CS 70 Fall 2019 — Discrete Mathematics and Probability Theory — Alistair Sinclair and Yun S. Song — HW 9

__Note:__ This homework consists of two parts. The first part (questions 1-6) will be graded and will determine your score for this homework. The second part (questions 7-8) will be graded if you submit them, but will not affect your homework score in any way. You should attempt the problems in the first part. You are strongly advised to attempt all the questions in the first part. You should attempt the problems in the second part only if you are interested and have time to spare.

For each problem, justify all your answers unless otherwise specified.

Part 1: Required Problems

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 innocent?

Solution:

Let A denote the event that the test indicates that the person is guilty, and B the event that the person is innocent. Note that

$$\mathbb{P}[B] = \frac{99}{100}, \quad \mathbb{P}[\overline{B}] = \frac{1}{100}, \quad \mathbb{P}[A \mid B] = \frac{1}{10}, \quad \mathbb{P}[A \mid \overline{B}] = \frac{4}{5}.$$

Using the Bayesian Inference Rule, we can compute the desired probability as follows:

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

3 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 $0 \leq k \leq n$?

(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, $\mathbb{P}(\bigcup_{k=1}^{n} A_k) \leq \sum_{k=1}^{n} \mathbb{P}(A_k) = np$.

(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

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

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,

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

4 Mario's Coins

Mario owns three identical-looking coins. One coin shows heads with probability $1/4$, another shows heads with probability $1/2$, and the last shows heads with probability $3/4$.

(a) Mario randomly picks a coin and flips it. He then picks one of the other two coins and flips it. Let X_1 and X_2 be the events of the 1st and 2nd flips showing heads, respectively. Are X_1 and X_2 independent? Please prove your answer.

(b) Mario randomly picks a single coin and flips it twice. Let Y_1 and Y_2 be the events of the 1st and 2nd flips showing heads, respectively. Are Y_1 and Y_2 independent? Please prove your answer.

(c) Mario arranges his three coins in a row. He flips the coin on the left, which shows heads. He then flips the coin in the middle, which shows heads. Finally, he flips the coin on the right. What is the probability that it also shows heads?

Solution:

(a) X_1 and X_2 are not independent. Intuitively, the fact that X_1 happened gives some information about the first coin that was chosen; this provides some information about the second coin that was chosen (since the first and second coins can't be the same coin), which directly affects whether X_2 happens or not.

To make this formal, we compute

$$\mathbb{P}(X_1) = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{3}{4}\right) = \frac{1}{2}$$

By symmetry, $\mathbb{P}(X_2) = \mathbb{P}(X_1)$, so

$$\mathbb{P}(X_1)\mathbb{P}(X_2) = \frac{1}{4}.$$

But if we consider the probability that both X_1 and X_2 happen, we have

$$\begin{aligned}\mathbb{P}(X_1 \cap X_2) &= \frac{1}{6}\Big[\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) + \\ &\qquad \left(\frac{1}{2}\right)\left(\frac{3}{4}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{2}\right)\Big] \\ &= \frac{22}{96} = \frac{11}{48}\end{aligned}$$

which is not equal to $1/4$, violating the definition of independence.

(b) Y_1 and Y_2 are not independent. Intuitively, the fact that Y_1 happens gives some information about the coin that was picked, which directly influences whether Y_2 happens or not.

To make this formal, we compute

$$\mathbb{P}(Y_1) = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{3}{4}\right) = \frac{1}{2}$$

By symmetry, $\mathbb{P}(Y_2) = \mathbb{P}(Y_1)$, so

$$\mathbb{P}(Y_1)\mathbb{P}(Y_2) = \frac{1}{4}$$

But if we consider the probability that both Y_1 and $Y_2 1$ happen, we have

$$\mathbb{P}(Y_1 \cap Y_2) = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right)^2 + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right)^2 + \left(\frac{1}{3}\right)\left(\frac{3}{4}\right)^2 = \frac{14}{48} = \frac{7}{24}$$

which is not equal to $1/4$, violating the definition of independence.

