CS 70 Spring 2025 — Discrete Mathematics and Probability Theory — Rao — HW 10

Due: Saturday, 4/5, 4:00 PM
Grace period until Saturday, 4/5, 6:00 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

Zachary Brandt
zbrandt@berkeley.edu

1 Probability Potpourri

Note 13
Note 14

Provide brief justification for each part.

(a) For two events A and B in any probability space, show that $\mathbb{P}[A \setminus B] \geq \mathbb{P}[A] - \mathbb{P}[B]$.

(b) Suppose $\mathbb{P}[D \mid C] = \mathbb{P}[D \mid \overline{C}]$, where $\overline{C}$ is the complement of C. Prove that D is independent of C.

(c) If A and B are disjoint, does that imply they're independent?

Solution:

(a) $\mathbb{P}[A \setminus B]$ is the probability of event A occuring and event B not occuring, i.e., $\mathbb{P}[A \setminus B] = \mathbb{P}[A] - \mathbb{P}[A \cap B]$. This is greater than or equal to $\mathbb{P}[A] - \mathbb{P}[B]$, because $\mathbb{P}[A \cap B]$ can only ever be as great as $\mathbb{P}[B]$ (or $\mathbb{P}[A]$), considering that it is the intersection of events A and B. When A perfectly coinsides with B, then $\mathbb{P}[A \cap B] = \mathbb{P}[B]$, but it is otherwise less than $\mathbb{P}[B]$.

$$\begin{aligned} \mathbb{P}[A \cap B] &\leq \mathbb{P}[B] \\ -\mathbb{P}[B] &\leq -\mathbb{P}[A \cap B] \\ \mathbb{P}[A] - \mathbb{P}[B] &\leq \mathbb{P}[A] - \mathbb{P}[A \cap B] \\ \mathbb{P}[A] - \mathbb{P}[B] &\leq \mathbb{P}[A \setminus B] \end{aligned}$$

(b) For D to be independent of C, it must be the case that $\mathbb{P}[D \mid C] = \mathbb{P}[D]$.

$$
\begin{aligned}
\mathbb{P}[D \mid C] &= \mathbb{P}[D \mid \overline{C}] \\
\mathbb{P}[D \mid C] &= \frac{\mathbb{P}[D \cap \overline{C}]}{\mathbb{P}[\overline{C}]} \\
\mathbb{P}[D \mid C] &= \frac{\mathbb{P}[D \cap \overline{C}]}{1-\mathbb{P}[C]} \\
\mathbb{P}[D \mid C](1-\mathbb{P}[C]) &= \mathbb{P}[D \cap \overline{C}] \\
\mathbb{P}[D \mid C] - \mathbb{P}[D \mid C] \cdot \mathbb{P}[C] &= \mathbb{P}[D \cap \overline{C}] \\
\mathbb{P}[D \mid C] - \mathbb{P}[D \cap C] &= \mathbb{P}[D \cap \overline{C}] \\
\mathbb{P}[D \mid C] &= \mathbb{P}[D \cap \overline{C}] + \mathbb{P}[D \cap C] \\
\mathbb{P}[D \mid C] &= \mathbb{P}[D]
\end{aligned}
$$

Therefore, D is independent of C.

(c) If A and B are disjoint, then knowing that one event happens provides information on the other. $\mathbb{P}[A \cap B] = 0$ if disjoint, and therefore, $\mathbb{P}[A \mid B] = 0 \neq A$. This does not imply they are independent of each other.

2 Independent Complements

Note 14

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: If $\overline{A}$ and $\overline{B}$ are independent, then $\mathbb{P}[\overline{A} \cap \overline{B}] = \mathbb{P}[\overline{A}] \cdot \mathbb{P}[\overline{B}]$. Also, from observation of a Venn diagram, $\overline{A} \cap \overline{B} = 1 - A \cup B$. Then,

