Random Graphs, Thresholds and Phase transitions

We give an example of the sort of question we will look at. Let $\mathcal{A}_\triangle$ be the set of all graphs which contain $\triangle$ as a subgraph, i.e. all graphs G which contain a set of three vertices $\{u, v, w\} \in V(G)$ such that $uv, uw, vw \in E(G)$. Let the random graph $G(n, p)$ be the graph with vertex set $[n] = 1, 2, ..., n$ in which each possible edge ij, $1 \leq i < j \leq n$, is present with probability p, independently of the others.

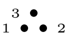
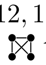
We are interested in how the probability that $G(n, p)$ contains a triangle changes for different values of p. Write $\mu_p(\mathcal{A}_\triangle, n)$ for the probability that the random graph $G(n, p)$ has a triangle. Clearly, for any n, $\mu_0(\mathcal{A}_\triangle, n) = 0$ and for $n \geq 3$, $\mu_1(\mathcal{A}_\triangle, n) = 1$. One can also show that for $p \leq p'$, $\mu_p(\mathcal{A}, n) \leq \mu_{p'}(\mathcal{A}, n)$. For edge probability $p = p(n)$, we investigate the behaviour of $\mu_{p(n)}(\mathcal{A}_\triangle, n)$ as $n \to \infty$. We will find that for $p(n)$ and $p'(n)$ 'not too far apart' that $\mu_{p(n)}(\mathcal{A}_\triangle, n) \to 0$ while $\mu_{p'(n)}(\mathcal{A}_\triangle, n) \to 1$ a sort of a 'phase transition' in the behaviour. These ideas will be made precise in this course as we investigate how 'fast' such phase transitions can occur. Material covered to include some theory of Boolean functions and the Margulis-Russo formula.

Lecture 1: Random Graphs and Thresholds

Random Graphs

For this lecture and the next we focus on a particular case of boolean analysis: graphs.

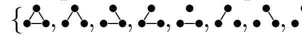
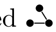
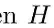
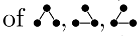
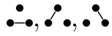
Define a graph $G = (V, E)$ to be a set of labelled vertices $[n] = \{1, 2, \ldots, n\}$ and set of two-subsets of vertices E which we call edges. Write $e(G)$ for the number of edges $|E|$. Technically the edge between vertices i and j should be denoted $\{i, j\}$ but we will use the standard shorthand ij or ji interchangably. We do not allow *loops* which are edges with both end points at the same vertex or *multiple edges* namely each pair of vertices has either zero or one edges between them). (Our graphs are *undirected* but it is possible to define *directed graphs* where each edges has a direction associated with it and $ij \neq ji$.)

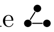
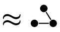
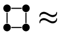
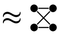
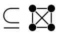
Each graph $G = ([n], E)$ can be associated with a boolean vector $x \in \{0, 1\}^{\binom{n}{2}}$, identify the $\binom{n}{2}$ positions in the vector x with the set of pairs of vertices in $[n]$ and each edge $e \in E$ is recorded by $x_e = 1$ and each non-edge by $x_e = 0$. For example (where vertex labels are always anticlockwise starting from bottom left e.g.  and the edges listed in lexicographic order e.g. $(12, 13, 23)$ and $(12, 13, 14, 23, 24, 34)$) the graph  corresponds to vector $(0, 1, 1)$, likewise  to $(1, 0, 0)$ and graph  to $(0, 1, 1, 1, 1, 1)$.

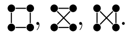
Given an integer n and a real number $0 \leq p \leq 1$, the random graph $G(n, p)$ is the graph with vertex set $[n] = 1, 2, ..., n$ in which each possible edge ij, $1 \leq i < j \leq n$, is present with probability p, independently of the others. The notation $G(n, p)$ indicates the probability space of graphs on $[n]$ with the probabilities above. We write $G \sim G(n, p)$ or to mean that G is a random graph with this distribution. For a graph H on n vertices we write $\mu_p(H) = \mu_p(H, n)$ for $\mathbb{P}(G(n, p) = H)$ and for a set of graphs on n vertices, $\mathcal{A}$, write $\mu_p(\mathcal{A}) = \mu_p(\mathcal{A}, n)$ for $\mathbb{P}(G(n, p) \in \mathcal{A})$.
For any given graph H on $[n]$, the probability of H depends only on the number of edges in H,

$$\mathbb{P}(G(n, p) = H) = p^{e(H)}(1 - p)^{\binom{n}{2} - e(H)}.$$

In the special case that $p = 1/2$, then all $\binom{n}{2}$ graphs on vertex set $[n]$ are equally likely.

Example As an example consider the probability space $G(3,p)$ where the set of possible graphs is $\{$$\}$. (Note that because the graphs are labelled  $\neq$ .) If we sample a graph $H \sim G(3,p)$ then H is  with probability p^3, for each of  the probability is $p^2(1-p)$, for each of  the probablity is $p(1-p)^2$ and finally for  the probability is $(1-p)^3$.

To study properties of random graphs we need a couple more notions from graph theory. We say that graphs H and G are *isomorphic*, denoted $H \approx G$ if there is a bijective function $\phi : V(H) \to V(G)$ such that $uv \in E(H)$ if and only if $\phi(u)\phi(v) \in E(G)$. For example  $\approx$  and  $\approx$  $\approx$ . Similarly, we say that graph H is a *subgraph* of graph G, denoted $H \subseteq G$ if there is an injective function $\phi : V(H) \to V(G)$ such that if $uv \in E(H)$ then $\phi(u)\phi(v) \in E(G)$. For example $\subseteq$ $\subseteq$  $\subseteq$  but  $\not\subseteq$ .

We also define some isomorphism classes of graphs. A graph G is a path on n vertices, denoted P_n, if its vertices can be (re)-labelled $v_1, \ldots, v_n$ such that $E(G) = \{v_i v_{i+1} : i \in [n-1]\}$. For example a path on 2 vertices is  and there are three paths on three vertices is . A graph G with $n \geq 3$ vertices, denoted C_n, is a cycle if its vertices can be (re)-labelled $v_1, \ldots, v_n$ such that $E(G) = \{v_i v_{i+1} : i \in [n]\}$ where the subscript addition is taken modulo n. For example a cycle on 3 vertices is  and there are three cycles on four vertices . A graph G is the complete graph on n vertices, denoted K_n, if $uv \in E(G)$ for all $u, v \in V(G)$. A graph G is the empty graph on n vertices, denoted $\overline{K}_n$, if $E(G) = \varnothing$. For example $K_4 =$  and $\overline{K}_4 =$ .

Thresholds in Random Graphs

We give an example of the sort of question we will look at. Let $\mathcal{A}_\triangle$ be the set of all graphs which contain  as a subgraph, i.e. all graphs G which contain a set of three vertices $\{u, v, w\} \in V(G)$ such that $uv, uw, vw \in E(G)$. We are interested in how the probability that $G(n,p)$ constains a triangle changes for different values of p. Clearly, for any n, $\mu_0(\mathcal{A}_\triangle, n) = 0$ and for $n \geq 3$, $\mu_1(\mathcal{A}_\triangle, n) = 1$. One can also show that for $p \leq p'$, $\mu_p(\mathcal{A}, n) \leq \mu_{p'}(\mathcal{A}, n)$. For edge probability $p = p(n)$, we investigate the behaviour of $\mu_{p(n)}(\mathcal{A}_\triangle, n)$ as $n \to \infty$. We will find that for $p(n)$ and $p'(n)$ 'not too far apart' that $\mu_{p(n)}(\mathcal{A}_\triangle, n) \to 0$ while $\mu_{p'(n)}(\mathcal{A}_\triangle, n) \to 1$ a sort of a 'phase transition' in the behaviour. These ideas will be made precise in this course as we investigate what are called monotone properties of graphs.

