CS-3334: Advanced Combinatorics **Fall 2022**

Lecture 3: Introduction to the Probabilistic Method

Lecturer: Kuan Yang *Scribe: Weihao Zhu*

We have introduced Paul Erdős's proof of the lower bound of $R(k, k)$ last time. This week, we will continue the introduction of the probabilistic method. We start from the review of probability.

3.1 Review of Probability and Basic Probabilistic Method

Here are some interesting problems we have discussed in class.

- What is the probability of throwing a six on a dice? Can we say that the probability is $\frac{1}{2}$ as there are just two outcomes, six or not six?
- Can we uniformly randomly pick a natural number?
- Consider an equilateral triangle inscribed in a circle. Suppose a chord of the circle is chosen at random. What is the probability that the chord is longer than a side of the triangle? (This is also known as **Bertrand paradox**.)

What do we really mean when we talk about "probability"? Before using the probabilistic method, we should clearly define what "probability" is.

3.1.1 Review of Probability

We present the definition of a probability space $(\Omega, \mathcal{F}, \mathbf{Pr}[\cdot])$ as follows:

- Ω is the set of "outcomes", which is also the sample space. It can be countable or uncountable.
- $\mathcal{F}$ is a σ-algebra (a set of all possible "events"), on which we can define probability.
- $\mathbf{Pr}[\cdot] : \mathcal{F} \to [0, 1]$ if a function such that
  - $\mathbf{Pr}[\emptyset] = 0, \mathbf{Pr}[\Omega] = 1$;
  - For any disjoint sets $A_1, \ldots, A_n, \ldots \in \mathcal{F}$, $\mathbf{Pr}[\cup A_i] = \sum \mathbf{Pr}[A_i]$.

We say $\mathcal{F}$ is a σ-algebra if it satisfies:

- $\emptyset \in \mathcal{F}$;
- $\forall A \in \mathcal{F}, A^{\complement} \in \mathcal{F}$;
- $\forall A_1, \ldots, A_n, \ldots \in \mathcal{F}, \cup A_i \in \mathcal{F}$.

The probabilistic method in combinatorics is based on the following fact:

$$\mathbf{Pr}[A] > 0 \implies A \neq \emptyset .$$

We may use the following tools to bound probabilities:

- Union Bound: For any countable sets $A_1, \ldots, A_n, \ldots$, $\mathbf{Pr}[\cup A_i] \leq \sum \mathbf{Pr}[A_i]$;
- Principle of inclusion and exclusion:

$$\mathbf{Pr}[\bigcup_{i=1}^{n} A_i] = \sum_{k=1}^{n}(-1)^{k+1} \sum_{I \in \binom{[n]}{k}} \mathbf{Pr}[\bigcap_{i \in I} A_i] = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbf{Pr}[\bigcap_{i \in I} A_i];$$

- Boole-Bonferroni Inequality:

$$\sum_{k=1}^{2t}(-1)^{k+1} \sum_{I \in \binom{[n]}{k}} \mathbf{Pr}[\cap A_i] \leq \mathbf{Pr}[\cup A_i] \leq \sum_{k=1}^{2t+1}(-1)^{k+1} \sum_{I \in \binom{[n]}{k}} \mathbf{Pr}[\cap A_i];$$

- Conditional probability: $\mathbf{Pr}[A|B] = \mathbf{Pr}[A \cap B]/\mathbf{Pr}[B]$.

Roughly speaking, a probability function is a weight function for each subset, and is countably additive. In principle, the finite probability arguments can be rephrased as counting proofs, but are usually more complicated without probabilities.

We first give two basic examples to show the power of the probabilistic method.

3.1.2 2-Colorable Hypergraphs

We say a k-uniform hypergraph $H = (V, E)$, where $E \subseteq \binom{V}{k}$, is 2-colorable if V can be colored with 2 colors such that no edge is monochromatic. For instance, when $k = 2$, it's easy to find that a 2-uniform hypergraph is a graph, and is 2-colorable if and only if it is bipartite.

