CS 70 Discrete Mathematics and Probability Theory
Fall 2020 **HW 10**

1 Independent Complements

Let Ω be a sample space, and let $A, B \subseteq \Omega$ be two independent events.

(a) Prove or disprove: $\overline{A}$ and $\overline{B}$ must be independent.

(b) Prove or disprove: A and $\overline{B}$ must be independent.

(c) Prove or disprove: A and $\overline{A}$ must be independent.

(d) Prove or disprove: It is possible that $A = B$.

Solution:

(a) True. $\overline{A}$ and $\overline{B}$ must be independent:

$$\begin{aligned}
\mathbb{P}[\overline{A} \cap \overline{B}] &= \mathbb{P}[\overline{A \cup B}] && \text{(by De Morgan's law)} \\
&= 1 - \mathbb{P}[A \cup B] && \text{(since } \mathbb{P}[\overline{E}] = 1 - \mathbb{P}[E] \text{ for all } E\text{)} \\
&= 1 - (\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]) && \text{(union of overlapping events)} \\
&= 1 - \mathbb{P}[A] - \mathbb{P}[B] + \mathbb{P}[A]\mathbb{P}[B] && \text{(using our assumption that } A \text{ and } B \text{ are independent)} \\
&= (1 - \mathbb{P}[A])(1 - \mathbb{P}[B]) \\
&= \mathbb{P}[\overline{A}]\mathbb{P}[\overline{B}] && \text{(since } \mathbb{P}[\overline{E}] = 1 - \mathbb{P}[E] \text{ for all } E\text{)}
\end{aligned}$$

(b) True. A and $\overline{B}$ must be independent:

$$\begin{aligned}
\mathbb{P}[A \cap \overline{B}] &= \mathbb{P}[A - (A \cap B)] \\
&= \mathbb{P}[A] - \mathbb{P}[A \cap B] \\
&= \mathbb{P}[A] - \mathbb{P}[A]\mathbb{P}[B] \\
&= \mathbb{P}[A](1 - \mathbb{P}[B]) \\
&= \mathbb{P}[A]\mathbb{P}[\overline{B}]
\end{aligned}$$

(c) False in general. If $0 < \mathbb{P}[A] < 1$, then $\mathbb{P}[A \cap \overline{A}] = \mathbb{P}[\varnothing] = 0$ but $\mathbb{P}[A]\mathbb{P}[\overline{A}] > 0$, so $\mathbb{P}[A \cap \overline{A}] \neq \mathbb{P}[A]\mathbb{P}[\overline{A}]$; therefore A and $\overline{A}$ are not independent in this case.

(d) True. To give one example, if $\mathbb{P}[A] = \mathbb{P}[B] = 0$, then $\mathbb{P}[A \cap B] = 0 = 0 \times 0 = \mathbb{P}[A]\mathbb{P}[B]$, so A and B are independent in this case. (Another example: If $A = B$ and $\mathbb{P}[A] = 1$, then A and B are independent.)

2 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

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

3 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

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

.

4 To Be Fair

Suppose you have a biased coin with $\mathbb{P}(\text{heads}) \neq 0.5$. How could you use this coin to simulate a fair coin? (*Hint*: Think about pairs of tosses.)

Solution:

Let's think about the experiment of throwing the biased coin twice as hinted towards in the hint: Its sample space is $\Omega = \{\text{HH}, \text{TT}, \text{HT}, \text{TH}\}$ with corresponding probability function $\mathbb{P}(\text{HH}) = p^2, \mathbb{P}(\text{TT}) = (1-p)^2$ and $\mathbb{P}(\text{HT}) = \mathbb{P}(\text{TH}) = p(1-p)$. Neither of these probabilities is $1/2$ which is what would be required for simulating a fair coin. However, we can generate new probabilities by looking at conditional probabilities! In particular, knowing that $\mathbb{P}(\text{HT}) = \mathbb{P}(\text{TH})$ informs us that $\mathbb{P}(\text{HT} \mid \{\text{HT}, \text{TH}\}) = \dfrac{\mathbb{P}(\text{HT})}{\mathbb{P}(\{\text{HT}, \text{TH}\})} = \dfrac{\mathbb{P}(\text{TH})}{\mathbb{P}(\{\text{HT}, \text{TH}\})} = \mathbb{P}(\text{TH} \mid \{\text{HT}, \text{TH}\})$, and since $\mathbb{P}(\text{HT} \mid \{\text{HT}, \text{TH}\}) + \mathbb{P}(\text{TH} \mid \{\text{HT}, \text{TH}\}) = 1$, it must be true that

