Random Graphs and Diffusions

Tuesday, December 3, 2019 8:17 PM

Erdos - Renyi Random Graph

$G(n,p)$ - n vertices
- each pair of vertices joined by an edge with probability p.

Sparse case $p = d/n$, average degree $\frac{d}{n}\cdot(n-1) \to d$.

Degree of one vertex
$\mathrm{Bin}(n-1, d/n)$

$$\mathbb{P}[\mathrm{Bin}(n-1, d/n) = k] = \binom{n-1}{k}\left(\frac{d}{n}\right)^k\left(1-\frac{d}{n}\right)^{n-k-1}$$

$$= \frac{(n-1)\cdot\ldots\cdot(n-k)}{n^k}\,\frac{d^k}{k!}\left(1-\frac{d}{n}\right)^n\left(1-\frac{d}{n}\right)^{k+1}$$

$$\to \frac{d^k}{k!}e^{-d} = \mathbb{P}[\mathrm{Pois}(d) = k]$$

Local neighbourhood

Number of k-cycles

$$\mathbb{E}\, N_k = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{2k}\cdot\left(\frac{d}{n}\right)^k$$

$$\to \frac{d^k}{2k}$$

So the graph is locally treelike.

Ex: $G(n, \frac{d}{n}) \xrightarrow{lwc} GWC(Pois(d))$

Proof:

Reveal graph with breadth first search.

$\# \text{neighbours} \sim Bin(n-1, d/n)$

$\approx Pois(d)$

$\sim Bin(n-3, d/n)$

Component Sizes

Theorem:

- $G(n, d/n)$. Components $C_1, C_2, \ldots$ ordered such that $|C_1| \geq |C_2| \geq \ldots$

  - If $d < 1$ then $|C_1| \leq O(\log n)$

  - If $d > 1$ $|C_1| = (c_d + o(1))n$

    $|C_2| \leq \log n$ $\quad c_d = \mathbb{P}[GW(d) \text{ survives}] > 0.$

  - If $d = 1$, $|C_1| \asymp n^{2/3}$

    $|C_2| \asymp n^{2/3}$

Reveal Components:

Proof: Breadth first search

- A_t - active vertices
-
- $U_t = n - A_t - t$ unexplored vertices

$A_0 = 1,$

$A_t = A_{t-1} + W_t$

where $W_t = Bin(U_t, d/n)$

let $\tilde{W}_t = Bin(n, d/n)$ IID

$W_t \leq \tilde{W}_t$

Then $A_t = 1 - t + S_t$

$\leq 1 - t + \tilde{S}_t$

$S_t = \sum_{i=1}^{t} W_i.$

time 0 $A_1 = 1$

time1 $A_1 = 2$

time 2 $A_3 = 1$

time 3 $A_4 = 1$

$A_5 = 0.$

$\mathbb{P}[\tilde{S} - t + 1 > 0] \leq e^{-cn}.$

So $\mathbb{P}[|C_1| > C \log n] \to 0.$

Supercritical Case: $G(n, d/n)$ $p > 1.$

$U_t \geq n - \tilde{S}_t \underset{w.h.p.}{\geq} n(1-\delta)$ for $t < \delta' n.$

such that $(1-\delta)p > 1 + \varepsilon$

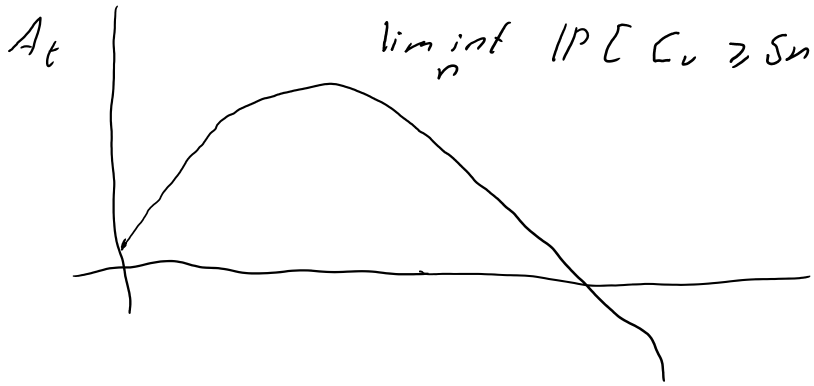
Set $\hat{W}_t \sim \mathrm{Bin}(n(1-\delta), \frac{d}{n}) \preccurlyeq W_t$ for $t < \delta' n$.

$A_t \geqslant 1 - t + \hat{S}_t$ a random walk with positive drift, so

$\mathbb{P}[\forall t \;\; \hat{S}_t - t \geqslant 0] > 0$

So $\mathbb{P}[\forall t \leqslant \delta n, \; A_t \geqslant 1] > 0$.

