Summer School on Probabilistic Methods

Taught by Joel Spencer
Notes by Dongryul Kim

June 26–30, 2017

This was a five-day lecture series on probabilistic methods in combinatorics, taught by Joel Spencer, New York University, and organized by Sang-il Oum, KAIST. It took place in NIMS, Daejeon, South Korea. More information about the summer school, as well as the official lecture notes can be found on the NIMS website, `http://camp.nims.re.kr/activities/eventpages/?id=207809`.

Contents

1 Lecture I: What is Erdős Magic?

Let's talk about April 1946. This was not a very good time for Korea. Korea was independent, but not really. This month is also when I was born, but this is not the reason I wrote down this moth.

Suppose the edges of K_n are colored red and blue. If n is sufficiently large, Ramsey's theorem tells that there is either a blue K_k or a red K_l. Define $R(k,l)$ as the minimal n such that this property holds.

Theorem 1.1 (Erdős 1946/47)**.** *If* $\binom{n}{k}2^{1-\binom{k}{2}} < 1$*, then* $n < R(k,k)$*.*

This theorem was proved in April 1946, and later published in 1947.

Proof. On K_n, color each edge red or blue, independently and each with equal probability. For each subset $S \subseteq \{1,\dots,n\}$, $|S| = k$, consider the event

$$\begin{aligned} \mathsf{Bad}_S: &\quad \text{all edges between the vertices in } S \text{ are of the same color,} \\ \mathsf{Bad}: &\quad \bigvee_S \mathsf{Bad}_S = \text{there exists a monochromatic } K_k. \end{aligned}$$

It is clear that $\Pr[\mathsf{Bad}_S] = 2^{1-\binom{k}{2}}$, and so we have an upper bound

$$\Pr[\mathsf{Bad}] \le \sum_S \Pr[\mathsf{Bad}_S] = \binom{n}{k}2^{1-\binom{k}{2}}.$$

Now comes Erdős magic. If the probability of Bad is less than 1, there *certainly absolutely 100% exists* a case that is not Bad. In other words, there is a coloring of the edges with no monochromatic K_k. □

We need to analyze this inequality in order to see what bound this gives on n. But we are only interested in it asymptotically. We can naïvely give an upper bound

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} < n^k.$$

Using this, we see that if

$$n^k 2^{-k^2/2} < 1,$$

or equivalently $n < 2^{k/2}$, then the inequality is satisfied. That is, we only have to focus on $n > 2^{k/2}$.

We can then give a better estimate on the binomial coefficient. Note that

$$\binom{n}{k} = \frac{n^k}{k!}\Big(1-\frac{1}{n}\Big)\Big(1-\frac{2}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big).$$

If we assume $k^2 \ll n$, then

$$\log\Big(1-\frac{1}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big) \approx -\frac{1}{n} - \cdots - \frac{k-1}{n} \approx \frac{k^2}{2n}.$$

So we get

$$\binom{n}{k} \sim \frac{n^k}{k!} e^{-k^2/2n}$$

if $k^2 \ll n$ indeed.

In our case, we have $n > 2^{k/2} \gg k^2$. So this estimate is tight and

$$\binom{n}{k} 2^{1-\binom{k}{2}} \approx c \frac{n^k}{\sqrt{k} k^k e^{-k}} 2^{k(k-1)/2}.$$

This is less than 1, asymptotically, when

$$n < \frac{k}{e\sqrt{2}} 2^{k/2}.$$

Theorem 1.2 (Erdős, 1946/47)**.** $R(k,k) > \frac{k}{e\sqrt{2}} 2^{k/2}$.

We now look at tournaments. (A tournament is a K_k with each edge oriented, or an complete oriented graph.) A tournament T has property S_k if for any k players $x_1, \ldots, x_k$, there is a player y (not equal to x_i) who wins all of them. (We exclude trivialities such as there being less than k players.)

There are some examples of small tournaments. For example, the graph on seven vertices and $i \to j$ if and only if $j - i$ is a quadratic residue modulo 7 is has property S_2. The question is, is there a graph with property S_{10}? You may or may not be able to solve this problem, but you'll definitely get a headache.

Take K_n, and give each edge a random orientation so that we get a random tournament. Given any k points, the probability that there is no point that wins all of them is

$$(1 - 2^{-k})^{n-k}.$$

So the probability that the tournament is Bad, i.e., does not satisfy S_k, is

$$\Pr[\mathsf{Bad}] \leq \binom{n}{k} (1 - 2^{-k})^{n-k}.$$

By Erdős magic, we again conclude that if the right hand side is less than 1, then there absolutely exists a tournament that satisfy S_k.

Now we need to do some asymptotic calculus to figure out the maximal k asymptotically. Here, it turns out that just using $\binom{n}{k} \leq n^k$ doesn't change the asymptotics. (This is something you don't know a priori.) So we want to find the maximal n satisfying

$$n^k (1 - 2^{-k})^{n-k} < 1.$$

Taking logarithms give

$$k \log n - 2^{-k} n < 0,$$

and so that right asymptotic is going to be $n \sim 2^k k^2 \log 2$.

Theorem 1.3 (Erdős, 1963)**.** *There is a tournament on* $2^k k^2 \log 2(1 + o(1))$ *vertices satisfying property* S_k*.*

On the other hand, any tournament satisfying S_k must have at least $2^{k+1} - 1$ vertices. This is not probabilistic. Take the player x_1 that wins at least half of the time, and look at the set of players who wins x_1. By how we picked x_1, this set is going to be at most half of the original set. In this set, take a player x_2 who wins at least half of the players of this set. If you keep on, there is going to be no player that wins all $x_1, \ldots, x_n$. This shows that the right extremal value for n is between 2^{k+1} and $(\log 2)k^2 2^k$. The right asymptotic is not known.

Let us go back to Ramsey numbers. There is a technique called alteration. This is first picking a random object and then fixing it if it is wrong.

Theorem 1.4. *For any positive integers* m, k, l*, and a real number* $0 \leq p \leq 1$*,*

$$R(k,l) \geq m - \binom{m}{k} p^{\binom{k}{2}} - \binom{m}{l} (1-p)^{\binom{l}{2}}.$$

Proof. Pick a random coloring of the edges of K_m, but here make the probability of being red p and the probability of being blue $1 - p$. Let X be the number of red K_k plus the number of blue K_l. Then the expected value of X is

$$\mathbb{E}[X] = \binom{m}{k} p^{\binom{k}{2}} + \binom{m}{l} (1-p)^{\binom{l}{2}}.$$

Now for each red K_k or blue K_l, pick a point from it, and remove it. Vertices might be selected more than once, but we know at least that there is going to be at least $m - X$ vertices. Also there cannot be a red K_k or a blue K_l. The expected of the remaining graph G is

$$\mathbb{E}[|G|] \geq m - \mathbb{E}[X] = m - \binom{m}{k} p^{\binom{k}{2}} - \binom{m}{l} (1-p)^{\binom{l}{2}}.$$

So by Erdős magic, there is a graph with no red K_k or blue K_l, with at least that many vertices. □

We now have to analyze the asymptotics, i.e., find the minimal n such that there exists such a p asymptotically. This is not easy. It is a good exercise in asymptotic calculus to find the right n for $l = 2k$.

2 Lecture II: More Erdős Magic

Let me continue and look at another favorite problem of Erdős. Here is a theorem.

