

Решения задач по математической статистике

КОНСТАНТИН ЗЮБИН

Задача 1а

Пусть случайные величины X_1, \dots, X_n имеют распределение $\text{Geom}(\theta)$. Найдём достаточную статистику. Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n P_\theta(X_i = x_i) = \prod_{i=1}^n \theta(1-\theta)^{x_i-1} I(x_i \in \mathbb{Z}_{\geq 1}) = \\ &= \theta^n (1-\theta)^{\sum_{i=1}^n x_i - n} \cdot I(x_1, \dots, x_n \in \mathbb{Z}_{\geq 1}). \end{aligned}$$

Положим $g(\theta, T) = \theta^n (1-\theta)^{T-n}$ и $h(X_1, \dots, X_n) = I(X_1, \dots, X_n \in \mathbb{Z}_{\geq 1})$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1 + \dots + x_n)h(x_1, \dots, x_n)$ и по критерию факторизации статистика $S_n = X_1 + \dots + X_n$ достаточна.

Задача 1б

Пусть случайные величины X_1, \dots, X_n имеют распределение $E(\frac{1}{\theta})$. Найдём достаточную статистику. Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\theta} e^{-\frac{x_i}{\theta}} I(x_i > 0) = \\ &= \theta^{-n} e^{-\frac{1}{\theta} \sum_{i=1}^n x_i} \cdot I(x_1, \dots, x_n > 0). \end{aligned}$$

Положим $g(\theta, T) = \theta^{-n} e^{-\frac{T}{\theta}}$ и $h(X_1, \dots, X_n) = I(X_1, \dots, X_n > 0)$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1 + \dots + x_n)h(x_1, \dots, x_n)$ и по критерию факторизации статистика $S_n = X_1 + \dots + X_n$ достаточна.

Задача 1в

Пусть случайные величины X_1, \dots, X_n имеют распределение $\mathcal{N}(\theta, \theta)$. Найдём достаточную статистику. Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x_i-\theta)^2}{2\theta}} = \\ &= \theta^{-\frac{n}{2}} e^{-\frac{1}{2\theta} \sum_{i=1}^n x_i^2} e^{-\frac{n\theta}{2}} (2\pi)^{-\frac{n}{2}} e^{\sum_{i=1}^n x_i}. \end{aligned}$$

Положим $g(\theta, T) = \theta^{-\frac{n}{2}} e^{-\frac{T}{2\theta}} e^{-\frac{n\theta}{2}}$ и $h(X_1, \dots, X_n) = (2\pi)^{-\frac{n}{2}} e^{\sum_{i=1}^n x_i}$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1^2 + \dots + x_n^2)h(x_1, \dots, x_n)$ и по критерию факторизации статистика $S_n^2 = X_1^2 + \dots + X_n^2$ достаточна.

Задача 1г

Пусть случайные величины X_1, \dots, X_n имеют распределение $\mathcal{N}(\theta, 1)$. Найдём достаточную статистику. Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\theta)^2}{2}} = \\ &= e^{\theta \sum_{i=1}^n x_i} e^{-\frac{n\theta^2}{2}} (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2}. \end{aligned}$$

Положим $g(\theta, T) = e^{\theta T} e^{-\frac{n\theta^2}{2}}$ и $h(X_1, \dots, X_n) = (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2}$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1 + \dots + x_n)h(x_1, \dots, x_n)$ и по критерию факторизации статистика $S_n = X_1 + \dots + X_n$ достаточна.

Задача 1д

Пусть случайные величины X_1, \dots, X_n имеют распределение $R[0, \theta]$. Найдём достаточную статистику. Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\theta} I(0 < x_i < \theta) = \\ &= \frac{1}{\theta^n} I(X_{(n)} < \theta) I(X_{(1)} > 0). \end{aligned}$$

Положим $g(\theta, T) = \frac{1}{\theta^n} I(T < \theta)$ и $h(X_1, \dots, X_n) = I(X_{(1)} > 0)$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, \max(x_1, \dots, x_n))h(x_1, \dots, x_n)$ и по критерию факторизации статистика $X_{(1)} = \max(X_1, \dots, X_n)$ достаточна.

