CS 70	Discrete Mathematics and Probability Theory	
Spring 2019	Babak Ayazifar and Satish Rao	HW 9

1 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 $\binom{n}{k} \leq n^k$.

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

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

2 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:

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, $i = 1, \ldots, n$, 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 $B_1, B_2, ..., B_n$ 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 $(i, j) \in \{1, \ldots, n\}^2$ 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 $(-1)^{k+1} \sum_{(i_1,\ldots,i_k)} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k})$, where the k-tuples in the summation range over all $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$ 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,

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

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 $n \to \infty$):

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

3 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 $k = n/2, \dots, n$ 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, $k = 1, \dots, n$, 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 $k \in \{1, \dots, n\}$ (note that $k \ne 0$ 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

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

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 Babak's Dice

Professor Ayazifar rolls three fair six-sided dice.

(a) Let X denote the maximum of the three values rolled. What is the distribution of X (that is, $\mathbb{P}[X = x]$ for $x = 1,2,3,4,6$)? You can leave your final answer in terms of "x". [*Hint*: Try to first compute $\mathbb{P}[X \le x]$ for $x = 1,2,3,4,5,6$].

(b) Let Y denote the minimum of the three values rolled. What is the distribution of Y?

Solution:

(a) Let X denote the maximum of the three values rolled. We are interested in $\mathbb{P}(X = x)$, where $x = 1,2,3,4,5,6$. First, define X_1, X_2, X_3 to be the values rolled by the first, second, and third dice. These random variables are i.i.d. and uniformly distributed between 1 and 6 inclusive.

Following the hint we first compute $\mathbb{P}[X \le x]$ for $x = 1,2,3,4,5,6$:

$$\mathbb{P}(X \le x) = \mathbb{P}(X_1 \le x)\mathbb{P}(X_2 \le x)\mathbb{P}(X_3 \le x) = \left(\frac{x}{6}\right)\left(\frac{x}{6}\right)\left(\frac{x}{6}\right) = \left(\frac{x}{6}\right)^3$$

Then, observing that $\mathbb{P}(X = x) = \mathbb{P}(X \le x) - \mathbb{P}(X \le x-1)$:

$$\mathbb{P}(X = x) = \left(\frac{x}{6}\right)^3 - \left(\frac{x-1}{6}\right)^3 = \frac{3x^2 - 3x + 1}{216} = \begin{cases} \frac{1}{216}, & x = 1 \\ \frac{7}{216}, & x = 2 \\ \frac{19}{216}, & x = 3 \\ \frac{37}{216}, & x = 4 \\ \frac{61}{216}, & x = 5 \\ \frac{91}{216}, & x = 6 \end{cases}$$

One can confirm that $\sum_{x=1}^{6} \mathbb{P}(X = x) = 1$.

(b) Similarly to the previous part, we first compute $\mathbb{P}[Y \geq y]$.

$$\mathbb{P}(Y \geq y) = \mathbb{P}(X_1 \geq y)\mathbb{P}(X_2 \geq y)\mathbb{P}(X_3 \geq y) = \left(\frac{6-(y-1)}{6}\right)\left(\frac{6-(y-1)}{6}\right)\left(\frac{6-(y-1)}{6}\right) = \left(\frac{7-y}{6}\right)^3 .$$

Then, observing that $\mathbb{P}(Y = y) = \mathbb{P}(Y \geq y) - \mathbb{P}(Y \geq y-1)$:

$$\mathbb{P}[Y = y] = \left(\frac{7-y}{6}\right)^3 - \left(\frac{6-y}{6}\right)^3 .$$

5 Testing Model Planes

Dennis is testing model airplanes. He starts with n model planes which each independently have probability p of flying successfully each time they are flown, where $0 < p < 1$. Each day, he flies every single plane and keeps the ones that fly successfully (i.e. don't crash), throwing away all other models. He repeats this process for many days, where each "day" consists of Dennis flying any remaining model planes and throwing away any that crash. Let X_i be the random variable representing how many model planes remain after i days. Note that $X_0 = n$. Justify your answers for each part.

(a) What is the distribution of X_1? That is, what is $\mathbb{P}[X_1 = k]$?

(b) What is the distribution of X_2? That is, what is $\mathbb{P}[X_2 = k]$? Show that X_2 follows a binomial distribution by finding some n' and p' such that $X_2 \sim \text{Binom}(n', p')$.

(c) Repeat the previous part for X_t for arbitrary $t \geq 1$.

(d) What is the probability that at least one model plane still remains (has not crashed yet) after t days? Do not have any summations in your answer.

(e) Considering only the first day of flights, is the event A_1 that the first and second model planes crash independent from the event B_1 that the second and third model planes crash? Recall that two events A and B are independent if $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$. Prove your answer using this definition.

