5. Random Graphs

- Recall: $\mathcal{G}_n$ = family of all (labelled, undirected) n-node graphs
  - $|\mathcal{G}_n| = 2^N$, where $N = \binom{n}{2}$
- The "Erdős - Rényi" ensembles of random graphs:
  - $\mathcal{G}(n,p)$ = all $G \in \mathcal{G}_n$ taken so that each edge has occurrence prob p, indep. of the other edges. [Gilbert 1959]
  - $\mathcal{G}(n,M)$ = all $G \in \mathcal{G}_n$ with exactly $M \le N$ edges, taken with uniform probability. [Erdős & Rényi 1960]
- Since for large n and given p, the number of edges in a $\mathcal{G}(n,p)$ random graph is heavily concentrated around $M = Np$, the ensembles $\mathcal{G}(n,p)$ and $\mathcal{G}(n,M)$ for $M = Np$ are "very similar" (?). Ensemble $\mathcal{G}(n,p)$ is easier to analyse and thus usually considered.
- Denote $G \models A$ ~ "graph G has property A".
- Random graphs $G \in \mathcal{G}(n,p)$ have property A <u>asymptotically almost surely</u> (a.a.s.) or <u>almost everywhere</u> (a.e.) if

$$\Pr_{n,p}(G \models A) \longrightarrow 1 \quad \text{as } n \longrightarrow \infty.$$

- Denote in general $q = 1-p$.

- For constant p, $0<p<1$, the features of $\mathcal{G}(n,p)$ graphs are quite regular, as illustrated by the following results:

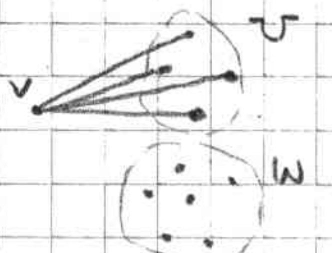
- <u>Theorem</u> 5.1 Let H be a fixed graph and $0<p<1$. Then a.e. $G \in \mathcal{G}(n,p)$ contains an induced copy of H.

<u>Proof.</u> Let $k = |H|$ = number of vertices in H. Then a graph G with $n = |G| \geq k$ vertices can be partitioned into $\lfloor n/k \rfloor$ disjoint sets of k vertices (with some left over). For each of these sets, the prob. that it forms an induced copy of H is some $r>0$.

Thus, the prob. that none of these sets forms an induced copy of H is

$$(1-r)^{\lfloor n/k \rfloor} \to 0 \quad \text{as } n \to \infty. \quad \square$$

- Let $k, l \in \mathbb{N}$. Say that a graph $G = (V, E)$ has property Q_{kl}, if $\forall U, W \in V$, $|U| \leq k$, $|W| \leq l$, $U \cap W = \emptyset$, G contains a vertex $v \in V \setminus (U \cup W)$ that is adjacent to all $u \in U$ and to no $w \in W$.



- <u>Lemma</u> 5.2 For every constant p, $0<p<1$, and all $k, l \in \mathbb{N}$, a.e. $G \in \mathcal{G}(n,p)$ has property Q_{kl}

<u>Proof.</u> For fixed $U, W, v \in V \setminus (U \cup W)$, the prob. that the condition is satisfied is

$$p^{|U|} q^{|W|} \geq p^k q^l.$$

The events are independent for so the prob. that <u>no</u> appropriate v exists for <u>given</u> U, W is

$$(1-p^{|U|} q^{|W|})^{n-|U|-|W|} \leq (1-p^k q^l)^{n-k-l}.$$

There are at most n^{k+l} (U,W)-pairs to be considered, so the prob. that some pair has no good v is bounded by:

$$n^{k+l}(\underbrace{1-p^k q^l}_{<1})^{n-k-l} \longrightarrow 0 \quad \text{as} \quad n \to \infty.$$

Thus in a.e. $G \in \mathcal{G}(n,p)$ all (U,W) have some good v. $\square$

- Corollary 5.3 Let p, $0<p<1$, be constant. Then:

(i) for any constant k, a.e. $G \in \mathcal{G}(n,p)$ has min degree $\geq k$;

(ii) a.e. $G \in \mathcal{G}(n,p)$ has diameter 2;

(iii) for any constant k, a.e. $G \in \mathcal{G}(n,p)$ is k-connected.

