More Practice Problems

Remember that the final for this class will take place on Thursday, not Friday. The final is during class time in the usual location. There will be no class on Friday.

1. You have a biased coin, which shows heads with probability $0 < p < 1$ and tails with probability $1 - p$.

   (a) How can you simulate a fair coin? The only requirement is that the expected number of times you flip the coin is finite.

   (b) In expectation, how many coin tosses are needed to simulate one fair toss?

   You may need to use the formula for the sum of an infinite geometric series to solve this problem.

   The sum of an infinite geometric series $a + a \cdot r + a \cdot r^2 + a \cdot r^3 \ldots$ is $\frac{a}{1-r}$ when $|r| < 1$.

   Solution:

   (a) Look at the results after two consecutive flips of this coin: HT, TH, HH, TT. It is clear that HT and TH are equally likely (both have probability $p \cdot (1 - p)$). So, what one can do to simulate a fair coin is the following. Flip this biased coin two times. If the result is one of HT, TH, then assume HT is heads and TH is tails. If the result is something else, repeat.

   (b) Let X be the random variable that denotes the number of two flips needed to simulate the fair coin. For $i \in \mathbb{Z}^+$, let X_i be the indicator random variable for the event that we need at least i two flips to terminate the process. Note that $X = \sum_{i=1}^{\infty} X_i$.

   Note that in order of the process to terminate, we must either get a HT or TH. Let the probability that we get a HT or TH be q.

   By the Linearity of Expectation, we have:

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{i=1}^{\infty} \mathbf{E}[X_i] \\
&= \sum_{i=1}^{\infty} \Pr[X_i = 1]
\end{aligned}$$

   But what is $\Pr[X_i = 1]$? Note that it is not the same value for any particular i – it is 1 when $i = 1$ and much much smaller when $i = n$. But the goal is to still be able to determine this for an arbitrary i.

$$
\begin{aligned}
\Pr[X_i = 1] = \ & \Pr[X_i = 1 | (X_1 = 1) \cap (X_2 = 1) \cap \cdots \cap (X_{i-1} = 1)] \\
& \cdot \Pr[X_{i_1} = 1 | (X_1 = 1) \cap (X_2 = 1) \cap \cdots \cap (X_{i-2} = 1)] \\
& \cdot \cdot \cdot \cdot \cdot \\
& \cdot \Pr[X_2 = 1 | X_1 = 1] \cdot \Pr[X_1 = 1]
\end{aligned}
$$

Note that $((X_1 = 1) \cap (X_2 = 1) \cap \cdots \cap (X_k = 1)) \equiv (X_k = 1)$:

$$
\begin{aligned}
= \ & \Pr[X_i = 1 | X_{i-1} = 1] \cdot \Pr[X_{i_1} = 1 | X_{i-2} = 1] \cdot \ldots \\
& \cdot \Pr[X_2 = 1 | X_1 = 1] \cdot \Pr[X_1 = 1]
\end{aligned}
$$

Note that the probability $\Pr[X_k = 1 | X_{k-1} = 1] = 1 - q$ since that is the probability that we do not get HT or TH:

$$
= \ (1-q)^{(i-1)}
$$

So we have that:

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{i=1}^{\infty} \Pr[X_1 = i] \\
&= \sum_{i=1}^{\infty} (1-q)^{(i-1)}
\end{aligned}
$$

Since $|1 - q| < 1$, we can use the sum of an infinite geometric series:

$$
\begin{aligned}
&= \sum_{i=1}^{\infty} (1-q)^{(i-1)} \\
&= \frac{1}{q}
\end{aligned}
$$

But what is q? From (a) we determined that the probability of HT and TH are both $p \cdot (1-p)$. Adding them together, we have that $q = 2p(1-p)$. Hence we have that $\mathbf{E}[X] = \frac{1}{2p(1-p)}$.

Finally, note that $\mathbf{E}[X]$ is the number of two coin flips needed to simulate a fair coin flip. The question asks for the expected number of single coin flips needed. The answer is therefore $2 \cdot \mathbf{E}[X]$.

2. You want to determine your lucky number. To do this, you flip 100 fair coins. You total the number of heads, H, and number of tails, T. You denote your lucky number to be $H - T$. What is the expected value of your lucky number?

Solution:

We seek $\mathbf{E}[H-T]$. By LOE, we have that $\mathbf{E}[H-T] = \mathbf{E}[H] - \mathbf{E}[T]$.

