Chapter 8. Random Graphs

The main reference for this section is Diestel Graph Theory, Chapter 11.

$|\Omega_n| = 2^{\binom{n}{2}}$.

Let Ω_n be the set of all graphs with vertex set $\{1, 2, \ldots, n\}$. We work in the uniform probability space on Ω_n.

$\Pr(G_0) = 2^{-\binom{n}{2}}$

We can choose a random element from Ω_n by flipping a fair coin, independently for each pair of distinct vertices $\{i, j\}$, and letting $ij \in E$ if and only if the coin comes up heads.

This is the uniform model of random graphs.

Recall: the expected number of edges in this model is $\frac{1}{2}\binom{n}{2}$ (Problem Sheet 3).

What if your coin comes up heads with probability p, for some $p \in [0,1]$ is fixed? Again, for each pair of distinct vertices $\{i,j\}$, independently, flip the coin and let $ij \in E$ if and only if the coin comes up heads.

Then $\Pr(ij \in E) = p$, independently for each $i \neq j$.

This gives a random graph model called the *binomial model* denoted $G(n,p)$, introduced by Gilbert, 1959.

Note: $G(n, \frac{1}{2})$ is the uniform model.

We write $G \in G(n,p)$ to mean that G is a random graph chosen from the binomial model.

For a fixed $G_0 \in \Omega_n$, the probability that the random graph G equals G_0 is

$$\Pr(G = G_0) = p^{|E(G_0)|}\,(1-p)^{\binom{n}{2}-|E(G_0)|}$$

Using independence !!

which depends only on $|E(G_0)|$.
This follows from Problem Sheet 3, Question 3: check!

For $G \in G(n,p)$, the expected number of edges of G is $p\binom{n}{2}$.
(Make sure you can prove this.)

$$\Pr(G_0 \text{ has } r \text{ edges}) = \binom{\binom{n}{2}}{r} p^r (1-p)^{\binom{n}{2}-r},$$

Binomial distrib $\binom{n}{2}$ trials & success prob p

$\ln n = \log_e(n)$ natural log

For fixed $p \in [0, 1]$, we have a sequence of probability spaces,

$$(G(n, p))_{n \in \mathbb{Z}^+}.$$

We can also let p be a function of n, where $p(n) \in [0, 1]$ for all $n \in \mathbb{Z}^+$. This gives the sequence of probability spaces

$$(G(n, p(n)))_{n \in \mathbb{Z}^+}.$$

We just write $G(n, p)$ for this

Example: let $p = 1/n$. Then the expected number of edges of $G \in G(n, \frac{1}{n})$ is:

$$\mathbb{E}(\#\text{edges}) = p\binom{n}{2} = \frac{1}{\cancel{n}} \cdot \frac{\cancel{n}(n-1)}{2} = \frac{n-1}{2}$$

$(\varepsilon > 0)$

FACT: If $p = (1-\varepsilon)\frac{\ln n}{n}$ then for $G \in G(n,p)$, $\Pr(G \text{ has an isolated vertex}) \to 1$

If $p = (1+\varepsilon)\frac{\ln n}{n}$ then $\Pr(G \text{ is connected}) \to 1$

Phase transition!!

#vertices in a max clique ≡ largest complete subgraph

Recall that $\omega(G)$ is the clique number of G, and $\alpha(G)$ is the independence number of G.

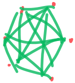
#vertices in a maximum independent set

Lemma 11.1.2

Let $G \in G(n,p)$. Then for any integer $k \geq 2$,

choice of $S \subseteq [n]$ with $|S| = k$

S



$$\Pr(\omega(G) \geq k) \leq \binom{n}{k} p^{\binom{k}{2}},$$

$$\Pr(\alpha(G) \geq k) \leq \binom{n}{k} (1-p)^{\binom{k}{2}}.$$

(union bound = Boole's inequality)

Proof. Let $G \in G(n,p)$. If G has a clique of order $\geq k$ then G has a clique of order k. For a set S of k vertices, let A_S be the event "$G[S]$ is a clique". Then $\Pr[A_S] = p^{\binom{k}{2}}$, using independence, since $\binom{k}{2}$ edges within S must be present. Hence

$$\Pr(\omega(G) \geq k) = \Pr\Big(\bigcup_{|S|=k} A_S\Big) \leq \sum_{|S|=k} \Pr(A_S) \quad \text{(by the union bound)}$$

$$= \binom{n}{k} p^{\binom{k}{2}}, \text{ as required.}$$

Proof for $\alpha(G)$ is an exercise: see Problem Sheet 8.