Definition 1.1 (monotone). A set of graphs $\mathcal{A}$ is *monotone* if $H \in \mathcal{A}$ and $H \subseteq G$ implies that $G \in \mathcal{A}$.

A function from $f : \{0,1\}^n \to \mathbb{R}$ is *monotone* if $f(x) \geq f(y)$ whenever $x \geq y$ i.e. for each i $x_i \geq y_i$.

Examples of monotone sets of graphs include the set of graphs containing  as a subgraph, the set of connected graphs (graphs which have a path along edges between any pair of vertices) and the set of all non-planar graphs (i.e. those that can't be drawn in the plane without edges crossing). Non-examples include the set of graphs with an odd number of edges and the set of -free graphs (those graphs *not* containing  as a subgraph).

Theorem 1.2. *For any monotone set of graphs $\mathcal{A}$ and $p' > p$,*

$$\mathbb{P}(G(n,p) \in \mathcal{A}) \leq \mathbb{P}(G(n,p') \in \mathcal{A})$$

Proof. Define $p_1 \in [0,1]$, by $p + (1-p)p_1 = p'$. Let $G \sim G(n,p)$ and $G_1 \sim G(n,p_1)$ and define the random graph $G_2 = G \cup G_1$, (this is the graph $([n], E(G) \cup E(G_1))$. Now each edge in G_2

occurs independently with probability $p + (1-p)p_1 = p'$ and hence $G_2 \sim G(n, p')$. Now because $\mathcal{A}$ is monotone

$$\mathbb{P}(G \in \mathcal{A}) \leq \mathbb{P}(G \cup G_1 \in \mathcal{A}) = \mathbb{P}(G_2 \in \mathcal{A}).$$

□

Definition 1.3 (threshold). The function $p^* = p^*(n)$ is a (coarse)[1] *threshold* for monotone $\mathcal{A}$ if $\mathbb{P}(G(n,p) \notin \mathcal{A}) \to 1$ for $p/p^* \to 0$ and $\mathbb{P}(G(n,p) \in \mathcal{A}) \to 1$ for $p/p^* \to \infty$.

Observe that if p^* is a threshold for $\mathcal{A}$ then $8p^*$ is also a threshold. We give some examples of monotone graph properties and their threshold functions.

graph property	threshold $p^*(n)$
contains 	$n^{-3/2}$
contains cycle	n^{-1}
contains △	n^{-1}
non-planar	n^{-1}
connected	$n^{-1}\log n$
contains 	$n^{-2/3}$

Threshold for a random graph containing a cycle

For a series of events $E_1, E_2, \ldots$ we say that E_n occurs *with high probability*, abbreviated *whp*, if $\mathbb{P}(E_n) \to 1$ as $n \to \infty$.

Theorem 1.4. *Let $\mathcal{A}_\circ$ be the set of graphs which contain a cycle as a subgraph then the function $p^* = \frac{1}{n}$ is a threshold for $\mathcal{A}_\circ$.*

Proof. Let $p = p(n)$ be any function such that $p/p^* \to 0$, i.e. such that $np \to 0$. Now sample the random graph $G_n \sim G(n,p)$. We want to show that whp G_n does not contain a cycle as a subgraph.

Let $X_n = X_n(G_n)$ be the random variable which counts the number of cycles in G_n. For example the number of cycles in the following graphs is #() = 1, #() = 0 and lastly #() = 3 as the graph  contains two s and the 4-cycle .

The probability that G_n has a cycle is at most the expectation of X_n:

$$\mathbb{P}(G_n \text{ has a cycle}) = \mathbb{P}(X_n > 0) = \sum_{k=1} \mathbb{P}(X_n = k) \leq \sum_{k=0} k\mathbb{P}(X_n = k) = \mathbb{E}(X_n),$$

and so it will be enough to show that $\mathbb{E}(X_n) \to 0$ as $n \to \infty$.

Let $\mathcal{S}$ be the set of all places in the graph where a cycle could occur. Explicitly, $\mathcal{S}_k$ is the set of all subsets of k vertices ordered up to rotation and orientation of the cycle and $\mathcal{S} = \cup_{k\geq 3}\mathcal{S}_k$. For $S \in \mathcal{S}$ define A_S to be the event that a cycle occurs on S in the random graph G_n. As expectation is linear,

$$\mathbb{E}(X_n) = \sum_{S\in\mathcal{S}} \mathbb{E}(1_{A_S}) = \sum_{k\geq 3}\sum_{S\in\mathcal{S}_k} \mathbb{P}(A_S) \qquad (1)$$

[1]The function $p^* = p^*(n)$ is a *sharp threshold* for monotone $\mathcal{A}$ if $\mathbb{P}(G(n,p) \notin \mathcal{A}) \to 1$ for $p < (1-\varepsilon)p^*$ and $\mathbb{P}(G(n,p) \in \mathcal{A}) \to 1$ for $p > (1+\varepsilon)p^*$.

For $S \in \mathcal{S}_k$ the probability that a cycle occurs on S is p^k as we need each of the k independent edges which form the cycle to be present in our random graph. We want to know $|\mathcal{S}_k|$. The number of ordered sets of size k is $\binom{n}{k}k!$ - which overcounts each $S \in \mathcal{S}_k$ by $2k$ times. Why $2k$? Once for each starting position on the cycle ($\times k$), and once for each direction of the cycle ($\times 2$). Hence[2] $\mathcal{S}_k = \binom{n}{k}k!/(2k) = \binom{n}{k}(k-1)!/2$. Thus by (1),

$$\mathbb{E}(X_n) = \sum_{i \geq 3} \binom{n}{k} \frac{(k-1)!}{2} p^k.$$

Now note that $\binom{n}{i}i! = n(n-1)\dots(n-i+1) \leq n^i$ and we get

$$\mathbb{E}(G_n) \leq \sum_{k \geq 3} n^k p^k = \frac{n^3p^3}{1-np},$$

which so $E(X_n)$ goes to zero for $np \to 0$. Hence as $\mathbb{P}(G_n \text{ has a cycle}) \leq \mathbb{E}(X_n)$ we have proven that whp G_n has no cycle, i.e. whp $G_n \notin \mathcal{A}_\circ$ for $p/p^* \to 0$.

For the second part of the proof we need to show that whp $G_n \in \mathcal{A}_\circ$ for $np \to \infty$. Recall (Q 2a) that any graph on n vertices with at least n edges must contain a cycle. We show that for $p = 3/n$ whp the number of edges in $G(n,p)$ is at least n. By Theorem 1.2 this implies for $p \geq 3/n$ that whp $G(n,p)$ contains a cycle as required.

