Математическая статистика.
Домашнее задание №11

Lev Khoroshansky

Задача №1(C).

По условию, $\mathbb{P}(X_i = k) = \dfrac{\theta^k}{k!}\, e^{-\theta}$ или же $\mathbb{P}(X_i) = \dfrac{\theta^{X_i}}{X_i!}\, e^{-\theta}$, откуда

$$L(\theta) = \ln\left(\frac{\theta^{X_1}}{X_1!}\, e^{-\theta} \dots \frac{\theta^{X_n}}{X_n!}\, e^{-\theta}\right) = (X_1 + \dots + X_n)\ln(\theta) - n \cdot \theta + C,$$

где $C = \mathrm{const}$ – логарифмы факториалов. Найдём частную производную и приравняем её к нулю:

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{X_1 + \dots + X_n}{\theta} - n = 0 \implies \widehat{\theta} = \overline{X}$$

Задача №2(C).

Мы знаем, что у распределения Пуассона $\mathbb{E}X_i = \mathbb{D}X_i = \theta$. Посчитаем дисперсию $\widehat{\theta}$ из предыдущей задачи:

$$\mathbb{D}\widehat{\theta} = \mathbb{D}\overline{X} = \frac{1}{n}\,\mathbb{D}X_i = \frac{\theta}{n}$$

Найдём информацию Фишера:

$$\begin{aligned} f(\theta) &= \ln[\mathbb{P}(X_i)] = X_i \ln(\theta) - \theta + C \\ f'_\theta(\theta) &= \frac{\partial f(\theta)}{\partial \theta} = \frac{X_i}{\theta} - 1 \\ I(\theta) &= \mathbb{E}[f'_\theta(\theta)]^2 = \frac{1}{\theta^2}\,\mathbb{D}X_1 = \frac{1}{\theta} \end{aligned}$$

Проверим эффективность:

$$\frac{\theta}{n} \geq \frac{1}{n \cdot \frac{1}{\theta}} = \frac{\theta}{n} \text{ – равенство выполняется.}$$

Следовательно, данная оценка является эффективной (несмещённость очевидна).

Задача №3.

Делаем всё то же самое, только для плотностей:

(a) $\mathcal{N}(\mu, \sigma^2) = \mathcal{N}(\theta, 1) \implies \rho_{X_i}(t_i) = \dfrac{1}{\sqrt{2\pi}\cdot 1}\,\exp\left(-\dfrac{(t_i-\theta)^2}{2\cdot 1^2}\right) = \dfrac{1}{\sqrt{2\pi}}\,\exp\left(-\dfrac{(t_i-\theta)^2}{2}\right)$

$$L(\theta) = \ln\left[(2\pi)^{-n/2}\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(t_i-\theta)^2\right)\right] = -\frac{1}{2}\sum_{i=1}^{n}(t_i-\theta)^2 + C$$
$$\frac{\partial L(\theta)}{\partial\theta} = \sum_{i=1}^{n}(\theta - t_i) = n\cdot\theta - (t_1+\dots+t_n) = 0 \implies \widehat{\theta} = \overline{t} = \overline{X}$$

(b) $\mathcal{N}(\mu, \sigma^2) = \mathcal{N}(\theta, 2\theta) \implies \rho_{X_i}(t_i) = \dfrac{1}{\sqrt{2\pi}\cdot\sqrt{2\theta}}\,\exp\left(-\dfrac{(t_i-\theta)^2}{2\cdot 2\theta}\right) = \dfrac{1}{2\sqrt{\pi\theta}}\,\exp\left(-\dfrac{(t_i-\theta)^2}{4\theta}\right)$

$$L(\theta) = \ln\left[(2\sqrt{\pi})^{-n}\cdot\theta^{-n/2}\cdot\exp\left(-\frac{1}{4\theta}\sum_{i=1}^{n}(t_i-\theta)^2\right)\right] = -\frac{n}{2}\ln(\theta) - \frac{1}{4\theta}\sum_{i=1}^{n}(t_i-\theta)^2 + C$$
$$\frac{\partial L(\theta)}{\partial\theta} = -\frac{n}{2\theta} + \frac{1}{4\theta^2}\sum_{i=1}^{n}\left(t_i^2-\theta^2\right) = \frac{1}{4\theta^2}\sum_{i=1}^{n}t_i^2 - \frac{n}{2\theta} - \frac{n}{4} = 0$$

Обозначим $\dfrac{1}{n}\sum\limits_{i=1}^{n}t_i^2 = \overline{T^2}$, тогда

$$\frac{n}{4\theta^2}\overline{T^2} - \frac{n}{2\theta} - \frac{n}{4} = 0 \;\to\; \theta^2 + 2\theta - \overline{T^2} = 0 \implies \widehat{\theta} = \sqrt{1+\overline{T^2}} - 1 = \sqrt{1+\overline{X^2}} - 1$$

(c) $\rho_{X_i}(t_i) = \dfrac{1}{\theta}\,\exp\left(-\dfrac{1}{\theta}\,t_i\right)$

$$L(\theta) = \ln\left[\theta^{-n}\cdot\exp\left(-\frac{1}{\theta}\sum_{i=1}^{n}t_i\right)\right] = -n\ln(\theta) - \frac{1}{\theta}\sum_{i=1}^{n}t_i$$
$$\frac{\partial L(\theta)}{\partial\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}t_i = 0 \implies \widehat{\theta} = \overline{t} = \overline{X}$$

Задача №8.

