EXPONENTIAL RANDOM GRAPH MODELS:
BASIC CONCEPTS

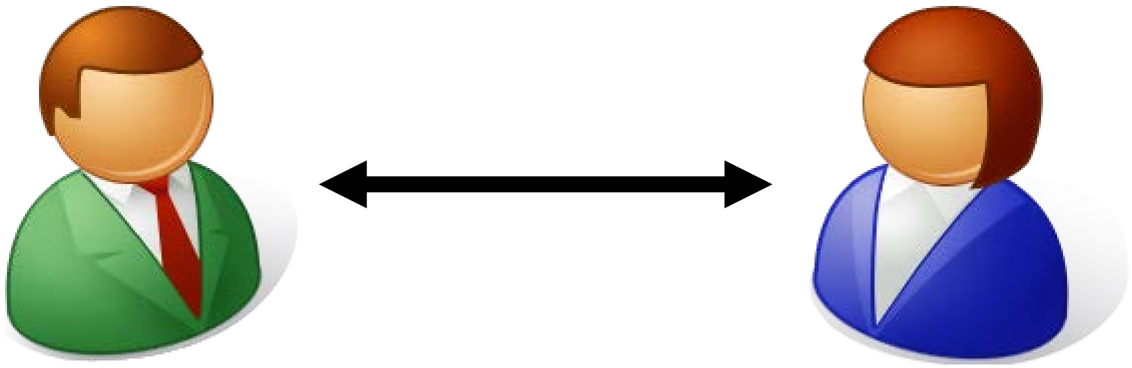
- ❑ Formation of social network structure
- ❑ Working with graph distributions

I eat (predominantly) vegetarian food

Individual Approach

- Ethical
- Economics
- Health
- Taste

Network Approach



<u>Individualizing</u> social structure is problematic
Emirbayer's (1997) *"Manifesto for a Relational Sociology"*

The social is by definition relational

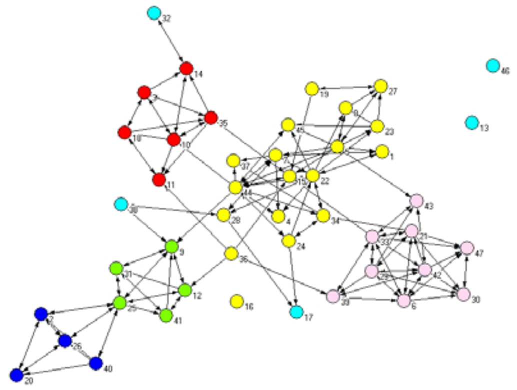
If we want to understand the social we need to understand social relations

We are “actors in social relations”
(Abbott, 1997: 1152)

To understand social relationships, we need a relational methodology
(not a methodology that assumes every individual is independent)

Elements for social network theory

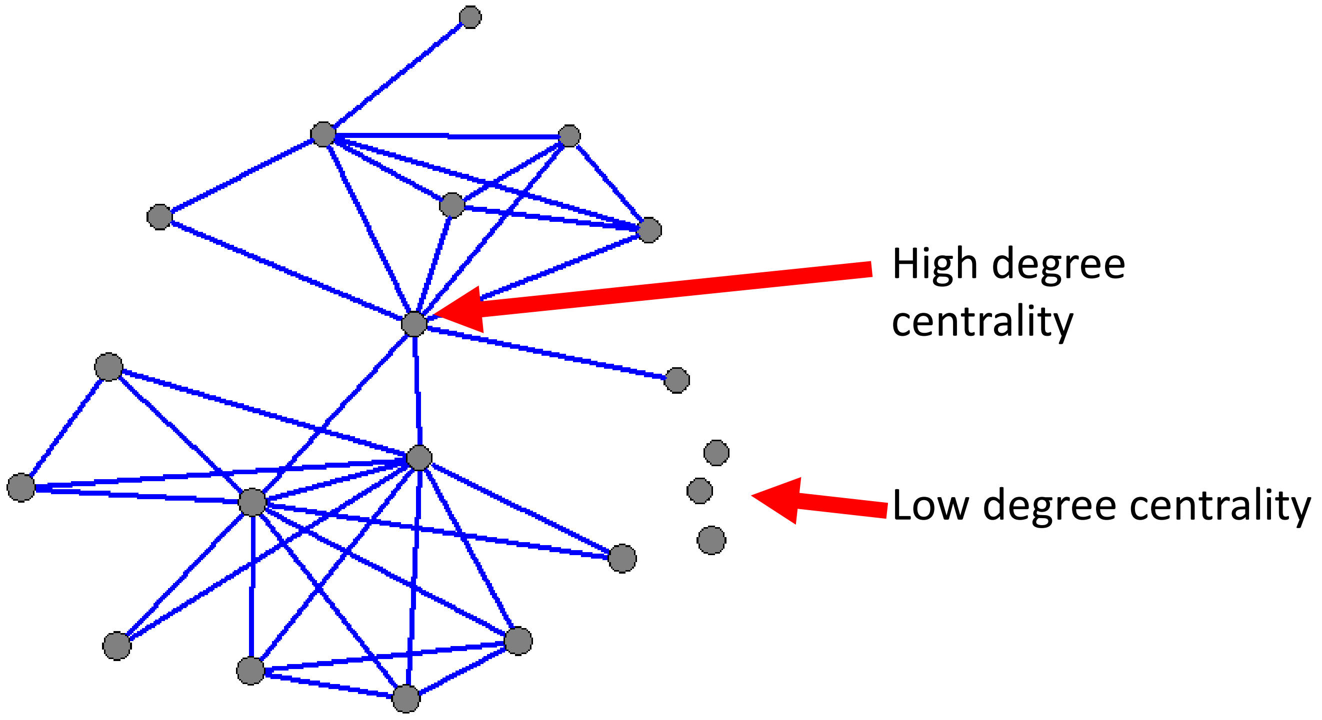
- Locally emergent
  - Local patterns form global structure
- Network ties self-organize
  - Through dependency among ties:
  - The presence of one tie may lead to another



