Lecture 17 Configuration model & Excess degree distribution

your friends have more friends than you do on average — why?

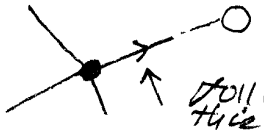
n, k_i — Given $\qquad \sum k_i = 2m$

$\Rightarrow$ $2m$ stubs (random)

Choosing two stubs at random & linking them until all $2m$ stubs have been used up $\Rightarrow$ Ensemble of networks with p_k

Notation $G(n,m)$

Recall p_k — probability that a node chosen at random has degree k



What is the probability that this node has degree k? p_k? No!

($k=0$: $p_0 \neq 0$ but no chance of reaching such a node)

Following an edge $\longrightarrow$ $(2m-1)$ stubs to reach

Prob. of ending at a node of degree k $= k/(2m-1)$

total # of nodes with degree k $\rightarrow$ $n p_k \cdot \dfrac{k}{2m-1} \approx \dfrac{k}{2m} n p_k = \dfrac{k p_k}{\langle k \rangle}$ — prob. of reaching ANY node of degree k

through neighbors — more likely to reach a high-degree node by following any one of the k edges

Consequence?

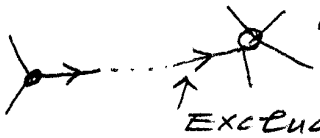
Ave. degree of neighbor $= \sum k \dfrac{k p_k}{\langle k \rangle} = \dfrac{\langle k^2 \rangle}{\langle k \rangle}$ $\qquad \langle k^2 \rangle$ - 2nd moment

$$\langle k^2 \rangle / \langle k \rangle - \langle k \rangle = \frac{1}{\langle k \rangle}\left(\langle k^2 \rangle - \langle k \rangle^2\right) = \frac{\sigma_k^2}{\langle k \rangle} > 0$$

Reason: when going through the nodes & averaging the degrees of the neighbors of each node, many of these neighbors appear in more than one average (k averages for a node of degree k)

High-degree nodes $\longrightarrow$ over represented

Excess degree $k-1$ $\qquad q_k$ — prob. of having excess degree k



$=$ prob. of reaching a node with degree $k+1$

$$\Rightarrow \quad q_k = (k+1) p_{k+1} / \langle k \rangle$$

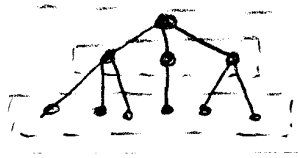
Generating functions $\qquad g_0(z) = \sum p_k z^k$

$\qquad g_1(z) = \sum q_k z^k$

$$g_1(z) = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} (k+1) p_{k+1} z^k = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} k p_k z^{k-1}$$

$$= \frac{1}{\langle k \rangle} g_0'(z) = g_0'(z) / g_0'(1)$$

2nd neighbors



# of 2nd neighbors $= \sum$(excess degree of 1st neighbors)

Prob. of having exactly k 2nd neighbors

Prob. of having k 2nd neighbors given that there are ℓ first neighbors

$$\Rightarrow p_k^{(2)} = \sum_{\ell=0}^{\infty} p_\ell \, \overbrace{P^{(2)}(k|\ell)}$$

ℓ first neighbors

$j_1, \cdots j_\ell$ – excess degrees

$$\Rightarrow P^{(2)}(k|\ell) = \sum_{j_1=0}^{\infty} \cdots \sum_{j_\ell=0}^{\infty} \prod_{r=1}^{\ell} q_{j_r} \, \delta\left(k, \sum_{r=1}^{\ell} j_r\right)$$

$$\Rightarrow \quad p_k^{(2)} = \sum_{\ell=0}^{\infty} p_\ell \sum_{j_1=0}^{\infty} \cdots \sum_{j_\ell=0}^{\infty} \delta\left(k, \sum_{r=1}^{\ell} j_r\right) \prod_{r=1}^{\ell} q_{j_r}$$

<u>GF</u> $$g_2(z) = \sum_{k=0}^{\infty} p_k^{(2)} z^k = \cdots = \sum_{\ell=0}^{\infty} p_\ell \left[\sum_{j=0}^{\infty} q_j z^j\right]^\ell$$

$$= \sum_{\ell=0}^{\infty} p_\ell \left[g_1(z)\right]^\ell = g_0(g_1(z))$$

Generalizing: $p_k^{(3)}$: $g_3(z) = g_0(g_1(g_1(z)))$
$= g_2(g_1(z))$

$$g_d(z) = g_{d-1}(g_1(z))$$

What is the ave. number of neighbors at distance d?

$d=2$: $$C_2 = g_2'(z)\big|_{z=1} = g_0'(g_1(z)) \cdot g'(z)\big|_{z=1}$$

$\boxed{g_1(z) = g_0'(z)/g_0'(1)}$

$$= \underbrace{g_0'(1)}_{\langle k \rangle} \, g_1'(1)$$

$$g_1'(1) = \sum_{k=0}^{\infty} k q_k = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} k(k+1) p_{k+1} = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} (k-1)(k) p_k = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$$

$$\Rightarrow C_2 = \langle k^2 \rangle - \langle k \rangle$$

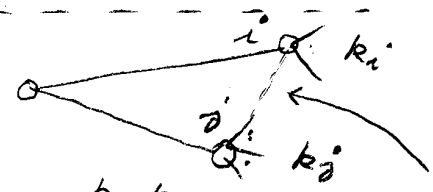
Arbitrary value of d

$$C_d = \left(\frac{C_2}{C_1}\right)^{d-1} \cdot C_1$$

$C_2 > C_1 \rightarrow$ Giant component

or $\langle k^2 \rangle - \langle k \rangle > \langle k \rangle \Rightarrow \langle k^2 \rangle > 2\langle k \rangle$

Clustering Coefficient $\mathbb{C}$



Prob. $= \dfrac{k_i k_j}{2m}$

$$\mathbb{C} = \sum_{k_i, k_j=0}^{\infty} q_{k_i} q_{k_j} \frac{k_i k_j}{2m} = \frac{1}{2m}\left(\sum_{k=0}^{\infty} k q_k\right)^2$$

$$= \frac{1}{2m\langle k \rangle^2}\left[\sum_{k=0}^{\infty} k(k+1) p_{k+1}\right]^2 = \frac{1}{2m\langle k \rangle^2}\left[\sum_{k=0}^{\infty} (k-1) k p_k\right]^2$$

$n\langle k \rangle = 2m$

$$= \frac{1}{n} \frac{\left(\langle k^2 \rangle - \langle k \rangle\right)^2}{\langle k \rangle^3}$$

Random Nets: Poisson $\rightarrow \mathbb{C} \sim \frac{1}{n} \rightarrow 0$

Scale free Nets: Power-law degree distri.

$$\mathbb{C} \rightarrow \frac{\langle k^2 \rangle}{n} \rightarrow \text{constant}$$

if $\langle k^2 \rangle$ diverges ($\gamma < 3$)