Instructions

- The homework is due on **Friday 2/3 at 5pm ET**.

- There are 3 paper-pencil-like problems and 1 coding problem.

- LaTeX is recommended but hand-written solutions will be accepted as long as they are legible. The coding part and the associated text should be coded in the Jupyter notebook provided to you on the class Git repo.

- No extension will be provided, unless for serious documented reasons.

- **Start early!**

- Study the material taught in class, and feel free to do so in small groups, but the solutions should be a product of your own work.

- This is not a multiple choice homework; reasoning, and mathematical proofs are required before giving your final answer.
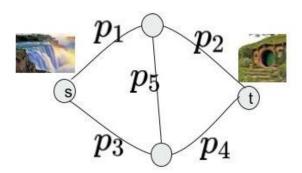
1   Is there a path from s to t? [15 points]

Figure 1:   How likely is water to reach t from s?

Consider the network shown in Figure 1. Recall that p_i is the probability pipe i breaks down, and that pipes break down independently.

1. What is the probability water can go from the water source s to the destination village t? Explain your answer. You do not need to simplify it algebraically.

2 PIN Cracker [30 Points]

John has a cell phone with a PIN that consists of 4 digits (0-9). Unfortunately John totally forgot his PIN, but at least he can try PIN numbers as many times as he wants to without blocking the device. He applies the following two strategies[1]:

(s_1) He tries a valid PIN uniformly at random each hour, till he enters the right one.

(s_2) He keeps track of the unsuccessful attempts, and chooses a PIN uniformly at random from the PIN numbers he has not tried yet.

1. [**10 points**] Write code that simulates strategies s_1 and s_2. You may assume that the correct PIN is 2022 in your code. Simulate each strategy 100, 200, 300, ..., 1 000 times, and report for each number of trials the average number of trials, and the standard deviation till John figures out his PIN. Present your empirical findings in two plots (one per strategy) with error bars, where the x-axis is the number of trials.

2. Let X_1, X_2 be the number of trials under strategies s_1 and s_2 respectively. Compute their expectations analytically:

    i) [**5 points**] $\mathbb{E}[X_1]$.
    ii) [**5 points**] $\mathbb{E}[X_2]$.

3. [**10 points**] Apply Markov's inequality to upper bound the probability $\mathbb{P}(X_i \geq 7\,000)$ for $i = 1, 2$. Does Markov's inequality always provide meaningful bounds?

---

[1]In reality, there exist better strategies to break a PIN https://www.popsci.com/technology/article/2012-09/infographic-day-fastest-way-crack-4-digit-pin-number/, but let's assume we employ only naive strategies here.

3   [25 points]

Provide clean proofs and calculations for the following problems.

1. (5 points) Two fair dice are rolled. Show that the event that their sum is 7 is independent of the score shown by the first die.

2. (10 points) For random variables X, Y, prove $\mathbb{E}(XY)^2 \le \mathbb{E}(X^2)\mathbb{E}(Y^2)$. When does equality hold?

3. (5 points) If one picks a numerical entry at random from tax documents, the first two significant digits, X, Y respectively, are found to have approximately the joint mass function

$$f(x,y) = \log_{10}(1 + \frac{1}{10x+y}), 1 \le x \le 9, 0 \le y \le 9.$$

Find the mass function of the first significant digit X. What do you observe, i.e., are digits 1 and 9 equally likely to be the first digit? What is the expected value of X?

4. (5 points) Find the conditional density function of Y given X when they have the joint density function $f(x,y) = \lambda^2 e^{-\lambda y}$ for $0 \le x \le y < \infty$.

4   Coding [30 points]

The coding part of the homework is available on Git here. Follow the instructions in the Jupyter notebook. After finishing it, download the notebook in **PDF format** which should include all the code and figures. If you have trouble converting it to PDF, you can download it as HTML, and then save it as PDF. Submit the math problems and coding problems **separately in two files** through Gradescope.