Random graphs

Let $0 < p < 1$, it may depend on n. Let $\mathcal{G}(n,p)$ be the probability space of all graphs G with $V(G) = \{1,2,\dots,n\}$, where G has probability

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

Edges exist with probability p, independently of each other.

Definition. We say that a.e. graph in $\mathcal{G}(n,p)$ has property Π if

$$\lim_{n\to\infty} P[G \text{ has } \Pi] = 1$$

Theorem. Let $0 < p < 1$ be fixed. Let k, ℓ be fixed. Then a.e. graph G in $\mathcal{G}(n,p)$ has the property that for all distinct vertices $x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_\ell$ there exists a vertex $v \in V(G) - \{x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_\ell\}$ such that v is adjacent to $x_1, x_2, \dots, x_k$ and not adjacent to $y_1, y_2, \dots, y_\ell$.

Proof. For fixed $x_1, \dots, x_k, y_1, \dots, y_\ell$ let us say that v *works* if $v \in V(G) - \{x_1, \dots, x_k, y_1, \dots, y_\ell\}$, and v is adjacent to $x_1, \dots, x_k$ and not adjacent to $y_1, \dots, y_\ell$.

The probability that a given v works is $p^k(1-p)^\ell$.

The probability it does not work is $1 - p^k(1-p)^\ell$.

The probability that no v works is $(1 - p^k(1-p)^\ell)^{n-k-\ell}$.

The probability that $\exists\, x_1, \dots, x_k, y_1, \dots, y_\ell$ such that no v works is

$$\le n^{k+\ell}(1 - p^k(1-p)^\ell)^{n-k-\ell} \to 0$$

as $n \to \infty$. $\square$

Consequences. When p is fixed:

(1) a.e. graph in $\mathcal{G}(n,p)$ has diameter ≤ 2
(2) for fixed k a.e. graph in $\mathcal{G}(n,p)$ is k-connected.

Proof. If G is not k-connected, then $\exists\, a, b, y_1, \dots, y_\ell$ where $\ell < k$, such that $\nexists\, a - b$ path in $G \backslash \{y_1, y_2, \dots, y_\ell\}$. Apply theorem to $a, b, y_1, y_2, \dots, y_\ell$. $\exists\, v$ adjacent to a, b, not equal to $a, b, y_1, \dots, y_\ell$, a contradiction.

(3) For every fixed graph H, a.e. graph in $\mathcal{G}(n,p)$ has an induced subgraph isomorphic to H.

Proof. Pick $v \in V(H)$. By induction a.e. graph in $\mathcal{G}(n,p)$ has an induced subgraph isomorphic to $H \backslash v$. Let $x_1, \dots, x_k$ be the vertices of $H \backslash v$ adjacent to v and let $y_1, \dots, y_\ell$ be the vertices of $H \backslash v$ not adjacent to v. Apply the theorem.

Markov's inequality. Let X be a non-negative random variable on a probability space with $0 < EX < \infty$. Then for all $t > 0$

$$P[X \geq tEX] \leq {}^{1}/_{t}$$

Proof.

$$EX = \int X \, dP \geq \int\limits_{[X \geq tEX]} X \, dP \geq tEX \int\limits_{[X \geq tEX]} 1 \, dP =$$

$$= tEX \cdot P[X \geq tEX]$$

and so $P[X \geq tEX] \leq {}^{1}/_{t}$.

Corollary. $P[X \geq z] \leq {}^{EX}/_{z}$ for every $z > 0$.

Proof. Let $z = tEX$.

Theorem (Erdös 1959) For every two integers k, l there exists a graph G with $\chi(G) \geq k$ and no cycles of length at most l.

Proof. We will consider $\mathcal{G}(n,p)$, where $p = p(n)$ will be determined later. We first prove that for a suitable choice of p a.e. graph in $\mathcal{G}(n,p)$ has few short cycles. Let $X_i(G)$ be the number of cycles in G of length exactly i. Then

$$X(G) \coloneqq \sum_{i=3}^{l} X_i(G)$$

is the number of cycles in G of length at most l. Now

$$EX = \sum_{i=3}^{l} EX_i = \sum_{i=3}^{l} \sum_{|A|=i} \sum_{\sigma} P[\sigma \text{ determines a cycle}] =$$

$$= \sum_{i=3}^{l} \binom{n}{i} \frac{1}{2} (i-1)!\, p^i \leq \frac{1}{2} \sum_{i=3}^{l} \frac{n^i}{i!} (i-1)!\, p^i \leq \sum_{i=3}^{l} (np)^i$$

