# 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.

Name: Caiwei Zahng, BUID: U97031910

## 1 Is there a path from $s$ to $t$? [15 points]

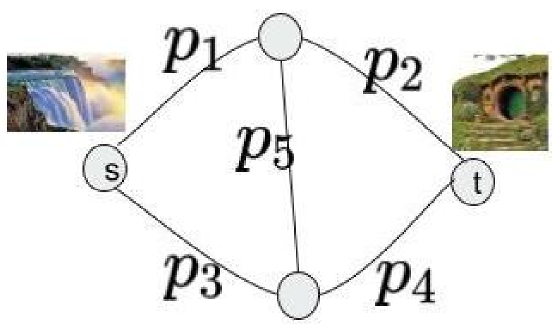



Figure 1: How likely is water to reach $t$ from $s$?

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.

### 1.1 Answer

# 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?

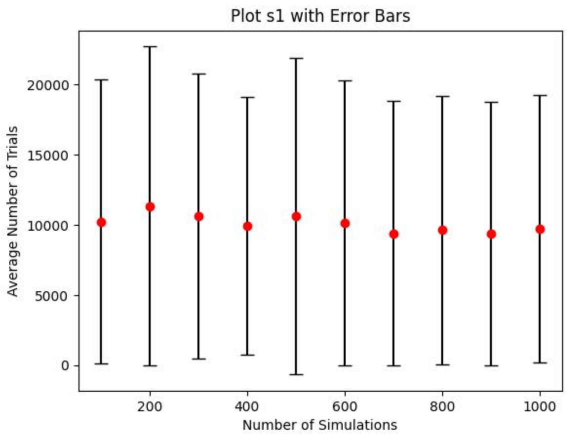
## 2.1 Answer

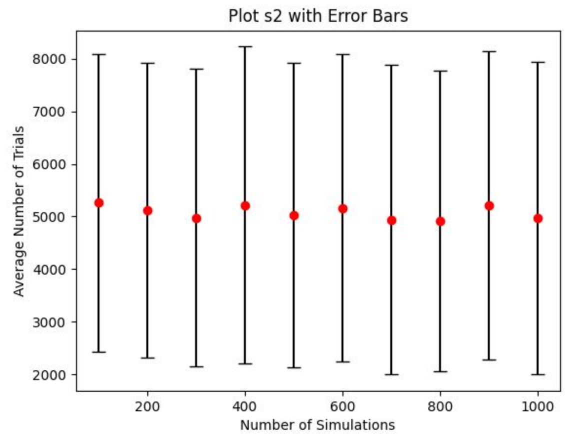
The plots record the empirical simulation findings, with the points record the average trials to solve the PIN with respect to different number of simulations, and the error bars represent the standard deviation of the trials with respect to each set of simulations.
(Plots are shown below)
As shown on the plot, using s1 strategies, we need to run about 10000 trails to get the right PIN, while using s2 strategies, we need to run about 5000 trails to get the right PIN.

## 2.2 Answer

$i)$ $s1$ follows a geometric distribution with the probability of successfully generate the right PIN equals $\frac{1}{10^4}$, so $X_1 \backsim G(\frac{1}{10^4})$. Based on the properties of geometric distribution, the expected values of the number of trails used untill sucess is $\mathbb{E}(X_1) = \frac{1}{p} = 10^4 = 10000$.

[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.




*ii)*

| $X_2$ | $P(X_2)$ |
|---|---|
| 1 | $\frac{1}{10^4}$ |
| 2 | $\frac{1}{10^4}$ |
| 3 | $\frac{1}{10^4}$ |
| $\cdots$ | $\cdots$ |
| 9999 | $\frac{1}{10^4}$ |
| 10000 | $\frac{1}{10^4}$ |

$\mathbb{E}(X_2) = \sum_{x=1}^{10000} x \cdot \frac{1}{10000} = \frac{1}{10000} \sum_{x=1}^{10000} x = \frac{1}{10000} \cdot \frac{(1+10000)\cdot 10000}{2} = \frac{1}{2} \cdot (1+10000) = 5000.5$

## 2.3 Answer

Markov's Inequality: $P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$
for i), $P(X_1 \geq 7000) \leq \frac{10000}{7000}$, which means $P(X_1 \geq 7000) \leq 1.42857$. However, based on the axiom of probability, the probability of any event is a number between 0 and 1, inclusive. So the Markov's Inequality gives us no meaningful interpretation for s1 strategy.
for ii), $P(X_2 \geq 7000) \leq \frac{5000.5}{7000}$, which means $P(X_2 \geq 7000) \leq 0.715$. In this case, the Markov's Inequality gives us meaningful bound that the probability of number of trails under strategy s2 goes over 7000 is less than or equal to about 0.715.