Let us determine $\mathbf{E}[H]$. Let us define the indicator random variables H_i for the event that we get a heads on flip i. Note that $H = \sum_i H_i$ and $\Pr[H_i = 1] = \frac{1}{2}$ for any i.

By LOE:

$$\begin{aligned}
\mathbf{E}[H] &= \sum_{i=1}^{100} \mathbf{E}[H_i] \\
&= \sum_{i=1}^{100} \Pr[H_i = 1] \\
&= \sum_{i=1}^{100} \frac{1}{2} \\
&= 50
\end{aligned}$$

Similarly we have that $\mathbf{E}[T] = 50$.

Hence, the expected value of your lucky number is $\mathbf{E}[H-T] = 50 - 50 = 0$.

3. Let $f : X \to Y$ be some function. If $S \subseteq X$, then define $f(S) = \{f(s) \mid s \in S\}$. Let $A, B \subseteq X$. Prove that $A \subseteq B$ implies $f(A) \subseteq f(B)$.

   Solution:

   Assume $A \subseteq B$. We would like to show $f(A) \subseteq f(B)$. Let y be an arbitrary element in $f(A)$. Then, we know there exists some $x \in A$ such that $f(x) = y$. Since $A \subseteq B$, $x \in B$. It follows that $y \in f(B)$ by definition of $f(B)$, and we are done.

4. Define a relation R on $\mathbb{N}$ where $(a, b) \in R$ if and only if a and b have no positive common factors other than 1. For each of the 5 properties of relations (reflexivity, symmetry, transitivity, anti-symmetry, irreflexivity) that we have studied, state and prove whether R has this property.

   Solution:

   Reflexivity: R is NOT reflexive.

   R would be reflexive iff $\forall x \in \mathbb{N}, (x, x) \in R$. However, consider $x = 2$. $(2, 2) \notin R$ since 2 and 2 DO have a positive common factor other than 1, namely 2. Thus, R is not reflexive.

   Irreflexivity: R is NOT irreflexive.

   R would be irreflexive iff $\forall x \in \mathbb{N}, (x, x) \notin R$. However, consider $x = 1$. $(1, 1) \in R$ because the only positive common factor between 1 and 1 is 1. Thus, R is not irreflexive.

   Symmetry: R IS symmetric.

Let $(x, y) \in R$ be arbitrary. Then x and y have no positive common factors other than 1. This means that y and x have no positive common factors other than 1. It follows that $(y, x) \in R$. Therefore, $(x, y) \in R \implies (y, x) \in R$. Thus, R is symmetric.

Antisymmetry: R is NOT antisymmetric.

R would be antisymmetric iff $(x, y) \in R \wedge (y, x) \in R \implies x = y$. However, consider $x = 2$ and $y = 3$. Since 2 and 3 have no positive common factors other than 1, it must be that $(2, 3) \in R$. Likewise, since 3 and 2 have no positive common factors other than 1, it must be that $(3, 2) \in R$. However, $2 \neq 3$. Thus, R is not antisymmetric.

Transitivity: R is NOT transitive.

R would be transitive iff $(x, y) \in R \wedge (y, z) \in R \implies (x, z) \in R$. However, consider $x = 2$, $y = 3$, $z = 4$. Since 2 and 3 have no positive common factors other than 1, it must be that $(2, 3) \in R$. Since 3 and 4 have no positive common factors other than 1, it must be that $(3, 4) \in R$. However, $(2, 4) \notin R$ since 2 and 4 DO have a positive common factor other than 1, namely 2. Thus, R is not transitive.

5. I dip a $3 \times 3 \times 3$ cube into paint so its entire surface is coated. I then disassemble the cube into 27 cubelets (of size $1 \times 1 \times 1$), select one uniformly at random, and place it in front of you on a table. From the five sides you can observe of the cubelet, no side is painted. What is the probability that the bottom side (that you cannot observe) is painted?

Solution:

Let the cube be C. Clearly the cubelet in front of you has no sides painted or exactly one side (the bottom side) painted. There are only seven cubelets that have this property (six cubelets on the center of each face of C, one cubelet at the center of C). However, each of these cubelets is not equally likely to be in front of you at this point. This is because if I picked up the center cubelet (no painted side), I could have placed it down on any of its six faces, and you would see all observable sides not painted. But if I picked up any of the other six cubelets, I have to place them down in exactly one way: with the painted side on the bottom. Thus, there are twelve total ways (outcomes), all equally likely, for me to pick up and place a cubelet in front of you when we distinguish different orientations of the cube. Since six of these twelve outcomes involved a cubelet with paint on the bottom side, the probability that the bottom has paint, and thus the answer to this question is $\frac{1}{2}$.

