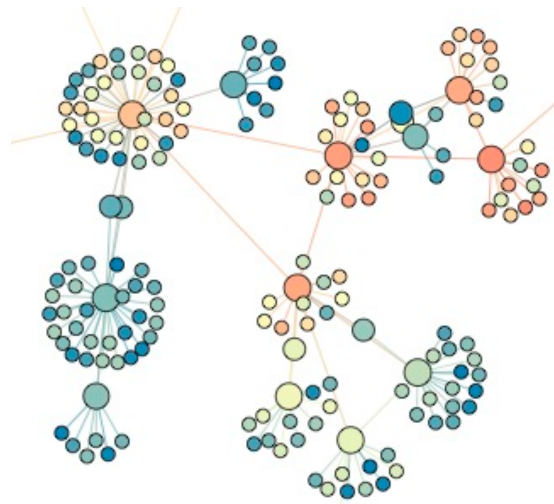


SNA 2A: Intro to Random Graphs

Lada Adamic

school of information

Network models

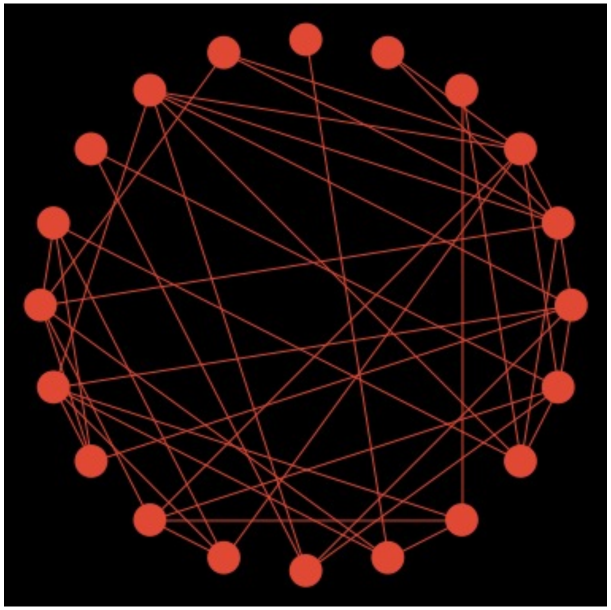
- Why model?
  - simple representation of complex network
  - can derive properties mathematically
  - predict properties and outcomes
- Also: to have a strawman
  - In what ways is your real-world network different from hypothesized model?
  - What insights can be gleaned from this?

Erdös and Rényi

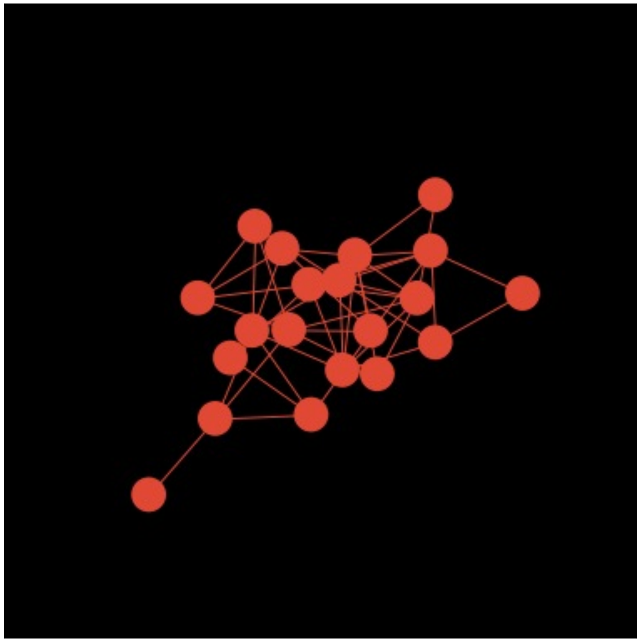
Erdös-Renyi: simplest network model

- Assumptions
  - nodes connect at random
  - network is undirected
- Key parameter (besides number of nodes N) : p or M
  - p = probability that any two nodes share and edge
  - M = total number of edges in the graph

