

YZV231E - Prob. & Stat. for Data Science

2023/2024 Fall

Homework 3

Professor: Murat Okatan
Res. Asst.: Fatih Bektaş
bektas18@itu.edu.tr

- Cheating is highly discouraged. It will be punished by a negative grade. Also disciplinary actions will be taken. Please do your homework on your own. Team work is not allowed. Pattern of your **solutions must belong only to you**.
- Upload your solutions through **Ninova**. Homeworks sent via e-mail and late submissions **will not be accepted**.
- You should write all your codes in **Python (3.8+)** language using **Jupyter notebook**, unless stated otherwise in the question. You can install Jupyter Notebook by following these steps on [this documentation](#). If you are not familiar with Jupyter Notebook, you can check [this tutorial](#). You may also use [Google Colab](#).
- Prepare a **report** using Latex including all your solutions, **codes** and their results.
- You **should use the provided Latex** template for the report. You may want to use [mathpix](#). You **cannot embed your handwritten solutions** into the document unless it's stated so. **Handwritten solutions will not be evaluated**.
- Mathematical calculations must be done **step-by-step**.
- This homework has a total of **100 points**. Each question has **specified** number of points.

1 Problems (70 Points)

Question 1 (10 pts)

The joint density function $f_{X,Y}(x, y)$ is given by

$$f_{X,Y}(x, y) = \begin{cases} \frac{12}{49}(2 + x + xy + 4y^2), & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

- Compute $f_X(x)$ for all $x \in \mathbb{R}^1$.
- Compute $f_Y(y)$ for all $y \in \mathbb{R}^1$.
- Determine whether or not X and Y are independent.

Question 2 (30 pts)

Suppose you roll one fair six-sided die and then flip as many coins as the number showing on the die. (For example, if the die shows 4, then you flip four coins.) Let X be the number showing on the die, and Y be the number of heads obtained.

- Formulate the joint PMF $f_{X,Y}[x, y]$.
- Compute $\text{Cov}(X, Y)$.
- Compute $\text{Corr}(X, Y)$.
- Interpret the correlation value.

Question 3 (10 pts)

Suppose X and Y are discrete, with

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{5}, & \text{if } (x, y) \in \{(2, 3), (3, 2), (3, 3), (2, 2), (3, 17)\}, \\ 0, & \text{otherwise.} \end{cases}$$

- Compute $E[X|Y = 3]$:
- Compute $E[Y|X = 3]$:
- Compute $E[X|Y]$:
- Compute $E[Y|X]$:

Question 4 (5 pts)

Let Z_n be the sum of the squares of the numbers showing when we roll n fair dice. Find (with proof) a number m such that $\frac{1}{n}Z_n \rightarrow m$. (Hint: Use the strong law of large numbers.)

Question 5 (15 pts)

Suppose the service time, in minutes, at a bank has the Exponential distribution with $\lambda = \frac{1}{2}$. Use the central limit theorem to estimate the probability that the average service time of the first n customers is less than 2.5 minutes, when:

- a) $n = 16$.
- b) $n = 36$.
- c) $n = 100$.
- d) What happens when n increases? Explain briefly why.

2 Coding (30 Points)

For this part, please download the Jupyter notebook and code the necessary parts.

3 Submission

You need to prepare a report for all the parts including coding part and its results. You need to submit the files shown below as zip file named with your number (i.e 123123123.zip).

- | | |
|-----------------------------------|-------------------------------|
| 1. Homework 3 - Coding Part.ipynb | # solution of the coding part |
| 2. report.pdf | # report file as pdf format |
| 3. report.tex | # report file as tex format |