Now it seems sensible to choose $p = p(n)$ so that $np = n^{\theta}$ for some $\theta > 0$. Thus the above is equal to

$$\sum_{i=3}^{l} n^{\theta i} \leq l n^{\theta l}$$

for n sufficiently large. If we choose $\theta < 1/l$, then $EX = o(n)$. By Markov's inequality

$$P[X \geq n/2] \leq \frac{2EX}{n} = o(1)$$

and so for all sufficiently large n we have $P[X \geq n/2] < 1/2$. That is,

(1) the probability that $G \in \mathcal{G}(n,p)$ has $\geq n/2$ cycles of length $\leq l$ is strictly less than ½

Next we give a lower bound on $\chi(G)$. For that we will use the inequality $\chi(G)\alpha(G) \geq n$. So we need an upper bound on $\alpha(G)$ and so we need an upper bound on $P[\alpha(G) \geq t]$.

$$P[\alpha(G) \geq t] = P[G \text{ has an independent set of size } t] \leq$$

$$\sum_A P[A \text{ is an independent set in } G] = \binom{n}{t}(1-p)^{\binom{t}{2}} <$$

$$< n^t(1-p)^{\binom{t}{2}} \leq n^t e^{-p\binom{t}{2}} = \left[ne^{-p(t-1)/2}\right]^t$$

where the second inequality uses $1 + x \leq e^x$.

If $t - 1 = \frac{3}{p}\log n$, then

$$\left[ne^{-p(t-1)/2}\right]^t = \left[n \cdot n^{-3/2}\right]^t = o(1)$$

Thus for $t = \frac{3}{p}\log n$ and sufficiently large n

(2) $P[\alpha(G) \geq t] < 1/2$

By (1) and (2) there exists a graph on n vertices with $\leq n/2$ cycles of length $\leq l$ and $\alpha(G) \leq t$. By deleting one vertex from each cycle of length $\leq l$ we arrive at an induced subgraph G' on at least $n/2$ vertices with no cycle of length $\leq l$ and $\alpha(G') \leq t$. Now

$$\chi(G') \geq \frac{|V(G')|}{\alpha(G')} \geq \frac{n/2}{t} \geq \frac{n/2}{\frac{3}{p}\log n} = \frac{n}{6n^{1-\theta}\log n} = \frac{n^{\theta}}{6\log n} \to \infty$$

and so for n sufficiently large we have $\chi(G') \geq k$, as desired.

The emergence of $\boldsymbol{K_4}$ subgraphs

It is reasonable to expect that for small p a.e. graph in $\mathcal{G}(n,p)$ has no K_4 subgraph, while for p close to 1 a.e. graph in $\mathcal{G}(n,p)$ has a K_4 subgraph. It is of interest to see when the change occurs. We will see that there is a sharp the threshold (a "phase transition").

For $A \subseteq \{1,2,\dots,n\}$ with $|A| = 4$ let

$$X_A(G) = \begin{cases} 1 & \text{if } A \text{ is a clique in } G \\ 0 & \text{otherwise} \end{cases}$$

Then the number of K_4 subgraphs in G is

$$X(G) \coloneqq \sum_{|A|=4} X_A(G)$$

We have

$$EX = \sum_{|A|=4} EX_A = \sum_{|A|=4} P[A \text{ is a clique}] = \binom{n}{4} p^6$$

and so by Markov's inequality

$$P[G \text{ has a } K_4 \text{ subgraph}] = P[X \geq 1] \leq EX \leq n^4 p^6$$

So if $n^4p^6 \to 0$, then a.e. graph in $\mathcal{G}(n,p)$ has no K_4 subgraph.

The emergence of K_4 subgraphs

It is reasonable to expect that for small p a.e. graph in $\mathcal{G}(n,p)$ has no K_4 subgraph, while for p close to 1 a.e. graph in $\mathcal{G}(n,p)$ has a K_4 subgraph. It is of interest to see when the change occurs. We will see that there is a sharp the threshold (a “phase transition”).

For $A \subseteq \{1,2,\ldots,n\}$ with $|A| = 4$ let

$$X_A(G) = \begin{cases} 1 & \text{if } A \text{ is a clique in } G \\ 0 & \text{otherwise} \end{cases}$$

Then the number of K_4 subgraphs in G is

$$X(G) := \sum_{|A|=4} X_A(G)$$

We have

$$EX = \sum_{|A|=4} EX_A = \sum_{|A|=4} P[A \text{ is a clique}] = \binom{n}{4} p^6$$

and so by Markov’s inequality

$$P[G \text{ has a } K_4 \text{ subgraph}] = P[X \geq 1] \leq EX \leq n^4 p^6$$

So if $n^4p^6 \to 0$, then a.e. graph in $\mathcal{G}(n,p)$ has no K_4 subgraph.

We will show if $n^4p^6 \to \infty$, then a.e. graph in $g(n,p)$ has a K_4 subgraph. It does not follow from $EX \to \infty$!!

