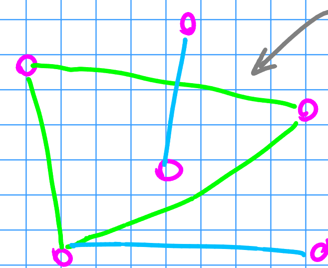
The Erdős-Rényi random graph $G(n,p)$

Small subgraphs

As a warm-up, we investigate the values of p that guarantee — with high probability — the existence of a triangle (i.e., a cycle of length 3) in the graph.

a triangle



Let N denote the number of triangles in $G(n,p)$. Then

$$\mathbb{E}N = \binom{n}{3} p^3$$

Then, by the "first moment method",

$$\mathbb{P}(N=0) = 1 - \mathbb{P}(N \geq 1) \geq 1 - \mathbb{E}N \geq 1 - \frac{(np)^3}{6}$$

Thus, if $p = o(\frac{1}{n})$, then

$G(n,p)$ does not contain any triangle w.h.p.

"with high probability"
i.e., with probability $\to 1$.

To get an upper bound, we use the "second moment method":

$$\mathbb{P}(N=0) = \mathbb{P}(N-\mathbb{E}N \le -\mathbb{E}N) \le \mathbb{P}(|N-\mathbb{E}N| \ge \mathbb{E}N)$$
$$\le \frac{\mathrm{var}(N)}{(\mathbb{E}N)^2},$$

so the graph contains a triangle w.h.p. whenever

$$\mathrm{var}(N) = o((\mathbb{E}N)^2).$$

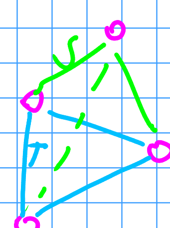
Writing $N = \sum'_T X_T$ where X_T is the indicator variable that triangle T is present,

sum over all $\binom{n}{3}$ triangles

$$\mathrm{var}(N) = \mathbb{E}(N^2) - (\mathbb{E}N)^2 = \mathbb{E}(N^2) - \left(\binom{n}{3}p^3\right)^2$$
$$= \sum'_{T,S} \underbrace{\left[\mathbb{E}(X_T \cdot X_S) - (\mathbb{E}X_T)(\mathbb{E}X_S)\right]}_{\mathrm{cov}(X_S, X_T)}$$

Note that if T and S don't have a common edge then X_T and X_S are independent and $\mathrm{cov}(X_S, X_T) = 0$.

So $$\mathrm{var}(N) = \sum'_T \left[\mathbb{E}X_T - (\mathbb{E}X_T)^2\right] \qquad \longleftarrow \text{terms when } S=T$$
$$+ \sum'_T \sum'_{S:\, S\cap T \text{ is one edge}} \left[\mathbb{E}(X_T X_S) - (\mathbb{E}X_T)^2\right]$$



$$= \binom{n}{3}(p^3 - p^6)$$
$$+ \binom{n}{3}\cdot(n-3)\cdot 3 \cdot p^5 \le (np)^3 + n^4 p^5 = o\left(\left(\binom{n}{3}p^3\right)^2\right)$$

whenever $np \to \infty$

We have proved that

If $np \to 0$ then $G(n,p)$ is triangle free w.h.p.

If $np \to \infty$ then $G(n,p)$ contains a triangle w.h.p.

$p_n = \frac{1}{n}$ is called a threshold function for the property.

What can we say when $np = c$ (constant)?
If the X_T were independent, N would have $\mathrm{Binom}\left(\binom{n}{3}, \left(\frac{c}{n}\right)^3\right)$ distribution which converges to a $\mathrm{Poisson}\left(\frac{c^3}{6}\right)$ distribution. (Recall that if Z is $\mathrm{Poisson}(\lambda)$ then $\mathbb{P}(Z=k) = \frac{\lambda^k}{k!} e^{-\lambda}$)

Exercise: Show that when $p = \frac{c}{n}$, the number N of triangles is asymptotically $\mathrm{Poisson}\left(\frac{c^3}{6}\right)$ in the sense that $\mathbb{P}(N=k) \xrightarrow{n\to\infty} \frac{(c^3/6)^k}{k!} e^{-c^3/6}$, $k = 0, 1, 2, \ldots$

In particular, $\mathbb{P}(N=0) \longrightarrow e^{-c^3/6}$

Hint: You may use the "method of moments": it suffices to prove that all moments of N converge to the corresponding moments of the Poisson distribution.

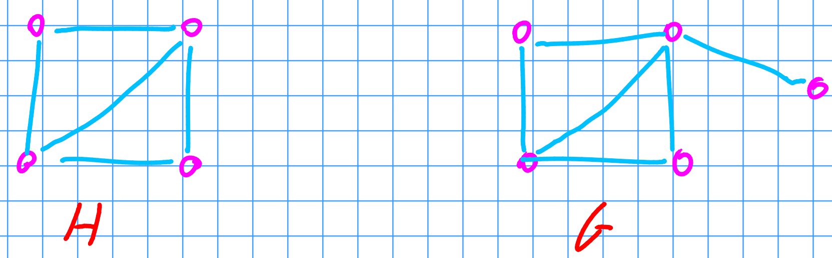
What can we say about other subgraphs? The first and second moment method extends, without problems, to K_k, cliques of size k. In that case

$$\mathbb{E}N_k = \binom{n}{k} p^{\binom{k}{2}} \approx \frac{1}{k!}\left(np^{\frac{k-1}{2}}\right)^k \longrightarrow \begin{cases} \infty & \text{if } p \gg n^{-\frac{2}{k-1}} \\ 0 & \text{if } p \ll n^{-\frac{2}{k-1}} \end{cases}$$

N_k: # of cliques of size k

The second moment method confirms this intuition.

