Lecture 5

1 Introduction to random graphs

In Lecture 3 we used a graph which was generated randomly in order to prove a lower bound on diagonal Ramsey numbers. Such graphs have numerous applications and thus call for a systematic study.

1.1 Basic models of random graphs

Definition 1.1. *The Erdős-Rényi random graph model $G(n,m)$ is the probability space $(\Omega, \mathbb{P})$ where $\Omega = \{G = ([n], E) : |E| = m\}$ and $\mathbb{P}$ is the uniform distribution, that is, $\mathbb{P}(G) = \binom{\binom{n}{2}}{m}^{-1}$ for every $G \in \Omega$.*

This model allows us to study the properties of a typical graph with a (large) given number of vertices and a given number of edges. However, working with this model is not very convenient due to dependencies (e.g., if some part of the graph contains many edges, then it is less likely that other parts will contain many edges as well). The next model we define, overcomes this problem.

Definition 1.2. *The binomial random graph model $G(n,p)$ is the probability space $(\Omega, \mathbb{P})$ where Ω consists of all labeled graphs with the vertex set $[n]$ and $\mathbb{P}(G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}$ for every $G \in \Omega$.*

Remark 1.3. **(a)** *A standard and very convenient way of generating a graph $G \sim G(n,p)$ is the following. For every $1 \leq i < j \leq n$, flip a coin whose outcome is heads with probability p, where all coin flips are mutually independent. The edge set of G consists of all pairs $1 \leq i < j \leq n$ for which the outcome of the coin flip was heads.*

(b) *$\mathbb{P}(G)$ depends solely on the number of edges of G (and not, say, on its structure). In particular, for every $0 \leq m \leq \binom{n}{2}$, the probability function $\mathbb{P}$ is uniform over all graphs with vertex set $[n]$ and with m edges.*

(c) *For $m \approx \binom{n}{2}p$ we expect $G(n,p)$ and $G(n,m)$ to behave similarly. While the two random graph models do differ, for our purposes in this course they are essentially the same (this can be made rigorous, but we will not do so in this course). In particular, it can be proved that*

$$\lim_{n\to\infty} \mathbb{P}(G(n,m) \in Q) = 1 \Longleftrightarrow \lim_{n\to\infty} \mathbb{P}(G(n,p) \in Q) = 1$$

for every monotone increasing graph property Q[1]*. This relation allows us to essentially forget about $G(n,m)$ and work solely with the much more convenient model $G(n,p)$.*

(d) *The probability space $G(n,1/2)$ is uniform over all graphs with vertex set $[n]$. Hence, for any graph property Q, proving that $\lim_{n\to\infty} \mathbb{P}(G(n,1/2) \in Q) = 1$ shows that "almost" every large graph satisfies Q.*

Example 1: We will first show that "almost" all large graphs have diameter at most 2 (i.e., every two vertices are connected by an edge or have a common neighbour). Let us denote this property by D_2. By Remark 1.3(d), it suffices to prove that $\lim_{n\to\infty} \mathbb{P}(G(n,1/2) \notin D_2) = 0$. Let $G \sim G(n,1/2)$ and for every $1 \le i < j \le n$, let A_{ij} denote the event "$dist_G(i,j) > 2$". For every $z \in V(G) \setminus \{i,j\}$, the probability that z is a common neighbour of i and j is $1/2 \cdot 1/2 = 1/4$ and thus the probability that z is not a common neighbour of i and j is $1 - 1/4 = 3/4$. The probability that i and j have no common neighbour in G is thus $(3/4)^{n-2}$; in particular, $\mathbb{P}(A_{ij}) \le (3/4)^{n-2}$ for every $1 \le i < j \le n$. We conclude that

$$\mathbb{P}(G \notin D_2) = \mathbb{P}\left(\bigcup_{1\le i<j\le n} A_{ij}\right) \le \sum_{1\le i<j\le n} \mathbb{P}(A_{ij}) \le \binom{n}{2}(3/4)^{n-2} = o(1).$$

It now follows from Remark 1.3(c) that $\lim_{n\to\infty} \mathbb{P}(G(n,m) \notin D_2) = 0$, for $m = \lceil \frac{n(n-1)}{4} \rceil$. One can of course prove this directly, but this turns out to be somewhat harder.

1.2 Staged exposure and monotonicity in $G(n,p)$

It is sometimes convenient to build a random graph in several stages. The following result shows how this can be done.

Proposition 1.4 (Staged exposure)**.** *Let k be a positive integer and suppose that $0 \le p, p_1, \ldots, p_k \le 1$ satisfy $1 - p = \prod_{i=1}^{k}(1 - p_i)$. Then $G(n,p)$ and $\bigcup_{i=1}^{k} G(n,p_i)$ form the same probability space.*

Proof. Fix some $1 \le i < j \le n$. Observe that in both $G(n,p)$ and $\bigcup_{i=1}^{k} G(n,p_i)$, the event that ij is an edge is mutually independent of the corresponding events for all other pairs of vertices. Moreover, the probability that ij is not an edge of $G(n,p)$ is $1-p$ and the probability that ij is not an edge of $\bigcup_{i=1}^{k} G(n,p_i)$ is $\prod_{i=1}^{k}(1-p_i) = 1-p$. $\square$

Proposition 1.4 is a simple but useful tool. In particular, it can be used to prove the monotonicity of $G(n,p)$ in p.