Задача 2

Пусть T — достаточная статистика. По критерию факторизации имеем $L(\theta, x_1, \dots, x_n) = g(\theta, T(x_1, \dots, x_n)) \cdot h(x_1, \dots, x_n)$. Тогда

$$\frac{\partial}{\partial \theta} (\ln L(\theta, x_1, \dots, x_n)) = \frac{\frac{\partial}{\partial \theta} g(\theta, T(x_1, \dots, x_n))}{g(\theta, T(x_1, \dots, x_n))}.$$

По теореме о неявной функции в окрестности всякого θ_0 , при котором вторая частная производная по θ не равна 0, можно выразить значения θ , удовлетворяющие условию $\frac{\partial}{\partial \theta} g(\theta, T(x_1, \dots, x_n)) = 0$, как функцию от $T(x_1, \dots, x_n)$. Таким образом, оценка максимального правдоподобия выражается через достаточную статистику.

Задача 3а

Найдём несмешённые статистики с наименьшей дисперсией для случайных величин X_1, \dots, X_n , распределённых по закону $Pois(\theta)$. Для начала найдём полную и достаточную статистику.

Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n P_\theta(X_i = x_i) = \prod_{i=1}^n e^{-\theta} \frac{\theta^{x_i}}{x_i!} I(x_i \in \mathbb{Z}_{\geq 0}) = \\ &= e^{-n\theta} \frac{\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} I(x_1, \dots, x_n \in \mathbb{Z}_{\geq 0}). \end{aligned}$$

Положим $g(\theta, T) = e^{-n\theta} \theta^T$ и $h(X_1, \dots, X_n) = \frac{I(x_1, \dots, x_n \in \mathbb{Z}_{\geq 0})}{\prod_{i=1}^n x_i!}$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1 + \dots + x_n)h(x_1, \dots, x_n)$ и по критерию факторизации статистика $T = X_1 + \dots + X_n$ достаточна.

Проверим, что статистика $X_1 + \dots + X_n$ полна. Пусть $E_\theta f(X_1 + \dots + X_n) = 0$. Отсюда для всех θ выполнено равенство

$$0 = \sum_{m=0}^{+\infty} \left(f(m) \sum_{x_1+\dots+x_n=m} L(\theta, x_1, \dots, x_n) \right) = \sum_{m=0}^{+\infty} \frac{f(m) \cdot n^m e^{-n\theta}}{m!} \theta^m.$$

После домножения на $e^{-n\theta}$ воспользуемся равенством коэффициентов разложения в ряд Тейлора. Тогда $f(m) = 0$ для всех целых m и, следовательно, f почти наверное равна 0 (относительно распределения $\text{Poiss}(\theta)$).

Теперь для каждой функции $a(\theta)$ достаточно подобрать функцию f такую, что $E_\theta(f(T)) = a(\theta)$. Тогда полученная статистика будет иметь наименьшую дисперсию. Значения f можно определить из разложения в ряд Тейлора, приведённого выше.

Пусть $a(\theta) = e^{-\theta}$. Имеем

$$\sum_{m=0}^{+\infty} \frac{f(m) \cdot n^m}{m!} \theta^m = e^{(n-1)\theta} = \sum_{m=0}^{+\infty} \frac{(n-1)^m}{m!} \theta^m.$$

Отсюда $f(m) = \left(\frac{n-1}{n}\right)^m$ и $f(X_1 + \dots + X_n) = \left(\frac{n-1}{n}\right)^{X_1+\dots+X_n}$.