Define $m(k)$ as the minimal number of edges in a k-uniform hypergraph that is not 2-colorable. When $k = 2$, it's simple to show that $m(2) = 3$ (triangle). When $k = 3$, we can prove that $m(3) = 7$ and **Fano plane** is the graph with minimal number of edges.

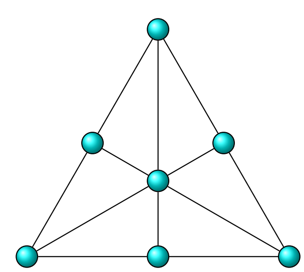

Figure 3.1: Fano plane

It is also known that $m(4) = 23$. However, we still don't know how large $m(k)$ is when $k \geq 5$.

In 1964, Paul Erdős derived a lower bound of $m(k)$ through the probabilistic method as follows.

Theorem 3.1 (Paul Erdős, 1964) $m(k) \geq 2^{k-1}$.

Proof: For any graph with $m < 2^{k-1}$ edges, we randomly color each vertex. For any edge, the probability that it is monochromatic is 2^{1-k}. Therefore, the probability that a monochromatic edge exists is no larger than $m \cdot 2^{1-k}$, which is smaller than 1. This completes the proof. ∎

In fact, a good upper bound is also obtained by him at the same time.

Theorem 3.2 (Paul Erdős, 1964) $m(k) = O(k^2 \cdot 2^k)$.

Proof: Fix the number of vertices as n, which will be determined later. We uniformly choose m edges from $\binom{[n]}{k}$ to form a k-uniform hypergraph with m edges. For any coloring $\chi : V \to \{0, 1\}$, define A_χ as the event that χ is a proper coloring in the random hypergraph. As we want to prove that there exists a k-uniform hypergraph with m edges that is not 2-colorable, it suffices to show that $\sum_\chi \mathbf{Pr}[A_\chi] < 1$.

If coloring χ colors a vertices with 0, and b vertices with 1, then for each edge e, we have

Note that $f(x) = \binom{x}{k}$ is convex.

$$\mathbf{Pr}[e \text{ is monochromatic}] = \frac{\binom{a}{k} + \binom{b}{k}}{\binom{n}{k}} \geq 2 \cdot \frac{\binom{n/2}{k}}{\binom{n}{k}}.$$

Define $p = \frac{\binom{n/2}{k}}{\binom{n}{k}}$. Therefore,

$$\mathbf{Pr}[A_\chi] = (1 - \mathbf{Pr}[e \text{ is monochromatic}])^m \leq (1 - 2p)^m,$$

which implies that

$$\sum_\chi \mathbf{Pr}[A_\chi] \leq 2^n \cdot (1 - 2p)^m < e^{n \ln 2 - 2mp}.$$

Obviously, $n \ln 2 - 2mp < 0$ suffices. Setting $n = k^2$, we can see that $m > n \ln 2/(2p) = O(k^2 \cdot 2^k)$, which completes the proof. ∎

3.1.3 List Chromatic Number

In this section, we will introduce the list chromatic number $ch(G)$, which is also known is the choice number. A *list coloring* of a graph is a proper coloring where each vertex is assigned a list of allowable colors. A graph G is said to be k-choosable, or k-list-colorable, if it has a proper coloring no matter how one assigns a list of k colors to each vertex. Then $ch(G)$ is defined as the minimum value of k such that G is k-choosable. It's easy to see that $\chi(G) \leq ch(G)$. However, the equality may not hold. Consider $K_{3,3}$ and the following allowable color lists: for the 3 vertices of the left part, assign color list $\{2, 3\}, \{1, 3\}, \{1, 2\}$ to them respectively, and assign the same three color lists to the vertices on the right.

The following proposition reveals the relationship between k-choosable bipartite graphs and 2-colorable hypergraphs.

