UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 01
Spring 2023

1. **Coin Flipping and Symmetry**

   Alice and Bob have $2n + 1$ fair coins, $n \geq 1$. Bob tosses $n + 1$ coins, while Alice tosses the remaining n coins. A fair coin lands on heads with probability $\frac{1}{2}$; assume that coin tosses are independent.

   a. Formulate this scenario in terms of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Describe explicitly what the outcomes are and what the probability measure of any event is defined to be.

   b. Show that the probability that after all coins have been tossed, Bob will have gotten more heads than Alice is $\frac{1}{2}$.
   Hint: Consider the event $A = \{\text{more heads in the first } n+1 \text{ tosses than the last } n \text{ tosses}\}$.

   Solution:

   a. We can let the sample space be $\Omega = \{H, T\}^{2n+1}$, the set of all possible configurations of $2n + 1$ coin tosses. We take $\mathcal{F} = 2^{\Omega}$, and the probability measure $\mathbb{P}$ is *uniform*, such that $\mathbb{P}(\{\omega\}) = 2^{-(2n+1)}$ for every outcome $\omega \in \Omega$. (Thus, $\mathbb{P}(A) = 2^{-(2n+1)}|A|$ for all $A \in \mathcal{F}$.)

   b. Define the events

   $$A = \{\text{there are more heads in the first } n+1 \text{ tosses than the last } n \text{ tosses}\},$$
   $$B = \{\text{there are more tails in the first } n+1 \text{ tosses than the last } n \text{ tosses}\}.$$

   $\mathbb{P}(A) = \mathbb{P}(B)$ by symmetry, as every outcome in A can be uniquely identified with an outcome in B by swapping heads and tails, and vice versa. We also note that $A \cap B = \varnothing$, as it is impossible for the first $n + 1$ tosses to have more heads *and* more tails than the last n tosses, and $A \cup B = \Omega$. Thus $\mathbb{P}(A) + \mathbb{P}(B) = 1$, and $\mathbb{P}(A) = \frac{1}{2}$.

   Alternatively, if the probability that Bob has more heads than Alice in the first n tosses is p, then the probability that Bob has fewer heads than Alice in the first n tosses is also p by symmetry, and the probability that they are tied after n tosses is $1 - 2p$.
   So, the probability that Bob wins is $p + (\frac{1}{2})(1 - 2p) = \frac{1}{2}$. Bob can win either by having more heads than Alice in the first n tosses, or by having the same number of heads as Alice in the first n tosses, then flipping heads on the last toss.

2. **Expanding the NBA**

The NBA is looking to expand to another city. In order to decide which city will receive a new team, the commissioner interviews potential owners from each of the N potential cities, $N \geq 1$, one at a time. The cities are interviewed in a uniformly random order.

a. Formulate this scenario in terms of a probability space. Describe explicitly what the outcomes are and what the probability measure of any event equals.

Unfortunately, the owners would like to know immediately after the interview whether their city will receive the team or not. The commissioner decides to use the following strategy: she will interview the first m owners and reject all of them, $m \in \{1, \dots, N\}$. After the mth owner is interviewed, she will pick the first city that is better than all previous cities. Assume that the commissioner has an objective scoring method, and each city receives a unique score.

b. What is the probability that the best city is selected?
You should arrive at an exact answer for the probability in terms of a summation. *Hint*: Consider the events $B_i = \{\text{the } i\text{th city is the best}\}$ and $A = \{\text{the best city is chosen}\}$.

c. Approximate your answer using $\sum_{i=1}^{n} i^{-1} \approx \ln n$, and find the optimal value of m that maximizes the probability that the best city is selected.
You may also use the fact that $\ln(n-1) \approx \ln n$.

