MSAI Probability Home Assignment 4-5
soft deadline: 18/10/2024 23:59 Moscow Time
hard deadline: 24/10/2024 18:00 Moscow Time

As announced earlier, grading for HWs consists of points and bonus points. Solving bonus (indicated with a star) problems is not required, but recommended. Solving all homeworks' normal problems correctly will give you a score of 7, solving all homeworks' bonus problems correctly will give you additional 2 points to the score.

Soft deadline is the intended deadline for this homework. Hard deadline is the date and time of homework discussion webinar, where we will discuss solutions to this homework. After solutions are released, no more homeworks are accepted.

Hand-written solutions are accepted if the handwriting is clear enough and scanned with sufficient quality, but LaTeX is always preferable. This homework includes a python task, which can be solved in Google Colab or in a local Jupyter Notebook. It is thus handy to solve everything (both LaTeX and code) in a single Jupyter Notebook.

Problem 1. (2 points) If $X \sim Bi(n,p)$ and $Y \sim Bi(m,p)$, and $X \perp Y$ (independent), what is $\mathbb{P}(X \mid X+Y=r)$, the PMF of the conditional distribution of X given $X+Y=r$? What is the name of this distribution?

Problem 2. (2 points) Find the PMF of distribution of a random variable X, which is equal to the number of failures in a series of Bernoulli trials with success probability p, which are carried out **not** a fixed number of times, but instead until there are r successes. This distribution is called Negative Binomial distribution.

Problem 3. (1 point) An airline overbooks a flight, selling more tickets for the flight than there are seats on the plane (figuring that it's likely that some people won't show up). The plane has 100 seats, and 110 people have booked the flight. Each person will show up for the flight with probability 0.9, independently. Find the probability that there will be enough seats for everyone who shows up for the flight.

Problem 4. (2 points) Consider two independent random variables $X \sim F_X$ and $Y \sim F_Y$. Find the CDF of random variables $Z_1 = \max(X,Y)$ (it means that for every outcome w we have $Z_1(w) = \max(X(w), Y(w))$, so Z_1 jumps between values of X and Y) and $Z_2 = \min(X,Y)$ (same reasoning applies).

Problem 5* . (1 bonus points) Use Python package SCIPY.STATS to plot the PMF of distribution from Problem 2 for different values of parameter p.

Problem 6* . (2 bonus points) A good student solves a problem correctly with probability 0.95, while a bad student — with probability 0.15. What is the minimal number of problems that the test should include so that the probability that good student does not pass the test does not exceed 0.01, and the probability that the bad student passes the test does not exceed 0.1? Passing the test means solving strictly more than half of the problems. You can use Python to solve this problem.

Problem 7* . (3 bonus points) Two players are playing a game. The first player says a number p_1 between 0 and 1. The second player, knowing the number of the first player, says a number p_2 between 0 and 1. Then, with probability p_1 the first number becomes zero, and with probability p_2 the second number becomes zero. The player whose number is greater, wins.

- (2 bonus points) What is the optimal winning strategy of player two?
- (1 bonus point) What is the optimal winning strategy of player one, if player two follows strategy from the previous point?

Hint: simply consider different cases in which the second player has to make the choice. From the cases, the solution is easy to get.