Решения задач по математической статистике

КОНСТАНТИН ЗЮБИН

Задача 1 («нулёвка»)

Пусть случайные величины $X_1, \dots, X_n \sim \mathcal{N}(\theta, 1)$ независимы и одинаково распределены. Вычислим информация Фишера. Имеем формулу для плотности $f_\theta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}$. Тогда

$$L(x_1, \dots, x_n, \theta) = f_\theta(x_1, \dots, x_n) = \prod_{j=1}^{n} f_\theta(x_j) = (2\pi)^{\frac{-n}{2}} \cdot e^{-\frac{1}{2}\sum\limits_{j=1}^{n}(x_j-\theta)^2}.$$

Далее,

$$\tfrac{\partial}{\partial\theta} \ln L(x_1, \dots, x_n, \theta)\big|_\theta = -\tfrac{1}{2} \cdot 2 \cdot \sum_{i=1}^{n} (\theta - x_j) = -n\theta + \sum_{i=1}^{n} x_j.$$

Окончательно,

$$\mathrm{E}_\theta \left(-n\theta + \sum_{i=1}^{n} X_j \right)^2 = n^2\theta^2 - 2n\theta \cdot \mathrm{E}_\theta \left(\sum_{i=1}^{n} X_j \right) + \mathrm{E}_\theta \left(\sum_{i=1}^{n} X_j \right)^2 =$$

$$= n^2\theta^2 - 2n\theta \cdot n\theta + n^2\theta^2 + n = n.$$

Задача 2 ($\mathcal{N}(\theta, \theta^2)$)

Вычислим информацию Фишера для независимых одинаково распределённых случайных величин $X_1, \dots, X_n \sim \mathcal{N}(\theta, \theta^2)$. Из вычислений выше имеем

$$\tfrac{\partial}{\partial\theta} \ln L(x_1, \dots, x_n, \theta)\big|_\theta = -\tfrac{n}{\theta} - \tfrac{x_1 + \ldots + x_n}{\theta^2} + \tfrac{x_1^2 + \ldots + x_n^2}{\theta^3}.$$

Тогда

$$\tfrac{\partial^2}{\partial^2\theta} \ln L(x_1, \dots, x_n, \theta)\Big|_\theta = \tfrac{n}{\theta^2} + 2 \cdot \tfrac{x_1 + \ldots + x_n}{\theta^3} - 3 \cdot \tfrac{x_1^2 + \ldots + x_n^2}{\theta^4}.$$

Отсюда

$$\mathrm{E}_\theta \left(\tfrac{\partial}{\partial\theta} \ln L(X_1, \dots, X_n, \theta)\big|_\theta \right)^2 = \mathrm{E}_\theta \left(- \tfrac{\partial^2}{\partial^2\theta} \ln L(X_1, \dots, X_n, \theta)\Big|_\theta \right) = \mathrm{E}_\theta(-\tfrac{n}{\theta^2} - 2 \cdot \tfrac{X_1 + \ldots + X_n}{\theta^3} + 3 \cdot \tfrac{X_1^2 + \ldots + X_n^2}{\theta^4}) =$$

$$= -\tfrac{n}{\theta^2} - \tfrac{2n\theta}{\theta^3} + \tfrac{3n(\theta^2+\theta^2)}{\theta^4} = \tfrac{3n}{\theta^2}.$$

Задача 2 ($\mathcal{N}(0, \theta)$)

Вычислим информацию Фишера для независимых одинаково распределённых случайных величин $X_1, \dots, X_n \sim \mathcal{N}(0, \theta)$. Из вычислений выше имеем

$$\tfrac{\partial}{\partial\theta} \ln L(x_1, \dots, x_n, \theta)\big|_\theta = -\tfrac{n}{2\theta} + \tfrac{x_1^2 + \ldots + x_n^2}{2\theta^2}.$$

Тогда

$$\tfrac{\partial^2}{\partial^2\theta} \ln L(x_1, \dots, x_n, \theta)\Big|_\theta = \tfrac{n}{2\theta^2} - \tfrac{x_1^2 + \ldots + x_n^2}{\theta^3}.$$