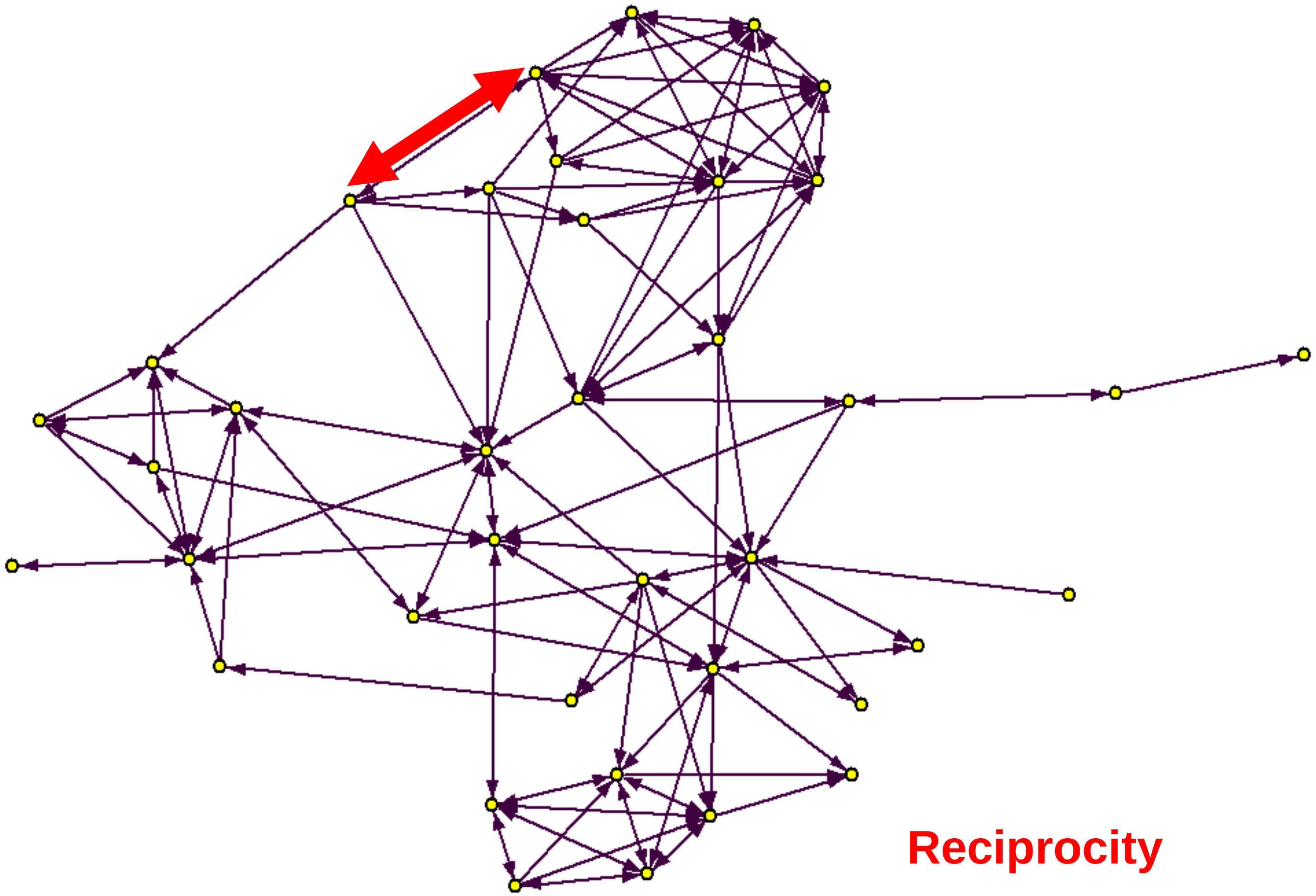
- Network patterns are evidence of ongoing structural processes
  - Static trace of dynamic social processes
- Multiple processes can operate simultaneously
- Social networks are structured, yet stochastic
  - Structure and randomness