Let $G_n \sim G(n, 3/n)$ and write Y_n for the number of edges in G_n. Notice $Y_n = \sum_{1 \leq i < j \leq n} 1_{ij \in E(G_n)}$ is the sum of $\binom{n}{2}$ independent random variables each of which is 1 with probability $p = 3/n$ and 0 with probability $1-p$. Thus Y has the binomial distribution[3] of $\text{bin}(\binom{n}{2}, p)$ with expectation $\mathbb{E}(Y_n) = \binom{n}{2}p$ and variance $\mathbb{V}(Y_n) = \binom{n}{2}p(1-p)$.

The expected number of edges is $\mathbb{E}(Y_n) = \binom{n}{2}\frac{3}{n} = \frac{3n}{2}(1-\frac{1}{n})$. For $n > 9$ if $|\frac{3n}{2}(1-\frac{1}{n}) - a| < n/3$ then $a > n$. Hence to show that whp the number of edges is at least n it is sufficient to show that whp $|\mathbb{E}(Y_n) - Y_n| < n/3$. But we can do this using Chebyshev's inequality

$$\mathbb{P}(|\mathbb{E}(Y_n) - Y_n| \geq n/3) \leq \frac{\mathbb{V}(Y_n)}{(n/3)^2} = \frac{3^3}{2n}\left(1 - \frac{1}{n}\right)\left(1 - \frac{3}{n}\right) \to 0.$$

Hence whp $e(G_n) \geq n$ and (consequently) whp G_n contains a cycle for $np \to \infty$. □

[2] For the purpose of the proof it would be enough to establish that $\mathcal{S}_k \leq \binom{n}{k}k!$.

[3] There are many ways to calculate the expectation and variance of the binomial random variable $\text{bin}(t,p)$ and this is not part of the course but for completeness we write out one method below.

For any random variable taking values in $\{0, 1, \dots, t\}$, one can construct the polynomial (known as the probability generating function of X), $f(x) = \sum_{i=1}^{t} \mathbb{P}(X = k)x^k$ and note that $f'(x)\big|_{x=1} = \mathbb{E}(X)$ and $f''(x)\big|_{x=1} = \mathbb{E}(X(X-1))$.

Hence for the random variable Y with distribution $\text{bin}(t,p)$ the probability generating function is $f(x) = \sum_{i=1}^{t}\binom{t}{k}p^k(1-p)^{t-k}x^k = (px-(1-p))^t$ which means

$$f'(x))\big|_{x=1} = tp(x + (1-p))^{t-1})\big|_{x=1} = np,$$

and

$$f''(x))\big|_{x=1} = t(t-1)p^2(x + (1-p))^{t-2})\big|_{x=1} = t(t-1)p.$$

Thus $\mathbb{E}(Y) = tp$ and the variance is

$$\mathbb{V}(Y) = \mathbb{E}(Y(Y-1)) + \mathbb{E}(Y) - \mathbb{E}(Y^2) = t(t-1)p^2 + np - t^2p^2 = tp(1-p).$$

Influence and Thresholds

In our work on random graphs we have been interested in finding the thresholds for monotone sets of graphs. This has meant an analysis of the function $\mu_p(\mathcal{A}) = \mathbb{P}(G(n,p) \in \mathcal{A})$. For non-trivial sets of graphs $\mathcal{A}$, the function satisfied $\mu_0(\mathcal{A}) = 0$ and $\mu_1(\mathcal{A}) = 1$ and for monotone $\mathcal{A}$ this function satisfies $\mu_{p'}(\mathcal{A}) \geq \mu_p(\mathcal{A})$ for $p' \geq p$. In this section we continue our study of this function $\mu_p(\mathcal{A})$. We will prove the Russo-Margulis lemma which allows us to calculate the derivate $\frac{d}{dp}\mu_p(\mathcal{A})$, i.e. the rate of change of the probability that a random graph $G_n \in \mathcal{A}$ as we change the edge probability p in $G_n \sim G(n,p)$. We will see that this derivative can be calculated in terms of what is called the *influence* of $\mathcal{A}$ which is an interesting property in its own right.

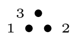
For this section we work in the general setting of a probability space over $\{0,1\}^n$.

We take the probability space Ω_n on $\{0,1\}^n$ where each bit is chosen to be 1 independently with probability p (otherwise 0). For any event $\mathcal{A}_n \subset \{0,1\}^n$ we write $\mu_p(\mathcal{A}_n)$ to be the probability that a randomly chosen $x \in \{0,1\}^n$ lies in the set $\mathcal{A}$. We write $\mu_p(x)$ to denote the probability of the event $\mathcal{A} = \{x\}$, notice

$$\mu_p(x) = p^{\sum_i x_i}(1-p)^{n-\sum_i x_i},$$

and

$$\mu_p(\mathcal{A}_n) = \sum_x \mu_p(x).$$

Recall that any vector $x \in \mathbb{F}_2^{\binom{n}{2}}$ can be associated with a graph on n vertices: identify the $\binom{n}{2}$ positions in the vector x with the set of pairs of vertices in $[n]$ and for each co-ordinate in x, $x_e = 1$ indicates that the edge e is present in the graph. Take the convention the graph is drawn with vertex labels increasing anticlockwise starting from bottom left e.g.  and that edges listed in lexicographic order e.g. $(12, 13, 23)$ and $(12, 13, 14, 23, 24, 34)$). Now the graph  corresponds to vector $(0,1,1)$, likewise  to $(1,0,0)$ and graph  to $(0,1,1,1,1,1)$. Hence the probability space defined includes the subcase of random graphs.

In this more general context the definitions of *monotone* carries over in the way you would expect[4]. We also define monotone functions.

Definition 1.5 (monotone). A function f is *monotone*, if $f(x) \geq f(y)$ whenever $x \geq y$ (i.e. $x_i \geq y_i$ for each i). A set $\mathcal{A}_n \in \{0,1\}^n$ is *monotone* if its indicator function $f_n = 1_{\mathcal{A}_n}$ is a monotone function. i.e. $f_n(x) = 1$ if $x \in \mathcal{A}$ and $f_n(x) = -1$ if $x \notin \mathcal{A}$.

In the language of voting schemes we want to say a voter has high influence if they are likely to be able to determine the outcome when we assume the rest of the popultion vote randomly. It will be on a scale of 0 to 1, where influence of 0 means they have no chance of their vote 'counting' and influence of 1 meaning that whatever the rest of the population vote the outcome would changed by the voter casting a different vote.

Definition 1.6 (pivotal). Given a function $f : \{0,1\}^n \to \mathbb{R}$ and $i \in [n]$ we say that i is pivotal for x if $f(x) \neq f(x \oplus i)$. For a set $A \subset \{0,1\}^n$ we say i is pivotal for x if it is pivotal for its indicator function 1_A.

For the n-bit vector $x = (x_1, \ldots, x_n)$ write $x \backslash \{x_i\}$ for the $(n-1)$-bit vector $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

[4]The definition of *non-trivial* does too. A monotone set $\mathcal{A}_n \in F_2^n$ is *non-trivial* if $\exists N$ such that $\forall n > N$, the n-vectors $\mathbf{0}$ and $\mathbf{1}$ satisfy $(0,0,\ldots,0) \notin \mathcal{A}_n$ and $(1,1,\ldots,1) \in \mathcal{A}_n$.