$\square$

Proof for $\alpha(G)$ is an exercise (Problem Sheet 8). $\square$

For $a \in \mathbb{R}$ and $r \in \mathbb{N}$, let

sequences of r distinct elements from $[a] = \{1, 2, \ldots, a\}$

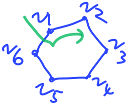
$$(a)_r = a(a-1)\cdots(a-r+1)$$

denote the falling factorial.

Lemma 11.1.4

Let $k \geq 3$ be an integer. The expected number of k-cycles in $G \in G(n,p)$ is

$$\frac{(n)_k}{2k}\, p^k.$$



Proof. Let X be the number of k-cycles in $G \in G(n,p)$.

Given a sequence $(v_1, v_2, \ldots, v_k)$ of k distinct vertices, the probability that this sequence describes a walk around a k-cycle is

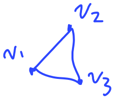
$$\Pr(v_1v_2, v_2v_3, \ldots, v_{k-1}v_k, v_kv_1 \in E(G)) = p^k,$$ using independence.

There are $(n)_k$ ways to choose this sequence of k distinct vertices.

Each cycle in G corresponds to exactly $2k$ such sequences corresponding to the choice of start-vertex and direction.

Hence, by linearity of expection,

$$\mathbb{E}X = \frac{(n)_k}{2k} p^k, \quad \text{as claimed.}$$



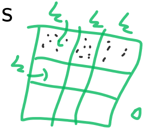
(v_1, v_2, v_3)
(v_1, v_3, v_2)
(v_2, v_1, v_3)
(v_2, v_3, v_1)
(v_3, v_1, v_2)
(v_3, v_2, v_1)

$\square$

Exercise: find the expected number of k-paths in $G \in G(n,p)$.
(See Problem Sheet 8.)

For a given graph property $\mathcal{P}$, we can ask how $\Pr(G \in \mathcal{P})$ behaves for $G \in G(n,p)$ as $n \to \infty$.

If $\Pr(G \in \mathcal{P}) \to 1$ as $n \to \infty$ then we say that $G \in \mathcal{P}$ holds asymptotically almost surely, abbreviated to "a.a.s.". (Note, Diestel says "almost every $G \in G(n,p)$ has $\mathcal{P}$".)



Proposition 11.3.1
For fixed $p \in (0,1)$ and every graph H, a.a.s. $G \in G(n,p)$ has an induced subgraph which is isomorphic to H.

Proof. Let $k = |V(H)|$. Suppose that $n \geq k$ and let $\mathcal{U} \subseteq \{1,2,\dots,n\} = [n]$ be a fixed set of k vertices. The probability that $G[\mathcal{U}] \cong H$ is some fixed constant $r \in (0,1)$ which depends only on H and p, but not on n. Now we can find $\lfloor n/k \rfloor$ disjoint sets of k vertices, $U_1, \dots, U_{\lfloor n/k \rfloor}$, within $V(G) = [n]$.

The probability that none of $U_1, \ldots, U_{\lfloor n/k \rfloor}$ induces a copy of H is

$$(1-r)^{\lfloor n/k \rfloor},$$

since the U_j are disjoint and hence the events $G[U_j] \not\cong H$ are independent of each other ($j = 1, \ldots, \lfloor n/k \rfloor$)

But $(1-r)^{\lfloor n/k \rfloor} \to 0$ as $n \to \infty$, since $1-r \in (0,1)$ and $\lfloor n/k \rfloor \to \infty$ as $n \to \infty$.

