CS70 Homework 11

Michael Pham

Spring 2023

Contents

1   Maybe Lossy Maybe Not

Let us say that Alice would like to send a message to Bob, over some channel. Alice has a message of length 6.

> **Problem 1.1.** Packets are dropped with probability p. If Alice sends 7 packets, what is the probability that Bob can successfully reconstruct Alice's message using polynomial interpolation.

Solution. We observe that in order for Bob to reconstruct a message of length 6, he'll need at least 6 packets; in other words, one of the packets can be dropped at most.

Then, we see that the probability that Bob can reconstruct Alice's message is $\binom{7}{1}p(1-p)^6 + (1-p)^7$. ∎

> **Problem 1.2.** Again, packets can be dropped with probability p. The channel may additionally corrupt 1 packet after deleting packets. Alice realizes this and sends 10 packets for a message of length 6. What is the probability that Bob receives enough packets to successfully reconstruct Alice's message using Berlekamp-Welch?

Solution. We observe that in this scenario, since there can be at most one corruption, then in order to guarantee that Bob can receive enough packets to always reconstruct Alice's message, then he needs at least eight packets (i.e. at most two packets are dropped).

The probability of this happening is:

$$\binom{10}{2}p^2(1-p)^8 + \binom{10}{1}p(1-p)^9 + \binom{10}{0}(1-p)^{10}$$

∎

> **Problem 1.3.** Again, packets can be dropped with probability p. This time, packets may be corrupted with probability q. A packet being dropped is independent of whether or not it is corrupted (i.e. a packet may be both corrupted and dropped). Consider the original scenario where Alice sends 7 packets for a message of length 6. What is probability that Bob can correctly reconstruct Alice's message using polynomial interpolation on all of the points he receives?

Solution. For this, we observe that there are two cases.

First, assuming that none of the packets are corrupted, then Bob can reconstruct the message if at most one packet is dropped. In other words, this scenario has a probability of $(1-q)^7((1-p)^7 + 7p(1-p)^6)$.

In the second scenario, the corrupted packet gets dropped as well. In this case, we see that the probability is $7q(1-q)^6p(1-p)^6$.

Thus, the total probability is $(1-q)^7((1-p)^7 + 7p(1-p)^6) + 7q(1-q)^6p(1-p)^6$. ∎

2   Class Enrollment

Lydia has just started her CalCentral enrollment appointment. She needs to register for a geography class and a history class. There are no waitlists, and she can attempt to enroll once per day in either class or both. The CalCentral enrollment system is strange and picky, so the probability of enrolling successfully in the geography class on each attempt is p_g and the probability of enrolling successfully in the history class on each attempt is p_h. Also, these events are independent.

> **Problem 2.1.** Suppose Lydia begins by attempting to enroll in the geography class everyday and gets enrolled in it on day G. What is the distribution of G?

Solution. We observe that the distribution of G is a geometric distribution, where $\Pr[G = n] = p_g(1-p_g)^{n-1}$, where n is the day on which Lydia gets enrolled. ∎

> **Problem 2.2.** Suppose she is not enrolled in the geography class after attempting each day for the first 7 days. What is $\Pr[G = i \mid G > 7]$, the conditional distribution of G given $G > 7$?

Solution. For this, we observe that $\Pr[G = i \mid G > 7] = \Pr[G = i - 7] = p_g(1 - p_g)^{i-8}$. ∎

> **Problem 2.3.** Once she is enrolled in the geography class, she starts attempting to enroll in the history class from day $G + 1$ and gets enrolled in it on day H. Find the expected number of days it takes Lydia to enroll in both the classes, i.e. $\mathbb{E}[H]$.

Solution. We observe that Lydia is expected to enroll in her geography class after $\frac{1}{p_g}$ days, and after that she is expected to enroll in her history class after an additional $\frac{1}{p_h}$ days. In other words, $\mathbb{E}[H] = \frac{1}{p_g} + \frac{1}{p_h}$. ∎

Suppose instead of attempting one by one, Lydia decides to attempt enrolling in both the classes from day 1. Let G be the number of days it takes to enroll in the geography class, and H be the number of days it takes to enroll in the history class.

> **Problem 2.4.** What is the distribution of G and H now? Are they independent?

Solution. The distribution of G and H are $\Pr[G = n] = p_g(1 - p_g)^{n-1}$ and $\Pr[H = n] = p_h(1 - p_h)^{n-1}$. They are independent. ∎

> **Problem 2.5.** Let A denote the day she gets enrolled in her first class and let B denote the day she gets enrolled in both the classes. What is the distribution of A?