(c) Let A be the coin with bias $1/4$, B be the fair coin, and C be the coin with bias $3/4$. There are six orderings, each with probability $1/6$: ABC, ACB, BAC, BCA, CAB, and CBA. Thus

$$\begin{aligned}
&\mathbb{P}(\text{Third coin shows heads} \mid \text{First two coins show heads}) \\
&= \frac{\mathbb{P}(\text{All three coins show heads})}{\mathbb{P}(\text{First two coins show heads})} \\
&= \frac{(1/4)(1/2)(3/4)}{\sum_{\text{Orderings}} \mathbb{P}(\text{First two coins show heads} \mid \text{Ordering})\mathbb{P}(\text{Ordering})} \\
&= \frac{(1/4)(1/2)(3/4)}{(1/6)\sum_{\text{Orderings}} \mathbb{P}(\text{First two coins show heads} \mid \text{Ordering})} \\
&= \frac{(1/4)(1/2)(3/4)}{(1/6)((1/4)(1/2)+(1/4)(3/4)+(1/2)(1/4)+(1/2)(3/4)+(3/4)(1/4)+(3/4)(1/2))} \\
&= \frac{3/32}{11/48} = \frac{9}{22}.
\end{aligned}$$

5 (Un)conditional (In)equalities

Let us consider a sample space $\Omega = \{\omega_1, \ldots, \omega_N\}$ of size $N > 2$, and two probability functions $\mathbb{P}_1$ and $\mathbb{P}_2$ on it. That is, we have two probability spaces: $(\Omega, \mathbb{P}_1)$ and $(\Omega, \mathbb{P}_2)$.

(a) If for every subset $A \subset \Omega$ of size $|A| = 2$ and every outcome $\omega \in \Omega$ it is true that $\mathbb{P}_1(\omega \mid A) = \mathbb{P}_2(\omega \mid A)$, then is it necessarily true that $\mathbb{P}_1(\omega) = \mathbb{P}_2(\omega)$ for all $\omega \in \Omega$? That is, if $\mathbb{P}_1$ and $\mathbb{P}_2$ are equal conditional on events of size 2, are they equal unconditionally? (*Hint*: Remember that probabilities must add up to 1.)

(b) If for every subset $A \subset \Omega$ of size $|A| = k$, where k is some fixed element in $\{2, \ldots, N\}$, and every outcome $\omega \in \Omega$ it is true that $\mathbb{P}_1(\omega \mid A) = \mathbb{P}_2(\omega \mid A)$, then is it necessarily true that $\mathbb{P}_1(\omega) = \mathbb{P}_2(\omega)$ for all $\omega \in \Omega$?

For the following two parts, assume that $\Omega = \left\{(a_1, \ldots, a_k) \mid \sum_{j=1}^{k} a_j = n\right\}$ is the set of configurations of n balls into k labeled bins, and let $\mathbb{P}_1$ be the probabilities assigned to these configurations by throwing the balls independently one after another into the bins, and let $\mathbb{P}_2$ be the probabilities assigned to these configurations by uniformly sampling one of these configurations.

(c) Let A be the event that all n balls land in exactly one bin. What are $\mathbb{P}_1(\omega \mid A)$ and $\mathbb{P}_2(\omega \mid A)$ for any $\omega \in A$? How about $\omega \in \Omega \setminus A$? Is it true that $\mathbb{P}_1(\omega) = \mathbb{P}_2(\omega)$ for all $\omega \in \Omega$?

(d) For the special case of $n = 9$ and $k = 3$, please give two outcomes B and C, so that $\mathbb{P}_1(B) < \mathbb{P}_2(B)$ and $\mathbb{P}_1(C) > \mathbb{P}_2(C)$.

Solution:

(a) Yes, this is indeed true. To see why, let's take the subset $A = \{\omega_i, \omega_j\}$ for some $i, j \in \{1, \ldots, N\}$ and compute: For any $k \in \{1,2\}$, we have $\mathbb{P}_k(\omega_i \mid A) = \frac{\mathbb{P}_k(\omega_i)}{\mathbb{P}_k(A)}$. Since this expression (by assumption) is the same for $k = 1$ and $k = 2$, we conclude that $\frac{\mathbb{P}_1(\omega_i)}{\mathbb{P}_2(\omega_i)} = \frac{\mathbb{P}_1(A)}{\mathbb{P}_2(A)}$. Repeating the reasoning for ω_j, we similarly find that $\frac{\mathbb{P}_1(\omega_j)}{\mathbb{P}_2(\omega_j)} = \frac{\mathbb{P}_1(A)}{\mathbb{P}_2(A)}$, and whence $\frac{\mathbb{P}_1(\omega_i)}{\mathbb{P}_1(\omega_j)} = \frac{\mathbb{P}_2(\omega_i)}{\mathbb{P}_2(\omega_j)}$. Since this is true for any $i, j \in \{1, \ldots, N\}$, we can sum over i to get

$$\frac{1}{\mathbb{P}_1(\omega_j)} = \sum_{i=1}^{N} \frac{\mathbb{P}_1(\omega_i)}{\mathbb{P}_1(\omega_j)} = \sum_{i=1}^{N} \frac{\mathbb{P}_2(\omega_i)}{\mathbb{P}_2(\omega_j)} = \frac{1}{\mathbb{P}_2(\omega_j)},$$

which shows that $\mathbb{P}_1\left(\omega_j\right) = \mathbb{P}_2\left(\omega_j\right)$ for all $j \in \{1, \ldots, N\}$.