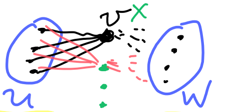
Hence a.a.s., one of $U_1, \ldots, U_{\lfloor n/k \rfloor}$ induces a copy of H

$\square$

Note: if p is fixed then $G(n,p)$ has $\sim p\binom{n}{2} = \Theta(n^2)$ (quadratic)

So each vertex has $\sim pn$ edges

This is a dense graph!



Given $i, j \in \mathbb{N}$, let $\mathcal{P}_{ij}$ be the property that given any disjoint vertex sets U, W with $|U| \leq i$ and $|W| \leq j$, the graph contains a vertex $v \notin U \cup W$ that is adjacent to all vertices in U but to none in W.

Lemma 11.3.2

For every constant $p \in (0,1)$ and all $i, j \in \mathbb{N}$, let $G \in G(n,p)$. Then a.a.s. $G \in \mathcal{P}_{ij}$.

Proof. Assume that $n \geq i+j+1$ For fixed disjoint set $U, W \subseteq [n]$ and $v \in [n] - (U \cup W)$ the probability that v is adjacent to all vertices of U and to no vertices of W is $p^{|U|}(1-p)^{|W|} \geq p^{i}(1-p)^{j}$, using independence

To simplify notation we write $q = 1-p$.

Hence the probability that no such v exists for the given sets $\mathcal{U}$, W is

$$\left(1-p^{|\mathcal{U}|}q^{|W|}\right)^{n-|\mathcal{U}|-|W|}$$

, since these events are independent for distinct $v \notin \mathcal{U} \cup W$

[No edge/non-edge choices are considered in more than one of these events]

Now $\left(1-p^{|\mathcal{U}|}q^{|W|}\right)^{n-|\mathcal{U}|-|W|} \leq \left(1-p^{i}q^{j}\right)^{n-|\mathcal{U}|-|W|}$

$$\leq \left(1-p^{i}q^{j}\right)^{n-i-j}$$

There are at most n^{i+j+2} pairs of disjoint sets $\mathcal{U}$, W with $|\mathcal{U}| \leq i$ and $|W| \leq j$, as

$$\sum_{s=0}^{i}\binom{n}{s} \leq \sum_{s=0}^{i} n^s = \frac{n^{i+1}-1}{n-1} \leq n^{i+1}$$

[also $2n^i$ works I think: check!]

& similarly for W.

Hence the probability that some U, W has no suitable v is at most

$$n^{i+j+2}\left(1-p^i q^j\right)^{n-i-j} \longrightarrow 0 \text{ as } n \to \infty,$$

since $1-p^i q^j \in (0,1)$.

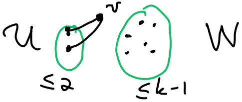
[check! eg using L'Hôpital's rule or taking logs.]

Hence a.a.s. P_{ij} holds, as required. $\square$

$\mathcal{P}_{ij}$: $\forall\, U, W \subseteq [n]$ if $|U| \le i$ and $|W| \le j$ then $\exists\, v \in [n] \setminus (U \cup W)$

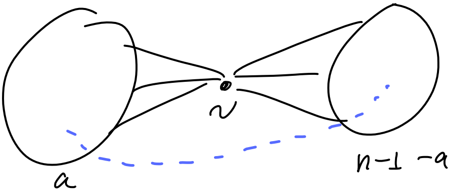
$(- - - -)$

$|[n]-W| \geqslant n-(k-1) = n-k+1$
$\geqslant (k+2)-k+1$
$= 3.$



Corollary 11.3.3
For every constant $p \in (0,1)$ and all $k \in \mathbb{N}$, a.a.s. $G \in G(n,p)$ is k-connected.

Proof. By Lemma 11.3.2, it is enough to show that every graph in $\mathcal{P}_{2,k-1}$ is k-connected when n is sufficiently large. Assume that $n \geqslant k+2$. [One more than is needed for k-connectivity.] Let W be any set of at most $k-1$ vertices. We want to prove that $G-W$ is connected. So let x, y be distinct vertices in $[n]-W$ and define $\mathcal{U} = \{x,y\}$. By definition of $\mathcal{P}_{2,k-1}$ there is a vertex v in $[n]-(\mathcal{U} \cup W)$ such that v is adjacent to both x and y. Hence xvy is a path between x and y in $G-W$, proving that $G-W$ is connected. $\square$



$(1-p)^{a(n-1-a)}$

very unlikely.

