WEEK 5 Lecture 1

CHAPTER 4
RANDOM GRAPHS

4.1 INTRODUCTION

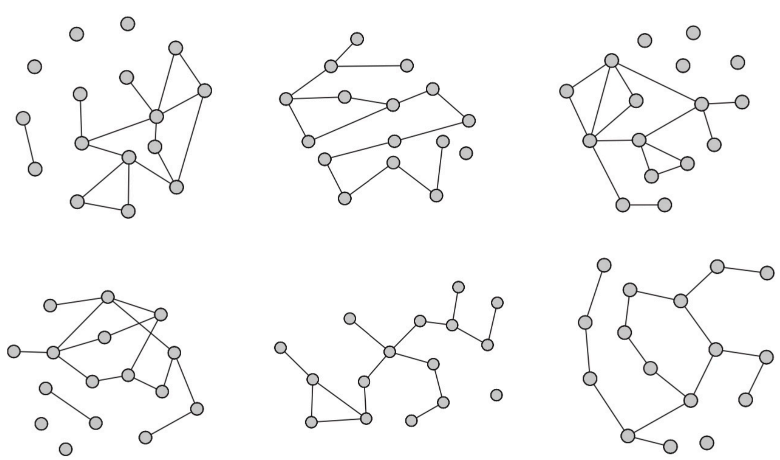
Real-world networks VS RANDOM GRAPHS with same N, L and disordered arrangement of links

2 pioneering papers by ERDŐS and RENYI in 1959, 1960
"On the evolution of random graphs"

PROBABILITY + GRAPH THEORY

IDEA: To study the properties of graphs a functions of the increasing # of random connections

ENSEMBLE of GRAPHS



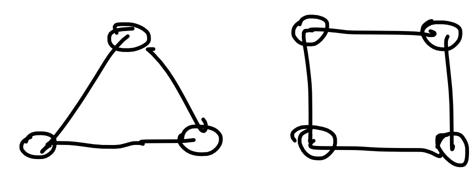
Six different realisations of model A with $N = 16$ and $K = 15$.

$N = 16$ $\quad L = 15$

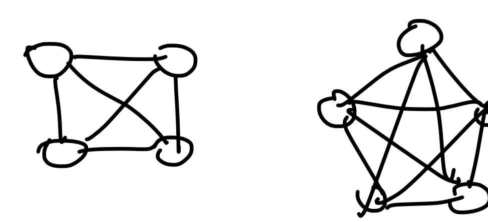
→ the average size of the LARGEST COMPONENT

→ the average # of subgraphs :

CYCLES



CLIQUES



4.2 RANDOM GRAPH ENSEMBLES

$G(N, L)$ ensemble considers every simple network with N nodes and L links with equal probability

DEF $G(N, L)$ ← fix L

ENSEMBLE A

The $\mathcal{G}(N,L)$ ensemble assigns to each simple network $G=(V,E)$ a probability

$$P(G) = \begin{cases} \frac{1}{Z} & \text{if } |V|=N \text{ and } |E|=L \\ 0 & \text{otherwise} \end{cases}$$

Z = # of simple networks with N nodes and L links

– The max # of links M in a simple network is

$$M = \binom{N}{2} = \frac{N(N-1)}{2}$$

binomial coefficient $C_n^k = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

– Hence Z is given by the # of ways in which we can choose L links out of M possibilities:

$$Z = \binom{M}{L} = \binom{\frac{N(N-1)}{2}}{L}$$

ENSEMBLE B

$\mathbb{G}(N, p)$ ensemble considers every simple network with N nodes obtained by connecting each pair of nodes with a probability p

DEF $\mathbb{G}(N, p)$ fix p : $0 \leq p \leq 1$

The $\mathbb{G}(N,p)$ ensemble assigns to each simple network $G=(V,E)$ with $|V|=N$ nodes a probability

$$\mathbb{P}(G) = p^{L} (1-p)^{\frac{N(N-1)}{2} - L} \quad \text{where } L = |E|$$

