CSE 694 – Prob. Analysis and Randomized Algo.

What is it about?

- Probabilistic thinking!

Administrative Stuff

- 5 assignments (to be done individually)
- 1 final presentation and report (I will assign papers and topic)

First few weeks

- Gentle introduction to concepts and techniques from probability theory
- Done via sample problems from many areas (networking, algorithms, combinatorics, coding, learning theory, etc.)

PTCF = *Probability Theory Concepts and Facts*

Outline

Example 1: Ramsey Numbers

- The Ramsey number $R(k,k)$ is the smallest integer n such that no matter how we assign red or blue to each edge of K_n, there must exist a monochromanic K_k.
- Analogy: $R(k,k)$ is the smallest n so that in any set of n people there must be **either** k mutual acquaintances, **or** k mutual strangers

Erdős' Quote

Imagine an alien force, vastly more powerful than us landing on Earth and demanding the value of $R(5,5)$ or they will destroy our planet. In that case, we should marshal all our computers and all our mathematicians and attempt to find the value. But suppose, instead, that they asked for $R(6,6)$, we should attempt to destroy the aliens.

Notes

- There are (much) more general Ramsey numbers. E.g., $R(a, b)$ is the smallest integer n such that no matter how we 2-color edges of K_n with red and blue, there exists either a red K_a or a blue K_b.
- Or multi-dimensional Ramsey numbers (the above is 2-dim)
- The problem is a generalization of the pigeonhole principle
- Intuition/interpretation:
  - when n is sufficiently large, there must be a monochromatic sub-clique of a given size
  - i.e., in a sufficiency large "space," local "patterns" must emerge. (this theme is manifested in different ways in this course)
  - problem is to find/estimate the threshold

Erdős' Theorem (1947)

Theorem

(i) *If* $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, *then* $R(k,k) > n$.

(ii) *Consequently,* $R(k,k) > \lfloor 2^{k/2} \rfloor$ *for all* $k \geq 3$.

To see (ii), let $n = \lfloor 2^{k/2} \rfloor$.
Then,

$$\binom{n}{k}2^{1-\binom{k}{2}} < \frac{n^k}{k!} \cdot \frac{2^{1+k/2}}{2^{k^2/2}} < \frac{2^{1+k/2}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1.$$

We will give two proofs of (i).

A Pigeonhole Principle Proof

We'll show that $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ implies, there exists a 2-edge-coloring of K_n **without** a monochromatic K_k.

- Let $[n]$ be the set of vertices
- Let $\Omega =$ set of all 2-edge-colorings of K_n
- For any $S \in \binom{[n]}{k}$, the number of colorings for which S is monochromatic is $2 \times 2^{\binom{n}{2}-\binom{k}{2}}$
- The number of colorings for which some $S \in \binom{[n]}{k}$ is monochromatic is at most

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

- But, the total number of colorings is $2^{\binom{n}{2}}$, and

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

Probabilistic Method Proof #1

- Pick a coloring $c \in \Omega$ uniformly at random.
- For any $S \in \binom{[n]}{k}$, let A_S be the event that S is monochromatic, then

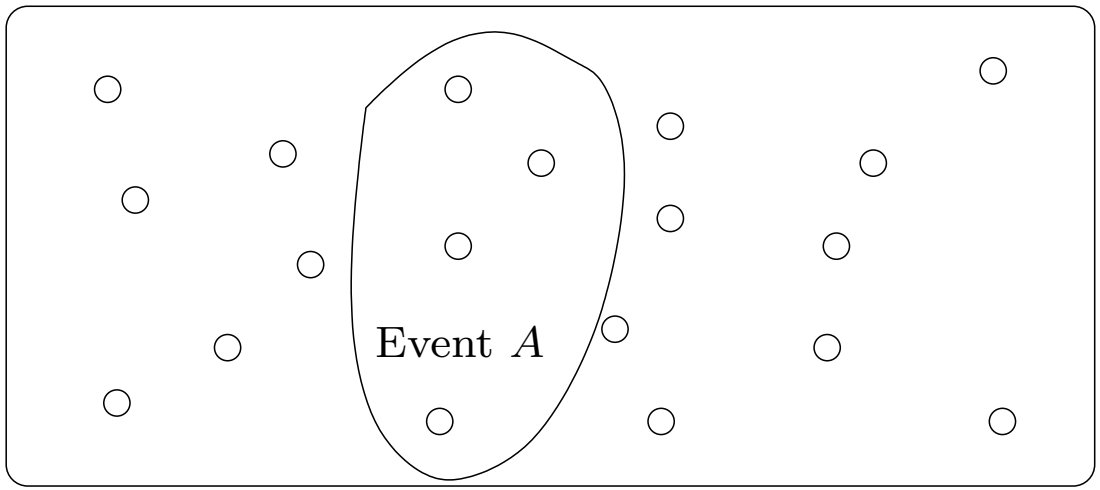
$$\text{Prob}[A_S] = \frac{\#\text{ colorings making } S \text{ mono.}}{\text{total } \#\text{ colorings}} = \frac{2 \times 2^{\binom{n}{2}-\binom{k}{2}}}{2^{\binom{n}{2}}} = 2^{1-\binom{k}{2}}$$