$\mathbb{E}(\#\text{cut vertices}) \leq n \times n \times (1-p)^{a(n-1-a)}$

Super-small!

$\square$

$\left[\chi(G) \geq \frac{n}{\alpha(G)}\right]$

Here log means the natural logarithm.

surprising high ??

Proposition 11.3.4

For every constant $p \in (0,1)$ and all $\varepsilon > 0$, a.a.s. $G \in G(n,p)$ satisfies

$$\chi(G) \geq \frac{\log(1/q)\, n}{(2+\varepsilon)\log n}, \quad = \Theta\left(\frac{n}{\log n}\right).$$

where $q = 1 - p$.

Proof. Let a be any fixed integer, $2 \leq a \leq n$.
Then by Lemma 11.1.2,

$$\Pr(\alpha(G) \geq a) \leq \binom{n}{a}(1-p)^{\binom{a}{2}} \leq n^a (1-p)^{\binom{a}{2}}$$

$$= q^{a\frac{\log n}{\log q} + \frac{a(a-1)}{2}}$$

$$= q^{\frac{a}{2}\left[\frac{2\log n}{\log q} + a - 1\right]}$$

$$= q^{\frac{a}{2}\left[a - 1 - \frac{2\log n}{\log(1/q)}\right]}$$

Set $a = \left\lceil \frac{(2+\varepsilon)\ln n}{\ln(1/q)} \right\rceil$. Then

$$\lim_{n\to\infty} \frac{a}{2}\left(a-1-\underbrace{\frac{2\log n}{\log(1/q)}}\right) \geq \lim_{n\to\infty} \frac{(2+\varepsilon)\log n}{2\log(1/q)}\left[\frac{\varepsilon \log n}{\log(1/q)} - 1\right]$$

$$= \infty$$

Hence $\Pr(\alpha(G) \geq a) \to 0$ as $n \to \infty$, since $q \in (0,1)$.

This shows that a a.s. $G \in G(n,p)$ has no independent set of order $\left\lceil \frac{(2+\varepsilon)\log n}{\log(1/q)} \right\rceil$, and hence a.a.s

$$\alpha(G) < \frac{(2+\varepsilon)\log n}{\log(1/q)}.$$

Therefore a.a.s. for $G \in G(n,p)$, $\Big\}$ $\chi(G) \geq \frac{n}{\alpha(G)} > \frac{\log(1/q)\, n}{(2+\varepsilon)\log n}$. $\square$

Recall that the girth of a graph is the length of its smallest cycle.

Now we work towards the proof of Erdős' result from 1959: there exist graphs with arbitrarily high girth and arbitrarily high chromatic number.

constant

constant

Lemma 11.2.1

Let k be a positive integer and let $p = p(n)$ be a function of n such that $p(n) \in [0, 1]$ and

$$p(n) \geq \frac{6k \log n}{n}$$

for sufficiently large n. Then for $G \in G(n, p)$, a.a.s.

$$\alpha(G) < \frac{n}{2k}.$$

Proof. Let $n, r \in \mathbb{Z}$, $n \geqslant r \geqslant 2$. By Lemma 11.1.2, for $G \in G(n,p)$ we have

$$\Pr(\alpha(G) \geqslant r) \leq \binom{n}{r}(1-p)^{\binom{r}{2}} \leq n^r (1-p)^{\binom{r}{2}}$$

$$= \left[n(1-p)^{\frac{r-1}{2}}\right]^r$$

$$\leq \left[n e^{-p(r-1)/2}\right]^r$$

Since $1-p \leq e^{-p}$.

