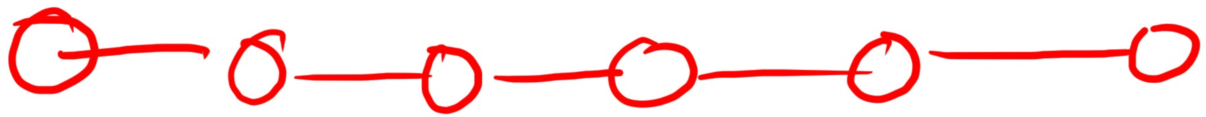
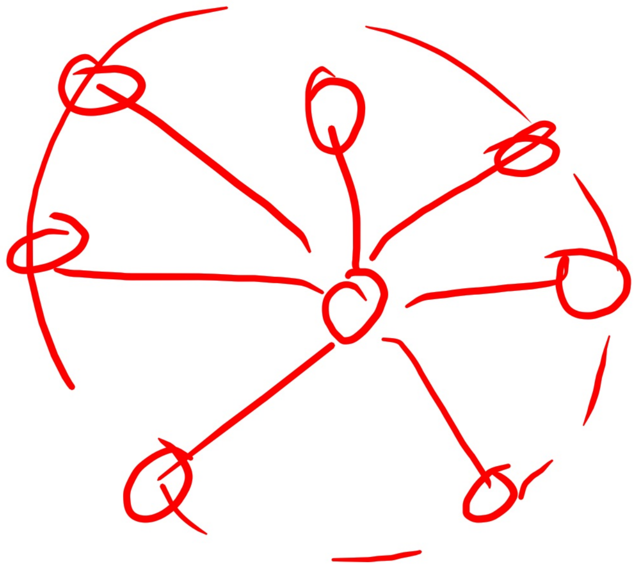
→ Networks: — Matrix Representation

Centrality ← node level description

Types

10^6 nodes ~ 7/8





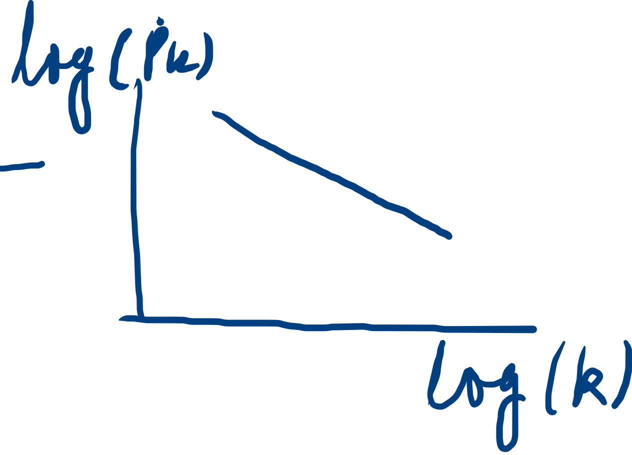
→ Group of nodes

Clustering Coeff.

Power-law

$$\boxed{p_k \sim k^{-\alpha}}$$

$$p_k = \frac{N_k}{N}$$



$$p_k = C k^{-\alpha}$$

Discrete

$$\sum_k p_k = 1$$

$$\sum_k k^{-\alpha} = \zeta(\alpha)$$

Riemman-Zeta fn.

Continuum

$$p(k) = C k^{-\alpha}$$

Normalization

$$\int p(k)\,dk = c \int_{k_{min}}^{\infty} k^{-\alpha}\,dk = C \left. \frac{k^{1-\alpha}}{1-\alpha} \right|_{k_{min}}^{\infty}$$

$$= \frac{C\, k_{min}^{1-\alpha}}{\alpha - 1} = 1$$

$$= \boxed{C = (\alpha - 1)\, k_{min}^{\alpha - 1}}$$

$$\langle k \rangle = \int_{k_{min}}^{k_{max}} k\, p(k)\, dk = C \int_{k_{min}}^{k_{max}} k\, k^{-\alpha}\, dk$$

$$= C \int_{k_{min}}^{k_{max}} k^{1-\alpha}\, dk$$

$$= C \left. \frac{k^{2-\alpha}}{2-\alpha} \right|_{k_{min}}^{k_{max}}$$

$$\langle k^2 \rangle = C \int_{k_{min}}^{k_{max}} k^2 \, p(k)\, dk = C \left. \frac{k^{3-\alpha}}{3-\alpha} \right|_{k_{min}}^{k_{max}}$$

