CS 252, Lecture 10: The Probabilistic Method

1 Introduction

Consider the following puzzle: Suppose that 12% of earth's surface is land, and the rest is water[1]. Irrespective of how the land and water is distributed on earth, it is always possible to inscribe a cube in earth such that all the vertices of the cube are on water!

We can prove this as follows: inscribe a cube inside earth uniformly at random i.e. rotate the cube randomly. Now, the probability that a given vertex of the cube lies on land is exactly equal to 0.12. Let the random variable X denote the number of points of the cube that lie on land. Let X_i, $1 \leq i \leq 8$ denote the indicator variable equal to 1 if the ith vertex of the cube lies on land, and equal to 0 otherwise. We have $\mathbb{E}[X_i] = \Pr(X_i = 1) = 0.12$. By linearity of expectation, we have $\mathbb{E}[X] = \sum_{i=1}^{8} \mathbb{E}[X_i] = 0.96 < 1$. That is, the average number of vertices of the cube on land is less than 1. Since X is an integer random variable, it implies that there exists a random configuration in which the number of vertices of the cube on land is equal to 0. Intuitively, this is true because not every one can be above average. More formally, we can use the Markov's inequality:

Lemma 1. *(Markov's Inequality) Let X be a non-negative random variable. For every $a > 0$, we have*

$$Pr(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Thus, $\Pr(X > 0.97) < 1$, or in other words, $\Pr(X = 0) > 0$. Hence, there exists a rotation of the cube such that all the vertices of the cube are on water.

We can actually directly also prove that $\Pr(X = 0) > 0$ as follows:

$$\begin{aligned}
\Pr(X = 0) &= 1 - \Pr(X \geq 1) \\
&= 1 - \Pr(\exists i : X_i \geq 1) \\
&\geq 1 - \sum_i \Pr(X_i \geq 1) = 1 - 0.96 > 0
\end{aligned}$$

where in the last step, we have used the union bound.

Essentially, what we have done is the following: when we are asked to show the existence of a cube with all the vertices on water, instead of directly showing it, we have constructed a probability space over

[1] Even though this is incorrect now, it could have been the case at some point in the distant past.

the possible cubes and have shown that there is a non-zero probability that a cube satisfies our required property. It may seem that we are doing unnecessary work here: why compute the probability of a cube having all vertices on water? why not directly show the existence of the cube with all the vertices on water, given a land water scenario? The issue is that showing the cube exists "constructively" is very difficult. Instead, we come up with a nice probability space where we can easily compute the probability of our required configuration, and then show that it's non-zero, and we are done!

This way of showing an existence of an object by constructing a clever probability space is known as the *probabilistic method*. It is one of the most powerful tools in Combinatorics and Theoretical Computer Science. It is also very general, applicable in many varied scenarios. In this lecture, we discuss three applications of it: first, in Ramsey theory, where the method first originated, and then second, MAX-CUT, a well studied problem in Computer Science, and then finally an application in Extremal Combinatorics.

2 Ramsey Theory

We start with a new puzzle: In any group of six people, there are either three mutual acquaintances, or three mutual non-acquaintances. Or in graph theoretic terms, in any graph on 6 vertices, there is either a clique of size 3 or an independent set of size 3. An alternate view is in terms of two coloring complete graphs: suppose that we color each edge of K_6 with either red or blue, then there is a monochromatic triangle.

We can prove it as follows: consider vertex 1, and look at the edges adjacent to 1. Since there are five edges, at least three of them should have the same color. Without loss of generality, let $(1,2),(1,3),(1,4)$ be colored red. Now, consider the triangle $(2,3,4)$. If any of the edges in this triangle is colored red, the corresponding edge together with 1 forms a red triangle. If all the edges of the triangle are colored blue, we of course have a monochromatic triangle.

Note that the above is tight: the graph C_5, the cycle on five vertices does not have any clique or independent set of size 3. We can generalize this to arbitrary k as follows:

Definition 2. (Ramsey Number) For an integer $k \geq 3$, the Ramsey number $R(k)$ is defined as the smallest integer n such that any graph on n vertices contains either a clique or an independent set of size k.

We have proved above that $R(3) = 6$. $R(4)$ is proved to be equal to 19. We can also show that $R(k) \leq \binom{2k-1}{k-1}$ by an inductive argument. Note that asymptotially, it is equivalent to $R(k) \leq O\left(\frac{4^k}{\sqrt{k}}\right)^2$.

What about lower bounds? In order to show $R(k) > m$, we need to show that there exists a graph on m vertices without a k sized clique or independent set. One construction that we can come up with is the following: consider a disjoint set of $k-1$ blocks $V_1, V_2, \ldots, V_{k-1}$ each containing $k-1$ vertices. For every pair of vertices u, v, they are adjacent if and only if they are in different blocks. There is no clique or independent set of size k in this graph. Thus, $R(k) \geq \Omega(k^2)$. For some time, it was conjectured that this is optimal. Then, Erdős gave a striking proof of $R(k)$ being exponential in k, introducing the probabilistic method.

[2]Note that $\binom{n}{\frac{n}{2}} \approx \frac{2^n}{\sqrt{n}}$.

Theorem 3. *(Erdős, 1947)* $R(k) \geq \Omega(2^{\frac{k}{2}}k)$.