If $p \geq \frac{6k\log n}{n}$ and $r \geq \frac{n}{2k}$ then

$$\frac{pr}{2} \geq \frac{\overset{3}{\cancel{6k}}\log n}{2\cancel{n}} \cdot \frac{\cancel{n}}{\cancel{2k}} = \frac{3\log n}{2}$$

$$\begin{aligned}
ne^{-p(r-1)/2} &= ne^{-pr/2 + p/2} \\
&\leq ne^{-\frac{3}{2}\log n + p/2} \\
&\leq ne^{-3/2\log n + \frac{1}{2}} \quad \text{since } p \leq 1 \\
&= n \cdot n^{-3/2} \cdot e^{1/2} \\
&= \sqrt{\frac{e}{n}} \to 0 \quad \text{as } n \to \infty
\end{aligned}$$

Since $p = p(n) \geq \frac{6k\log n}{n}$ for sufficiently large n, we take $r = \lceil n/2k \rceil$ to conclude that

$$\lim_{n\to\infty} \Pr\left(\alpha(G) \geq \frac{n}{2k}\right) = \lim_{n\to\infty} \Pr(\alpha(G) \geq r) = 0.$$

$\square$

Recall Markov's inequality: if $X : \Omega \to \mathbb{N}$ is a nonnegative integer-valued random variable on a set Ω, and $k > 0$, then

$$\Pr(X \geq k) \leq \frac{\mathbb{E}X}{k}.$$

We can now prove Erdős' famous result.

Theorem 11.2.2. (Erdős, 1959)
For every integer $k \geq 3$ there exists a graph H with girth $g(H) > k$ and chromatic number $\chi(H) > k$.

Proof.

Fix ε with $0 < \varepsilon < \frac{1}{k}$ and let $p = p(n) = n^{\varepsilon - 1}$.

$(np = n^{\varepsilon} \to \infty)$

Let $X(G)$ be the number of cycles in $G \in G(n,p)$ with length $\leq k$.

By Lemma 11.1.4 and linearity of expectation,

$$\mathbb{E}X = \sum_{i=3}^{k} \frac{(n)_i}{2i} p^i \leq \frac{1}{2} \sum_{i=3}^{k} (np)^i \leq \frac{k-2}{2} (np)^k,$$

(# terms)

as $np = n^{\varepsilon} > 1$.

Using Markov's inequality,

$$\Pr\left(X \geq \frac{n}{2}\right) \leq \frac{\mathbb{E}X}{n/2} \leq (k-2)\, n^{k-1} p^k$$

$$= (k-2)\, n^{k-1}\, n^{(\varepsilon-1)k}$$

$$= (k-2)\, n^{k\varepsilon - 1}.$$

Note $k\varepsilon < 1$ by choice of ε. Hence

$$\lim_{n\to\infty} \Pr\left(X \geq \frac{n}{2}\right) = 0. \quad (*)$$

[That is, a.a.s. $X(G) < \frac{n}{2}$.]

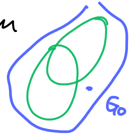
Note also that $p = n^{\varepsilon-1} \geq \frac{6k\log n}{n}$ for large enough n, as k is constant.

Hence by Lemma 11.2.1 and $(*)$, we can choose n large enough so that

$$\Pr\left(X \geq \frac{n}{2}\right) < \frac{1}{2} \quad \text{and} \quad \Pr\left(\alpha(G) \geq \frac{n}{2k}\right) < \frac{1}{2}.$$

This shows that for some fixed graph G_0 on n vertices we have $\alpha(G_0) < \frac{n}{2k}$ and G_0 has fewer than $\frac{n}{2}$ cycles of length $\leq k$.

Construct H from G_0 by deleting one vertex from every cycle in G_0 of length $\leq k$.



Then $|H| \geq \frac{n}{2}$ and by construction, $g(H) > k$.

Also $\alpha(H) \leq \alpha(G_0) < \frac{n}{2k}$

Since every independent set in H is also an independent set in G_0.

Therefore

$$\chi(H) \geq \frac{|H|}{\alpha(H)} \geq \frac{n/2}{\alpha(H)} > \frac{n/2}{n/(2k)} = k,$$

completing the proof. $\square$

We delete strictly less than $n/2$ vertices from G_0 to get H and $|G_0| = n$, so $|H| > n - \left(\frac{n}{2}\right) = \frac{n}{2}$

[End of Chapter 8. Try Problem Sheet 8.]

[End of MATH5425 Graph Theory lectures!!!]