$$
\begin{aligned}
\mathbb{P}[\overline{A}] \cdot \mathbb{P}[\overline{B}] &= (1-\mathbb{P}[A]) \cdot (1-\mathbb{P}[B]) \\
&= (1-\mathbb{P}[A]) \cdot (1-\mathbb{P}[B]) \\
&= 1 - \mathbb{P}[B] - \mathbb{P}[A] + \mathbb{P}[A] \cdot \mathbb{P}[B] \\
\text{Since } A \text{ and } B \text{ are independent} \quad &= 1 - \mathbb{P}[B] - \mathbb{P}[A] + \mathbb{P}[A \cap B] \\
&= 1 - (\mathbb{P}[B] + \mathbb{P}[A] - \mathbb{P}[A \cap B]) \\
&= 1 - \mathbb{P}[A \cup B] \\
&= \mathbb{P}[\overline{A} \cap \overline{B}]
\end{aligned}
$$

(b) True: If A and $\overline{B}$ are independent, then $\mathbb{P}[A \cap \overline{B}] = \mathbb{P}[A] \cdot \mathbb{P}[\overline{B}]$. Also, from observation, like of a Venn Diagram, $A - A \cap B = A \cap \overline{B}$. Then,

$$\begin{aligned} \mathbb{P}[A] \cdot \mathbb{P}[\overline{B}] &= \mathbb{P}[A] \cdot (1 - \mathbb{P}[B]) \\ &= \mathbb{P}[A] - \mathbb{P}[A] \cdot \mathbb{P}[B] \\ \text{Since } A \text{ and } B \text{ are independent} \quad &= \mathbb{P}[A] - \mathbb{P}[A \cap B] \\ &= \mathbb{P}[A \cup \overline{B}] \end{aligned}$$

(c) False: For A and $\overline{A}$ to be independent, $\mathbb{P}[A \mid \overline{A}] = \mathbb{P}[A \cap \overline{A}] / \mathbb{P}[\overline{A}] = \mathbb{P}[A]$. But since $\mathbb{P}[A \cap \overline{A}]$ is equal to zero, $\mathbb{P}[A \mid \overline{A}] \neq 0$, and the events are not independent.

(d) False: For A and B to be independent, then $\mathbb{P}[A \mid B] = \mathbb{P}[A]$. But since $\mathbb{P}[A \cap B] = 1$, because $A = B$, then $\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A]}{\mathbb{P}[B]} = 1 \neq \mathbb{P}[A] \quad (\text{when } \mathbb{P}[A] \neq 1)$.

3 Cliques in Random Graphs

Note 13
Note 14

Consider the 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.

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

(b) A k-clique in a 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 E_S be the event that a set S forms a clique. What is the probability of E_S for a particular set S of k vertices?

(c) Suppose that $V_1 = \{v_1, \ldots, v_\ell\}$ and $V_2 = \{w_1, \ldots, w_k\}$ are two arbitrary sets of vertices. What conditions must V_1 and V_2 satisfy in order for E_{V_1} and E_{V_2} to be independent? Prove your answer.

(d) Prove that $\binom{n}{k} \leq n^k$. (You might find this useful in part (e)).

(e) Prove that the probability that the graph contains a k-clique, for $k \geq 4 \log_2 n + 1$, is at most $1/n$. *Hint:* Use the union bound.

Solution:

(a) The size of the sample space is $2^{\binom{n}{2}}$. The number of distinct pairs of vertices there are is $\binom{n}{2}$, and the two can either have an edge connecting the two or not.

(b) For a set S with k vertices, all vertices must be connected to each other for there to be a k-clique. This can only happen one way, and therefore, $\mathbb{P}[E_S] = 1/2^{\binom{k}{2}}$.

(c) If there are no vertices of G that are both in each of the subsets V_1 and V_2, then $V_1 \cap V_2 = \emptyset$, and $\mathbb{P}[E_{V_1} \cap E_{V_2}] = \mathbb{P}[E_{V_1}] \cdot \mathbb{P}[E_{V_2}]$, since the edges are simply added with one half probability without any overlap. The probabilities are

$$\mathbb{P}[E_{V_1}] = \left(\frac{1}{2}\right)^{\binom{|V_1|}{2}}$$
$$\mathbb{P}[E_{V_2}] = \left(\frac{1}{2}\right)^{\binom{|V_2|}{2}}$$

If there are however two vertices that are both in V_1 and V_2, then there will be one less coin toss for both E_{V_1} and E_{V_2}, plus the only one coin toss for the edge that connects both of these vertices. Therefore, the events are also independent if there is only one shared vertex, because it takes two to form an edge to reduce the total number of coin tosses.

