CS 70 Discrete Mathematics and Probability Theory
Fall 2018 Alistair Sinclair and Yun Song **HW 9**

1 Double-Check Your Intuition

(a) (i) Let $X \sim \text{Bin}(5, 1/4)$. Let Y be a random variable equal to $5 - X$. What are the distributions of X and Y?

(ii) Let Z be a random variable denoting the result of a die roll (so $1 \le Z \le 6$ uniformly at random). What is $\mathbb{E}[Z^2]$?

For each of the problems below, if you think the answer is "yes" then provide a proof. If you think the answer is "no", then provide a counterexample.

(b) If A and B are integer-valued random variables such that for every integer i, $\mathbb{P}(A = i) = \mathbb{P}(B = i)$, then is $\mathbb{P}(A = B) > 0$?

(c) If C is an integer-valued random variable, then is $\mathbb{E}[C^2] = \mathbb{E}[C]^2$?

(d) If X and Y are random variables and $\mathbb{E}[X] > 100\,\mathbb{E}[Y]$, then is $\mathbb{P}(X > Y) > 1/100$?

(e) If X and Y are random variables taking positive values, then is $\mathbb{E}[\frac{X}{X+Y}] = \frac{\mathbb{E}[X]}{\mathbb{E}[X+Y]}$?

(f) If A, B, C are events such that $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$, then are A, B, C mutually independent?

(g) Is an event A never independent with itself?

(h) If A and B are independent events, then are $\overline{A}$ and $\overline{B}$ independent?

Solution:

(a) (i) $\mathbb{P}(X = 0) = \binom{5}{0}(3/4)^5 = 243/1024$
$\mathbb{P}(X = 1) = \binom{5}{1}(1/4)^1(3/4)^4 = 405/1024$
$\mathbb{P}(X = 2) = \binom{5}{2}(1/4)^2(3/4)^3 = 270/1024$
$\mathbb{P}(X = 3) = \binom{5}{3}(1/4)^3(3/4)^2 = 90/1024$
$\mathbb{P}(X = 4) = \binom{5}{4}(1/4)^4(3/4)^1 = 15/1024$
$\mathbb{P}(X = 5) = \binom{5}{5}(1/4)^5 = 1/1024$
$\mathbb{P}(Y = 0) = 1/1024, \mathbb{P}(Y = 1) = 15/1024, \mathbb{P}(Y = 2) = 90/1024, \mathbb{P}(Y = 3) = 270/1024, \mathbb{P}(Y = 4) = 405/1024, \mathbb{P}(Y = 5) = 243/1024$ because $\mathbb{P}(Y = k) = \mathbb{P}(X = 5 - k)$

(ii) $\mathbb{E}[Z^2] = (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2)/6 = 91/6$

(b) No. Let A be 0 with probability $1/2$ and 1 with probability $1/2$. Let $B = 1 - A$. Then A and B are never equal but B takes the values 0, 1 with probability $1/2$ each.

(c) No. Let $C = Z$ from part (a). Then $\mathbb{E}[C] = 3.5^2 = 12.25 \neq \mathbb{E}[C]^2 = 91/6$.

(d) No. Let $Y = 1$ and let X be a random variable with $\mathbb{P}(X = 100000000) = 1/1000$ and $\mathbb{P}(X = 0) = 999/1000$. Then $\mathbb{P}(X > Y) = 1/1000 < 1/100$.

(e) No. Let $Y = 1$ and $X \sim 1 + \text{Bernouilli}(1/2)$. Then $\mathbb{E}[\frac{X}{X+Y}] = (1/2)(1/2) + (1/2)(2/3) = 5/12$ but $\frac{\mathbb{E}[X]}{\mathbb{E}[X+Y]} = \frac{1+1/2}{1/2+2} = 3/5$.

(f) No. Let A be an event with probability 0 and let B be some event with probability $1/2$ and let $C = B$. Then $\mathbb{P}(A \cap B \cap C) \leq \mathbb{P}(A) = 0 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ but B and C are clearly not independent.

(g) No. If $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$, then $\mathbb{P}(A \cap A) = \mathbb{P}(A) = \mathbb{P}(A)^2$.

(h) Yes. We have $\mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\overline{A \cup B}) = 1 - \mathbb{P}(A \cup B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B) = (1 - \mathbb{P}(A))(1 - \mathbb{P}(B)) = \mathbb{P}(\bar{A})\mathbb{P}(\bar{B})$.

2 Airport Revisited

(a) Suppose that there are n airports arranged in a circle. A plane departs from each airport, and randomly chooses an airport to its left or right to fly to. What is the expected number of empty airports after all planes have landed?

(b) Now suppose that we still have n airports, but instead of being arranged in a circle, they form a general graph, where each airport is denoted by a vertex, and an edge between two airports indicates that a flight is permitted between them. There is a plane departing from each airport and randomly chooses a neighboring destination where a flight is permitted. What is the expected number of empty airports after all planes have landed? (Express your answer in terms of $N(i)$ - the set of neighboring airports of airport i, and $\deg(i)$ - the number of neighboring airports of airport i).