Definition 1.7 (influence of i-th bit, total influence). The influence of the i-th bit of a function f, is the probability that for a randomly chosen $x\backslash\{x_i\}$ changing the i-th co-ordinate of x changes f.

$$I_i^p(f) = \mu_p(\{x \ : \ x \neq f(x \oplus i)\}).$$

The influence of i-th bit of a set $\mathcal{A}$ is the influence of $f = 1_{\mathcal{A}}$. The *total influence* is the sum over all co-ordinates $I^p(f) = \sum_i I_i^p(f)$.

Notice that for a monotone set $\mathcal{A}$ the influence of the i-th bit is

$$I_i^p(\mathcal{A}) = \mu_p(\{x \ : \ (x_1, \ldots, x_{i-1}, 0, x_i, \ldots, x_n\} \notin \mathcal{A} \ \& \ (x_1, \ldots, x_{i-1}, 1, x_i, \ldots, x_n\} \in \mathcal{A}\}).$$

Example: In the parity function each co-ordinate has influence 1. For the dictator function $f = \mathrm{Dict}_1(\mathrm{f})$ the first co-ordinate has influence 1 the others have influence 0.

Lecture 2: Influence and Fourier analysis

Recap and influence of a set

We work with random strings of length n where we choose $x \in \{0,1\}^n$ by letting each bit independently be '1' with probability p and '0' with probability $1-p$. Equivalently - for $x \in \{0,1\}^n$, set $\mu_p(x) = p^{\sum_i x_i}(1-p)^{n-x_i}$ and for $\mathcal{A} \subset \{0,1\}^n$ set

$$\mu_p(\mathcal{A}) = \sum_{x \in \mathcal{A}} \mu_p(x) = \sum_{x \in \{0,1\}^n} \mathbf{1}_{\mathcal{A}}(x)\mu_p(x).$$

We will be interested in monotone non-trivial $\mathcal{A}_n \subset \{0,1\}^n$. To recap $\mathcal{A}_n$ is monotone if $x \leq y$ and $x \in \mathcal{A}_n$ together imply $y \in \mathcal{A}_n$. Recall that for $x, y \in \{0,1\}^n$ we say $x \leq y$ if for each i, $x_i \leq y_i$. Also to recap $\mathcal{A}_n$ is non-trivial if $\mathcal{A}_n \neq \varnothing$ and $\mathcal{A}_n \neq \{0,1\}^n$.

By yesterdays lecture we know for monotone non-trivial $\mathcal{A}_n$ that $\mu_0(\mathcal{A}_n) = 0$, $\mu_1(\mathcal{A}_n) = 1$ and for $p' \geq p$ $\mu_{p'}(\mathcal{A}_n) \geq \mu_p(\mathcal{A}_n)$. Observe that for fixed n and each $x \in \{0,1\}^n$ $\mu_p(x)$ is a polynomial in p with degree at most $\binom{n}{2}$, so $\mu_p(\mathcal{A})$ is also a polynomial in p with degree at most $\binom{n}{2}$. In particular for fixed n $\mu_p(\mathcal{A})$ is continuous and differentiable. Intuitively value $\frac{d}{dp}\mu_p(\mathcal{A})$ gives insight into how close together we may take p and p' so that $\mu_p(\mathcal{A}_n)$ is near 0 and $\mu_{p'}(\mathcal{A}_n)$ is near 1.

Recall also the influence of a function at co-ordinate i was defined be $I_i^p(f) = \mu_p(\{x : f(x) \neq f(x \oplus i)\})$ and total influence $I^p(f) = \sum_i I_i^p(f)$. Similarly define the influence of a set $\mathcal{A}_n \in \{0,1\}^n$ at co-ordinate i to be

$$I_i^p(\mathcal{A}) = \mu_p(\{x : (x \in \mathcal{A} \text{ and } x \oplus i \notin \mathcal{A}) \text{ or } (x \notin \mathcal{A} \text{ and } x \oplus i \in \mathcal{A})\})$$

and the total influence to be $I^p(\mathcal{A}_n) = \sum_i I_i^p(\mathcal{A}_n)$.

Rate of change of monotone events bounded above by influence

Lemma 2.8 (Russo-Margulis). *Let $\mathcal{A} \in \{0,1\}^n$ be a monotone event. Then*

$$\frac{d\,\mu_p(\mathcal{A})}{dp} = I^p(\mathcal{A}).$$

Proof. We consider the slightly more general case where each bit x_i is chosen to be '1' independently with probability p_i, writing $I_i^{(p_1,\ldots,p_n)}(\mathcal{A})$ for the influence of the i-th bit, i.e. the probability that the i-th bit is influential given bits $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ are chosen to be '1' independently with probabilities $p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n$ respectively.

Hence it will suffice to show that

$$\frac{d\mu_{(p_1,\ldots,p_n)}(\mathcal{A})}{dp_i} = I_i^{(p_1,\ldots,p_n)}p(\mathcal{A}),$$

WLOG take $i = 1$. Now, let $X \in F_2^{n-1}$ and $Y \in F_2^{n-1}$ be defined as follows,

$$X = \{(x_2, \ldots, x_n) : f(0, x_2, \ldots, x_n) = 1 \text{ and } f(1, x_2, \ldots, x_n) = 1\}.$$

$$Y = \{(x_2, \ldots, x_n) : f(0, x_2, \ldots, x_n) = -1 \text{ and } f(1, x_2, \ldots, x_n) = 1\}.$$

We can express the probability of the event $\mathcal{A}$ in terms of X and Y,

$$\mu_{(p_1,\ldots,p_n)}(\mathcal{A}) = \mu_{(p_2,\ldots,p_n)}(X) + \mu_{p_1}(x_1 = 1)\mu_{(p_2,\ldots,p_n)}(Y).$$

Note that Y is the pivotal set for f, and hence $\mathbb{P}_{(p_2,\ldots,p_n)}(Y) = I_1^p(f)$ and so

$$\mathbb{P}_{(p_1,\ldots,p_n)}(\mathcal{A}) = \mathbb{P}_{(p_2,\ldots,p_n)}(X) + p_1 I_1^p(f).$$

Now take the derivative of $\mathbb{P}_{(p_1,\ldots,p_n)}(\mathcal{A})$ with respect to p_1 and we are done.

□

Fourier

From now on we consider only the case $p = 1/2$ so each n-bit string is equally likely. We introduce the notions of fourier coefficients and characters in two different ways.

Intro 1 Begin by defining an inner product. For $f, g : \{0,1\}^n \to \{-1, 1\}$ define

$$< f, g >= 2^{-n} \sum_{x \in \{0,1\}^n} f(x)g(x) = \mathbb{E}_{1/2}(fg)$$

We claim that there is a set $\{\chi_S(x)\}_{S\subset[n]}$ form an orthonormal basis. Assuming the claim we can then write $f(x) = \sum_{S\subset[n]} \hat{f}(S)\chi_S(x)$ where $\hat{f}(S) =< f, \chi_S >$. The $\hat{(f)}(S)$ we call the fourier coefficients. It remains to prove the claim.

