CS 70 Spring 2022 — Discrete Mathematics and Probability Theory — Koushik Sen and Satish Rao — HW 8

1 Probability Warm-Up

(a) Suppose that we have a bucket of 30 red balls and 70 blue balls. If we pick 20 balls uniformly out of the bucket, what is the probability of getting exactly k red balls (assuming $0 \le k \le 20$) if the sampling is done **with** replacement, i.e. after we take a ball out the bucket we return the ball back to the bucket for the next round?

(b) Same as part (a), but the sampling is **without** replacement, i.e. after we take a ball out the bucket we **do not** return the ball back to the bucket.

(c) If we roll a regular, 6-sided die 5 times. What is the probability that at least one value is observed more than once?

Solution:

(a) Let A be the event of getting exactly k red balls. Then treating all balls as distinguishable, we have a total of 100^{20} possibilities to draw a sequence of 20 balls. In order for this sequence to have exactly k red balls, we need to first assign them one of $\binom{20}{k}$ possible locations within the sequence. Once done so, we have 30^k ways of actually choosing the red balls, and 70^{20-k} possibilities for choosing the blue balls. Thus in total we arrive at

$$\mathbb{P}[A] = \frac{\binom{20}{k} \cdot 30^k \cdot 70^{20-k}}{100^{20}} = \binom{20}{k}\left(\frac{3}{10}\right)^k\left(\frac{7}{10}\right)^{20-k}.$$

(b) We note that the size of the sample space is now $\binom{100}{20}$, since we are choosing 20 balls out of a total of 100. To find $|A|$, we need to be able to find out how many ways we can choose k red balls and $20-k$ blue balls. So we have that $|A| = \binom{30}{k}\binom{70}{20-k}$. So

$$\mathbb{P}[A] = \frac{\binom{30}{k}\binom{70}{20-k}}{\binom{100}{20}}.$$

(c) Let B be the event that at least one value is observed more than once. We see that $\mathbb{P}[B] = 1 - \mathbb{P}[\overline{B}]$. So we need to find out the probability that the values of the 5 rolls are distinct. We see that $\mathbb{P}[\overline{B}]$ simply the number of ways to choose 5 numbers (order matters) divided by the sample space (which is 6^5). So

$$\mathbb{P}[\overline{B}] = \frac{6!}{6^5} = \frac{5!}{6^4}.$$

So,

$$\mathbb{P}[B] = 1 - \frac{5!}{6^4}.$$

2 Past Probabilified

In this question we review some of the past CS70 topics, and look at them probabilistically.
For the following experiments,

i. Define an appropriate sample space Ω.

ii. Give the probability function $\mathbb{P}[\omega]$.

iii. Compute $\mathbb{P}[E_1]$.

iv. Compute $\mathbb{P}[E_2]$.

(a) Fix a prime $p > 2$, and uniformly sample twice with replacement from $\{0, \ldots, p-1\}$ (assume we have two $\{0, \ldots, p-1\}$-sided fair dice and we roll them). Then multiply these two numbers with each other in $(\bmod p)$ space.
$E_1 =$ The resulting product is 0.
$E_2 =$ The product is $(p-1)/2$.

(b) Make a graph on n vertices by sampling uniformly at random from all possible edges, (assume for each edge we flip a coin and if it is head we include the edge in the graph and otherwise we exclude that edge from the graph).
$E_1 =$ The graph is complete.
$E_2 =$ vertex v_1 has degree d.

(c) Create a random stable matching instance by having each person's preference list be a random permutation of the opposite entity's list (make the preference list for each individual job and each individual candidate a random permutation of the opposite entity's list). Finally, create a uniformly random pairing by matching jobs and candidates up uniformly at random (note that in this pairing, (1) a candidate cannot be matched with two different jobs, and a job cannot be matched with two different candidates (2) the pairing does not have to be stable).
$E_1 =$ All jobs have distinct favorite candidates.
$E_2 =$ The resulting pairing is the candidate-optimal stable pairing.

