Setting

let C_1 be the largest component of $G_{n,p}$, $p = \frac{k}{n}$.

We showed in class that

$$k < 1 \Rightarrow |C_1| = o(\log n)$$
$$k > 1 \Rightarrow |C_1| = \omega(n).$$

What happens when $k = 1$?

Results

Turns out that if $p = \frac{1}{n}$, $|C_1| \sim n^{2/3}$

This paper shows

Thm 1: for $A > 8$, $n > 1000$?

$$P(|C_1| > An^{\frac{2}{3}}) \leq \frac{4}{A} e^{-\frac{A^2(A-4)}{32}}$$

Thm 2: for $0 < \delta < \frac{1}{10}$, $n > \frac{200}{\delta^{3/5}}$

$$P(|C_1| < \lfloor \delta n^{2/3} \rfloor) \leq 15\delta^{3/5}$$

which gives us very specific bounds.

$A = 8.1$: $P(|C_1| > 8.1 n^{2/3}) \leq .00011$

$\delta = 10^{-5}$: $P(|C_1| < 10^{-5} n^{2/3}) \leq .015$

so for $n > 200{,}000$, we "probably" have

$$10^{-5} n^{2/3} < |C_1| < 8.1 n^{2/3}$$

Note that these bounds don't depend on n, so don't get better probabilities as $n \to \infty$. You get better probabilities by relaxing the constants on the bounds.

Something that is true, but not proven in this paper, is that in fact there is some fixed constant multiple of $n^{2/3}$ that $|C_1|$ approaches.
i.e. $\frac{|C_1|}{n^{2/3}}$ does converge to some fixed value.

More generally,

Thm 3: (Luczak, Pittel, Wierman 1994)

Let $p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$ where $\lambda \in \mathbb{R}$ is fixed. Then for any $m \in \mathbb{N}$, the sequence

$$\left(\frac{|C_1|}{n^{2/3}}, \frac{|C_2|}{n^{2/3}}, \ldots, \frac{|C_m|}{n^{2/3}} \right)$$

converges in distribution to a random vector with positive components.

the proofs of thm's 1 and 2 are calculation heavy, so instead, we will prove

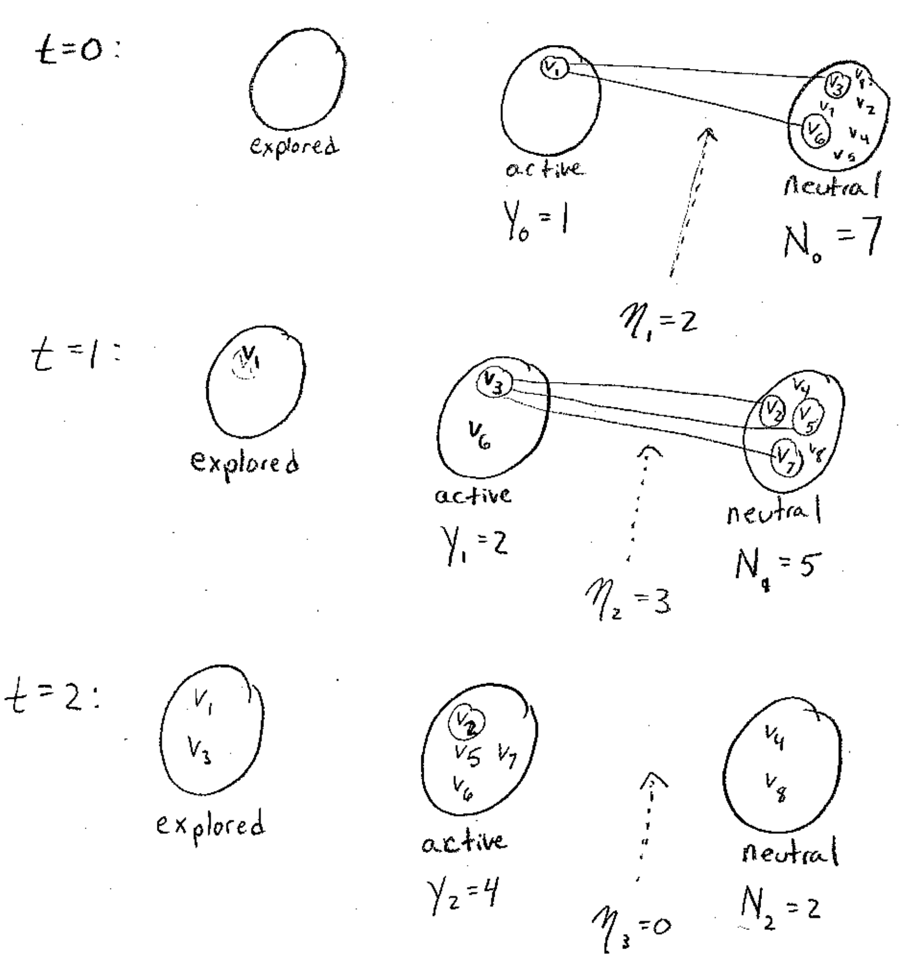
$$P(|C_1| > An^{2/3}) \leq \frac{3}{A^{3/2}}$$

which is not as good as thm1, since we lose the nice exponential factor, but as a consolation prize, we get an extra factor of $\frac{1}{\sqrt{A}}$.

Exploration Process & Variable naming

for a vertex v, $C(v)$ is the component containing v.

label the vertices $v_1, \dots, v_n$ so that $v = v_1$.



at every step we "explore" the active vertex of lowest index and "activate" all its neutral neighbors.

Note that

$$Y_t = 1 + \sum_{i=1}^{t} (\eta_i - 1)$$

and that η_t is a random variable, $\eta_t \sim Bi(N_{t-1}, p)$.