Solution. We observe here that A is essentially when either G or H occurs; from here, we can think of it in the following way: for each "step", we see that the probability of either of the events occurring is $1 - (1 - p_h)(1 - p_g)$. Then, we see that $\Pr[A = n] = ((1 - p_h)(1 - p_g))^{n-1}(1 - (1 - p_h)(1 - p_g))$. ∎

> **Problem 2.6.** What is the expected number of days it takes Lydia to enroll in both classes now, i.e. $\mathbb{E}[B]$?

Solution. We observe that $\Pr[B = n] = \Pr[G < n]\Pr[H = n] + \Pr[H < n]\Pr[G = n] + \Pr[G = n]\Pr[H = n]$.
From here, we can rewrite it as a summation, which is:

$$\Pr[B = n] = \left(\sum_{i=1}^{n-1}(1 - p_g)^{i-1}p_g \right)\left((1 - p_h)^{n-1}p_h\right) + \left(\sum_{i=1}^{n-1}(1 - p_h)^{i-1}p_h \right)\left((1 - p_g)^{n-1}p_g\right) + (1 - p_h)^{n-1}p_h(1 - p_g)^{n-1}p_g$$

Now, we see that $\mathbb{E}[B] = \sum_{n=1}^{\infty} k\Pr[B = n]$. This expression simplifies down to being:

$$\mathbb{E}[B] = \frac{1}{p_h} + \frac{1}{p_g} - \frac{1}{p_h + p_g - p_g p_h}$$

A different way to think of the final answer is that we're finding $\mathbb{E}[G] + \mathbb{E}[H] - \mathbb{E}[H \cup G]$. ∎

> **Problem 2.7.** What is the expected number of classes she will be enrolled in by the end of 30 days?

Solution. We observe that the expected number of classes that she'll be enrolled in at the end of 30 days is:

$$\sum_{i=1}^{30}(1 - p_g)^{i-1}(p_g)(1 - p_h)^{30} + \sum_{i=1}^{30}(1 - p_h)^{i-1}(p_h)(1 - p_g)^{30} + 2\sum_{i=1}^{30}\sum_{j=1}^{30}(1 - p_g)^{i-1}p_g(1 - p_h)^{j-1}p_h$$

∎

3   Swaps and Cycles

We'll say that a permutation $\pi = (\pi(1), \ldots, \pi(n))$ contains a *swap* if there exist $i, j \in \{1, \ldots, n\}$ so that $\pi(i) = j$ and $\pi(j) = i$, where $i \neq j$.

> **Problem 3.1.** What is the expected number of swaps in a random permutation?

Solution. To begin with, let $X_i = 1$ if the ith number is in a swap, and $X_i = 0$ otherwise. Now, we denote S to be the number of switches. We see then that:

$$S = \frac{X_1 + \ldots + X_n}{2}$$

Now, from here, we want to find $\mathbb{E}[S]$. To do this, we observe the following:

$$\mathbb{E}[S] = \mathbb{E}\left[\frac{X_1 + \ldots + X_n}{2}\right]$$
$$= \frac{1}{2}\mathbb{E}[X_1 + \ldots + X_n]$$
$$= \frac{n}{2}\mathbb{E}[X_1]$$

Now, in order to find $\mathbb{E}[X_1]$, we can first find $\Pr[1 \text{ is in a switch with } 2]$, then multiply by $n - 1$. We see then that there are $n!$ ways to order the numbers, and out of those $n!$ ways, $(n-2)!$ of them are where 1 and 2 are swapped. Thus, we have that $\frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$.

Thus, our final value for $\mathbb{E}[S]$ is:

$$\mathbb{E}[S] = \frac{n}{2}(n-1)\left(\frac{1}{n(n-1)}\right)$$
$$= \frac{1}{2}$$

■

> **Problem 3.2.** What about the variance?

Solution. We observe that $\mathrm{Var}[S] = \mathbb{E}[S^2] - \mathbb{E}[S]^2$.