(b) Yes, it indeed would. There are two ways of verifying this. The first one is to observe that if $A' \subset A$ and $\omega \in A'$, then $\mathbb{P}_1(\omega \mid A') = \mathbb{P}_1(\omega \mid A' \cap A) = \frac{\mathbb{P}_1(\omega \mid A)}{\mathbb{P}_1(A' \mid A)} = \frac{\mathbb{P}_2(\omega \mid A)}{\mathbb{P}_2(A' \mid A)} = \mathbb{P}_2(\omega \mid A')$, where the second equality follows from the product rule (Theorem 13.1): $\mathbb{P}_1(A) \cdot \mathbb{P}_1(A' \mid A) \cdot \mathbb{P}_1(\omega \mid A \cap A') = \mathbb{P}_1(\{\omega\} \cap A \cap A') = \mathbb{P}_1(\omega) = \mathbb{P}_1(A)\mathbb{P}_1(\omega \mid A)$. That is, if $\mathbb{P}_1$ and $\mathbb{P}_2$ coincide conditional on some event A, they also coincide conditional on any smaller event A'. In particular, if they coincide on all events of size k, they also coincide on all events of size 2, which we have already dealt with in part (a).

The second way to convince ourselves that part (b) is true, is to observe that none of the arguments used in part (a) really relied on A having size 2, and so the very same reasoning carries through for A of size k.

(c) There are exactly k outcomes in A (namely, $(n,0,\ldots,0), (0,n,0,\ldots), \ldots, (0,\ldots,0,n)$; i.e. each bin could be the full one), and all of them are equally likely under either $\mathbb{P}_1$ or $\mathbb{P}_2$. That is, if $\omega \in A$, then $\mathbb{P}_1(\omega) = \left(\frac{1}{k}\right)^n$, and $\mathbb{P}_2(\omega) = \left[\binom{n+k-1}{k-1}\right]^{-1}$. Consequently, for $\omega \in A$,

$$\mathbb{P}_1(\omega \mid A) = \frac{k^{-n}}{k \cdot k^{-n}} = \frac{1}{k} \qquad \mathbb{P}_2(\omega \mid A) = \frac{\binom{n+k-1}{k-1}^{-1}}{k \cdot \binom{n+k-1}{k-1}^{-1}} = \frac{1}{k}.$$

If $\omega \notin A$, then $\mathbb{P}_1(\omega \mid A) = \mathbb{P}_2(\omega \mid A) = 0$, and so $\mathbb{P}_1(\omega \mid A)$ and $\mathbb{P}_2(\omega \mid A)$ coincide for all $\omega \in \Omega$. This, however, does *not* imply that $\mathbb{P}_1$ and $\mathbb{P}_2$ are the same! Indeed, when computing the probability of $\omega \in A$ above, we saw that $\mathbb{P}_1(\omega) \neq \mathbb{P}_2(\omega)$ (remember that the assumption of part (b) was that the conditional probabilities coincide for *all* events of size k, here we have only shown equality conditional on *one* such event).

(d) Intuitively, throwing balls independently one after another makes it much less likely that all balls stack up in one bin as opposed to spreading out more evenly. This suggests taking, e.g., $A = \{\text{all balls land in bin 1}\}$, whose probability we already computed in part (c). That is, to show that $\mathbb{P}_1(A) < \mathbb{P}_2(A)$, we need to show that $k^{-n} < \binom{n+k-1}{k-1}^{-1}$. Plugging in $k = 3$ and $n = 9$, we have

$$k^{-n} = 3^{-9} = 3^{-2} \cdot 3^{-3} \cdot 3^{-4} = \frac{1}{9 \cdot 27 \cdot 3^4} < \frac{1}{5 \cdot 11} = \frac{2}{11 \cdot 10} = \binom{11}{2}^{-1} = \binom{n+k-1}{k-1}^{-1},$$

as desired.

Conversely, the same reasoning suggests that evenly distributed balls are much more likely under $\mathbb{P}_1$ than under $\mathbb{P}_2$. And indeed, letting $B = \{\text{each bin has exactly three balls}\}$, we have

$$\mathbb{P}_1(B) = \binom{9}{3}\binom{6}{3}\binom{3}{3}\left(\frac{1}{3}\right)^9 = \frac{9!}{(3!)^3 \cdot 3^9} = \frac{7!}{3^{10}} > \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3}{3^6} \cdot \frac{2}{11 \cdot 10} = \frac{7}{3} \cdot \frac{6 \cdot 5}{3^3} \cdot \frac{4}{3}\mathbb{P}_2(B),$$

and since the last factor is bigger than 1, we have $\mathbb{P}_1(B) > \mathbb{P}_2(B)$ as promised (of course, we could have also just plugged all these powers and factorials into a calculator to get the same result).