$$\alpha \in [2, 3]$$

$$\langle k^r \rangle \sim C \left. \frac{k^{r+1-\alpha}}{r+1-\alpha} \right|_{k_{min}}^{k_{max}}$$

$$\langle k \rangle \pm \frac{\sigma_k}{\sqrt{n}}$$

$$p(ak) = a^{()}\, k^{-\alpha}$$

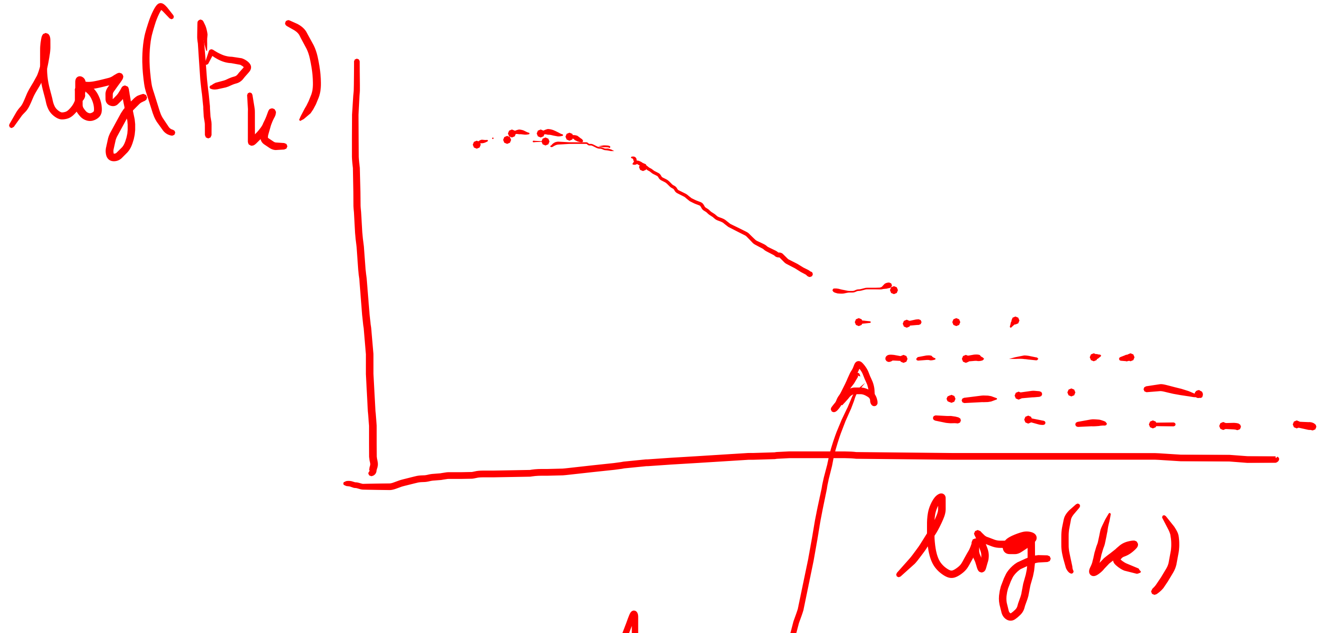
PARETO'S $\boxed{80:20}$

$\boxed{\text{FRACTALS}}$

→ Draw the distribution

$\frac{N_k}{N}$ ← $\frac{\text{\# of nodes with degree } k}{\text{total \# of nodes}}$ $k = 1, \ldots k_{max}$

Clauset Newman (2004)



linear binning

logarithmic binning

$b_0 = 1$

$b_1 = 2, 3$

$b_2 = 4, 5, 6, 7$ — — — —

CDF

Rank-frequency Method.

		r_i/n
10	1	
9	2	
8	3	
8	4	
6	5	
3	6	
2	7	

$$\frac{\int_{x_{1/2}}^{\infty} p(x)\,dx}{\int_{x_{min}}^{\infty} p(x)\,dx} = \frac{1}{2}$$

$$\frac{\int_{x_{1/2}}^{\infty} x\,p(x)\,dx}{\int_{x_{min}}^{\infty} x\,p(x)\,dx} = 2^{-\left(\frac{\alpha-2}{\alpha-1}\right)}$$

$$k_{max} \sim f(N) \qquad p(k) = C e^{-k}$$

$$\int_{k_{max}}^{\infty} p(k)\, dk = \frac{1}{N}.$$

$$C = (\alpha-1)\, k_{min}^{\alpha-1}$$

$$C \int_{k_{max}}^{\infty} k^{-\alpha}\, dk = C\, \frac{k_{max}^{1-\alpha}}{\alpha-1}$$

$$= k_{min}^{\alpha-1} \cdot k_{max}^{1-\alpha} = \frac{1}{N}.$$

$$k_{max} = k_{min} N^{-\frac{1}{1-\alpha}}$$

$$= k_{min} N^{\frac{1}{\alpha-1}}$$

→ hubs ⟶ power-law degree dist.

