CS 70 Discrete Mathematics and Probability Theory
Spring 2017 Rao

HW 9

1 Sundry

Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. (In case of homework party, you can also just describe the group.) How did you work on this homework? Working in groups of 3-5 will earn credit for your "Sundry" grade.

Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up. (Signature here)

2 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 the probability that the graph contains a k-clique for $k = 4\lceil \log n \rceil + 1$ is at most $1/n$.

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) 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,

$$\Pr\left[\bigcup_{S\subseteq V,|S|=k} A_S\right] \le \sum_{S\subseteq V,|S|=k} \Pr[A_S] = \sum_{S\subseteq V,|S|=k} \frac{1}{2^{\binom{k}{2}}}.$$

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

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

3 Student Request Collector

After a long night of debugging, Alvin has just perfected the new homework party/office hour queue system. CS 70 students sign themselves up for the queue, and TAs go through the queue, resolving requests one by one. Unfortunately, our newest TA (let's call him TA Bob) does not understand how to use the new queue: instead of resolving the requests in order, he always uses the Random Student button, which (as the name suggests) chooses a random student in the queue for him. To make matters worse, after helping the student, Bob forgets to click the Resolve button, so the student still remains in the queue! For this problem, assume that there are n total students in the queue.

(a) Suppose that Bob has already helped k students. What is the probability that the Random Student button will take him to a student who has not already been helped?

(b) Let X_i^r be the event that TA Bob has not helped student i after pressing the Random Student button a total of r times. What is $\Pr[X_i^r]$? Assume that the results of the Random Student button are independent of each other. Now approximate the answer using the inequality $1 - x \le e^{-x}$.

(c) Let T_r represent the event that TA Bob presses the Random Student button r times, but still has not been able to help all n students. (In other words, it takes TA Bob longer than r Random Student button presses before he manages to help every student). What is T_r in terms of the events X_i^r? (*Hint*: Events are subsets of the probability space Ω, so you should be thinking of set operations...)

(d) Using your answer for the previous part, what is an upper bound for $\Pr[T_r]$? (You may leave your answer in terms of $\Pr[X_i^r]$. Use the inequality $1 - x \le e^{-x}$ from before.)

(e) Now let $r = \alpha n \ln n$. What is an upper bound for $\Pr[X_i^r]$?

(f) Calculate an upper bound for $\Pr[T_r]$ using the same value of r as before. (This is more formally known as a bound on the tail probability of the distribution of button presses required to help every student. This distribution will be explored in more detail later, in the context of random variables.)

(g) What value of r do you need to bound the tail probability by $1/n^2$? In other words, how many button presses are needed so that the probability that TA Bob has not helped every student is at most $1/n^2$?

Solution:

(a) There are $n-k$ students who have not been helped, and the probability that one of these students is chosen is $(n-k)/n$ or $1-k/n$.

(b) The probability that student i is chosen by the Random Student button is $1/n$, so the complement of this probability is $1-1/n$. Using the assumption of independence:

$$\Pr[X_i^r] = \left(1 - \frac{1}{n}\right)^r \leq e^{-r/n}$$

(c) T_r is the event that TA Bob has pressed the button r times, but has not been able to help either student 1, or student 2, or student 3, ... This is the union: $T_r = \bigcup_{i=1}^n X_i^r$.

(d) Use the union bound. $\Pr[T_r] \leq n \cdot \Pr[X_i^r] \leq ne^{-r/n}$.

(e) Plug in for r. $\Pr[X_i^r] \leq e^{-r/n} = e^{-\alpha \ln n} = n^{-\alpha}$.

(f) A quick application of the union bound derived in the previous parts: $\Pr[T_r] \leq n \cdot \Pr[X_i^r] = n^{1-\alpha}$.

(g) Set $1-\alpha = -2$, which is $\alpha = 3$. This gives $r = 3n\ln n$. (Side-note: This problem is more commonly known as the coupon collector's problem. Once we cover random variables, we will see that the expected number of button presses required to help every student is $\Theta(n \log n)$.)

4 Combinatorial Coins

Allen and Alvin are flipping coins for fun. Allen flips a fair coin k times and Alvin flips $n-k$ times. In total there are n coin flips.

(a) Use a combinatorial proof to show that

$$\sum_{i=0}^{k} \binom{k}{k-i}\binom{n-k}{i} = \binom{n}{k}.$$

You may assume that $n-k \geq k$.

(b) Prove that the probability that Allen and Alvin flip the same number of heads is equal to the probability that there are a total of k heads.

