EECS 376: Foundations of Computer Science
University of Michigan, Spring 2024

Discussion Worksheet 10

1 Linearity of Expectation

1. **From Winter 2018 Final Exam**
   In an undirected graph, a *3-painting* is an assignment of one of three colors to each vertex. (Adjacent vertices do *not necessarily* need to have different colors.) Given a 3-painting of an undirected graph, an edge is called *colorful* if its endpoints are assigned different colors.

   Describe and analyze an efficient randomized algorithm that, given an undirected graph as input, outputs a 3-painting that *in expectation* has at least $\frac{2}{3}$ of the maximum possible number of colorful edges, i.e., this gives a 2/3 approximation in expectation.

Solution: Let $G = (V, E)$ be a graph with $|V| = n$. Let c_1, c_2, c_3 be our colors. We will randomly color each vertex of the graph with one of these colors. We claim that in expectation, two thirds of the edges of G will be colorful. For each $e \in E$, let X_e be a random indicator variable that evaluates to 1 if edge e is colorful, and 0 otherwise. Let X be a random variable that evaluates to the number of colorful edges in G, so $X = \sum\limits_{e \in E} X_e$. We claim that $\mathbb{E}[X] = \frac{2}{3}|E|$. For some $e \in E$, note that there are 9 possible paintings of the two endpoints of e. We see that 3 of these have the same color on both endpoints. Thus $P[X_e = 1] = \frac{6}{9} = \frac{2}{3}$, and therefore $E[X_e] = 1 \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} = \frac{2}{3}$. Thus by the linearity of expectation, we have

$$\mathbb{E}[X] = \sum_{e \in E} \mathbb{E}[X_e] = |E| \cdot \frac{2}{3}.$$

Let C be the maximum number of colorful edges in G. Note that $|E| \geq C$, so $\mathbb{E}[X] = \frac{2}{3}|E| \geq \frac{2}{3}C$, so we have our result.

2. Let A be an array of length n of a random permutation of n distinct integers. Compute the expected number of increasing subarrays (i.e., contiguous subsequences) in A of length k.

Solution: To find the expected number of increasing subarrays of length k, we can consider whether a particular subarrays of length k is increasing and then use Linearity of Expectations to find the expected number of increasing subarrays in the entire array.

Formally, let us define the indicator random variable X_i to be:

$$X_i = \begin{cases} 1 & A \text{ starting at position } i \text{ until position } i + k - 1 \text{ is increasing} \\ 0 & \text{otherwise.} \end{cases}$$

Since the last position at which a length-k subarray can start is at position $n - k + 1$, then it makes sense to either set $X = X_1 + X_2 + \cdots + X_{n-k+1}$, or alternatively to set $X = X_1 + X_2 + \cdots + X_n$, but define $X_{n-k+1}, X_{n-k+2}, \ldots X_n = 0$. For the sake of this solution, we will use the first of these two alternatives. Additionally, for any given i, $\mathbf{E}[X_i] = \frac{1}{k!}$ because there are $k!$ permutations over k distinct elements, and only 1 of those permutations is increasing.

$$\begin{aligned}
\mathbf{E}[X] &= \mathbf{E}[X_1 + X_2 + \cdots + X_{n-k+1}] && && (1)\\
&= \sum_{i=1}^{n-k+1} \mathbf{E}[X_i] && \text{By Linearity of Expectations} && (2)\\
&= \sum_{i=1}^{n-k+1} \frac{1}{k!} && \text{Only permutation is increasing} && (3)\\
&= \frac{n-k+1}{k!} && && (4)
\end{aligned}$$

Therefore, the expected number of increasing subarrays in A of length k is $\frac{n-k+1}{k!}$.

3. A group of n students, all of whom have distinct heights, line up in a single-file line uniformly at random to get a group picture taken. If a student has any students in front of them who is taller than them, then they will not be seen in the picture. For this reason, every student files one complaint to the photographer for each taller student who is in front of them since each one of these students would individually block the original student from being seen. Compute the expected number of complaints that the photographer will receive.