what they look like



after spring layout



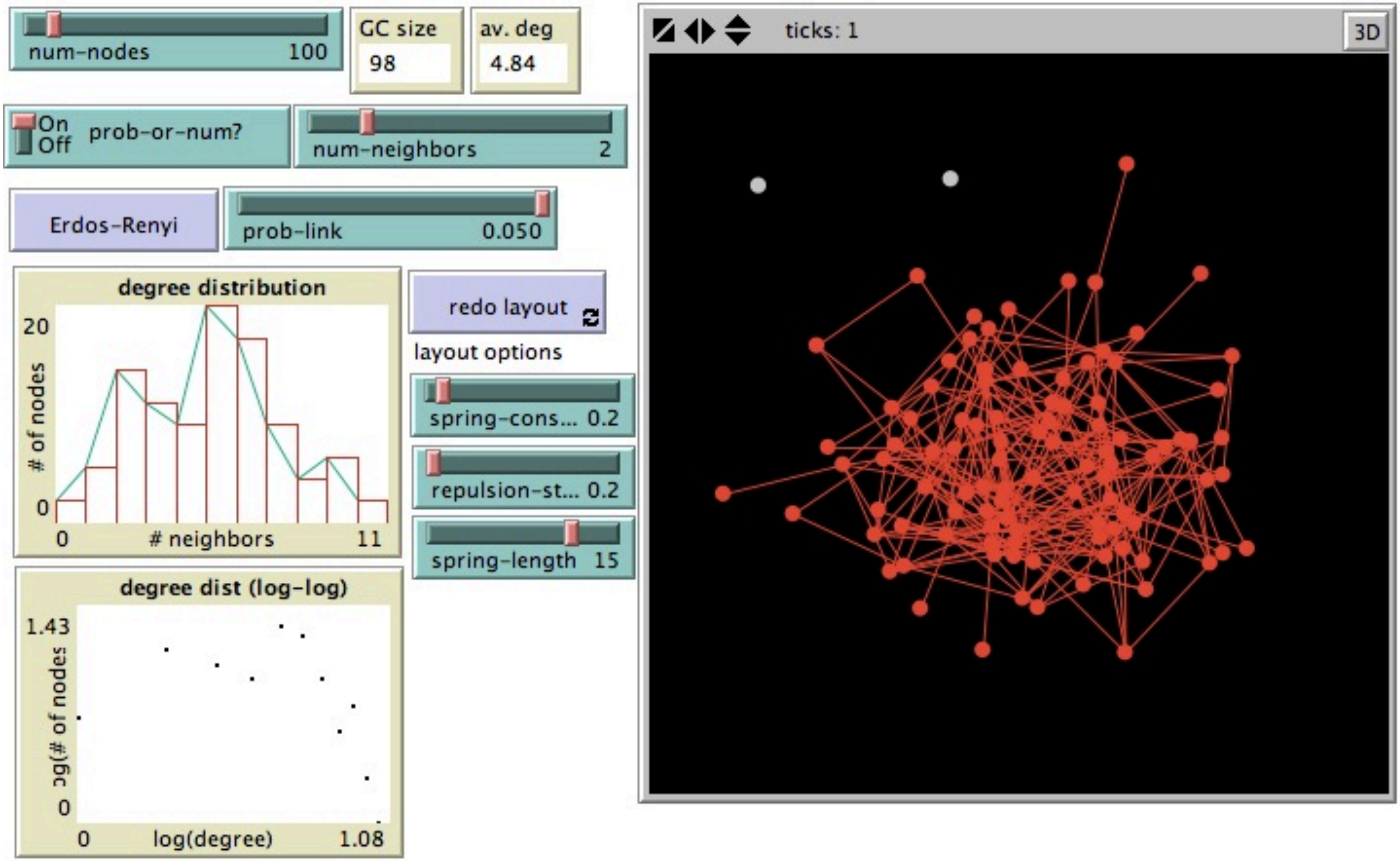
Degree distribution

- (N,p)-model: For each potential edge we flip a biased coin
  - with probability p we add the edge
  - with probability (1-p) we don't

Quiz Q:

- As the size of the network increases, if you keep ***p***, the probability of any two nodes being connected, the same, what happens to the average degree
  - a) stays the same
  - b) increases
  - c) decreases

http://ladamic.com/netlearn/NetLogo501/ErdosRenyiDegDist.html



http://www.ladamic.com/netlearn/NetLogo501/ErdosRenyiDegDist.html

Degree distribution

- What is the probability that a node has 0,1,2,3… edges?
- Probabilities sum to 1

How many edges per node?