(f) Considering only the first day of flights, let A_2 be the event that the first model plane crashes *and* exactly two model planes crash in total. Let B_2 be the event that the second plane crashes on the first day. What must n be equal to in terms of p such that A_2 is independent from B_2? Prove your answer using the definition of independence stated in the previous part.

(g) Are the random variables X_i and X_j, where $i < j$, independent? Recall that two random variables X and Y are independent if $\mathbb{P}[X = k_1 \cap Y = k_2] = \mathbb{P}[X = k_1]\mathbb{P}[Y = k_2]$ for all k_1 and k_2. Prove your answer using this definition.

Solution:

(a) Since Dennis is performing n trials (flying a plane), each with an independent probability of "success" (not crashing), we have $X_1 \sim \text{Binom}(n, p)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$, for $0 \le k \le n$.

(b) Each model plane independently has probability p^2 of surviving both days. Whether a model plane survives both days is still independent from whether any other model plane survives both days, so we can say $X_2 \sim \text{Binom}(n, p^2)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^{2k} (1-p^2)^{n-k}$, for $0 \le k \le n$.

(c) By extending the previous part we see each model plane has probability p^t of surviving t days, so $X_t \sim \text{Binom}(n, p^t)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^{tk} (1-p^t)^{n-k}$, for $0 \le k \le n$.

(d) We consider the complement, the probability that no model planes remain after t days. By the previous part we know this to be $\mathbb{P}[X_t = 0] = \binom{n}{0} p^{t(0)} (1-p^t)^{n-0} = (1-p^t)^n$. So the probability of at least model plane remaining after t days is $1 - (1-p^t)^n$.

(e) No. $\mathbb{P}[A_1 \cap B_1]$ is the probability that the first three model planes crash, which is $(1-p)^3$. But $\mathbb{P}[A_1]\mathbb{P}[B_1] = (1-p)^2(1-p)^2 = (1-p)^4$. So $\mathbb{P}[A_1 \cap B_1] \ne \mathbb{P}[A_1]\mathbb{P}[B_1]$ and A_1 and B_1 are not independent.

(f) $\mathbb{P}[A_1 \cap B_1]$ is the probability that only the first model plane and second model plane crash, which is $(1-p)^2 p^{n-2}$. $\mathbb{P}[A_1]$ is the probability that the first model plane crashes, and exactly one of the remaining $n-1$ model planes crashes, so $\mathbb{P}[A_2] = (1-p) \cdot \binom{n-1}{1} (1-p) p^{n-1-1} = (n-1)(1-p)^2 p^{n-2}$. Trivially, we have $\mathbb{P}[B_2] = 1-p$, so $\mathbb{P}[A_2]\mathbb{P}[B_2] = (n-1)(1-p)^3 p^{n-2}$ which is equal to $\mathbb{P}[A_2 \cap B_2] = (1-p)^2 p^{n-2}$ only when $(n-1)(1-p) = 1$, or when $n = \frac{1}{1-p} + 1$.

(g) No. Let $k_1 = 0$ and $k_2 = 1$. Then $\mathbb{P}[X_i = k_1 \cap X_j = k_2] = 0$ because the sequence of X_i is by definition non-increasing. But $\mathbb{P}[X_i = k_1] > 0$ and $\mathbb{P}[X_j = k_2] > 0$ so $\mathbb{P}[X_i = k_1]\mathbb{P}[X_j = k_2] > 0$.

6 Cookie Jars

You have two jars of cookies, each of which starts with n cookies initially. Every day, when you come home, you pick one of the two jars randomly (each jar is chosen with probability $1/2$) and eat one cookie from that jar. One day, you come home and reach inside one of the jars of cookies, but you find that is empty! Let X be the random variable representing the number of remaining cookies in non-empty jar at that time. What is the distribution of X?

Solution: Assume that you found jar 1 empty. The probability that $X = k$ and you found jar 1 empty is computed as follows. In order for there to be k cookies remaining, you must have eaten a cookie for $2n-k$ days, and then you must have chosen jar 1 (to discover that it is empty). Within those $2n-k$ days, exactly n of those days you chose jar 1. The probability of this is $\binom{2n-k}{n} 2^{-(2n-k)}$. Furthermore, the probability that you then discover jar 1 is empty the day after is $1/2$. So, the probability that $X = k$ and you discover jar 1 empty is $\binom{2n-k}{n} 2^{-(2n-k+1)}$. However, we assumed that we discovered jar 1 to be empty; the probability that $X = k$ and jar 2 is empty is the same by symmetry, so the overall probability that $X = k$ is:

$$\mathbb{P}(X = k) = \binom{2n-k}{n} \frac{1}{2^{2n-k}}, \qquad k \in \{0, \dots, n\}.$$