(d)

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{(n-k)! \cdot k!} \\ &= \frac{(n) \times (n-1) \times \cdots \times (n-k+1)}{k!} \\ &\leq n^k\end{aligned}$$

The numerator of the fraction multiples n by itself k times but is subtracted from by many other terms and factored by $k!$ and is therefore at least less than or equal to n^k.

(e) For a graph to contain a k-clique, there must be a subset of vertices of size $|S| = k$ where all vertices are connected. We know from before that the probability of E_S is $1/2^{\binom{k}{2}}$. There are $\binom{n}{k}$ ways to pick the vertices to form $|S|$. Therefore, the probability that the graph contains a k-clique is

$$\binom{n}{k} \times \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

Since each S has that probability and there are $\binom{n}{k}$ of them. We know from the previous problem that $\binom{n}{k} \leq n^k$, so

$$\begin{aligned}
\binom{n}{k} \times \left(\frac{1}{2}\right)^{\binom{k}{2}} &\leq n^k \times \left(\frac{1}{2}\right)^{\binom{k}{2}} \\
&\leq \frac{n^k}{2^{\binom{k}{2}}} \\
&= \frac{n^k}{2^{\frac{k!}{(k-2)!\cdot 2}}} \\
&= \frac{n^k}{2^{\frac{k\cdot(k-1)}{2}}} \\
\text{still less than} \quad &\leq n^{(4\log_2 n+1)} \div \left(2^{\frac{(4\log_2 n+1)\cdot((4\log_2 n+1)-1)}{2}}\right) \\
&= n^{(4\log_2 n+1)} \div \left(2^{(4\log_2 n+1)\cdot(2\log_2 n)}\right) \\
&= n^{(4\log_2 n+1)} \div \left(n^{(4\log_2 n+1)\cdot(2)}\right) \\
&= 1 \div \left(n^{(4\log_2 n+1)}\right) \\
\text{again less than} \quad &\leq \frac{1}{n}
\end{aligned}$$

4 Poisoned Smarties

Note 14

Supposed there are 3 people who are all owners of their own Smarties factories. Burr Kelly, being the brightest and most innovative of the owners, produces considerably more Smarties than her competitors and has a commanding 50% of the market share. Yousef See, who inherited her riches, lags behind Burr and produces 40% of the world's Smarties. Finally Stan Furd, brings up the rear with a measly 10%. However, a recent string of Smarties related food poisoning has forced the FDA investigate these factories to find the root of the problem. Through her investigations, the inspector found that 2 Smarties out of every 100 at Kelly's factory was poisonous. At See's factory, 5% of Smarties produced were poisonous. And at Furd's factory, the probability a Smarty was poisonous was 0.1.

(a) What is the probability that a randomly selected Smarty will be safe to eat?

(b) If we know that a certain Smarty didn't come from Burr Kelly's factory, what is the probability that this Smarty is poisonous?

(c) If a randomly selected Smarty is poisonous, what is the probability it came from Stan Furd's Smarties Factory?

Solution:

(a) The probability that a randomly selected Smarty is safe to eat will be $0.5 \times 0.98 + 0.4 \times 0.95 + 0.1 \times 0.90 = 0.96$.

(b) The probability that a certain Smarty is poisonous, given that it is not from Burr Kelly's factory, is

$$\begin{aligned}\mathbb{P}[P \mid \overline{B}] &= \frac{\mathbb{P}[P \cap \overline{B}]}{\mathbb{P}[\overline{B}]} \\ &= \frac{\mathbb{P}[P \cap \overline{B}]}{1 - \mathbb{P}[Y] - \mathbb{P}[S]} \\ &= \frac{0.4 \times 0.05 + 0.1 \times 0.1}{0.5} \\ &= 0.06\end{aligned}$$