Solution:

(a) Let X_i be the indicator variable denoting whether airport i ends up empty. This can happen if and only if planes from both of its neighboring airports are flying elsewhere, and this happens with a probability of $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$. Hence the expected number of empty airports is

$$\mathbb{E}[X_1 + \ldots + X_n] = \frac{n}{4}$$

(b) Similar to the previous part, we now have $\mathbb{E}[X_i] = P(X_i = 1) = \prod_{j \in N(i)} \left(1 - \frac{1}{\deg(j)}\right)$. Hence

$$\mathbb{E}[X_1+\ldots+X_n]=\sum_{i=1}^{n}\prod_{j\in N(i)}\left(1-\frac{1}{\deg(j)}\right)$$

3 Fizzbuzz

(a) Fizzbuzz is a classic software engineering interview question. You are given a natural number n, and for each integer i from 1 to n you have to print either "fizzbuzz" if i is divisible by 15, "fizz" if i is divisible by 3 but not 15, "buzz" if i is divisible by 5 but not 15, or the integer itself if i is not divisible by 3 or 5.

If n is a multiple of 15, then how many printed lines will contain an integer?

(Hint: If you pick a line uniformly at random, then what is the probability that the printed line contains an integer?)

(b) Recall the Euler totient function from Homework 4:

$$\phi(n)=\left|\{i:1\leq i\leq n,\texttt{gcd}(i,n)=1\}\right|.$$

Suppose $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$ where $p_1,p_2,\ldots,p_k$ are distinct primes and $\alpha_1,\alpha_2,\ldots,\alpha_k$ are positive integers. Prove that

$$\frac{\phi(n)}{n}=\prod_{j=1}^{k}\left(1-\frac{1}{p_j}\right)$$

Solution:

(a) Let's sample a random integer i from $\{1,2,\ldots,n\}$. Let A be the event that i is divisible by 3 and let B be the event that j is divisible by 5. Since n is divisible by 15, then $\mathbb{P}(A)=1/3$, $\mathbb{P}(B)=1/5$, and $\mathbb{P}(A\cap B)=1/15$ because the last probability denotes the probability i is divisible by $3\cdot 5=15$ (we use the fact that 3 and 5 are coprime).

Note that $1-\mathbb{P}(A\cup B)$ is the probability that i is not divisible by 3 or 5, i.e. the probability that line i is an integer.

We have

$$\begin{aligned}\mathbb{P}(A\cup B)&=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B)\\&=\frac{1}{3}+\frac{1}{5}-\frac{1}{15}\\&=\frac{7}{15}\end{aligned}$$

so $1-\mathbb{P}(A\cup B)=\frac{8}{15}$. Thus $\frac{8n}{15}$ lines will contain an integer.

(b) Sample i uniformly at random from $\{1,2,\ldots,n\}$. Let A_m, for $1 \le m \le k$ denote the event "i is divisible by p_m". Since n is divisible by p_m, then $Pr(A_m) = 1/p_m$. Note that for any $\{m_1, m_2, \ldots, m_t\} \subseteq \{1,2,\ldots,k\}$ we have

$$\mathbb{P}(A_{m_1} \cap \cdots \cap A_{m_t}) = \frac{1}{p_{m_1} p_{m_2} \cdots p_{m_t}}$$

because $p_{m_1}, p_{m_2}, \cdots, p_{m_t}$ are all distinct primes.

Let B denote the event "i is relatively prime to n". Then

$$\begin{aligned}
\mathbb{P}(B) &= 1 - \mathbb{P}(A_1 \cup A_2 \cdots \cup A_k) \\
&= 1 - \sum_{m=1}^{k} \mathbb{P}(A_m) + \sum_{m_1<m_2}^{k} \mathbb{P}(A_{m_1} \cap A_{m_2}) - \sum_{m_1<m_2<m_3}^{k} \mathbb{P}(A_{m_1} \cap A_{m_2} \cap A_{m_3}) + \cdots \\
&= 1 - \sum_{m=1}^{k} \frac{1}{p_m} + \sum_{m_1<m_2}^{k} \frac{1}{p_{m_1} p_{m_2}} - \sum_{m_1<m_2<m_3}^{k} \frac{1}{p_{m_1} p_{m_2} p_{m_3}} + \cdots \\
&= \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right)
\end{aligned}$$

by the Principle of Inclusion-Exclusion.

Since $\mathbb{P}(B) = \phi(n)/n$, then the result follows.

Extra: We did not cover the Chinese Remainder Theorem in class, but it can also be used to solve this problem. Note that the integers mod n are in one-to-one correspondence with all the possible $(y_1, y_2, \ldots, y_t)$ where y_i is taken mod $p_i^{m_i}$. Thus, the probability some integer in $\{1,2,\ldots,n\}$ is coprime to n is equal to the probability that $y_i \neq 0$ mod $p_i^{m_i}$ for all i. This is the product of $(1 - 1/p_i)$ over all i.

4 Cliques in Random Graphs

Consider 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. So for example if $n = 2$, then with probability $1/2$, $G = (V,E)$ is the graph consisting of two vertices connected by an edge, and with probability $1/2$ it is the graph consisting of two isolated vertices.

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

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

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

Optional: Can you come up with a combinatorial proof? Of course, an algebraic proof would also get full credit.

(d) 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 (c).

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

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

$$\leq n \cdot (n-1) \cdots (n-k+1) \tag{2}$$

$$\leq n^k \tag{3}$$

We can also translate the proof above into a combinatorial proof. The number of ways there are to read k books out of n (**??**) 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).

(d) 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

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

5 Balls and Bins, All Day Every Day

You throw n balls into n bins uniformly at random, 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 n 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 bins 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, $\mathbb{P}(\bigcup_{i=1}^{n} A_k) \le \sum_{i=1}^{n} \mathbb{P}(A_k) = np$.

(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 the bin contains k balls and $k \in \{1, \ldots, n\}$ (note that $k \neq 0$ since we know at least one ball has landed in this bin). 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,

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