Solution:

(a) On the right-hand side of the equation, we are calculating the total number of ways to choose k items from a group of n items. On the left-hand side, we first partition the whole group into two subgroups with k items and $n-k$ items. In order to choose k items in total from those two groups, we pick $k-i$ items from the first group and i items from the second group. Therefore, the total number of ways to do so is simply a summation with all possible values of i.

(b) Let's first figure out the probability that there are a total of k heads. Suppose X is a random variable that denotes the total number of heads in the n flips. It can be seen clearly that X follows a binomial distribution with parameter $p = 1/2$ and n. Then the probability that there are a total of k heads can be calculated using binomial distribution:

$$\Pr(X = k) = \binom{n}{k}\left(\frac{1}{2}\right)^n.$$

Now, let's do a mathematical derivation to prove that it's the same as the probability that Allen and Alvin flip the same number of heads. Let A be the random variable representing the number of heads Allen gets, and let B be the random variable representing the number of heads Alvin gets.

$$\begin{aligned}
\Pr(\text{Allen and Alvin flip the same number of heads}) &= \sum_i \Pr(A = i, B = i) \\
&= \sum_i \binom{k}{i}\left(\frac{1}{2}\right)^k \binom{n-k}{i}\left(\frac{1}{2}\right)^{n-k} \\
&= \sum_i \binom{k}{i}\binom{n-k}{i}\left(\frac{1}{2}\right)^n \\
&= \sum_i \binom{k}{k-i}\binom{n-k}{i}\left(\frac{1}{2}\right)^n \\
&= \binom{n}{k}\left(\frac{1}{2}\right)^n.
\end{aligned}$$

For the last equality, we used the identity from part (a), $\sum_i \binom{k}{k-i}\binom{n-k}{i} = \binom{n}{k}$. In order to use the identity from part (a), we need the assumption that $n-k \geq k$. However, since the probability of obtaining exactly k heads is the same as the probability of obtaining exactly $n-k$ heads for a fair coin, and Alvin and Allen are interchangeable in the problem, we can interchange $n-k$ and k if necessary to ensure that $n-k \geq k$, so the above derivation works in the case $n-k < k$ case as well.

5 Geometric Distribution

Two faulty machines, M_1 and M_2, are repeatedly run synchronously in parallel (i.e., both machines execute one run, then both execute a second run, and so on). On each run, M_1 fails with probability

p_1 and M_2 fails with probability p_2, all failure events being independent. Let the random variables X_1, X_2 denote the number of runs until the first failure of M_1, M_2 respectively; thus X_1, X_2 have geometric distributions with parameters p_1, p_2 respectively. Let X denote the number of runs until the first failure of *either* machine. Show that X also has a geometric distribution, with parameter $p_1 + p_2 - p_1 p_2$.

Solution:

We have that $X_1 \sim \text{Geom}(p_1)$ and $X_2 \sim \text{Geom}(p_2)$. Also, X_1, X_2 are independent r.v.'s. We also use the following definition of the minimum:

$$\min(x,y) = \begin{cases} x & \text{if } x \leq y; \\ y & \text{if } x > y. \end{cases}$$

Now, for all $k \in \{1,2,\ldots\}$, $\min(X_1,X_2) = k$ is equivalent to $(X_1 = k) \cap (X_2 \geq k)$ or $(X_2 = k) \cap (X_1 > k)$. Hence,

$$\begin{aligned} \Pr[X = k] &= \Pr[\min(X_1,X_2) = k] \\ &= \Pr[(X_1 = k) \cap (X_2 \geq k)] + \Pr[(X_2 = k) \cap (X_1 > k)] \\ &= \Pr[X_1 = k] \cdot \Pr[X_2 \geq k] + \Pr[X_2 = k] \cdot \Pr[X_1 > k] \end{aligned}$$

(since X_1 and X_2 are independent)

$$= [(1-p_1)^{k-1} p_1](1-p_2)^{k-1} + [(1-p_2)^{k-1} p_2](1-p_1)^k$$

(since X_1 and X_2 are geometric)

$$\begin{aligned} &= ((1-p_1)(1-p_2))^{k-1} (p_1 + p_2(1-p_1)) \\ &= (1 - p_1 - p_2 + p_1 p_2)^{k-1} (p_1 + p_2 - p_1 p_2). \end{aligned}$$

But this final expression is precisely the probability that a geometric r.v. with parameter $p_1 + p_2 - p_1 p_2$ takes the value k. Hence $X \sim \text{Geom}(p_1 + p_2 - p_1 p_2)$, and $\mathbf{E}[X] = (p_1 + p_2 - p_1 p_2)^{-1}$.