→ small world ⟶ short avg. path length

→ high clustering

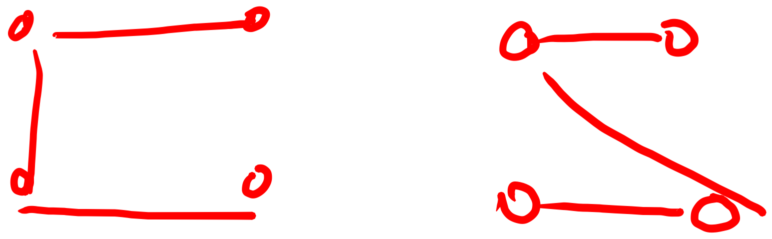
→ sparse networks

Rule

Random Graphs. (1959/1960)

Erdős-Rényi model

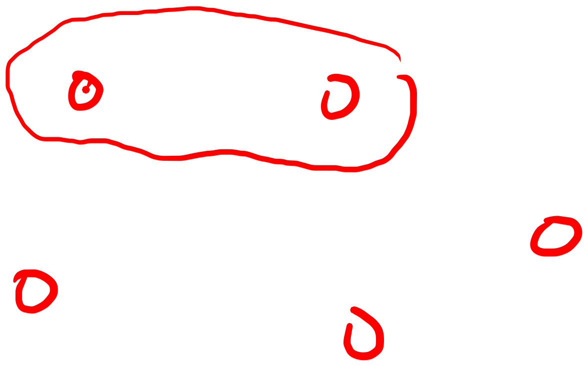
$G(n,m)$



Gilbert

$G(n,p)$

p



$G(n, p)$

$\llcorner$ prob. of connecting node pairs

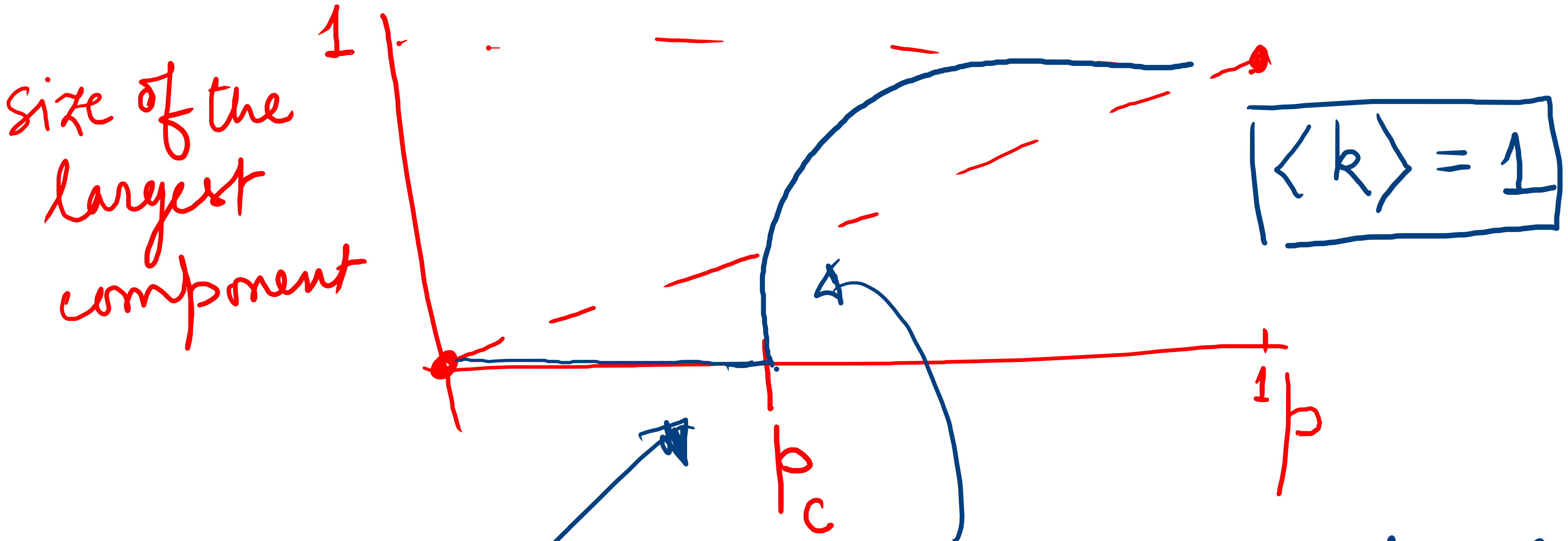
Average number of links.