Proposition 3.3 *If there exists a non-2-colorable k-uniform n-edge hypergraph, then $ch(K_{n,n}) > k$.*

Proof: Let $H = (V, E)$ be a non-2-colorable k-uniform hypergraph where $|E| = n$. Label vertices in $K_{n,n}$ by u_e and v_e, and assign color list e of size k. If $K_{n,n}$ has a proper coloring, let C be the set of used colors among n vertices in the left part. Then, for any vertex in H, if it belongs to C, color it by 0. Otherwise color it by 1. Clearly for each edge $e \in E$, the color of u_e is in C while the color of v_e is not in C. So it forms a 2-coloring of hypergraph H, which leads to a contradiction. ■

Corollary 3.4 $ch(K_{n,n}) > (1 - o(1)) \log_2 n$.

Since $m(k) = O(k^2 \cdot 2^k)$.

Theorem 3.5 *If $n < 2^{k-1}$, then $ch(K_{n,n}) \leq k$.*

Proof: For each color, uniformly i.i.d. mark it as L or R. For any vertex in the left/right part of $K_{n,n}$, we only use colors marked L/R to color it. For each vertex, the probability that there is no valid color for it is 2^{-k}. As long as $2n \cdot 2^{-k} < 1$, the probability that there exists valid marking is greater than zero, which implies that a valid marking and a proper coloring exist. ■

Corollary 3.6 $ch(K_{n,n}) = (1 \pm o(1)) \log_2 n$.

Actually, it has been proved that $ch(G) > (1 + o(1)) \log_2 d$ where d is the average degree of graph G. The proof is based on the hypergraph container method, which we may discuss in the future.

3.2 *Linearity of Expectation*

Expectation is also a powerful tool in combinatorics. We list some relative definitions and properties first.

- Random variable: $X : \Omega \to R$;
- Event: $A = X^{-1}(a)$;
- Conditional expectation: $\mathbf{E}[X|Y]$ (which is function $f(y) = \mathbf{E}[X|Y = y]$);
- Law of total expectation: $\mathbf{E}[X] = \mathbf{E}[\mathbf{E}[X|Y]]$;
- Averaging principle: $\mathbf{E}[X] = a \implies X \geq a / X \leq a$ is possible;

- Linearity of expectations: Let $X = c_1 X_1 + \ldots + c_n X_n$, then $\mathbf{E}[X] = c_1 \mathbf{E}[X_1] + \ldots + c_n \mathbf{E}[X_n]$. (Note that we do not need to ensure that these random variables are independent.)

  We now introduce some applications of expectations.

3.2.1 Hamiltonian Paths in Tournaments

Theorem 3.7 (Szele, 1943) *There exists a tournament (a directed graph where each pair of vertices has exactly one directed edge between them) of size n with at least $n! \cdot 2^{1-n}$ Hamiltonian paths.*

Proof: Pick a random tournament. Define X as the number of Hamiltonian paths. For each permutation π, let X_π be 1 if $\pi(1) \to \pi(2) \to \cdots \to \pi(n)$ is a path in the tournament. Otherwise, let X_π be 0.

Therefore,

$$X = \sum_\pi X_\pi \implies \mathbf{E}[X] = \sum_\pi \mathbf{E}[X_\pi] = n! \cdot 2^{1-n},$$

which completes the proof. ■

We usually call X_π an indicator random variable. The expectation of X_π is exactly the probability of the event it indicates.

This theorem was considered the first use of the probabilistic method. Szele conjectured that the maximal number of Hamiltonian paths is $n! / (2 - o(1))^n$, which was proved by Noga Alon in 1990.

3.2.2 Sum-free Sets

Now, we will introduce a "cute" result from Paul Erdős.

Theorem 3.8 (Paul Erdős,1965) *Let $A = \{a_1, \ldots, a_n\}$ be a set of n non-zero integers. There is a subset $B \subseteq A$ such that B is a sum-free set (i.e., no $a, b, c \in B$ with $a + b = c$) of size at least $n/3$.*

Proof: For $\theta \in [0, 1]$, let $S_\theta = \{n \in A : \{n\theta\} \in (\frac{1}{3}, \frac{2}{3})\}$, where $\{x\} \in [0, 1)$ is defined as the *fractional part* of a real number x. If S_θ is not sum-free, then there exists $a + b = c$ in S_θ and $a\theta + b\theta = c\theta$, which leads to a contradiction as $(\frac{1}{3}, \frac{2}{3})$ is sum-free for fractional parts. Therefore, S_θ is sum-free.

