

Lecture 2: April 3, 2013

Madhur Tulsiani

Scribe: Shubendu Trivedi

1 Coin tosses continued

We return to the coin tossing example from the last lecture again:

Example 1.1 *Given, that $\mathbb{P}[\text{heads}] = p$, what is $\mathbb{E}[Z]$, where $Z = \#\text{tosses till the first heads?}$*

We saw that $\mathbb{E}[Z] = \frac{1}{p}$. This of course holds when all the coin tosses are independent. Such a Z is called a Geometric Random Variable (In general: Suppose we have probability p of success in one try, and if we can make independent attempts many times over, then the geometric random variable counts the number of attempts that would be needed to obtain the first success in expectation).

A side remark: In the last lecture we used that for x satisfying $|x| \leq 1$, $\sum_{i=1}^{\infty} i \cdot x^{i-1} = \frac{1}{(1-x)^2}$. This can be derived by differentiating both sides of the equality $\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$, which can be derived by defining the partial sums S_n of the series on the left, as follows:

$$\begin{aligned} S_n &= 1 + x + x^2 + \dots + x^{n-1} \\ xS_n &= x + x^2 + \dots + x^n \\ \hline (1-x)S_n &= 1 - x^n \\ \therefore S_n &= \frac{1 - x^n}{1 - x} \end{aligned}$$

Hence, we have,

$$\lim_{n \rightarrow \infty} S_n = \frac{1}{1-x} \text{ if } |x| \leq 1$$

Another thing that we swept under the proverbial rug in the last lecture: What is the basic event in the case of this example? How can we define a probability if the set is of potentially infinite size? We need to make sure that the problem is well posed by defining a valid probability space.

As mentioned earlier, here if we try to assign a probability to each possible outcome (which is an infinite sequence of coin tosses), it will simply be 0. However, we can consider a collection of events which is closed under union, intersection and complementation, to which we will assign probabilities. For all $i \in \mathbb{N}$, and all sequences of i bits denoted by $b \in \{0, 1\}^i$, we can define the event

$$E_{i,b} \equiv \text{first } i \text{ bits are according to } b.$$

Thus, $\mathbb{P}[E_{i,b}] = p^{\#\text{ of 1s in } b} (1-p)^{\#\text{ of 0s in } b}$. We will use the collection \mathcal{E} generated by unions, complements and intersections of all such events, to which we can easily assign probabilities.

We now return to the case with a finite number of coin tosses.

Example 1.2 Consider n independent tosses of a coin which comes up heads with probability p . Define the following random variable:

$$Y = \begin{cases} 1 & \text{if \# of heads is odd} \\ 0 & \text{if \# of heads is even} \end{cases}$$

We want to compute $\mathbb{E}[Y]$.

Note that unlike in the previous lecture, here we have no linearity to exploit and add up. So what do we do in such a case? Also, note that $\mathbb{E}[Y] = \frac{1}{2}$ if $\mathbb{P}[\text{heads}] = \frac{1}{2}$ (by symmetry). However, the case $p \neq \frac{1}{2}$ is more interesting.

Since we have n tosses, let us call this $\mathbb{E}[Y_n]$. We have,

$$\mathbb{E}[Y_n] = p \mathbb{E}[Y_n|H] + (1-p) \mathbb{E}[Y_n|T].$$

In this setting, one “trick” that we utilize is that instead of a variable that takes values 0 and 1, let’s make the variable take values 1 and -1 . We take

$$\tilde{Y} = \begin{cases} -1 & \text{if \# of heads is odd} \\ +1 & \text{if \# of heads is even} \end{cases}$$

The following claim $\tilde{Y} = 1 - 2Y$ is easy to check.

Now let us define another random variable.

$$X_i = \begin{cases} -1 & \text{if } i\text{th coin is heads} \\ +1 & \text{if } i\text{th coin is tails} \end{cases}$$

We have: $\tilde{Y} = X_1 X_2 X_3 \dots X_n$. Therefore $\mathbb{E}[\tilde{Y}] = \mathbb{E}[X_1 X_2 \dots X_n]$.

