10
Basic Probabilistic Methods

We now introduce a powerful tool to prove *existence*: the *probabilistic method*.

10.1 Probabilistic counting

We use the following definition for a probability space $(\Omega, \mathcal{F}, \mathbf{Pr}[\cdot])$:

- 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 simple fact:

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

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.

Paul Erdős is considered as the father of the probabilistic method. We start from his classic result on ramsey numbers.

Theorem 10.1 (Paul Erdős, 1947)**.** *$R(k,k) > n$ if*

$$\binom{n}{k} 2^{1-\binom{k}{2}} < 1 .$$

Proof. Color the edges of K_n independently and uniformly at random. Fix a set $S \in \binom{[n]}{k}$. Let $\mathcal{E}_S$ be the event that S induces a monochro-

matic K_k. It's easy to show that $\mathbf{Pr}[\mathcal{E}_S] = 2^{1-\binom{k}{2}}$. Thus, we have

$$\mathbf{Pr}[\exists \text{ a monochromatic } K_k] = \mathbf{Pr}[\cup \mathcal{E}_S] \leq \sum_S \mathbf{Pr}[\mathcal{E}_S] = \binom{n}{k} 2^{1-\binom{k}{2}} < 1.$$

It implies that the probability that no monochromatic K_k exists is not zero, which completes the proof. □

Here we use a simple but widely applied tool: the *union bound*.

Proposition 10.2 (Union bound). *For any countable sets $A_1, \ldots, A_n, \ldots$, $\mathbf{Pr}[\cup A_i] \leq \sum \mathbf{Pr}[A_i]$.*

We give more examples of basic probabilistic counting.

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.

It is also known as *property B*.

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

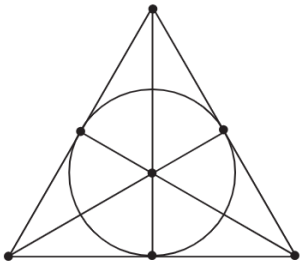

Figure 10.1: Fano plane, consisting of 7 vertices and 7 3-uniform edges

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

Theorem 10.3 (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. □

The probabilistic method can also give upper bounds.

Theorem 10.4 (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 χ colors a vertices with 0, and b vertices with 1, then we have for any edge e

Note that $x \mapsto \binom{x}{k}$ is a convex function on x.

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

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. □

A related problem is to determine the *list chromatic number* $ch(G)$, which is also known as 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 10.5. *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 10.6. $ch(K_{n,n}) > (1 - o(1)) \log_2 n$.

Proof. By Theorem 10.4, $m(k) = O(k^2 \cdot 2^k)$. □

Theorem 10.7. *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 10.8. $ch(K_{n,n}) = (1 \pm o(1)) \log_2 n$.

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.

10.2 *Linearity of expectation*

The *linearity of expectation* is also a powerful tool in combinatorics. Let $X = c_1 X_1 + \ldots + c_n X_n$, then $\mathbb{E}[X] = c_1 \mathbb{E}[X_1] + \ldots + c_n \mathbb{E}[X_n]$. (Note that we do not need to guarantee that these random variables are independent.)

Theorem 10.9 (Szele, 1943)**.** *There exists a tournament 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.

Remark 10.10. This theorem was considered the first use of the probabilistic method. Szele conjectured that the maximum number is

$\frac{n!}{(2-o(1))^n}$, which was proved by Noga Alon in 1990.

Now we consider a continuous probability space and introduce the following "cute" result from Paul Erdős.

Theorem 10.11 (Paul Erdős,1965)**.** *Let $A = \{a_1, \ldots, a_n\}$ be a set of n nonzero 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. □

Remark 10.12. 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.

As we mentioned above, a probability function is a weight function for each subset, and thus a probabilistic argument is equivalent to a counting argument. Now we reformulate some proofs we introduced before in terms of the probabilistic method.

Probabilistic proof of Erdős-Ko-Rado Theorem. Let π be a random permutation of $[n]$. Consider a circle and all contiguous blocks of size k. That is, $C_\pi = \{\{\pi((i+j) \bmod n) : j \in [k]\} : i \in [n]\}$. (We assume that $\pi(0) = \pi(n)$ here.)

