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

# 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  $\mathbb{P}[A \setminus B] \ge \mathbb{P}[A] \mathbb{P}[B]$ .
- (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) It can be helpful to first draw out a Venn diagram:



We can see here that  $\mathbb{P}[A] = \mathbb{P}[A \cap B] + \mathbb{P}[A \setminus B]$ , and that  $\mathbb{P}[B] = \mathbb{P}[A \cap B] + \mathbb{P}[B \setminus A]$ .

Looking at the RHS, we have

$$\begin{split} \mathbb{P}[A] - \mathbb{P}[B] &= (\mathbb{P}[A \cap B] + \mathbb{P}[A \setminus B]) - (\mathbb{P}[A \cap B] + \mathbb{P}[B \setminus A]) \\ &= \mathbb{P}[A \setminus B] - \mathbb{P}[B \setminus A] \\ &\leq \mathbb{P}[A \setminus B] \end{split}$$

(b) Using the total probability rule, we have

$$\mathbb{P}[D] = \mathbb{P}[D \cap C] + \mathbb{P}[D \cap \overline{C}] = \mathbb{P}[D \mid C] \cdot \mathbb{P}[C] + \mathbb{P}[D \mid \overline{C}] \cdot \mathbb{P}[\overline{C}].$$

But we know that  $\mathbb{P}[D \mid C] = \mathbb{P}[D \mid \overline{C}]$ , so this simplifies to

$$\mathbb{P}[D] = \mathbb{P}[D \mid C] \cdot (\mathbb{P}[C] + \mathbb{P}[\overline{C}]) = \mathbb{P}[D \mid C] \cdot 1 = \mathbb{P}[D \mid C],$$

which defines independence.

(c) No; if two events are disjoint, we cannot conclude they are independent. Consider a roll of a fair six-sided die. Let A be the event that we roll a 1, and let B be the event that we roll a 2. Certainly A and B are disjoint, as P[A ∩ B] = 0. But these events are not independent: P[B | A] = 0, but P[B] = 1/6.

Since disjoint events have  $\mathbb{P}[A \cap B] = 0$ , we can see that the only time when disjoint *A* and *B* are independent is when either  $\mathbb{P}[A] = 0$  or  $\mathbb{P}[B] = 0$ .

2 Independent Complements

Note 14 Let  $\Omega$  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:

$$\mathbb{P}[\overline{A} \cap \overline{B}] = \mathbb{P}[\overline{A \cup B}]$$
 (by De Morgan's law)  

$$= 1 - \mathbb{P}[A \cup B]$$
 (since  $\mathbb{P}[\overline{E}] = 1 - \mathbb{P}[E]$  for all  $E$ )  

$$= 1 - (\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B])$$
 (union of overlapping events)  

$$= 1 - \mathbb{P}[A] - \mathbb{P}[B] + \mathbb{P}[A] \mathbb{P}[B]$$
 (since  $A$  and  $B$  are independent)  

$$= (1 - \mathbb{P}[A])(1 - \mathbb{P}[B])$$
 (since  $\mathbb{P}[\overline{E}] = 1 - \mathbb{P}[E]$  for all  $E$ )

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

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

- (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 P[A] = P[B] = 0, then P[A ∩ B] = 0 = 0 × 0 = P[A] P[B], so A and B are independent in this case. (Another example: If A = B and P[A] = 1, then A and B are independent.)

# 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  $V_1 = \{v_1, \dots, v_\ell\}$  and  $V_2 = \{w_1, \dots, w_k\}$  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  $k \ge 4\log_2 n + 1$ , is at most 1/n. *Hint:* Use the union bound.