Solution:

(a) i. This is essentially the same as throwing two $\{0, \ldots, p-1\}$-sided dice, so one appropriate sample space is $\Omega = \{(i, j) \ : \ i, j \in \mathrm{GF}(p)\}$.

ii. Since there are exactly p^2 such pairs, the probability of sampling each one is $\mathbb{P}[(i, j)] = 1/p^2$.

iii. Now in order for the product $i \cdot j$ to be zero, at least one of them has to be zero. There are exactly $2p-1$ such pairs, and so $\mathbb{P}[E_1] = \frac{2p-1}{p^2}$.

iv. For $i \cdot j$ to equal $(p-1)/2$ it doesn't matter what i is as long as $i \neq 0$ and $j \equiv i^{-1}(p-1)/2 \pmod{p}$. Thus $|E_2| = \left|\left\{(i,j) \;:\; j \equiv i^{-1}(p-1)/2\right\}\right| = p-1$, and whence $\mathbb{P}[E_2] = \frac{p-1}{p^2}$.
Alternative Solution for $\mathbb{P}[E_2]$*:* The previous reasoning showed that $(p-1)/2$ is in no way special, and the probability that $i \cdot j = (p-1)/2$ is the same as $\mathbb{P}[i \cdot j = k]$ for any $k \in \mathrm{GF}(p)$. But $1 = \sum_{k=0}^{p-1} \mathbb{P}[i \cdot j = k] = \mathbb{P}[i \cdot j = 0] + (p-1)\mathbb{P}[i \cdot j = (p-1)/2] = \frac{2p-1}{p^2} + (p-1)\mathbb{P}[i \cdot j = (p-1)/2]$, and so $\mathbb{P}[E_2] = \left(1 - \frac{2p-1}{p^2}\right)/(p-1) = \frac{p-1}{p^2}$ as desired.

(b) i. Since any n-vertex graph can be sampled, Ω is the set of all graphs on n vertices.

ii. As there are $N = 2^{\binom{n}{2}}$ such graphs, the probability of each indivdual one g is $\mathbb{P}[g] = 1/N$ (by the same reasoning that every sequence of fair coin flips is equally likely!).

iii. There is only one complete graph on n vertices, and so $\mathbb{P}[E_1] = 1/N$.

iv. For vertex v_1 to have degree d, exactly d of its $n-1$ possible adjacent edges must be present. There are $\binom{n-1}{d}$ choices for such edges, and for any fixed choice, there are $2^{\binom{n}{2}-(n-1)}$ graphs with this choice. So $\mathbb{P}[E_2] = \frac{\binom{n-1}{d} 2^{\binom{n}{2}-(n-1)}}{2^{\binom{n}{2}}} = \binom{n-1}{d}\left(\frac{1}{2}\right)^{n-1}$.

(c) i. Here there are two random things we need to keep track of: The random preference lists and the random pairing. A person i's preference list can be represented as a permutation σ_i of $\{1,\ldots,n\}$, and the pairing itself is encoded in another permutation ρ of the same set (indicating that job i is paired with candidate $\rho(i)$). So $\Omega = \{(\sigma_1,\ldots,\sigma_{2n},\rho) \;:\; \sigma_i,\rho \in S_n\}$, where S_n is the set of permutations of $\{1,\ldots,n\}$.

ii. $|\Omega| = (n!)^{2n+1}$, and so $\mathbb{P}[\mathscr{P}] = 1/|\Omega|$ for each $\mathscr{P} \in \Omega$.

iii. For E_1, we observe that there are $n!$ possible configurations of all jobs having distinct favourite candidates, and that each job has $(n-1)!$ ways of ordering their non-favourite candidates, so $|E_1| = \underbrace{n!}_{\text{distinct favourites}} \cdot \underbrace{[(n-1)!]^n}_{\text{ordering of non-favourites}} \cdot \underbrace{(n!)^n}_{\text{candidate's preferences}} \cdot \underbrace{n!}_{\rho}$. Consequently, $\mathbb{P}[E_1] = n!\left(\frac{(n-1)!}{n!}\right)^n = \frac{n!}{n^n}$.

iv. No matter what $\sigma_1,\ldots,\sigma_{2n}$ are, there is exactly one candidate-optimal pairing, and so $\mathbb{P}[E_2] = \frac{(n!)^{2n}}{(n!)^{2n+1}} = \frac{1}{n!}$.

3 Peaceful rooks

A friend of yours, Eithen Quinn, is fascinated by the following problem: placing m rooks on an $n \times n$ chessboard, so that they are in peaceful harmony (i.e. no two threaten each other). Each rook is a chess piece, and two rooks threaten each other if and only if they are in the same row or column. You remind your friend that this is so simple that a baby can accomplish the task. You forget however that babies cannot understand instructions, so when you give the m rooks to your baby niece, she simply puts them on random places on the chessboard. She however, never puts two rooks at the same place on the board.