По условию, $\mathbb{P}(X_i = k) = \dfrac{1}{\theta}$ или же $\mathbb{P}(X_i) = \dfrac{1}{\theta}$, но это справедливо только при $X_i \le \theta$, тогда

$$L(\theta) = -n\ln(\theta)\cdot\mathrm{Ind}(\forall i\;\; X_i \le \theta)$$
$$\frac{\partial L(\theta)}{\partial\theta} = -\frac{n}{\theta} < 0 \quad \forall\theta \ge 1 \text{ (если не учитывать индикаторы)}$$

Видно, что свой максимум $L(\theta)$ достигает в самой левой точке, но это справедливо только при том условии, что $\theta \ge X_{(n)}$, откуда получаем $\widehat{\theta} = X_{(n)}$, что более-менее логично.

Задача №4.

Обозначим $\mathbb{D}X_i = \theta^2 = A$ и найдём $\mathbb{D}\Theta$:

$$\mathbb{D}\Theta = \frac{1}{n^2}\,\mathbb{D}(X_1^2 + \dots + X_n^2) = \frac{1}{n}\,\mathbb{D}X_1^2 = \frac{1}{n}\,\mathbb{E}X_1^4 - \frac{1}{n}\,(\mathbb{E}X_1^2)^2$$

Найдём $\mathbb{E}X_1^4$: для начала заметим, что $\dfrac{X_1}{\sqrt{A}} = Y \sim \mathcal{N}(0;1)$, откуда $\mathbb{E}X_1^4 = A^2 \cdot \mathbb{E}Y^4$.
Теперь задача свелась к поиску 4 момента у стандартной нормальной величины:

$$\varphi_Y(y) = \exp\left(-\frac{y^2}{2}\right), \quad \varphi_Y^{(k)}(0) = (i)^k \cdot \mathbb{E}Y^k \implies \varphi_Y^{(4)}(0) = \mathbb{E}Y^4$$

Разложим характеристическую функцию в ряд Тейлора в окрестности нуля:

$$\varphi_Y(y) = 1 - \frac{1}{2}\,y^2 + \frac{1}{8}\,y^4 + \dots$$

Откуда $\dfrac{1}{8} = \dfrac{1}{4!}\,\mathbb{E}Y^4 \implies \mathbb{E}Y^4 = 3$, следовательно, $\mathbb{E}X_1^4 = 3A^2$, тогда

$$\mathbb{D}\Theta = \frac{1}{n}\,(3A^2 - A^2) = \frac{2}{n}\,A^2$$

Далее нам нужно найти информацию Фишера:

$$\rho_{X_1}(t) = \frac{1}{\sqrt{2\pi A}}\,\exp\left(-\frac{t^2}{2A}\right) \implies L_1(A) = \ln[\rho_{X_1}(t)] = -\frac{1}{2}\,\ln(A) - \frac{t^2}{2A} + C \implies \frac{\partial L_1(A)}{\partial A} = -\frac{1}{2A} + \frac{t^2}{2A^2}$$

$$I(A) = \mathbb{E}\left[-\frac{1}{2A} + \frac{t^2}{2A^2}\right]^2 = \frac{1}{4A^4}\,\mathbb{E}[t^2 - A]^2 = [t \leftrightarrow X_1] = \frac{1}{4A^4}\,\mathbb{D}X_1^2 = \frac{1}{2A^2}$$

Проверяем неравенство Рао-Крамера:

$$\frac{2}{n}\,A^2 \geq \frac{1}{n \cdot \dfrac{1}{2A^2}} = \frac{2}{n}\,A^2$$

Равенство достигается, откуда следует эффективность оценки (опять же, несмещённость очевидна из-за нулевого ожидания).

Задача №10.

По условию, $\rho_X(t) = \dfrac{1}{2}\,\exp(-|t-\theta|)$, тогда

$$L(\theta) = \ln\left[\frac{1}{2^n}\,\exp\left(-\sum_{i=1}^{n}|t_i - \theta|\right)\right] = -\sum_{i=1}^{n}|t_i - \theta| + C$$

$$\frac{\partial L(\theta)}{\partial \theta} = -\sum_{i=1}^{n}\frac{t_i - \theta}{|t_i - \theta|} = -\sum_{i=1}^{n}\operatorname{sgn}(t_i - \theta)$$

Если $n = 2k+1$, то подойдёт $\widehat{\theta} = X_{(k+1)}$, так как в половине случаев будет -1, при равенстве 0, а в остальной половине $+1$, что даст 0 в сумме.

Если $n = 2k$, то мы можем либо взять $\widehat{\theta} = X_{(k)}$, либо $\widehat{\theta} = X_{(k+1)}$.