We can also do this problem with familiar notation. We also consider the following events.

A: event that the bottom side of the cubelet in front of you is painted.

B: event that all five observable sides of the cubelet are not painted.

C: event that we select the center cubelet.

D: event that we select the cubelet at the center of a face.

F: event that we select a cubelet that is not one of the above

Note that $A \cap B$ represents the event that you pick up one of the six cubelets with one side painted and also place it with its painted side down. Then:

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

Note that $A \cap B = B \cap D$:

$$= \frac{\Pr[B \cap D]}{\Pr[B \cap C] + \Pr[B \cap D] + \Pr[B \cap F]}$$
$$= \frac{\Pr[B|D] \cdot \Pr[D]}{\Pr[B|C] \cdot \Pr[C] + \Pr[B|D] \cdot \Pr[D] + \Pr[B|F] \cdot \Pr[F]}$$

Note that $\Pr[B|F] = 0$:

$$= \frac{\Pr[B|D] \cdot \Pr[D]}{\Pr[B|C] \cdot \Pr[C] + \Pr[B|D] \cdot \Pr[D]}$$

Note that $\Pr[B|D] = \frac{1}{6}$ (since there is one side that you can put down such that the 5 observable sides are not painted). Note that $\Pr[D] = \frac{6}{27}$ (since there are 6 such cubelets). Note that $\Pr[B|C] = 1$, and $\Pr[C] = \frac{1}{27}$.

$$\begin{aligned}\Pr[A|B] &= \frac{\Pr[B|D] \cdot \Pr[D]}{\Pr[B|C] \cdot \Pr[C] + \Pr[B|D] \cdot \Pr[D]} \\ &= \frac{\frac{1}{6} \times \frac{6}{27}}{\frac{1}{6} \times \frac{6}{27} + 1 \times \frac{1}{27}} \\ &= \frac{1}{2}\end{aligned}$$

6. You go to a vending machine with 11 different candy bars, but you only like one type. The vending machine only has one candy bar left of each type, and since it is broken, it randomly releases a candy bar every time you pay (it always releases a candy bar). You will get the candy bar you like on your nth try.

   (a) What is the expected value of n?

   (b) Now lets say that the vending machine has an infinite supply of each of the candy bars. Do you think the expected value of n now is higher or lower than before? What is the new expected value?

   You may need to use the formula for the sum of an infinite geometric series to solve this problem.

   The sum of an infinite geometric series $a + a \cdot r + a \cdot r^2 + a \cdot r^3 \ldots$ is $\frac{a}{1-r}$ when $|r| < 1$.

Solution:

(a) We seek $\mathbf{E}[n]$:

$$\mathbf{E}[n] = \sum_{k=1}^{11} \Pr[n = k] \times k$$

For any arbitrary k, we can construct the outcomes in $n = k$ as follows:

Step 1: Choose a position in the order for your favorite candy bar – 11 ways

Step 2: Permute the remaining 10 candy bars into the remaining ordering – 10! ways

Note that $|\Omega| = 11!$ since it is just a permutation of the 11 candy bars. Since the probability space is uniform, we have that $\Pr[n = k] = \frac{10!}{11!} = \frac{1}{11}$. Hence:

$$\begin{aligned}
\mathbf{E}[n] &= \sum_{k=1}^{11} \Pr[n = k] \times k \\
&= \sum_{k=1}^{11} \frac{1}{11} \times k \\
&= \frac{1}{11} \times \sum_{k=1}^{11} k \\
&= \frac{1}{11} \times \frac{11 \times 12}{2} = 6
\end{aligned}$$

(b) For $i \in \mathbb{Z}^+$, let n_i be the indicator random variable for the event that we need to get at least i candy bars to get the one that we like. Note that $n = \sum_{i=1}^{\infty} n_i$.

