Точната формула за оценка се формира в зависимост от резултатите. За приблизителна, може да използвате 2 + брой точки. Време за работа: 3 часа. Успех.

Ще считаме, че навсякъде работим върху вероятностно пространство $(\Omega, \mathcal{F}, \mathbb{P})$.

Задача 1. Законът на Бенфорд[1], известен още като „законът за първа цифра", е емпирично наблюдение, че в много реални бази от данни, първата ненулева цифра (при това не само в десетично представяне!) не е равномерно разпределена, а е най-често 1. Десетичното разпределние на Бенфорд пък е разпределение върху $\{1, 2, \ldots, 9\}$, за което вероятността на цифрата d е $\log_{10}(1 + 1/d)$.

Намерете с точност 0.01 вероятността първата ненулева цифра в десетичното представяне на X да бъде 1 за:

1. (0.25 т.) $X \sim N(0, 1)$;
2. (0.25 т.) $X \sim Exp(1)$;
3. (0.75 т.) $X = U_1/U_2$ за независими $U_1, U_2 \sim U(0, 1)$. Намерете $\mathbb{E}X$.
4. *бонус: (0.5 т.)* X = *„случайна положителна степен на 2"*. Формално, намерете $\lim_{n\to\infty} \mathbb{P}(X_n$ започва с 1) за $X_n = 2^{U_n}$, $U_n \sim U(\{1, 2, \ldots, n\})$.

Сравнете с разпределението на Бенфорд.[2]

Задача 2. Нека $X_1, X_2 \sim \Gamma(3, 2)$, т.е $f_X(x) = 4x^2e^{-2x}\mathbb{1}_{\{x>0\}}$.

1. (0.5 т.) Намерете плътността $Y = X_1/(X_1 + X_2)$.
2. (0.25 т.) Намерете $Cor(X_1, X_2)$ и $Cor(X_1 + X_2, Y)$.
3. (0.25 т.) Независими ли са $X_1 + X_2$ и Y?

Задача 3. Според компания за производство на чипове, само 1 на всеки 1000 чипа е неизправен.

1. (0.5 т.) Как бихте оценили вероятността от 100 чипа да има поне 1 неизправен чрез ЦГТ? Какви други начини бихте предложили?
2. (0.5 т.) Неравенството на Бери-Есен гласи, че ако $X_1, \ldots, X_n$ са iid сл. вел. и

$$\mu := \mathbb{E}X_1, \sigma := \sqrt{DX}, \rho := \mathbb{E}\left[|X - \mu|^3\right] < \infty, Z_n := \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}},$$

то

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(Z_n \le x) - \Phi(x)| \le \frac{\rho}{2\sigma^3\sqrt{n}}.$$

Как това може да е полезно при евентуално решение на 1.? При какви n, грешката при приближение чрез ЦГТ би била под 0.001, ако приемете че информацията, дадена от компанията е вярна?

Задача 4. Нека съвместната плътност на X и Y е $f_{X,Y}(x, y) = cx^3 + 1$ за $x, y \ge 0, x + 4y \le 1$ и 0 извън тази област, където c е някаква константа. Намерете:

1. (0.5 т.) c, плътността на X и очакването на Y;
2. (0.25 т.) $\mathbb{E}(Y|X = 1/2)$;
3. (0.25 т.) плътността на случайната величина $Z = X + 2Y$.

[1] Забелязан от Newcomb в логаритмични таблици през 1881 и при по-голям набор от данни от Benford през 1934.

[2] Ако имате време, може да опитате да пресметнете вероятностите и за цифрите, по-големи от 1 и да сравните с разпределението на Бенфорд.