For any $S \in \mathcal{F}$, let X_S be 1 if $S \in C_\pi$ and 0 otherwise. Therefore,

$$\mathbf{E}[X_S] = \mathbf{Pr}[S \in C_\pi] = \frac{n}{\binom{n}{k}}.$$

Since $\mathcal{F}$ is an intersecting family, we have $\sum X_S = |\mathcal{F} \cap C_\pi| \leq k$. To prove this, note that for every $S \in C_\pi$, there exists $2(k-1)$ other subsets in C_π intersecting S, but they can be paired off into $k-1$ distinct pairs, and two subsets in each pair are disjoint. So $\sum X_S = |\mathcal{F} \cap C_\pi| \leq k$, and thus $\sum \mathbf{E}[X_S] \leq k$, which completes the proof. □

Probabilistic proof of LYM inequality. Consider a random permutation π of $[n]$. Construct a chain

$$C_\pi = \{\emptyset, \{\pi(1)\}, \{\pi(1), \pi(2)\}, \ldots, \{\pi(1), \ldots, \pi(n)\}\}.$$

For any $S \in \mathcal{F}$, let X_S be 1 if $S \in C_\pi$ and 0 otherwise. Therefore,

$$\mathbf{E}[X_S] = \mathbf{Pr}[S \in C_\pi] = \frac{|S|!(n-|S|)!}{n!} = \frac{1}{\binom{n}{|S|}}.$$

Since $\mathcal{F}$ is an anti-chain, we have $\sum X_S = |\mathcal{F} \cap C_\pi| \leq 1$, which implies that $\sum \mathbf{E}[X_S] \leq 1$. This completes the proof. □

We can also give an alternate proof for Turán's theorem.

Theorem 10.13 (Caro-Wei inequality)**.** *For any graph G,*

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d(v)+1}.$$

Proof. Consider a random permutation π of V. Let I be the set of vertices that appear before all its neighbors. Obviously, I is an independent set.

For any vertex v, $\mathbf{Pr}[v \in I] = \frac{1}{d(v)+1}$, which implies that $\mathbb{E}[|I|] = \sum_v (d(v)+1)^{-1}$. This completes the proof. □

Notice that if we take the component of graph G, we have the following corollary.

Corollary 10.14. *For any graph G,*

$$\omega(G) \geq \sum_{v \in V(G)} \frac{1}{n - d(v)} \geq \frac{1}{1 - \frac{2m}{n^2}}.$$

The last inequality above is due to Jensen's inequality. After rearranging, we have $m \leq (1 - \frac{1}{r}) \cdot \frac{n^2}{2}$ if graph G is K_{k+1}-free, which is the Turán's Theorem (cf. Theorem 8.3).

10.3 *Crossing numbers and discrete geometry*

Define $cr(G)$ as the minimal number of crossings in a drawing of graph G with n vertices and m edges. Recall that in Section 5.5, 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 give a lower bound of $cr(G)$.

By the Euler's formula (Theorem 5.42), we know that $|E| \leq 3|V| - 6$ for any planar graph (Corollary 5.43). 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 10.15 (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 10.16 (Crossing lemma)**.** *$cr(G) \geq \frac{1}{64} \cdot m^3/n^2$ 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 $cr(G) \geq \frac{1}{64} \cdot \frac{m^3}{n^2}$. □

Actually, the crossing lemma is not well-known in a long time, until 1997 when László Székely surprisingly applied it to some geometric problems.

We first give a proof of Szemerédi–Trotter theorem (Theorem 9.3).

Proof of Szemerédi–Trotter theorem. Construct a graph $G = (V, E)$ such that $V = P$, and $E = \{(p_1, p_2) : \exists \ell \in L, p_1, p_2 \in \ell \text{ are adjacent on } \ell\}$.

Obviously, we have $|E| = \sum_{\ell \in L}(|P \cap L| - 1) \geq \frac{1}{2}(I(P, L) - |L|)$. Note that two lines share at most a common point, so the crossing number of this graph is at most $|L|^2$. Based on the crossing lemma, we have

$$|L|^2 \geq cr(G) \gtrsim \frac{|E|^3}{|V|^2} \gtrsim \frac{I(P, L)^3}{|P|^2}$$

if $I(P, L) \geq 8|P|$, which completes the proof of the theorem. □

Another application is the *unit distance problem*. Given n points on a plane, how many pairs of points are at distance 1?

Example 10.17. In a grid of size $\sqrt{n} \times \sqrt{n}$, there are $O(n)$ such pairs.

In 1946, Paul Erdős gave the following conjecture:

Conjecture 10.18 (Paul Erdős, 1946)**.** For any n points, the number of unit distances is at most $n^{1+o(1)}$.

However, Paul Erdős only derived an upper bound of $O(n^{1.5})$. In 1973, Józsa and Szemerédi improved this bound to $o(n^{1.5})$. In 1984, Beck and Spencer gave a bound of $O(n^{1.44})$. The best known bound is $O(n^{4/3})$, which was given by Spencer, Szemerédi and Trotter in 1984. Here, we will introduce the proof by Székely in 1997 as follows.

Proof of Conjecture 10.18. For each point p, if there are at least 2 points having unit distance to p, draw a unit circle centered at p. Then, for any pair (p, q), if there exists more than one arc between them, we only keep one and erase others.

Now, we have a drawing of a graph with n vertices and m edges, where $m \geq (\#\text{unit distance} - n)/2$. Since any two circles intersect at most two points, we have $cr \leq 2 \cdot \binom{n}{2}$, which implies $\#\text{unit distance} \lesssim n^{4/3}$. □