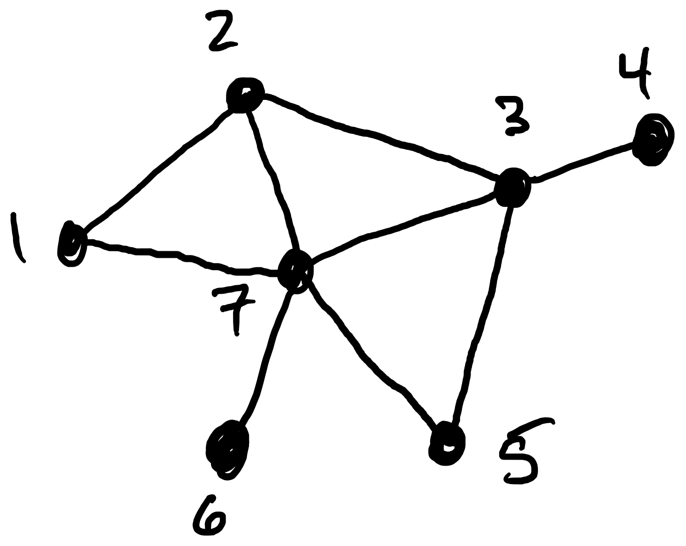
degree distribution
degree sequence
power law

SYSM 6302

CLASS 8

Degree Sequence

A list of all node degrees, repeating entries as necessary.



node	degree
1	2
2	3
3	4
4	1
5	2
6	1
7	5

Degree Sequence: 1, 1, 2, 2, 3, 4, 5

Doesn't need to be in order, but usually is

Degree Distribution

The distribution of degrees of nodes

(directed graphs have in- and out-degree distributions)

(as though degree is a random variable

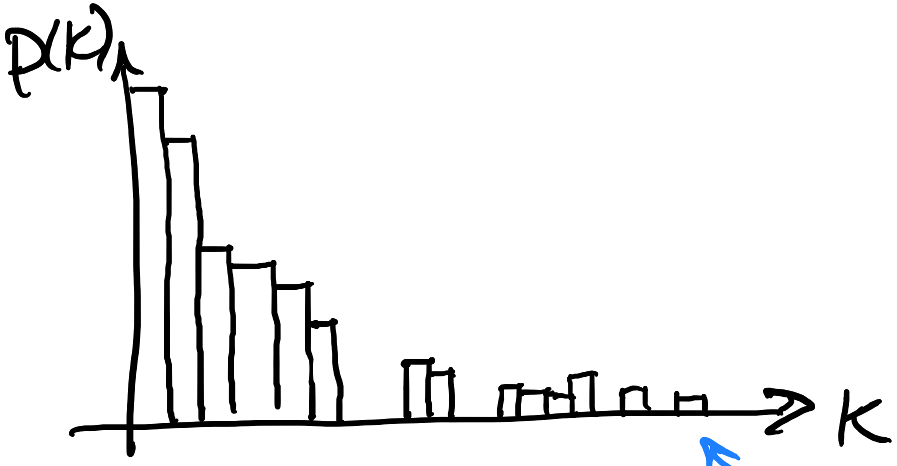
$P(k) \sim$ probability of a node having degree k)

In practice: a frequency histogram of the degree sequence.

↳ when normalized by # of nodes, acts like a probability distribution

$1,1,2,2,3,4,5 \Rightarrow P(1)=\frac{2}{7}, P(2)=\frac{2}{7}, P(3)=\frac{1}{7}, P(4)=\frac{1}{7}, P(5)=\frac{1}{7}$

Most realworld networks have right skewed degree distributions:



More low degree nodes than high degree nodes

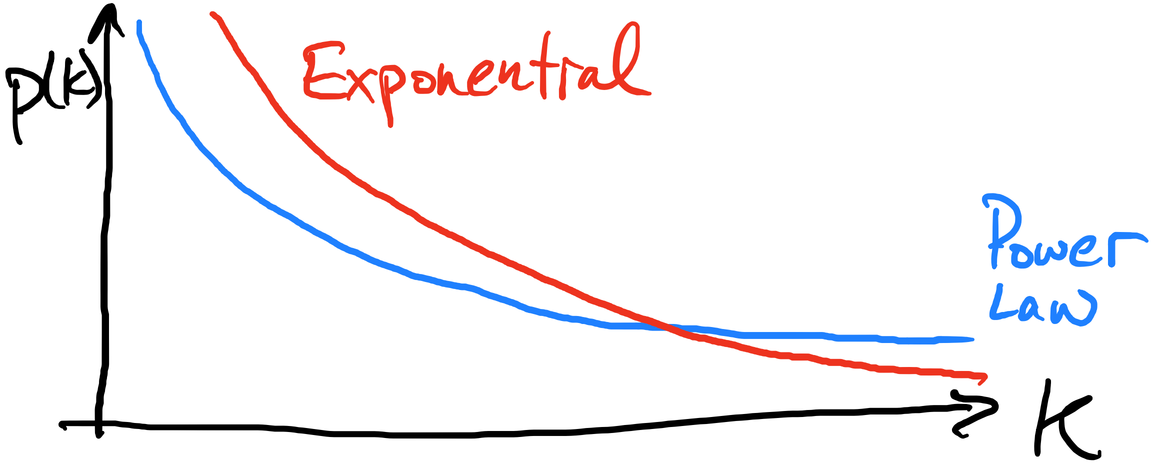
high degree nodes are often called "hubs"

Hubs account for many of the "networked" phenomena we observe, such as the small-world effect.

Power Laws

A network's degree distribution follows a power law if it is of the form: $p(k) = Ck^{-\alpha}$ $\quad 2 \leq \alpha \leq 3$

Power law distributions are an example of "heavy tailed" distributions because they are not bounded (in the tail) by the exponential distribution.



$$p(k) = \beta e^{-\beta k} < p(k) = Ck^{-\alpha}$$

as $k \to \infty$

Power laws having heavy tails means that there are still fewer high degree nodes than low degree nodes (right skewed), but there are more hubs present than in other networks.

Networks that have power law degree distributions are often called Scale-free networks.

$$P(k) = C k^{-\alpha} \implies P(ak) = C(ak)^{-\alpha} = C a^{-\alpha} k^{-\alpha}$$