Solution:

a. A good choice of sample space Ω is the set of all permutations of $1, \dots, N$. Each outcome, for example $(2, 5, 1, N, \dots)$, is a possible order in which the commissioner interviews the N cities. The probability measure $\mathbb{P}$ is the uniform measure, such that $\mathbb{P}(\{\omega\}) = \frac{1}{N!}$ for every $\omega \in \Omega$. And, as Ω is finite, we can take $\mathcal{F} = 2^{\Omega}$.

b. Let B_i be the event that the ith city is the best among the N cities, $i = 1, \dots, N$, and let A be the event that the best city is indeed chosen by the commissioner. Then, by the law of total probability,

$$\mathbb{P}(A) = \sum_{i=1}^{N} \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{P}(A \mid B_i).$$

Next, we see that $\mathbb{P}(A \mid B_i) = 0$ for $i = 1, \dots, m$: if the best city is among the first m cities, there is no chance of picking the best city. For $i = m+1, \dots, N$, we have $\mathbb{P}(A \mid B_i) = \frac{m}{i-1}$. This is the key observation: Let city j be the second-best city *among the first i cities*. Then, given B_i, city i will be chosen by the commissioner if and only if city j is one of the first m cities interviewed, and this has probability $\frac{m}{i-1}$. Therefore

$$\mathbb{P}(A) = \frac{m}{N} \sum_{i=m+1}^{N} \frac{1}{i-1}.$$

c. Now, we turn towards approximation.

$$\mathbb{P}(A) \approx \frac{m}{N}(\ln N - \ln m) = -\frac{m}{N} \ln \frac{m}{N}.$$

If we let $x := \frac{m}{N}$, then $\mathbb{P}(A) \approx -x \ln x$. Differentiating with respect to x suggests that the optimal value is $x = \frac{1}{e}$, so we should reject the first $\frac{N}{e}$ cities. Substituting in this value of x, the optimal probability is $\mathbb{P}(A) \approx \frac{1}{e}$.

Note: This problem is a famous example from optimal stopping theory commonly known as the **secretary problem**. (In the usual formulation, a boss is interviewing secretaries instead of a commissioner interviewing city representatives.)

In fact, one may use a dynamic programming approach to see why the policy outlined here is in fact the optimal policy. If you are interested, the details of such an approach can be found in *Dynamic Programming and the Secretary Problem* by Beckmann.

3. **Passengers on a Plane**

There are n passengers in a plane with n assigned seats, but after boarding, the passengers take the seats randomly. n is a positive integer. Assuming all seating arrangements are equally likely, what is the probability that no passenger is in their assigned seat? Also find the limit of this probability as $n \to \infty$.

Hint: Use the principle of inclusion-exclusion and the power series $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Solution: It may be easier to first find the probability of the complementary event, that at least one passenger sits in their assigned seat. If A_i is the event that passenger i sits in their assigned seat, this is

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \cdots + (-1)^{n-1}\, \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right).$$

Consider the second summation. Any particular $A_i \cap A_j$ has probability $\frac{(n-2)!}{n!}$: passengers i and j are fixed, and the remaining $n-2$ passengers can be arranged in $(n-2)!$ ways. And, there are $\binom{n}{2}$ choices of $i < j$. Generalizing to the kth summation, the above is equal to

$$\sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \frac{(n-k)!}{n!} = \sum_{k=1}^{n} (-1)^{k-1} \frac{1}{k!}.$$

Thus, the probability that no passenger is in their assigned seat is

$$1 - \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \to \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1} \approx 0.368.$$

Alternate solution (optional). Instead of using inclusion–exclusion, we can consider recursion. Let B_n be the event that no passengers sit in their assigned seat on a plane with n seats, and let $p_n = \mathbb{P}(B_n)$. We observe that $p_0 = 1$, as B_0 is vacuously true, and $p_1 = 0$. For $n \geq 2$,

$$p_n = \frac{n-1}{n} p_{n-1} + \frac{1}{n} p_{n-2}.$$

Why is this true? Suppose passenger 1 sits in seat $i \neq 1$. If passenger i does not sit in seat 1, then we take the probability that none of the $n-1$ passengers $\neq 1$ sit in their own seat. Otherwise, passenger i is in seat 1, and we only need the probability that the remaining $n-2$ passengers are not in their assigned seats. To find p_n, define $d_n = p_n - p_{n-1}$, where

$$d_n = \left[\left(1 - \frac{1}{n}\right) p_{n-1} + \frac{1}{n} p_{n-2}\right] - p_{n-1} = -\frac{1}{n}(p_{n-1} - p_{n-2}) = -\frac{1}{n} d_{n-1}.$$