# 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 $\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$. 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

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

   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 $0 \leq x \leq y < \infty$.

## 3.1 Answer

Two events A and B are independent if and only if $P(A|B) = P(A)$
Rolling two dice, let A be the event that the sum of the two dice is 7 and B be the event that the first die shows a score $x$.
Therefore, $P(A) = 6/36 = 1/6$ and since the second die can only show one score to get a sum of 7, $P(A|B = x) = \frac{P(A \cap B)}{P(B)} = \frac{1/36}{1/6} = 1/6$, where $x$ can be 1,2,3,4,5,6. In this way, $P(A) = P(A|B)$, which proves that the event that the sum of two dice being 7 is independent of the score shown by the first die.

## 3.2 Answer

Based on Cauchy-Schwarz inequality, for any two vectors x and y in a Euclidean space, $(x \cdot y)^2 \leq x^2 y^2$. Now considering $X, Y$ as two vectors, and the dot product of X and Y is given by the expected value of $XY$. Therefore, $(\mathbb{E}(XY))^2 <= (\mathbb{E}(X^2))(\mathbb{E}(Y^2))$.
The equality holds when $X$ and $Y$ are linearly dependent, which means that there exists a non-zero constant c such that $Y = cX$ .

## 3.3 Answer

Using Marginal Probability Mass Function of $X$, $f_X(x) = \sum_y f(x, y)$ as $0 \leq y \leq 9$ .

$$f_X(x) = log_{10}(1 + \frac{1}{10x}) + log_{10}(1 + \frac{1}{10x+1}) + \cdots + log_{10}(1 + \frac{1}{10x+9})$$

$$f_X(x) = log_{10}(\frac{10x+1}{10x} \cdot \frac{10x+2}{10x+1} \cdot \cdots \cdots \cdot \frac{10x+10}{10x+9})$$

$$f_X(x) = log_{10}(\frac{10x+10}{10x})$$

$$f_X(x) = log_{10}(\frac{x+1}{x})$$

To see whether digits 1 and 9 are equally likely to be the first digit, we just need to compare $f_X(1)$ and $f_X(9)$.
Based on the formula above, $f_X(1) = log_{10}(2)$, while $f_X(9) = log_{10}(\frac{10}{9})$, which is not equal. Therefore we can conclude that they are not equally likely to be the first digit.
For expected value of X, we can use density function that $\mathbb{E}(X) = \sum_x xf_X(x)$, for $1 \leq x \leq 9$.

$$\mathbb{E}(X) = \sum_1^9 xlog_{10}(\frac{x+1}{x})$$

$$\mathbb{E}(X) = 1log_{10}(\frac{2}{1}) + 2log_{10}(\frac{3}{2}) + 3log_{10}(\frac{4}{3}) + \cdots + 9log_{10}(\frac{10}{9})$$

$$\mathbb{E}(X) = log_{10}(2) - log_{10}(1) + 2log_{10}(3) - 2log_{10}(2) + 3log_{10}(4) - 3log_{10}(3) + \cdots + 9log_{10}(10) - 9log_{10}(9)$$

$$\mathbb{E}(X) = -log_{10}(1) - log_{10}(2) - log_{10}(3) - \cdots - log_{10}(9) + 9log_{10}(10)$$

$$\mathbb{E}(X) = 9 - log_{10}(1 \cdot 2 \cdot 3 \cdots 9)$$

$$\mathbb{E}(X) = 9 - 5.5524 = 3.4476.$$

Therefore, the expected value of first digits is about 3.4476.

## 3.4 Answer

Conditional density formula: $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$, where $f_X(x)$ is the marginal density function of $X$: $f_X(x) = \int_x^\infty f(x, y)dy$.
Substituting the given joint density function $f(x, y) = \lambda^2 e^{-\lambda y}$ for $0 \leq x \leq y < \infty$ into the expression:

$$f_{Y|X}(y|x) = \frac{\lambda^2 e^{-\lambda y}}{\int_x^\infty \lambda^2 e^{-\lambda y} dy}$$

$$f_{Y|X}(y|x) = \frac{\lambda^2 e^{-\lambda y}}{\frac{\lambda^2}{\lambda} e^{-\lambda x}}$$

$$f_{Y|X}(y|x) = \lambda e^{\lambda x - \lambda y}$$

So the conditional density function of $Y$ given $X$ is $\lambda e^{\lambda x - \lambda y}$

# 4 Coding [30 points]

The coding part of the homework is available on Git here[2]. 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.

[2]https://github.com/tsourolampis/cs365-spring23/blob/main/hw1/HW1coding.ipynb