Theorem 2.1. *Consider sets* $|A_i| = n$ *for* $1 \leq i \leq m$*. If* $m \leq 2^{n-1}$*, there exists a red-blue coloring of the elements so that no* A_i *is monochromatic.*

Proof. Color the points randomly. The probability of A_i being monochromatic is 2^{1-n}. So the probability that one of the sets is monochromatic is at most $m2^{1-n} < 1$. □

This is a very pure form of the probabilistic method. The question is, what if there are more sets? Define the function $m(n)$ as $m(n) > m$ if there does exist a two-coloring. For example, consider all the n-element subset of a set Ω with $|\Omega| = 2n - 1$. This shows that

$$m(n) \leq \binom{2n-1}{n} \sim c\frac{4^n}{\sqrt{n}}.$$

Erdős came up with a way of constructive a random family. Let $|\Omega| = v$, where v is a parameter. Take random subsets $A_1, \ldots, A_m \subseteq \Omega$ with $|A_i| = n$. Fix a coloring $\chi : \Omega \to \{R, B\}$, with a red points and b blue points. The probability that a set A is monochromatic is

$$\Pr[A \text{ monochromatic}] = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{v}{n}} \geq \frac{2\binom{v/2}{n}}{\binom{v}{n}}.$$

So the probability that all A are not monochromatic is

$$\Pr[\text{all } A \text{ not monochromatic}] \leq \Big[1 - \frac{2\binom{v/2}{n}}{\binom{v}{n}}\Big]^m.$$

But there are 2^v possible colorings, and so the probability that there is a coloring that makes no A_i monochromatic is

$$\Pr[\exists \chi \text{ s.t. no } A_i \text{ monochromatic}] \leq 2^v \Big[1 - \frac{2\binom{v/2}{n}}{\binom{v}{n}}\Big]^m.$$

So the theorem is, by Erdős magic, that if $2^v[1-2\binom{v/2}{n}/\binom{v}{n}]^m < 1$ then there exist $A_1, \ldots, A_m$ with no coloring. If you work out, the optimal m is around

$$m \sim \frac{v \log 2}{2\binom{v/2}{n}/\binom{v}{n}}.$$

Now we need to optimize v. Note that

$$\frac{2\binom{v/2}{n}}{\binom{v}{n}} = 2\prod_{i=0}^{n-1} \frac{v/2 - i}{v - i} = 2^{1-n} \prod_{i=0}^{n-1} \frac{v - 2i}{v - i}.$$

Here if $v = cn^2$, then

$$\log \prod_{i=0}^{n-1} \frac{v - 2i}{v - i} = \sum_{i=0}^{n-1} \log\Big(1 - \frac{i}{v - i}\Big) \sim \sum_{i=0}^{n-1} -\frac{i}{v} \sim -\frac{1}{2c}.$$

The right choice of c is $c = 1/2$ and get

$$m(n) \leq \frac{e \log 2}{4} n^2 2^n.$$

This is the best result known.

The lower bound can be improved. Here is a recent argument by Kozik–Cherkashin. They are going to define a randomized algorithm. Erdős magic then tells us that if the probability that the algorithm fails is less than 1, then there exists a coloring. Let $m = 2^{n-1}k$, and suppose $A_1, \ldots, A_m \subseteq \Omega$. Here is the surprisingly dumb algorithm.

(1) Order the elements of Ω randomly.

(2) Sequentially color the elements. If coloring v red creates a red A, then color v blue. Else, color v red.

The analysis of the algorithm is quite subtle. The algorithm fails when it creates a blue set f. Look at the first point in the ordering of f. Why was this colored blue? It is because it couldn't be colored red, i.e., there was a set e that was already colored red except for that element. Let us call this event $\mathsf{Blame}[e, f]$.

Now we want to bound $\Pr[\mathsf{Blame}[e, f]]$. We may assume that the elements of Ω are ordered randomly by assigning a uniform i.i.d. variables. Then

$$\Pr[\mathsf{Blame}[e, f]] \leq \int_0^1 y^{n-1}(1-y)^{n-1} dy \sim \frac{c}{2^{2n}\sqrt{n}}.$$

It follows that

$$\Pr\Big[\bigvee_{e,f} \mathsf{Blame}[e, f]\Big] \leq m^2 c n^{-1/2} 2^{-2(n-1)} = c k^2 n^{-1/2}.$$

Therefore the conclusion is that if $k < c_1 n^{1/4}$, the probability that the algorithm fails is less than 1 and hence there is a coloring.

Theorem 2.2. $c_1 2^{n-1} n^{1/4} \leq m(n) \leq c_2 2^{n-1} n^2$.

The lower bound actually can be improved to $c 2^{n-1} n^{1/2} / \log^c n$, with a more careful analysis of the same algorithm. This is due to Kozik–Cherkashin.

Now I would like to talk about asymptotics of the binomial coefficients. The exact formula is

$$\binom{n}{k} = \frac{n^k}{k!} A \text{ with } A = \prod_{i=0}^{k-1} \Big(1 - \frac{i}{n}\Big).$$

Here, we also have

$$B = \log A = \sum_{i=0}^{k-1} \log\Big(1 - \frac{i}{n}\Big).$$

If i is small, $\log(1 - \frac{i}{n}) = -\frac{i}{n} +$ "small". So if $k = o(\sqrt{n})$ then $B \sim 0$ and $A \sim 1$. For larger k, $k = o(n^{2/3})$, we need to look at the next term in the Taylor series.

We are then going to get $B \sim -\frac{k^2}{2n}$ and $A \sim e^{-\frac{k^2}{2n}}$. If $k = o(n^{3/4})$, we have $B \sim -\frac{k^2}{2n} - \frac{k^3}{6n^2}$ and $A \sim e^{-\frac{k^2}{2n} - \frac{k^3}{6n^2}}$. At times calculations become extremely delicate and we need these.

3 Lecture III: Asymptopia

Let Z be a random variable with zero mean. By Markov's inequality,

$$\Pr[Z \geq a] = \Pr[e^{\lambda Z} \geq e^{\lambda a}] \leq \mathbb{E}[e^{\lambda Z}]e^{-\lambda a}.$$

This is called the Chernoff bound. They key is that this works for every λ, so you pick the best λ.

Now let $X_i = \pm 1$ uniformly and i.i.d., and let $Z = \sum_{i=1}^N X_i$. In this case, because the X_i's are mutually independent,

$$\mathbb{E}[e^{\lambda Z}] = \mathbb{E}\left[\prod_{i=1}^N e^{\lambda X_i}\right] = \prod_{i=1}^N \mathbb{E}[e^{\lambda X_i}] = (\cosh \lambda)^n.$$

Here we have a nice inequality

$$\frac{e^\lambda + e^{-\lambda}}{2} = 1 + \frac{1}{2}\lambda^2 + \cdots \leq 1 + \frac{\lambda^2}{2} + \cdots = e^{\lambda^2/2}.$$

So using this inequality we can write

$$\Pr[Z \geq a\sqrt{n}] \leq e^{\lambda^2 \frac{n}{2} - \lambda a\sqrt{n}} = e^{-a^2/2},$$

after picking $\lambda = a/\sqrt{n}$. This result is useful because it works for every a and every n. The central limit theorem works for constant a and growing n, but here we can let a grow with n.

More generally, suppose $Z = \sum_{i=1}^N Y_i$ where Y_i are mutually independent with $\mathbb{E}[Y_i] = 0$, $\mathrm{Var}[Y_i] = \sigma_i^2$, and $\sigma = \sum_{i=1}^N \sigma_i^2 = \mathrm{Var}[Z]$. The Chernoff bound gives

$$\Pr[Z \geq a\sigma] \leq \mathbb{E}[e^{\lambda Z}]e^{-\lambda a\sigma}.$$

We can't completely analyze $\mathbb{E}[e^{\lambda Y_i}]$, but we hope

$$\mathbb{E}[e^{\lambda Y_i}] = 1 + \frac{\lambda^2}{2}\sigma_i^2 + (\text{stuff}) < e^{\frac{\lambda^2}{2}\sigma_i^2(1+\epsilon)},$$

which oftentimes work. When this does work, we would have

$$\Pr[Z \geq a\sigma] \leq e^{\frac{a^2}{2}(1+\epsilon)}e^{-a^2} = e^{-\frac{a^2}{2}(1-\epsilon)}.$$

When $a = o(\sigma)$, then $\lambda = o(1)$ and the (stuff) is often small.