By the Linearity of Expectation, we have:

$$\begin{aligned}
\mathbf{E}[b] &= \sum_{i=1}^{\infty} \mathbf{E}[n_i] \\
&= \sum_{i=1}^{\infty} \Pr[n_i = 1]
\end{aligned}$$

But what is $\Pr[X_i = 1]$? Note that it is not the same value for any particular i – it is 1 when $i = 1$ and much much smaller when $i = 1000$. But the goal is to still be able to determine this for an arbitrary i.

$$
\begin{aligned}
\Pr[n_i = 1] = \ & \Pr[n_i = 1 | (n_1 = 1) \cap (n_2 = 1) \cap \cdots \cap (n_{i-1} = 1)] \\
& \cdot \Pr[n_{i_1} = 1 | (n_1 = 1) \cap (n_2 = 1) \cap \cdots \cap (n_{i-2} = 1)] \\
& \cdots\cdots \\
& \cdot \Pr[n_2 = 1 | n_1 = 1] \cdot \Pr[n_1 = 1]
\end{aligned}
$$

Note that $((n_1 = 1) \cap (n_2 = 1) \cap \cdots \cap (n_k = 1)) \equiv (n_k = 1)$:

$$
\begin{aligned}
= \ & \Pr[n_i = 1 | n_{i-1} = 1] \cdot \Pr[n_{i_1} = 1 | n_{i-2} = 1] \cdot \ldots \\
& \cdot \Pr[n_2 = 1 | n_1 = 1] \cdot \Pr[n_1 = 1]
\end{aligned}
$$

Note that the probability $\Pr[n_k = 1 | n_{k-1} = 1] = \frac{10}{11}$ since we are equally likely to get our favorite candy bar on each turn:

$$
= \left(\frac{10}{11}\right)^{(i-1)}
$$

So we have that:

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{i=1}^{\infty} \Pr[X_1 = i] \\
&= \sum_{i=1}^{\infty} \left(\frac{10}{11}\right)^{(i-1)}
\end{aligned}
$$

Since $|\frac{10}{11}| < 1$, we can use the sum of an infinite geometric series:

$$
= 11
$$

7. Let G be a connected graph where all vertices are of even degree. Prove that G has no *cut edges*. A *cut edge* is an edge, that if removed, would increase the number of connected components of the graph.

   Solution:

   Suppose, for the sake of contradiction, that G does have a cut edge $e = \{v_1, v_2\}$. Since G is connected, $G - \{e\}$ has exactly two connected components. Note that each vertex in $G - \{e\}$ has the same degree as in G except v_1 and v_2, whose degrees are each one less than in G. Thus, v_1 is the only vertex of odd degree in its connected component in $G - \{e\}$. Therefore, there are an odd number of odd-degree vertices in that connected component. However, we know that there must be an even number of odd degreed vertices because each connected component of a graph is itself a graph, and hence contains even number of odd degree vertices. This is a contradiction, and we are done.

8. Let $T = (V, E)$ be a tree with $n \geq 2$ vertices. For any two vertices u, v in the tree, let $d(u, v)$ be the length of the path between u and v.

   Prove that for any vertex $u \in V$,

$$\sum_{v \in V} d(u, v) \leq \binom{n}{2}$$

Solution:

We prove this by performing induction on the number of vertices n.

Base Case: When $n = 2$, we can write $V = \{u_0, u_1\}$. Without loss of generality, let $u = u_0$. Then:

$$d(u_0, u_0) + d(u_0, u_1) = 0 + 1 = 1 = \binom{2}{2}$$

Induction Hypothesis: Suppose that for all trees with k vertices (for some $k \geq 2$), for any vertex u, the inequality holds.

Induction Step: Let $T = (V, E)$ be a tree with $k + 1$ vertices. Let x be an arbitrary leaf in T. Consider the tree T' that is constructed by removing ℓ from T. Now let u be an arbitrary vertex of V. We have two cases:

- Suppose $u \neq x$, which means that u is a vertex in T'. By the inductive hypothesis applied to T', we know that:

$$\sum_{v \in (V - \{x\})} d(u, v) \leq \binom{k}{2}$$

  Notice that

$$\sum_{v \in V} d(u, v) = d(u, x) + \sum_{v \in (V - \{x\})} d(u, v) \leq d(u, x) + \binom{k}{2}$$