- Each node has (N – 1) tries to get edges
- Each try is a success with probability p
- The binomial distribution gives us the probability that a node has degree k:

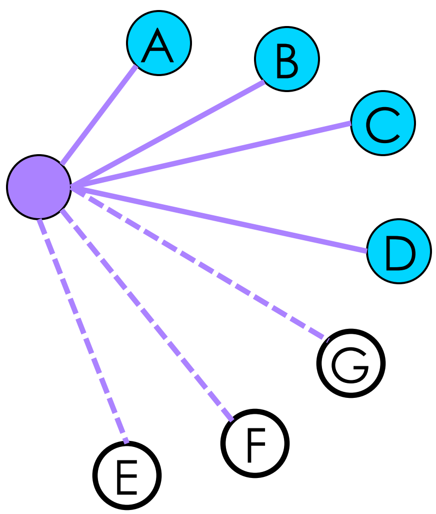
$$B(N-1;k;p) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Quiz Q:

- The maximum degree of a node in a simple (no multiple edges between the same two nodes) N node graph is
  - a) N
  - b) N - 1
  - c) N / 2

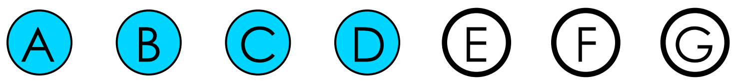
Explaining the binomial distribution

- 8 node graph, probability ***p*** of any two nodes sharing an edge
- What is the probability that a given node has degree 4?

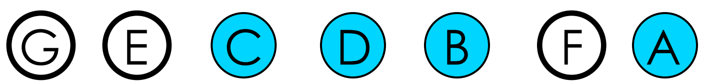


Binomial coefficient: choosing 4 out of 7

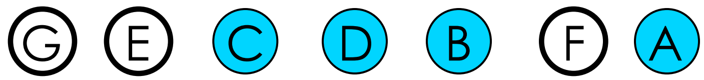
Suppose I have 7 blue and white nodes, each of them uniquely marked so that I can distinguish them. The blue nodes are ones I share an edge with, the white ones I don't.



How many different samples can I draw containing the same nodes but in a different order (the order could be e.g. the order in which the edges are added (or not)? e.g.



binomial coefficient explained



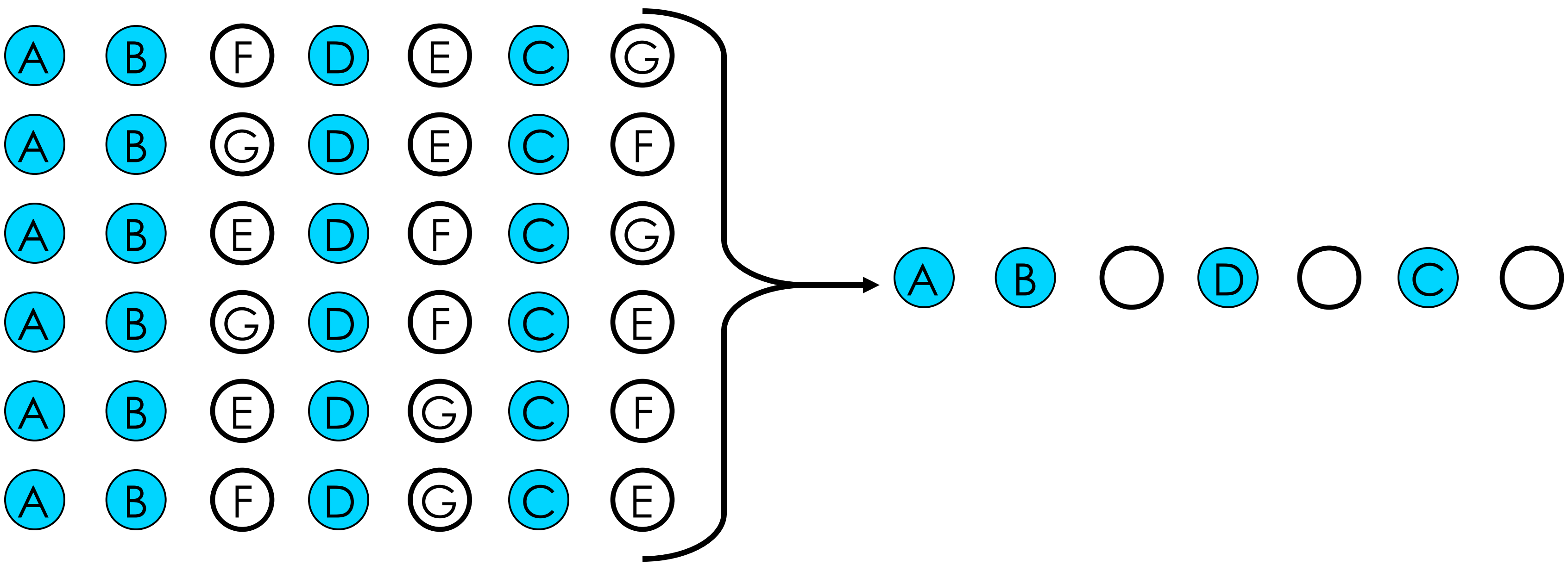
If order matters, there are **7!** different orderings:

I have 7 choices for the first spot, 6 choices for the second (since I've picked 1 and now have only 6 to choose from),
5 choices for the third, etc.

7! = 7 * 6 * 5 * 4 * 3 * 2 * 1

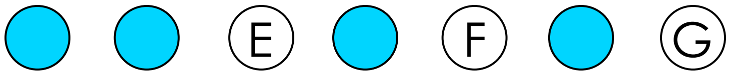
binomial coefficient

Suppose the order of the nodes I don't connect to (white) doesn't matter.
All possible arrangements (3!) of white nodes look the same to me.



Instead of 7! combinations, we have 7!/3! combinations

binomial coefficient explained



The same goes for the blue nodes, if we can’t tell them apart, we lose a factor of 4!

binomial coefficient explained

number of ways of choosing k items out of (n-1)

$$= \frac{\text{number of ways of arranging } \boldsymbol{n-1} \text{ items}}{(\text{\# of ways to arrange } \boldsymbol{k} \text{ things})*(\text{\# ways to arrange } \boldsymbol{n}-1-\boldsymbol{k} \text{ things})}$$

$$= \frac{\boldsymbol{n-1}!}{\boldsymbol{k}!\,(\boldsymbol{n}-1-\boldsymbol{k})!}$$

Note that the binomial coefficient is symmetric – there are the same number of ways of choosing ***k*** or ***n-1-k*** things out of ***n-1***

Quiz Q:

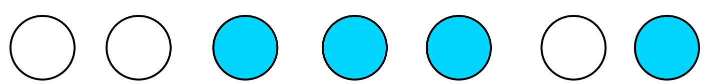
- What is the number of ways of choosing 2 items out of 5?
  - 10
  - 120
  - 6
  - 5

Now the distribution

- p = probability of having edge to node (blue)
- (1-p) = probability of not having edge (white)
- The probability that you connect to 4 of the 7 nodes in some particular order (two white followed by 3 blues, followed by a white followed by a blue) is

P(white)*P(white)*P(blue)*P(blue)*P(blue)*P(white)*P(blue)

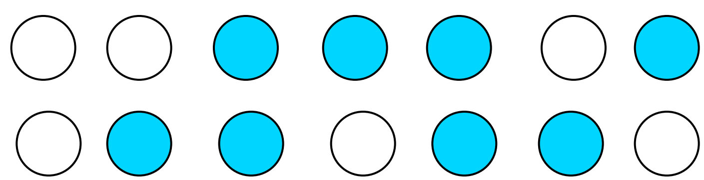
$= p^4 * (1-p)^3$



Binomial distribution

- If order doesn't matter, need to multiply probability of any given arrangement by number of such arrangements:

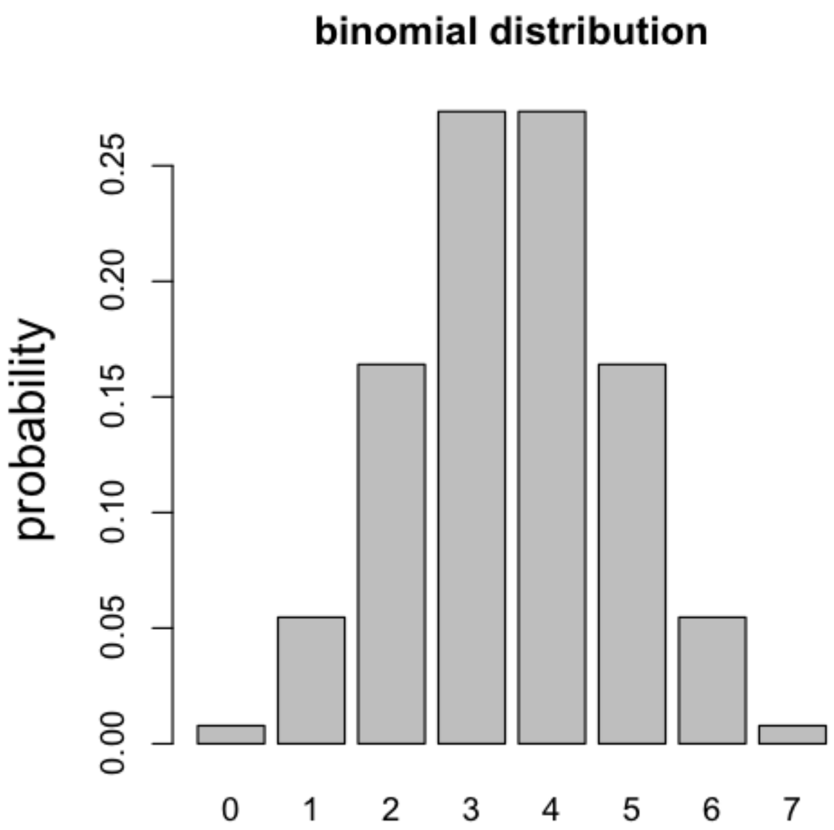
$$B(7;4;p) = \binom{7}{4} p^4 (1-p)^3$$



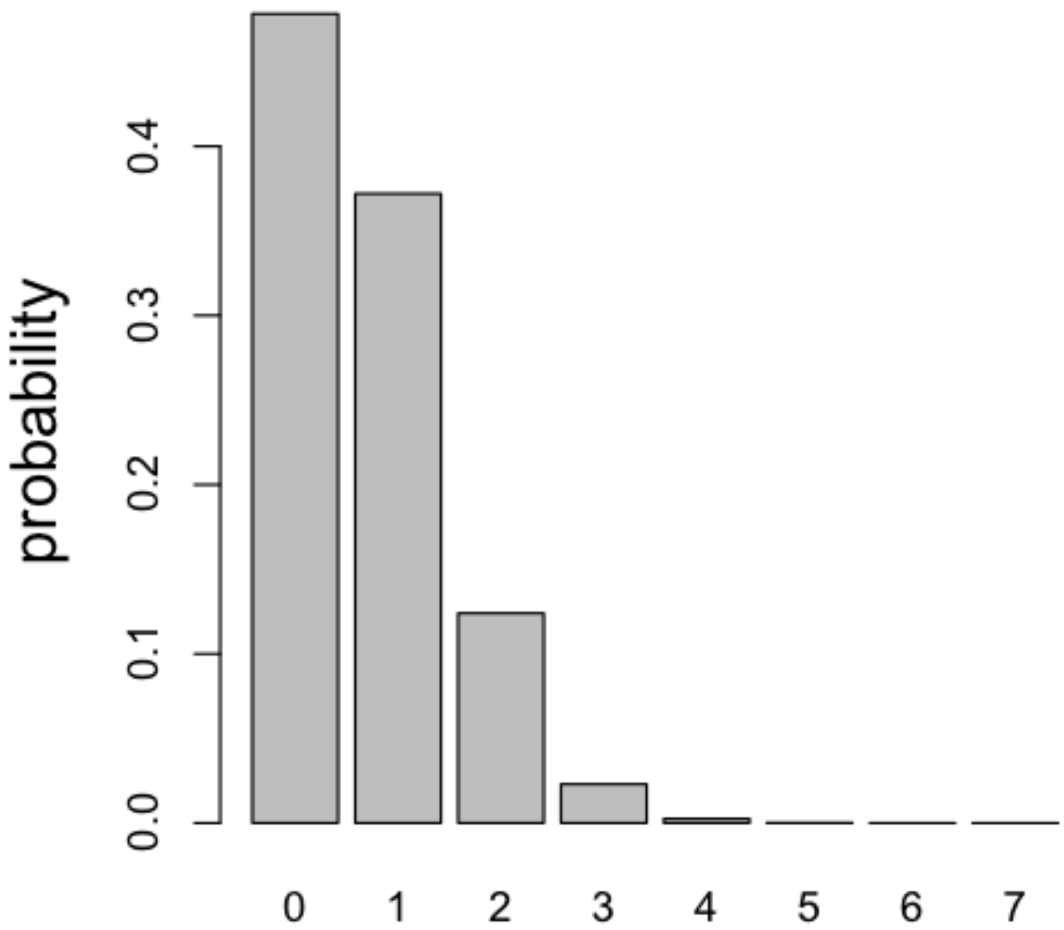
+

….