Let me look at another example, and we move on the applications. Consider the Gaussian normal distribution N. The Chernoff bound is

$$\Pr[N \geq a] \leq \mathbb{E}[e^{\lambda N}]e^{-\lambda a} = e^{\frac{\lambda^2}{2} - \lambda a} = e^{-a^2/2}$$

where we set $\lambda = a$.

Consider a tournament T_n and a ranking (a permutation) σ. We want to find the best ranking, the σ that minimizes

$$\mathsf{Fit}[T_N, \sigma] = \#\text{nonupset} - \#\text{upset}.$$

Now define $F(N) = \min_{T_N} \max_\sigma \mathsf{Fit}[T_N, \sigma]$.

Theorem 3.1 (Erdős–Moon)**.** *$F(n) \le n^{3/2}\sqrt{\log n}$, i.e., there exists a T_n such that all σ have $\mathsf{Fit}[T_n, \sigma] \le n^{3/2}\sqrt{\log n}$.*

We will see what β is.

Proof. Take a random tournament T_n. For a fixed σ, the distribution for $\mathsf{Fit}[T_n, \sigma]$ is

$$\mathsf{Fit}[T_n, \sigma] \sim \sum_{i=1}^{\binom{n}{2}} X_i.$$

The Chernoff bound gives

$$\Pr[\mathsf{Fit}[T_n, \sigma] \ge a\sqrt{\tbinom{n}{2}}] \le e^{-a^2/2}.$$

So if

$$\Pr\Big[\bigvee_\sigma \mathsf{Fit}[T_n, \sigma] \ge a\sqrt{\tbinom{n}{2}}\Big] \le n! e - a^2/2 < 1,$$

then we can use Erdős magic. This happens when $a > \sqrt{2n \log n}$. □

So there are tournaments that can't be ranked with more than 51% accuracy. The $\log n$ can actually be removed, and then the bound is correct.

Theorem 3.2 (Spencer)**.** *$F(n) = \Theta(n^{3/2})$.*

Suppose that X_i is a general random variable with $\mathbb{E}[X_i] = 0$, but assume that we know $|X_i| \le 1$ always. In this case we again have

$$\mathbb{E}[e^{\lambda X_i}] \le \mathbb{E}[\cosh(\lambda) + \sinh(\lambda) X_i] = \cosh(\lambda).$$

Hence by the old argument,

$$\Pr[Z \ge a\sqrt{n}] \le e^{-a^2/2}.$$

Theorem 3.3. *Consider n vectors $\vec{v}_i \in \mathbb{R}^d$ with $|\vec{v}_i| \le 1$. Then there signs $\epsilon_1, \dots, \epsilon_n \in \{-1, +1\}$ such that*

$$|\epsilon_1 \vec{v}_1 + \cdots + \epsilon_N \vec{v}_N| \le \sqrt{N}.$$

Proof. Consider uniform i.i.d. variables $\epsilon_1, \ldots, \epsilon_N \in \{-1, +1\}$. Set

$$Z = \left|\sum_{i=1}^{N} \epsilon_i \vec{v}_i\right|^2 = \sum_{i,j=1}^{N} (\epsilon_i \vec{v}_i) \cdot (\epsilon_j \vec{v}_j) = \sum_{i,j=1}^{N} \epsilon_i \epsilon_j (\vec{v}_i \cdot \vec{v}_j).$$

Its expectation can be computed as

$$\mathbb{E}[Z] = \sum_{i,j=1}^{N} \vec{v}_i \cdot \vec{v}_j \mathbb{E}[\epsilon_i \epsilon_j] = \sum_i \vec{v}_i \cdot \vec{v}_i \leq N.$$

Then we use a form of Erdős magic that says that a value less than or equal to the expectation can be realized. □

Consider a game played on an $n \times n$ array of lights. Paul sets each of the lights either on or off, and Carole only can pull switches on each row and each column. When a switch is pulled, the lights on the corresponding columns and rows are changed. The payoff for Carole is the number of lights on minus the number of lights off. What is the maximum payoff Carole can get? Also why did a spell Carole in this funny way?

4 Lecture IV: Random Graphs

The reason for Carole is that it is an anagram of "oracle". Now we give strategies for both Paul and Carole. Write $x_i = 1$ if Carole doesn't switch row i and $x_i = -1$ if Carole does switch row i. Likewise define y_i for column i. Then the payoff is going to be

$$\sum_{i,j} a_{ij} x_i y_j.$$

For Paul, take $a_{ij} = \pm 1$ uniform i.i.d., and for each fixed $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, define

$$\mathsf{Bad}(\vec{x}, \vec{y}) : \sum_{i,j} a_{ij} x_i y_j > \beta.$$

Note that the random variable $\sum a_{ij} x_i y_j$ has the distribution

$$\sum_{i,j} a_{ij} x_i y_j \sim S_{n^2},$$

where we use the standard notation $S_N = \sum_{i=1}^{N} X_i$. The Chernoff bound tells us

$$\Pr[S_{n^2} \geq an] \leq e^{-a^2/2}$$

and thus

$$\Pr\left[\bigvee_{\vec{x},\vec{y}} \sum_{i,j} a_{ij} x_i y_j > \beta\right] < 4^n e^{-a^2/2} = 1$$

when $a = \sqrt{2n \log 4}$. So if we let $\beta = an = c_1 n^{3/2}$, we can use Erdős magic. This means that Paul initially set the lights so that Carole does not take more than $c_1 n^{3/2}$.

It is interesting that there is also a probabilistic strategy for Carole. Fix any $a_{ij} = \pm 1$. Take $y_1, \ldots, y_n = \pm 1$ uniformly and independently. Now the row sum has distribution

$$R_i = \sum_{j=1}^{n} a_{ij} y_j \sim S_n.$$

This is not good because the expectation of S_n is zero, but $\mathbb{E}[|R_i|] = \mathbb{E}[|S_n|] \sim c\sqrt{N}$. Select x_i so that $x_i R_i = |R_i|$, i.e., she flips the columns randomly and then flips the rows to make them positive. Now the expected payoff is

$$\mathbb{E}\Big[\sum |R_i|\Big] \sim cn^{3/2}.$$

Again by Erdős magic, there exists a $y_1, \ldots, y_n$ so that $\sum_i |R_i| \geq cn^{3/2}$. I like to call it "chaos from order". There are ways to make Carole's strategy algorithmic.

Consider the random graph $G(n, p)$. This is actually a probability space with edges draw with probability p mutually independently. We are going to look at the event

$$A: \text{ there exists a } K_4.$$

Erdős and Rényi discovered that there is a threshold value for p where the probability for A jumps from 0 to 1.

First consider, for any $|S| = 4$,

$$X_S = \begin{cases} 1 & \text{if } S \text{ is } K_4, \\ 0 & \text{else,} \end{cases} \qquad X = \sum_{|S|=4} X_S.$$

Then we immediately have $\mathbb{E}[X] = \binom{n}{4} p^6$. If $p \sim cn^{-2/3}$, then

$$\mathbb{E}[X] = \binom{n}{4} p^6 = \mu \sim \frac{c^6}{24}.$$

If $p(n) \ll n^{-2/3}$, then we get

$$\Pr[A] = \Pr[X \geq 1] \leq \mathbb{E}[X] \to 0.$$

But what about the other side? Knowing that $\mathbb{E}[X] \to \infty$ doesn't mean that $\Pr[X = 0] \to 0$. To say this, we look at the variance. Suppose $\mu = \Theta(n^4 p^6) \to \infty$. If the variance is small enough, the distribution is concentrated near μ and so the probability that $X = 0$ is small. In particular, we have Chebyshev's inequality

$$\Pr[|X - \mu| \geq \lambda \sigma] \leq \lambda^{-2}.$$

So if $\mu \to \infty$ and $\sigma^2 = o(\mu^2)$ then $\Pr[X = 0] \to 0$.