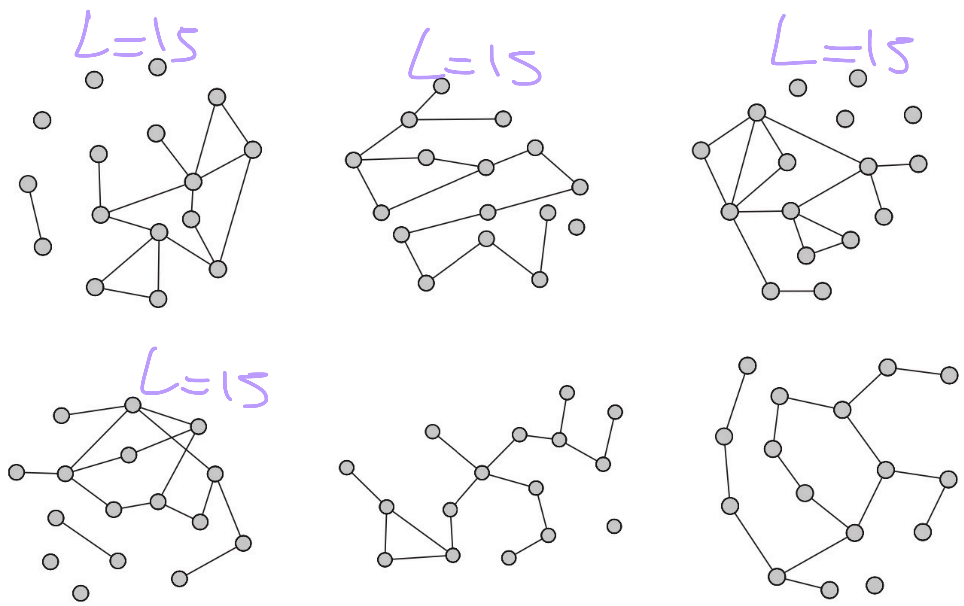
Each network in $\mathbb{G}(N,p)$ has N fixed but L can take different values, and it can be seen as the result of $M = \frac{N(N-1)}{2}$ indipendent coin tossings, one for each pair of nodes, with success probability (of drawing a link)

probability that L pairs of nodes are linked

probability that the remaining $M-L$ pairs are not linked

4

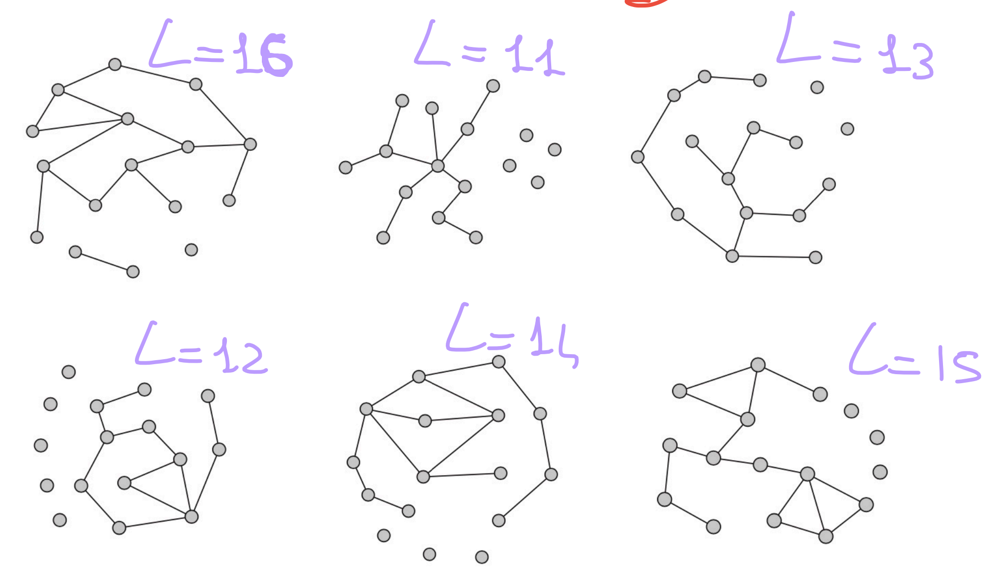
$G(N,L)$ ENSEMBLE A



Six different realisations of model A with $N = 16$ and $K = 15$.

$N=16$ $L=15$

$G(N,p)$ ENSEMBLE B



Six different realisations of model B, with $N = 16$ and $p = 0.125$.