Proof (i), (ii) Immediate from Lemma 5.2.

(iii) In a.e. $G \in \mathcal{G}(n,p)$, no two vertices u_1, u_2 can be separated by a cutset of size $k-1$, because one may choose in Lemma 5.2 $U = \{u_1, u_2\}$, $W = \{w_1, \dots, w_{k-1}\}$ for arbitrary $w_1, \dots, w_{k-1}$, and obtain a path $u_1 - v - u_2$ connecting u_1, u_2 and avoiding $w_1, \dots, w_{k-1}$. $\square$

- More generally, it is possible to prove (Alon & Spencer Sec. 10.7):

Theorem 5.4 Let A be any first-order property of graphs (i.e. A is expressible in terms of quantification over vertices, edge relation $E(u,v)$, and identity). Then for any constant p, $0<p<1$, either $G \models A$ a.e. or $G \models \text{not } A$ a.e. $\square$

- This is called a zero-one law for random graphs (w.r.t. first-order properties and constant p).

Some characteristic numbers

- Given a graph $G = (V, E)$, denote:

  $\alpha(G)$ = size of largest indep. set in G ("independence number")
  $\omega(G)$ = " " " clique " ("clique number")
  $\chi(G)$ = smallest number of colours needed for colouring vertices of G so that no two adjacent ones get the same colour ("chromatic number")

- Simple relations:

  - $\omega(G) = \alpha(\bar{G})$, where $\bar{G} = (V, (V\times V)\setminus E)$

  - $\chi(G) \geq \frac{|V|}{\alpha(G)}$ [Since each colour-class of vertices is an independent set.]

- To analyse these numbers we shall use the first- and second-moment methods.

  Recall: X an integer-valued random variable (actually dependent on n, so $X^{(0)}, X^{(1)}, X^{(2)}, \ldots$)

  First-moment method: If $E[X] \to 0$, then $X = 0$ a.e.

  Second-moment method: If $Var[X] = o(E[X]^2)$, then $X > 0$ a.e.
  – In fact, $X \sim E[X]$ a.e.

- Further refinements of the 2nd-moment method:

  - Suppose $X = X_1 + \cdots + X_m$, where X_i is an indicator for event A_i. Write $i \sim j$ if $i \neq j$ and $A_i \not\perp A_j$.
  Denote:

$$\Delta = \sum_{i \sim j} Pr(A_i \,\&\, A_j)$$

- Then for $i \sim j$:

$$\mathrm{Cov}[X_i, X_j] = E[X_i X_j] - E[X_i]E[X_j] \le E[X_i X_j] = \Pr(A_i \,\&\, A_j)$$

and for $i \neq j$, $i \not\sim j$: $\mathrm{Cov}[X_i, X_j] = 0$. Thus

$$\begin{aligned} \mathrm{Var}[X] &= \sum_i \mathrm{Var}[X_i] + \sum_{i \neq j} \mathrm{Cov}[X_i, X_j] \\ &\le \sum_i E[X_i] + \sum_{i \sim j} \Pr(A_i \,\&\, A_j) \\ &= E[X] + \Delta. \end{aligned}$$

Lemma Δ. If $E[X] \to \infty$ and $\Delta = o(E[X]^2)$, then $X > 0$ a.e. In fact, $X \sim E[X]$ a.e. $\square$

- Say that $X_1, \ldots, X_m$ are symmetric if for every $i \neq j$ there is an automorphism of the probability space that sends event A_i to event A_j. (A_i and A_j are e.g. "the same except for numbering of components".)

Then one can write:

$$\Delta = \sum_{i \sim j} \Pr(A_i \,\&\, A_j) = \sum_i \Pr(A_i) \underbrace{\sum_{j \sim i} \Pr(A_j | A_i)}_{\Delta^*},$$

where the sum Δ^* is independent of i. Thus in this case:

$$\Delta = \Delta^* \sum_i \Pr(A_i) = \Delta^* E[X].$$

Lemma Δ^*. If $E[X] \to \infty$ and $\Delta^* = o(E[X])$, then $X > 0$ a.e. In fact, $X \sim E[X]$ a.e. $\square$

- Consider then the clique number $\omega(G)$ of a $\mathcal{G}(n,p)$ random graph G.