Now, if we have independence, then we’d have:

$$\mathbb{E}[\tilde{Y}] = \mathbb{E}[X_1] \cdot \mathbb{E}[X_2] \cdots \mathbb{E}[X_n]$$

Also, $\mathbb{E}[X_i] = 1 - 2p \therefore \mathbb{E}[\tilde{Y}] = (1 - 2p)^n$. Since we need $\mathbb{E}[Y]$, we have:

$$\mathbb{E}[Y] = \frac{1 - (1 - 2p)^n}{2}$$

Basically, the “trick” of using $+1$ and -1 as values for a random variable in place of 0 and 1, was able to capture the parity of the event. Also, note that here we *did* need to use that all n coin tosses are independent.

2 Toy Problem: Coupon Collection

The model is the following: There are n kinds of items/coupons and at each time step T , we get one random coupon (independently of the others) out of the n total. We define a random variable, T which is the time when we first have all the n types of coupons. We want to find $\mathbb{E}[T]$.

We can make the following claim:

$$T = \sum_{i=1}^n X_i$$

Where, X_i is the time to get from the $i - 1$ to the i type of coupon. Thus we have,

$$\mathbb{E}[T] = \sum_i \mathbb{E}[X_i]$$

Clearly, we have $\mathbb{E}[X_1] = 1, \mathbb{E}[X_2] = \frac{n}{n-1}, \mathbb{E}[X_3] = \frac{n}{n-2}$ and so on.

This is because, X_i represents a Geometric Random Variable with success probability $\frac{n}{n-i+1}$ of getting a new coupon. Thus, at the i th step:

$$\mathbb{E}[X_i] = \frac{n}{n-i+1}$$

$$\therefore \mathbb{E}[T] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \dots + n = n \cdot H(n)$$

where $H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$ is the n^{th} harmonic number. It is known (see Wikipedia for example) that $H_n = n \ln n + \Theta(1)$. Thus, we have that $\mathbb{E}[T] = n \ln n + \Theta(n)$.

3 A Simple Randomized Algorithm for Max-Cut

Using the ideas discussed so far, in this section we design an algorithm for Max-Cut.

The problem is as follows: Given a graph $G = (V, E)$, we want to divide the vertex set V into two set S and \bar{S} such that the number of edges between S and \bar{S} is as large as possible.

For each $i \in V$, assign i into S with probability $\frac{1}{2}$ and into \bar{S} with probability $\frac{1}{2}$. Let us define a random variable Z , such that $Z = \#$ of edges cut. Clearly, $Z = \sum_e X_e$, where:

$$X_e = \begin{cases} 1 & \text{if } e \text{ is cut} \\ 0 & \text{otherwise} \end{cases}$$

Since we are picking each vertex with probability $\frac{1}{2}$ to be in either S or \bar{S} , for any edge in E , the probability that it is cut is $\frac{1}{2}$. Thus if the number of edges $|E| = m$, then by linearity of expectation the expected number of edges cuted, $\mathbb{E}[Z] = \frac{m}{2}$.

Now suppose we also want to find the sets S and \bar{S} *deterministically*, while ensuring that the number of edges between and S and \bar{S} is at least $m/2$. We will proceed as follows: We line up the vertices v_1, v_2, \dots, v_n and then *deterministically* decide if each v_i should be in S or \bar{S} , while maintaining

$$\mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i] \geq \frac{m}{2}.$$

That is, suppose v_1, v_2, \dots, v_i are placed into S or \bar{S} deterministically, we have to analyze the expectation of the cut-size from now on.

$$\begin{aligned} & \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i] \\ &= \frac{1}{2} \cdot \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i \ \& \ v_{i+1} \in S] + \frac{1}{2} \cdot \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i \ \& \ v_{i+1} \in \bar{S}]. \end{aligned}$$

It is easy to see that at least with one of the two choices of keeping v_{i+1} in S or \overline{S} , the expected cut size does not decrease. Therefore at least one of the following is true:

$$\begin{aligned} \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i \text{ \& } v_{i+1} \in S] &\geq \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i] \\ \text{or} \\ \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i \text{ \& } v_{i+1} \in \overline{S}] &\geq \mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i] \end{aligned}$$

Thus, our strategy is to look at the expected size of the cut in both the cases and put v_{i+1} in the set that leads to a greater expected cut-size. Note that this decision is completely deterministic.