Пусть $a(\theta) = e^{-\theta} \frac{\theta^n}{n!}$. Имеем

$$\sum_{m=0}^{+\infty} \frac{f(m) \cdot n^m}{m!} \theta^m = e^{(n-1)\theta} \cdot \frac{\theta^n}{n!} = \sum_{m=0}^{+\infty} \frac{(n-1)^m}{m! n!} \theta^{m+n} = \sum_{m=n}^{+\infty} \frac{(n-1)^{m-n}}{(m-n)! n!} \theta^m.$$

Отсюда $f(m) = C_m^n \left(\frac{1}{n}\right)^n \left(\frac{n-1}{n}\right)^{m-n} I(m \geq n)$ и

$$f(X_1 + \dots + X_n) = C_{X_1+\dots+X_n}^n \left(\frac{1}{n}\right)^n \left(\frac{n-1}{n}\right)^{X_1+\dots+X_n-n} I(X_1 + \dots + X_n \geq n).$$

Пусть $a(\theta) = \theta^3$. Имеем

$$\sum_{m=0}^{+\infty} \frac{f(m) \cdot n^m}{m!} \theta^m = e^{n\theta} \theta^3 = \sum_{m=0}^{+\infty} \frac{n^m}{m!} \theta^{m+3} = \sum_{m=3}^{+\infty} \frac{n^{m-3}}{(m-3)!} \theta^m.$$

Отсюда $f(m) = \frac{m(m-1)(m-2)}{n^3}$ и

$$f(X_1 + \dots + X_n) = \frac{(X_1+\dots+X_n)(X_1+\dots+X_n-1)(X_1+\dots+X_n-2)}{n^3}.$$

Задача 3б

Найдём несмешённые статистики с наименьшей дисперсией для случайных величин X_1, \dots, X_n , распределённых по закону $\text{Geom}(\theta)$. Для начала найдём полную и достаточную статистику.

Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n P_\theta(X_i = x_i) = \prod_{i=1}^n \theta(1-\theta)^{x_i-1} I(x_i \in \mathbb{Z}_{\geq 1}) = \\ &= \theta^n (1-\theta)^{\sum_{i=1}^n x_i - n} I(x_1, \dots, x_n \in \mathbb{Z}_{\geq 1}). \end{aligned}$$

Положим $g(\theta, T) = \theta^n (1-\theta)^{T-n}$ и $h(X_1, \dots, X_n) = I(x_1, \dots, x_n \in \mathbb{Z}_{\geq 1})$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1 + \dots + x_n) h(x_1, \dots, x_n)$ и по критерию факторизации статистика $T = X_1 + \dots + X_n$ достаточна.

Проверим, что статистика $X_1 + \dots + X_n$ полна. Пусть $E_\theta f(X_1 + \dots + X_n) = 0$. Отсюда для всех θ выполнено равенство

$$0 = \sum_{m=0}^{+\infty} \left(f(m) \sum_{x_1+\dots+x_n=m} L(\theta, x_1, \dots, x_n) \right) = \sum_{m=0}^{+\infty} \frac{f(m)(m-1)!}{(m-n)!(n-1)!} \theta^n (1-\theta)^{m-n}.$$

После деления на $\theta^n (1-\theta)^{-n}$ воспользуемся равенством коэффициентов разложения в ряд Тейлора в точке 1. Тогда $f(m) = 0$ для всех целых m и, следовательно, f почти наверное равна 0 (относительно распределения $\text{Geom}(\theta)$).

Теперь для каждой функции $a(\theta)$ достаточно подобрать функцию f такую, что $E_\theta(f(T)) = a(\theta)$. Тогда полученная статистика будет иметь наименьшую дисперсию. Значения f можно определить из разложения в ряд Тейлора, приведённого выше.

