CS 70 Discrete Mathematics and Probability Theory
Fall 2021 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 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:

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

3 Easter Eggs

You made the trek to Soda for a Spring Break-themed homework party, and every attendee gets to leave with a party favor. You're given a bag with 20 chocolate eggs and 40 (empty) plastic eggs. You pick 5 eggs (uniformly) without replacement.

(a) What is the probability that the first egg you drew was a chocolate egg?

(b) What is the probability that the second egg you drew was a chocolate egg?

(c) Given that the first egg you drew was an empty plastic one, what is the probability that the fifth egg you drew was also an empty plastic egg?

Solution:

(a) $\mathbb{P}(\text{chocolate egg}) = \frac{20}{60} = \frac{1}{3}$.

(b) Long calculation using Total Probability Rule: let C_i denote the event that the ith egg is chocolate, and P_i denote the event that the ith egg is plastic. We have

$$\begin{aligned}\mathbb{P}(C_2) &= \mathbb{P}(C_1 \cap C_2) + \mathbb{P}(P_1 \cap C_2) \\ &= \mathbb{P}(C_1)\mathbb{P}(C_2 \mid C_1) + \mathbb{P}(P_1)\mathbb{P}(C_2 \mid P_1) \\ &= \frac{1}{3} \cdot \frac{19}{59} + \frac{2}{3} \cdot \frac{20}{59} \\ &= \frac{1}{3}.\end{aligned} \tag{1}$$

Short calculation: By symmetry, this is the same probability as part (a), $1/3$. This is because we don't know what type of egg was picked on the first draw, so the distribution for the second egg is the same as that of the first. To see this rigorously observe that $\mathbb{P}[C_2 \cap P_1] = \mathbb{P}[P_2 \cap C_1]$ and, thus:

$$\begin{aligned}\mathbb{P}[C_2] &= \mathbb{P}[C_2 \cap C_1] + \mathbb{P}[C_2 \cap P_1] \\ &= \mathbb{P}[C_2 \cap C_1] + \mathbb{P}[P_2 \cap C_1] \\ &= \mathbb{P}[C_1] \ .\end{aligned}$$

(c) By symmetry, since we don't know any information about the 2nd, 3rd, or 4th eggs, $\mathbb{P}(\text{5th egg = plastic} \mid \text{1st egg = plastic}) = \mathbb{P}(\text{2nd egg = plastic} \mid \text{1st egg = plastic}) = 39/59$. Rigorously, notice that $\mathbb{P}[C_5 \cap P_2 \mid P_1] = \mathbb{P}[P_5 \cap C_2 \mid P_1]$ and therefore:

$$\begin{aligned}\mathbb{P}[P_5 \mid P_1] &= \mathbb{P}[P_5 \cap C_2 \mid P_1] + \mathbb{P}[P_5 \cap P_2 \mid P_1] \\ &= \mathbb{P}[C_5 \cap P_2 \mid P_1] + \mathbb{P}[P_5 \cap P_2 \mid P_1] \\ &= \mathbb{P}[P_2 \mid P_1] \ .\end{aligned}$$

One could also brute force this with Total Probability Rule (like in the previous part), but the calculation is quite tedious.

4 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 $\mathbb{P}(E_1)$ given event E_1.

iv. Compute $\mathbb{P}(E_2)$ given event 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 $\mathbb{P}(E_2) = \frac{(n!)^{2n}}{(n!)^{2n+1}} = \frac{1}{n!}$.

5 Ball-and-Bin Counting Problems

Say you have 5 bins, and randomly throw 7 balls into them.

1. What is the probability that the first bin has precisely 3 balls in it?
2. What is the probability that the third bin has at least 3 balls in it?
3. What is the probability that at least one of the bins has precisely 3 balls in it?

Solution:

1. First, choose 3 balls out of 7 to put in the first bin. Each of these 3 balls has $Pr = 1/5$ to be in the first bin, and each of the rest $7-3=4$ balls has $Pr = 4/5$ to NOT be in the first bin. Thus, the answer is

$$\binom{7}{3}\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^4.$$

2. Extend part (a) results into the sum of exactly 3/4/5/6/7 balls

$$\sum_{k=3}^{7} \binom{7}{k} \left(\frac{1}{5}\right)^{k} \left(\frac{4}{5}\right)^{7-k}.$$

3. Use inclusion-exclusion. Let A_i be the event that bin i has exactly 3 balls. Then $\sum_{i=1}^{5} \mathbb{P}[A_i] = 5\binom{7}{3}(1/5)^3(4/5)^4$. We have to subtract the events $A_i \cap A_j$, of which there are $\binom{5}{2}$. We have $\mathbb{P}[A_i \cap A_j] = 7!/(3!)^2(1/5)^6(3/5)$.

   The reasoning behind $\frac{7!}{3!3!}$ is because: we want the 7 balls to be split in exactly 3 balls in A_i, 3 balls in A_j, and last ball wherever else. First assume that all 7 balls are lined up, and A_i takes first three balls, A_j takes the next three balls, so you have 7! ways to order all the balls initially. Then, reduce the over-counted parts since the order of the 3 balls in A_i doesn't matter (divide by 3! since there are this many ways to order the 3 balls), and similarly, the order of the 3 balls in A_j doesn't matter either. Thus, you have $\frac{7!}{3!3!}$.

   Therefore our answer is

$$5\binom{7}{3}\left(\frac{1}{5}\right)^{3}\left(\frac{4}{5}\right)^{4} - \binom{5}{2}\frac{7!}{3!3!}\left(\frac{1}{5}\right)^{6}\frac{3}{5}.$$