(c) The probability that a Smary came from Stan Furd, given that it is poisonous is,

$$\begin{aligned}\mathbb{P}[S \mid P] &= \frac{\mathbb{P}[S \cap P]}{\mathbb{P}[P]} \\ &= \frac{\mathbb{P}[P \mid S] \times \mathbb{P}[S]}{\mathbb{P}[P]} \\ &= \frac{0.1 \times 0.1}{1 - 0.96} \\ &= 0.25\end{aligned}$$

5 Symmetric Marbles

Note 14

A bag contains 4 red marbles and 4 blue marbles. Rachel and Brooke play a game where they draw four marbles in total, one by one, uniformly at random, without replacement. Rachel wins if there are more red than blue marbles, and Brooke wins if there are more blue than red marbles. If there are an equal number of marbles, the game is tied.

(a) Let A_1 be the event that the first marble is red and let A_2 be the event that the second marble is red. Are A_1 and A_2 independent?

(b) What is the probability that Rachel wins the game?

(c) Given that Rachel wins the game, what is the probability that all of the marbles were red?

Now, suppose the bag contains 8 red marbles and 4 blue marbles and we add a tiebreaker to the game: if there are an equal number of red and blue marbles among the four drawn, Rachel wins if the third marble is red, and Brooke wins if the third marble is blue.

(d) What is the probability that the third marble is red?

(e) Given that there are k red marbles among the four drawn, where $0 \leq k \leq 4$, what is the probability that the third marble is red? Answer in terms of k.

(f) Given that the third marble is red, what is the probability that Rachel wins the game?

Solution:

(a) Since we are drawing without replacement, the outcome of the first draw should influence the outcome of the second. If the first draw is a red marble, the probability of A_2 should be lower compared to if we first drew a blue marble.

(b) Rachel wins the game when there are more red marbles than blue marbles, i.e., if all the marbles are red, or if all the marbles except one are red, otherwise the game ends in a tie or in a loss for Rachel. There is only one way for all marbles to be red, which happens with probability $1 \div \binom{8}{4}$. There are four ways for three of the marbles to be red and four different blue marbles to choose from, therefore, the total probability of Rachel winning is

$$\left(\binom{4}{3} \times \binom{4}{1} + 1\right) \div \binom{8}{4} = \frac{4 \times 4 + 1}{70} = \frac{17}{70}$$

.

(c) The probability that all marbles were red, given that Rachel won is

$$\mathbb{P}[\text{all red} \mid \text{Rachel won}] = \frac{\mathbb{P}[\text{all red} \cap \text{Rachel won}]}{\mathbb{P}[\text{Rachel won}]} = \frac{1 \div \binom{8}{4}}{17 \div \binom{8}{4}} = \frac{1}{17}$$

(d) The probability that third marble is red can be split into three cases: the first, where the first two marbles are red as well, the second where one of the first two is red, and the other is blue, and the third, where both of the first two marbles are blue.

$$\begin{aligned} \text{RR}: &\quad \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} = \frac{336}{1320} \\ \text{RB}: &\quad 2 \times \frac{8}{12} \times \frac{4}{11} \times \frac{7}{10} = \frac{448}{1320} \\ \text{BB}: &\quad \frac{4}{12} \times \frac{3}{11} \times \frac{8}{10} = \frac{96}{1320} \end{aligned}$$

The total probability is then $\frac{336+448+96}{1320} = \frac{880}{1320} = \frac{2}{3}$.

(e) From the previous problem's answer we can see that there is some symmetry involved, in that, the probability that the third marble is red is the same as the probability that the first marble is red ($8 \div 12$). Just consider the first marble drawn as the third position. Since there are k red marbles drawn, with each one having the same probability of being the third as the being the first the probability is then $\frac{k}{4}$.

(f) For Rachel to lose the game if the third marble is red is for all other marbles to be blue. Then the probability of her winning is one minus the probability of her losing (considering the tie breaker rule). The probability of this is $\frac{4}{12} \times \frac{3}{11} \times \frac{8}{10} \times \frac{2}{9} = \frac{192}{11880}$. The probability that the third marble is red is $\frac{2}{3}$ from before. Therefore, the conditional probability is then

$$\mathbb{P}[\text{Rachel loses} \mid \text{Third marble is red}] = \frac{\frac{192}{11880}}{\frac{2}{3}} = \frac{4}{165}$$

Then, $\mathbb{P}[\text{Rachel wins} \mid \text{Third marble is red}] = 1 - \frac{4}{165} = \frac{161}{165}$.