The base case is $d_1 = -1$. Repeating the recurrence relation, we have

$$d_n = -\frac{1}{n} d_{n-1} = \frac{1}{n(n-1)} d_{n-2} = \cdots = \frac{(-1)^{n-1}}{n!} d_1 = \frac{(-1)^n}{n!}.$$

Finally, we can write p_n as a telescoping sum of differences, taking $d_0 := p_0$:

$$p_n = (p_n - p_{n-1}) + \cdots + (p_1 - p_0) + p_0 = \sum_{i=0}^{n} d_i = \sum_{i=0}^{n} \frac{(-1)^i}{i!}.$$

4. **Upperclassmen**

You meet two students in the library. Assume that each student is an upperclassman or underclassman with equal probability, and each student takes EECS 126 with probability $\frac{1}{10}$, independent of each other and independent of their class standing. What is the probability that both students are upperclassmen, given at least one of them is an upperclassman currently taking EECS 126?

Solution: We define the following events:

- A is the given event that at least one student is an upperclassman taking EECS 126.
- B is the event that both students are upperclassmen.
- U is the event that at least one student is an upperclassman.
- E is the event that at least one student is taking EECS 126.

We wish to find $\mathbb{P}(B \mid A)$, which we can rewrite by Bayes' rule as

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B) \cdot \mathbb{P}(B)}{\mathbb{P}(A)}.$$

To find $\mathbb{P}(A \mid B)$, we notice that given B (both students are upperclassmen), the event A can be reduced to the event E that at least one student is taking EECS 126. The condition that at least one student is an upperclassman becomes redundant. Then, as taking EECS 126 is independent of being an upperclassman,

$$\mathbb{P}(A \mid B) = \mathbb{P}(E \mid B) = \mathbb{P}(E) = 1 - \mathbb{P}(E^c) = 1 - \left(\frac{9}{10}\right)^2.$$

The probability of B is straightforward by independence:

$$\mathbb{P}(B) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

Now, to find $\mathbb{P}(A)$, we observe that A^c is the event that none of the two students is an upperclassman taking EECS 126. By the given independences,

$$\begin{aligned}\mathbb{P}(A) = 1 - \mathbb{P}(A^c) &= 1 - (\mathbb{P}(\text{is not an upperclassman taking EECS 126}))^2 \\ &= 1 - \left(1 - \frac{1}{2} \cdot \frac{1}{10}\right)^2.\end{aligned}$$

Simplifying our final expression, we get

$$\mathbb{P}(B \mid A) = \frac{\left(1 - \frac{9}{10}\right)^2 \cdot \frac{1}{4}}{1 - \left(\frac{19}{20}\right)^2} = \frac{\frac{19}{400}}{\frac{39}{400}} = \frac{19}{39}.$$

5. **Conditional Probability Space**

   Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and fix an event $B \in \mathcal{F}$ with probability $\mathbb{P}(B) > 0$.

   a. Show that $\mathcal{F}|_B \coloneqq \{A \cap B : A \in \mathcal{F}\}$ is a σ-algebra on the sample space B (*not* Ω).

   b. Consider the function $\mathbb{P}(\cdot \mid B)\colon \mathcal{F}|_B \to [0,1]$, which takes as input $A \in \mathcal{F}|_B$ and outputs its conditional probability $\mathbb{P}(A \mid B)$. Show that $(B, \mathcal{F}|_B, \mathbb{P}(\cdot \mid B))$ is a probability space satisfying Kolmogorov's axioms.