$$\langle m \rangle = \binom{n}{2} p$$

<u>Average degree</u>

$$\langle k \rangle = \frac{2\langle m \rangle}{n} = (n-1)p \simeq np$$

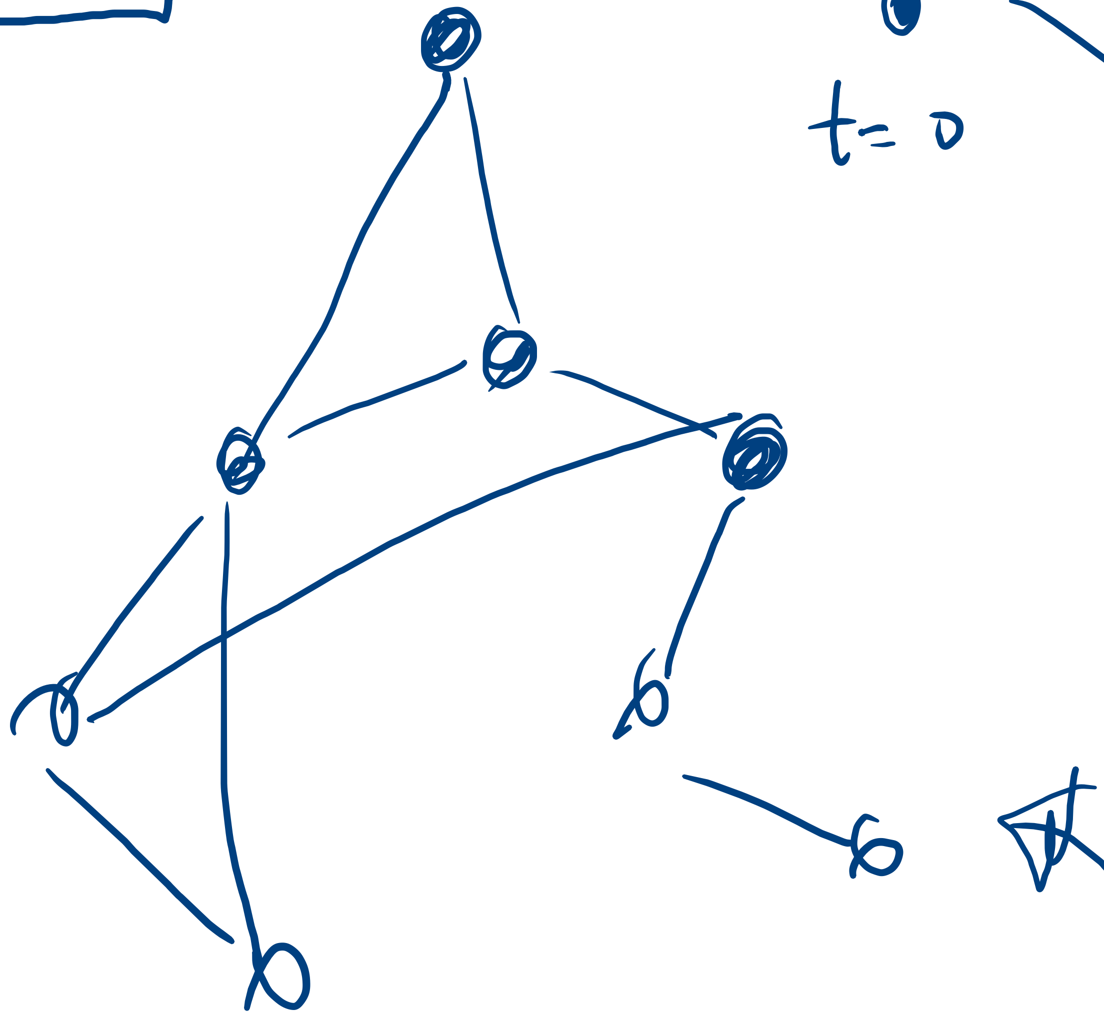
<u>Density</u> $$\rho = p = \frac{\langle m \rangle}{\binom{n}{2}}$$