Proposition 1.5 (Monotonicity)**.** *Let Q be a monotone increasing graph property and let $0 \le p_1 \le p_2 \le 1$. Then $\mathbb{P}(G(n,p_1) \in Q) \le \mathbb{P}(G(n,p_2) \in Q)$.*

[1] A *graph property* is simply a set of graphs. If Q is a graph property and $G \in Q$, then we say that G *satisfies* Q. A graph property Q is said to be *monotone increasing* if it is closed under the addition of edges, that is, if $G \in Q$ and $H \supseteq G$ imply that $H \in Q$.

Proof. Let $p_0 = 1 - \frac{1-p_2}{1-p_1}$ and observe that $0 \leq p_0 \leq 1$ and $1 - p_2 = (1-p_0)(1-p_1)$. For every $i \in \{0,1,2\}$ let $G_i \sim G(n, p_i)$. Clearly $G_1 \subseteq G_0 \cup G_1$. Since Q is monotone increasing, it thus follows that

$$\mathbb{P}(G_1 \in Q) \leq \mathbb{P}(G_0 \cup G_1 \in Q) = \mathbb{P}(G_2 \in Q),$$

where the equality holds by Proposition 1.4. □

1.3 Graphs with large girth and large chromatic number

Before stating and proving the result of this section, we need several definitions and simple claims.

Definition 1.6. *The* girth *of a graph G, which we denote by $g(G)$, is the length (i.e., the number of edges) of a shortest cycle in G. If G is a forest, then we set $g(G) = \infty$.*

Definition 1.7. *Let $G = (V, E)$ be a graph and let $S \subseteq V$. The set S is called* independent *if it spans no edges, that is, if $xy \notin E$ for every $x, y \in S$. The* independence number *of G, denoted by $\alpha(G)$, is the size of a largest independent set in G.*

Definition 1.8. *The chromatic number of a graph $G = (V, E)$, denoted by $\chi(G)$, is the smallest positive integer k for which there exists a function $c : V \to \{1, \ldots, k\}$ (customarily known as a proper k-colouring of G) such that $c(u) \neq c(v)$ whenever $uv \in E$.*

Observe that if a graph G contains a copy of K_d, then $\chi(G) \geq d$. It is natural to ask whether the converse implication holds as well, that is, whether $\chi(G) \geq d$ implies $K_d \subseteq G$. This is trivially true for $d \in \{1, 2\}$, but, as the following result demonstrates, is very far from being true in general.

Theorem 1.9 (Erdős 1959)**.** *For all integers k and ℓ there exists a graph G such that $g(G) > \ell$ and $\chi(G) > k$.*

Before proving Theorem 1.9, we state and prove a simple lower bound on the chromatic number of a graph.

Claim 1.10. *Let G be a graph on n vertices. Then $\chi(G) \geq n/\alpha(G)$.*

Proof. Let $k = \chi(G)$ and let $c : V(G) \to \{1, \ldots, k\}$ be a proper k-colouring of G. For every $1 \leq i \leq k$ let $A_i = \{u \in V(G) : c(u) = i\}$. Observe that $A_1 \cup \ldots \cup A_k$ is a partition of $V(G)$ and that A_i is independent for every $1 \leq i \leq k$. Hence

$$n = |V(G)| = \sum_{i=1}^{k} |A_i| \leq k \cdot \alpha(G).$$

We conclude that $\chi(G) = k \geq n/\alpha(G)$ as claimed. □

Proof of Theorem 1.9. Let $\theta = \frac{1}{2\ell}$ and let $G \sim G(n,p)$, where n is sufficiently large and $p = n^{\theta-1}$. We call a cycle of G *short* if its length is at most ℓ. Let X count the number of short cycles in G. Then

$$\mathbb{E}(X) = \sum_{i=3}^{\ell} \frac{n(n-1)\cdot \ldots \cdot (n-i+1)}{2i} \cdot p^i \le \sum_{i=3}^{\ell} n^i p^i = \sum_{i=3}^{\ell} n^{\theta i} = O\left(n^{\theta \ell}\right) = o(n),$$

where the last equality holds since $\theta < 1/\ell$. It then follows by Markov's inequality that

$$\mathbb{P}\left(X \ge n/2\right) = o(1). \tag{1}$$

Now, set $t = \lceil 3 \ln n/p \rceil$ and observe that

$$\begin{aligned}\mathbb{P}\left(\alpha(G) \ge t\right) &\le \binom{n}{t}(1-p)^{\binom{t}{2}} \le n^t e^{-pt(t-1)/2} \le \left(n \cdot e^{-p(t-1)/2}\right)^t \\ &\le \left(n \cdot e^{-1.4 \ln n}\right)^t = o(1),\end{aligned} \tag{2}$$

Combining (1) and (2) implies that

$$\mathbb{P}\left(X \ge n/2 \text{ or } \alpha(G) \ge t\right) = o(1).$$

In particular, there exists a graph H on n vertices which contains at most $n/2$ short cycles such that $\alpha(H) \le t$. Deleting one arbitrary vertex from every short cycle of H yields a graph H' on at least $n/2$ vertices such that $g(H') > \ell$ and $\alpha(H') \le t$. Using Claim 1.10 we conclude that

$$\chi(H') \ge \frac{|V(H')|}{\alpha(H')} \ge \frac{n/2}{\lceil 3 \ln n/p \rceil} \ge n^{\theta/2} > k,$$

where the last two inequalities hold for sufficiently large n. $\square$