Choose θ u.a.r. from $[0, 1]$. Thus, $\mathbf{Pr}[n \in S_\theta] = \frac{1}{3}$ as $\{n\theta\}$ u.a.r. By linearity, $\mathbf{E}[|S_\theta|] = n/3$, which completes the proof. ■

This problem was used in an exam for Chinese mathematics olympiad training team. Up till now, the best lower bound we have known is $(n + 2)/3$, which was proved by Jean Bourgain in 1977.

3.2.3 Crossing Number

Define $cr(G)$ as the minimal number of crossings in a drawing of graph G with n vertices and m edges. Recall that in the first lecture, we have introduced that $K_{3,3}$ is not a planar graph. It's easy to show that $cr(K_{3,3}) = 1$. In this section, we will show a lower bound of $cr(G)$.

We first give an "easy bound" by the **Euler's formula**.

We claim that for all planar graphs (may be disconnected), $|E| \leq 3|V|$. Recall the **Euler's formula** that $v - e + f = 2$ for every connected planar graph. Computing the number of incident pairs (e, f), it's easy to show that $3f \leq 2e$. Plugging it back, we have $|E| \leq 3|V| - 6$ for all connected planar graphs. As a dis-connected planar graph can be divided into several connected ones, $|E| \leq 3|V|$ for all planar graphs.

Is it true for an isolated vertex?

We ignore constants to avoid some counterexmples.

For any graph, we consider its drawing. For each crossing, remove an edge incident to it. Then the remaining graph is planar. Therefore, $|E| - cr(G) \leq 3|V|$, which implies that $cr(G) \geq m - 3n$.

However this bound is not tight, as it only shows that $cr(G) = \Omega(n^2)$ when $m = \Omega(n^2)$, while the upper bound of $cr(G)$ is $\binom{m}{2} = \Omega(n^4)$. In 1973, Erdős and Guy conjectured that $cr(G) \geq c \cdot m^3/n^2$ for some constant $c > 0$. In 1982, the inequality was proved when $c = \frac{1}{100}$.

Theorem 3.9 (Ajtai-Chvátal-Newborn-Szemerédi, 1982) $cr(G) \geq \frac{1}{100} \cdot m^3/n^2$.

The constant factor was improved to $\frac{1}{64}$ later, and the proof was based on the probabilistic method.

Theorem 3.10 (Chazelle-Sharir-Welzl) *$cr(G) \geq \frac{1}{64} \cdot m^3/n^2$ as long as $m \geq 4n$.*

Proof: For each graph $G = (V, E)$ and a drawing, pick a real number $p \in (0, 1)$ (to be determined later). For each vertex $v \in V$, we remove it with probability $1 - p$. Thus, we obtain an induced subgraph $G' = (V', E')$. Obviously, we have

The proof which we now present arose from e-mail conversations between Bernard Chazelle, Micha Sharir and Emo Welzl.

$$\begin{aligned} \mathbf{E}[|V'|] &= pn, \\ \mathbf{E}[|E'|] &= p^2 m, \\ \mathbf{E}[cr(G')] \leq \mathbf{E}[\text{number of remaining crossings}] &= p^4 cr(G). \end{aligned}$$

Note that the easy bound $cr(G) \geq m - 3n$ holds for any graph G. Therefore,

$$\begin{aligned} & \mathbf{E}[cr(G') - (|E'| - 3|V'|)] \geq 0 \\ \Longrightarrow\ & p^4 cr(G) - p^2 m + 3pn \geq 0 \\ \Longrightarrow\ & cr(G) \geq p^{-3}(pm - 3n). \end{aligned}$$

Assume that $m \geq 4n$ and set $p = 4n/m$, we can find that $cr(G) \geq \frac{1}{64} \cdot \frac{m^3}{n^2}$. ■