The only question is how to compute the quantity $\mathbb{E}[Z \mid \text{decision for } v_1, \dots, v_i]$ efficiently. This is easy since given the decision for v_1, \dots, v_i , the edges for which both vertices are already on the same side contribute 0 to the expectation and the edges for which both vertices are on different sides contribute 1. The remaining edges contribute 1/2 to the expectation. Thus, we can given the decision for v_1, \dots, v_i , we can easily compute the contribution of each edge.

This leads to an algorithm which *de-randomizes* the randomized algorithm by making the sequentially in a way that ensures that the *conditional expectation* given our choices so far is large. This is sometimes also referred to as the *Method of Conditional Expectations*.

4 The Probabilistic Method: Independent Sets

Now we do one more application of expectations which is often called the *Probabilistic Method*. It is often used to show the existence of objects with certain properties without necessarily constructing them. In the previous section we used probabilistic reasoning to show that a cut exists, but then later also showed how to find such a cut.

Again, consider a graph $G = (V, E)$. Now, we want to define an independent set $S \subseteq V$, such that no edge lies completely within the set S . That is, $\forall e = (v_i, v_j)$, either $v_i \notin S$ or $v_j \notin S$.

We are interested in finding a large independent set. Let's say that $\text{deg}(v_i)$ is the number of edges containing v_i . The following result is due to Caro and Wei.

Theorem 4.1 *Let $G = (V, E)$ be a graph with n vertices. Then there exists an independent set S such that*

$$|S| \geq \sum_{i=1}^n \frac{1}{\text{deg}(v_i) + 1} \geq \frac{n}{\max_i(\text{deg}(v_i)) + 1}.$$

The main trick in such kind of problems is to set up the right kind of probabilistic experiment, the analysis is usually quite easy.

In this question, we can't do everything independently unlike in some previous questions. Suppose that we do - and hence pursue the following idea: Put each v_i in S with probability p . We can't guarantee that we would not pick up both the endpoints of an edge to keep in S . However, this idea can also be made to work and we will come back to it in a bit. We first prove the theorem using a different idea.

Proof: Pick a random ordering of the vertices v_1, v_2, \dots, v_n and we pick v_i if it appears before all its neighbors in the ordering. This is clearly an independent set since for any edge (v_i, v_j) , the vertex which appears later in the ordering will certainly not be picked. The next question is to analyze the size of this independent set. Or we want to look at $\mathbb{E}[|S|]$. We have $|S| = \sum_i X_i$, where

$$X_i = \begin{cases} 1 & \text{if } v_i \in S \\ 0 & \text{otherwise} \end{cases}$$

We would want to calculate $\mathbb{E}[X_i]$. To do it, if we choose a random order, we've shrunk our probability space to the neighborhood of v_i alone. If we have a random order, what is the probability that i appears first?

$$\mathbb{E}[X_i] = \frac{1}{\deg(v_i) + 1}$$

This immediately gives that

$$\mathbb{E}[|S|] = \sum_{i=1}^n \frac{1}{\deg(v_i) + 1},$$

and hence there must exist an independent set S with the above size. ■

We can now salvage our “wrong” idea discussed earlier. The problem here is: No matter what p is, we might end up picking up an edge. What one can do is to throw away such edges and let us say we get a set S' from S by throwing away edges. Now we want to know what is the expected size of S' . This method is called the Method of Alterations (since we *alter* and object to make it satisfy the desired properties). We have

$$\mathbb{E}[S'] = \mathbb{E}[|S| - 2 \cdot \text{number of edges deleted}].$$

We can use linearity to compute the above expectation. We have that $\mathbb{E}[|S|] = p \cdot n$. Also, each edge is deleted if and only if both its vertices are in S . This happens with probability p^2 . Thus,

$$\mathbb{E}[S'] = p \cdot n - 2 \cdot p^2 \cdot |E|.$$

Suppose the maximum degree is d , then the number of edges $|E| \leq \frac{dn}{2}$, thus we have:

$$\mathbb{E}[S'] \geq p \cdot n - p^2 \cdot nd$$

Now suppose we choose $p = \frac{1}{2d}$, then

$$\mathbb{E}[S'] \geq \frac{n}{2d} - \frac{n}{4d} = \frac{n}{4d}$$

And thus, we still get that there exists an independent set with size at least $\frac{n}{4d}$.