  Since there are k vertices in T', it must be the case that $d(u, x) \leq k$. Thus

$$\sum_{v \in V} d(u, v) \leq k + \binom{k}{2}$$

  Recall that $\binom{k}{2} = \frac{k(k-1)}{2}$, which means that $k + \binom{k}{2} = \frac{k(2+k-1)}{2} = \frac{k(k+1)}{2} = \binom{k+1}{2}$. Thus

$$\sum_{v \in V} d(u, v) \leq k + \binom{k}{2} = \binom{k+1}{2}$$

- Suppose $u = x$. Since x is a leaf, there is a unique vertex y such that (x, y) is an edge in T. Notice then that

$$\sum_{v \in V} d(x, v) = d(x, x) + \sum_{v \in (V - \{x\})} d(u, v) = 0 + \sum_{v \in (V - \{x\})} (1 + d(y, v)) = k + \sum_{v \in (V - \{x\})} d(y, v)$$

By the inductive hypothesis we know that $\sum_{v\in(V-\{x\})} d(y,v) \leq \binom{k}{2}$, and therefore

$$\sum_{v\in V} d(x,v) \leq k + \binom{k}{2} = \binom{k+1}{2}$$

as before.

9. A CMSC 250 angel tells you in a dream that every connected graph has a connected subgraph that is a tree, which retains all the vertices of the original graph (called a *spanning tree*). The angel also tells you a procedure that allows you to find that exact subgraph given any connected graph, G. The following is a procedure: We will keep adding edges to a subgraph H of G so that at the end H is a spanning tree of G. Initially H has no edges and $V(H) := V(G)$. While H has more than 1 component, find an edge in G that has endpoints in two different components of H and add it to H. Prove the following properties:

   (a) If H has more than 1 component, there is some edge in G whose endpoints lie in different components of H.

   (b) At all times H is an acyclic graph.

   (c) When this procedure terminates, H will be a spanning tree of G.

   Solution:

   (a) Suppose for the sake of contradiction that H has more than one component and each edge in G has endpoints in the same component of H. Then there are no paths in G between vertices in different components of H, contradicting the fact that G is connected. So we must have edges in G with endpoints in different components.

   (b) The proof is by induction on the number of edges n added to H.

   Base Case: When $n = 0$, the subgraph H has no edges in it, and therefore is acyclic.

   Induction Hypothesis: The subgraph H is acyclic after the addition of $n = k$ edges, for some $k \geq 0$.

   Induction Step: Consider the subgraph H after the addition of $k + 1$ edges. Note that by the induction hypothesis, the subgraph was acyclic prior to the addition of the $k + 1^{th}$ edge. Now let us consider the addition of the $k + 1^{th}$ edge $\{u, v\}$. Suppose for contradiction that $\{u, v\}$ causes a cycle to be formed in H. This cycle must include the edge $\{u, v\}$ because there was no cycle before it was added. Then before the addition of $\{u, v\}$, there must have been a path in H between u and v. However, u and v were in different components of H, meaning that there was no such path, a contradiction. Thus the addition of the $k+1^{th}$ edge does not cause a cycle and we have proved the inductive step.

(c) As long as H contains more than one component, we will find an edge in G to add to H by part (a). Thus we will only stop when H becomes a single component, at which point it will be connected and acyclic by part (b), i.e., a tree.

10. A 10 digit number is chosen randomly where each of the digits is with equal probability equal to one of the digits 1 to 9 and where each digit is chosen independently of the other digits. Let N be the number of digits missing from the randomly selected 10 digit number. For example if the number if 1231452832, then we are missing the digits 6, 7, 9 and so $N = 3$. Find $E[N]$.

Solution:

Let N_i be the indicator random variable for the event that the number i is missing from the 10 digit number. Note that $N = \sum_i N_i$.

By Linearity of Expectation, we have that:

$$\begin{aligned}\mathbf{E}[N] &= \sum_i \mathbf{E}[N_i] \\ &= \sum_i \Pr[N_i = 1]\end{aligned}$$

Note that $|N_i = 1| = 8^{10}$, since we are just selecting 10 digits, each of which has 8 options (since it cannot be i). Note that $|\Omega| = 9^{10}$, since we are selecting 10 digits, each of which has 9 options. Since the probability space is uniform, we have that $\Pr[N_i = 1] = \left(\frac{8}{9}\right)^{10}$.

Plugging this in, we have:

$$\begin{aligned}\mathbf{E}[N] &= \sum_i \Pr[N_i = 1] \\ &= \sum_i \left(\frac{8}{9}\right)^{10} \\ &= 9 \cdot \left(\frac{8}{9}\right)^{10}\end{aligned}$$