$\liminf_n \mathbb{P}[C_v \geqslant \delta n] > 0$



Rescaling of S_t:

$\mathbb{E}\, W_t = \mathbb{E}\, \mathrm{Bin}(n - S_t, d/n)$

If $f(t) = \frac{1}{n} A_{nt} = \frac{1}{n} S_{nt} - t$

$g(t) = \frac{1}{n} S_{nt}$

$g'(t) = (n - n g(t)) \cdot \frac{d}{n} = d(1 - g(t))$

$f'(t) = d(1 - g(t)) - 1$

$= d(1 - f(t) - t) - 1.$

Solution: $f(t) = 1 - e^{-dt} - t. \Rightarrow 1 - t = e^{-dt}$

Branching process approach,

probability generating function

$g(\theta) = \sum_k \theta^k p_k$

For Poisson(d) $g(\theta) = \sum_{n=0}^{\infty} \theta^k \frac{d^k e^{-d}}{k!} = e^{\theta d} e^{-d} = e^{-d(1-\theta)}$

Extinction probability is smallest Θ_* s.t.

$$\Theta_* = e^{-d(1-\Theta_*)}$$

SDE:

Let $X_t = \frac{1}{\sqrt{n}}(S_{nt} - n g(t))$

$$X_{t+\frac{1}{n}} - X_t \cong \frac{1}{\sqrt{n}}(W_t - g'(t))$$

$$\mathbb{E}\left(X_{t+\frac{1}{n}} - X_t \mid \mathcal{F}_t\right] = n \cdot \frac{1}{\sqrt{n}} \mathbb{E}\left(Bin(n - S_{nt}, d/n) - d(1-g(t)) \mid \mathcal{F}_t \right\}$$

$$= \sqrt{n}\, \mathbb{E}\left(-\frac{d}{n}(S_{nt} - n g(t))\right)$$

$$= -d X_t$$

$$n \operatorname{Var}\left(X_{t+\frac{1}{n}} - X_t \mid \mathcal{F}_t \right] = n \operatorname{Var}\left(\frac{1}{\sqrt{n}} W_t \right) \cong d(1-g(t))$$

$$dX_t = -d \cdot X_t \, dt + \sqrt{d(1-g(t))}\, dB_t.$$

Solution is

$$X_t = \int_0^t \sqrt{d(1-g(s))}\, e^{-d(t-s)}\, dB_s.$$

So X_t is approximately normal.

The component size is also normal.

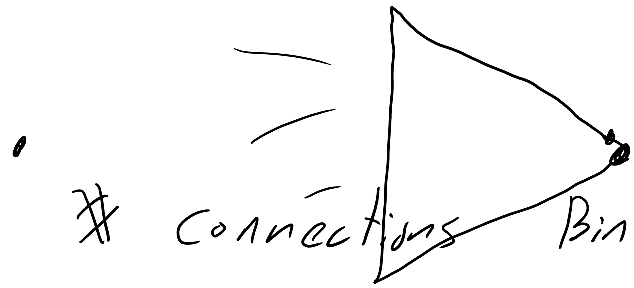
$$\forall t \quad \mathbb{P}\left[\hat{S}_t - t \le \left((1-\delta)p - \tfrac{\varepsilon}{2}\right) t \right] \le e^{-ct}$$

so

$$\exists D \text{ s.t. } \mathbb{P}\left[D \log n \le |C_v| \le \delta n \right] \le n^{-2}.$$

— At most 1 giant component

$n^{2/3}$ $n^{2/3}$



connections $\mathrm{Bin}(n^{4/3}, d/n)$

Concentration

Let A_v be the event

$A_v = \{ |C_v| \geq D \log n \}$

If $\{v_1, \dots, v_m\}$ random vertices, $m \approx n^{\varepsilon}$ then $\{A_{v_i}\}$ close in TV distance to a product measure since # revealed vertices is $O(n^{\varepsilon} \log n)$

$$\mathbb{P}\left[\left|\frac{1}{m}\sum_{i=1}^{m} 1_{A_{v_i}} - \mathbb{P}[A_{v_i}]\right| > \varepsilon\right] \to 0.$$

$$\Rightarrow \mathbb{P}\left[\left||C_1| - \mathbb{P}[A_{v_i}]\right| > \varepsilon\right] \to 0.$$

$$\mathbb{P}[A_{v_i}] = O\left(\frac{1}{n^2}\right) + \mathbb{P}[|C_v| \geq L]$$

$$- \mathbb{P}[L \leq |C_v| \leq \delta n]$$

$$\lim_{L\to\infty} \lim_{n\to\infty} \mathbb{P}[|C_v| < L] = 1 - \mathbb{P}[\text{Survives}]$$

$$\lim_{L\to\infty} \mathbb{P}[L \leq |C_v| \leq \delta n] \leq \lim_{L\to\infty} e^{-cL} = 0$$

So $\mathbb{P}[A_{v_i}] \xrightarrow[n\to\infty]{} \mathbb{P}[\text{Survives}]$

$$\frac{1}{n}|C_1| \xrightarrow[n\to\infty]{P} \mathbb{P}[\text{Survives}]$$

Critical Case

- Each time a component runs out add a new vertex.

