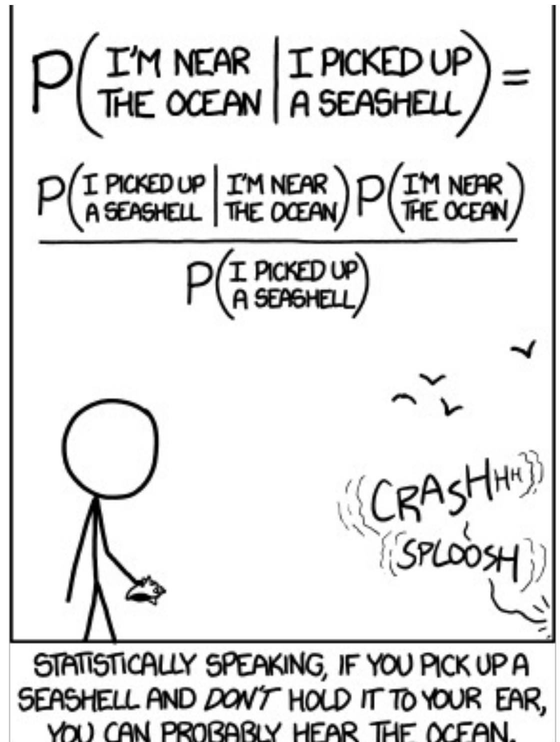
DS8104: Network Science

Class 5: Configuration Models



https://alexandergates.net https://uvads8104.github.io/

MCMC

Markov chain monte carlo (MCMC)

Goal: select a statistical sample to approximate a hard combinatorial problem

Applications: solve integration and optimization problems

- calculate expectations
- marginalization
- minimization
- model selection

Metropolis & Ulam, 1949

MCMC

Goal (rephrased): you want to explore some massive state space in a principled way

Start from a random point in that space

Rule for taking a step away from your current point

(sometimes) rejection rule, to know when your step is outside of the acceptable space

Difficulty: convergence? How many steps are necessary to converge to the equilibrium distribution?

MCMC in networks

Define a space of network-types with some property
(i.e. having a specific degree distribution, connected, etc.)

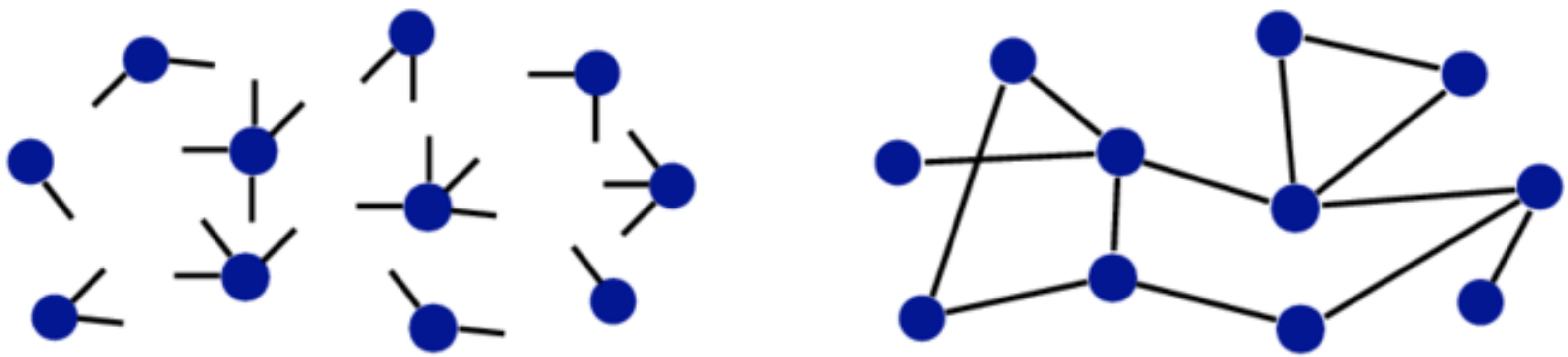
We start with some network we know is in the space
(often the empirical observed network)

Rule for modifying the network that produces a new network with the same properties

Produces a sample of networks satisfying one property, which we can then use to test other properties

Configuration model

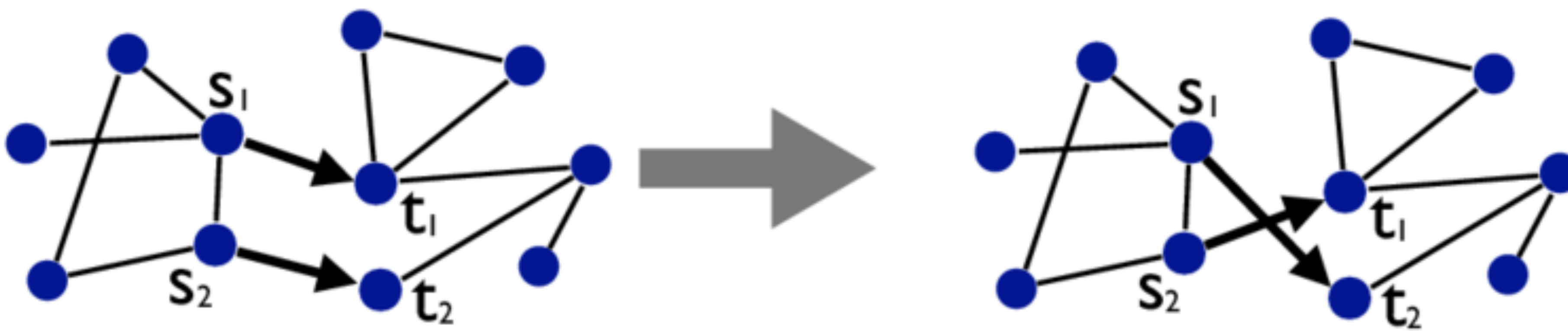
Motivation: The network density and degree sequence are *super, super, super, super* important for most other network properties!



Fix the degree sequence:
4,4,3,3,2,2,2,2,2,1,1

Configuration model

Edge swap rule: sample from the space of networks with the same degree sequence



Configuration model

Analytically tractable for most properties!!

The probability that any two nodes i and j are connected is the product of their degrees:

$$p_{ij} = \frac{k_i k_j}{2m - 1}$$

Guaranteed giant component:

$$\langle k^2 \rangle - 2\langle k \rangle > 0$$

Clustering coefficient:

$$C = \frac{1}{n} \frac{[\langle k^2 \rangle - \langle k \rangle]^2}{\langle k \rangle^3}$$