(a) Assuming your niece picks the places uniformly at random, what is the chance that she places the $(i+1)^{\text{st}}$ rook such that it doesn't threaten any of the first i rooks, given that the first i rooks don't threaten each other?

(b) What is the chance that your niece actually accomplishes the task and does not prove you wrong?

(c) Now imagine that the rooks can be stacked on top of each other, then what would be the probability that your niece's placements result in peace? Assume that two rooks threaten each other if they are in the same row or column. Also two pieces stacked on top of each other are obviously in the same row and column, therefore they threaten each other.

(d) Explain the relationship between your answer to the previous part and the birthday paradox. In particular if we assume that 23 people have a 50% chance of having a repeated birthday (in a 365-day calendar), what is the probability that your niece places 23 stackable pieces in a peaceful position on a 365×365 board?

Solution:

(a) After having placed i rooks in a peaceful position, i of the rows and i of the columns are taken. So for the next rook we have $n-i$ choices for the row and $n-i$ choices for the column in order to remain in a peaceful position. The total number of board cells left is n^2-i. So the chance that the next rook keeps the peace is $\frac{(n-i)^2}{n^2-i}$.

(b) The product over $i = 0, \ldots, m-1$ gives us the final answer. So the answer is

$$\prod_{i=0}^{m-1} \frac{(n-i)^2}{n^2-i} = \frac{(n!)^2(n^2-m)!}{(n^2)!((n-m)!)^2}$$

(c) The only thing that changes from the previous part is that when placing the i-th piece, we no longer have n^2-i possibilities, but n^2 possibilities. So the answer changes to

$$\prod_{i=0}^{m-1} \frac{(n-i)^2}{n^2} = \frac{(n!)^2}{((n-m)!)^2 n^{2m}}$$

(d) l the columns must be different.

All the rows being different is simply the birthday paradox. Similarly all the columns being different is another birthday paradox. So if the probability that m persons have different birthdays in an n-day calendar is p, then the probability that m rooks end up in a peaceful position on an $n \times n$ chessboard is p^2. Of course this can be verified by hand. The answer to the previous part is

$$\frac{(n!)^2}{((n-m)!)^2 n^{2m}} = \left(\frac{n!}{(n-m)!n^m}\right)^2$$

The expression inside the parenthesis is the answer to the birthday paradox.

So if the probability p is 0.5 (which roughly happens for $n = 365$ and $m = 23$), then the probability that rooks end up in a peaceful position is $p^2 = 0.25$.

4 Five Up

Say you toss a coin five times, and record the outcomes. For the three questions below, you can assume that order matters in the outcome, and that the probability of heads is some p in $0 < p < 1$, but *not* that the coin is fair ($p = 0.5$).

(a) What is the size of the sample space, $|\Omega|$?

(b) How many elements of Ω have exactly three heads?

(c) How many elements of Ω have three or more heads?
(*Hint:* Argue by symmetry.)

For the next three questions, you can assume that the coin is fair (i.e. heads comes up with $p = 0.5$, and tails otherwise).

(d) What is the probability that you will observe the sequence HHHTT? What about HHHHT?

(e) What is the chance of observing at least one head?

(f) What about the chance of observing three or more heads?

For the final three questions, you can instead assume the coin is biased so that it comes up heads with probability $p = \frac{2}{3}$.

(g) What is the chance of observing the outcome HHHTT? What about HHHHT?

(h) What about the chance of at least one head?

(i) What about the chance of ≥ 3 heads?

Solution:

(a) Since for each coin toss, we can have either heads or tails, we have 2^5 total possible outcomes.

(b) Since we know that we have exactly 3 heads, what distinguishes the outcomes is at which point these heads occurred. There are 5 possible places for the heads to occur, and we need to choose 3 of them, giving us the following result: $\binom{5}{3}$.

(c) We can use the same approach from part (b), but since we are asking for 3 or more, we need to consider the cases of exactly 4 heads, and exactly 5 heads as well. This gives us the result as: $\binom{5}{3}+\binom{5}{4}+\binom{5}{5}=16$.
To see why the number is exactly half of the total number of outcomes, denote the set of outcomes that has 3 or more heads as A. If we flip over every coin in each outcome in set A, we get all the outcomes that has 2 or less head. Denote the new set as A'. Then we know that A and A' have the same size and they together cover the whole sample space. Therefore, $|A|=|A'|$ and $|A|+|A'|=2^5$, which gives $|A|=2^5/2$.