We can compute the variance as

$$\mathrm{Var}[X] = \sum \mathrm{Var}[X_S] + \sum_{S \neq T} \mathrm{Cov}[X_S, X_T].$$

Firstly, $\sum \mathrm{Var}[X_S] \leq \sum \mathbb{E}[X_S] = \mu = o(\mu^2)$. Also, the covariance of independent variables is zero, so $\mathrm{Cov}[X_S, X_T] = 0$ unless $|S \cap T| = 2$ or 3. If it is nonzero, we can bound

$$\mathrm{Cov}[X_S, X_T] \leq \mathbb{E}[X_S X_T] = \Pr[S \text{ and } T \text{ are both } K_4].$$

This shows that

$$\mathrm{Var}[X] \leq \mu + O(n^6 p^{11}) + O(n^5 p^9) = o((n^4 p^6)^2),$$

if you figure things out. So if $p(n) \gg n^{-2/3}$ then $\Pr[A] \to 1$.

Definition 4.1. A **threshold function** for A is a function $p_0(n)$ such that

(i) if $p(n) \ll p_0(n)$ then $\Pr[A] \to 0$,

(ii) if $p(n) \gg p_0(n)$ then $\Pr[A] \to 1$.

So the threshold function for the event A is $p_0(n) = n^{-2/3}$.

Many interesting events have threshold functions, but sometimes the expectation does not give the right answer. For example, consider the fish. The expectation is around $n^5 p^7$ but the threshold is not $p_0(n) = n^{-5/7}$, because there is not even a K_4 in that case.

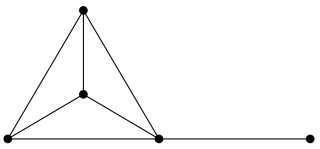


Figure 1: The fish

Definition 4.2. A graph H with v vertices is **(strictly) balanced** if for all subgraphs H' with v' vertices and e' edges, we have

$$\frac{e'}{v'} \leq \frac{e}{v}.$$

Theorem 4.3. *If H is balanced (which is actually more than strictly balanced), then $p = n^{-v/e}$ is a threshold function for containing H as a subgraph.*

The proof is exactly the same as that of for K_4.

Consider the interesting event

$$A: \text{ the graph is connected.}$$

Theorem 4.4 (Erdős–Rényi)**.** *If* $p = \dfrac{\log n}{n} + \dfrac{c}{n}$ *then* $\Pr[A] \to e^{-e^{-c}}$.

Denote by X_i the event that i is isolated, and define $X = \sum_i X_i$. Then in this case,

$$\mathbb{E}[X_i] = (1-p)^{n-1} \sim e^{-pn} = e^{-logn-c} = e^{-c}n^{-1}.$$

Then $\mu = \mathbb{E}[X] = e^{-c}$. It turns out that this is like a Poisson distribution and so the probability that there is no isolated point is like $e^{-\mu}$.

5 Lecture V: The Erdős–Rényi Phase Transition I

Here is a picture of what happens for different regions of p.

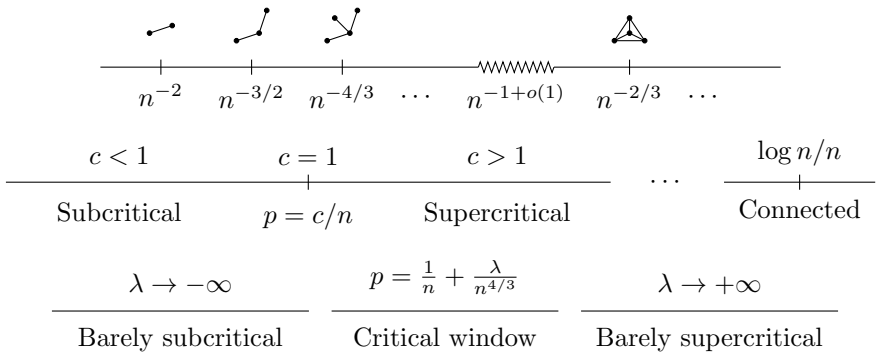


Figure 2: Different regions of p

This is like the bible. At day -2, let there be an edge. At day $-3/2$, trees with three vertices appear. At different times different stuff appears. In the second picture, we focus on $p = c/n$. This is a very natural thing to consider, because the average degree is going to be c. It turns out that the behavior for $c < 1$ is quite different from the behavior for $c > 1$. The region $c < 1$ is called the **subcritical region** and $c > 1$ is called the **supercritical region**. This is like ice turning into water. For some time Erdős and Rényi thought the cases $c < 1$, $c = 1$, $c > 1$ deal eveyrthing. But one can open up the case $c = 1$ and parametrize $p = N^{-1} + \lambda N^{-4/3}$. Constant λ is called the **critical window** and the cases $\lambda \to -\infty$ and $\lambda \to +\infty$ are called **barely subcritical** and **barely supercritical**. I will try to justify why $-3/4$ is the right parametrization.

Let us look again at connectivity. Let $p = \log n/n + cn^{-1}$ and define X_i as the characteristic function of the event of i being isolated. We have seen that

$$\mu = \mathbb{E}[X] \sim n(1-p)^{n-1} \sim e^{-c}.$$

By the inclusion-exclusion principle,

$$\Pr[X=0] = 1 - S_1 + S_2 - S_3 + \cdots \pm S_n, \quad S_r = \sum_{\{i_1,\ldots,i_r\}} \mathbb{E}[X_{i_1}\cdots X_{i_r}].$$

For fixed r and $n \to \infty$, we have

$$S_r = \binom{n}{r}(1-p)^{r(n-1)-\binom{r}{2}} \to \frac{[n(1-p)^{n-1}]^r}{r!}1 \to \frac{\mu^r}{r!}.$$

But we have a subtle issue, because this works for r fixed. To make the argument work, we use the **Bonferoni inequalities**, which say that the truncations of the inclusion-exclusion series alternately overestimate and underestimate the actual value. So for any r,

$$1 - S_1 + S_2 - \cdots - S_{2r-1} \le \Pr[X=0] \le 1 - S_1 + S_2 - \cdots + S_{2r}.$$

As we send $n \to \infty$ first and then $r \to \infty$, we get $\Pr[X=0] \to e^{-\mu}$.

We need to also take care of 2-point components, 3-point components, etc. This is a bit technical, and let's only consider the 2-point component. The expected number of a 2-point component is

$$\mathbb{E}[\#(\text{2-point component})] = \binom{n}{2}p(1-p)^{2(n-2)} \sim [n(1-p)^{n-1}]^2\frac{p}{2} = \frac{\mu^2}{2}\frac{\log n}{n}$$

which goes to 0 because of the p term. You need to add them up for k-point components.

Let us now look at $p = c/n$. For the random graph, let us write the connected components as

$$|C_1| = L_1 \ge |C_2| = L_2 \ge \cdots.$$

Also, let us call a connected graph **simple** if it is a tree or unicyclic, **complex** otherwise, and define its **complexity** as $e - v + 1$. For $c < 1$, the subcritical region, we have:

- $L_1 \sim L_2 = \Theta(\log n)$,
- all components are simple,
- the number of unicyclic components are bounded.