Proof. We will use the probabilistic method. The first goal is to construct a probability space on graphs on n vertices. The most natural way to do this is the following, known as Erdős-Renyi random graph $G(n, \frac{1}{2})$: for each pair (i, j) such that $1 \leq i < j \leq n$, independently add an edge between (i, j) with probability equal to $\frac{1}{2}$. Let $G = (V, E)$ denote the random graph sampled according to this distribution. Let $S \subseteq V$ be a subset of vertices of size k. The probability that S is a clique or an independent set is equal to $2^{1-\binom{k}{2}}$. Thus, by union bound, the probability that there exists a clique or an independent set of size k in G is at most

$$\binom{n}{k} 2^{1-\binom{k}{2}} \leq 2\left(\frac{ne}{k}\right)^k 2^{-\binom{k}{2}} \leq \frac{1}{3}$$

where we set n to be $k2^{\frac{k}{2}}\frac{1}{10e}$. Hence, the probability that the graph does not have a k-clique or k-independent set is non-zero. Thus, there exists a graph with $\Omega(2^{\frac{k}{2}}k)$ vertices that does not have any clique or independent set of size k. □

A couple remarks: The idea of probabilistic method is also used (implicitly) by another great mind Claude Shannon around the same time in 1948 to prove the existence of good error correcting codes. The study of random graphs that we used, picking each edge independently, is a rich field on its own, with many applications.

The above construction shows that there exists a graph on n vertices that does not have a clique or independent size of size roughly $2\log n$. One might wonder if there is a way to construct such a graph efficiently. Note that we can always brute force over all the graphs of a given size to find the graph that we need, but we are interested if we can find one such graph in time polynomial in n. And surprisingly, this is a major open problem! It is not just a mathematical puzzle, it has connections to so called "randomness extractors", which deal with outputting perfectly random bits from biased sources. There has been a great amount of progress on this problem in the last $3-4$ years, culminating in the construction of graphs on n vertices with no clique or independent set of size $(\log n)^{\log\log n}$.

3 MAX-CUT

In the MAX-CUT problem, we are given a graph $G = (V, E)$, and the goal is to find a cut $S \subseteq V$ that maximizes edges that cross S i.e. $E(S, \overline{S})$. Computing the MAX-CUT of a given graph is a well studied **NP**-complete problem. However, in every graph, there is a cut for which at least half the edges cross. We will prove this using the probabilistic method.

Theorem 4. *For every graph G, there is a cut for which half the edges in G cross.*

Proof. As before, we take a random cut and prove that the probability that half the edges cross it is non-zero. To be precise, we take each vertex independently with probability $\frac{1}{2}$ to S. Let X be the random

variable denoting the number of edges crossing the cut S. For an edge $e = (u, v)$, X_e be the indicator random variable equalling 1 if S cuts e, and 0 otherwise. We have

$$\mathbb{E}[X_e] = \Pr(u \in S \wedge v \notin S) + \Pr(u \notin S \wedge v \in S) = \frac{1}{2}$$

By linearity of expectation, we have $\mathbb{E}[X] = \sum_{e \in E} \mathbb{E}[X_e] = \frac{|E|}{2}$. Thus, by Markov's inequality, $X \geq \frac{|E|}{2}$ with non-zero probability. □

Note that this gives a trivial randomized algorithm to approximate the MAX-CUT of a graph with approximation ratio $\frac{1}{2}$. This was the best algorithm for a long time until in 1994 when Goemans and Williamson gave an algorithm using semidefinite programming that achieves 0.878 approximation factor.

4 Applications in Extremal Combinatorics

What are the maximum possible number of edges in a graph without any triangles? What are the maximum possible number of subsets of $\{1, 2, \ldots, n\}$ where in no subset is contained in other? In Extremal Combinatorics, we study these questions like these, on the extremal properties of combinatorial objects.

Now, we demonstrate that probabilistic method can also be useful to "prove" theorems, not just showing the existence of certain objects. Let $\mathcal{F}$ be a family of subsets of $\{1, 2, \ldots, n\}$. $\mathcal{F}$ is said to be an antichain if no set in $\mathcal{F}$ is contained in a different set in $\mathcal{F}$. How large can $\mathcal{F}$ be?

Theorem 5. *(Sperner's theorem) The size of any antichain on $\{1, 2, \ldots, n\}$ is at most $\binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Proof. At first, it seems that there is no intuitive probability space associated with the problem. But we can consider the following inspired space: Let Ω be the probability space of uniformly random orderings of $\{1, 2, \ldots, n\}$. For a set $S \in \mathcal{F}$, let E_S be the event that in the random ordering, the union of first $|S|$ elements is equal to S. Note that

$$\Pr(E_S) = \frac{|S|!(n - |S|)!}{n!} = \frac{1}{\binom{n}{|S|}}.$$

Furthermore, since for every pair $S, T \in \mathcal{F}$, neither is contained in other, we can deduce that $\Pr(E_S \cap E_T) = 0$. Thus, by union bound, we get that

$$\sum_{S \in \mathcal{F}} \frac{1}{\binom{n}{|S|}} \leq 1.$$

Since for every j, $\binom{n}{j} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$, we get that

$$|\mathcal{F}| \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1$$

which proves our required claim. □

Note that the above bound is in fact tight: we can just consider all the subsets of $\{1, 2, \ldots, n\}$ that are of size equal to $\lfloor \frac{n}{2} \rfloor$.