(d) Since each coin toss is an independent event, the probability of each of the coin tosses is $\frac{1}{2}$ making the probability of this outcome $\frac{1}{2^5}$. This holds for both cases since both heads and tails have the same probability.

(e) We will use the complementary event, which is the event of getting no heads. The probability of getting no heads is the probability of getting all tails. This event has a probability of $\frac{1}{2^5}$ by a similar argument to the previous part. Since we are asking for the probability of getting at least one heads, our final result is: $1-\frac{1}{2^5}$.

(f) Since each outcome in this probability space is equally likely, we can divide the number of outcomes where there are 3 or more heads by the total number of outcomes to give us: $\frac{\binom{5}{3}+\binom{5}{4}+\binom{5}{5}}{2^5}$

(g) By using the same idea of independence we get for HHHTT: $\frac{1}{3}\times\frac{1}{3}\times\frac{2}{3}\times\frac{2}{3}\times\frac{2}{3}=\frac{2^3}{3^5}$
For HHHHT, we get:
$\frac{1}{3}\times\frac{2}{3}\times\frac{2}{3}\times\frac{2}{3}\times\frac{2}{3}=\frac{2^4}{3^5}$

(h) Similar to the unbiased case, we will first find the probability of the complement event, which is having no heads. The probability of this is $\frac{1}{3^5}$, which makes our final result $1-\frac{1}{3^5}$

(i) In this case, since we are working in a nonuniform probability space (getting 4 heads and 3 heads don't have the same probability), we need to separately consider the events with different numbers of heads to find our result. This will get us:

$$\binom{5}{3}\frac{2^3}{3^5}+\binom{5}{4}\frac{2^4}{3^5}+\binom{5}{5}\frac{2^5}{3^5}$$

5 Flipping Coins

Consider the following scenarios, where we apply probability to a game of flipping coins. In the game, we flip one coin each round. The game will not stop until two consecutive heads appear.

(a) What is the probability that the game ends by flipping exactly five coins?

(b) Given that the game ends after flipping the fifth coin, what is the probability that three heads appear in the sequence?

(c) If we change the rule that the game will not stop until three consecutive tails or three consecutive heads appear, what is the probability that the game stops by flipping at most six coins?

Solution:

(a) If the game ends by flipping exactly five coins, we know the flipping results of last three coins must be $\{T, H, H\}$. For the first two coins, the results can be $\{H, T\}, \{T, H\}$ or $\{T, T\}$. So the probability equals to $\frac{3}{4} \times \frac{1}{8} = \frac{3}{32}$.

(b) Given the condition that the game ends after flipping the fifth coin, the only possible sequances containing three heads are $\{H, T, T, H, H\}$ and $\{T, H, T, H, H\}$. Let A denote the event that the game ends after flipping the fifth coin, B denote the event that three heads appear when the game ends. Then $\mathbb{P}[A \cap B] = \frac{2}{32}$ and $\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]} = \frac{2}{3}$.

(c) We only consider the case that the game ends only because of three consecutive heads appearing. If the game ends in three rounds, the result must be $\{H, H, H\}$ and the probability is $\frac{1}{8}$. If the game ends in four rounds, the results must be $\{T, H, H, H\}$ and the probability is $\frac{1}{16}$. If the game ends in five rounds, the results must be $\{H, T, H, H, H\}$ or $\{T, T, H, H, H\}$ and the probability is $\frac{1}{16}$. If the game ends in six rounds, the last four coins must be $\{T, H, H, H\}$ and the first two coins cannot be $\{T, T\}$, so the probability is $\frac{3}{4} \times \frac{1}{16} = \frac{3}{64}$. The total probability is

$$\frac{1}{8} + 2 \times \frac{1}{16} + \frac{3}{64} = \frac{19}{64}$$

. By symmetry, the probability that the game ends in the sixth round because of three consecutive tails is also $\frac{19}{64}$. So the final answer is

$$2 \times \frac{19}{64} = \frac{19}{32}$$

.

6 Lie Detector

A lie detector is known to be $4/5$ reliable when the person is guilty and $9/10$ reliable when the person is innocent. If a suspect is chosen from a group of suspects of which only $1/100$ have ever committed a crime, and the test indicates that the person is guilty, what is the probability that he is guilty?