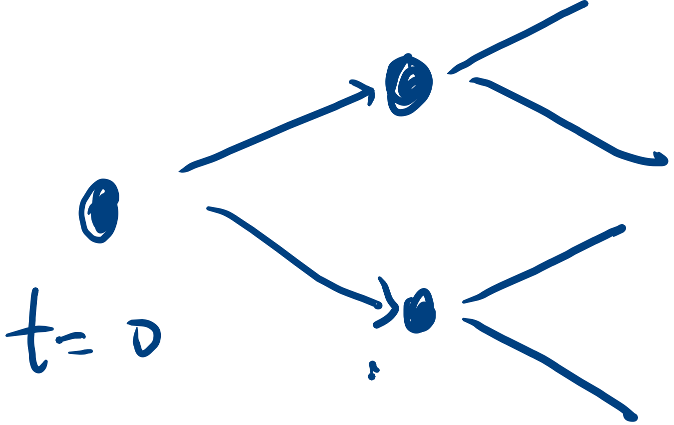
Phase Transition
Bifurcation

Emergence of a Giant component.

$\boxed{C = p}$



Tree



$$\langle k \rangle \simeq 1$$

$$\langle k \rangle \simeq np$$

$$p = \frac{\langle k \rangle}{n}$$

$$p \ll 1$$

Average path length

$$n = 1 + \langle k\rangle + \langle k\rangle^2 + \dots + \langle k\rangle^d$$

$$= \frac{\langle k\rangle^{d+1} - 1}{\langle k\rangle - 1}$$

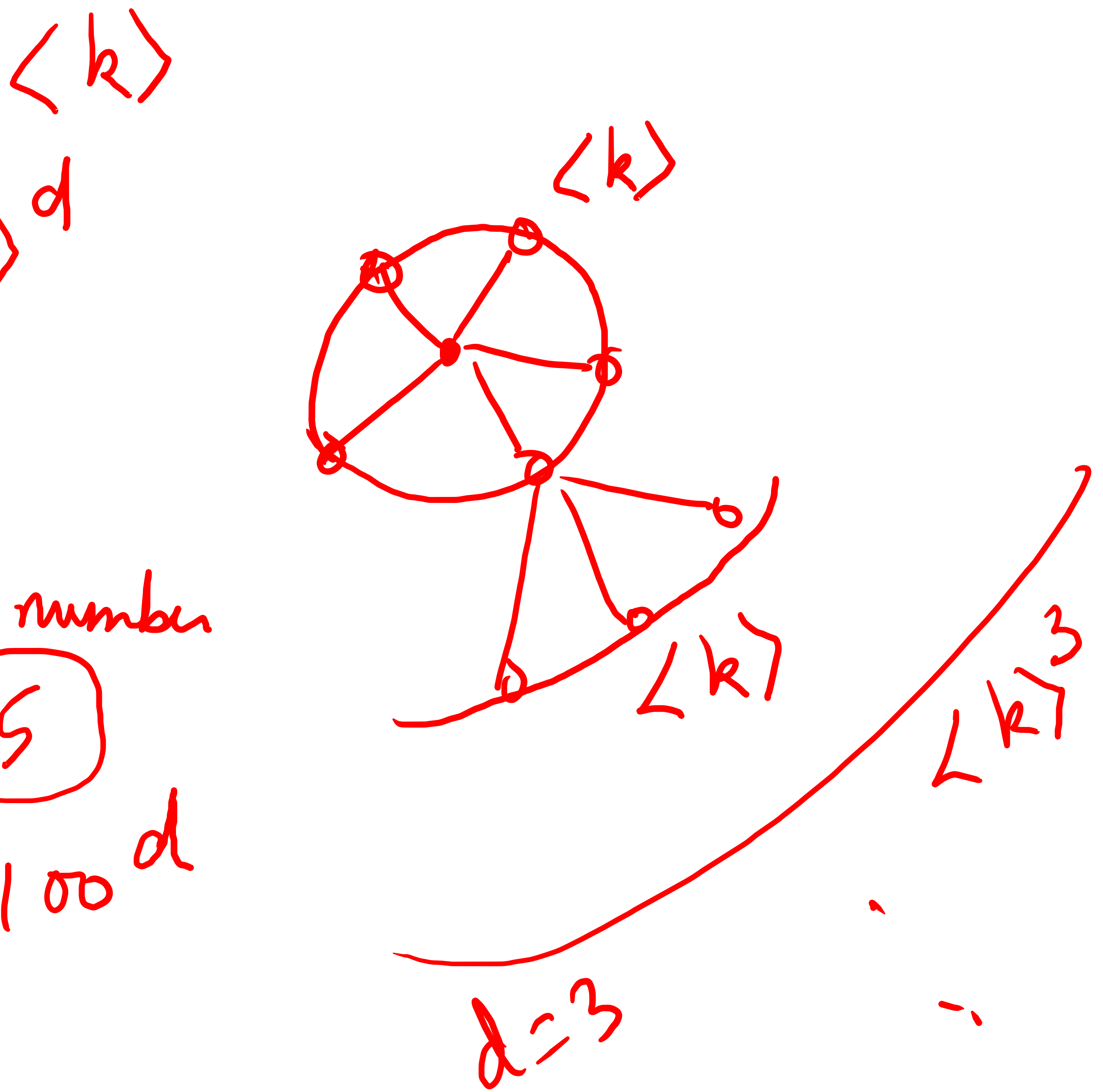
$$\simeq \langle k\rangle^d$$

$$d \simeq \frac{\ln n}{\ln \langle k\rangle}$$

Dunbar's number

75

$$n = 100^d$$



Degree dist

$$P(k) = p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

Binomial Dist.