6 Socks

Note 13
Note 14

Suppose you have n different pairs of socks (n left socks and n right socks, for $2n$ individual socks total) in your dresser. You take the socks out of the dresser one by one without looking and lay them out in a row on the floor. In this question, we'll go through the computation of the probability that no two matching socks are next to each other.

(a) We can consider the sample space as the set of length $2n$ permutations. What is the size of the sample space Ω, and what is the probability of a particular permutation $\omega \in \Omega$?

(b) Let A_i be the event that the ith pair of matching socks are next to each other. Calculate $\mathbb{P}[A_i]$.

(c) Calculate $\mathbb{P}[A_1 \cap \ldots \cap A_k]$ for an arbitrary $k \geq 2$. (Hint: try using a counting based approach.)

(d) Putting these all together, calculate the probability that there is at least one pair of matching socks next to each other. Your answer can (and should) be expressed as a summation. (Hint: use Inclusion/Exclusion.)

(e) Using your answer from the previous part, what is the probability that no two matching socks are next to each other? (This should follow directly from your answer to the previous part, and also can be left as a summation.)

Solution:

(a) The size of the sample space is $|\Omega| = (2n)!$, and any particular permutation has a probability of $\mathbb{P}[\omega] = \frac{1}{(2n)!}$.

(b) If two of the $2n$ socks are next to each other as a matching pair, that means there are only $(2n-1)!$ ways to rearrange the socks (because the pair has to stick together). Also, the problem doesn't specify that the left one has to come first necessarily, so there are 2 different ways to have them stay together, i.e., there are $2 \times (2n-1)!$ ways to order the socks with a pair staying together. Therefore, the $\mathbb{P}[A_i] = \frac{2 \times (2n-1)!}{(2n)!}$.

(c) For an arbitrary k, the probability $\mathbb{P}[A_i \cap \cdots \cap A_k]$ is

$$\frac{2^k \times (2n-k)!}{(2n)!}$$

since the pairs of socks have to stick together, consolidating them as one item each to be ordered among the remaining socks. Additionally, the pairs can have the left and right socks be in any order presumably, and so there are 2^k ways to order again.

(d) The probability that there is at least one pair of matching socks next to each other in the permuation is the probability that any of $A_1, A_2, \ldots A_n$ happen. As such, all the permutations in each of those sets is valid, so the probability is then

$$\begin{aligned}
\mathbb{P}[A_1 \cup A_2 \cup \cdots \cup A_n] &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq \{1,\ldots,n\} : |S|=k} \mathbb{P}[\cup_{i \in S} A_i] \\
&= \sum_{k=1}^{n} \mathbb{P}[A_k] - \sum_{i,j \in \{1,\ldots,n\} : i<j} \mathbb{P}[A_i \cap A_j] + \sum_{i,j,k \in \{1,\ldots,n\} : i<j<k} \mathbb{P}[A_i \cap A_j \cap A_k] - \ldots \\
&= \sum_{k=1}^{n} \frac{2 \times (2n-1)!}{(2n)!} - \sum_{i,j \in \{1,\ldots,n\} : i<j} \frac{2^2 \times (2n-2)!}{(2n)!} + \ldots \\
\text{(highlight)} \quad &= \binom{n}{1} \frac{2 \times (2n-1)!}{(2n)!} - \binom{n}{2} \frac{2^2 \times (2n-2)!}{(2n)!} + \ldots \\
&= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \frac{2^k \times (2n-k)!}{(2n)!}
\end{aligned}$$

Where on the highlighted line I make the observation that the sums are summing equal probabilities, each some binomial amount. The first summation, sums n or $\binom{n}{1}$ of the same probaility, while the second sum sums $\binom{n}{2}$, and so on.

(e) The last part answered the question as to what the probability that there is at least one matching pair. This question asks for the complement, that there are no matching pairs. Therefore, the probaility is

$$1 - \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \frac{2^k \times (2n-k)!}{(2n)!}$$