An alternative, slightly cleaner approach is to work with the *tail probabilities* of the geometric distribution, rather than with the usual point probabilities as above. In other words, we can work with $\Pr[X \geq k]$ rather than with $\Pr[X = k]$; clearly the values $\Pr[X \geq k]$ specify the values $\Pr[X = k]$ since $\Pr[X = k] = \Pr[X \geq k] - \Pr[X \geq (k+1)]$, so it suffices to calculate them instead. We then get the following argument:

$$\begin{aligned} \Pr[X \geq k] &= \Pr[\min(X_1,X_2) \geq k] \\ &= \Pr[(X_1 \geq k) \cap (X_2 \geq k)] \\ &= \Pr[X_1 \geq k] \cdot \Pr[X_2 \geq k] && \text{since } X_1, X_2 \text{ are independent} \\ &= (1-p_1)^{k-1}(1-p_2)^{k-1} && \text{since } X_1, X_2 \text{ are geometric} \\ &= ((1-p_1)(1-p_2))^{k-1} \\ &= (1 - p_1 - p_2 + p_1 p_2)^{k-1}. \end{aligned}$$

This is the tail probability of a geometric distribution with parameter $p_1 + p_2 - p_1 p_2$, so we are done.

6 Poisson Distribution

(a) It is fairly reasonable to model the number of customers entering a shop during a particular hour as a Poisson random variable. Assume that this Poisson random variable X has mean λ. Suppose that whenever a customer enters the shop they leave the shop without buying anything with probability p. Assume that customers act independently, i.e. you can assume that they each simply flip a biased coin to decide whether to buy anything at all. Let us denote the number of customers that buy something as Y and the number of them that do not buy anything as Z (so $X = Y + Z$). What is the probability that $Y = k$ for a given k? How about $\Pr[Z = k]$? Prove that Y and Z are Poisson random variables themselves.

Hint: You can use the identity

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

(b) Prove that Y and Z are independent.

(c) Assume that you were given two independent Poisson random variables X_1, X_2. Assume that the first has mean λ_1 and the second has mean λ_2. Prove that $X_1 + X_2$ is a Poisson random variable with mean $\lambda_1 + \lambda_2$.

Hint: Recall the binomial theorem.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Solution:

(a) We consider all possible ways that the event $Y = k$ might happen: namely, $k + j$ people enter the shop ($X = k + j$) and then exactly k of them choose to buy something. That is,

$$\begin{aligned}
\Pr[Y = k] &= \sum_{j=0}^{\infty} \Pr[X = k+j] \cdot \Pr[Y = k \mid X = k+j] \\
&= \sum_{j=0}^{\infty} \left(\frac{\lambda^{k+j}}{(k+j)!} e^{-\lambda} \right) \cdot \left(\binom{k+j}{k} p^j (1-p)^k \right) \\
&= \sum_{j=0}^{\infty} \frac{\lambda^{k+j}}{(k+j)!} e^{-\lambda} \cdot \frac{(k+j)!}{k!\,j!} p^j (1-p)^k \\
&= \frac{(\lambda(1-p))^k e^{-\lambda}}{k!} \cdot \sum_{j=0}^{\infty} \frac{(\lambda p)^j}{j!} \\
&= \frac{(\lambda(1-p))^k e^{-\lambda}}{k!} \cdot e^{\lambda p} \\
&= \frac{(\lambda(1-p))^k e^{-\lambda(1-p)}}{k!}.
\end{aligned}$$

Hence, Y follows the Poisson distribution with parameter $\lambda(1-p)$. The case for Z is completely analogous:

$$\Pr[Z=k] = \frac{(\lambda p)^k e^{-\lambda p}}{k!}$$

and Z follows the Poisson distribution with parameter λp.

(b) If Y and Z are independent, then $\Pr(Y=y, Z=z) = \Pr(Y=y)\cdot\Pr(Z=z)$:

$$\begin{aligned}
\Pr(Y=y,Z=z) &= \sum_{x=0}^{\infty} \Pr(X=x, Y=y, Z=z) \\
&= \sum_{x=0}^{\infty} \Pr(Y=y, Z=z \mid X=x)\Pr(X=x) \\
&= \Pr(Y=y, Z=z \mid X=y+z)\Pr(X=y+z) \\
&= \frac{\cancel{(y+z)!}}{y!z!} p^z (1-p)^y \frac{e^{-\lambda}\lambda^{y+z}}{\cancel{(y+z)!}} \\
&= \frac{e^{-\lambda(1-p)}(\lambda(1-p))^y}{y!}\cdot\frac{e^{-\lambda p}(\lambda p)^z}{z!} \\
&= \Pr(Y=y)\cdot\Pr(Z=z)
\end{aligned}$$