- The probability that some $S \in \binom{[n]}{k}$ is monochromatic is

$$\text{Prob}\left[\bigcup_S A_S\right] \leq \sum_S \text{Prob}[A_S] = \binom{n}{k} 2^{1-\binom{k}{2}} < 1$$

- Thus, there must be some coloring for which no S is monochromatic!

PTCF: Simple Probability Space



Sample Space Ω

- Ω is a finite set of all possible outcomes of some experiment
- Each outcome occurs equally likely
- A subset A of outcomes is an event
  - Think of it as a set of outcomes satisfying a certain property
- $\text{Prob}[A] = \frac{|A|}{|\Omega|}$: the fraction of outcomes in A
- In most cases, **not** a good way to think about probability spaces

PTCF: The Union Bound

Lemma

Let $A_1, A_2, \ldots$ be any finite or countably infinite sequence of events. Then,

$$\text{Prob}\left[\bigcup_{i\geq 1} A_i\right] \leq \sum_{i\geq 1} \text{Prob}[A_i]$$

Note:

- this bound holds for **any** probability space (not just simple spaces).
- the bound is simple but extremely useful!

Probabilistic Method Proof #2 (much better than #1!)

- Color each edge of K_n with either red or blue with probability $1/2$
- For any $S \in \binom{[n]}{k}$, let A_S be the event that S is monochromatic, then

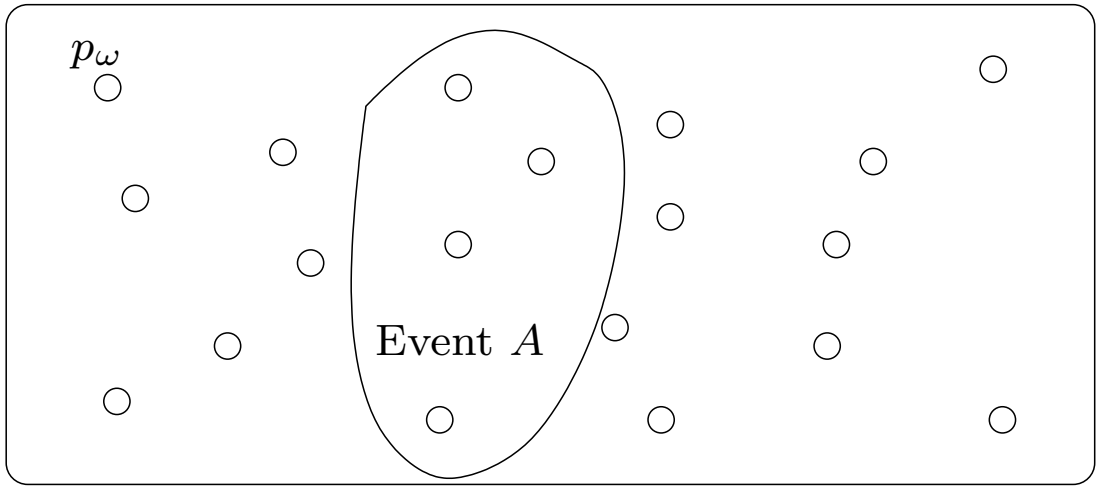
$$\text{Prob}[A_S] = \text{Prob}[S \text{ is blue}] + \text{Prob}[S \text{ is red}] = 2 \times \frac{1}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$$

- The probability that some $S \in \binom{[n]}{k}$ is monochromatic is

$$\text{Prob}\left[\bigcup_S A_S\right] \leq \sum_S \text{Prob}[A_S] = \binom{n}{k} 2^{1-\binom{k}{2}} < 1$$

- Thus, there must be some coloring for which no S is monochromatic!