#### **Solution:**

- (a) Between every pair of vertices, there is either an edge or there isn't. 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, so we have

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

CS 70, Spring 2025, HW 10

3

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

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

(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

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

## 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) Let S be the event that a smarty is safe to eat. Let BK be the event that a smarty is from Burr Kelly's factory. Let YS be the event that a smarty is from Yousef See's factory. Finally, let SF be the event that a smarty is from Stan Furd's factory. By total probability, we have

$$\mathbb{P}[S] = \mathbb{P}[BK] \mathbb{P}[S \mid BK] + \mathbb{P}[YS] \mathbb{P}[S \mid YS] + \mathbb{P}[SF] \mathbb{P}[S \mid SF]$$
  
$$= \frac{1}{2} \cdot \frac{49}{50} + \frac{2}{5} \cdot \frac{19}{20} + \frac{1}{10} \cdot \frac{9}{10}$$
  
$$= \frac{49}{100} + \frac{38}{100} + \frac{9}{100}$$
  
$$= \frac{96}{100} = \frac{24}{25} = 0.96$$

Therefore the probability that a Smarty is safe to eat is 0.96.

(b) Let *P* be the event that a smarty is poisonous.

$$\mathbb{P}[P \mid \overline{BK}] = \frac{\mathbb{P}[\overline{BK} \cap P]}{\mathbb{P}[\overline{BK}]}$$

Since BK, YS, SF are a partition of the entire sample space, we know that if BK did not occur, then either YS occurred, or SF occurred:

$$= \frac{\mathbb{P}[YS \cap P]}{\mathbb{P}[\overline{BK}]} + \frac{\mathbb{P}[SF \cap P]}{\mathbb{P}[\overline{BK}]}$$
$$= \frac{\mathbb{P}[P \mid YS] \mathbb{P}[YS]}{1 - \mathbb{P}[BK]} + \frac{\mathbb{P}[P \mid SF] \mathbb{P}[SF]}{1 - \mathbb{P}[BK]}$$
$$= \frac{\frac{1}{20} \cdot \frac{2}{5}}{\frac{1}{2}} + \frac{\frac{1}{10} \cdot \frac{1}{10}}{\frac{1}{2}} = 2 \cdot \frac{2}{100} + 2 \cdot \frac{1}{100}$$
$$= \frac{6}{100} = \frac{3}{50} = 0.06$$

(c) From Bayes' Rule, we know that:

$$\mathbb{P}[SF \mid P] = \frac{\mathbb{P}[P \mid SF] \mathbb{P}[SF]}{\mathbb{P}[P]}.$$

In part (a), we calculated the probability that any random Smarty was safe to eat; here, notice that  $\mathbb{P}[P] = 1 - \mathbb{P}[S]$ . This means we have

$$\mathbb{P}[SF \mid P] = \frac{\mathbb{P}[P \mid SF] \mathbb{P}[SF]}{1 - \mathbb{P}[S]}$$
$$= \frac{\frac{1}{10} \cdot \frac{1}{10}}{1 - \frac{24}{25}} = \frac{\frac{1}{100}}{\frac{1}{25}}$$
$$= \frac{25}{100} = \frac{1}{4} = 0.25$$

### 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  $0 \le k \le 4$ , 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) They are not independent; removing one red marble lowers the probability of the next marble being red.
- (b) Let *p* be the probability that Rachel wins. Since there are an equal number of red and blue marbles, by symmetry, the probability that Rachel wins and the probability that Brooke wins is the same. Thus, the probability that there is a tie is 1 p p = 1 2p.

We now compute the probability that there is a tie. For there to be a tie, two of the four marbles need to be red. There are  $\binom{8}{4}$  ways to pick 4 marbles, and  $\binom{4}{2}\binom{4}{2}$  to pick 2 red and blue marbles, respectively, giving a probability of

$$\frac{\binom{4}{2}\binom{4}{2}}{\binom{8}{4}} = \frac{36}{70} = \boxed{\frac{18}{35}}$$

We conclude that  $1 - 2p = \frac{18}{35}$ . Solving for *p* gives  $p = \boxed{\frac{17}{70}}$ .

