UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Problem Set 2
Spring 2021

1. **Among Us**

In the game of Among Us, there are 9 players. 4 of them are imposters and 5 of them are crewmates. There is also a deck of 17 cards containing 11 "sabotage" cards and 6 "task" cards. Imposters want to play sabotage cards, and crewmates want to play task cards. Here's how the play proceeds.

- A captain and a first mate are chosen uniformly at random from the 9 players.
- The captain draws 3 cards from the deck and gives 2 to the first mate, discarding the third.
- The first mate chooses one to play.

Now suppose you are the first mate, but the captain gave you 2 sabotage cards. Being a crewmate, you wonder, did the captain just happen to have 3 sabotage cards, or was the captain an imposter who secretly discarded a task card. In this scenario, what's the probability that the captain is an imposter? Let's assume that imposter captains always try to discard task cards, and crewmate captains always try to discard sabotage cards.

Solution: Let's define the following.

- Let F be the event that the captain is an imposter.
- Let O be the event that you are handed 2 sabotage cards.

By Bayes Rule,

$$P(F \mid O) = \frac{P(O \mid F)P(F)}{P(O \mid F)P(F) + P(O \mid F^{\complement})P(F^{\complement})}$$

You know that $P(F) = \frac{4}{8}$ and $P(F^{\complement}) = \frac{4}{8}$ (you are a crewmate). For $P(O \mid F)$, there are two cases.

- The captain drew 3 sabotage cards.
- The captain drew 2 sabotage cards and 1 task card, but discarded the task card.

.
On the other hand, for $P(O \mid F^{\complement})$, there is only one case.

- The captain drew 3 sabotage cards.

.
The probability that the captain draws 3 sabotage cards is

$$\frac{\binom{11}{3}\binom{6}{0}}{\binom{17}{3}} = \frac{\frac{11 \cdot 10 \cdot 9}{6} \cdot 1}{\frac{17 \cdot 16 \cdot 15}{6}}$$

$$= \frac{33}{136}$$

The probability that the captain draws 2 sabotage cards and 1 task card is

$$\frac{\binom{11}{2}\binom{6}{1}}{\binom{17}{3}} = \frac{\frac{11\cdot 10}{2}\cdot 6}{\frac{17\cdot 16\cdot 15}{6}} = \frac{66}{136}$$

Putting everything together, we get

$$P(F \mid O) = \frac{(\frac{33}{136}+\frac{66}{136})\cdot\frac{4}{8}}{(\frac{33}{136}+\frac{66}{136})\cdot\frac{4}{8}+\frac{33}{136}\cdot\frac{4}{8}} = \frac{99}{99+33} = \frac{3}{4}$$

2. Lightbulbs

Consider an $n \times n$ array of switches. Each row i of switches corresponds to a single lightbulb L_i, so that L_i lights up if at least i switches in row i are flipped on. All of the switches start in the "off" position, and each are flipped "on" with probability p, independently of all others. What is the expected number of lightbulbs that will be lit up? Express your answer in closed form without any summations.

Solution: Each row i of switches can be represented by a random variable $X_i \sim \text{Binomial}(n, p)$ for $1 \le i \le n$. We are interested in the expectation

$$\mathbb{E}[\mathbb{1}_{X_1 \ge 1} + \mathbb{1}_{X_2 \ge 2} + \cdots + \mathbb{1}_{X_n \ge n}].$$

By linearity, this becomes

$$\sum_{i=1}^{n} \mathbb{E}[\mathbb{1}_{X_i \ge i}] = \sum_{i=1}^{n} P(X_i \ge i) = \sum_{i=1}^{n} P(X \ge i),$$

where $X \sim \text{Binomial}(n, p)$. Using the tail sum formula, this is just $\mathbb{E}[X] = np$.

3. Random Bipartite Graph

Consider a random bipartite graph with, K left nodes and M right nodes. Each of the $K \cdot M$ possible edges of this graph is present with probability p independently.

(a) Find the distribution of the degree of a particular right node.

(b) Now, pick a left node u and right node v. Conditioned on the event that the edge (u, v) is present, what is the distribution of the degree of the right node v? Is it the same as in part (a)?

(c) We call a right node with degree one a *singleton*. What is the average number of singletons in a random bipartite graph?

(d) Find the average number of left nodes that are connected to at least one singleton.

Solution:

(a) Consider a particular right node, let's say right node i. Each of the left nodes are connected to right node i with probability p independently of the others. Thus, the number of edges connected to right node i has a binomial distribution with parameters K and p. So

$$P(\text{degree of right node } i = x) = \binom{K}{x} p^x (1-p)^{K-x},\ 0 \le x \le K.$$

(b) When an edge is selected, one left node is definitely connected to the corresponding right node. The other $K-1$ left nodes are connected to the right node with probability p. Thus,

$$P(\text{degree of right node corresponding to a selected edge} = x) = \binom{K-1}{x-1} p^{x-1} (1-p)^{K-x},\ 1 \le x \le K.$$

(c) The probability of a right node being a singleton is $\binom{K}{1} p(1-p)^{K-1}$. Thus, by linearity of expectation the average number of singletons is $MKp(1-p)^{K-1}$.

