Точната формула за оценка се формира в зависимост от резултатите. За приблизителна, може да използвате 2 + брой точки. Време за работа: 210 минути. Успех.

Задача 1. (1.5 т.) Нека ξ и η са независими сл. вел., $\xi \sim Exp(4)$ и $\eta \sim U(-1, 1)$.

1. (0.5 т.) Намерете $\mathbb{E}[\xi\eta], \mathbb{E}[\xi|\eta = 0]$, $\mathrm{Corr}(\xi, \eta)$ и $P(\xi \leq 0.25)$.
2. (0.5 т.) Намерете плътността на $\xi^2 + 3$ и ξ/η.
3. (0.5 т.) Нека $Y_1, Y_2, Y_3, Y_4 \sim Exp(1)$ са независими. Нека $Y_{(i)}$ е i-тото по големина от тях. Намерете очакването и дисперсията на $Y_{(3)}$.

Задача 2. (1 т.) По едно и също време двама клиенти пристигат в система за обслужване, която има две опашки: ако T е времето за обслужване:

$$T \sim \begin{cases} Exp(1), & \text{(бавна опашка) с вероятност } 0.3, \\ Exp(3), & \text{(бърза опашка) с вероятност } 0.7. \end{cases}$$

1. (0.5 т.) Нека T_i е времето за обслужване на i-тия клиент. Намерете разпределението $T_1 + T_2$.
2. (0.25 т.) Намерете $E(T_1 + T_2)$?
3. (0.25 т.) За големи n, оценете вероятността сумата от n такива независими времена да надвишава $5n$.

Задача 3. (1 т.) Нека $X_1, X_2, \ldots$ са независими, еднакво разпределени сл. вел. с плътност

$$f_X(x) = \begin{cases} 2x, & \text{ако } 0 < x < 1, \\ 0, & \text{иначе.} \end{cases}$$

1. (0.5 т.) Приблизително колко наблюдения n са нужни, за да е вярно с 97% сигурност, че средната извадкова стойност $\overline{X}_n := (X_1 + \cdots + X_n)/n$ ще бъде между 0.5 и 0.7?
2. (0.5 т.) Колко горе-долу очаквате да е $\max\limits_{1 \leq k \leq n} X_k$ при $n = 10\,000$?

Задача 4. (0.75 т.) Нека X и Y са независими сл. вел., $X \sim \mathcal{N}(2, 4)$ и $Y \sim \mathcal{N}(2, 9)$.

1. (0.25 т.) Намерете съвместната им плътност и $P(X < Y)$.
2. (0.5 т.) Пресметнете $E(|X - Y|)$ и $|E(X) - E(Y)|$. А бихте ли могли да кажете кое от тях е по-голямо без пресмятане на точните им стойности?

Задача 5. (1 т.) Нека ABC е правоъгълен триъгълник с катети $AC = BC = 1$. През *случайно* (равномерно) избрана точка P от катета AC е прекарана права, успоредна на хипотенузата AB. Нека Q и R са пресечните точки на тази права с катетите.

1. (0.25 т.) Намерете очакването на дължината на QR.
2. (0.5 т.) Намерете вероятността две независимо случайно избрани точки от вътрешността на ABC да лежат от различни страни на QR.
3. (0.25 т.) Нека A, B са избрани случайно върху окръжност с център т. O и радиус 1. Какво е очакваното лице на триъгълника AOB? Какво е очакването на $AO \cdot AB \cdot \sin(OAB)/2$? Обяснете защо двете са/не са равни.

Задача 6. *(Бонус 1 т.)*

В „класическата" (частично наричана „честотна") статистика при оценка на неизвестен параметър λ най-често използваме методи като максимално правдоподобие или метод на моментите, без да моделираме самия λ като случайна величина. За сметка на това, при *Бейсовия подход*, предполагаме, че λ има някакво *априорно* (prior, предварително, от лат.: *от по-рано, преди)*) разпределение, например $\Gamma(\alpha, \beta)$, и комбинираме наблюдаваните данни с това априорно знание, за да получим *апостериорното* (posterior, от лат.: *от по-късно)*) разпределение на λ. Така можем да генерираме или прогнозираме нови наблюдения, които са от разпределение, в което сме отчели както вече наблюдаваните данни.

Удобни от изчислителна гледна точка са модели, при които апостериорното разпределение е от същия тип като априорното, като разликата се изразява в промяна на параметрите (чрез данните сме „научили" стойности, които по-добре описват реалността). По-долу е представен един такъв случай: при експоненциален модел $Exp(\lambda)$ и априорно гама разпределение за неизвестния параметър λ, апостериорното също е от тип гама.

Нека $X_1, \dots, X_n$ са независими наблюдения над експоненциално разпределена случайна величина,

$$X \mid \lambda \sim \mathrm{Exp}(\lambda),$$

където параметърът $\lambda > 0$ е неизвестен. В рамките на Бейсовата постановка приемаме, че априорно λ има $\Gamma(\alpha, \beta)$ разпределение, т.е. има плътност

$$f_\lambda(x) \;=\; \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, \exp(-\beta x), \quad x > 0.$$

1. (0.25 т.) Докажете, че апостериорното разпределение на λ след наблюдавани $X_1, \dots, X_n$ е

$$\lambda \mid X_1, \dots, X_n \sim \Gamma\Big(\alpha + n,\ \beta + \sum_{i=1}^{n} X_i\Big).$$

2. (0.25 т.) Намерете прогнозното разпределение на едно ново наблюдение X_{n+1} (т.е. условната вероятност $\mathbb{P}\big(X_{n+1} \in A \mid X_1, \dots, X_n\big)$) и покажете, че в някои случаи то може да се изрази с т.нар. *Pareto* разпределение Y:

$$F_Y(x) = 1 - \left(\frac{a}{x}\right)^\alpha, \quad x > a,$$

за някакъв параметър a. Последното е един от класическите примери за разпределение с тежка опашка. (*може да интегрирате по x разпределението* $\mathbb{P}(X_{n+1} \in A \mid \lambda = x) \times f_{\lambda \mid X_1,\dots,X_n}(x \mid X_1, \dots, X_n)$.)

3. (0.5 т.) Намерете $\mathrm{E}(\lambda \mid X_1, \dots, X_n)$ и $\mathrm{Var}(\lambda \mid X_1, \dots, X_n)$. Напомняме, че условната дисперсия на сл. вел. Y при дадено X се дефинира като

$$D(Y \mid X) \;=\; E\big[(Y - E(Y \mid X))^2 \mid X\big].$$

От последната дефиниция докажете също, че

$$D(Y) \;=\; E\big[D(Y \mid X)\big] \;+\; D\big(E[Y \mid X]\big).$$

4. *(бонус на бонуса)* Как бихте интерпретирали разликата между класическия доверителен интервал и Бейсовия (highest posterior density) интервал за λ в този пример?