For $c > 1$, the supercritical region, we have:

- there is a giant component $L_1 \sim yn$, where $1 - y = e^{-cy}$,
- $L_2 = \Theta(\log n)$,
- C_1 is complex and all others are simple.

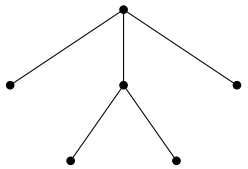
The **Poisson distribution** $\textsc{Po}(c)$ is a discrete variable defined by

$$\Pr[\textsc{Po}(c) = k] = e^{-c}\frac{c^k}{k!}.$$

This is interesting because it is a limit of the binomial distribution. As $n \to \infty$, we have

$$\textsc{Bin}\left[n - o(n), \frac{c}{n}\right] \to Po(c).$$

There is a process called the **Galton–Watson process**, which was motivated by the question of whether the British aristocracy will eventually die out. There is a root node, and each node has a number of children, independently distributed as $\textsc{Po}(c)$. To analyze this, consider i.i.d. variables $Z_1, Z_2, \ldots \sim \textsc{Po}(c)$. We are going to do a breath-first search, and assume that a person dies after it makes children. Set Z_i to be the children of the ith node, and $Y_0 = 1$, $Y_i = Y_{i-1} + Z_i - 1$. Then the society ends when Y_t becomes 0, i.e., when $T = \min\{t : Y_t = 0\}$ is finite. Now this is a random walk. If $c < 1$, there is a negative drift and so $\Pr[T < \infty] = 1$ and if $c > 1$, there is positive drift and $\Pr[T = \infty] > 0$.



Z_i		3	0	2	0	0	0
Y_i	1	3	2	3	2	1	0

Figure 3: Example of a Galton–Watson process

We can exactly compute $y = \Pr[T = \infty]$ for $c > 1$. If we define $z = 1 - y$, then

$$z = \sum_{k=0}^{\infty} \frac{e^{-c}c^k}{k!} z^k = e^{-c}e^{cz}.$$

So y is the positive root of $1 - y = e^{-cy}$.

We can compute the probability of T being k is

$$\Pr[T_c^{\textsc{Po}} = k] = \frac{e^{-ck}(ck)^{k-1}}{k!}.$$

At the critical $c = 1$, we have

$$\Pr[T_1^{\textsc{Po}} = k] = \frac{e^{-k}k^{k-1}}{k!} \sim \frac{1}{\sqrt{2\pi}} k^{-3/2}.$$

This does go to 0, but it has a heavy tail. In particular, $\mathbb{E}[T] = \infty$.

Now what does this have to do with random graphs? Consider the random graph $G(n, c/n)$ and look at the size of the component of v. At the beginning, v is connected to $\textsc{Bin}(n-1, p) \approx \textsc{Po}(c)$ number of vertices, and the next is connected to $\textsc{Bin}(n-4, p) \approx \textsc{Po}(c)$ number of vertices, and so on. But there is an ecological limitation, because the total number of vertices is n. This has to be analyzed.

6 Lecture VI: The Erdős–Rényi Phase Transition II

Let us write $T_{n,p}^{\text{GR}} = |C(v)|$, where GR stands for graph. When the ecological constraint is insignificant, we have

$$T_{n,p}^{\text{GR}} = |C(v)| \approx T_c^{\text{PO}}.$$

But we always have

$$\Pr[T_{n,p}^{\text{GR}} \geq a] \leq \Pr[T_c^{\text{PO}} \geq a].$$

This is because you can imagine T_c^{PO} as doing the random graph, but replenishing the reservoir by adding virtual points in to make the size of the reservoir always n. In the subcritical case $c < 1$, we have

$$\Pr[T_{n,p}^{\text{GR}} \geq a] \leq \Pr[T_c^{\text{PO}} \geq a] \leq e^{-\kappa a}$$

for some constant κ. Pick a large constant K such that $K\kappa > 1$ and pick $a = K \log n$. Then

$$\Pr[|C(v)| \geq K \log n] = o(n^{-1})$$

and this means that $L_1 = \max|C(v)| \leq K \log n$.

What about the supercritical case? Either the process dies out early, or it gets infinite. The picture is that the infinite ones all connect up and form the giant component.

Theorem 6.1. $\Pr[|C(v)| = t] \leq \Pr[\text{BIN}[n-1, (1-p)^t] = n - t]$.

Proof. Consider the process as points trying to join the tree. Again we do a breath-first search starting at v, but at each step of the search, the points outside the tree flips a coin to decide whether to be a child of the current node. A necessary condition for the connected component having size t is that the points outside $C(v)$ got t consecutive tails and the points inside $C(v)$ got at least one heads in the first t flips. □

There is no middle ground for that probability. That value is $o(n^{-1})$ unless either $t = O(\log n)$ or $t \sim yn$. If t is small, we have $1 - (1 - c/n)^t \sim tc/n$. Then the probability for

$$\text{BIN}\left[n, \frac{tc}{n}\right] = t$$

drops exponentially in t. In particular, it is $o(n^{-1})$ if $t \geq K \log n$.

If $t \sim wn$, then $1 - (1 - c/n)^{wn} \sim 1 - e^{-cw}$. Then we are looking at the probability that

$$\text{BIN}[n-1, 1 - e^{-cw}] = wn.$$

This is exponentially small unless $1 - e^{-cw} = w$, which is $w = y(c)$. This implies that the components are either of size yn or $O(\log n)$.

We still need to exclude the possibility of there being only small components. But we have a good approximation of the probability that the component is

small by the Galton–Watson process. This is going to be around the probability that the Galton–Watson process is finite. We know that this is around $1 - y(c)$. Thus the small components should populate around $(1 - y(c))n$ points. Making this precise becomes a bit technical, but this is the rough argument.

There was a third picture, which was discovered in the 1980s. You can make a computer simulation for this. When $\lambda = -4$, you will see an asteroid field. When $\lambda = +4$, there is going to be a dominant component, a Ceres growing to a Jupiter. As λ grows, the asteroids are sucked to the Jupiter. In fact, the size of the second largest component sometimes decreases because it is merged to the largest component.

There is some interesting physics. Consider two components of size $cn^{2/3}$ and $dn^{2/3}$. The probability that these two components collide is going to be

$$cn^{2/3}dn^{2/3}(d\lambda)n^{-4/3} = cd(d\lambda).$$

In the computer you should be able to see these colliding by tracking the top ten components. It is a fast program.

I still need to justify why $n^{-4/3}$ is the important scaling. Recall that in the Galton–Watson process,

$$\Pr[T_c^{\text{Po}} = k] = \Pr[T_1^{\text{Po}} = k]\frac{(ce^{1-c})^k}{c}.$$

Say that $c = 1 + \epsilon$. We are interested when start seeing the different between c and 1, i.e., criticality and near-criticality. Taylor expansion gives

$$(1+\epsilon)e^{-\epsilon} = (1+e)\Big(1 - \epsilon + \frac{\epsilon^2}{2} + \cdots\Big) = 1 - \frac{\epsilon^2}{2} + \cdots.$$

So

$$\Pr[T_c^{\text{Po}} = k] \sim \frac{1}{\sqrt{2\pi}}k^{-3/2}e^{-(\epsilon^2/2)k}.$$

This means that the behavior is polynomial when $k \ll \epsilon^{-2}$ and exponential when $k \gg \epsilon^{-2}$. On the other hand, the large component was around yn, and with $c = 1 + \epsilon$, you can check that $y \sim 2\epsilon$.