For $S \subset [n]$ define $\chi_S : \{0,1\}^n \to \{-1, 1\}$ by

$$\chi_S(x) = (-1)^{\sum_{i\in S}}.$$

One can check that $\chi_S(x)\chi_T(x) = \chi_{S\triangle T}(x)$ (where $S\triangle T$ is the set difference $(S\backslash T) \cup (T\backslash S)$.) Now for $S = T$ the set difference is empty so

$$< \chi_S, \chi_S >= 2^{-n} \sum_{x\in\{0,1\}^n} (-1)^{\sum_{j\in\varnothing} x_j} = 1.$$

For $S \neq T$, let $i \in S\triangle T$ and we will split the sum over $x \in \{0,1\}^n$ into those x for which $x_i = 0$ and those x for which $x_i = 1$ and see these sums cancel.

$$< \chi_S, \chi_T >= 2^{-n}\Big(\sum_{x\in\{0,1\}^n:\, x_i=0} (-1)^0(-1)^{\sum_{j\in S\triangle T\backslash\{i\}} x_j} + \sum_{x\in\{0,1\}^n:\, x_i=0} (-1)^{-1}(-1)^{\sum_{j\in S\triangle T\backslash\{i\}} x_j}\Big) = 0.$$

and hence we have that $\{\chi_S(x)\}_{S\subset[n]}$ form an orthonormal basis as claimed.

Intro 2 We make a sequence of rearrangements which arrive at the same result. First let $\mathbf{1}_{=x}(y)$ be '1' if $x = y$ and '0' otherwise. Then we may write

$$f(x) = \sum_{y\in\{0,1\}^n} f(y)\mathbf{1}_{=x}(y). \tag{2}$$

Now note that because f takes only values -1 and 1 we get

$$\mathbf{1}_{=x}(y) = \prod_{i=1}^{n} \mathbf{1}_{=x_i}(y_i) = \prod_{i=1}^{n} \left(\frac{1 + (-1)^{x_i+y_i}}{2}\right).$$

And thus,

$$\mathbf{1}_{=x}(y) = \frac{1}{2^n} \sum_{A\subset[n]} (-1)^{\sum_{i\in A} x_i+y_i}. \tag{3}$$

The last step is from expanding out the brackets (the set A corresponds to the term where for $i \notin A$ take the '1' from i-th bracket and for $i \in A$ we take the $(-1)^{x_i+y_i}$ term from the i-th bracket. But now by (2) and (3) we may write our function $f(x)$ as

$$f(x) = \sum_{y \in \{0,1\}^n} f(y) \mathbf{1}_{=x}(y) = \sum_{A \subset [n]} \underbrace{\left(\frac{1}{2^n} \sum_{y \in \{0,1\}^n} f(y)(-1)^{\sum_{i \in A} y_i} \right)}_{\hat{f}(A)} \Big(\underbrace{(-1)^{\sum_{i \in A} x_i}}_{\chi_A(x)} \Big).$$

and we have the same fourier decomposition as earlier.

Lecture 3: Influence and Fourier analysis

The first half of this lecture we revise Fourier notions introduced yesterday and in the second bit we show a result bounding the derivative of $\mu_p(\mathcal{A}_n)$ for monotone $\mathcal{A}$.

We pause and do a small in-lecture exercise to help internalise the definitions (if you are reading this you should pause and do this exercise too, answers in the footnote[5]).

Exercise -1. For each of the following boolean functions $f : \{0,1\}^n \to \{-1,1\}$, aka voting schemes, find a set S such that the function is expressible in terms of that character, i.e. $f(x) = \chi_S(x)$ or $f(x) = -\chi_S(x)$.

(a) The dictator function, $\mathrm{Dict}_1(\mathrm{x}) = 1$ if $x_1 = 1$ and $\mathrm{Dict}_1(\mathrm{x}) = -1$ if $x_1 = 0$.

(b) The parity function, $\mathrm{Par}(\mathrm{x}) = 1$ if $\sum_i x_i$ is odd and $\mathrm{Par}(\mathrm{x}) = -1$ if $\sum_i x_i$ is even.

(c) The XOR function of the first two inputs defined by $\mathrm{XOR}_{\{1,2\}} = 1$ if $x_1 \neq x_2$ and $\mathrm{XOR}_{\{1,2\}} = -1$ if $x_1 = x_2$.

(d) The constant function $f(x) = 1$.

We also pause to show the following.

Exercise 0. For $f : \{0,1\}^n \to \{-1,1\}$ show

(a) the expectation satisfies $\mathbb{E}_{1/2}(f) = \hat{f}(\varnothing)$ and

(b) the variance satisfies $\mathbb{V}_{1/2}(f) = \sum_{S \neq \varnothing} \hat{f}(S)$

We now make a pagebreak and then write out the answers to Exercise 0 to give the reader a chance to the exercise themselves.

[5]The functions can be written characters or their negatives as: (a) $-\chi_{\{1\}}(x)$, (b) $-\chi_{[n]}(x)$, (c) $-\chi_{\{1,2\}}(x)$ and (d) $\chi_{\varnothing}(x)$.

Exercise 0. For part (a), observe that $\chi_\varnothing(x) = 1$ for all x. Thus (all sums over $x \in \{0,1\}^n$) we get $\mathbb{E}_{1/2}(f) = 2^{-n}\sum_x f(x)\cdot 1 = 2^{-n}\sum_x f(x)\cdot\chi_\varnothing(x) = \hat{f}(\varnothing)$.

For the variance in part (b) first calculate $\mathbb{E}(f^2)$

$$\mathbb{E}(f^2) = 2^{-n}\sum_x\Big(\sum_S \hat{f}(S)\chi_S(x)\Big)\Big(\sum_T \hat{f}(T)\chi_T(x)\Big) = \sum_{S,T}\hat{f}(T)\hat{f}(S)\ \underbrace{2^{-n}\sum_x \chi_S(x)\chi_T(x)}_{<\chi_S,\chi_T>}$$

and hence because $\{\chi_S\}_S$ form an orthonormal basis we have $\mathbb{E}(f^2) = \sum_S \hat{f}(S)^2$. And so we are done by part (a).

We record one last property we will need. Observe that for $f : \{0,1\} \to \{-1,1\}$ we have $f(x)^2 = 1$ for all x. So in particular $\mathbb{E}(f^2) = 1$. Thus by our calculations in the exercise we get the following usually called Parseval's identity.

$$\sum_S \hat{f}(S)^2 = 1.$$

We may now prove the following.

Theorem 3.9. *For a monotone $\mathcal{A} \subset \{0,1\}^n$,*

$$\frac{d}{dp}\mu_p(\mathcal{A})\Big|_{p=1/2} \le \sqrt{n}.$$

Proof. By the Margulis-Russo result it is equivalent to show that for any monotone $\mathcal{A} \subset \{0,1\}^n$ the total influence of $\mathcal{A}$ at $p = 1/2$ satisfies

$$I^{1/2}(\mathcal{A}) \le \sqrt{n}.$$

We define a function $f : \{0,1\}^n \to \{-1,1\}$ which acts a little like the indicator for $\mathcal{A}$ by

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathcal{A} \\ -1 & \text{if } x \notin \mathcal{A}. \end{cases}$$

Notice $\mathcal{A}$ a monotone set implies that f is a monotone function. This monotonicity of f is important. For any $y \in \{0,1\}^n$ and $i \in [n]$ then $f(y \oplus i) \ne f(y)$ if and only if

$$f(y_1, \dots, y_{i-1}, 1, y_{i+1}, \dots, y_n) = 1 \quad \text{and} \quad f(y_1, \dots, y_{i-1}, 0, y_{i+1}, \dots, y_n) = -1.$$

We are nearly ready to start the calculations of the proof. But first we recall the behaviour of a particular character in fourier analysis.