if p = 0.5

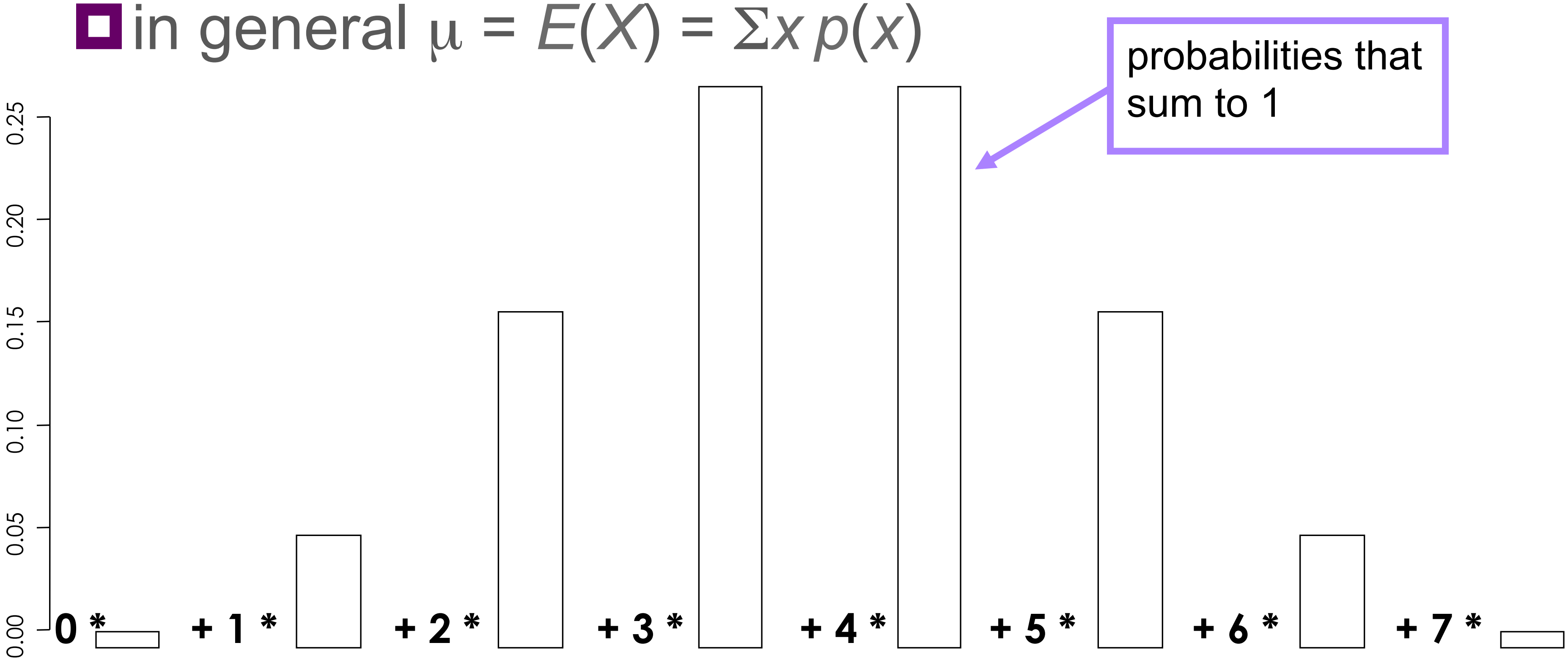

binomial distribution
probability
0.00
0.05
0.10
0.15
0.20
0.25
0
1
2
3
4
5
6
7

p = 0.1



What is the mean?

- Average degree **z = (n-1)*p**
- in general $\mu = E(X) = \Sigma x\, p(x)$