Now to talk about "no middle ground", the small components and the giant component has to have significantly different size. The small component is when $T_{1+\epsilon}^{\text{Po}} < \infty$. This is around $|C| < K\epsilon^{-2}$. The giant component has probability around $\Pr[T_{1+\epsilon}^{\text{Po}} = \infty] = y(1+\epsilon) \sim 2\epsilon$. So the maximal component has size $|C_{\max}| \sim 2\epsilon n$. For these two to be similar, we need $\epsilon = \lambda n^{-1/3}$, and they become distinguished when $\lambda \to \infty$.

7 Lecture VII: Games Mathematicians Play

Mathematicians play games, and my favorite game is the Liar game. There are two players, Paul and Carole, and there are three parameters N, Q, and K. Carole thinks of a number $x \in \{1, \ldots, N\}$, and Paul has Q chances of asking a

yes/no question. Paul wins if he has a strategy of always guessing the number correctly, and Carole tries to prevent this. This problem is simple; Paul splits splits splits, so Paul wins if and only if $N \leq 2^Q$.

But Carole is allowed to lie at most K times. (The case $K = 1$ is called the Diplomat's game, because a diplomat can lie but only once.) For fixed Q and K, we can try to find the maximal N such that Paul wins.

Theorem 7.1. *If $K = 1$ and $2^Q < N(Q + 1)$ then Carole wins.*

Proof. We fix a strategy for Paul, and Carole plays randomly. She doesn't listens to Paul, flips a coin and says yes or no. But if Carole "cheats", we consider it as Carole losing. (For instance, Carole can't say yes, no, yes, no to "Is x 11?" asked four times.) For $1 \leq i \leq N$, consider the random variable

$$X_i = \begin{cases} 1 & \text{if } i \text{ viable at the end,} \\ 0 & \text{if not.} \end{cases}$$

Then Carole wins if and only if $X = \sum_{i=1}^{N} X_i \geq 2$. The expectation of X_i is

$$\mathbb{E}[X_i] = \Pr[\textsc{Bin}[Q, \tfrac{1}{2}] \leq 1] = \frac{Q+1}{2^Q}.$$

So $\Pr[X > 1] > 0$.

Now this is a perfect information game, so either Paul has a strategy that wins all the time, or Carole has a strategy that wins all the time. But this argument shows that Carole always has a positive chance of winning, i.e., Paul cannot have a strategy that wins all the time. This shows that Carole must have a positive change of winning. □

But this doesn't tell us what the strategy is. Let's do the derandomization and find what the strategy for Carole is. In the middle of the game, let x be the number of possible numbers with no lie, and let y be the number of possibilities with one lie. Let's call this state $\vec{P} = (x, y)$. The initial state is going to be $(N, 0)$. Consider the weight function, with R questions remaining,

$$W_R(x, y) = \frac{x(R+1)}{2^R} + \frac{y}{2^R}.$$

The interpretation is, if Carole suddenly starts flipping coins again, the expected number of viable answers at the end.

Suppose Paul plays (a, b), i.e., he selects a numbers from the x no-lie numbers and b numbers from the y one-lie numbers and asks if the number is in that set. If Carole said yes, the state would change to

$$\vec{P}^{\text{yes}} = (a, b + (x - a)),$$

and if she said no, the state would change to

$$\vec{P}^{\text{no}} = (x - a, (y - b) + a).$$

The claim is that

$$W_R(\vec{P}) = \frac{1}{2}(W_{R-1}(\vec{P}^{\text{yes}}) + W_{R-1}(\vec{P}^{\text{no}})).$$

You can calculate and check this, but you can go back to the probability interpretation of the weights.

Now the strategy for Carole is, play so that $W_{R-1}(\vec{P}^{\text{new}}) \geq W_R(\vec{P}^{\text{old}})$! This is possible because at each time this is the average. At the start of the game, we are assuming that $W_Q(N, 0) > 1$. So $W_0(x, y) = x + y > 1$. Then $x + y \geq 2$ because it is an integer.

We can also turn this around and give a strategy for Paul. Suppose $N(Q + 1) \leq 2^Q$. The idea is that Paul asks questions so that $W_{R-1}(\vec{P}^{\text{yes}}) \approx W_{R-1}(\vec{P}^{\text{no}})$. If x and y are even, Paul can set $a = x/2$ and $b = y/2$. But there is a complication when the numbers are odd. Look at $N = Q = 5$. Here the best split Paul can make is 2 and 3. Then if Carole plays $(3, 2)$, the weight is $W_4(3, 2) = 17/16 > 1$.

Let us now look at the Vector Balancing game. Paul and Carole play N rounds with $\vec{P}^{\text{init}} = \vec{0} \in \mathbb{R}^N$. At the ith round, Paul picks a vector $\vec{i} \in \{-1, +1\}^N$ and Carole replace $\vec{P}^{\text{new}} = \vec{P}^{\text{old}} + \vec{v}_i$. The payoff for Paul is $|\vec{P}^{\text{final}}|_\infty$.

Theorem 7.2. *Carole can make so that Paul gets at most $\beta = \sqrt{2N \log N}$ points.*

Proof. Again, Carole plays randomly. Then we will get $x_i \sim S_N$ and so

$$\Pr[|S_n| \geq \beta] < e^{-\beta^2/2N} = \frac{1}{N}.$$

Consider

$$I_i = \begin{cases} 1 & \text{if } |x_i^{\text{final}}| \geq \beta, \\ 0 & \text{otherwise,} \end{cases} \qquad Z = \sum_{i=1}^{N} I_i.$$

Then $\mathbb{E}[Z] < 1$ and so there exist moves of Carole so that $Z < 1$. □

Again you can derandomize to get a deterministic strategy.

8 Lecture VIII: Needles in Exponential Haystacks

There is a set Ω with $|\Omega| = n$ and for each point $r \in \Omega$ there is an information $\mathsf{Info}(r)$, mutually independent. (For example, red/blue or true/false.) There are sets $A_\alpha \subseteq \Omega$, $\alpha \in I$ and bad events Bad_α. We want there to be an event with no bad, i.e.,

$$\bigwedge_{\alpha \in I} \overline{\mathsf{Bad}_\alpha} \neq \emptyset. \qquad (*)$$

We write $\alpha \sim \beta$ if $A_\alpha \cap A_\beta \neq \emptyset$, which is when Bad_α and Bad_β are related.

Theorem 8.1 (Lovász Local Lemma, symmetric)**.** *If all* $\Pr[\mathsf{Bad}_\alpha] \le p$*, each* α *has at most* d *many* β *with* $\alpha \sim \beta$*, and* $edp < 1$*, then* $(*)$*.*

For example, if $A_\alpha \subseteq \Omega$ have size $|A_\alpha| = n$ and Bad_α is the event that A_α is monochromatic, then $p = 2^{1-n}$. So if each A_a overlaps at most $d < 2^{n-1}/e$ other sets, then there is a coloring with no monochromatic sets.

There was a recent breakthrough by Robin Moser, who gave an algorithm for finding this needle in an haystack. Here is the $\mathsf{FIX-IT}$ algorithm.

(1) Randomly assign each variable.

(2) While set is bad,

(3) Select one of the bad sets and reassign everything that is in that set.

Let us call Log the list (in order) of sets reassigned in (3), and let TLog the length of Log. If we can prove that $\mathbb{E}[\mathsf{TLog}] < \infty$, then we get a good algorithm and there exists a solution.