$$\chi_{\{i\}}(y) = (-1)^{\{i\}\cdot y} = \begin{cases} 1 & \text{if } y_i = 0 \\ -1 & \text{if } y_i = 1. \end{cases}$$

Now we can begin the calculations. The game plan is to relate $\hat{f}(\{i\})$ to the influence of f.

$$\hat{f}(\{i\}) = \frac{1}{2^n}\sum_{y\in\{0,1\}^n} f(y)\chi_{\{i\}}(y) = \frac{1}{2^n}\sum_{y\in\{0,1\}^n} f(y)(1_{y_i=0}(y) - 1_{y_1=1}(y)) \tag{4}$$

Notice in (4) the second equality follows by writing out $\chi_{\{i\}}(y)$ in terms of the indicator functions $1_{y_i=0}(y)$ and $1_{y_i=1}(y)$. We can now expand out the sum in (4) to get that

$$\hat{f}(\{i\}) = \frac{1}{2^n} \sum_{y \setminus \{y_i\} \in \{0,1\}^{n-1}} f(y_1, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_n\} - f(y_1, \ldots, y_{i-1}, 1, y_{i+1}, \ldots, y_n\}. \quad (5)$$

The equation (5) rearranges nicely. If $f(y) = f(y \oplus i)$ then $f(y_1, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_n\} - f(y_1, \ldots, y_{i-1}, 1, y_{i+1}, \ldots, y_n\} = 0$ or if $f(y) \neq f(y \oplus i)$ then f monotone implies we have $f(y_1, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_n\} - f(y_1, \ldots, y_{i-1}, 1, y_{i+1}, \ldots, y_n\} = -2$. The number of times this difference of two will be recorded in (5) is half the number of such y.

$$\hat{f}(\{i\}) = \frac{1}{2^n} \times (-2) \times (|\{y \ : \ f(y) \neq f(y \oplus i)\}|/2) = -\frac{1}{2^n} |\{y \ : \ f(y) \neq f(y \oplus i)\}| \qquad (6)$$

We have now written $\hat{f}(\{i\})$ in terms of the influence of the i-th bit. Notice that (6) calculates the probability of picking a y (under $p = 1/2$) such that $f(y) \neq f(y \oplus i)$. Hence,

$$\hat{f}(\{i\}) = -I_i^{\frac{1}{2}}(f). \qquad (7)$$

Now apply some fourier analysis. By a Corollary of Parseval (see notes from first part of course) $\sum_S \hat{f}(S)^2 = 1$. We also recall the Cauchy-Schwarz inequality for real $a_1, \ldots a_n, b_1, \ldots, b_n$ which says $(\sum_i a_i b_i)^2 \leq (\sum_i a_i^2)(\sum_i b_i^2)$ i.e. $\sum_i b_i^2 \geq (\sum_i a_i b_i)^2 / (\sum_i a_i^2)$. Then (taking $a_i = 1$ and $b_i = \hat{f}(\{i\})$ to apply Cauchy-Shwartz in the second inequality),

$$1 = \sum_S \hat{f}(S)^2 \geq \sum_{i \in [n]} \hat{f}(\{i\})^2 \geq \frac{1}{n} (\sum_i \hat{f}(\{i\}))^2 = \frac{1}{n} (\sum_i I_i^{\frac{1}{2}}(f))^2 = \frac{1}{n} (I(f))^2. \qquad (8)$$

This, (8), is exactly what we want. It says $I(f) \leq \sqrt{n}$. □

Observe that the monotone condition is necessary. If we allow any possible set $\mathcal{A}$ then we could take $\mathcal{A} = \{x \ \sum_i x_i \text{ is even }\}$, i.e. so that $f = 1_{\mathcal{A}}$ is the parity function which has influence n.

We will also see that the lower bound in the theorem is tight up to constatnts. In this next example we show that the majority function, which is monotone, has total influence of $\Theta(\sqrt{n})$. For this calculation we need stirling's formula which gives the approximate growth rate of the factorial function: $n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n (1 + O(1/n))$.

Example 3.10 (Influence of Majority). For odd n, denote by $\mathrm{Maj_n}$ the function that returns 1 if more 1's than 0's in x and -1 otherwise. We show $I^{\frac{1}{2}}(\mathrm{Maj_n}) = \Theta(\sqrt{\mathrm{n}})$. As all co-ordinates have the same influence it is enough to show the influence of the first co-ordinate is what we want, i.e. $I_1^{\frac{1}{2}}(\mathrm{Maj_n}) = \Theta(\frac{1}{\sqrt{\mathrm{n}}})$.

Recall that $x \oplus 1$ denotes the vector x after the first co-ordinate has been flipped (e.g. $(1,1,0) \oplus 1 = (0,1,0)$).

$$\begin{aligned}
I_1^{\frac{1}{2}}(\mathrm{Maj_n}) &= \mu_{\frac{1}{2}}(\{x \ : \ f(x \oplus 1) \neq f(x)\}) \\
&= \frac{1}{2^n} |\{x \ : \ f(x \oplus 1) \neq f(x)\}| \\
&= \frac{1}{2^{n-1}} \{(x_2, \ldots, x_n) \ : \ \text{ exactly half the } x_i \text{ are } 1 \ \}) \\
&= \frac{1}{2^{n-1}} \binom{n-1}{\frac{n-1}{2}}
\end{aligned}$$

We can now substitute our bound on the middle binomial, everything cancels (except the desired square root) and we are finished the calculation. (The m above being $(n-1)/2$, note that n odd guarantees that this is an integer).

$$\begin{aligned}
I_1^{\frac{1}{2}}(\mathrm{Maj_n}) &= \frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}} \\
&= \Theta\left(\frac{1}{2^{n-1}}4^{\frac{n-1}{2}}\frac{1}{\sqrt{n}}\right) \\
&= \Theta\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Exercises

Exercise 1. Let $\mathcal{A}_\triangle$ be the set of all graphs which contain  as a subgraph. Fix a constant $0 < p \leq 1$, and show that $\mathbb{P}(G(n,p) \in \mathcal{A}_\triangle) \to 1$.

Exercise 2. Prove the following (for the second part it may help to use Chebyshev's inequality: for X be a random variable and $t > 0$; $\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \mathbb{V}[X]/t^2$):

Let $X_1, X_2, \ldots$ be a sequence of random variables each taking non-negative integer values. If $\mathbb{E}[X_n] \to 0$ then

$$\mathbb{P}(X_n = 0) \to 1,$$

and if $\mathbb{E}[X_n] > 0$ for each n, and $\mathbb{V}[X_n]/(\mathbb{E}[X_n])^2 \to 0$ then

$$\mathbb{P}(X_n = 0) \to 0.$$

Exercise 3. Show whp $np \to \infty$ implies whp G_n contains  i.e. a 3-cycle.
Let Y_n count the number of  in G_n and for any 3-subset of vertices $S \subset V(G)$ let A_S be the event that G_n restricted to the vertices S is a .

(a) Show by linearity of expectation that:

$$\mathbb{V}(Y_n) = \sum_{S,T \in \binom{[n]}{3}} \Big(\mathbb{P}(A_S \,\&\, A_T) - \mathbb{P}(A_S)\mathbb{P}(A_T) \Big)$$

where $\binom{[n]}{3}$ denotes the set of sets of three vertices in the graph.