   Solution:

   a. We simply check that $\mathcal{F}|_B$ satisfies the definition of being a σ-algebra:

      i. *Nonempty.* Because $\Omega \in \mathcal{F}$, we have $\Omega \cap B = B \in \mathcal{F}|_B$ as well, so $\mathcal{F}|_B$ is nonempty.

      ii. *Closed under complements.* Suppose that $C \in \mathcal{F}|_B$, which means there exists some $A \in \mathcal{F}$ such that $C = A \cap B$. We want to show that $B \setminus C \in \mathcal{F}|_B$, where

      $$B \setminus C = B \setminus (A \cap B) = B \cap A^c,$$

      which belongs to $\mathcal{F}|_B$ because $A^c \in \mathcal{F}$.

      iii. *Closed under countable unions.* Let $C_1, C_2, \ldots \in \mathcal{F}|_B$, which means there exist $A_1, A_2, \ldots \in \mathcal{F}$ such that $C_n = A_n \cap B$ for each n. We know that $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$, so we observe that

      $$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} (B \cap A_n) = B \cap \bigcup_{n=1}^{\infty} A_n = B \cap A \in \mathcal{F}|_B.$$

   b. Let us check that $\mathbb{P}(\cdot \mid B)$ satisfies Kolmogorov's axioms.

      i. *Nonnegativity.* We know that $\mathbb{P}(B) > 0$ and $\mathbb{P}(A) \geq 0$ for all $A \in \mathcal{F}$, so $\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\,\mathbb{P}(B) = \mathbb{P}(A)/\,\mathbb{P}(B) \geq 0$ for all $A \in \mathcal{F}|_B$.

      ii. *Unit measure.* We see that $\mathbb{P}(B \mid B) = \mathbb{P}(B \cap B)/\,\mathbb{P}(B) = \mathbb{P}(B)/\,\mathbb{P}(B) = 1$. Note that the sample space for $\mathbb{P}(\cdot \mid B)$ is B, not the original Ω. (This reflects the fact that we are restricting to only outcomes where B is true, i.e. *conditioning* on B.)

      iii. *Countable additivity.* Let $A_1, A_2, \ldots$ be pairwise disjoint events in $\mathcal{F}|_B$. Note that $A_n = A_n \cap B$ for every n. We check that

      $$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n \,\middle|\, B\right) = \frac{1}{\mathbb{P}(B)}\,\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \frac{1}{\mathbb{P}(B)} \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n \mid B).$$

6. **Independence and Pairwise Independence**

A collection of events $\{A_i\}_{i\in I}$ is said to be *pairwise independent* if $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j)$ for all distinct indices $i \neq j$.

You flip a fair coin 99 times, where the result of each flip is independent of all other flips. For $i = 1, \dots, 99$, let A_i be the event that the ith flip comes up heads. Let B be the event that in total, an *odd* number of heads are seen. Show that the events $A_1, \dots, A_{99}, B$ are pairwise independent but *not* independent.

Solution: A good choice of sample space is $\Omega = \{H, T\}^{99}$, the set of all 2^{99} possible configurations of 99 coin flips, or equivalently $\{0,1\}^{99}$, the set of all 99-bit binary strings. We first check that $A_1, \dots, A_{99}, B$ are pairwise independent: for $1 \leq i < j \leq 99$,

$$\mathbb{P}(A_i \cap A_j) = \frac{2^{97}}{2^{99}} = \frac{2^{98}}{2^{99}} \cdot \frac{2^{98}}{2^{99}} = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j).$$

We also see that $\mathbb{P}(B) = \frac{1}{2}$ by symmetry: there is a unique correspondence between outcomes with an odd number of flips and outcomes with an even number of flips, found by simply switching the result of the first flip. $\mathbb{P}(A_i \cap B) = \frac{1}{4}$ by the same argument, now applied to the outcomes in A_i. So, for $1 \leq i \leq 99$,

$$\mathbb{P}(A_i \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_i) \cdot \mathbb{P}(B).$$

However, $A_1, \dots, A_{99}, B$ are not mutually independent:

$$\mathbb{P}\left(\bigcap_{i=1}^{99} A_i \cap B\right) = \frac{1}{2^{99}} \neq \left(\frac{1}{2}\right)^{100} = \prod_{i=1}^{99} \mathbb{P}(A_i) \cdot \mathbb{P}(B).$$

This formalizes the idea that B is "determined by" the events $A_1, \dots, A_{99}$.