Отсюда

$$\mathrm{E}_\theta \left(\tfrac{\partial}{\partial\theta} \ln L(X_1, \dots, X_n, \theta)\big|_\theta \right)^2 = \mathrm{E}_\theta \left(- \tfrac{\partial^2}{\partial^2\theta} \ln L(X_1, \dots, X_n, \theta)\Big|_\theta \right) = \mathrm{E}_\theta(-\tfrac{n}{2\theta^2} + \tfrac{x_1^2 + \ldots + x_n^2}{\theta^3}) =$$

$$= -\tfrac{n}{2\theta^2} + \tfrac{n\theta}{\theta^3} = \tfrac{n}{2\theta^2}.$$

Задача 2 ($\mathcal{N}(\theta,\theta)$)

Вычислим информацию Фишера для независимых одинаково распределённых случайных величин $X_1,\ldots,X_n \sim \mathcal{N}(\theta,\theta)$. Из вычислений выше имеем

$$\tfrac{\partial}{\partial\theta}\ln L(x_1,\ldots,x_n,\theta)\big|_\theta = \tfrac{-n}{2} + \tfrac{-n}{2\theta} + \tfrac{x_1^2+\ldots+x_n^2}{2\theta^2}.$$

Тогда

$$\tfrac{\partial^2}{\partial^2\theta}\ln L(x_1,\ldots,x_n,\theta)\Big|_\theta = \tfrac{n}{2\theta^2} - \tfrac{x_1^2+\ldots+x_n^2}{\theta^3}.$$

Отсюда

$$\mathrm{E}_\theta\left(\tfrac{\partial}{\partial\theta}\ln L(X_1,\ldots,X_n,\theta)\big|_\theta\right)^2 = \mathrm{E}_\theta\left(-\tfrac{\partial^2}{\partial^2\theta}\ln L(X_1,\ldots,X_n,\theta)\Big|_\theta\right) = \mathrm{E}_\theta(-\tfrac{n}{2\theta^2} + \tfrac{x_1^2+\ldots+x_n^2}{\theta^3}) =$$

$$= -\tfrac{n}{2\theta^2} + \tfrac{n\theta}{\theta^3} = \tfrac{n}{2\theta^2}.$$

Задача 3

Пусть случайные величины $X_1,\ldots,X_n \sim \mathrm{Cauchy}(\theta)$ независимы и имеют распределение Коши. Плотность имеет вид $f_\theta(x) = \frac{1}{\pi((x-\theta)^2+1)}$. Вычислим информацию Фишера. Имеем

$$L(x_1,\ldots,x_n,\theta) = f_\theta(x_1,\ldots,x_n) = \prod_{j=1}^n f_\theta(x_j) = \tfrac{1}{\pi^n}\cdot\tfrac{1}{\prod_{j=1}^n((x_j-\theta)^2+1)}.$$

Далее,

$$\tfrac{\partial}{\partial\theta}\ln L(x_1,\ldots,x_n,\theta)\big|_\theta = \tfrac{\partial}{\partial\theta}\left(-n\ln\pi - \sum_{j=1}^n \ln((x_j-\theta)^2+1)\right)\Bigg|_\theta = \sum_{j=1}^n \tfrac{2(x_j-\theta)}{(x_j-\theta)^2+1}.$$

Имеем

$$\mathrm{E}_\theta\,\tfrac{2(X_j-\theta)}{(X_j-\theta)^2+1} = \int\limits_{-\infty}^{+\infty}\tfrac{2(x_j-\theta)dx_j}{\pi((x_j-\theta)^2+1)^2} = 0,$$

поскольку подынтегральная функция после сдвига на θ становится нечётной. Из независимости для $j\neq k$ получаем

$$\mathrm{E}_\theta\,\tfrac{2(X_j-\theta)}{(X_j-\theta)^2+1}\cdot\tfrac{2(X_k-\theta)}{(X_k-\theta)^2+1} = \mathrm{E}_\theta\,\tfrac{2(X_j-\theta)}{(X_j-\theta)^2+1}\cdot\mathrm{E}_\theta\,\tfrac{2(X_k-\theta)}{(X_k-\theta)^2+1} = 0\cdot 0 = 0.$$