$$\mathbb{P}(\text{HT} \mid \{\text{HT}, \text{TH}\}) = \mathbb{P}(\text{TH} \mid \{\text{HT}, \text{TH}\}) = 1/2.$$

That is, to simulate a fair coin we can throw the biased coin twice until we observe either HT or TH, since the probability of either showing up first is exactly $1/2$.

5 Identity Theft

A group of n friends go to the gym together, and while they are playing basketball, they leave their bags against the nearby wall. An evildoer comes, takes the student ID cards from the bags, randomly rearranges them, and places them back in the bags, one ID card per bag.

(a) What is the probability that no one receives his or her own ID card back?

Hint: Use the inclusion-exclusion principle.

(b) What is the limit of this proability as $n \to \infty$?

Hint: $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Solution:

(a) We are looking for the probability of the event that no one receives his or her own ID card back. It is easier to consider the complement of the above event, which is the event that at least one person receives his or her ID card back. Let A_i, $i = 1, \dots, n$, be the event that the ith friend receives his or her own ID card back, so the event we are considering now is $A_1 \cup \cdots \cup A_n$. We will compute this probability using the generalized inclusion-exclusion formula. Recall that for events a set of n events $B_1, B_2, \dots, B_n$ this is

$$\mathbb{P}[\bigcup_{i=1}^{n} B_i] = \sum_{i=1}^{n} \mathbb{P}[B_i] - \sum_{i,j} \mathbb{P}[B_i \cap B_j] + \sum_{i,j,k} \mathbb{P}[B_i \cap B_j \cap B_k] - \ldots \pm \mathbb{P}[\bigcap_{i=1}^{n} B_i] \ .$$

- First, we add $\mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n)$. Here, $\mathbb{P}(A_i)$ is the probability that the ith friend receives his or her own ID card back, which is $1/n$. So, we add $n \cdot (1/n) = 1$.
- Next, we subtract $\sum_{(i,j)} \mathbb{P}(A_i \cap A_j)$, where the sum runs over all $(i, j) \in \{1, \dots, n\}^2$ with $i < j$. Note that $\mathbb{P}(A_i \cap A_j)$ is the probability that both friend i and friend j receive their own ID cards back, which has probability $(n-2)!/n!$. (To see this, observe that once we have decided that friends i and j will receive their own ID cards back, there are $(n-2)!$ ways to permute the ID cards of the $n-2$ other friends, and there are $n!$ total permutations of the n ID cards.) So, we subtract $\sum_{(i,j)} (n-2)!/n!$, but the summation has $\binom{n}{2}$ terms, so we subtract a total of

$$\binom{n}{2} \frac{(n-2)!}{n!} = \frac{n!}{2!(n-2)!} \cdot \frac{(n-2)!}{n!} = \frac{1}{2!}.$$

- At the kth step of the inclusion-exclusion process, we add $(-1)^{k+1} \sum_{(i_1,\dots,i_k)} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k})$, where the k-tuples in the summation range over all $(i_1, \dots, i_k) \in \{1, \dots, n\}^k$ with $i_1 < \cdots < i_k$. To compute $\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k})$, note that we have decided that k friends will receive their own ID cards back, the remaining $n-k$ ID cards can be permuted in $(n-k)!$ ways, and there are $n!$ total permutations, so the probability is $(n-k)!/n!$. The summation has a total of $\binom{n}{k}$ terms, so we add

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

Now, adding up all of these probabilities together, we have

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots + \frac{(-1)^{n+1}}{n!}.$$

Recall that $A_1 \cup \cdots \cup A_n$ is the *complement* of the event we were originally interested in. So,

$$\mathbb{P}(\text{no friends receive their own ID cards back}) = 1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

(b) Recall the power series for e^x:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Therefore, we have the approximation (which gets better as $n \to \infty$):

$$\mathbb{P}(\text{no friends receive their own ID cards back}) \approx e^{-1} = \frac{1}{e} \approx 0.368.$$

6 Balls and Bins, All Day Every Day

Suppose n balls are thrown into n labeled bins one at a time, where n is a positive *even* integer.