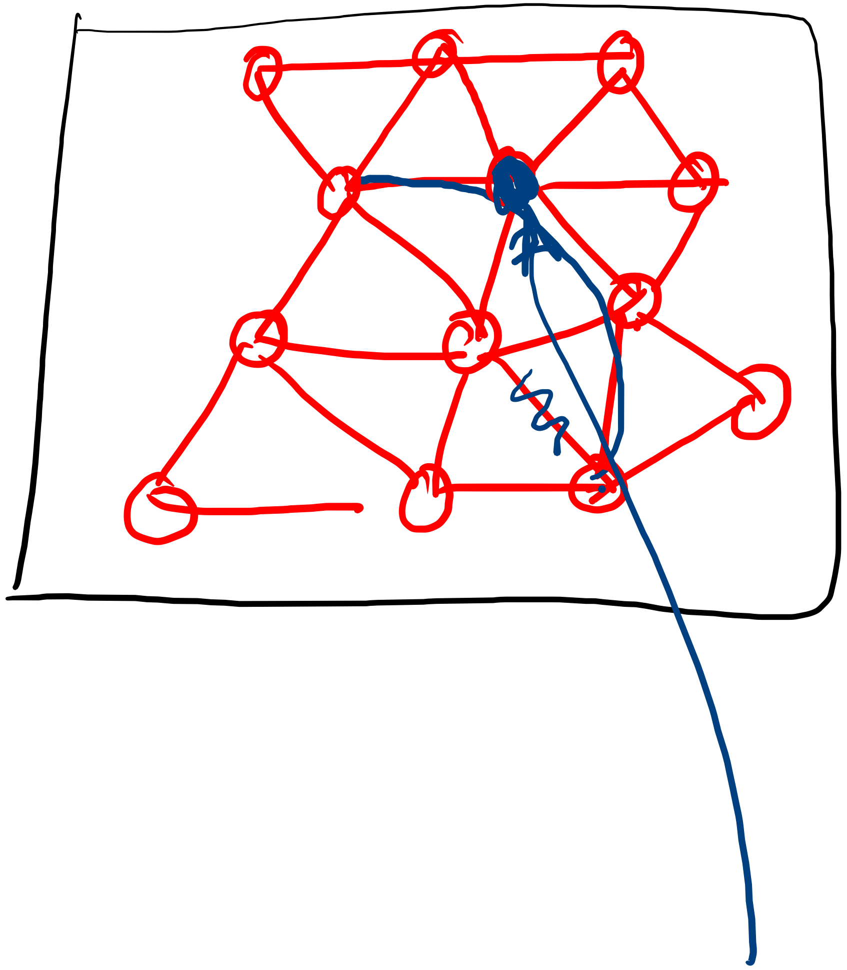
$\langle k \rangle = np$

$\lim n \to \infty$

$$\simeq \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}$$

mean $= \langle k \rangle$

variance $= \langle k \rangle$



Small World network

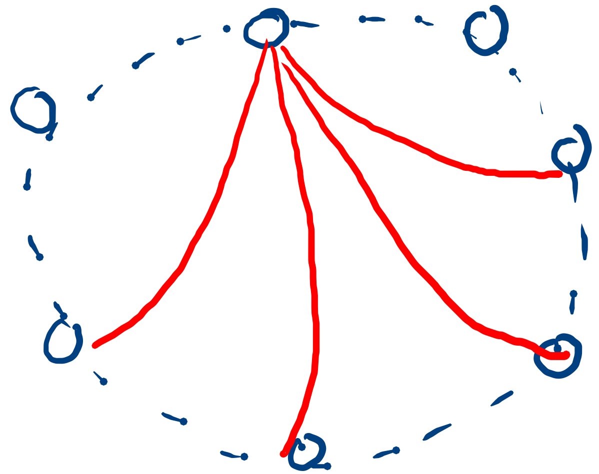
Collective phenomena . . . . .