Также получаем

$$\mathrm{E}_\theta\left(\tfrac{2(X_j-\theta)}{(X_j-\theta)^2+1}\right)^2 = \int\limits_{-\infty}^{+\infty}\tfrac{4(x_j-\theta)^2dx_j}{\pi((x_j-\theta)^2+1)^3} = \int\limits_{-\infty}^{+\infty}\tfrac{4x_j^2dx_j}{\pi(x_j^2+1)^3} = \tfrac{8}{\pi}\int\limits_0^{+\infty}\tfrac{x_j^2dx_j}{(x_j^2+1)^3} =$$

$$= \{x_j = \operatorname{tg}\varphi\} = \tfrac{8}{\pi}\int\limits_0^{\frac{\pi}{2}}\sin^2\varphi\cos^{-2+6-2}\varphi\;d\varphi = \tfrac{8}{\pi}\int\limits_0^{\frac{\pi}{2}}\sin^2\varphi\cos^2\varphi\;d\varphi = \tfrac{2}{\pi}\int\limits_0^{\frac{\pi}{2}}\sin^2(2\varphi)\;d\varphi =$$

$$= \tfrac{2}{\pi}\int\limits_0^{\frac{\pi}{2}}\tfrac{1-\cos(4\pi)}{2}d\varphi = \tfrac{2}{\pi}\cdot\tfrac{\pi}{4} = \tfrac{1}{2}.$$

Введём обозначение $g(x,\theta)=\frac{2(x-\theta)}{(x-\theta)^2+1}$. Тогда

$$\mathrm{E}_\theta\left(\sum_{j=1}^n g(X_j,\theta)\right)^2=\sum_{j=1}^n \mathrm{E}_\theta\, g(X_j,\theta)^2+2\sum_{j<k}\mathrm{E}_\theta\, g(X_j,\theta)g(X_k,\theta)=\tfrac{n}{2}+0=\tfrac{n}{2}.$$

Задача 4а

Рассмотрим случайные величины с распределением $\mathrm{E}(\frac{1}{\theta})$, где $\theta>0$. Имеем формулу для плотности $f_\theta(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}I(x\geqslant 0)$. Тогда

$$L(\theta,x_1,\ldots,x_n)=f_\theta(x_1,\ldots,x_n)=\prod_{j=1}^n f_\theta(x_j)=\tfrac{1}{\theta^n}\cdot e^{-\frac{\sum\limits_{j=1}^n x_j}{\theta}}$$

при $x_1,\ldots,x_n\geqslant 0$, и $f_\theta(x_1,\ldots,x_n)=0$, если хотя бы одна из координат отрицательна.

Вычислим информацию Фишера. Имеем

$$\tfrac{\partial}{\partial\theta}\ln L(x_1,\ldots,x_n,\theta)\big|_\theta=\tfrac{\partial}{\partial\theta}\left(-n\ln\theta-\tfrac{1}{\theta}\sum_{j=1}^n x_j\right)\Bigg|_\theta=-\tfrac{n}{\theta}+\tfrac{1}{\theta^2}\sum_{j=1}^n x_j=\sum_{j=1}^n\tfrac{x_j-\theta}{\theta^2}.$$

Обозначим через $g(x,\theta)=\frac{x-\theta}{\theta^2}$. Тогда (в последней формуле при $j\neq k$)

$$\mathrm{E}_\theta\, g(X_j,\theta)^2=\mathrm{E}_\theta\left(\tfrac{X_j^2}{\theta^4}-\tfrac{2X_j}{\theta^3}+\tfrac{1}{\theta^2}\right)=\tfrac{\theta^2+\theta^2}{\theta^4}-\tfrac{2\theta}{\theta^3}+\tfrac{1}{\theta^2}=\tfrac{1}{\theta^2};$$

$$\mathrm{E}_\theta\, g(X_j,\theta)=\mathrm{E}_\theta\,\tfrac{X_j-\theta}{\theta^2}=\tfrac{\theta-\theta}{\theta^2}=0;$$

$$\mathrm{E}_\theta\,(g(X_j,\theta)g(X_k,\theta))=\mathrm{E}_\theta\, g(X_j,\theta)\,\mathrm{E}_\theta\, g(X_k,\theta)=0\cdot 0=0.$$