(d) First we find the probability of a left node being connected with at least one singleton. Consider the ith left nodel.

$$\begin{aligned}
&P(\text{left nodel } i \text{ is connected to at least one singleton}) \\
=&1 - P(\text{left node } i \text{ is not connected to any singleton}) \\
=&1 - \prod_{j=1}^{M} [1 - P(\text{left node } i \text{ is connected to right node } j \text{ and right node } j \text{ is a singleton})] \\
=&1 - [1 - p(1-p)^{K-1}]^M,
\end{aligned}$$

where the second equality is because of the independence of all the connections (edges). Then by linearity of expectation, the average number of left nodes that are connected to at least one singleton is

$$K[1 - (1 - p(1-p)^{K-1})^M].$$

Other solutions: We can also use conditioning to solve this problem but the solution is more complicated. We still find the probability of a left node being in at least one singleton first. As we can see, the degree of a right node is a binomial distribution with parameters M and p. Now we condition on the degree of a right node being d, with $1 \le d \le M$. Consider one of these d edges. The probability that the left node connected to this edge is a singleton is $\binom{K-1}{0} p^0 (1-p)^{K-1} = (1-p)^{K-1}$ by part (c). Thus, conditioned on the degree of a left node being d, the probability that the left node is in not connected to any singletons is the probability that none of the d left nodes connected to the right node is a singleton that is:

$$(1 - (1-p)^{K-1})^d.$$

Then we know that the probability that a left node is connected to at least one singleton, conditioned on the degree of the left node being d is:

$$1 - (1 - (1 - p)^{K-1})^d.$$

By law of total probability, we have

$$\begin{aligned}
&P(\text{a left node is connected to at least one singleton}) \\
=&\sum_{d=1}^{M}[1 - (1 - (1 - p)^{K-1})^d]\binom{M}{d}p^d(1-p)^{M-d} \\
=&[1 - (1-p)^M] - \sum_{d=1}^{M}\binom{M}{d}(p - p(1-p)^{K-1})^d(1-p)^{M-d} \\
=&[1 - (1-p)^M] - [(1 - p(1-p)^{K-1})^M - (1-p)^M] \\
=&1 - (1 - p(1-p)^{K-1})^M.
\end{aligned}$$

Therefore, by linearity of expectation, the average number of left nodes that are connected to at least one singleton is

$$K[1 - (1 - p(1-p)^{K-1})^M].$$

4. **Compact Arrays**

Consider an array of n entries, where n is a positive integer. Each entry is chosen uniformly randomly from $\{0, \ldots, 9\}$. We want to make the array more compact, by putting all of the non-zero entries together at the front of the array. As an example, suppose we have the array

$$[6, 4, 0, 0, 5, 3, 0, 5, 1, 3].$$

After making the array compact, it now looks like

$$[6, 4, 5, 3, 5, 1, 3, 0, 0, 0].$$

Let i be a fixed positive integer in $\{1, \ldots, n\}$. Suppose that the ith entry of the array is non-zero (assume that the array is indexed starting from 1). Let X be a random variable which is equal to the index that the ith entry has been moved after making the array compact. Calculate $\mathbb{E}[X]$ and $\operatorname{var}(X)$.

Solution:

Let X_j be the indicator that the jth entry of the original array is 0, for $j \in \{1, \ldots, i-1\}$. Then, the ith entry is moved backwards $\sum_{j=1}^{i-1} X_j$, positions, so

$$\mathbb{E}[X] = i - \sum_{j=1}^{i-1}\mathbb{E}[X_j] = i - \frac{i-1}{10} = \frac{9i+1}{10}.$$

The variance is also easy to compute, since the X_j are independent. Then, $\operatorname{var}(X_j) = (1/10)(9/10) = 9/100$, so

$$\operatorname{var}(X) = \operatorname{var}\left(i - \sum_{j=1}^{i-1} X_j\right) = \sum_{j=1}^{i-1}\operatorname{var}(X_j) = \frac{9(i-1)}{100}.$$

5. **Clustering Coefficient**

This problem will explore an important probabilistic concept of clustering that is widely used in machine learning applications today. Consider n students, where n is a positive integer. For each pair of students $i, j \in \{1, \dots, n\}$, $i \neq j$, they are friends with probability p, independently of other pairs. We assume that friendship is mutual. We can see that the friendship among the n students can be represented by an undirected graph G. Let $N(i)$ be the number of friends of student i and $T(i)$ be the number of triangles attached to student i. We define the **clustering coefficient** $C(i)$ for student i as follows:

$$C(i) = \frac{T(i)}{\binom{N(i)}{2}}.$$

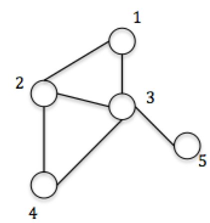

Figure 1: Friendship and clustering coefficient.

The clustering coefficient is not defined for the students who have no friends. An example is shown in Figure **??**. Student 3 has 4 friends (1, 2, 4, 5) and there are two triangles attached to student 3, i.e., triangle 1-2-3 and triangle 2-3-4. Therefore $C(3) = 2/\binom{4}{2} = 1/3$.

Find $\mathbb{E}[C(i) \mid N(i) \geq 2]$.

Solution:

First, we compute $\mathbb{E}[C(i) \mid N(i) = k]$, for $k \in \{2, \dots, n-1\}$. Suppose that student i has friends $f_1, \dots, f_k$. We can see that $T(i)$ equals the number of friend pairs among $\{f_1, \dots, f_k\}$. Since there are $\binom{k}{2}$ possible pairs and each pair of students are friends with probability p, independently of other pairs, we know that the expected number of friend pairs among $\{f_1, \dots, f_k\}$ is $\binom{k}{2}p$. Then we have

$$\mathbb{E}[C(i) \mid N(i) = k] = \frac{\binom{k}{2}p}{\binom{k}{2}} = p.$$

Since this is true for all $k \geq 2$, we have $\mathbb{E}[C(i) \mid N(i) \geq 2] = p$.