Science

Watts & Strogatz

$$\frac{6}{\binom{6}{2}} = \frac{6}{15} = \boxed{0.4}$$

$$\boxed{\frac{\langle k \rangle}{n-1}} \neq \boxed{p = C_{RG}}$$

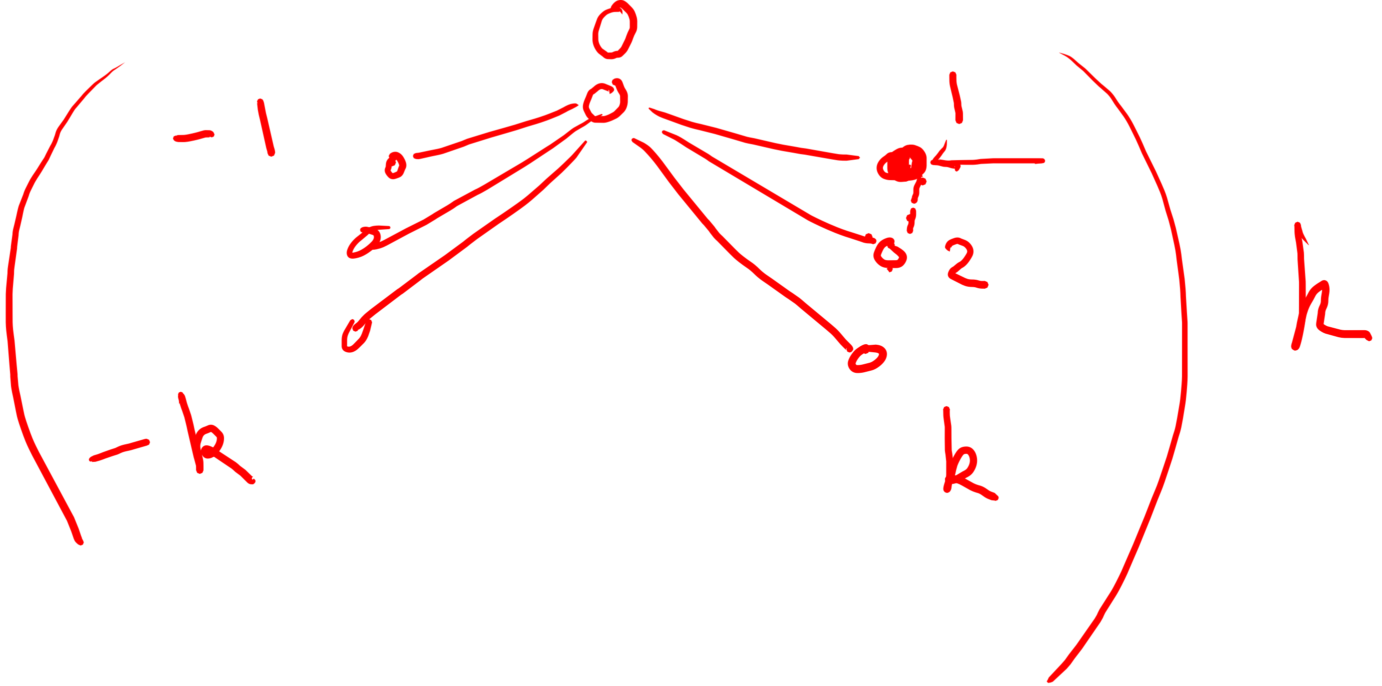


Ring graph of n nodes with each vertex having degree z

$$\boxed{z = 2k}$$

Number of pair of neighbours $\binom{2k}{2}$

$$(k+1)(k-1)$$



$$(k-2)+(k-3)+\cdots+0$$

$$= \frac{(k-2)(k-1)}{2}$$

$$(k+1)(k-1) + \frac{(k-2)(k-1)}{2}$$

$$= \frac{(k-1)}{2}\left[2k + \cancel{2} + k - \cancel{2}\right]$$

$$= \frac{3k(k-1)}{2}$$

$$C = \frac{\frac{3\cancel{k}(k-1)}{2}}{\frac{\cancel{2}\cancel{k}(2k-1)}{\cancel{2}}} = \frac{3(k-1)}{2(2k-1)}$$

$$= \frac{3\left(\frac{z}{2}-1\right)}{2(z-1)} = \frac{3}{4}\frac{(z-2)}{(z-1)}$$

$C(p)/C(0)$

Data $\longrightarrow$ Network

Network level studies

Distributions

Averages

Density

Why?

$\longrightarrow$ Models

G(n, p)
Random Graph

Watts - Strogatz

Barabasi Albert Model

node level properties

Centrality Measures

Group of nodes

cliques

k-plex

k-clan

k-core

k-components

$\boxed{\text{Mixing.}}$

Mixing — Assortative / Disassortative

Homophily

$$Q = \frac{m_c - \langle m_c \rangle}{m}$$

Modularity

$Q = 0$ random

$Q > 0$ assortative

$Q < 0$ disassortative