Solution:

Let A denote the event that the test indicates that the person is guilty, and B the event that the person is actually guilty. Note that

$$\mathbb{P}[B] = \frac{1}{100}, \quad \mathbb{P}[\overline{B}] = \frac{99}{100}, \quad \mathbb{P}[A \mid B] = \frac{4}{5}, \quad \mathbb{P}[A \mid \overline{B}] = \frac{1}{10}.$$

By Bayes' Rule and the Total Probability Rule the desired probability is

$$\mathbb{P}[B \mid A] = \frac{\mathbb{P}[B]\mathbb{P}[A \mid B]}{\mathbb{P}[A]} = \frac{\mathbb{P}[B]\mathbb{P}[A \mid B]}{\mathbb{P}[B]\mathbb{P}[A \mid B] + \mathbb{P}[\overline{B}]\mathbb{P}[A \mid \overline{B}]} = \frac{(1/100)(4/5)}{(1/100)(4/5) + (99/100)(1/10)} = \frac{8}{107}$$

7 PIE Extended

One interesting result yielded by the Principle of Inclusion and Exclusion (PIE) is that for any events $A_1, A_2, \ldots, A_n$ in some probability space,

$$\sum_{i=1}^{n} \mathbb{P}[A_i] - \sum_{i<j\leq n} \mathbb{P}\left[A_i \cap A_j\right] + \sum_{i<j<k\leq n} \mathbb{P}\left[A_i \cap A_j \cap A_k\right] - \cdots + (-1)^{n-1}\mathbb{P}[A_1 \cap A_2 \cap \cdots \cap A_n] \geq 0$$

(Note the LHS is equal to $\mathbb{P}[\bigcup_{i=1}^{n} A_i]$ by PIE, and probability is nonnegative).

Prove that for any events $A_1, A_2, \ldots, A_n$ in some probability space,

$$\sum_{i=1}^{n} \mathbb{P}[A_i] - 2\sum_{i<j\leq n} \mathbb{P}\left[A_i \cap A_j\right] + 4\sum_{i<j<k\leq n} \mathbb{P}\left[A_i \cap A_j \cap A_k\right] - \cdots + (-2)^{n-1}\mathbb{P}[A_1 \cap A_2 \cap \cdots \cap A_n] \geq 0$$

(Hint: consider defining an event B to represent "an odd number of $A_1, \ldots, A_n$ occur")

Solution:

For events $A_1, A_2, \ldots, A_n$ on a probability space, let B represent the event "an odd number of $A_1, \ldots, A_n$ occur."

Upon showing

$$\mathbb{P}[B] = \sum_{i=1}^{n} \mathbb{P}[A_i] - 2\sum_{i<j\leq n} \mathbb{P}\left[A_i \cap A_j\right] + 4\sum_{i<j<k\leq n} \mathbb{P}\left[A_i \cap A_j \cap A_k\right] - \cdots + (-2)^{n-1}\mathbb{P}[A_1 \cap \cdots \cap A_n]$$

the desired result then follows immediately from nonnegativity of probability.

We proceed to show this equality by combinatorial argument, analogous to the proof of the Principle of Inclusion-Exclusion (likewise, one could also proceed via induction). For any outcome ω, consider the index set $M \subseteq \{1, \ldots, n\}$ such that

$$M = \{i \in \{1, \ldots, n\} \mid \omega \in A_i\}$$

In other words, M is the set of all indices i such that outcome ω is in event A_i. Let $|M| = m$. For any $S \subseteq M$, $\omega \in \bigcap_{i\in S} A_i$. Since there are $\binom{m}{k}$ possible subsets S of size k, and each corresponding probability term has a multiplicative factor of $(-2)^{k-1}$ on the RHS, we have that the probability of outcome ω is counted a total of

$$\begin{aligned}
\sum_{k=1}^{m} \sum_{S\subseteq M, |S|=k} (-2)^{k-1} &= \sum_{k=1}^{m} (-2)^{k-1}\binom{m}{k} \\
&= -\frac{1}{2}\left[\sum_{k=0}^{m} (-2)^{k}\binom{m}{k} - (-2)^0\binom{m}{0}\right] \\
&= -\frac{1}{2}\left[(1+(-2))^m - 1\right] \\
&= \frac{1-(-1)^m}{2} \\
&= \begin{cases} 1, & m \text{ is odd} \\ 0, & \text{otherwise} \end{cases}
\end{aligned}$$

times on the RHS. The second line is algebraic manipulation, and the third line uses the binomial theorem.

Therefore, each outcome ω that is in an odd number of events $A_1, \ldots, A_n$ has its probability counted exactly once and every other outcome's probability is zeroed out in the RHS, concluding the proof.