Пусть $a(\theta) = \theta$. Имеем

$$\sum_{m=0}^{+\infty} \frac{f(m)(m-1)!}{(m-n)!(n-1)!} (1-\theta)^m = \frac{(1-\theta)^n}{\theta^{n-1}} = \sum_{m=0}^{+\infty} \frac{(m+n-2)!}{m!(n-2)!} (1-\theta)^{m+n} = \sum_{m=n}^{+\infty} \frac{(m-2)!}{(m-n)!(n-2)!} (1-\theta)^m.$$

Отсюда $f(m) = \frac{n-1}{m-1} I(m \geq n)$ и $f(X_1 + \dots + X_n) = \frac{n-1}{X_1+\dots+X_n-1} I(X_1 + \dots + X_n \geq n)$.

Пусть $a(\theta) = \theta(1-\theta)$. Имеем

$$\sum_{m=0}^{+\infty} \frac{f(m)(m-1)!}{(m-n)!(n-1)!} (1-\theta)^m = \frac{(1-\theta)^{n+1}}{\theta^{n-1}} = \sum_{m=0}^{+\infty} \frac{(m+n-2)!}{m!(n-2)!} (1-\theta)^{m+n+1} = \sum_{m=n+1}^{+\infty} \frac{(m-3)!}{(m-n-1)!(n-2)!} (1-\theta)^m.$$

Отсюда $f(m) = \frac{(m-n)(n-1)}{(m-1)(m-2)} I(m \geq n+1)$ и

$$f(X_1 + \dots + X_n) = \frac{(X_1+\dots+X_n-n)(n-1)}{(X_1+\dots+X_n-1)(X_1+\dots+X_n-2)} I(X_1 + \dots + X_n \geq n+1).$$

Пусть $a(\theta) = \theta^2$.

$$\sum_{m=0}^{+\infty} \frac{f(m)(m-1)!}{(m-n)!(n-1)!} (1-\theta)^m = \frac{(1-\theta)^n}{\theta^{n-2}} = \sum_{m=0}^{+\infty} \frac{(m+n-3)!}{m!(n-3)!} (1-\theta)^{m+n} = \sum_{m=n}^{+\infty} \frac{(m-3)!}{(m-n)!(n-3)!} (1-\theta)^m.$$

Отсюда $f(m) = \frac{((n-1)(n-2))}{(m-1)(m-2)} I(m \geq n)$ и

$$f(X_1 + \dots + X_n) = \frac{((n-1)(n-2))}{(X_1+\dots+X_n-1)(X_1+\dots+X_n-2)} I(X_1 + \dots + X_n \geq n).$$

Задача 3в

Найдём несмешённые статистики с наименьшей дисперсией для случайных величин X_1, \dots, X_n , распределённых по закону $R[0, \theta]$. Для начала найдём полную и достаточную статистику.

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\theta} I(0 < x_i < \theta) = \\ &= \frac{1}{\theta^n} I(X_{(n)} < \theta) I(X_{(1)} > 0). \end{aligned}$$

Положим $g(\theta, T) = \frac{1}{\theta^n} I(T < \theta)$ и $h(X_1, \dots, X_n) = I(X_{(1)} > 0)$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, \max(x_1, \dots, x_n)) h(x_1, \dots, x_n)$ и по критерию факторизации статистика $X_{(1)} = \max(X_1, \dots, X_n)$ достаточна (мы уже проверяли это в задаче 1д).

Вычислим функцию распределения $X_{(n)}$ и её плотность. Имеем

$$F_{X_{(n)}, \theta}(x) = P_\theta(X_{(n)} \leq x) = \prod_{i=1}^n P(X_i \leq x) = \frac{x^n}{\theta^n} I(0 < x < \theta) + I(x \geq \theta),$$

$$f_{X_{(n)}, \theta} = \frac{nx^{n-1}}{\theta^n} I(0 \leq x \leq \theta).$$

Проверим, что статистика $X_{(n)}$ полна. Пусть $E_\theta f(X_{(n)}) = 0$. Отсюда для всех θ выполнено равенство

$$0 = \int_{-\infty}^{+\infty} f(x) f_{\theta, X_{(1)}} dx = \int_0^\theta \frac{nf(x)x^{n-1}dx}{\theta^n}.$$

Тогда

$$0 = \int_0^\theta f(x)x^{n-1}dx$$

и после дифференцирования по θ получаем $f(\theta)\theta^{n-1} = 0$ и $f(\theta) = 0$. Таким образом, f почти наверное равна 0 (относительно распределения $R[0, \theta]$) и $X_{(n)}$ полна.