(c) Let *A* be the event that there are 3 red marbles drawn, and let *B* be the event that there are 4 red marbles drawn. We wish to compute

$$\mathbb{P}[B \mid (A \cup B)] = rac{\mathbb{P}[B \cap (A \cup B)]}{\mathbb{P}[A \cup B]} = rac{\mathbb{P}[B]}{\mathbb{P}[A] + \mathbb{P}[B]}.$$

Similar to the calculation in part (b), the probability that there are 3 red marbles drawn is  $\frac{\binom{4}{3}\binom{4}{1}}{\binom{8}{4}} = \frac{16}{70}$ , and the probability that there are 4 red marbles drawn is  $\frac{\binom{4}{4}\binom{4}{0}}{\binom{8}{4}} = \frac{1}{70}$ , giving a final answer of  $\frac{\frac{1}{70}}{\frac{16}{70}+\frac{1}{70}} = \boxed{\frac{1}{17}}$ .

(d) By symmetry, the probability that the third marble is red is the same as the probability that the first marble is red, or the same as any marble being red. One way to see this is to imagine drawing the four marbles in order, then moving the first marble drawn to the third position. This is another way to draw four marbles that yields the same distribution.

There are 8 red marbles, and 12 marbles in total. Thus, the probability that the third marble is red is  $\frac{8}{12} = \left[\frac{2}{3}\right]$ .

- (e) We are given that there are k red marbles among the 4 drawn. By symmetry, each marble has the same probability of being red, so the probability that the third marble is red is  $\frac{k}{4}$ .
- (f) The only way for Rachel to lose the game given that the third marble is red is if all the other marbles are blue. The probability that the third marble is red and all the other marbles are blue is  $\frac{4}{12} \cdot \frac{3}{11} \cdot \frac{8}{10} \cdot \frac{2}{9} = \frac{8}{495}$ , and the probability that the third marble is red is  $\frac{8}{12} = \frac{2}{3}$ , so the probability that Rachel loses given that the third marble is red is  $\frac{\frac{8}{495}}{\frac{2}{3}} = \frac{4}{165}$ , and the probability that Rachel wins given that the third marble is red is  $\frac{161}{165}$ .
- 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  $\Omega$ , and what is the probability of a particular permutation  $\omega \in \Omega$ ?
- (b) Let  $A_i$  be the event that the *i*th pair of matching socks are next to each other. Calculate  $\mathbb{P}[A_i]$ .
- (c) Calculate  $\mathbb{P}[A_1 \cap ... \cap A_k]$  for an arbitrary  $k \ge 2$ . (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) We have a uniform sample space of size (2n)!.
- (b) Consider the *i*th matching pair as a single, condensed unit. As an example, in for n = 3, an original permutation could look like 132213. Let us condense both the 2's together, and label it as *B*. Then, a resulting string would look like 13*B*13. Then, there are 2n 1 'units' left that we can order, and thus (2n 1)! ways to order them. Also, when we condensed them, either the left sock or the right sock could've came first, so there are 2 ways to condense this pair. Thus, the probability is  $\frac{2(2n-1)!}{(2n)!}$ .
- (c) We will employ an analogous strategy to the previous part. We will consider all k of these matching socks. There are  $2^k$  ways to condense them. Once condensed, there are (2n-k)! ways to order the remaining units. Thus, the probability is  $2^k \frac{(2n-k)!}{(2n)!}$ .
- (d) We look for:

$$\mathbb{P}[A] = \mathbb{P}[A_1 \cup A_2 \cup \dots \cup A_n]$$
  
=  $\sum_{i=1}^n \mathbb{P}[A_i] - \sum_{1 \le i < j \le n} \mathbb{P}[A_i \cap A_j] + \cdots$   
=  $\sum_{i=1}^n 2 \cdot \frac{(2n-1)!}{(2n)!} - \sum_{1 \le i < j \le n} 2^2 \cdot \frac{(2n-2)!}{(2n)!} + \cdots$   
=  $\binom{n}{1} 2^1 \cdot \frac{(2n-1)!}{(2n)!} - \binom{n}{2} 2^2 \cdot \frac{(2n-2)!}{(2n)!} + \cdots$   
=  $\sum_{k=1}^n (-1)^{k+1} \binom{n}{k} 2^k \cdot \frac{(2n-k)!}{(2n)!}$ 

(e) This is just the complement of the previous part, which becomes

$$1 - \mathbb{P}[A] = \sum_{k=0}^{n} (-1)^{k} \binom{n}{k} \frac{2^{k}(2n-k)!}{(2n)!}.$$