UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 1

Spring 2021

1. **Choosing from Any Jar Makes No Difference**

   Each of k jars contains w white and b black balls. A ball is randomly chosen from jar 1 and transferred to jar 2, then a ball is randomly chosen from jar 2 and transferred to jar 3, etc. Finally, a ball is randomly chosen from jar k. Show that the probability that the last ball is white is the same as the probability that the first ball is white, i.e., it is $w/(w+b)$.

   Solution:

   We derive a recursion for the probability p_i that a white ball is chosen from the ith jar. We have, using the total probability theorem,

   $$p_{i+1} = \frac{w+1}{w+b+1} p_i + \frac{w}{w+b+1}(1-p_i) = \frac{1}{w+b+1} p_i + \frac{w}{w+b+1},$$

   starting with the initial condition $p_1 = w/(w+b)$. Thus, we have

   $$p_2 = \frac{1}{w+b+1} \cdot \frac{w}{w+b} + \frac{w}{w+b+1} = \frac{w}{w+b}.$$

   More generally, this calculation shows that if $p_{i-1} = w/(w+b)$, then $p_i = w/(w+b)$. Thus, we obtain $p_i = w/(w+b)$ for all i.

2. **Coin Flipping & Symmetry**

   Alice and Bob have $2n+1$ fair coins (where $n \geq 1$), each coin with probability of heads equal to $1/2$. Bob tosses $n+1$ coins, while Alice tosses the remaining n coins. Assuming independent coin tosses, show that the probability that, after all coins have been tossed, Bob will have gotten more heads than Alice is $1/2$.

   Hint: Consider the event $A = \{\text{more heads in the first } n+1 \text{ tosses than the last } n \text{ tosses}\}$.

   Solution:

   If we let Ω be the sample space consisting of all possible $2n+1$ tosses, then Ω is a uniform probability space by assumption. Define the events

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

   By symmetry, $P(A) = P(B)$, and we note that $A \cap B = \varnothing$ since it is impossible for the first $n+1$ tosses to have more heads *and* more tails than the last n tosses, and $A \cup B = \Omega$. So, $P(A) + P(B) = 1$ and hence $P(A) = 1/2$.

   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, and the probability that they are tied after n tosses is $1-2p$. So, the probability that Bob wins is $p+(1/2)(1-2p)=1/2$ since Bob can either win 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 and then flipping heads on the last toss.

3. **Passengers on a Plane**

   There are N passengers in a plane with N assigned seats (N is a positive integer), but after boarding, the passengers take the seats randomly. Assuming all seating arrangements are equally likely, what is the probability that no passenger is in their assigned seat? Compute the probability when $N \to \infty$.

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

   Solution:

   First, let us calculate the probability that at least one passenger sits in his or her assigned seat using inclusion-exclusion. Let A_i, $i=1,\dots,N$, be the event that passenger i sits in his or her assigned seat. We first add the probabilities of the single events (of which there are N), and the probability of each event is $(N-1)!/N!$ (indeed there are $(N-1)!$ ways to permute the remaining passengers once a specific passenger is fixed, and $N!$ total permutations, so the probability is $(N-1)!/N!$); next, we subtract the probabilities of the pairwise intersections of events (of which there are $\binom{N}{2}$), and the probability of each event is $(N-2)!/N!$ (there are $(N-2)!$ ways to permute the passengers other than the fixed two); continuing on, we see that

   $$P\Big(\bigcup_{i=1}^{N} A_i\Big) = \sum_{j=1}^{N}(-1)^{j+1}\binom{N}{j}\frac{(N-j)!}{N!} = \sum_{j=1}^{N}(-1)^{j+1}\frac{1}{j!}.$$

   Now, the event that no passenger sits in his or her assigned seat is the complement of the event just discussed:

   $$1-P\Big(\bigcup_{i=1}^{N} A_i\Big) = 1-\sum_{j=1}^{N}(-1)^{j+1}\frac{1}{j!} = \sum_{j=0}^{N}\frac{(-1)^j}{j!}.$$

   Taking the limit as $N \to \infty$, the expression converges to $\sum_{j=0}^{\infty}(-1)^j/j!$, and using the expression for the power series of the exponential function, we conclude that the probability converges to $\exp(-1) \approx 0.368$.

