Introduction to Probability Theory
10th week: Random variables continuous and discrete.

Fact 1. *For a random variable X with the CDF F, we have*

$$\mathbb{P}(a < X \leq b) = F(b) - F(a).$$

Fact 2. *Let $A \subseteq \mathbb{R}$. Then, for a discrete random variable X we have*

$$\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x),$$

while for a continuous random variable Y with density f

$$\mathbb{P}(Y \in A) = \int_{x \in A} f(x)dx.$$

Theorem 1. *If random variables $X_1, \ldots, X_k$ are independent, then*

$$\operatorname{Var}(X_1 + X_2 + \cdots + X_k) = \operatorname{Var} X_1 + \operatorname{Var} X_2 + \cdots + \operatorname{Var} X_k.$$

Below we list some basic information on the most popular discrete distributions.

distribution	parameters	$\mathbb{P}(X = k)$	for $k =$	$\mathbb{E}(X)$	$\operatorname{Var}(X)$
binomial	n, p	$\binom{n}{k} p^k (1-p)^{n-k}$	$0, 1, \ldots, n$	np	$np(1-p)$
Poisson	λ	$\frac{\lambda^k}{k!} e^{-\lambda}$	$0, 1, 2, \ldots$	λ	λ
geometric	p	$(1-p)^{k-1} p$	$1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
hipergeometric	N, m, n	$\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	$0, 1, 2, \ldots, n$	$n\frac{m}{N}$	$n\frac{m}{N}\frac{N-m}{N}\frac{N-n}{N-1}$

A random variable X has **exponential distribution** with parameter $\lambda > 0$, if its CDF F is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x \leq 0. \end{cases}$$

Then for the density of X we have

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x \leq 0, \end{cases}$$

$\mathbb{E}X = 1/\lambda$, and $\operatorname{Var} X = 1/\lambda^2$.

Homework assignment for Quiz 7, May 9th.

Problem 1. The CDF F of a random variable X is given by

$$F(x) = \begin{cases} 0 & \text{for} \quad x \leq 0 \\ x^3 & \text{for} \quad 0 \leq x \leq 1 \\ 1 & \text{for} \quad x \geq 1. \end{cases}$$

Find the denisty f of X, $\mathbb{E}X$, and compute $\mathbb{P}(1/2 \leq X \leq 2)$ and $\mathbb{P}(X \leq \mathbb{E}X)$.

Problem 2. The CDF $F_{X,Y}$ of a random variable (X, Y) is given by

$$F_{X,Y}(x, y) = \begin{cases} (1 - e^{-x})(1 - e^{-y}) & \text{for} \quad x, y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal CDF's F_X and F_Y, and the densities $f_{X,Y}$, f_X, f_Y.
(b) Are X and Y independent?
(c) Compute $\mathbb{P}(-3 \leq X \leq 2)$ and $\mathbb{P}(X > 1)$.

Problem 3. The density of a random variable (X, Y) is given by

$$f_{X,Y}(x,y) = \begin{cases} c & \text{in the triangle with vertices } (0,0), (0,1), (1,1) \ , \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the CDF $F_{X,Y}$ and the densities of marginal distributions f_X and f_Y.
(b) Are X and Y independent?
(c) Compute $\mathbb{P}(1/4 \leq X \leq 3/2)$.
(d) Find $\mathrm{Cov}(X, Y)$.

Problem 4. The average number of accidents per year at some crossing in Poznań is 3.8. Find the probability that in 2020 there will be more than two accidents at this crossing. Use Poisson distribution and a calculator.

Problem 5. The time (measured in hours) it takes a CS student to find in the internet a nice, affordable notebook is exponentially distributed with parameter $\lambda = 1.5$. Find the probability that a student will find a notebook in more than 2 but less than 3 hours.

Problems we shall do in class on May 9th

Problem 1. The density of (X, Y) is given by

$$f_{X,Y}(x) = \begin{cases} cy & \text{for } (x,y) \text{ inside the triangle with vertices in } (0,0), (0,4), (1,4). \\ 0 & \text{otherwise.} \end{cases}$$

(a) Compute c.
(b) Find the density f_X, the CDF F_X and the expectation $\mathbb{E}X$.
(c) Are X and Y independent? Do you need to compute the precise form of f_Y or F_Y to verify it?
(d) Using F_X, find $\mathbb{P}(1/2 \leq X \leq 3/2)$.
(e) Find $\mathbb{P}(1/2 \leq X \leq 3/2)$, using f_X.
(f) Find $\mathbb{P}(X \geq 7/8)$, and the estimate it using Markov's inequality.

Problem 2. According to police statistics, at some crossing there are in average 0,9 traffic accidents per year. Find the probability that in the next year the will be at least three accidents at that crossing. Then estimate this probability using Markov's and Chebyshev's inequalities.

Problem 3. At five different places in the town five different yoga gurus ask pedestrians to do head-stand. The average number of persons one yoga teacher can convince to do it is 1.4 per day. Find the probability that at a given day precisely three yoga teachers will convince at least two people to perform this exercise.
