{(n)}$ 的概率密度为 $f(x) = n(x-\theta)^{n-1}I(\theta, \theta+1)$，有 $D(X_i) = \frac{1}{12}$，$E(X_{(n)}^2) = \int_{\theta}^{\theta+1} x^2 n(x-\theta)^{n-1}\,\mathrm{d}x = n(\frac{1}{n+2} + \frac{2\theta}{n+1} + \frac{\theta^2}{n})$，$D(X_{(n)}) = \frac{n}{(n+1)^2(n+2)}$，故 $D(\hat{\theta}_1) = \frac{1}{n^2}\sum D(X_i) = \frac{1}{12n}$，$D(\hat{\theta}_2) = \frac{n}{(n+1)^2(n+2)} < \frac{1}{(n+1)^2}$，当 $n \geq 10$ 时有 $D(\hat{\theta}_2) < D(\hat{\theta}_1)$，逐项代入计

算得 $n \geq 8$ 即可保证 $D(\hat{\theta}_2) < D(\hat{\theta}_1)$，即 $n \geq 8$ 时 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效。有效性的定义中并未对 n 加以限制，上述结果恐怕不能证明 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效？ ■

A14Q3.

- 易得 $\ln L(\theta) = -\sum(X_i - \theta) = n(\theta - \bar{X})$，随 θ 增加而增加，由条件 $X_i \geq \theta$ 可得极大似然估计 $\hat{\theta} = \min(X_i)$（或写作 $X_{(1)}$）。

  $$F(x) = \int_\theta^x \mathrm{e}^{-(x-\theta)}\,\mathrm{d}x = 1 - \mathrm{e}^{-(x-\theta)}(x \geq \theta)$$
  $$f(x_{(1)}) = n(1 - F(x_{(1)}))^{n-1} f(x_{(1)}) = n\mathrm{e}^{-n(x_{(1)}-\theta)}$$
  $$E(X_{(1)}) = \int_\theta^\infty n x_{(1)} \mathrm{e}^{-n(x_{(1)}-\theta)}\,\mathrm{d}x_{(1)} = \theta + \frac{1}{n}$$

  因此 $\hat{\theta}_1$ 满足 $E(\hat{\theta}_1) = \theta$，是 θ 的无偏估计。

- $E(X) = \theta + 1$，$\hat{\theta}_2 = \bar{X} - 1$。$E(\hat{\theta}_2) = \theta$，是无偏估计。

- $$D(\hat{\theta}_1) = D(X_{(1)}) = \int_\theta^\infty n(x_{(1)} - \theta - \frac{1}{n})^2 \mathrm{e}^{-n(x_{(1)}-\theta)}\,\mathrm{d}x_{(1)} = \frac{1}{n^2}$$
  $$D(\hat{\theta}_2) = D(\bar{X}) = \frac{1}{n} D(X) = \frac{1}{n} \int_\theta^\infty (x - \theta - 1)^2 \mathrm{e}^{-(x-\theta)}\,\mathrm{d}x = \frac{1}{n}$$

  由切比雪夫不等式，$P(|\hat{\theta}_1 - \theta| > \epsilon) \leq \frac{1}{n^2\epsilon^2} \to 0$，由定义 $\hat{\theta}_1$ 是相合估计。同理可得 $\hat{\theta}_2$ 也是相合估计。

  $\hat{\theta}_2$ 的另一个做法：由于 $E(X)$ 存在，由辛钦大数定律 $\bar{X}$ 依概率收敛于 $E(X)$，故 $\hat{\theta}_2 = \bar{X} - 1$ 是相合估计。

  $D(\hat{\theta}_1) < D(\hat{\theta}_2)$，$\hat{\theta}_1$ 更有效。

■

A14Q4.

1. $E(X) = \frac{3}{4}\theta$，记 $y = \max(x_1, x_2)$，有 $f(y) = \frac{6y^5}{\theta^6}$，$E(Y) = \frac{6}{7}\theta$。$E(T_1) = \frac{4}{3}E(X) = \theta$，$E(T_2) = \frac{7}{6}E(Y) = \theta$，所以两者都是无偏估计。
2. $E(X^2) = \frac{3}{5}\theta^2$，$E(T_1^2) = \frac{31}{30}\theta^2$，$E((T_1 - \theta)^2) = E(T_1^2) - \theta^2 = \frac{1}{30}\theta^2$。$E(Y^2) = \frac{3}{4}\theta^2$，$E((T_2 - \theta)^2) = \frac{49}{36}E(Y^2) - \theta^2 = \frac{1}{48}\theta^2$。可见 T_2 的均方误差更小。

3. $E(T_c)=\dfrac{6}{7}c\theta$，$E(T_c^2)=\dfrac{3}{4}c^2\theta^2$，$MSE((T_c-\theta)^2)=E(T_c^2)-2\theta E(T_c)+\theta^2=(\dfrac{3}{4}c^2-\dfrac{12}{7}c+1)\theta^2$，$c=\dfrac{8}{7}$ 时取最小值。

■

A14Q5. $\bar{X}\sim N(\mu,\dfrac{\sigma}{\sqrt{n}})$，要求有 $-z_{0.025}<\dfrac{\bar{X}-\mu}{0.025/\sqrt{200}}<z_{0.025}$，查表得 $z_{0.025}=1.96$，μ 的置信区间为 $(0.0775,0.0845)$。■

A14Q6. 取枢轴量为 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，置信度为 $1-\alpha$ 的置信区间的长度为 $\dfrac{2\sigma}{\sqrt{n}}z_{\alpha/2}\leq L$，$n$ 应当满足 $n\geq\left(\dfrac{2\sigma}{L}z_{\alpha/2}\right)^2$。■