4. **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 is a positive integer), one at a time. 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. The cities are interviewed in a uniformly random order. What is the probability that the best city is selected? Assume that the commissioner has an objective method of scoring each city and that each city receives a unique score.

   You should arrive at an exact answer for the probability in terms of a summation. 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 can also say $\ln(n-1) \approx \ln n$.

Hint: Consider the events $B_i = \{\text{i-th city is the best}\}$ and $A = \{\text{best city is chosen}\}$.

Solution:

Let B_i, $i = 1, \ldots, N$, be the event that the ith city is the best of the N cities, and let A be the event that the best city is picked by the commissioner. Then,

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

using the law of total probability. Next, $P(A \mid B_i) = 0$ for $i = 1, \ldots, m$ because if the best city is among the first m cities, there is no chance of picking the best city. Also, $P(A \mid B_i) = m/(i-1)$ for $i = m+1, \ldots, N$ because $P(A \mid B_i)$ is the probability that second-best city among the first i cities is within the first m cities. Therefore,

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

Now, we turn towards approximation.

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

Letting $x := m/N$, then $P(A) \approx -x\ln x$, and differentiating with respect to x suggests that the optimal value is $x = 1/\mathrm{e}$, so we should reject the first N/e cities. Plugging in this value for x into $P(A) = -x\ln x$ gives the optimal probability as $P(A) \approx 1/\mathrm{e}$.

Note: This problem is a famous example from optimal stopping theory and is commonly known as the secretary problem (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.

5. **Coupling: Choosing a Sample Space**

Consider a *random* graph with 100 vertices; since there are $\binom{100}{2}$ possible edges, there are $2^{\binom{100}{2}}$ possible graphs. So, $\Omega = \{\text{all } 2^{\binom{100}{2}} \text{ possible graphs}\}$ and we define the probability of a graph $G \in \Omega$ to be $p^{E(G)}(1-p)^{\binom{100}{2}-E(G)}$, where $E(G)$ is the number of edges present in the graph. We call this the **random graph with parameter** p (abbreviated $\mathrm{RG}(p)$).

A **triangle** is a set of three vertices in the graph with all possible edges between them. Prove that the probability that there are at least ten triangles in the random graph is greater in $\mathrm{RG}(p)$ than under $\mathrm{RG}(q)$, if $p > q$. [*Hint*: Think about using a sequence of two coin flips to generate $RG(q)$ followed by $RG(p)$ on the same vertex set.]

Solution:

We will define a new sample space, which corresponds to a particular method of generating such a random graph: for each possible edge in the graph, we assign it a random number in the interval $[0,1]$, and then we define $\mathrm{RG}(p)$ to be the set of edges assigned a value $< p$. Formally, the new sample space is $\Omega = [0,1]^{\binom{100}{2}}$, that is, an outcome $\omega \in \Omega$ is a set of $\binom{100}{2}$

numbers between 0 and 1. Think about why this sample space is a valid way of generating RG(p)!

The reason why this helps us solve the problem is that under this new sample space, RG(p) actually *contains* RG(q). If an edge appears in RG(q), then the edge is assigned a value $< q$; but then it is certainly assigned a value $< p$ (since $p > q$) and so it appears in RG(p). So, if we let

$$E_p = \text{the event that there are at least ten triangles in } \mathrm{RG}(p),$$

then $E_p \supseteq E_q$, because if there are at least ten triangles in RG(q), then there must be at least ten triangles in RG(p). Hence, $P(E_p) \geq P(E_q)$, which is what we wanted to show!

This is a technique known as **coupling**.