(b) After some case analysis and from (a) show: $\mathbb{V}(Y_n) \leq n^4 p^5 + n^3 p^3$.

(c) From (b) conclude that whp $Y_n > 0$.

Exercise 4. Show that the function $p^*(n) = \frac{1}{n^{2/3}}$ is a threshold for $G(n,p)$ containing  as a subgraph.

Exercise 5. Given $k \in \mathbb{N}$, let $\mathcal{P}_k$ be the set of graphs which have a path on k vertices as a subgraph.

(a) Find the threshold function for $\mathcal{P}_3$ (notice $\mathcal{P}_3$ is the set of graphs containing the path  as a subgraph).

(b) Find the threshold for $\mathcal{P}_4$.

(c) Let $k \in \mathbb{N}$ be a constant. Find the threshold for $\mathcal{P}_k$ in terms of k and n.

Exercise 6. We can define an interated majority function for $n = 3^k$. The base case is $\text{Imaj}_1(x_1, x_2, x_3) = \text{Maj}_3(x_1, x_2, x_3)$ and

$$\text{Imaj}_k(x) = \text{Maj}_3(\text{Imaj}_{k-1}(x_1, \ldots, x_{3^{k-1}}), \text{Imaj}_{k-1}(x_{3^{k-1}+1}, \ldots, x_{2.3^{k-1}}), \text{IMaj}_{k-1}(x_{2.3^{k-1}+1}, \ldots, x_{3^k})).$$

(a) Calculate the influence of the i-th bit $I_i^p(\text{Imaj}_2)$ and total influence $I^p(\text{Imaj}_2)$.

(b) For $p = 1/2$ calculate $I_i^p(\text{Imaj}_k)$ and $I^p(\text{Imaj}_k)$.

Exercise 7. Suppose $\mathcal{A}$ is non-trivial monotone and let $p_c(n)$ be such that

$$\mathbb{P}(G(n, p_c(n)) \in \mathcal{A}_n) = \frac{1}{2}$$

and then show that for $p_b(n) = 1 - (1 - p_c(n))^k$ we have

$$\mathbb{P}(G(n, p_b(n)) \in \mathcal{A}_n) = 1 - \frac{1}{2^k}.$$

Sketches/solutions for some of the exercises.

Exercise 1. Let $\mathcal{A}_\triangle$ be the set of all graphs which contain $\triangle$ as a subgraph. Fix a constant $0 < p \leq 1$, and show that $\mathbb{P}(G(n,p) \in \mathcal{A}_\triangle) \to 1$.

Solution: Group the vertices into $n/3$ sets of three. Now for the random graph to have no triangles, in particular, each group of three must not induce a triangle. For any fixed set of three vertices there is a p^3 chance that there is a triangle. Thus calculating the probability that there is no triangle in the graph we get $\mathbb{P}(G(n,p) \notin \mathcal{A}_\triangle) \leq (1-p^3)^{n/3} \to 0$.

Exercise 2. Prove the following (for the second part it may help to use Chebyshev's inequality: for X be a random variable and $t > 0$; $\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \mathbb{V}[X]/t^2$):

Let $X_1, X_2, \ldots$ be a sequence of random variables each taking non-negative integer values. If $\mathbb{E}[X_n] \to 0$ then

$$\mathbb{P}(X_n = 0) \to 1,$$

and if $\mathbb{E}[X_n] > 0$ for each n, and $\mathbb{V}[X_n]/(\mathbb{E}[X_n])^2 \to 0$ then

$$\mathbb{P}(X_n = 0) \to 0.$$

Exercise 3. Show whp $np \to \infty$ implies whp G_n contains $\triangle$ i.e. a 3-cycle. Let Y_n count the number of $\triangle$ in G_n and for any 3-subset of vertices $S \subset V(G)$ let A_S be the event that G_n restricted to the vertices S is a $\triangle$.

Solution: The solution follows in a similar way to the threshold calculations of $\boxtimes$ for which we give a full solution.

Exercise 4. Show that the function $p^*(n) = \frac{1}{n^{2/3}}$ is a threshold for $G(n,p)$ containing $\boxtimes$ as a subgraph.

Solution:
Let p be such that $p/p^* \to 0$, i.e. $n^{2/3}p \to 0$ and let $G_n \sim G(n,p)$. Let $Y_n = Y_n(G_n)$ count the number of $\boxtimes$ in G_n. For each set S of 4 vertices from $[n]$, let A_S be the event that G_n restricted to the vertices S is a $\boxtimes$. We can write Y_n in terms of these indicator random variables.

$$Y_n = \sum_{S \in \binom{[n]}{4}} 1_{A_S}$$

As expectation is linear, the expected number of $\boxtimes$ in G_n is

$$\mathbb{E}(Y_n) = \sum_{S \in \binom{[n]}{4}} \mathbb{E}(1_{A_S}) \leq n^4 p^6 = (n^{2/3}p)^6.$$

Hence $\mathbb{E}(Y_n) \to 0$ for $n^{2/3}p \to 0$. Observe $\mathbb{P}(G_n \text{ contains a } \boxtimes) = \mathbb{P}(Y_n > 0) \leq \mathbb{E}(Y_n)$ and so for $n^{2/3}p \to 0$ whp G_n does not contain $\boxtimes$ as a subgraph.

Now it remains to show that for $p/p^* \to \infty$, i.e. for $n^{2/3}p \to \infty$ that whp $G_n \sim G(n,p)$ contains a $\boxtimes$. For this part of the proof we calculate the variance of Y_n by writing $Y_n = \sum_S 1_{A_S}$ and expanding. Write $\sum_S$ for $\sum_{S \in \binom{[n]}{4}}$.

$$\mathbb{V}(Y_n) = \mathbb{E}(Y_n^2) - \mathbb{E}(Y_n)^2 \quad = \quad \mathbb{E}\left(\left(\sum_S 1_{A_S}\right)^2\right) - \left(\sum_S \mathbb{E}(1_{A_S})\right)^2.$$

We can rearrange a little to get an expression for the variance in terms of the probabilities of the events A_S and A_T

$$\begin{aligned}\mathbb{V}(Y_n) &= \mathbb{E}\Big(\sum_S 1_{A_S} \sum_{T\in[n]^3} 1_{A_T}\Big) - \sum_S \mathbb{E}(1_{A_S}) \sum_T \mathbb{E}(1_{A_T}) \\ &= \sum_{S,T} \Big(\mathbb{E}(1_{A_S} 1_{A_T}) - \mathbb{E}(1_{A_S})\mathbb{E}(1_{A_T})\Big) \\ &= \sum_{S,T} \Big(\mathbb{P}(A_S \ \& \ A_T) - \mathbb{P}(A_S)\mathbb{P}(A_T)\Big). \qquad (9)\end{aligned}$$

If $S \cap T = \varnothing$, i.e. the vertex subsets S and T are disjoint then the events A_S and A_T are independent. Notice this is also true if S and T intersect in one vertex because they still share no edges in common. Hence if $|S \cap T| \leq 1$ then $\mathbb{P}(A_S \ \& \ A_T) = \mathbb{P}(A_S)\mathbb{P}(A_T)$ and these terms cancel in the expression for the variance (9) above.