Теперь для каждой функции $a(\theta)$ достаточно подобрать функцию f такую, что $E_\theta(f(T)) = a(\theta)$. Тогда полученная статистика будет иметь наименьшую дисперсию.

Пусть $a(\theta) = \frac{1}{\theta^2}$. Имеем

$$\int_0^\theta f(x)x^{n-1}dx = \frac{1}{n}\theta^{n-2}.$$

Продифференцируем по θ . Для $n \geq 3$ получаем $f(\theta)\theta^{n-1} = \frac{n-2}{n}\theta^{n-3}$ и $f(\theta) = \frac{n-2}{n}\theta^{-2}$. Отсюда $f(X_{(n)}) = \frac{n-2}{n(X_{(n)})^2}$.

Пусть $a(\theta) = \ln \theta$. Имеем

$$\int_0^\theta f(x)x^{n-1}dx = \frac{1}{n}\theta^n \ln \theta.$$

Продифференцируем по θ . Получаем $f(\theta)\theta^{n-1} = \theta^{n-1} \ln \theta + \frac{1}{n}\theta^{n-1}$ и $f(\theta) = \ln \theta + \frac{1}{n}$. Отсюда $f(X_{(n)}) = \ln X_{(n)} + \frac{1}{n}$.

Пусть $a(\theta) = e^\theta$. Имеем

$$\int_0^\theta f(x)x^{n-1}dx = \frac{1}{n}\theta^n e^\theta.$$

Продифференцируем по θ . Получаем $f(\theta)\theta^{n-1} = \theta^{n-1}e^\theta + \frac{1}{n}\theta^n e^\theta$ и $f(\theta) = e^\theta(1 + \frac{\theta}{n})$. Отсюда $f(X_{(n)}) = e^{X_{(n)}}(1 + \frac{X_{(n)}}{n})$.

Задача 3г

Найдём несмешённые статистики с наименьшей дисперсией для случайных величин X_1, \dots, X_n , распределённых по закону $E(\frac{1}{\theta})$. Для начала найдём полную и достаточную статистику.

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\theta} e^{-\frac{x_i}{\theta}} I(x_i > 0) = \\ &= \theta^{-n} e^{-\frac{1}{\theta} \sum_{i=1}^n x_i} \cdot I(x_1, \dots, x_n > 0). \end{aligned}$$

Положим $g(\theta, T) = \theta^{-n} e^{-\frac{T}{\theta}}$ и $h(X_1, \dots, X_n) = I(X_{(1)} > 0)$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, x_1 + \dots + x_n)h(x_1, \dots, x_n)$ и по критерию факторизации статистика $S_n = X_1 + \dots + X_n$ достаточна (мы ранее проверяли это в задаче 1б).

Случайная величина S_n имеет гамма-распределение $\Gamma(n, \frac{1}{\theta})$ и плотность $f_{S_n, \theta}(x) = \frac{x^{n-1} e^{-\frac{x}{\theta}}}{\Gamma(n)\theta^n} I(x > 0)$

Проверим, что статистика S_n полна. Пусть $E_\theta f(S_n) = 0$. Отсюда для всех θ выполнено равенство

$$0 = \int_{-\infty}^{+\infty} f(x) f_{\theta, X_{(1)}} dx = \int_0^{+\infty} \frac{f(x) x^{n-1} e^{-\frac{x}{\theta}}}{\Gamma(n)\theta^n} dx.$$

Тогда

$$0 = \int_0^{+\infty} f(x) x^{n-1} e^{-\frac{x}{\theta}} dx.$$

В правой части равенства написано преобразование Лапласса от функции $f(x)x^{n-1}$ (или преобразование Меллина от функции $f(x)e^{-\frac{x}{\theta}}$). Из его инъективности (или инъективности преобразования Меллина) получаем $f(x)x^{n-1} = 0$ и $f(x) = 0$ (почти наверное). Поэтому статистика S_n полна.