m_c : number of links between nodes that belong to the same class.

$$\frac{1}{2} \sum_{ij} A_{ij} \, \delta_{c_i, c_j} \qquad c_i : i = 1, \ldots k$$

$$\langle m_c \rangle = \frac{1}{2} \sum_{ij} \frac{k_i k_j}{2m} \delta_{c_i, c_j}$$

$$\boxed{Q = \frac{1}{2m} \sum_{ij} \underbrace{\left(A_{ij} - \frac{k_i k_j}{2m} \right)}_{\text{Modularity matrix}} \delta_{c_i, c_j}}$$

e_r : fraction of edges that link nodes of type r

		Women R_1	R_2
Men	R_1	0.4	0.1
	R_2	0.2	0.3

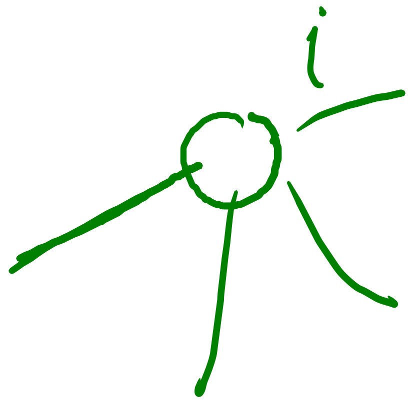
$$\delta_{c_i, c_j} = \sum_r \delta_{c_i, r} \, \delta_{c_j, r}$$

$$\frac{1}{2m} \sum_{ij} A_{ij} \, \delta_{c_i, c_j} = \frac{1}{2m} \sum_{ij} A_{ij} \sum_r \delta_{c_i, r} \, \delta_{c_j, r}$$

$$= \sum_r \underbrace{\frac{1}{2m} \sum_{ij} A_{ij} \, \delta_{c_i, r} \, \delta_{c_j, r}}_{e_r}$$

a_r = fraction of edges that link to a node of type r

$$\frac{1}{2m}\sum_i k_i \, \delta_{c_i,r}$$



term 2 in Q.

$$\frac{1}{2m}\sum_{ij}\frac{k_i k_j}{2m}\,\delta_{c_i,c_j} = \frac{1}{2m}\sum_{i,j}\frac{k_i k_j}{2m}\sum_r \delta_{c_i,r}\,\delta_{c_j,r}$$

$$= \sum_r \frac{1}{2m}\sum_i k_i\,\delta_{c_i,r}\;\frac{1}{2m}\sum_j k_j\,\delta_{c_j,r} = \sum_r a_r^2$$

$$Q = \sum_r \left(e_r - a_r^2\right)$$