We know that $\mathbb{E}[S]^2 = \frac{1}{4}$. Now, to find $\mathbb{E}[S^2]$, we first see that:

$$S = \frac{1}{2}(X_1 + \ldots + X_n)$$
$$S^2 = \frac{1}{4}(X_1 + \ldots + X_n)^2$$
$$\mathbb{E}[S^2] = \frac{1}{4}\left(\sum_{i=1}^{n}\mathbb{E}[X_i^2] + 2\sum_{i<j}\mathbb{E}[X_iX_j]\right)$$

Now, we know that $\sum_{i=1}^{n}\mathbb{E}[X_i^2] = n\left(\frac{1}{n}\right) = 1$. In order to find $\mathbb{E}[X_iX_j]$, we see that there are two cases:

1. X_i and X_j are in the same swap. In this case, the probability is $\frac{1}{n(n-1)}$ from the last part.

2. X_i and X_j are in different swaps. In this case, we see first we have $\frac{n-2}{n}$ choices, then second we have $\frac{1}{n-1}$ choice. The third pick, we have $\frac{n-3}{n-2}$ choices, and the fourth we have $\frac{1}{n-3}$ choices. Thus, the total probability is $\frac{(n-2)(1)(n-3)(1)}{n(n-1)(n-2)(n-3)} = \frac{1}{n(n-1)}$

Therefore, we see then that the probability that X_i and X_j are both in a swap is equal to $\frac{1}{n(n-1)} + \frac{1}{n(n-1)} = \frac{2}{n(n-1)}$.

With this in mind then, we see the following:

$$\mathbb{E}[S^2] = \frac{1}{4}\left(1 + 2\binom{n}{2}\frac{2}{n(n-1)}\right)$$
$$= \frac{1}{4}(1+2)$$
$$= \frac{3}{4}$$

And with that, we have that $\mathrm{Var}[S] = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$. ∎

> **Problem 3.3.** In the same spirit as above, we'll say that π contains a k-*cycle* if there exist $i_1,\ldots,i_k \in \{1,\ldots,n\}$ with $\pi(i_1) = i_2, \pi(i_2) = i_3, \ldots, \pi(i_k) = i_1$. Compute the expectation of the number of k-cycles.

Solution. In order to find the expectation of the number of k-cycles, $\mathbb{E}[S_k]$, we observe that we have $\binom{n}{k}$ tuples to form a k-cycle.

From here, we see that the probability of forming a k-cycle is:

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

We then multiply this by $(n-k)!$ because the remaining numbers can be in any order. Putting this all together, we have then that $\mathbb{E}[S_k] = \binom{n}{k}\frac{(n-k)!(k-1)!}{n!} = \frac{1}{k}$. ∎

4   Throwing Frisbees

Shahzar and his $n-1$ friends stand in a circle and play the following game: Shahzar throws a frisbee to one of the other people in the circle randomly, with each person being equally likely, and thereafter, the person holding the frisbee throws it to someone else in the circle, again uniformly at random. The game ends when someone throws the frisbee back to Shahzar.

Problem 4.1. What is the expected number of times the frisbee is thrown through the course of the game?

Solution. We observe that after Shahzar has thrown the frisbee, we can think of the scenario as being like a geometric distribution, where it ends after reaching Shahzar again; this has a probability of $\frac{1}{n-1}$. Then, we see that the expected value $\mathbb{E}[S] = 1 + \frac{1}{\frac{1}{n-1}} = 1 + n - 1 = n$ ∎

Problem 4.2. What is the expected number of people that never get the frisbee during the game?

Solution. Let X_i denote that the ith person did not receive a frisbee. We see then that the total number of people who didn't receive it is $X = X_1 + \ldots + X_{n-1}$.

Now, from here, we observe the following:

$$\mathbb{E}[X] = \mathbb{E}[X_1 + \ldots + X_{n-1}]$$
$$= \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_{n-1}]$$

From here, we see that $\mathbb{E}[X_i] = \left(1 - \frac{1}{n-1}\right)^n$. Then, we see that:

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_{n-1}]$$
$$= (n-1)\mathbb{E}[X_1]$$
$$= (n-1)\left(1 - \frac{1}{n-1}\right)^n$$

∎

5   Balls and Bins

> **Problem 5.1.** Throw n balls into m bins, where m and n are positive integers. Let X be the number of bins with exactly one ball. Compute $\mathrm{Var}[X]$. Your final answer should not contain any summations.

Solution. First, we will want to find $\mathbb{E}[X]$. To do this, we can let $X_i = 1$ if the ith bin has one ball in it, and $X_i = 0$ otherwise.