PTCF: Discrete Probability Space



Sample Space Ω

- Each $\omega \in \Omega$ is assigned a number $p_\omega \in [0,1]$, such that $\sum_{\omega \in \Omega} p_\omega = 1$.
- For any event A, $\text{Prob}[A] = \sum_{\omega \in A} p_\omega$.
- In the simple space, $p_\omega = \frac{1}{|\Omega|}, \forall \omega$
- Note: this is **not** the most general definition, but suffices for now.

PTCF: How do we "assign" the p_ω?

- Could think of it as a mathematical function, like saying "give each outcome ω a number p_ω equal to $1/|\Omega|$"
- That's **not** the probabilistic way of thinking!
- Probabilistic way of thinking:
  - An experiment is an *algorithm* whose outcome is not deterministic
  - For example, algorithms making use of a random source (like a bunch of "fair" coins)
  - Ω is the set of all possible outputs of the algorithm
  - p_ω is the "likelihood" that ω is output

Example 2: Sperner Lemma

Lemma (Sperner, 1928)

The maximum size of a family $\mathcal{F}$ of subsets of $[n]$ whose members do not contain one another is $\binom{n}{\lfloor n/2 \rfloor}$.

- The collection of $\lfloor n/2 \rfloor$-subsets of $[n]$ satisfies the condition
- Suffices to show that, for any such $\mathcal{F}$, $|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$.
- Fix $F \in \mathcal{F}$, choose a permutation $\pi \in S_n$ uniformly at random
- Let A_F be the event that $F = \{\pi_1, \ldots, \pi_k\}$ for some k, then

$$\text{Prob}[A_F] = \frac{k!(n-k)!}{n!} = \frac{1}{\binom{n}{k}} \geq \frac{1}{\binom{n}{\lfloor n/2 \rfloor}}$$

- The A_F are mutually exclusive (why?), hence

$$1 \geq \text{Prob}\left[\bigcup_{F \in \mathcal{F}} A_F\right] = \sum_{F \in \mathcal{F}} \text{Prob}[A_F] \geq \frac{|\mathcal{F}|}{\binom{n}{\lfloor n/2 \rfloor}}$$

Example 3: Non-Adaptive Group Testing

- A $t \times n$ matrix $\mathbf{A}$ is called d-disjunct iff the union of any d columns does not contain another column
- Columns are codewords of superimposed codes
- Rate of the code is $R(\mathbf{a}) = \frac{\log n}{t}$
- Want codes with high rates. But, as $n \to \infty$ and $d \to \infty$

$$\frac{1}{d^2 \log e}(1 + o(1)) \leq \limsup_{\mathbf{A}} R(\mathbf{A}) \leq \frac{2 \log d}{d^2}(1 + o(1))$$

(From Dyachkov, Rykov (1982), and Dyachkov, Rykov and Rashad (1989))

- We'll prove the lower bound

Existence of Good d-disjunct Matrix

- Set a_{ij} to 1 with probability p
- The probability that $\mathbf{A}$ is **not** d-disjunct is at most

$$(d+1)\binom{n}{d+1}\left[1-p(1-p)^d\right]^t \leq$$

$$(d+1)\binom{n}{d+1}\left[1-\frac{1}{d+1}(1-\frac{1}{d+1})^d\right]^t$$

- This is < 1 as long as

$$t \geq 3(d+1)\ln\left[(d+1)\binom{n}{d+1}\right]$$

- In particular, for large n, there exist d-disjunct matrices with rate

$$\frac{\log n}{t} \approx \frac{1}{3(d+1)^2}$$

Key Ideas We've Learned

- In a sufficiently large "space," locally nice "patterns" often emerge
- To show the existence of some combinatorial object, set up some probability space and show that it exists with probability > 0
- The above is essentially a pigeonhole principle kind of proof, casted in probabilistic language
- We will see throughout the course that the probabilistic language is crucial!
- Thinking about probabilities "locally" is better than "globally"