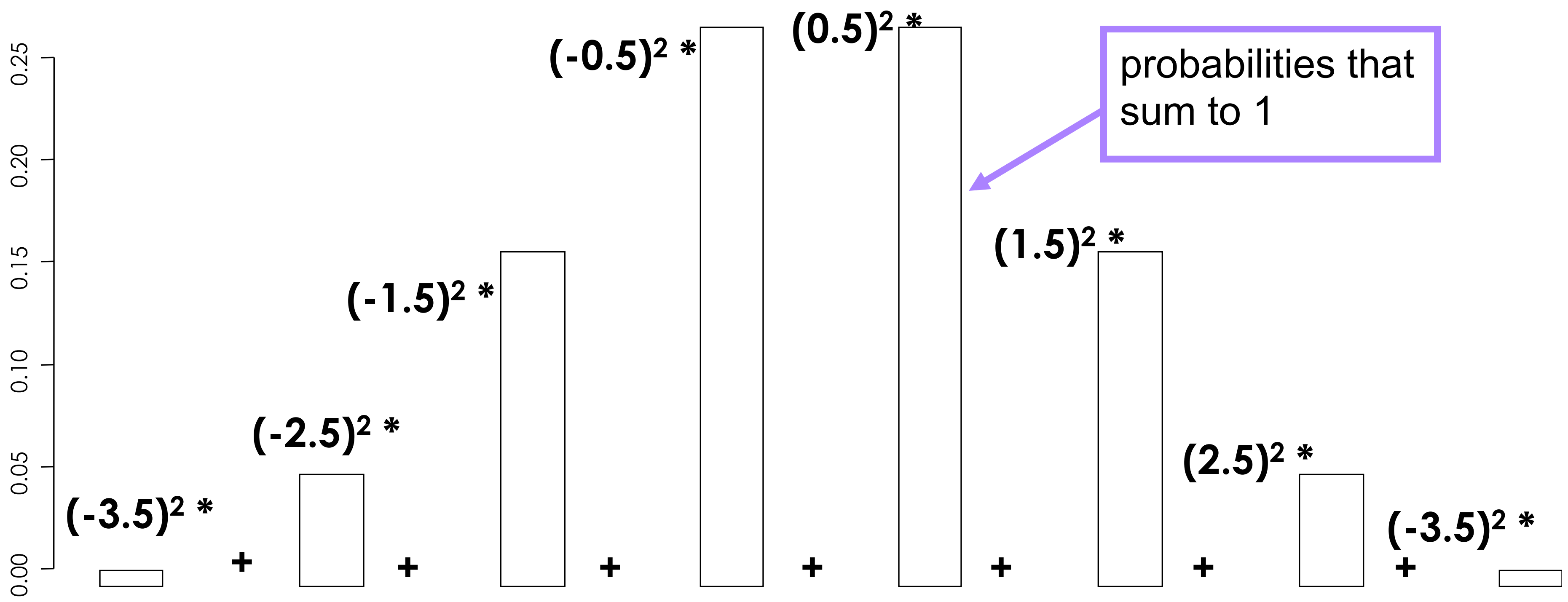
$\mu = 3.5$

Quiz Q:

- What is the average degree of a graph with 10 nodes and probability $p = 1/3$ of an edge existing between any two nodes?
  - 1
  - 2
  - 3
  - 4

What is the variance?