Consider the example $\Omega = \{1, \dots, 8\}$ and $A = \{1, 2, 3\}$, $\dots$, $F = \{6, 7, 8\}$. Suppose $\mathsf{Log} = ADCFECBF$. We can do Tetris on this structure:

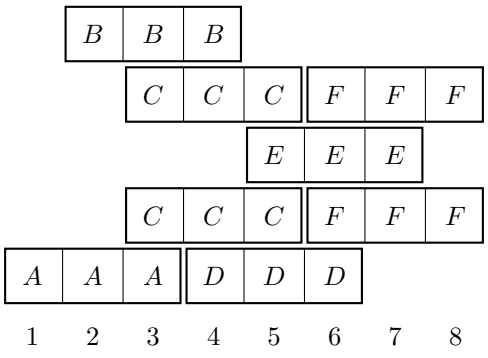


Figure 4: Example of a Tetris

We can define the "pyramid" as the string of blocks that support the last element. In this case, the pyramid is going to be $ADCFEF$. Note that the pyramids of the prefixes of Log are all distinct. This implies

$$\mathbb{E}[\mathsf{TLog}] = \sum_{s \text{ string}} \Pr[s \text{ is a pyramid of a prefix}].$$

To analyze the probability, we assume that each x_j chooses countably many assignments in advance. The probability of $X_1 \cdots X_t$ being a pyramid of a prefix is at most p^t, because all the assignments are independent.

Here there is an interesting algebra. Consider the free semialgebra generated by the tetris blocks, with the relation that X and Y commute if they have no

overlap. It can be shown that two elements are the same if and only if they have the same tetris picture. This implies

$$\mathbb{E}[\mathsf{TLog}] \leq \sum_{s \in \text{algebra}} p^{-\text{length}(s)}.$$

When is the right hand side finite? If we denote by $w_n[X]$ the sum over s of length at most n ending in X, then it can be proved that

$$w_n[X] \leq p[X] w_{n-1}[X] + p[X] \prod_{X \sim Y} (1 + w_{n-1}[Y]).$$

So in the symmetric case, if there exists a $w \geq p$ with

$$w \geq pw + p(1+w)^d,$$

then $w_n[X] \leq x$ for all n and X. This condition is satisfied if $p \sim 1/ed$ and $w \sim 1/d$.

Even if the adversary is prescient, i.e., knows what the coin flips are going to be, the adversary can't stop the algorithm from terminating.

Theorem 8.2 (Spencer, 1985)**.** *Consider $S_1, \ldots, S_n \subseteq \{1, \ldots, n\}$. For a set $S \subseteq \{1, \ldots, n\}$ and a coloring $\chi : \{1, \ldots, n\} \to \{-1, +1\}$, define its discrepancy as*

$$\text{disc}(S) = \left| \sum_{j \in S} \chi(j) \right|.$$

Then there exists a χ such that $\text{disc}(S_i) \leq 6\sqrt{n}$ for all $1 \leq i \leq n$.

I conjectured that you can't find this coloring χ in polynomial time. But in 2010 this was disproved by Bansal.

Let me outline the argument. We look at a vector formulation. Vectors $\vec{r}_1, \ldots, \vec{r}_n \in \mathbb{R}^n$ with $|\vec{r}_i|_\infty \leq 1$ are given. We want to show that if $\vec{z} \in [-1, +1]^n$ (in particular $\vec{z} = \vec{0}$) then there exists a $\vec{x} \in \{-1, +1\}^n$ with

$$|\vec{r}_i \cdot (\vec{x} - \vec{z})| \leq K\sqrt{n}$$

for all $1 \leq i \leq n$.

In Phase I, we find an $\vec{x} \in [-1, +1]^n$ with at least $n/2$ coordinates ± 1. Start at $\vec{x} = \vec{z}$, and we are going to move $\vec{x}$ in a "controlled" Brownian motion. If $|x_i| \geq 1 - \epsilon$, we "freeze" the coordinate x_i. Define

$$L_j = [n^{-1/2} \vec{r}_j] \cdot [\vec{x} - \vec{z}]$$

so that we would want all $|L_j| \leq K$.

We control the Brownian motion by restricting the space V of allowable moves $\vec{y} = (y_1, \ldots, y_n)$. If i is frozen, we set $y_i = 0$. We also set $\vec{y}$ orthogonal to the current $\vec{x}$, and also $\vec{y}$ orthogonal to $\vec{r}_j$ for j with the top $n/4$ values of $|L_j|$. After analyzing this martingale, we see that the probability of ending with

fewer than $n/5$ many j ever have $|L_j| \geq K$. But because we always freeze the top $n/4$ coordinates, this implies that all j has $|L_j| \leq K$.

In Phase s, we set $m = 2^{1-s}n$. We are going to start $\vec{z}$ with at most m coordinates frozen and end with at most $m/2$ coordinates frozen. Effectively we can get $|\vec{r}_j| \leq m$. This sum converges.

9 Lecture IX: Zero-One Laws I

From yesterday recall the game on a vector $\vec{P} \in \mathbb{R}^n$, where Paul picks a vector $\vec{v} \in \{-1, +1\}^n$, Carole either adds or subtracts that vector from $\vec{P}$, and the payoff for Paul is $|\vec{P}^{\text{final}}|_\infty$. We showed that Carole can make the payoff be at most $\sqrt{2n \log n}$.

There is also a strategy for Paul. We are going to derandomize the ramdomized strategy. Define the weight function as

$$W_r(\vec{P}) = \sum_{i=1}^{n} \Pr[|x_i + S_r| \geq \beta].$$

Paul wants to pick $\vec{v}$ so that $W_{r-1}(\vec{P} + \vec{v})$ and $W_{r-1}(\vec{P} - \vec{v})$ are "close" enough. We don't know completely if we can do this.

Let us define, for each coordinate,

$$\Delta_i = \Pr[|x_i + 1 + S_{r-1}| \geq \beta] - \Pr[|x_i - 1 + S_{r-1}| \geq \beta].$$

If $\vec{v} = (v_1, \ldots, v_n)$, we will have

$$W_{r-1}(\vec{P} + \vec{v}) - W_{r-1}(\vec{P} - \vec{v}) = \sum_{i=1}^{n} v_i \Delta_i.$$

But look at Δ_i. The event that $|x_i + 1 + S_{r-1}| \geq \beta$ but not $|x_i - 1 + S_{r-1}|$ is when S_{r-1} is exactly one value. (The parity of S_{r-1} is the parity of $r - 1$.) So we get

$$|\Delta_i| \leq c r^{-1/2}.$$

If Paul picks v_i sequentially, he can make $|\sum v_i \Delta_i| \leq c r^{-1/2}$. Then whatever Carole does, the new weight is going to be at least

$$W_{r-1}(\vec{P}^{\text{new}}) \geq W_r(\vec{P}^{\text{old}}) - c_1 r^{-1/2}.$$

Every round, we lose around $r^{-1/2}$ and so Paul can play so that

$$W_0(\vec{P}^{\text{final}}) \geq W_n(\vec{0}) - \sum_{r=1}^{n} c_1 r^{-1/2} \geq n \Pr[|S_n| \geq \beta] - c_2 n^{1/2}.$$

Theorem 9.1. *If $n \Pr[|S_n| \geq \beta] > c_2 n^{1/2}$ then Paul wins. The optimal β is around $(1 + o(1))\sqrt{n \log n}$.*

The value is then between $\sqrt{n \log n}$ and $\sqrt{2n \log n}$. The right constant is not known.

Now let us turn to the Ehrenfeucht game. There are two players, Spoiler and Duplicator. There are two graphs G_1 and G_2, and there are k rounds. On round i, Spoiler picks either $x_i \in G_1$ or $y_i \in G_2$, and Duplicator picks a point in the other graph. The goal of Duplicator is to pick the points so that x_i and x_j are adjacent if and only if y_i and y_j are adjacent. Duplicator wins if she manages to completely duplicate the graph. We are going to call this game $\mathsf{EHR}[G_1, G_2; k]$. This is a perfect information game, so either Spoiler wins or Duplicator wins. So why is this an interesting game? Glad you asked.