Centrality: How important a node is

Degree centrality: the degree of each node

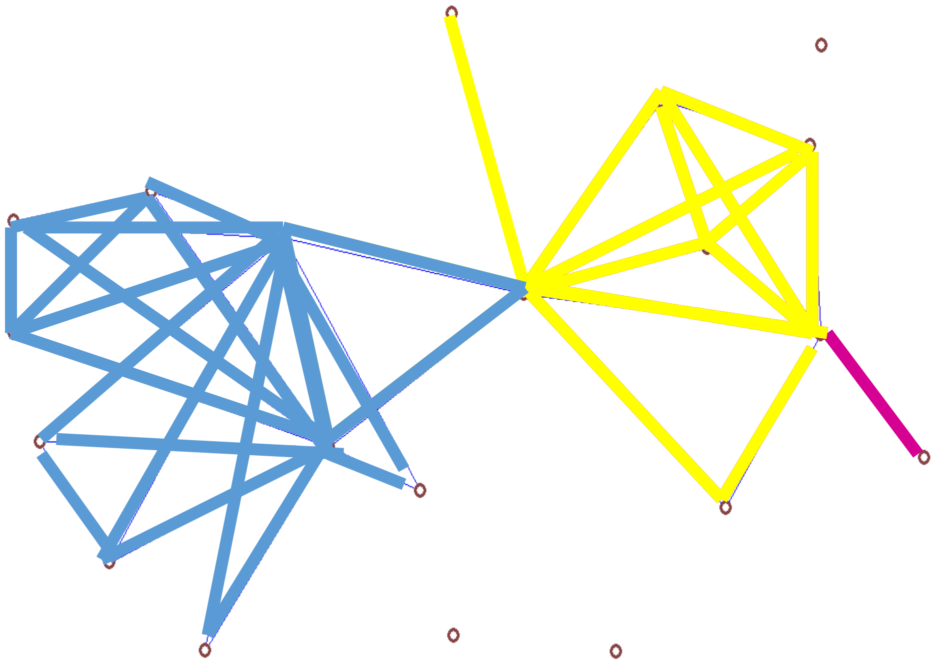


After hours socialising network



Reciprocity

Clustering in networks

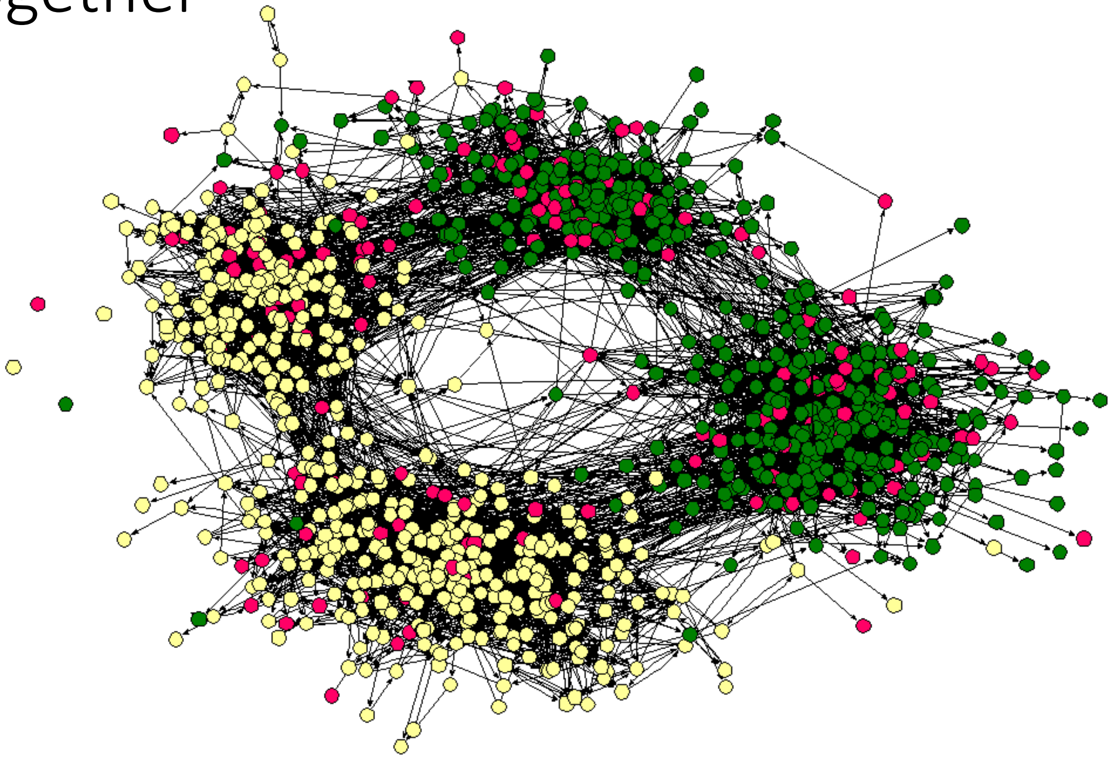


Friendship network

(Robins, 2002)

Homophily

- Birds of a feather flock together

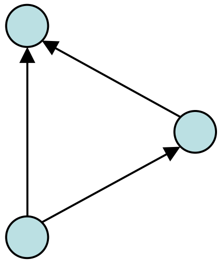


High School friendship
Moody, 2001 – colours indicate white/black/other

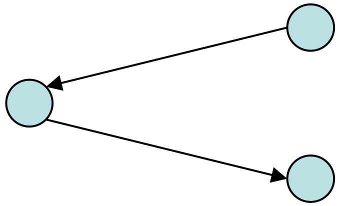
Local, everyday social "rules"

- You scratch my back, I'll scratch yours

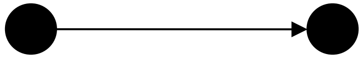
- A friend of a friend is a friend



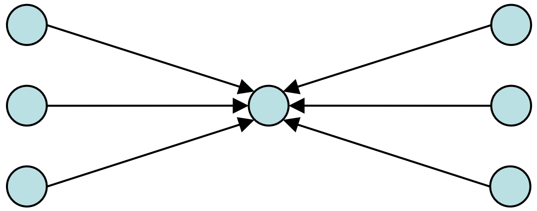
- Brokerage



- Birds of a feather flock together



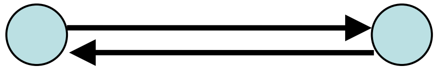
- Follow the crowd



No single rule explains why ALL network ties occur

Dependency

- You scratch my back, I'll scratch yours



  - One tie may follow the other in time
  - The models may express the outcomes of an implicit longitudinal process