Now, let $X = X_1 + \ldots + X_m$, where X is the number of bins that have exactly one ball. Then, in order to find $\mathbb{E}[X]$, we do the following:

$$\mathbb{E}[X] = \mathbb{E}[X_1 + \ldots + X_m]$$
$$= m\mathbb{E}[X_1]$$

Now, to find $\mathbb{E}[X_1]$, we observe that we want one ball to go into the bin, and the rest go to other bins. The probability of this happening is $n\left(\frac{1}{m}\right)\left(\frac{m-1}{m}\right)^{n-1}$.

Therefore, we have that $\mathbb{E}[X] = m(n)\left(\frac{1}{m}\right)\left(\frac{m-1}{m}\right)^{n-1} = n\left(\frac{m-1}{m}\right)^{n-1}$.

Now, to find $\mathrm{Var}[X]$, we observe that $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

To find $\mathbb{E}[X]^2$, we observe that this is simply $m(n)\left(\frac{1}{m}\right)\left(\frac{m-1}{m}\right)^{n-1} = n\left(\frac{m-1}{m}\right)^{n-1}$

To find $\mathbb{E}[X^2]$, we first see that:

$$X = X_1 + \ldots + X_m$$
$$X^2 = (X_1 + \ldots + X_m)^2$$
$$\mathbb{E}[X^2] = \sum_{i=1}^{m} \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_i X_j]$$

In order to find $\mathbb{E}[X_i X_j]$, we want to find the probability that both X_i and X_j have one ball in them. We see that the probability of this happening is $\binom{n}{2}\left(\frac{2}{m}\right)\left(\frac{1}{m}\right)\left(\frac{m-2}{m}\right)^{n-2}$.

Then, from here, we see that:

$$2\sum_{i<j} \mathbb{E}[X_i X_j] = 2\binom{m}{2}\binom{n}{2}\left(\frac{2}{m}\right)\left(\frac{1}{m}\right)\left(\frac{m-2}{m}\right)^{n-2}$$
$$= 4\binom{m}{2}\binom{n}{2}\left(\frac{1}{m}\right)^2\left(\frac{m-2}{m}\right)^{n-2}$$

Putting it all together, we get the following:

$$\mathrm{Var}[X] = n\left(\frac{m-1}{m}\right)^{n-1} + 4\binom{m}{2}\binom{n}{2}\left(\frac{1}{m}\right)^2\left(\frac{m-2}{m}\right)^{n-2} - n^2\left(\frac{m-1}{m}\right)^{2n-2}$$

■

6   Will I Get My Package?

A delivery guy in some company is out delivering n packages to n customers, where n is a natural number greater than 1. Not only does he hand each customer a package uniformly at random from the remaining packages, he opens the package before delivering it with probability $1/2$. Let X be the number of customers who receive their own packages unopened.

> **Problem 6.1.** Compute the expectation $\mathbb{E}[X]$.

Solution. Let $X_i = 1$ if the ith person receives their own package unopened, and 0 otherwise. From here, we see that $X = X_1 + \ldots X_n$.

Then, we observe that:

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

From here, we observe that the probability a person receives their own package is $\frac{1}{n}$, and the probability that it's unopened is $\frac{1}{2}$. Then, with this in mind, we see that:

$$\mathbb{E}[X] = n\left(\frac{1}{2n}\right)$$
$$= \frac{1}{2}$$

∎

> **Problem 6.2.** Compute the variance $\mathrm{Var}[X]$.

Solution. Now, we want to compute the variance. Again, recall that $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Now, we want to find $\mathbb{E}[X^2]$. Observe first that:

$$X = X_1 + \ldots + X_n$$
$$X^2 = (X_1 + \ldots + X_n)^2$$
$$\mathbb{E}[X^2] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_i X_j]$$

We first observe that $\sum_{i=1}^{n} \mathbb{E}[X_i^2] = \frac{1}{2}$. Now, in order to find $\mathrm{Pr}[X_i = 1 \cap X_j = 1]$, we find the probability that both i and j receive their own packages and they're both unopened:

$$\mathrm{Pr}[X_i = 1 \cap X_j = 1] = \frac{1}{2} \cdot \frac{1}{n} \cdot \frac{1}{2} \cdot \frac{1}{n-1}$$
$$= \frac{1}{4n(n-1)}$$

Therefore, we see then that:

$$2\sum_{i<j} \mathbb{E}[X_i X_j] = 2\binom{n}{2}\left(\frac{1}{4n(n-1)}\right)$$
$$= \frac{1}{4}$$

With this, we observe that $\mathbb{E}[X^2] = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$.

Putting it all together, we have the following:

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{3}{4} - \frac{1}{4} \\
&= \frac{1}{2}
\end{aligned}$$

■