Instructions

- The homework is due on **Friday 2/4 at 5pm ET**.

- There are 4 problems, and 3 pages in total.

- 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.

- The code necessary to generate the plots in Problems 2.1, and 3.1, should be included in the PDF of your solutions; see `https://www.overleaf.com/learn/latex/Code_listing` for several ways this can be done in LaTeX, that is the recommended language to typeset your solutions.

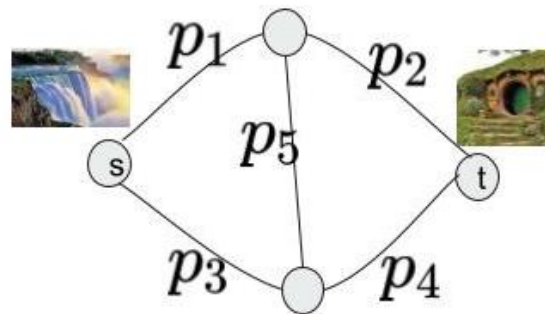# 1   Is there a path from s to t? [15 points]

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

Consider the network shown in Figure **??**. 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   Mixture of Gaussians [40 points]

Let X, Y be two independent normal RVs, with means $\mu_x = 100, \mu_y = 300$ and standard deviations $\sigma_x = \sigma_y = 10$. Consider the RV U defined by

$$U = \frac{1}{2}(X + Y).$$

Alternatively, consider the RV Z that is generated as follows:

(a) With probability $\frac{1}{2}$ we sample Z from $N(\mu = 100, \sigma^2 = 10)$.

(b) With probability $\frac{1}{2}$ we sample Z from $N(\mu = 300, \sigma^2 = 10)$.

1. **[10 points]** Simulate the sampling, and produce two histograms (one for U and one for Z) over $10\,000$ samples for each U, Z.

2. **[10 points]** Compute the expected values of U, Z.

3. **[20 points]** Compute the variances of U, Z.

   Hint: To compute the variance of Z, it will be helpful express Z as a function of X, Y and an indicator variable $\mathbf{I}$ that indicates from which of the two distributions (a),(b) the realization of Z was drawn.

4   Gambling [15 points]

Suppose a casino has a finite fortune F (e.g., 10.5M\$) that is willing to put on a table against you. You play the following game against the casino using a fair coin: if heads (H) appear for the first time after k tosses, the game is over, and you get paid 2^k dollars. We agree to toss the coin n times or until it shows up heads, whatever happens first.

1. How much money would you be willing to put on this bet if $F = 10\,500\,000\$$, $n = 30$? Report your answer with 4 decimal digits of accuracy.