Multiple social processes

Tie

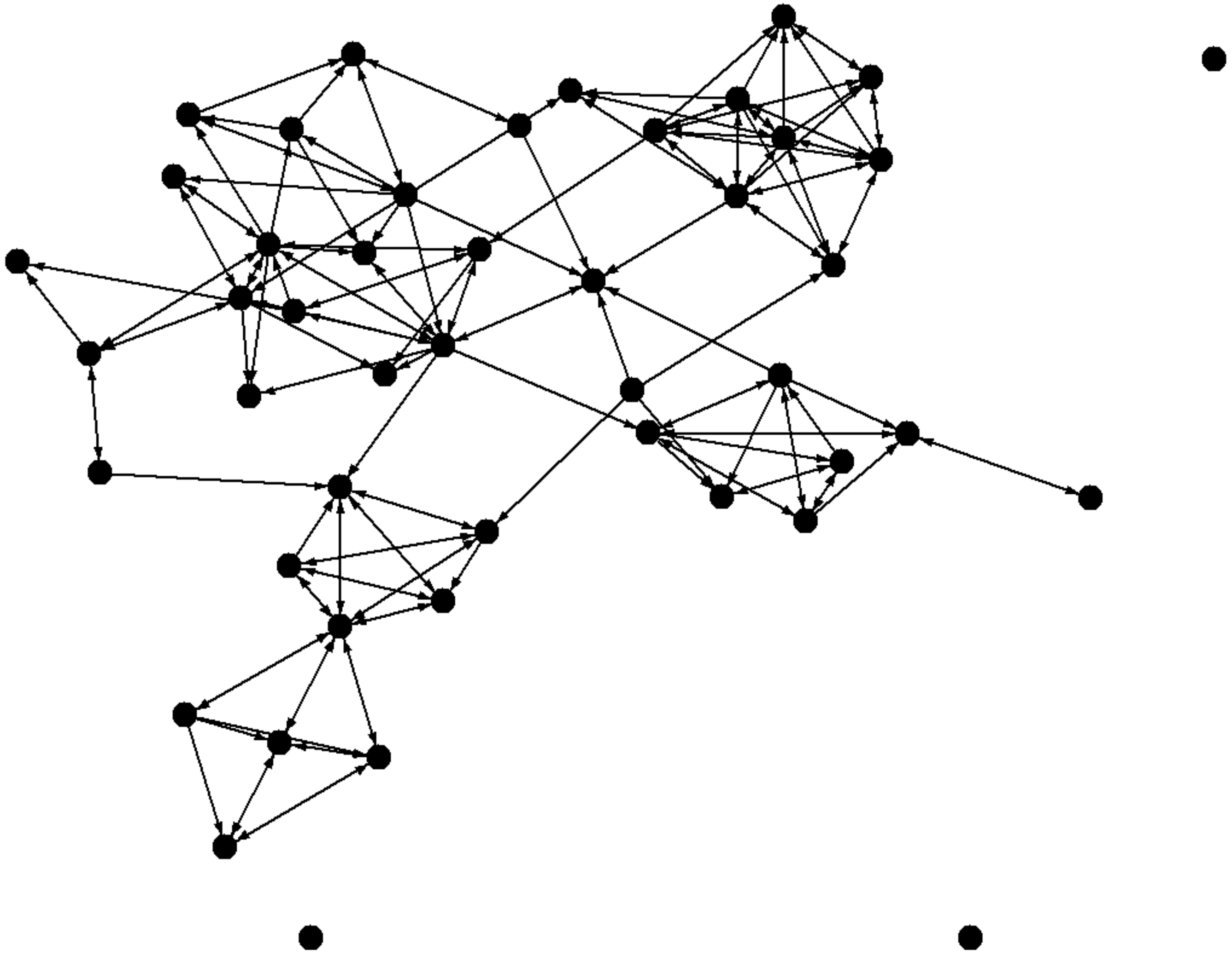
Reciprocity

Activity

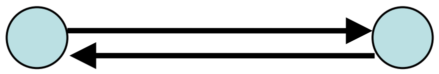
Popularity

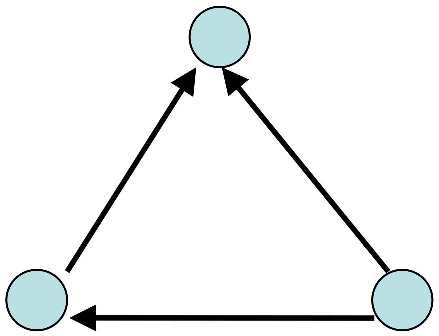
Triads

Brokerage



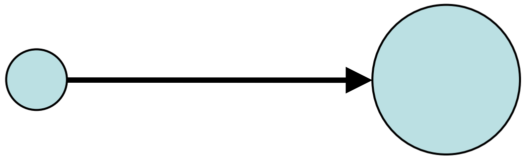
Network self-organisation





Ties occur due to the presence (or absence) of other ties

Actor-relation effects



"I trust my colleague
who has lots of experience"

Ties occur due to the presence of actor attributes

Exogenous effects

Formal reporting lines affect informal social ties

"I trust my boss"

Ties occur due to the presence of other "fixed" or "formal" ties

Which social process?

- Would anyone suggest that a network is explained ONLY by:
  - Reciprocity?
  - Transitive closure?
  - Preferential attachment?
  - Brokerage?
  - Homophily?

- If not, then we need a model that can examine a network for ALL of these processes at the same time

Exponential random graph models (ERGM)

(Frank & Strauss, 1986; Wasserman & Pattison, 1996; Robins et al, 2009; Snijders et al, 2006

Parameter for Q
(how important is Q in the model)

Why do social network ties occurs?