Definition. Let $f, g: \mathbb{N} \to \mathbb{R}$. We define $f \ll g$ to mean that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

Thus if $p \ll \frac{1}{n^{2/3}}$, then a.e. graph in $\mathcal{G}(n,p)$ has no K_4 subgraph and we want to show that if $p \gg \frac{1}{n^{2/3}}$, then a.e. graph in $\mathcal{G}(n,p)$ has a K_4 subgraph.

Lemma (Chebyshev's inequality) If X is a random variable on a probability space, then for all $\epsilon > 0$.

$$P[|X - EX| \geq \epsilon] \leq \frac{\operatorname{var}(X)}{\epsilon^2}$$

Definition.

$$\operatorname{var}(X) := E|X - EX|^2 = E[X^2 - 2XEX + E^2X] =$$

$$= EX^2 - 2EX \cdot EX + E^2X = EX^2 - E^2X.$$

Proof. Apply Markov to $Y := |X - EX|^2$.

THM. (a) If $p \ll \frac{1}{n^{2/3}}$, then a.e. $G \in \mathcal{G}(n,p)$ has no K_4 subgraph.
(b) If $\frac{1}{n^{2/3}} \ll p$, then a.e. $G \in \mathcal{G}(n,p)$ has a K_4 subgraph.

Proof. (a) done

(b) Assume $\frac{1}{n^{2/3}} \ll p$. Recall $EX = \binom{n}{4}p^6 \leq \frac{1}{24}n^4p^6$. Need to show

$$P[X = 0] \to 0 \text{ as } n \to \infty$$

$$P[X=0] \le P[|X-EX| \ge EX] \le \frac{\mathrm{var}(X)}{E^2X} = \frac{EX^2 - E^2X}{E^2X}$$

where the second inequality uses Chebyshev's inequality.

$$EX^2 = E\left(\sum_{|A|=4} X_A\right)^2 = E\left[\sum_{|A|=4} X_A + \sum_{\substack{(A,B)\\ A\neq B}} X_A X_B\right] =$$

$$= EX + \sum_{\substack{(A,B)\\ A\neq B}} E(X_A X_B) =$$

$$= EX + \sum_{A\cap B=\emptyset} E(X_A X_B) + \sum_{|A\cap B|=1} E(X_A X_B) + \sum_{|A\cap B|=2} E(X_A X_B) +$$

$$+ \sum_{|A\cap B|=3} E(X_A X_B)$$

$$\sum_{A\cap B=\emptyset} E(X_A X_B) = \sum_{A\cap B=\emptyset} p^{12} = \binom{n}{4}\binom{n-4}{4} p^{12} \le E^2X$$

$$\sum_{|A\cap B|=1} E(X_A X_B) = \sum_{|A\cap B|=1} p^{12} = \binom{n}{4}\cdot 4 \cdot \binom{n-4}{3} p^{12} = o(E^2X)$$

$$\sum_{|A\cap B|=2} E(X_A X_B) = \sum_{|A\cap B|=2} p^{11} = \binom{n}{4}\binom{4}{2}\binom{n-4}{2} p^{11} = o(E^2X)$$

$$\sum_{|A\cap B|=3} E(X_A X_B) = \sum_{|A\cap B|=3} p^9 = \binom{n}{4}\binom{4}{3}\binom{n-4}{1} p^9 = o(E^2 X)$$

$$P[X=0] \le \frac{EX^2 - E^2X}{E^2X} \le \frac{EX + o(E^2X)}{E^2X} = o(1)$$

as desired.

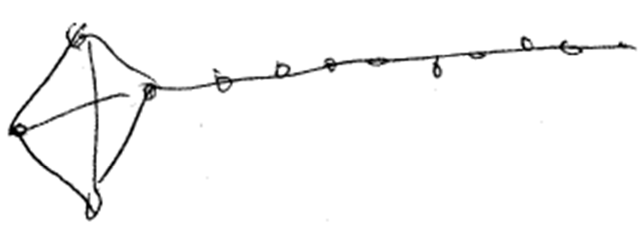
What if we replace K_4 by K_t? Let

$$\rho(K_t) = \frac{|E(K_t)|}{|V(K_t)|} = \frac{\binom{t}{2}}{t}$$

Then the threshold will be $\frac{1}{n^{1/\rho(K_t)}}$.

What if we replace K_t by a graph H?

Example. K_4 has threshold $\frac{1}{n^{2/3}}$.



has $\rho = 1 + \varepsilon$. Does that mean the threshold is $\frac{1}{n^{1/(1+\varepsilon)}}$? No!

Note that $\frac{1}{n^{2/3}} \gg \frac{1}{n^{1/(1+\varepsilon)}}$

Theorem (Erdős-Rényi 1960) Let H be a fixed graph and let α be the maximum edge-density among all induced subgraphs of H. Then $p = \frac{1}{n^{1/\alpha}}$ is a threshold for the event that $G \in \mathcal{G}(n, p)$ contains a subgraph isomorphic to H.