Теперь для каждой функции $a_k(\theta) = \theta^k$ подберём несмещённую статистику T_k для a_k , а после вычислим условное матожидание $E_\theta(T_k | S_n)$. Имеем $E(X_1^k) = k! \theta^k$ и $E(\frac{1}{k!} \overline{X^k}) = \theta^k$.

Из симметричности и линейности

$$\begin{aligned} E_\theta(\frac{1}{k!} \overline{X^k} | S_n) &= \frac{1}{k!} E_\theta(X_1^k | S_n) = \frac{1}{k!} \int_{-\infty}^{+\infty} x^k \cdot f_{X_1|S_n, \theta}(x | y) dx = \frac{1}{k!} \int_{-\infty}^{+\infty} x^k \cdot \frac{f_{X_1, S_n, \theta}(x, y)}{f_{S_n, \theta}(y)} dx = \\ &= \frac{1}{k!} \int_{-\infty}^{+\infty} x^k \cdot \frac{f_{S_n|X_1, \theta}(y|x) f_{X_1, \theta}(x)}{f_{S_n, \theta}(y)} dx = \frac{1}{k!} \int_{-\infty}^{+\infty} x^k \cdot \frac{f_{S_{n-1}, \theta}(y-x) f_{X_1, \theta}(x)}{f_{S_n, \theta}(y)} dx = \\ &= \frac{1}{k!} \int_{-\infty}^{+\infty} x^k \cdot \frac{(y-x)^{n-2} e^{-\frac{y-x}{\theta}}}{\Gamma(n-1)\theta^{n-1}} I(y-x > 0) \cdot \frac{e^{-\frac{x}{\theta}}}{\theta} I(x > 0) \cdot \frac{\Gamma(n)\theta^n}{y^{n-1} e^{-\frac{y}{\theta}}} I(y > 0) dx = \\ &= \frac{n-1}{k!} \int_{-\infty}^{+\infty} \frac{x^k (y-x)^{n-2}}{y^{n-1}} I(y-x > 0) \cdot I(x > 0) \cdot I(y > 0) dx = \frac{n-1}{k!} I(y > 0) \int_0^y x^k (1 - \frac{x}{y})^{n-2} d\frac{x}{y} = \\ &= \frac{(n-1)y^k}{k!} I(y > 0) \int_0^1 t^k (1-t)^{n-2} dt = \frac{(n-1)y^k}{k!} I(y > 0) B(k+1, n-1) = \\ &= y^k \cdot \frac{(n-1)k!(n-2)!}{k!(n+k-1)!} I(y > 0) = y^k \frac{(n-1)!}{(n+k-1)!} I(y > 0). \end{aligned}$$

Отсюда $E_\theta(T_k | S_n) = (S_n)^k \frac{(n-1)!}{(n+k-1)!} I(S_n > 0)$.

Задача 4а

Пусть X_1, \dots, X_n распределены по закону $R[\theta_1, \theta_2]$. Имеем

$$\begin{aligned} L(\theta, x_1, \dots, x_n) &= \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n \frac{1}{\theta_2 - \theta_1} I(\theta_1 < x_i < \theta_2) = \\ &= \frac{1}{(\theta_2 - \theta_1)^n} I(X_{(n)} < \theta_2) I(X_{(1)} > \theta_1). \end{aligned}$$

Положим $g(\theta, T) = \frac{1}{(\theta_2 - \theta_1)^n} I(T_2 < \theta_2) I(T_1 > \theta_1)$, где T — вектор, и $h(X_1, \dots, X_n) = 1$. Тогда $L(\theta, x_1, \dots, x_n) = g(\theta, (\min(x_1, \dots, x_n), \max(x_1, \dots, x_n))) h(x_1, \dots, x_n)$ и по критерию факторизации статистика $(X_{(1)}, X_{(n)})$ достаточна.