(i.e. ERGM is a tie-based model)

A probability distribution of graphs

Sum over all configurations Q
(dependence assumption determines which configurations)

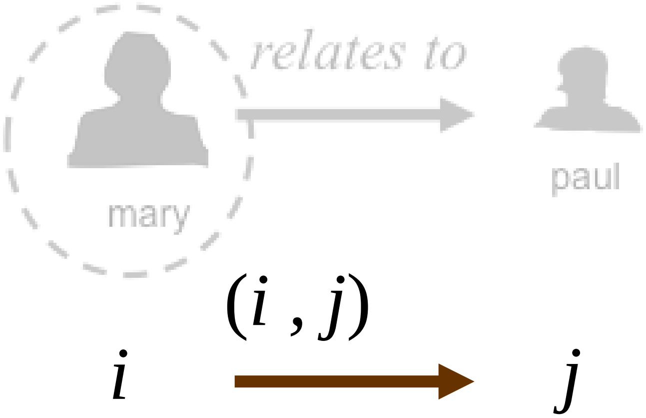
A normalising constant
(probability of all graphs must add to 1)

DEPENDENCE AND SIMPLE ERGMs

- ❑ The general form of an ERGM
- ❑ Bernoulli models
- ❑ Markov models

Some notation

We conceive of the Graph as a collection of

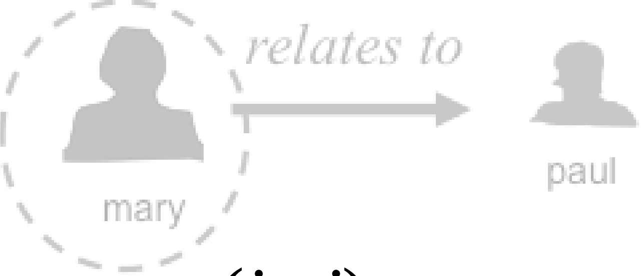


Tie variables: $\{X_{ij}: i,j \in V\}$

$$X_{ij} = \begin{cases} 1 & \text{if } i \rightarrow j \\ 0 & \text{otherwise} \end{cases}$$

Some notation

We conceive of the Graph as a collection of



$(i\ ,\ j)$

$i \longrightarrow j$

Tie variables: $\{X_{ij}: i,j \in V\}$

$$x_{ij} = \begin{cases} 1 & \text{if } i \to j \\ 0 & \text{otherwise} \end{cases}$$

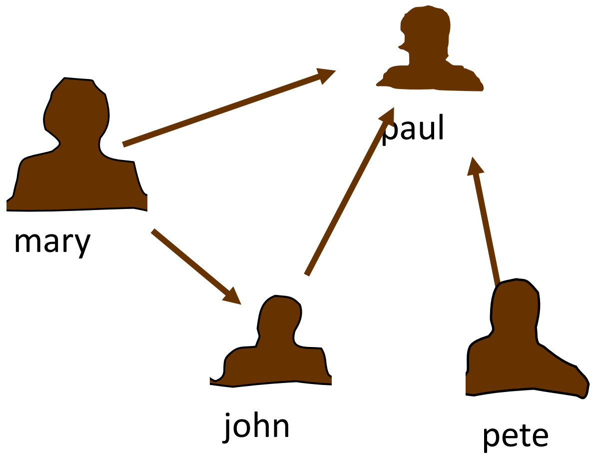
Generally binary

on $\longrightarrow$ $x_{ij} = 1$

off $\not\longrightarrow$ $x_{ij} = 0$

Some notation

We conceive of the Graph as a collection of



Tie variables: $\{X_{ij}: i,j \in V\}$

$$x_{ij} = \begin{cases} 1 & \text{if } i \rightarrow j \\ 0 & \text{otherwise} \end{cases}$$

$x =$

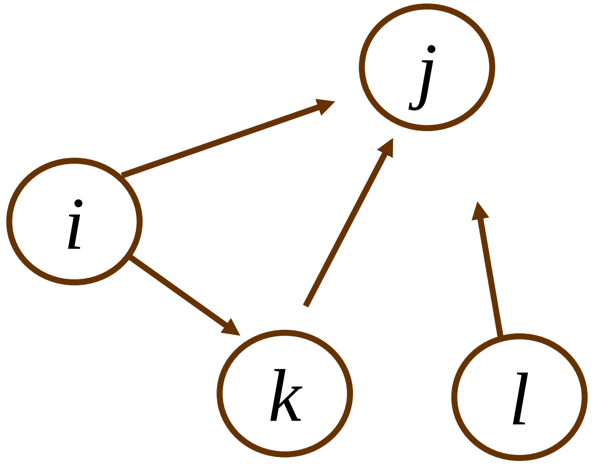
i	-	x_{ij}	x_{ik}	x_{il}
j	x_{ji}	-	x_{jk}	x_{jl}
k	x_{ki}	x_{kj}	-	x_{kl}
l	x_{li}	x_{lj}	x_{lk}	-