(a) What is the probability that exactly k balls land in the first bin, where k is an integer $0 \le k \le n$?

(b) What is the probability p that at least half of the balls land in the first bin? (You may leave your answer as a summation.)

(c) Using the union bound, give a simple upper bound, in terms of p, on the probability that some bin contains at least half of the balls.

(d) What is the probability, in terms of p, that at least half of the balls land in the first bin, or at least half of the balls land in the second bin?

(e) After you throw the balls into the bins, you walk over to the bin which contains the first ball you threw, and you randomly pick a ball from this bin. What is the probability that you pick up the first ball you threw? (Again, leave your answer as a summation.)

Solution:

(a) The probability that a particular ball lands in the first bin is $1/n$. We need exactly k balls to land in the first bin, which occurs with probability $(1/n)^k$, and we need exactly $n-k$ balls to land in a different bin, which occurs with probability $(1-1/n)^{n-k}$, and there are $\binom{n}{k}$ ways to choose which of the k balls land in first bin. Thus, the probability is $\binom{n}{k}(1/n)^k(1-1/n)^{n-k}$.

(b) This is the summation over $k = n/2, \ldots, n$ of the probabilities computed in the first part, i.e., $\sum_{k=n/2}^{n} \binom{n}{k}(1/n)^k(1-1/n)^{n-k}$.

(c) The event that some bin has at least half of the balls is the union of the events A_k, $k = 1, \ldots, n$, where A_k is the event that bin k has at least half of the balls. By the union bound, $\mathbb{P}(\bigcup_{k=1}^{n} A_k) \le \sum_{k=1}^{n} \mathbb{P}(A_k) = np$.

(d) The probability that the first bin has at least half of the balls is p; similarly, the probability that the second bin has at least half of the balls is also p. There is overlap between these two events, however: the first bin has half of the balls and the second bin has the second half of the balls. The probability of this event is $\binom{n}{n/2}n^{-n}$: there are n^n total possible configurations for the n balls to land in the bins, but if we require exactly $n/2$ of the balls to land in the first bin and the remaining balls to land in the second bin, there are $\binom{n}{n/2}$ ways to choose which balls land in the first bin. By the principle of inclusion-exclusion, our desired probability is $p+p-\binom{n}{n/2}n^{-n} = 2p-\binom{n}{n/2}n^{-n}$.

(e) Condition on the number of balls in the bin. First we calculate the probability $\mathbb{P}(A_k)$, where A_k is the event that, in addition to the first ball you threw, an additional $k-1$ of the other $n-1$ balls landed in this bin, which by the reasoning in Part (a) has probability

$$\mathbb{P}(A_k) = \binom{n-1}{k-1}(1/n)^{k-1}(1-1/n)^{n-k} \ .$$

If we let B be the event that we pick up the first ball we threw, then

$$\mathbb{P}(B \mid A_k) = 1/k$$

since we are equally likely to pick any of the k balls in the bin. Thus the overall probability we are looking for is, by an application of the law of total probability,

$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(A_k \cap B) = \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{P}(B \mid A_k) = \sum_{k=1}^{n} \frac{1}{k}\binom{n-1}{k-1}\left(\frac{1}{n}\right)^{k-1}\left(1-\frac{1}{n}\right)^{n-k}.$$

7 Cliques in Random Graphs

In last week's homework you worked on a graph $G=(V,E)$ on n vertices which is generated by the following random process: for each pair of vertices u and v, we flip a fair coin and place an (undirected) edge between u and v if and only if the coin comes up heads. Now consider:

(a) What is the size of the sample space?

