UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

**Problem Set 2**
Fall 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

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

Putting everything together, we get

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

**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. 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 $i$th 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 $i$th entry has been moved after making the array compact. Calculate $\mathbb{E}[X]$ and $\text{var}(X)$.

**Solution:**

Let $X_j$ be the indicator that the $j$th entry of the original array is 0, for $j \in \{1, \ldots, i-1\}$. Then, the $i$th 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, $\text{var}(X_j) = (1/10)(9/10) = 9/100$, so

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

4. **Borel-Cantelli & Strong Law**

In this problem, we walk through a proof of the strong law (assuming finite 4th moments) that relies only on basic probability. In class we covered the *Borel-Cantelli lemma*, which states that for events $(A_n)_{n=1}^{\infty}$, if $\sum_{n=1}^{\infty} P(A_n) < \infty$, then

$$P(A_n \text{ i.o.}) = 0,$$

where we define the event $\{A_n \text{ i.o.}\} = \cap_{n \geq 1} \cup_{m \geq n} A_n$ as the event where infinitely many $A_n$ occur.

(a) Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^4 < \infty$ (and so we also have finite second and third moments). Let $S_n = X_1 + \cdots + X_n$, and compute $\mathbb{E}[S_n^4]$. Write your answer in terms of the moments $\mathbb{E}[X_i^2], \mathbb{E}[X_i^3], \mathbb{E}[X_i^4]$.

(b) Fix an $\epsilon > 0$, and use Markov's inequality to show that, for any $n$,

$$P(|S_n/n| > \epsilon) \leq O(n^{-2}).$$

(c) Finally, use Borel-Cantelli to conclude that $P(\lim_{n\to\infty} S_n/n = 0) = 1$. This a weaker (the full theorem assumes only finite first moments) form of the *strong law of large numbers*.

**Solution:**

(a) We expand:

$$\mathbb{E}S_n^4 = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right)^4 = \mathbb{E}\sum_{1 \leq i,j,k,l \leq n} X_i X_j X_k X_l.$$

Terms of the form $\mathbb{E}[X_i^3 X_j]$, $\mathbb{E}[X_i^2 X_j X_k]$, and $\mathbb{E}[X_i X_j X_k X_l]$ are just 0 by independence. We are left with

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i^4\right] + \mathbb{E}\left[\sum_{i \neq j} X_i^2 X_j^2\right] = n\mathbb{E}[X_1^4] + 3n(n-1)\mathbb{E}[X_1^2]\mathbb{E}[X_2^2].$$

(b) By Markov's inequality and the previous part, we have

$$P(|S_n/n| > \epsilon) < \epsilon^{-4}\mathbb{E}(S_n/n)^4 = O(\epsilon^{-4}n^{-2}).$$

(c) Letting $A_n = \{|S_n/n| > \epsilon\}$, we get from the Borel-Cantelli lemma that $P(|S_n/n| > \epsilon \text{ i.o.}) = 0$. Since $\epsilon$ is arbitrary, this implies almost sure convergence.

5. **Bounds for the Coupon Collector's Problem**

Recall the coupon collector's problem, where each box contains a single coupon, and there are $n$ different types of coupons. We let $X$ be a random variable which is equal to the number of boxes bought until one of every type of coupon is obtained.

The expected value of $X$ is $nH_n$, where $H_n$ is the *harmonic number of order* $n$ which is defined as

$$H_n \triangleq \sum_{i=1}^{n} \frac{1}{i},$$

and satisfies the inequalities

$$\ln n \le H_n \le \ln n + 1.$$

(a) Use Markov's inequality in order to show that

$$P(X > 2nH_n) \le \frac{1}{2}.$$

(b) Use Chebyshev's inequality in order to show that

$$P(X > 2nH_n) \le \frac{\pi^2}{6(\ln n)^2}.$$

*Note:* You can use the identity

$$\sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}.$$

(c) Define appropriate events and use the union bound in order to show that

$$P(X > 2nH_n) \le \frac{1}{n}.$$

*Note:* The sequence $a_n = (1 - 1/n)^n$, for $n = 1, 2, 3, \ldots$, is strictly increasing and $\lim_{n\to\infty} a_n = 1/e$.

**Solution:**

(a)

$$\begin{aligned} P(X > 2nH_n) &\le \frac{\mathbb{E}[X]}{2nH_n} \\ &= \frac{1}{2}. \end{aligned}$$

(b) Recall that $X$ can be written as $X = X_1 + \cdots + X_n$, where the $X_i$s are independent random variables with $X_i \sim \text{Geometric}(\frac{n-i+1}{n})$. Therefore

$$\begin{aligned}
\operatorname{var}(X) &= \sum_{i=1}^{n} \operatorname{var}(X_i) \\
&< \sum_{i=1}^{n} \Big(\frac{n}{n-i+1}\Big)^2 \\
&= n^2 \sum_{i=1}^{n} \frac{1}{i^2} \\
&< \frac{\pi^2 n^2}{6}.
\end{aligned}$$

So using Chebyshev's inequality we have that

$$\begin{aligned}
P(X > 2nH_n) &\le P(|X - nH_n| > nH_n) \\
&\le \frac{\operatorname{var}(X)}{(nH_n)^2} \\
&< \frac{\pi^2}{6H_n^2} \\
&\le \frac{\pi^2}{6(\ln n)^2}.
\end{aligned}$$

(c) Let $A_i$ be the event that we fail to get box $i$ after $2nH_n$ tries.

$$\begin{aligned}
P(A_i) &= \Big(\frac{n-1}{n}\Big)^{2nH_n} \\
&= \Big[\Big(1 - \frac{1}{n}\Big)^n\Big]^{2H_n} \\
&< e^{-2H_n} \\
&\le e^{-2\ln n} \\
&= \frac{1}{n^2}.
\end{aligned}$$

Now using the union bound we can conclude that

$$\begin{aligned}
P(X > 2nH_n) &= P(A_1 \cup \cdots \cup A_n) \\
&\le \sum_{i=1}^{n} P(A_i) \\
&\le \frac{1}{n}.
\end{aligned}$$