$N=16$ $p=0.125$

$$p \cdot \frac{N(N-1)}{2} = 15$$

$G(N,L)$ and $G(N,p)$ are ASYMPTOTICALLY EQUIVALENT

For $L = p\,\dfrac{N(N-1)}{2}$ and $N \to \infty$ then the 2 ensembles $G(N,L)$ and $G(N,p)$ have the same statistical properties

Hence we focus on $G(N,p)$

4.3 DISTRIBUTION P_L

PROPOSITION

The probability P_L that a network in $G(N,p)$ has L links follows the BINOMIAL DISTRIBUTION

M

$$P_L = \binom{\frac{N(N-1)}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2} - L}$$

i.e. $L \sim \mathrm{Bin}\left(\frac{N(N-1)}{2}, p\right)$

BINOMIAL DISTRIBUTION

Proof

probability that L pairs of nodes are linked

probability that the remaining pairs are NOT linked

# of ways we can choose the L linked pairs of nodes out of the M possibilities

Since $L \sim Bin\left(\frac{N(N-1)}{2}, p\right)$

$$\langle L \rangle = \frac{N(N-1)}{2} \cdot p$$

average # of links in a network of the ensemble $G(N,p)$

4.4 DEGREE DISTRIBUTION

PROPOSITION

The degree distribution $P(k)$ of the $G(N,p)$ ensemble is binomial

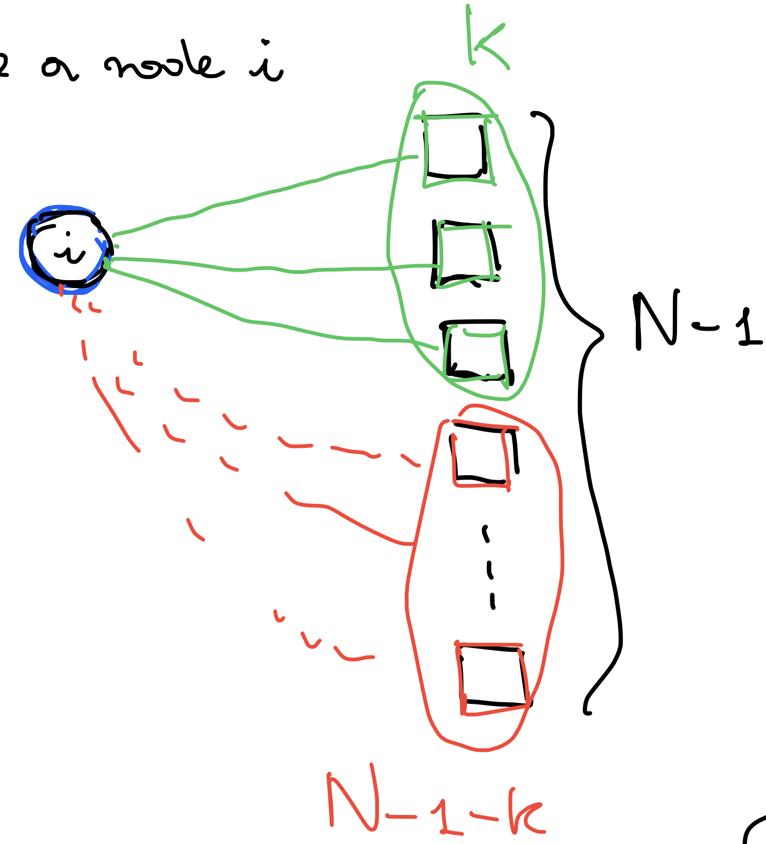
$$P_B(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

i.e. $k \sim Bin(N-1, p)$

C^k_{N-1}

probability of linking two nodes

Proof: Take a node i



p^k is the probability of the existence of k links

$(1-p)^{N-1-k}$ is the probability for the absence of the remaining links

$$C^k_{N-1} = \binom{N-1}{k} = \# \text{ of different ways}$$

of selecting the end points of the k links among the $N-1$ possibilities

GENERATING FUNCTIONS

DEF

Given a degree distribution $P(k)$, its GENERATING FUNCTION $G(x)$ is defined by

$$G(x) = \sum_{k=0}^{\infty} P(k)\, x^k$$

PROPOSITION

The generating function $G(x)$ of $P(K)$ has the following properties:

① $G(1) = 1$

② $G^{[m]}(1) \equiv \left.\frac{d^m G(x)}{dx^m}\right|_{x=1} = \langle k(k-1)\dots(k-m+1)\rangle$

m-th FACTORIAL MOMENT of $P(k)$

$\langle k\rangle = G'(1)$

$\langle k(k-1)\rangle = G''(1) \equiv G^{[2]}(1)$

$$G(x) = \sum_{k=0}^{\infty} P(k)\, x^k$$

Proof

① $G(1) = \sum_{k=0}^{\infty} P(k)\, 1^k = \sum_{k=0}^{\infty} P(k) = 1$

② $m=1$ $\quad \frac{dG(x)}{dx} = \frac{d}{dx}\sum_{k=0}^{\infty} P(k)\, x^k = \sum_{k=0}^{\infty} P(k)\, \frac{d}{dx} x^k =$

$$= \sum_{k=0}^{\infty} P(k)\, k\, x^{k-1}$$

$$\left.\frac{dG(x)}{dx}\right|_{x=1} = \sum_{k=0}^{\infty} P(k)\, k = \langle k \rangle$$