(b) A k-clique in graph is a set S of k vertices which are pairwise adjacent (every pair of vertices is connected by an edge). For example a 3-clique is a triangle. Let's call the event that S forms a clique E_S. What is the probability of E_S for a particular set S of k vertices?

(c) For two sets of vertices $V_1 = \{v_1,\dots,v_\ell\}$ and $V_2 = \{w_1,\dots,w_k\}$, are E_{V_1} and E_{V_2} independent?

(d) Prove that $\binom{n}{k} \le n^k$.

(e) Prove that the probability that the graph contains a k-clique, for $k \ge 4\log n + 1$, is at most $1/n$. (The log is taken base 2). *Hint:* Apply the union bound and part (d).

Solution:

(a) Between every pair of verticles, there is either an edge or not. Since there are two choices for each of the $\binom{n}{2}$ pairs of vertices, the size of the sample space is $2^{\binom{n}{2}}$.

(b) For a fixed set of k vertices to be a k-clique, all of the $\binom{k}{2}$ pairs of those vertices have to be connected by an edge. The probability of this event is $1/2^{\binom{k}{2}}$.

(c) E_{V_1} and E_{V_2} are independent if and only if V_1 and V_2 share at most one vertex: If V_1 and V_2 share at most one vertex, then since edges are added independently of each other, we have

$$\mathbb{P}(E_{V_1} \cap E_{V_2}) = \mathbb{P}(\text{all edges in } V_1 \text{ and all edges in } V_2 \text{ are present}) = \left(\frac{1}{2}\right)^{\binom{|V_1|}{2}} \cdot \left(\frac{1}{2}\right)^{\binom{|V_2|}{2}} = \mathbb{P}(E_{V_1}) \cdot \mathbb{P}(E_{V_2}).$$

Conversely, if V_1 and V_2 share at least two vertices, then their intersection $V_3 = V_1 \cap V_2$ has at least 2 elements, and whence

$$\mathbb{P}(E_{V_1} \cap E_{V_2}) = \left(\frac{1}{2}\right)^{\binom{|V_3|}{2}} \cdot \left(\frac{1}{2}\right)^{\binom{|V_1|}{2} - \binom{|V_3|}{2}} \cdot \left(\frac{1}{2}\right)^{\binom{|V_2|}{2} - \binom{|V_3|}{2}} = \left(\frac{1}{2}\right)^{\binom{|V_1|}{2} + \binom{|V_2|}{2} - \binom{|V_3|}{2}} \neq \mathbb{P}(E_{V_1}) \cdot \mathbb{P}(E_{V_2}).$$

(d) The algebraic solution is an application of the definition of $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \tag{1}$$
$$\leq n \cdot (n-1) \cdots (n-k+1) \tag{2}$$
$$\leq n^k \tag{3}$$

(e) Let A_S denote the event that S is a k-clique, where $S \subseteq V$ is of size k. Then, the event that the graph contains a k-clique can be described as the union of A_S's over all $S \subseteq V$ of size k. Using the union bound,

$$\mathbb{P}\left[\bigcup_{S \subseteq V, |S|=k} A_S\right] \leq \sum_{S \subseteq V, |S|=k} \mathbb{P}[A_S] = \sum_{S \subseteq V, |S|=k} \frac{1}{2^{\binom{k}{2}}}.$$

Now, since there are $\binom{n}{k}$ ways of choosing a subset $S \subseteq V$ of size k, the right-hand side of the above equality is

$$\frac{\binom{n}{k}}{2^{\binom{k}{2}}} = \frac{\binom{n}{k}}{2^{k(k-1)/2}} \leq \frac{n^k}{\left(2^{(k-1)/2}\right)^k} \leq \frac{n^k}{\left(2^{(4\log n + 1 - 1)/2}\right)^k} = \frac{n^k}{\left(2^{2\log n}\right)^k} = \frac{n^k}{n^{2k}} = \frac{1}{n^k} \leq \frac{1}{n}.$$