(c) To show that X_1+X_2 is a Poisson random variable with mean $\lambda_1+\lambda_2$, we have show that

$$\Pr[(X_1+X_2)=i] = \frac{(\lambda_1+\lambda_2)^i}{i!}e^{-(\lambda_1+\lambda_2)}.$$

We proceed as follows:

$$\begin{aligned}
\Pr[(X_1+X_2)=i] &= \sum_{k=0}^{i} \Pr[X_1=k, X_2=(i-k)] \\
&= \sum_{k=0}^{i} \frac{\lambda_1^k}{k!}e^{-\lambda_1}\cdot\frac{\lambda_2^{i-k}}{(i-k)!}e^{-\lambda_2} \\
&= e^{-\lambda_1}e^{-\lambda_2}\sum_{k=0}^{i}\frac{1}{k!(i-k)!}\lambda_1^k\lambda_2^{i-k} \\
&= \frac{e^{-\lambda_1}e^{-\lambda_2}}{i!}\sum_{k=0}^{i}\frac{i!}{k!(i-k)!}\lambda_1^k\lambda_2^{i-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{i!}\sum_{k=0}^{i}\binom{i}{k}\lambda_1^k\lambda_2^{i-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{i!}(\lambda_1+\lambda_2)^i
\end{aligned}$$

To go from the second-to-last line to the last line, we use the binomial expansion.

7 Poisson Coupling

Consider the following discrete joint distribution for $p \in [0,1]$.

$$\begin{aligned}
\Pr(X=0,Y=0) &= 1-p, \\
\Pr(X=1,Y=y) &= \frac{e^{-p}p^y}{y!}, \qquad y=1,2,\ldots, \\
\Pr(X=1,Y=0) &= e^{-p}-(1-p), \\
\Pr(X=x,Y=y) &= 0, \qquad \text{otherwise.}
\end{aligned}$$

(a) Recall that all valid distributions satisfy two important properties. Argue that this distribution is a valid joint distribution.

(b) Show that X has the Bernoulli distribution with probability p.

(c) Show that Y has the Poisson distribution with parameter $\lambda = p$.

(d) Show that $\Pr(X \neq Y) \leq p^2$.

Now, let X_i, $i=1,2,\ldots$ be a sequence of random variables with probabilities p_i, $i=1,2,\ldots$. Similarly, let Y_i be a Poisson random variable with parameter $\lambda = p_i$, $i=1,2,\ldots$. The X_i and Y_i are coupled, so that they have the joint distribution described above (with $p=p_i$), but for $i \neq j$, (X_i,Y_i) and (X_j,Y_j) are independent.

We will now introduce a coupling argument which shows that the distribution of $\sum_{i=1}^n X_i$ approaches a Poisson distribution with parameter $\lambda = p_1 + \cdots + p_n$.

(e) A common way to measure the "distance" between two probability distributions is known as the total variation norm, and it is given by

$$d(X,Y) = \frac{1}{2}\sum_{k=0}^{\infty} |\Pr(X=k)-\Pr(Y=k)|.$$

Show that $d(X,Y) \leq \Pr(X \neq Y)$. [*Hint*: Use the Law of Total Probability to split up the events according to $\{X=Y\}$ and $\{X \neq Y\}$.]

(f) Show that $\Pr(\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i) \leq \sum_{i=1}^n \Pr(X_i \neq Y_i)$. [*Hint*: Maybe try the Union Bound.]

(g) Finally, for the X_i and Y_i defined above, show that $d(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i) \leq \sum_{i=1}^n p_i^2$.

Solution:

(a) We need to verify that the probabilities sum to 1. Indeed,

$$\begin{aligned}
\Pr(X=0,Y=0)+\Pr(X=1,Y=0)+\sum_{y=1}^{\infty}\Pr(X=1,Y=y) &= e^{-p}+\sum_{y=1}^{\infty}\frac{e^{-p}p^y}{y!} \\
&= \cancel{e^{-p}}+1-\cancel{e^{-p}} = 1.
\end{aligned}$$

Also, the probabilities are non-negative, since $e^{-p} \geq 1-p$ always.

