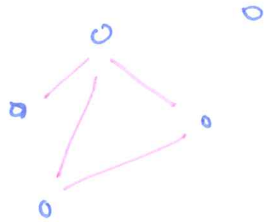
Random Graphs

Mike Steel
Oct 7 2011

Simplest model.

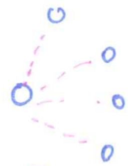
Erdős-Renyi



$G_{n,p}$ $\qquad G_p$

$G_{n,k}$

G_p : $\deg(v) \sim Bin(n-1, p)$



$\approx Po(\lambda)$ $\quad \lambda = np$ [Sometimes $p = p_n$, so $\lambda = \lambda_n$]

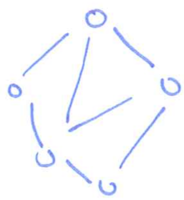
$G_{n,k}$ matching $k = \binom{n}{2} \cdot p$. (in G_p # edges random)

in $G_{n,k}$ $\deg(v)$ random but # edges $= k$.

Why useful?

1/. Maths: The "probabilistic method"

The "Sociology theorem"



6 people.
3 friends OR
3 strangers

S "good" if all pairs of nodes in S connected. or no pair of nodes connected.

Ramsey theory: R_k = Smallest n : all G_n have ≥ 1 good set of size k

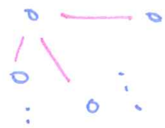
$R_3 = 6$, $R_4 = 18$, $R_5 = ??$ $(\in 43, 49)$

Question: Is there a graph on 100 nodes with no good set of size 10? (i.e. is $R_{10} > 100$?)

Even verifying a candidate hard!

Yes!

Proof 1: Consider $G_{100,1/2}$



(coin toss)

$$X_S = \begin{cases} 1, & \text{if } S \text{ good} \\ 0, & \text{else} \end{cases}$$

$$X = \#\text{ good sets in } G_{n,1/2} = \sum_S X_S.$$

$$E(X) = E\left(\sum_S I_s\right) = \sum_S E(I_s) \quad (\text{note } I_s \text{ \underline{not} independent})$$
$$= \sum_S \mathbb{P}(I_s = 1)$$
$$= \sum_S 2 \cdot \left(\tfrac{1}{2}\right)^{\binom{10}{2}} = \underbrace{\binom{100}{10}}_{1.7\times 10^{13}} \cdot \underbrace{2^{-\binom{10}{2}+1}}_{1/1.8\times 10^{13}}$$
$$< 1$$

So $E(X) < 1$. But $I_s \in 0, 1$ so $\underbrace{\mathbb{P}(\exists S : I_S = 0)}_{>0}$

so $\exists G$: $I_S = 0$ for G.

Proof 2: A_s: $I_S = 1$ (i.e. S is good)

$$\mathbb{P}\left(\bigcup_S A_S\right) \le \sum_S \mathbb{P}(A_S). \;\dots\; < 1.$$

$$\mathbb{P}\left(\bigcap_S \bar{A}_S\right) > 0. \Rightarrow \bigcap_S \bar{A}_S \neq \phi$$

This type of argument shows $R(k,k) \ge 2^{k/2}$.

$$\left[\frac{n^k}{k!} > \binom{n}{k} \ge 2^{\frac{k(k-1)}{2} - 1} \Rightarrow n \ge 2^{k/2}\right]$$

Why useful?

2). Very nice structure (properties hold a.s. as n grows)
(– can be useful for applications) See Appendix 1A.

Example
→ * Evolution of $G_{n,k}$ as k grows (Pic)
→ * $\mathbb{P}(G_p$ connected$)$

Suppose $p = \frac{\log n + c}{n}$

$$\mathbb{P}(\deg(v) = 0) = (1-p)^{n-1} \sim \left(1 - \frac{(\log n + c)}{n}\right)^n$$

$$\sim e^{-\log n} e^{-c} = \frac{e^{-c}}{n}.$$

$$\mathbb{P}(\deg(v) > 0 \text{ for all } v)$$

$$= \mathbb{P}\left(\bigcap_v \deg(v) > 0\right)$$

$$\simeq \prod_v \mathbb{P}(\deg(v) > 0)$$

$$= \left(1 - \frac{e^{-c}}{n}\right) = e^{-e^{-c}}$$

$\deg(v) > 0$ for all v is necessary but not sufficient for G_p to be connected.