So by this observation and (9),

$$\mathbb{V}(Y_n) \leq \sum_{|S\cap T|=\{2,3,4\}} \mathbb{P}(A_S \ \& \ A_T). \qquad (10)$$

We now consider the three options: $|S \cap T| = 2, 3, 4$. For each of these, for $S, T \in \binom{[n]}{4}$ with the given intersection we want to calculate $\mathbb{P}(A_S \ \& \ A_T)$. For $|S \cap T| = 2$, one edge is shared. There are 10 other edges that need to be present in order to have $\boxtimes$ on both S and on T. Hence $\mathbb{P}(A_S \ \& \ A_T) = p^{11}$ for $|S \cap T| = 2$. Similarly, for $|S \cap T| = 3$, we get $\mathbb{P}(A_S \ \& \ A_T) = p^9$ and for $|S \cap T| = 4$, we get $\mathbb{P}(A_S \ \& \ A_T) = \mathbb{P}(A_S) = p^6$.

The aim is to find an upper bound for the right hand side of (10). Hence we want to know how many $S, T \in \binom{[n]}{4}$ for each of the possible overlaps. When S and T overlap on 2 vertices, the number of ways to pick them is to first pick the set of vertices in S then pick the two vertices in S that will ovelap with T, and lastly pick the last two vertices in T (the ones that don't overlap with S). This makes $\binom{n}{4}\binom{4}{2}\binom{n}{2}$. Actually all we need is that the number of $S, T \in \binom{[n]}{4}$ which overlap on two vertices is at most n^6. Similarly the number that overlap on three vertices is at most n^5 and the number overlapping on all four vertices is at most n^4.

We can now that calculate an explicit upper bound on our variance. From (9),

$$\mathbb{V}(Y_n) \leq n^6 p^{11} + n^5 p^9 + n^4 p^6. \qquad (11)$$

Now we have a good upper bound on the variance. What we actually want to show is that whp G_n contains a $\boxtimes$. In other words we want to show whp $Y_n > 0$.

We use the following non-obvious idea. I have some b for which I know $b > 0$ and I want to use this to show that $a > 0$. Notice it is enough to show that $|b - a| < b$.

Let's go. By some re-arraning and Chebyshev,

$$\mathbb{P}(Y_n > 0) \geq \mathbb{P}\Big(|Y_n - \mathbb{E}(Y_n)| < \mathbb{E}(Y_n)/2\Big) = 1 - \mathbb{P}\Big(|Y_n - \mathbb{E}(Y_n)| \geq \mathbb{E}(Y_n)/2\Big) \geq 1 - \frac{4\mathbb{V}(Y_n)}{\mathbb{E}(Y_n)^2}.$$

The problem is now reduced to terms we have already calculated. By (11),

$$\mathbb{P}(Y_n > 0) \geq 1 - \frac{n^6 p^{11} + n^5 p^9 + n^4 p^6}{\binom{n}{4} p^6}. \qquad (12)$$

For $n^{2/3}p \to \infty$ the fraction in (12) goes to zero. Hence for $n^{2/3}p \to \infty$ whp $G(n,p)$ contains a  as a subgraph.

Exercise 5. Given $k \in \mathbb{N}$, let $\mathcal{P}_k$ be the set of graphs which have a path on k vertices as a subgraph.

(a) Find the threshold function for $\mathcal{P}_3$ (notice $\mathcal{P}_3$ is the set of graphs containing the path  as a subgraph).

(b) Find the threshold for $\mathcal{P}_4$.

(c) Let $k \in \mathbb{N}$ be a constant. Find the threshold for $\mathcal{P}_k$ in terms of k and n.

Exercise 6. We can define an interated majority function for $n = 3^k$. The base case is $\mathrm{Imaj}_1(x_1, x_2, x_3) = \mathrm{Maj}_3(x_1, x_2, x_3)$ and

$$\mathrm{Imaj}_k(x) = \mathrm{Maj}_3(\mathrm{Imaj}_{k-1}(x_1, \dots, x_{3^{k-1}}), \mathrm{Imaj}_{k-1}(x_{3^{k-1}+1}, \dots, x_{2.3^{k-1}}), \mathrm{IMaj}_{k-1}(x_{2.3^{k-1}+1}, \dots, x_{3^k})).$$

(a) Calculate the influence of the i-th bit $I_i^p(\mathrm{Imaj}_2)$ and total influence $I^p(\mathrm{Imaj}_2)$.

(b) For $p = 1/2$ calculate $I_i^p(\mathrm{Imaj}_k)$ and $I^p(\mathrm{Imaj}_k)$.

Exercise 7. Suppose $\mathcal{A}$ is non-trivial monotone and let $p_c(n)$ be such that

$$\mathbb{P}(G(n, p_c(n)) \in \mathcal{A}_n) = \frac{1}{2}$$

and then show that for $p_b(n) = 1 - (1 - p_c(n))^k$ we have

$$\mathbb{P}(G(n, p_b(n)) \in \mathcal{A}_n) = 1 - \frac{1}{2^k}.$$

Solution Consider the union of k copies of $G(n, p_0)$, for some k which we will decide later. Let $H = ([n], E(G_1) \cup \ldots \cup E(G_k))$ where each $G_i \sim G(n, p_0)$. Here the graphs G_i are all defined on the same vertex set $[n]$, and H is the random graph on this vertex set with edge set the union of the edge sets of the G_i. For any given $i \neq j \in [n]$ the probability ij is not in the edge set of H is exactly the probability that the edge ij does not appear in any of the G_i, which is $(1-p)^k$. This means H is the random graph where each edge is present independently with probability $1 - (1-p)^k$. Thus $H \sim G(n, 1 - (1-p)^k)$.

The next idea is to notice that $\mathcal{A}$ monotone means that $H \in \mathcal{A}$ if $\exists i$ such that $G_i \in \mathcal{A}$. Thus,

$$\mathbb{P}(H \in \mathcal{A}) \leq 1 - \mathbb{P}(\forall i, G_i \notin \mathcal{A}) = 1 - \mathbb{P}(G_1 \notin \mathcal{A})^k = 1 - \frac{1}{2^k}. \tag{13}$$

Appendix: Probability Recap

Lemma 5.11 (Markov's inequality)**.** *If X is a random variable taking only non-negative values and $t > 0$, then $\mathbb{P}(X \geq t) \leq \mathbb{E}[X]/t$.*

Proof. (of Markov's inequality) Let $1_{X \geq t}$ be the indicator function of the event that $X \geq t$. Then always (with probability 1), the random variable X satisfies the relation $X \geq t 1_{X \geq t}$. Now take the expectation of both sides to get

$$\mathbb{E}(X) \geq t\mathbb{E}(1_{X \geq t}) = t\mathbb{P}(X \geq t).$$

□

Recall the variance $\mathbb{V}[X]$ of a random variable X is defined by

$$\mathbb{V}[X] = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Lemma 5.12 (Chebyshev Inequality)**.** *Let X be a random variable and let $t > 0$. Then*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathbb{V}[X]}{t^2}.$$

Proof. This follows from Markov's inequality. We consider the probability of the event that the difference between X and its expectation it at least t. As t is positive,

$$\mathbb{P}\big(|X - \mathbb{E}[X]| \geq t\big) = \mathbb{P}\big((X - \mathbb{E}[X])^2 \geq t^2\big).$$

Then Markov's inequality applies to show this is less than $\mathbb{E}(X - \mathbb{E}(X))^2/t^2$ which is simply $\mathbb{V}[X]/t^2$ and we are done. □