I want talk about first order logic and graphs. In this language, there are variables $x, y, z, \ldots$, two relations $x = y, x \sim y$ ($\sim$ means adjacent), the usual boolean stuff $\vee, \wedge, \neg, \ldots$, and quantifiers $\forall, \exists$. But very critically, the variables can only be vertices, not sets. Let us look at a few examples.

$$\begin{aligned} \text{no isolated point :} &\quad \forall x\ \exists y\ x \sim y \\ \text{there is a } K_4 \text{ :} &\quad \exists x\ \exists y\ \exists z\ \exists w\ x \sim y \wedge \cdots \wedge z \sim w \\ \text{radius at most 2 :} &\quad \forall x\ \forall y\ [y = x \vee y \sim x \vee \exists z (x \sim z \wedge z \sim y)] \end{aligned}$$

Connectivity is not expressible by first order logic. For a sentence A, we consider the **quantifier depth** $\mathrm{qd}(A)$, which is the maximal number of nested quantifiers. For instance, the sentence

$$A = \exists x\ \exists y\ (y = x \vee y \sim x) \vee \forall z\ \exists w\ (w \sim z)$$

has $\mathrm{qd}(A) = 2$.

Theorem 9.2. *Duplicator wins $\mathsf{EHR}[G_1, G_2; k]$ if and only if G_1 and G_2 satisfy the same A of quantifier quantifier depth at most k.*

The proof gets a bit technical, because it will need to actually define first order logic. But let me look at an example. Let A be the "radius at most 2", which has quantifier depth 3. Suppose A is true in G_1 but false on G_2. The claim is that Spoiler wins $\mathsf{EHR}[G_1, G_2; 3]$. First Spoiler picks two points $y_1, y_2 \in G_2$, that has distance at least 3. Spoiler then should have picked two points in G_1, not connected. Then Spoiler picks the point $x_3 \in G_1$ that is connected to both x_1 and x_2. Then Duplicator loses.

Taking this as a black box, let us now look at random graphs.

Theorem 9.3 (Fagin; Glebskii et al.)**.** *Fix a number $0 < p < 1$ and let A be any first order statement. Then*

$$\lim_{n \to \infty} \Pr[G(n,p) \vDash A] = 0 \text{ or } 1.$$

Here $\vDash A$ can be thought of as "satisfies A". What this means is

Theorem 9.4. *Let $G_1 \sim G(n,p)$ and $G_2 \sim G(m,p)$ be independent random graphs. Fix a k. Then*

$$\lim_{n,m \to \infty} \Pr[\textit{Duplicator wins } \mathsf{EHR}[G_1, G_2; k]] = 1.$$

I'm going to prove the second statement and then show you why the second implies the first.

10 Lecture X: Zero-One Laws II

Consider the sentence

$$A_\ell : \forall a, b, a + b = \ell \, \forall x_1, \ldots, x_a, y_1, \ldots, y_b \\ \exists z \; z \sim x_1 \wedge \cdots \wedge z \sim x_n \wedge \neg z \sim y_1 \wedge \cdots \wedge \neg z \sim y_b.$$

We claim that

$$\lim_{n\to\infty} \Pr[G(n,p) \vDash A_{k-1}] = 1.$$

This is because the probability of A_{k-1} not being satisfied is

$$\Pr[G(n,p) \nvDash A_{k-1}] \leq 2^{k-1} \binom{n}{k-1} (1 - 2^{1-k})^{n-k+1} \to 0.$$

It immediately follows that

$$\Pr[G(n,p) \vDash A_{k-1} \text{ and } G(m,p) \vDash A_{k-1}] \to 0$$

as $n, m \to \infty$.

For the Ehrenfeucht game $\mathsf{EHR}[G(n,p), G(m,p), k]$, Duplicate can use a very unsophicated strategy. She never looks ahead and just duplicate, which is always going to be possible. This can always be done. This prove the second theorem.

Now let's see how this implies the first result. Suppose not and assume that

$$\lim[G(n,p) \vDash A] = \alpha \in (0,1).$$

(Here I'm slightly cheating and you should use lim sup and lim inf instead.) If this happens,

$$\lim_{n,m\to\infty} \Pr[G(n,p) \vDash A \text{ and } G(m,p) \vDash \neg A] = \alpha(1-\alpha) > 0$$

and

$$\Pr[\text{Duplicator wins } \mathsf{EHR}[G(n,p), G(m,p), \mathrm{qd}(A)]] \\ \geq \Pr[G(n,p) \vDash A \text{ and } G(m,p) \nvDash A] \to \alpha(1-\alpha) > 0.$$

So we get a contradiction.

For logicians, this is the end of the story. But for us working in random graphs, p constant is only one case. There are threshold functions, and at these functions, you won't have a zero one law. I haven't proved this, but exactly at the threshold function, the probability that there is a K_4 is some nonzero non-one constant.

So for p that is not a threshold function, there is going to be some zero-one law. People came up with threshold functions, and many of them seem to be a rational power of n. So I made a conjecture, and later with Shelah managed to prove it.

Theorem 10.1 (Shelah–Spencer, 1988)**.** *Let $p(n) = n^{-\alpha}$, where $0 < \alpha < 1$ and $\alpha \notin \mathbb{Q}$. Let A be any first order statement. Then*

$$\lim_{n\to\infty} \Pr[G(n, n^{-\alpha}) \vDash A] = 0 \text{ or } 1.$$

I won't prove it because this is not a two-week summer school. We also only have to show that the probability that Duplicator wins EHR goes to 1. But here Duplicator needs to look ahead.

For example, take $\alpha = \frac{1}{2} + \epsilon$ an irrational number. Then

$$\Pr[\forall x, y\ \exists z\ x \sim z \sim y] \to 0.$$

Suppose Duplicator is not smart enough and just duplicates. Here is a strategy for Spoiler. Secretly Spoiler finds a shape $x_1 \sim x_3 \sim x_2$ in G_1 and first picks x_1, x_2. Duplicator will not see this trap and pick arbitrary $y_1, y_2 \in G_2$. Then Spoiler picks x_3 and poor Duplicator loses. This shows that Duplicator has to be smart and look out for the "dangerous" traps. It is quite complicated.

Recall that came where $\vec{P} = 0 \in \mathbb{R}^n$ initially and we can either add or subtract a vector $\vec{v}$ that hands Carole. We can interpret this as an on-line algorithm, i.e.,

Let us look at the Tenure game. There are a few pawns at level 0 (tenures), 1 (associate professors), 2 (senior assistant professors), 3 (assistant professors), etc. Paul is the chair of the department, and he wants there to be at least one tenure, in which case he wins. Carole wants there to be no tenure.

At each year, Paul write to Carole to promote some subset of the faculty. Carole then can either

- promote the list and fire the rest, or
- fire the list and promotes the list.

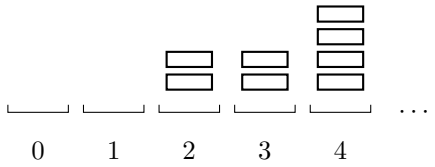


Figure 5: An initial position where Paul wins

For an initial position $(x_1, \ldots, x_n)$, if $\sum x_i 2^{-i} < 1$ then Carole wins. Her strategy is going to be playing randomly. The derandomization is going to be the weight function

$$W = \sum_i x_i 2^{-i},$$

can Carole plays so that $W^{\text{new}} \le W^{\text{old}}$. In this game, the nice thing is that Paul can exactly balance and so the weight is exact.

Theorem 10.2. *Paul wins if and only if* $\sum x_i 2^{-i} \geq 1$.

This depends on a lemma that if there are coins of value 1/2, 1/4, ... that add up to at least 1, they can be split into two pile each of which add up to at least 1/2.