But !

Classic Result Theorem (E-R) $p = \frac{\log n + c}{n}.$

$$\mathbb{P}(G_p \text{ connected}) \longrightarrow e^{-e^{-c}} \quad \text{as } n \to \infty.$$

Example : People on earth $n = 6 \times 10^{10}$, but lets say 10^{10}

$\lambda = np = 100$ friends (average)

$np = 100 = 10 \log_e 10 + c \Rightarrow c \simeq 77. \quad e^{-e^{-77}} \simeq 1.$

What about

Δ_G = max min path length between any two nodes

"Six degrees of separation"

Heuristic: If $\deg(v) \leq d \;\; \forall v$ then

$$n \leq 1 + d + d^2 + \dots + d^\Delta < d^{\Delta+1}$$

$$\Rightarrow \Delta \geq \frac{\log n}{\log d} - 1.$$

Take $d = np$

$$\Delta \gtrsim \frac{\log(n)}{\log(np)} \; (?)$$

(Chung + Lu)

Theorem If $p \geq c\frac{\log n}{n}$, $c > 2$ then:

$$\Delta_G \sim \frac{\log(n)}{\log(np)}$$

World again! $n = 10^{10}$ $np = 100$ $\left(c = \frac{100}{10\log_e 10} \gg 2\right)$

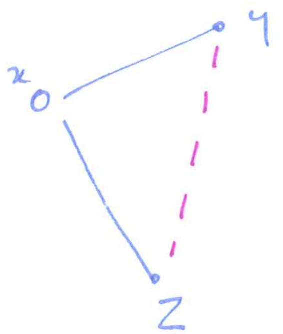
$$\Delta_G \sim \frac{\log 10^{10}}{\log 10^2} = 5$$ (+ connected)

"Six degrees of separation"

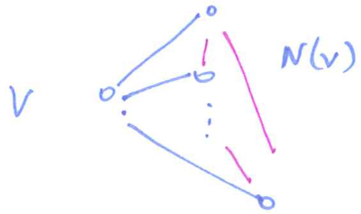
Av. path distance $\overline{\Delta_G}$.

Problem with E-R model: (Independence assumption)

'real networks' have high transitivity



Clustering coefficient



local

$$C_v = \frac{\# \text{ edges in } N(v)}{\binom{k_v}{2}}$$

$k_v = |N(v)|$.

global

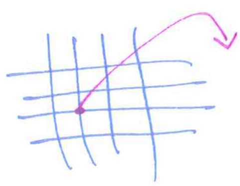
C  vs

$$\bar{C} = \underset{v}{\text{av}} \{C_v\} = \frac{1}{n}\sum_v C_v.$$

For E-R graphs $\bar{C}$, C are typically small. Appendix 1B.

But for social networks, (eg WWW) or other realistic networks (eg PP interaction nets) $\bar{C}$ large

Watts + Strogatz model "small world networks"



$\bar{\Delta}_G$ still $\sim \log n$

but $\bar{C}$ larger.

Appendix 1

A/ Evolution of random graph

$k \uparrow$ or $p = p_n$.

$p \le n^{-\beta}$ $\quad \beta > 2$
a.s. no edges at all.

$\beta = -3/2$ a.s. has an edge

$\beta = n^{-1}$ first cycles appear
one component outgrows other and at

$p \approx \frac{\log n}{n}$ devours them so graph connected

$p = (1+\varepsilon)\frac{\log n}{n}$ G_p has a.s. a Hamiltonian cycle.

$\vdots$

B/.

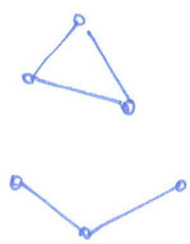
C_v

$X \sim Bin(n-1, p)$
Conditional on $X = k$

$$C_v = \frac{\binom{k}{2} p}{\binom{k}{2}} = p = \lambda/n.$$

$$\left(= \frac{10^2}{10^{10}} = 10^{-8} \right)$$

C



$$\frac{\binom{n}{3} p^3}{3\binom{n}{3} p^2 (1-p) + \binom{n}{3} p^3} \sim \frac{\lambda}{3n}.$$