Solution: Let X_{ij} be an indicator random variable where:

$$X_{ij} = \begin{cases} 1 & \text{if } i < j \text{ and the student in position } i \text{ is taller than student in position } j \\ 0 & \text{otherwise.} \end{cases}$$

Note that for convention, we assume position 1 to be the front of the line and position n to be the end of the line. The number of complaints that will be filed is exactly equal to the number of X_{ij}s that equal 1.

For any $i < j$, it is equally likely that the student in position i is taller than the student in position j as it is that the student in position i is shorter than the student in position j. One explanation for this is that the students lined up uniformly at random. Another explanation is that for any ordering of students for which the student in position i is taller than the student in position j, there is exactly one other ordering for which the student in position i is shorter than the student in position j and all other students are in the exact same position (just swap these two students). Now we compute:

$$\begin{aligned}\mathbf{E}\left[\sum_{1\le i<j\le n} X_{ij}\right] &= \sum_{1\le i<j\le n} \mathbf{E}[X_{ij}] \\ &= \sum_{1\le i<j\le n} \frac{1}{2} \\ &= \binom{n}{2}\cdot\frac{1}{2} \\ &= \frac{n\cdot(n-1)}{4}\end{aligned}$$

4. Balls-and-bins. Let there be k balls and n bins. Every ball is independently put into one bin uniformly at random.

(a) What is the probability that there exists a bin that contains at least two balls? (This problem has a name: birthday problem)

Solution: We solve this by finding the complement: the probability that every bin has at most one ball. Imagine putting the balls in one by one. Every ball has to avoid all the non-empty bins. Thus the probability is

$$\frac{n}{n}\times\frac{n-1}{n}\times\frac{n-2}{n}\times\ldots\times\frac{n-k+1}{n}=\frac{n!}{(n-k)!n^k}.$$

Then the probability that there exists a bin with at least two balls is

$$1-\frac{n!}{(n-k)!n^k}.$$

(b) What is the probability of having w balls in the first bin, for $w=0,1,\ldots,k$?

Solution: For some fixed w balls of the k, there is a $\left(\frac{1}{n}\right)^w$ probability of those w balls landing in the first bin and a $\left(\frac{n-1}{n}\right)^{k-w}$ probability of the remaining balls *not* landing in the first bin. There are $\binom{k}{w}$ such choices of w balls. The overall probability of having w balls in the first bin is

$$\binom{k}{w}\left(\frac{1}{n}\right)^w\left(\frac{n-1}{n}\right)^{k-w}.$$

Note: this process is "binomially distributed".

(c) How many non-empty bins are there in expectation?

Solution: Let X_j to be the indicator that the jth bin is non-empty. Then

$$\Pr(X_j = 1) = 1 - \left(\frac{n-1}{n}\right)^k.$$

The number of non-empty bins is equal to $\sum_{j=1}^{n} X_j$ whose expectation is

$$\begin{aligned}\mathsf{E}\sum_{j=1}^{n} X_j &= \sum_{j=1}^{n} \Pr(X_j = 1)\\ &= \sum_{j=1}^{n}\left(1 - \left(\frac{n-1}{n}\right)^k\right)\\ &= n\left(1 - \left(\frac{n-1}{n}\right)^k\right)\end{aligned}$$

2 Markov's Inequality

5. Suppose we flip a fair coin 376 times and want to bound the probability that at least 203 heads are flipped. Let X be the random variable representing the number of heads flipped. Which of the following is the best possible bound we can obtain using Markov's inequality?

   A. $\Pr[X \geq 203] \leq \frac{203}{188}$
   B. $\Pr[X \geq 203] \geq \frac{203}{188}$
   C. $\Pr[X \geq 203] \leq 1$
   D. $\Pr[X \geq 203] \leq \frac{188}{203}$
   E. $\Pr[X \geq 203] \geq \frac{188}{203}$

Solution: Markov's inequality says that $\Pr[Z \geq a] \leq \frac{\mathbf{E}[Z]}{a}$, for any non-negative random variable Z and any $a > 0$. By definition, X is non-negative and $203 > 0$. So, by Markov's inequality, the probability of there being at least 203 heads is less than or equal to the expected number of heads divided by 203. Since the coin is fair, $\mathbf{E}[X] = 376/2 = 188$. Therefore, $\Pr[X \geq 203] \leq \frac{188}{203}$.