$=$

i	-	1	1	0
j	0	-	0	0
k	0	1	-	0
l	0	1	0	-

Some notation

We conceive of the Graph as a collection of



Tie variables: $\{X_{ij} : i,j \in V\}$

$$x_{ij} = \begin{cases} 1 & \text{if } i \rightarrow j \\ 0 & \text{otherwise} \end{cases}$$

$x =$

i	-	x_{ij}	x_{ik}	x_{il}
j	x_{ji}	-	x_{jk}	x_{jl}
k	x_{ki}	x_{kj}	-	x_{kl}
l	x_{li}	x_{lj}	x_{lk}	-

$=$

i	-	1	1	0
j	0	-	0	0
k	0	1	-	0
l	0	1	0	-

Some notation

Regard each network tie as a *random variable* (often binary)

Notation

X_{ij} = 1 if there is a network tie from person *i* to person *j*
= 0 if there is no tie
for *i*, *j* members of some set of *actors N*.

A *directed network:* X_{ij} and X_{ji} are distinct.
A *non-directed network:* $X_{ij} = X_{ji}$

X … matrix of all variables
x … matrix of observed ties (the network)

(For nodal attributes, we use Y_i as the variable to indicate the measure of the attribute for node *i*.)

Exponential random graph models

(Frank & Strauss, 1986; Wasserman & Pattison, 1996)

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left\{\sum_{Q}\lambda_Q z_Q(\mathbf{x})\right\}$$

The summation is over all “configurations” *Q*

Local subgraphs that are hypothesized as the ‘building blocks’ of the network

$$z_Q(\mathbf{x}) = \prod_{X_{ij}\in Q} x_{ij} = 1 \text{ if } Q \text{ is observed in graph}$$

λ_Q parameter for the presence of Q

κ is a normalizing quantity.

Exponential random graph models

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left\{\sum_{Q}\lambda_Q z_Q(\mathbf{x})\right\}$$

Parameter for Q
(how important is Q in the model)

Statistic for Q
(Is Q in the observed graph **x**?)

A probability distribution of graphs

Sum over all configurations Q
(dependence assumption determines which configurations)

A normalising constant
(probability of all graphs must add to 1)

What are we trying to do?

Estimate model parameters

- ❑ Positive parameter estimates indicate more configurations observed in the network than expected by chance.
- ❑ Negative parameter estimates indicate fewer configurations than expected by chance.

We want to know how the global network structure might have been built up out of small local substructures.

The parameter estimates permit us to make inferences about this.

Dependences

Once we move beyond simple random graph models, we introduce dependences among network tie variables

These express various types of network self organization.

A dependence assumption picks out certain types of network patterns – network configurations – that are possible in the model.

In other words, we assume that the network is built up of these configurations.

Markov random dependences (and graphs)

(Frank & Strauss, 1986, *JASA*)

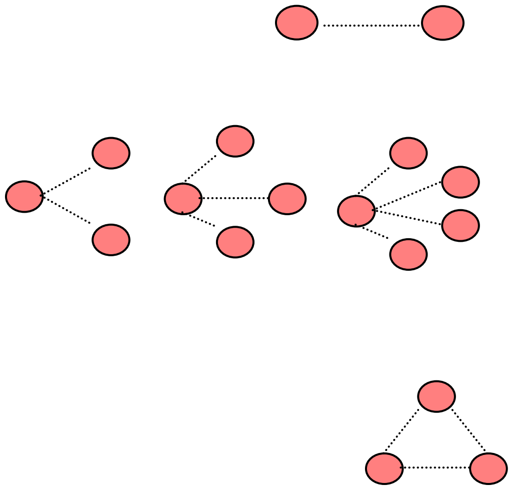
- Frank and Strauss drew on the work of Besag (1974) in spatial statistics
  - In particular, the *Hammersley-Clifford theorem* that sets out constraints on model form implied by dependence assumptions.
  - Dependence graph

- They proposed a network dependence assumption (Markov dependence):
  - Two tie variables are conditionally independent unless they share a node.