Далее,

$$\mathrm{E}_\theta\left(\sum_{j=1}^n\tfrac{x_j-\theta}{\theta^2}\right)^2=\sum_{j=1}^n \mathrm{E}_\theta\, g(X_j,\theta)^2+2\sum_{j<k}\mathrm{E}_\theta\, g(X_j,\theta)g(X_k,\theta)=\tfrac{n}{\theta^2}+0=\tfrac{n}{\theta^2}.$$

Задача 4б

Рассмотрим случайные величины с распределением $\mathrm{Geom}(\theta)$, где $\theta\in(0,1)$. Имеем

$$L(\theta,x_1,\ldots,x_n)=\mathrm{P}_\theta(X_1=x_1,\ldots,X_n=x_n)=\prod_{j=1}^n\mathrm{P}(X_j=x_j)=\theta^n(1-\theta)^{x_1+\ldots+x_n-n}.$$

Вычислим информацию Фишера. Имеем

$$\tfrac{\partial}{\partial\theta}\ln L(x_1,\ldots,x_n,\theta)\big|_\theta=\tfrac{\partial}{\partial\theta}\left(n\ln\theta+(-n+\sum_{j=1}^n x_j)\ln(1-\theta)\right)\Bigg|_\theta=\tfrac{n}{\theta}-\tfrac{-n+\sum\limits_{j=1}^n x_j}{1-\theta}=\tfrac{n}{\theta(1-\theta)}-\sum_{j=1}^n\tfrac{x_j}{1-\theta}.$$

Обозначим через $g(x,\theta)=\frac{1-x\theta}{\theta(1-\theta)}$. Тогда (в последней формуле при $j\neq k$)

$$\mathrm{E}_\theta\, g(X_j,\theta)^2=\mathrm{E}_\theta\,\tfrac{1-2X_j\theta+X_j^2\theta^2}{\theta^2(1-\theta)^2}=\tfrac{1-2\frac{\theta}{\theta}+\frac{\theta^2(2-\theta)}{\theta^2}}{\theta^2(1-\theta)^2}=\tfrac{1-\theta}{\theta^2(1-\theta)^2}=\tfrac{1}{\theta^2(1-\theta)};$$

$$\mathrm{E}_\theta\, g(X_j,\theta)=\mathrm{E}_\theta\,\tfrac{1-X_j\theta}{\theta(1-\theta)}=\tfrac{1-\frac{\theta}{\theta}}{\theta^2(1-\theta)^2}=0;$$

$$\mathrm{E}_\theta\,(g(X_j,\theta)g(X_k,\theta))=\mathrm{E}_\theta\, g(X_j,\theta)\,\mathrm{E}_\theta\, g(X_k,\theta)=0\cdot 0=0.$$

Далее,

$$\mathrm{E}_\theta \left(\sum_{j=1}^{n} \tfrac{x_j-\theta}{\theta^2}\right)^2 = \sum_{j=1}^{n} \mathrm{E}_\theta\, g(X_j,\theta)^2 + 2\sum_{j<k} \mathrm{E}_\theta\, g(X_j,\theta) g(X_k,\theta) = \tfrac{n}{\theta^2(1-\theta)} + 0 = \tfrac{n}{\theta^2(1-\theta)}.$$

Задача 4в

Рассмотрим случайные величины с распределением $\Gamma(\alpha, \frac{1}{\theta})$, где $\theta > 0, \alpha \leqslant 1$. Имеем формулу для плотности $f_\theta(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} e^{-\frac{x}{\theta}} I(x \geqslant 0)$. Тогда

$$L(\theta, x_1, \ldots, x_n) = f_\theta(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_\theta(x_j) = \tfrac{(x_1 \ldots x_n)^{\alpha-1}}{\Gamma(\alpha)^n \theta^n} \cdot e^{-\frac{\sum\limits_{j=1}^{n\alpha} x_j}{\theta}}.$$

