统计学习题03参考解答

1. (a)

由 X_{i+1} 与 X_i 独立，可知

$$X_{i+1}-X_i \sim N(0, 2\sigma^2).$$

因此

$$\begin{aligned}
&E(X_{i+1}-X_i)^2 \\
&= Var(X_{i+1}-X_i) + [E(X_{i+1}-X_i)]^2 \\
&= 2\sigma^2 + 0 \\
&= 2\sigma^2.
\end{aligned}$$

从而有

$$\begin{aligned}
&E\left(\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\right) \\
&= \sum_{i=1}^{n-1} E(X_{i+1}-X_i)^2 \\
&= \sum_{i=1}^{n-1} 2\sigma^2 \\
&= 2(n-1)\sigma^2.
\end{aligned}$$

为使 $c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 成为 σ^2 的无偏估计，应有

$$E\, c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2 = \sigma^2$$

即应有

$$\sigma^2 = c\,\mathbb{E}\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2 = 2c(n-1)\sigma^2.$$

从而可取

$$c = \frac{1}{2(n-1)}.$$

(b) 由于 $X_i-\bar{X}$ 是独立同分布正态随机变量 $X_1, X_2, \cdots, X_n$ 的线性组合：

$$X_i-\bar{X} = \left(1-\frac{1}{n}\right)X_i - \frac{1}{n}\sum_{j\neq i}X_j, \qquad (1)$$

所以 $X_i-\bar{X}$ 服从正态分布，又有

$$\mathbb{E}(X_i-\bar{X}) = \mathbb{E}X_i - \mathbb{E}\bar{X} = \mu-\mu=0$$

且由(1)可知

$$\begin{aligned}
&Var(X_i-\bar{X}) \\
&= \left(1-\frac{1}{n}\right)^2 Var(X_i) + \frac{1}{n^2}\sum_{j\neq i} Var(X_j) \\
&= \left(1-\frac{1}{n}\right)^2\sigma^2 + \frac{n-1}{n^2}\sigma^2 \\
&= \frac{n-1}{n}\sigma^2.
\end{aligned}$$

[注：$X_i-\bar{X}$ 的方差或用如下方法直接计算.

$$Var(X_i - \bar{X})$$

$$= Var X_i + Var \bar{X} - 2cov(X_i, \bar{X})$$

$$= \sigma^2 + \frac{1}{n}\sigma^2 - \frac{2}{n}\sum_{j=1}^{n} cov(X_i, X_j)$$

$$= \sigma^2 + \frac{1}{n}\sigma^2 - \frac{2}{n} cov(X_i, X_i)$$

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

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

其中用到当 $i \neq j$ 时，X_i 与 X_j 独立，从而

$$cov(X_i, X_j) = 0.$$]

因此

$$X_i - \bar{X} \sim N\left(0, \frac{n-1}{n}\sigma^2\right).$$

所以

$$\xi_i := \frac{X_i - \bar{X}}{\sqrt{\frac{n-1}{n}\sigma^2}} \sim N(0,1), \quad i = 1, 2, \cdots, n.$$

注意到如果 $\xi \sim N(0,1)$，则有

$$\mathbb{E}|\xi| = \sqrt{\frac{2}{\pi}} \tag{2}.$$

事实上，

$$E|\xi| = \int_{-\infty}^{+\infty} |\xi| \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) d\xi$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \xi \exp\left(-\frac{1}{2}\xi^2\right) d\xi$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} -d\left(\exp\left(-\frac{1}{2}\xi^2\right)\right)$$

$$= \sqrt{\frac{2}{\pi}} .$$

因此

$$E|X_i-\bar{X}| = E\left(\sqrt{\frac{n-1}{n}}\sigma|\xi_i|\right) = \sqrt{\frac{2(n-1)}{n\pi}}\sigma.$$

为使 $k\sum_{i=1}^{n}|X_i-\bar{X}|$ 成为 σ 的无偏估计，应有

$$\sigma = E\, k\sum_{i=1}^{n}|X_i-\bar{X}|$$

$$= k\sum_{i=1}^{n} E|X_i-\bar{X}|$$

$$= k\cdot n\cdot \sqrt{\frac{2(n-1)}{n\pi}}\sigma$$

$$= k\cdot \sqrt{\frac{2n(n-1)}{\pi}}\sigma .$$

从而可取

$$k = \sqrt{\frac{\pi}{2n(n-1)}} .$$

3. 由于 $Var\,\hat{\theta} > 0$，可见

$$\mathbb{E}(\hat{\theta})^2 = Var(\hat{\theta}) + (\mathbb{E}\hat{\theta})^2$$

$$= Var(\hat{\theta}) + \theta^2 > \theta^2$$

即 $\hat{\theta}^2$ 不是 θ^2 的无偏估计.

4. 参数为θ时样本的联合分布密度为

$$f(X|\theta) = \mathbb{P}(X=1|\theta)\cdot \mathbb{P}(X=2|\theta)\cdot \mathbb{P}(X=1|\theta)$$

$$= \theta\cdot(1-\theta)\cdot\theta$$

$$= \theta^2(1-\theta),$$

参数θ当作随机变量Θ的先验分布密度为

$$f_{\Theta}(\theta) = \mathbb{1}_{[0,1]}(\theta),$$

故Θ的后验密度为

$$f_{\Theta|X}(\theta|x) \propto f(X|\theta)\,f_{\Theta}(\theta) = \theta^2(1-\theta)\mathbb{1}_{[0,1]}(\theta).$$

因为

$$\int_{-\infty}^{+\infty}\theta^2(1-\theta)\mathbb{1}_{[0,1]}(\theta)\,d\theta = \int_0^1(\theta^2-\theta^3)\,d\theta = \frac{1}{3}-\frac{1}{4} = \frac{1}{12}.$$

故Θ的后验密度为 $12\theta^2(1-\theta)\mathbb{1}_{[0,1]}(\theta)$.