Up to a constant, invariant to scale

$$\implies P(ak) \propto P(k)$$

proportional to

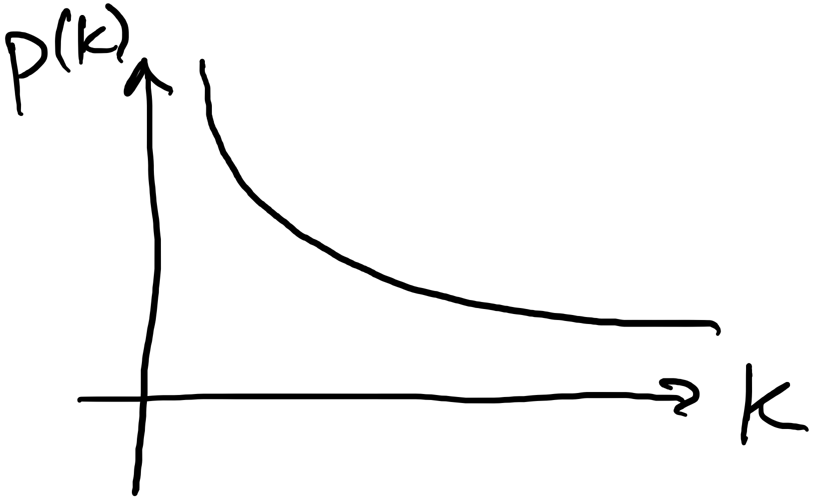
Log Magic of Power Laws

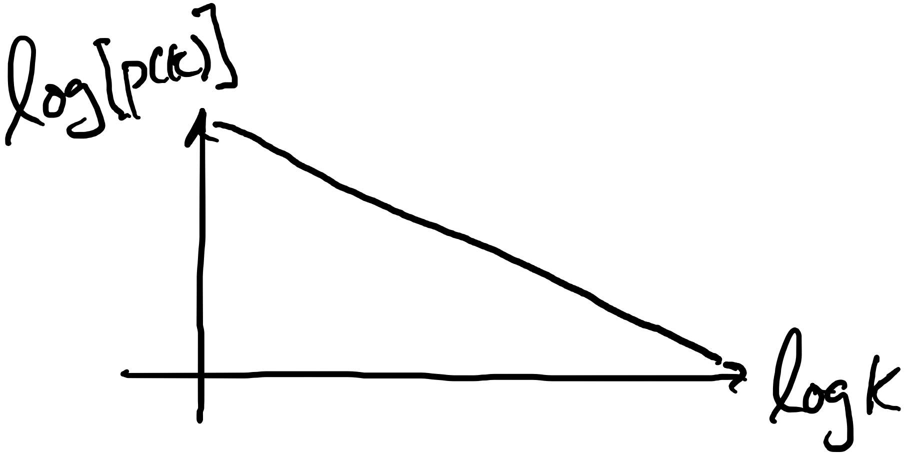
$$p(k) = ck^{-\alpha} \implies \log[p(k)] = \log[ck^{-\alpha}]$$

$$= \log c + \log k^{-\alpha}$$

$$\underbrace{\log[p(k)]}_{y} = \underbrace{\log c}_{b} \underbrace{-\alpha}_{+m} \underbrace{\log k}_{x}$$

$$y = b + m\,x$$





It's a line!

Cumulative distribution

"Fluid" approximation

$$P(k) = C \sum_{k=K}^{\infty} k^{-\alpha} \simeq \int_{K}^{\infty} k^{-\alpha}\, dk \qquad (\alpha > 1)$$

$$= \frac{C}{\alpha - 1} K^{-(\alpha-1)}$$

also a line!
with slope $\alpha - 1$

CDF nicely smooths out the variation usually present in the PDF.

Fitting Power Laws

- → Eyeball test: does the PDF/CDF look straight on a Log-Log plot?

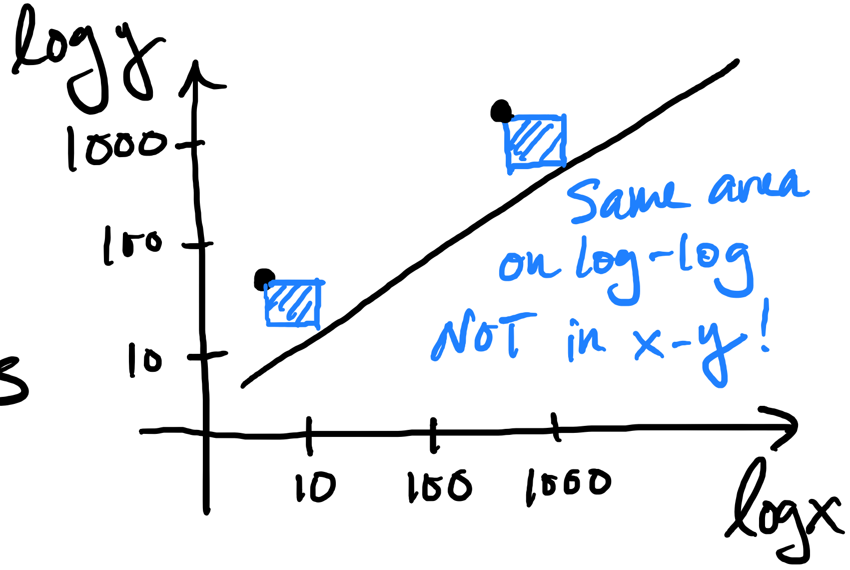
- → Most "scale-free graphs" are not a perfect match to a power law → only a portion looks straight (follows a power law)

- → Power Law is a tail phenomena, so we look for alignment to the power law distribution <u>in the tail</u>.

If it looks straight:

→ could run a linear regression

↳ but this gives slightly biased results



→ Use a maximum-likelihood result:

$$\alpha = 1 + N\left[\sum \ln\left(\frac{k_i}{k_{min} - \frac{1}{2}}\right)\right]^{-1}$$

degree sequence, so degrees can be repeated

nodes with degree $k \geq k_{min}$

minimum degree for which the power law holds