(b) We know that $\Pr(X=0)=\Pr(X=0,Y=0)=1-p$, and that $\Pr(X=x)=0$ for any $x\notin\{0,1\}$. Then $\Pr(X=0)+\Pr(X=1)=1$, so $\Pr(X=1)=p$. This is a sufficient approach, but to be fully explicit, we can verify through direct calculation that $\Pr(X=1)=p$:

$$\Pr(X=1)=\Pr(X=1,Y=0)+\sum_{y=1}^{\infty}\Pr(X=1,Y=y)=e^{-p}-(1-p)+\sum_{y=1}^{\infty}\frac{e^{-p}p^y}{y!}$$
$$=\cancel{e^{-p}}-1+p+1\cancel{-e^{-p}}=p.$$

Hence, X has the Bernoulli distribution with probability of success p.

(c) We see that $\Pr(Y=0)=\Pr(X=0,Y=0)+\Pr(X=1,Y=0)=e^{-p}$, and for $y=1,2,\dots$ we have

$$\Pr(Y=y)=\Pr(X=1,Y=y)=\frac{e^{-p}p^y}{y!}.$$

This is indeed the Poisson distribution with rate $\lambda=p$.

(d) We can recognize that $\Pr(X\neq Y)=1-\Pr(X=Y)$:

$$\begin{aligned}\Pr(X\neq Y)&=1-\Pr(X=Y)=1-\Pr(X=0,Y=0)-\Pr(X=1,Y=1)\\&=1-(1-p)-\frac{e^{-p}p^1}{1!}\\&=p-e^{-p}p\\&=p(1-e^{-p})\le p^2.\end{aligned}$$

In the last line, we are using $1-e^{-p}\le p$. Note that this follows from $e^{-p}\ge 1-p$ by rearranging the inequality.

Alternatively, we can compute $\Pr(X\neq Y)$ directly:

$$\Pr(X\neq Y)=\Pr(X=1,Y=0)+\Pr(X=1,Y\ge 2)=e^{-p}-(1-p)+\sum_{y=2}^{\infty}\frac{e^{-p}p^y}{y!}$$
$$=e^{-p}-(1-p)+1-e^{-p}-pe^{-p}=p(1-e^{-p})\le p^2.$$

(e) One has

$$\begin{aligned}d(X,Y)&=\frac{1}{2}\sum_{k=0}^{\infty}|\Pr(X=k)-\Pr(Y=k)|\\&=\frac{1}{2}\sum_{k=0}^{\infty}|\Pr(X=k,X=Y)+\Pr(X=k,X\neq Y)-\Pr(Y=k,X=Y)\\&\qquad-\Pr(Y=k,X\neq Y)|.\end{aligned}$$

Note that the event $\{X = k, X = Y\}$ is the same as $\{Y = k, X = Y\}$ (they both equal the event $\{X = Y = k\}$). Hence, these terms cancel and we have

$$\begin{aligned} d(X,Y) &= \frac{1}{2}\sum_{k=0}^{\infty} |\Pr(X = k, X \neq Y) - \Pr(Y = k, X \neq Y)| \\ &\leq \frac{1}{2}\left(\sum_{k=0}^{\infty} \Pr(X = k, X \neq Y) + \sum_{k=0}^{\infty} \Pr(Y = k, X \neq Y)\right) = \frac{1}{2}(\Pr(X \neq Y) + \Pr(X \neq Y)). \end{aligned}$$

We see that the factor of $1/2$ disappears and we are left with

$$d(X,Y) \leq \Pr(X \neq Y). \tag{1}$$

(f) Note that the event $\{\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i\} \subseteq \{\exists i\ X_i \neq Y_i\}$, since if the two summations $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n Y_i$ are different, then there must be at least one term which is different between the summations. Now, we can write

$$\Pr\left(\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i\right) \leq \Pr(X_i \neq Y_i \text{ for some } i) = \Pr\left(\bigcup_{i=1}^n \{X_i \neq Y_i\}\right).$$

Now, we apply the Union Bound to the term on the right to obtain

$$\Pr\left(\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i\right) \leq \sum_{i=1}^n \Pr(X_i \neq Y_i). \tag{2}$$

(g) Thanks to the inequalities we have proved, we can write down

$$d\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right) \underbrace{\leq}_{(1)} \Pr\left(\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i\right) \underbrace{\leq}_{(2)} \sum_{i=1}^n \Pr(X_i \neq Y_i) \leq \sum_{i=1}^n p_i^2.$$

The last inequality is from part (d).

This is known as Le Cam's Theorem. It provides precise bounds on how far the sum of independent Bernoulli random variables is from a Poisson distribution.