Random Walk

This is difficult to analyze, so we will compare it to a random walk.

let $\{\xi_i\}$ be iid, distributed like $Bi(n, p)$ and coupled with $\{\eta_i\}$ so that $\xi_i \geq \eta_i$.

$$S_t = 1 + \sum_{i=1}^{t} (\xi_i - 1)$$

So S_t is a random walk on $\mathbb{N}$, starting at 1 and finishing when we get to 0, or pass H. At every step, we will either step back 1, stay in place, or step forward by some amount.

Note, however, that $E[\xi_i] = n \cdot p = 1$, so $\{S_t\}$ is a martingale.

How will we use this random walk?

If $|C(v)| > H^2$, then either our random walk ended at H or it is still going at time H^2. We will bound that probability by $\frac{3}{H}$ and take $H = \sqrt{An^{2/3}}$ so that

$$P(|C(v)| > An^{2/3}) < \frac{3}{\sqrt{An^{2/3}}}$$

Calculations

Let γ be the stopping time

$$\gamma = \min\{t : S_t = 0 \text{ or } S_t \geq H\}.$$

Since "in expectation $\{S_t\}$ is staying put," you might think its likely that $S_\gamma = 0$. In fact that is true.

$$1 = E[S_\gamma] \leq H \cdot P(S_\gamma \geq H)$$

optional stopping theorem

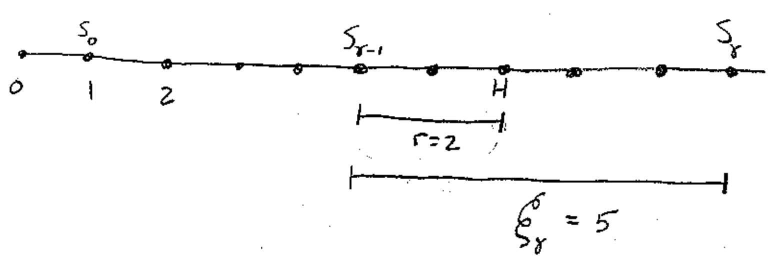
Markov's inequality

so $$P(S_\gamma \geq H) \leq \frac{1}{H}.$$

Bounding the probability that our walk is still going at time H^2 is harder.

Lemma 5: Let $X \sim Bi(n,p)$. Then ($S_\gamma - H$ given $S_\gamma \geq H$) is stochastically dominated by X.

Pf: let $r = H - S_{\gamma-1}$. $S_\gamma - H = \xi_\gamma - r < \xi_\gamma$ which is $Bi(n,p)$.



Cor 6: If f is an increasing function

$$E[f(S_\gamma - H) \mid S_\gamma \geq H] \leq E[f(X)]$$

write $S_\gamma^2 = (H + S_\gamma - H)^2 = H^2 + 2H(S_\gamma - H) + (S_\gamma - H)^2$

and use $f(x) = 2Hx + x^2$

to get

$$E[S_\gamma^2 \mid S_\gamma \geq H] = E[H^2 + f(S_\gamma - H) \mid S_\gamma \geq H]$$
$$\leq H^2 + E[f(X)]$$
$$= H^2 + E[2HX + X^2]$$
$$= H^2 + 2H \cdot 1 + 2 - \tfrac{1}{n}$$
$$\leq H^2 + 3H \qquad \text{(for } H \geq 2\text{)}.$$

Note that $S_t^2 - (1-\frac{1}{n})t$ is also a martingale (this is not too hard to check) so we can apply the optional stopping theorem.

$$S_0^2 - (1-\tfrac{1}{n}) \cdot 0 = E[S_\gamma^2 - (1-\tfrac{1}{n})\gamma]$$
$$1 = E[S_\gamma^2] - (1-\tfrac{1}{n})E[\gamma]$$
$$1 + (1-\tfrac{1}{n})E[\gamma] = E[S_\gamma^2].$$

So

$$1 + (1-\tfrac{1}{n})E[\gamma] = E[S_\gamma^2] = E[S_\gamma^2 \mid S_\gamma = 0] \cdot P(S_\gamma = 0) + E[S_\gamma^2 \mid S_\gamma \geq H] \cdot P(S_\gamma \geq H)$$
$$\leq 0 + (H^2 + 3H) \cdot \tfrac{1}{H}$$
$$= H + 3$$

*** Algebraic Manipulation ***

when $3 \leq H \leq n-3$

$$E[\gamma] \leq H + 3$$

So $P[\gamma \geq H^2] \leq \frac{E[\gamma]}{H^2} \leq \frac{H+3}{H^2} \leq \frac{H+H}{H^2} = \frac{2}{H}$

Wrap up of calculations

So if the component of the graph we are exploring has size at least H^2, then our exploration process either ends at $\gamma \geq H$ or is still going at $t = H^2$. Thus

$$P(|C(v)| \geq H^2) \leq P(S_\gamma \geq H) + P(\gamma \geq H^2) \leq \frac{1}{H} + \frac{2}{H} = \frac{3}{H}.$$

taking $T = H^2$

$$P(|C(v)| \geq T) \leq \frac{3}{\sqrt{T}} \qquad \text{for} \qquad 9 \leq T \leq (n-3)^2.$$

Let N_T = number of vertices contained in a component of size at least T.

$$P(|C_1| > T) \leq P(|N_T| > T) \underset{\text{markov}}{\leq} \frac{1}{T} \cdot E[|N_T|] \underset{\text{linearity of expectation}}{=} \frac{1}{T} \cdot n \cdot P[C(v) > T]$$

$$\leq \frac{3n}{T^{3/2}}$$

Taking $T = An^{2/3}$ we get

$$P(|C_1| > An^{2/3}) \leq \frac{3n}{A^{3/2} n} = \frac{3}{A^{3/2}}$$