Note: *This concludes the first part of the homework. The problems below are optional, will not affect your score, and should be attempted only if you have time to spare.*

Part 2: Optional Problems

6 Boy or Girl Paradox

You know Mr. Smith has two children, at least one of whom is a boy. Assume that gender is independent and uniformly distributed, so for any child, the probability that they are a boy is the same as the probability they are a girl, which is $1/2$.

(a) What is the probability that both children are boys?

(b) Now suppose you knock on Mr. Smith's front door and you are greeted by a boy who you correctly deduce to be Mr. Smith's older child. What is the probability that he has two boys? Compare your answer to the answer in part (a).

(c) Show that in a family of n children, of which n_g are girls and n_b are boys, every girl has more brothers than every boy.

(d) Conditional on the youngest child being a girl, what is the probability of her having exactly k brothers? Similarly, conditional on the youngest child being a boy, what is the probability of him having exactly k brothers?

Solution:

(a) Let B_1 be the event that the first child is a boy, and B_2 be the event that the second child is a boy. We are asked to find $\mathbb{P}[(B_1 \cap B_2) \mid (B_1 \cup B_2)]$:

$$\begin{aligned}
\mathbb{P}[(B_1 \cap B_2) \mid (B_1 \cup B_2)] &= \frac{\mathbb{P}[(B_1 \cap B_2) \cap (B_1 \cup B_2)]}{\mathbb{P}[B_1 \cup B_2]} \\
&= \frac{\mathbb{P}[B_1 \cap B_2]}{\mathbb{P}[B_1] + \mathbb{P}[B_2] - \mathbb{P}[B_1 \cap B_2]} \\
&= \frac{(1/2)(1/2)}{1/2 + 1/2 - 1/4} = \frac{1/4}{3/4} \\
&= \frac{1}{3}
\end{aligned}$$

Note: It is tempting to think that because the children's genders are independent, the probability of the second child being a boy given that the first is a boy is simply $1/2$. While this is true, when we write it out in terms of events, we can see that this is not the quantity that we want. See part (b) for more details.

(b) In this part, we want to find $\mathbb{P}[(B_1 \cap B_2) \mid B_1]$:

$$\begin{aligned}
\mathbb{P}[(B_1 \cap B_2) \mid B_1] &= \frac{\mathbb{P}[(B_1 \cap B_2) \cap B_1]}{\mathbb{P}[B_1]} \\
&= \frac{\mathbb{P}[B_1 \cap B_2]}{\mathbb{P}[B_1]} \qquad (1) \\
&= \frac{1/4}{1/2} = \frac{1}{2}
\end{aligned}$$

Note the distinction between this part and part (a), and that a common mistake in determining the answer to part (a) is solving part (b) instead.

(c) Every girl clearly has n_b brothers. However, a boy cannot be his own brother, so every boy has exactly $n_b - 1 < n_b$ brothers.

(d) Since children's genders come independently at probability $1/2$, the probability of the youngest child, regardless of their gender, having exactly k brothers is $\binom{n-1}{k}\left(\frac{1}{2}\right)^{n-1}$. That is, the probability of any child having a certain number of brothers is independent of their gender. However, within a family girls will always have more brothers than boys. This is an instance of Simpson's Paradox.

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 $V_1 = \{v_1, \ldots, v_\ell\}$ and $V_2 = \{w_1, \ldots, w_k\}$, are E_{V_1} and E_{V_2} independent?

(d) Prove that $\binom{n}{k} \leq n^k$. *Optional:* Can you come up with a combinatorial proof? Of course, an algebraic proof would also get full credit.

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

Solution:

(a) There are two choices for each of the $\binom{n}{2}$ pairs of vertices, so 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 $\binom{n}{k}$:

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

$$\leq n \cdot (n-1) \cdots (n-k+1) \tag{3}$$

$$\leq n^k \tag{4}$$

We can also translate the proof above into a combinatorial proof. The number of ways there are to pick k books out of a book shelf with n books to read is smaller than the number of ways you can read k books in n days, where each day you can read any of these k books. (**??**) is less than or equal to the number of ways to place k books from n on the shelf without placement (**??**) (as if you have n books but only k spaces on your bookshelf) is less than or equal to the number of ways to place k books from n on the shelf with replacement (**??**) (as if you own a bookstore and you have k spaces to place n different book titles).

(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

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