- The exp. number of k-cliques in G is

$$f(k) = \binom{n}{k} p^{\binom{k}{2}}.$$

Thus if $k = k(n,p)$ is such that $f(k) \to 0$, then G a.a.s has no k-cliques, and so $\omega(G) < k$.

- For simplicity, consider the case $p = \frac{1}{2}$. Then

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}} \sim n^k 2^{-\frac{k^2}{2}},$$

so $\omega(G) \lesssim 2 \log_2 n$. On the other hand, the following holds:

- <u>Theorem</u> 5.5 Let $k = k(n)$ satisfy $k \sim 2 \log_2 n$, $f(k) \to \infty$. Then $\omega(G) \geq k$ for a.e. $G \in \mathcal{G}(n, 1/2)$.

<u>Proof.</u> For each vertex set S, $|S| = k$, define indicator

$$X_S(G) \sim \text{"} S \text{ is a clique in } G \text{"}$$

and consider $X = \sum_{|S|=k} X_S$. Then $\omega(G) \geq k$ iff $X(G) > 0$.

Since we are assuming that $E[X] = f(k) \to \infty$, to apply Lemma Δ^* it suffices to bound the sum

$$\Delta^* = \sum_{T \sim S} \Pr(X_T \mid X_S).$$

Now for fixed S, $T \sim S$ iff $|T \cap S| = i$ for some $2 \leq i \leq k-1$. Thus

$$\Delta^* = \sum_{i=2}^{k-1} \binom{k}{i} \binom{n-k}{k-i} 2^{-\left(\binom{k}{2} - \binom{i}{2}\right)}$$
$$= E[X] \cdot \sum_{i=2}^{k-1} g(i),$$

where

$$g(i) = \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} \cdot 2^{\binom{i}{2}}.$$

At $i=2$ this can be bounded:

$$g(2) = \frac{\binom{k}{2}\binom{n-k}{k-2}}{\binom{n}{k}} \cdot 2 \sim \frac{k^4}{n^2} \sim o(n^{-1})$$

And similarly at $i = k-1$:

$$g(k-1) = \frac{k(n-k)}{\binom{n}{k}} \cdot 2^{\binom{k-1}{2}} =$$

$$\left|\ \binom{k}{2} = \binom{k-1}{1} + \binom{k-1}{2}\right.$$

$$= \frac{k(n-k)}{\binom{n}{k} \cdot 2^{-\binom{k}{2}}} \cdot 2^{-(k-1)}$$

$$\sim \frac{2kn}{E[X]} \cdot 2^{-k} \sim o(n^{-1})$$

$$\left|\ \begin{array}{l} k \sim 2\log_2 n \\ E[X] \to \infty \end{array}\right.$$

Also for the other terms, $2 < i < k-1$,

$$g(i) = o(n^{-1})$$

[proof omitted], and so

$$\Delta^* = o(E(X)).$$

Lemma Δ^* then yields the result $X(G) > 0$, i.e. $\omega(G) \geq k$ a.e.

$\square$

- In fact the clique number for $\mathcal{G}(n,p)$ random graphs is strongly concentrated:

<u>Theorem</u> 5.6 Let $k = k(n,p)$ be the largest integer s.th.

$$\binom{n}{k} p^{\binom{k}{2}} \geq \ln n.$$

Then $\omega(G) \in \{k, k+1\}$ for a.e. $G \in \mathcal{G}(n,p)$. $\square$

- Let us consider then the chromatic number $\chi(G)$. Assume again for simplicity that $p = 1/2$.

<u>Theorem</u> 5.7 For a.e. $G \in \mathcal{G}(n, 1/2)$:

$$\chi(G) \sim \frac{n}{2\log_2 n}.$$

<u>Proof.</u> From Thm 5.6 we know that $\omega(G) \sim 2\log_2 n$ a.e. Since for $p = 1/2$, graphs G and $\bar{G}$ have the same distribution, also $\alpha(G) = \omega(\bar{G}) \sim 2\log_2 n$ a.e. Thus

$$\chi(G) \geq \frac{n}{\alpha(G)} \sim \frac{n}{2\log_2 n} \quad \text{a.e.}$$

For the converse inequality,

[to be filled in later]