- N_t = # components visited by time t, $N_0 = 1$.

$$A_t = N_t - t + S_t \geq 1.$$

Recall $S_{t+1} - S_t \sim Bin(U_t, \frac{1}{n})$

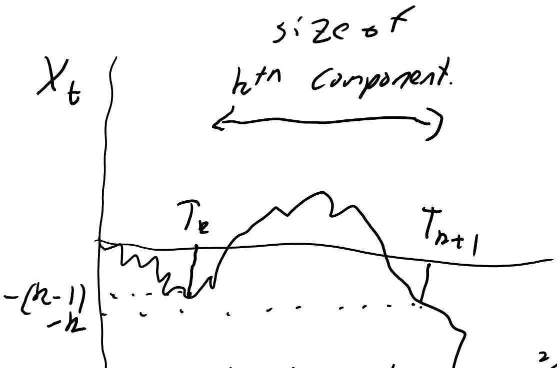
$$= Bin(n - S_t - N_t, \frac{1}{n})$$

$$X_t = S_t - t$$

$$N_t = \min_{0 \le u \le t} X_u \quad + 1$$

Let $T_k = \min\{ t : X_t = -k+1 \}$

Component size k is $T_{k+1} - T_k$



Let $t \le n^{2/3+\varepsilon}$.

$$\mathbb{E}[X_{t+1} - X_t \mid \mathcal{F}_t] = \mathbb{E}(\mathrm{Bin}(U_t, \tfrac{1}{n}) - 1 \mid \mathcal{F}_t)$$

$$= \frac{U_t}{n} - 1 = -\frac{1}{n}(S_t + N_t)$$

$$\mathrm{Var}[S_{t+1} - (t+1) \mid \mathcal{F}_t] = \mathrm{Var}(\mathrm{Bin}(U_t, \tfrac{1}{n}) \mid \mathcal{F}_t)$$

$$= U_t \, \tfrac{1}{n} (1 - \tfrac{1}{n})$$

$Z_t = X_t + \frac{1}{n} \sum_{i=0}^{t-1} S_i + N_i$ is a martingale

and $\mathrm{Var}(Z_t) = \sum_{i=0}^{t-1} \mathbb{E}\, U_i \frac{1}{n}(1 - \frac{1}{n}) \le t$

$S_t \approx t$ for $t = o(n)$.

$N_t \approx$ mi ... $t-1$

$$N_{n^{3/4}} = \min_{0 \le t \le n^{3/4}} X_t = \min_t Z_t - \frac{1}{n} \sum_{i=0}^{t-1} (S_i + N_i)$$

$$\geq \min_t Z_t - \frac{n^{3/4}}{n} (S_{n^{3/4}} + n^{3/4})$$

$$\mathbb{P}[N_{n^{3/4}} \geq -n^{4/7}]$$

$$= \mathbb{P}[\min_{t \le n^{3/4}} X_t < -n^{4/7}]$$

$$\leq \mathbb{P}[\max_{t \le n^{3/4}} -Z_t \geq \tfrac{1}{2} n^{4/7}] \longrightarrow 0$$ Doobs maximal inequality

$$+ \mathbb{P}[\frac{n^{3/4}}{n} (S_{n^{3/4}} + n^{3/4}) \geq \tfrac{1}{2} n^{4/7}]$$

$\longrightarrow 0$ L.D. for S_t.

So for $t \simeq n^{2/3}$, $S_t + N_t \approx t$.

<u>Rescaling</u>:

$$Y_t = n^{-1/3} (S_{n^{2/3} t} - n^{2/3} t)$$

$\boxed{\text{Martingale CLT.}}$ $n^{-1/3} Z_{n^{2/3} t} \longrightarrow B_t$ (std B.M.)

since $\sum_{i=1}^{t} \mathbb{E}(M_{i+1} - M_i)^2 \approx t$.

(and no big jumps)

- $Y_t - n^{-1/3} Z_{n^{2/3} t}$

$$= - n^{-1/3} \frac{1}{n} \sum_{i=0}^{n^{2/3} t - 1} S_i + N_i$$

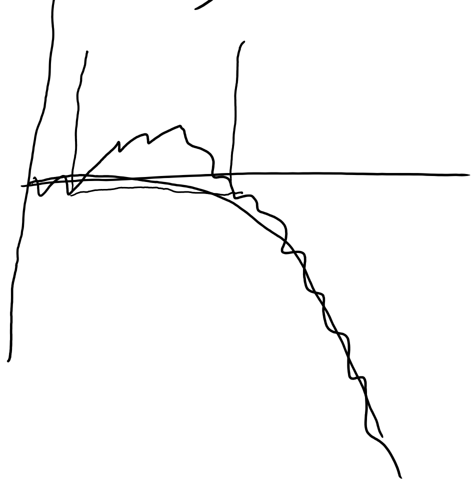
$$\approx - n^{-4/3} \sum_{i=0}^{n^{2/3} t - 1} i$$

$$\approx - n^{-4/3} \cdot \frac{1}{2} \left(n^{2/3} t\right)^2$$

$$= -\frac{1}{2} t^2$$

$$Y_t \xrightarrow{d} B_t - t^2/2$$

long $O(n^{2/3})$ excursions.



Cluster sizes

$$Z_t = Y_t - \min_{s \le t} Y_s$$

$n^{-2/3}$(Cluster sizes) $\iff$ excursion lengths