Задача 46

Докажем неполноту (? см. ниже) статистики из предыдущего пункта, построив ненулевые последовательности $\{a_n\}$ и $\{b_n\}$ такие, что $E_\theta(a_n X_{(1)} + b_n X_{(n)}) = 0$ для любого $\theta = (\theta_1, \theta_2)$.

Вычислим функции распределения и плотности величин $X_{(1)}$ и $X_{(n)}$.

$$\begin{aligned} F_{X_{(n)}, \theta}(x) &= P_\theta(X_{(n)} \leq x) = P_\theta(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = \\ &= \left(\frac{x-\theta_1}{\theta_2-\theta_1}\right)^n I(\theta_1 < x < \theta_2) + I(x \geq \theta_2), \\ f_{X_{(n)}, \theta}(x) &= \frac{n(x-\theta_1)^{n-1}}{(\theta_2-\theta_1)^n} I(\theta_1 < x < \theta_2). \\ F_{X_{(1)}, \theta}(x) &= P_\theta(X_{(1)} \leq x) = 1 - P_\theta(X_{(1)} > x) = 1 - P_\theta(X_1 > x, \dots, X_n > x) = \\ &= 1 - \prod_{i=1}^n (1 - P(X_i \leq x)) = 1 - \left(\frac{\theta_2-x}{\theta_2-\theta_1}\right)^n I(\theta_1 < x < \theta_2) + I(x \leq \theta_1), \\ f_{X_{(1)}, \theta}(x) &= \frac{n(\theta_2-x)^{n-1}}{(\theta_2-\theta_1)^n} I(\theta_1 < x < \theta_2). \end{aligned}$$

Теперь вычислим их матожидания:

$$\begin{aligned} E_\theta(X_{(n)}) &= \frac{n}{(\theta_2-\theta_1)^n} \int_{\theta_1}^{\theta_2} x(x-\theta_1)^{n-1} dx = \frac{n}{(\theta_2-\theta_1)^n} \left(\int_{\theta_1}^{\theta_2} (x-\theta_1)^n dx + \theta_1 \int_{\theta_1}^{\theta_2} (x-\theta_1)^{n-1} dx \right) = \\ &= \frac{n}{(\theta_2-\theta_1)^n} \left(\frac{(\theta_2-\theta_1)^{n+1}}{n+1} + \frac{\theta_1(\theta_2-\theta_1)^n}{n} \right) = \frac{n}{n+1} \theta_2 + \frac{1}{n+1} \theta_1. \\ E_\theta(X_{(1)}) &= \frac{n}{(\theta_2-\theta_1)^n} \int_{\theta_1}^{\theta_2} x(\theta_2-x)^{n-1} dx = \frac{n}{(\theta_2-\theta_1)^n} \left(- \int_{\theta_1}^{\theta_2} (\theta_2-x)^n dx + \theta_2 \int_{\theta_1}^{\theta_2} (\theta_2-x)^{n-1} dx \right) = \\ &= \frac{n}{(\theta_2-\theta_1)^n} \left(- \frac{(\theta_2-\theta_1)^{n+1}}{n+1} + \frac{\theta_2(\theta_2-\theta_1)^n}{n} \right) = \frac{n}{n+1} \theta_1 + \frac{1}{n+1} \theta_2. \end{aligned}$$

Теперь $E_\theta(a_n X_{(1)} + b_n X_{(n)}) = 0$ для любого $\theta = (\theta_1, \theta_2)$ тогда и только тогда, когда $a_n n + b_n = 0$ и $a_n + nb_n = 0$. Таким образом, при $n = 1$ получаем $a_1 = -b_1$ и при больших n выполнено $a_n = b_n = 0$ (вроде, хотели получить ненулевые a_k, b_k ?)