$m=2$ $$\frac{d^2G(x)}{dx^2} = \sum_{k=0}^{\infty} P(k) \frac{d^2}{dx^2} x^k = \sum_{k=0}^{\infty} P(k)\, k(k-1)\, x^{k-2}$$

$$\left.\frac{d^2G(x)}{dx^2}\right|_{x=1} = \sum_{k=0}^{\infty} P(k)\, k(k-1) = \langle k(k-1) \rangle$$

and so on

$\boxed{\text{BINOMIAL DISTRIBUTION}}$

$$P(k) = Bin(N-1, p) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

$k = 0, 1, \ldots, N-1$

$$G_B(x) = \sum_{k=0}^{\infty} P(k)\, x^k$$

$$= \sum_{k=0}^{N-1} \binom{N-1}{k} p^k x^k (1-p)^{N-1-k}$$

Recall the NEWTON BINOMIAL FORMULA

$$(a+b)^M = \sum_{k=0}^{M} \binom{M}{k} a^k b^{M-k}$$

We use it with $a = px$ $\quad b = 1-p \quad M = N-1$

$$\boxed{G_B(x) = (\underbrace{px}_{a} + \underbrace{1-p}_{b})^{\overbrace{N-1}^{M}}}$$

GENERATING FUNCTION of Binomial $P(k)$

We can now use property ② of generating functions to calculate the moments of $P(k)$

$$\langle k \rangle = G_B'(1) = (N-1)(px+1-p)^{N-2} \cdot p \Big|_{x=1} = (N-1)(\cancel{p}+1-\cancel{p})^{N-2} \cdot p$$

$$\boxed{\langle k \rangle = (N-1)\,p}$$

$$\langle k(k-1) \rangle = G_B^{[2]}(1) = (N-1)(N-2)(px+1-p)^{N-3} \cdot p^2 \Big|_{x=1} =$$

$$= (N-1)(N-2)\, p^2$$

$$\boxed{\langle k(k-1)\rangle = (N-1)(N-2)\, p^2}$$

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

$$\langle k\rangle = (N-1)\,p$$

$$\langle k^2\rangle = \langle k\rangle + \langle k^2 - k\rangle = \langle k\rangle + (N-1)(N-2)\,p^2$$

$$= (N-1)\,p + (N-1)(N-2)\,p^2 =$$

$$= (N-1)\,p\,\left[1 + (N-2)\,p\right]$$

VARIANCE $\sigma^2 \equiv \langle k^2\rangle - \langle k\rangle^2 = (N-1)\,p\,\left[1 + (\cancel{N}-2)\,p - (\cancel{N}-1)\,p\right]$

$$= (N-1)\,p\,\left[1 - 2p + p\right] = (N-1)\,p\,(1-p)$$

$$\boxed{\sigma^2 = (N-1)\,p\,(1-p)}$$

Analogously $\boxed{\langle k(k-1)\cdots(k-m+1)\rangle = \dfrac{(N-1)!}{(N-m-1)!}\, p^m}$

POISSON DISTRIBUTION

$$\boxed{P(k) = \text{Poiss}(c) = \frac{c^k e^{-c}}{k!} \qquad k = 0, 1, \ldots}$$

$$G_p(x) = \sum_{k=0}^{\infty} P(k)\, x^k = e^{-c} \sum_{k=0}^{\infty} \frac{x^k c^k}{k!}$$

RECALL the TAYLOR series of the exponential

$$e^{\alpha} = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!}$$

We use it with $\alpha = x \cdot c$ We get:

$$\boxed{G_p(x) = e^{-c} \cdot e^{cx}}$$

GENERATING FUNCTION of POISSON $P(k)$

We can calculate the MOMENTS of $\text{Poiss}(c)$

$$\langle k \rangle = G_p'(1) = e^{-c}\, c\, e^{cx}\Big|_{x=1} = \cancel{e^{-c}}\, c\, \cancel{e^{c}} = c$$

$$\boxed{\langle k \rangle = c}$$

$$\langle k(k-1) \rangle = G_p''(1) = e^{-c}\, c^2 e^{cx} \Big|_{x=1} = c^2$$

$$\boxed{\langle k(k-1) \rangle = c^2}$$

Analogously

$$\boxed{\langle k(k-1)\cdots(k-m+1) \rangle = c^m}$$

IMPORTANT: Please do not forget to submit QUIZ 2