6. Consider a class with 376 students, and let Y be the random variable representing the number of students that get a B or better. Suppose that $\mathbf{E}[Y] = 280$. Which of the following is a correct bound for the probability that strictly more than 203 students get a grade of B or better?

   A. $\Pr[Y > 203] \leq \frac{203}{376}$
   B. $\Pr[Y > 203] \geq \frac{280-203}{376-203}$
   C. $\Pr[Y > 203] \leq \frac{203}{280}$
   D. $\Pr[Y > 203] \geq \frac{280-203}{376-280}$

E. $\Pr[Y > 203] \leq \frac{280}{376}$

Solution: The "reverse" Markov inequality holds that $\Pr[Z > a] \geq \frac{\mathbf{E}[Z]-a}{B-a}$, where Z is a random variable that is at most B and $a < B$. Y is a random variable that is at most 376 (since in a class of 376 students, at most 376 can get a B or better), and $203 > 0$. Furthermore, it is given that $\mathbf{E}[Y] = 280$. Then, by Reverse Markov's inequality, $\Pr[Y > 203] \geq \frac{280-203}{376-203}$.

3 Pre-Crypto Review

7. Apply the Extended Euclidean Algorithm to compute the modular inverse of 12 modulo 31.

Solution: Finding 12 modulo 31 is equivalent to finding an integer x such that $12x \equiv 1 \pmod{31}$. This can be accomplished using the Extended Euclidean Algorithm through either a non-tabular or a tabular approach (g, a, b, x, y table). Both approaches will result in the same answer and are both acceptable. Each of the two approaches is performing the same computations, the only difference being the format of those computations. Since an example of the tabular method was discussed already during discussion, this solution will only present the non-tabular solution.

In step 1 of the non-tabular method, the Euclidean Algorithm is run on inputs 12 and 31. In step 2, the computations from step 1 are reversed for the purpose of re-expressing 1 in terms of 12 and 31. In step 3, modular arithmetic is applied to the result from step 2 to compute the modular inverse of 12 modulo 31.

Step 1: Euclidean Algorithm

$$\begin{aligned}
31 &= 2 \cdot 12 + 7 \\
12 &= 1 \cdot 7 + 5 \\
7 &= 1 \cdot 5 + 2 \\
5 &= 2 \cdot 2 + 1
\end{aligned}$$

Step 2: Re-expressing 1 in terms of 12 and 31

$$\begin{aligned}
1 &= 5 - 2 \cdot 2 \\
&= 5 - 2 \cdot (7 - 1 \cdot 5) = 3 \cdot 5 - 2 \cdot 7 \\
&= 3 \cdot (12 - 1 \cdot 7) - 2 \cdot 7 = 3 \cdot 12 - 5 \cdot 7 \\
&= 3 \cdot 12 - 5 \cdot (31 - 2 \cdot 12) = 13 \cdot 12 - 5 \cdot 31
\end{aligned}$$

Step 3: Computing the modular inverse

$$\begin{aligned} 1 &= 13 \cdot 12 - 5 \cdot 31 \\ 1 &\equiv 13 \cdot 12 - 5 \cdot 31 \pmod{31} \\ &\equiv 13 \cdot 12 \pmod{31} \end{aligned}$$

Therefore, the inverse of 12 modulo 31 is 13. Of note, had the inverse been negative such as -18, while it is not incorrect to say that the inverse is -18, typically modular inverses modulo n are standardized to be within the set $\{0, 1, \dots, n-1\}$. In this way, if the inverse was found to be -18, then one should report $-18 + 31 = 13$ as the inverse. The reason why -18 being an inverse of 12 implies $-18 + 31$ is also an inverse of 12 is because $-18 \cdot 12 \equiv 1 \pmod{31} \rightarrow -18 \cdot 12 + 31 \cdot 12 \equiv 1 \pmod{31}$ since adding multiples of 31 to either side does not change the equivalence.