- variance in degree
  σ^2=**(n-1)*p*(1-p)**
- in general $\sigma^2 = E[(X-\mu)^2] = \Sigma\ (x-\mu)^2\ p(x)$



Approximations

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

Binomial

limit *p* small

$$p_k = \frac{z^k e^{-z}}{k!}$$

Poisson

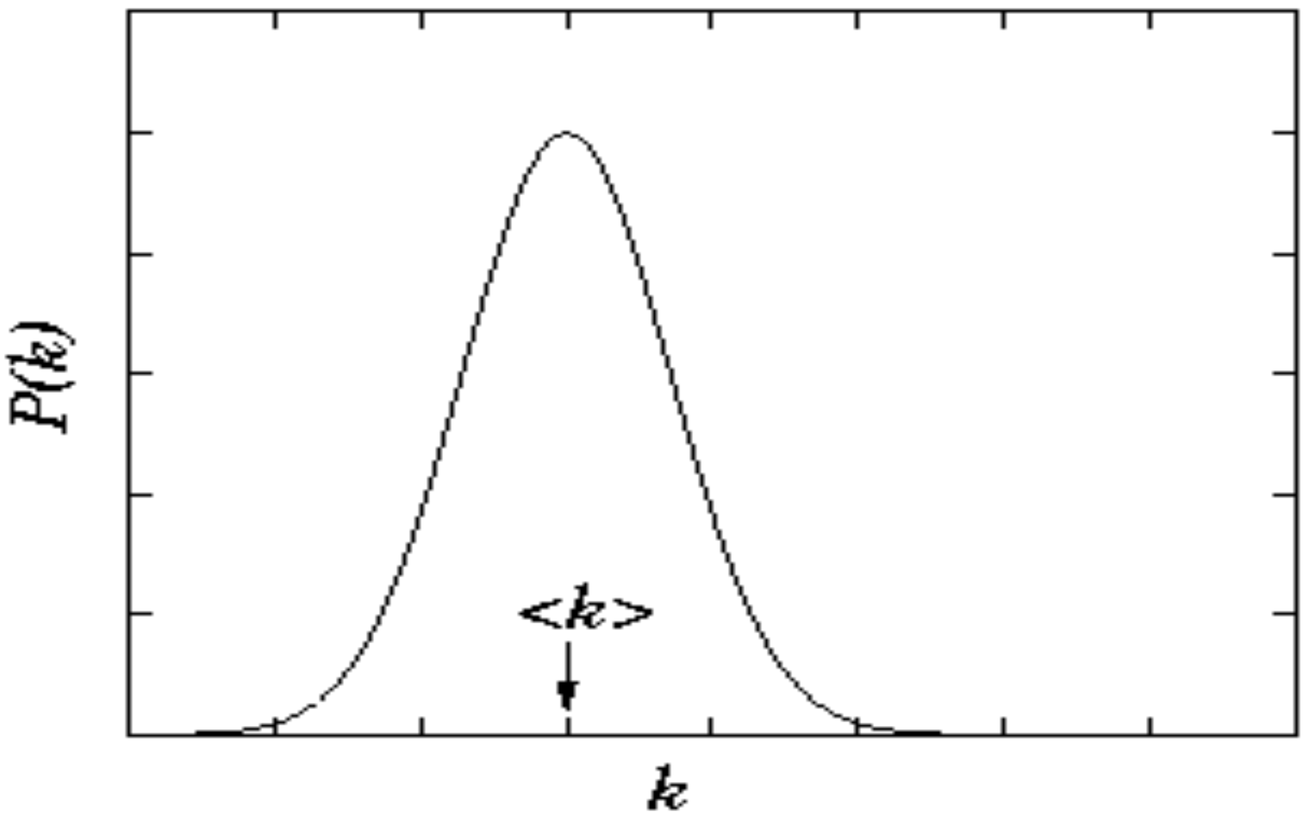
limit large *n*

$$p_k = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-z)^2}{2\sigma^2}}$$

Normal

Poisson distribution

Poisson distribution



What insights does this yield? No hubs

- You don't expect large hubs in the network