Вычислим информацию Фишера и оценку максимального правдоподобия. Имеем

$$\tfrac{\partial}{\partial\theta} \ln L(x_1, \ldots, x_n, \theta)\big|_\theta = \tfrac{\partial}{\partial\theta} \left((\alpha-1) \sum_{j=1}^{n} (\ln x_j) - n \ln \Gamma(\alpha) - n\alpha \ln\theta + -\tfrac{1}{\theta} \sum_{j=1}^{n} x_j \right)\Bigg|_\theta =$$

$$= -\tfrac{n\alpha}{\theta} + \tfrac{1}{\theta^2} \sum_{j=1}^{n} x_j = \tfrac{1}{\theta^2} \left(-n\alpha\theta + \sum_{j=1}^{n} x_j \right) = \sum_{j=1}^{n} \tfrac{x_j - \alpha\theta}{\theta^2}.$$

При $\theta = \overline{X}$ производная обращается в 0, при меньших значениях положительна, а при больших — отрицательна. Следовательно, в точке $\overline{X}$ достигается максимум и $\overline{X}$ — оценка максимального правдоподобия.

Вычислим матожидание и дисперсию случайной величины X_j.

$$\mathrm{E}_\theta\, X_j = \int\limits_0^{+\infty} \tfrac{u^\alpha}{\Gamma(\alpha)\theta^\alpha} e^{-\frac{u}{\theta}} du = \tfrac{\theta}{\Gamma(\alpha)} \int\limits_0^{+\infty} s^\alpha e^{-s} ds = \tfrac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \cdot \theta = \alpha\theta,$$

$$\mathrm{E}_\theta\, X_j^2 = \int\limits_0^{+\infty} \tfrac{u^{\alpha+1}}{\Gamma(\alpha)\theta^\alpha} e^{-\frac{u}{\theta}} du = \tfrac{\theta^2}{\Gamma(\alpha)} \int\limits_0^{+\infty} s^{\alpha+1} e^{-s} ds = \tfrac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \cdot \theta^2 = \alpha(\alpha+1)\theta^2.$$

Обозначим через $g(x, \theta) = \frac{x_j - \alpha\theta}{\theta^2}$. Тогда (в последней формуле при $j \neq k$)

$$\mathrm{E}_\theta\, g(X_j, \theta)^2 = \mathrm{E}_\theta\, \tfrac{X_j^2 - 2X_j\alpha\theta + \alpha^2\theta^2}{\theta^4} = \tfrac{\alpha(\alpha+1)\theta^2 - 2\alpha^2\theta^2 + \alpha^2\theta^2}{\theta^4} = \tfrac{\alpha}{\theta^2};$$

$$\mathrm{E}_\theta\, g(X_j, \theta) = \mathrm{E}_\theta\, \tfrac{X_j - \alpha\theta}{\theta^2} = \tfrac{\alpha\theta - \alpha\theta}{\theta^2} = 0;$$

$$\mathrm{E}_\theta \left(g(X_j, \theta) g(X_k, \theta) \right) = \mathrm{E}_\theta\, g(X_j, \theta)\, \mathrm{E}_\theta\, g(X_k, \theta) = 0 \cdot 0 = 0.$$

Далее,

$$\mathrm{E}_\theta \left(\sum_{j=1}^{n} \tfrac{x_j - \alpha\theta}{\theta^2}\right)^2 = \sum_{j=1}^{n} \mathrm{E}_\theta\, g(X_j,\theta)^2 + 2\sum_{j<k} \mathrm{E}_\theta\, g(X_j,\theta) g(X_k,\theta) = \tfrac{n\alpha}{\theta^2} + 0 = \tfrac{n\alpha}{\theta^2}.$$

Задача 5

Пусть случайные вектора $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \ldots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ независимы и таковы, что их компоненты X_i, Y_i независимы и распределены по закону $\mathrm{E}(\frac{1}{\theta_1})$ и $\mathrm{E}(\frac{1}{\theta_2})$, соответственно. Вычислим информационную матрицу. Поскольку плотности случайных X_i не зависят от θ_2, а плотности Y_i — от θ_1, то в информационная матрица будет диагональной, а на диагонали будут стоять информации Фишера для соответствующих компонент. Таким образом, матрица имеет вид

$$\begin{pmatrix} \frac{n}{\theta_1^2} & 0 \\ 0 & \frac{n}{\theta_2^2} \end{pmatrix}.$$