However, $\mathbb{E}N$ does not always provide the correct answer. Consider



By the first moment method,

$$\mathbb{P}(G(n,p) \text{ doesn't contain } H) \geq 1 - \mathbb{E}N_H = 1 - \Theta(n^4 p^5)$$
$$\to 1 \text{ if } p = o(n^{-4/5})$$

Take, for example, $p = n^{-9/11} = o(n^{-4/5})$, so there are no copies of H in $G(n,p)$.

However, $\mathbb{E}N_G = \Theta(n^5 p^6) = \Theta(n^{1/11}) \to \infty$

Even though the expected number of copies of G goes to infinity, there are no copies of G w.h.p.!

In order to determine the correct threshold function for the subgraph containment problem, denote by v_H and e_H the number of vertices and edges in a graph H and for a graph G define

$$m(G) = \max\left\{\frac{e_H}{v_H} : \underbrace{H \subseteq G}_{\text{subgraph}}\right\}$$

Theorem (Bollobás '81)

For any fixed graph G,

$$\mathbb{P}(G \subseteq G(n,p)) \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } p \ll n^{-1/m(G)} \\ 1 & \text{if } p \gg n^{-1/m(G)} \end{cases}$$

Proof The first part follows from the first moment method: let $H \subseteq G$ such that

$$m(G) = \frac{e_H}{v_H} \quad \text{(a subgraph of max density)}$$

Then

$$\mathbb{P}(G \subseteq \mathcal{G}(n,p)) \le \mathbb{P}(H \subseteq \mathcal{G}(n,p))$$
$$\le \mathbb{E} N_H = O(n^{v_H} p^{e_H}) = O\left(\left(np^{m(G)}\right)^{v_H}\right) \to 0$$
if $np^{m(G)} \to 0$.

The second half follows from the second moment method. Just as before,

$$\mathbb{P}(G \not\subseteq \mathcal{G}(n,p)) = \mathbb{P}(N_G = 0) \le \frac{\operatorname{var}(N_G)}{(\mathbb{E} N_G)^2}$$

But

$$\operatorname{var}(N_G) = \sum_{\substack{G', G'' \\ \text{copies of } G}}' \operatorname{cov}(\underbrace{X_{G'}}_{\text{indicator of } G'}, X_{G''})$$

The key observation is that for each $H \subseteq G$ there are

$$\Theta\Big(\underbrace{n^{v_H}}_{\substack{\#\text{ of ways of} \\ \text{choosing } H}} \cdot \underbrace{n^{2(v_G - v_H)}}_{\substack{\#\text{ of ways of} \\ \text{choosing the remaining} \\ \text{vertices}}}\Big)$$

pairs G', G'' such that $G' \cap G''$ is isomorphic to H.

So

$$\operatorname{var}(N_G) = \sum_{H \subseteq G}' n^{2v_G - v_H} \cdot \left(p^{2e_G - e_H} - p^{2e_G}\right)$$

$$\asymp \max_{H \subseteq G} n^{2v_G - v_H} \cdot p^{2e_G - e_H}$$

$$\asymp \max_{H \subseteq G} \frac{(\mathbb{E} N_G)^2}{\mathbb{E} N_H}$$

"same order of magnitude"

Thus,

$$\mathbb{P}(G \not\subseteq \mathcal{G}(n,p)) = O\left(\max_{H \subseteq G} \frac{1}{\mathbb{E} N_H}\right) = O\left(\max_{H \subseteq G} n^{-v_H} p^{-e_H}\right)$$

Finally, note that if $np^{m(G)} \to \infty$ then

$$np^{e_H/v_H} \to \infty \text{ for all } H \subseteq G \text{ and therefore}$$

$$\max_H n^{-v_H} p^{-e_H} = \max_H \left(np^{e_H/v_H}\right)^{-v_H} \to 0 \text{ as desired.} \blacksquare$$

Trees

When T is a tree on k vertices, then clearly $m(T) = \frac{e_T}{v_T} = \frac{k-1}{k}$ and therefore when $n^{-1-\frac{1}{k-1}} \ll p \ll n^{-1-\frac{1}{k}}$ then $\mathcal{G}(n,p)$ contains all trees of size k but no tree of size $k+1$. (Here k is fixed!)

In this entire range (i.e., for $p = o(\frac{1}{n})$) the graph is a forest: The expected number of cycles X_c of any length is

$$\mathbb{E}X_c = \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k \leq \sum_{k=3}^{n} \frac{(pn)^k}{2k} \to 0$$

if $np \to 0$,

so by the first moment method,

$$\mathbb{P}(X_c \geq 1) \leq \mathbb{E}X_c \to 0.$$

Thus, for any fixed $k \geq 2$, if

$$n^{-1-\frac{1}{k-1}} \ll p \ll n^{-1-\frac{1}{k}},$$

then $\mathcal{G}(n,p)$ contains only trees (i.e., it is a forest) and the size of the largest component equals k w.h.p.

Nothing "happens" between $n^{-100/99}$ and $n^{-101/100}$.