We will analyze 4 complex networks models in this course.

name	mechanism
ER	randomness
BA	preferential attachment
WS	Small World
BE	Core-periphery

Today we look at Erdos-Renyi (ER) which has a randomness mechanism

It was originally published in 1960 (cold war)

All we have to know about ER is condensated in (Erdos-Renyi for dummies slide)

Erdos and Renyi were two mathematicians.

The idea is we have a set of nodes, and each couple of node is connected with a certain probability p.
We fixate the number of nodes, and we generate links in a total random ways.

The probability of the degree is a binomial distribution, that can then be approsimated to a Poissonian

The probability of finding a node with high degree is very very low (due to the exponential)

$$p(d) \simeq \frac{z^{d} \cdot e^{-z}}{d!}$$

$$z = p \cdot (n-1)$$

number of nodes -1

probability of having a link

$$\binom{M}{K} = \frac{M!}{K!\,(M-K)!}$$

We can have two ways of construct an ER model, let's see the type A:

(ER is symmetric, we don't consider direction in the links)

We fixate N (the number of nodes) and K the number of links:

$$M = \frac{N(N-1)}{2}$$ this is the maximum number of links with N nodes

$$0 \leq K \leq M$$

We cannot have K larger than M!

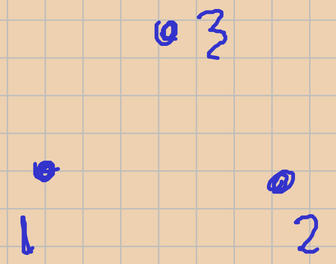
We then generate the graph G by generating the K links totally random between the nodes.

We have N and K, but we can have a certain number of graphs with those charasteristics:

$$C_K^M = \binom{M}{K} = \frac{M!}{K!\,(M-K)!}$$ this is the combination

Computationally: we select the number of nodes N and calculate M, then we compute all the M possible links (NO DIRECTIONALS, (a,b) is the same as (b,a) and only one is considered) in an array long M.
We choose K (less than M), so we choose randomly K values from that array and generate the graph this way!

Ex)



$$N=3 \Rightarrow M=\frac{3\cdot 2}{2}=3$$

If $K=2$

$$M=\underbrace{\{(1,3),(2,3),(1,3)\}}$$

we choose k(so 2) links from that list

By using the combinatorial formua, we can calculate how many different networks we can find

$$\binom{M}{K}=\frac{M!}{K!(M-K)!}=\frac{6}{2(1)}=3$$

The probability associated to generate a specific graph is:

$$\boxed{P_G = \frac{1}{C_M^K}} = \frac{1}{3} \quad \text{in our example.}$$

We can calculate the probability associated to taking a specific link

$$p = \frac{1}{M} = \frac{1}{3} \quad \text{in our example}$$

In this example P and p are the same, in general they are not.

$$\boxed{p \in [0,1]}$$

ER model B:
we fixate the number of nodes and select p, the probability to have each link!

p must be between 0 (no connections) and 1 (all nodes connected)

There must be a link (no pun intended) between a ER generated with A model and B model

$$G^{ER}_{N,K} \qquad\qquad G^{ER}_{N,p}$$

ER model A: we select N and K

ER model B: we select N and p

With model B, the probability of exctracting one single network from all the possibles is the "bearnoulli trial", so the prob. to obtain a network with N nodes and K links is:

$$\underbrace{P^{N,K}_{G}} = p^{K}(1-p)^{M-K} \cdot \binom{M}{K}$$

Probability of obtaining a graph with N nodes and K links

The average expacted value of links K is

$$\boxed{K=pM}$$

The average of the bernoulli distribution.

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

This is the variance associated with the average value K expected

The two models behave similarly when $p = \frac{K}{M}$

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

this is the probability of having k successes in n tries

$p =$ p=probability of success (to get a certain face in a dice is 1/6)
k=how many success we want (we want to analyze the prob. of having k succ.)
n=the number of tries we have

We can write the probability of a specific node i to have degree k

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

here k is how many success we want!

This is the probability that the i node has degree k.

We use N-1 (number of nodes -1) because we have as much "tries" as the remaining nodes (we can't have a self link!)
p is the probability of success because it's the prob. of having a link! (link -> degree)

$$\boxed{\begin{aligned} \langle k \rangle &= p(N-1) \\ \sigma_k &= \sqrt{p(1-p)(N-1)} \end{aligned}}$$

this is the average degree of each node

this is the standard deviation associated with a degree having k degree

Those two are interesting to compute!

- In the limits N->00 and p->0 that binomial distribution is approximated to a poissonian!

$$P_k = \frac{e^{-z} z^k}{k!} \qquad z = Np$$

$$\Rightarrow \begin{cases} \langle k \rangle = pN \\ \sigma_k = \sqrt{pN} \end{cases}$$

the std. dev. is the square root of the mean z

Becouse $N \approx N-1$ if $N \to \infty$

Giant component: it's a large conn. component, is a connected component whose size scales linearly with the number of nodes N: if we have a larger network -> we have a larger giant component.

If we increase p, more nodes are connected.
Is there a value of p where every node is connected (NB: not every node connected with every node, that is just p=1)

- If we have a ER graph generated (model B), we can write the probability of having a link as: $p = \frac{c}{N}$, N -> number of nodes, c = constant

If we want to know the size S_1 of the largest component

$$\begin{cases} \text{if } c < 1 \rightarrow S_1 \propto \ln N \\ \text{if } c = 1 \rightarrow S_1 \propto N^{2/3} \\ \text{if } c > 1 \rightarrow S_1 \propto N \end{cases}$$

if $c>1$ $S_2 \propto \ln N$

the size of the second largest component

- If $c>1$ the largest component is a GIANT COMPONENT! (the size scales linearly with N), the second largest scales with ln(N)

  If $c>1$ we have a giant component!

It is known that:

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

The average degree is equal to c!!!
So if the average degree is larger than 1, the largest component is a giant component

Since $\langle k \rangle = pN \Rightarrow \langle k \rangle > 1 \Rightarrow pN > 1 \Rightarrow \boxed{p > \frac{1}{N}}$

this is the condition on the probability for having the largest component as a giant component.