Markov random graphs

- Suppose that edges are conditionally dependent if and only if they share a node. (Frank & Strauss, 1986)

- Frank and Strauss showed that configurations in this model comprised edges, stars and triangles.



edge

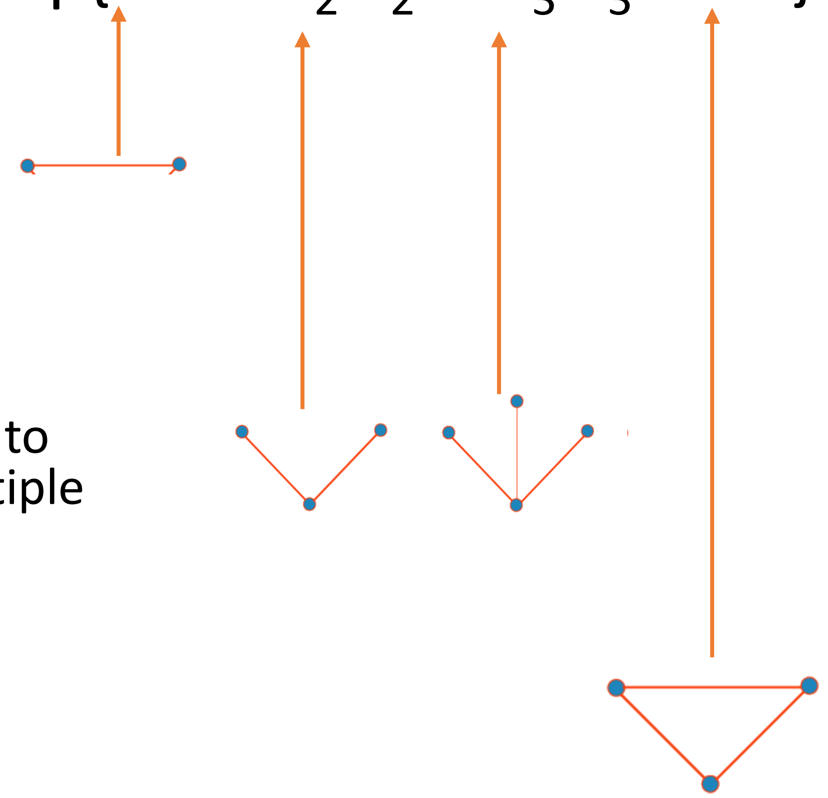
2-star *3-star* *4-star* …

triangle

A Markov random graph model:

Undirected networks

$$\Pr(\boldsymbol{X} = \boldsymbol{x}) = (1/\kappa) \exp\{\theta L + \sigma_2 S_2 + \sigma_3 S_3 + \tau T\}$$



- *Edge parameter (θ)*

  L … number of edges

- *Star parameters (σ_k)*

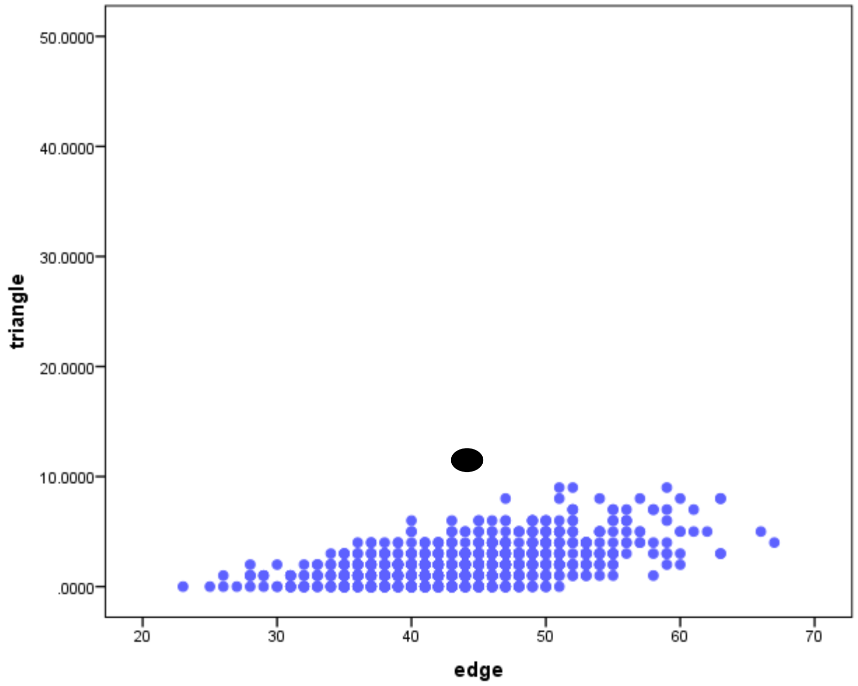
  Propensities for individuals to have connections with multiple network partners

- *Triangle parameter (τ)*

  represents network closure

If θ is the only nonzero parameter, this is a Bernoulli random graph model.

Simulated results from Bernoulli graph model

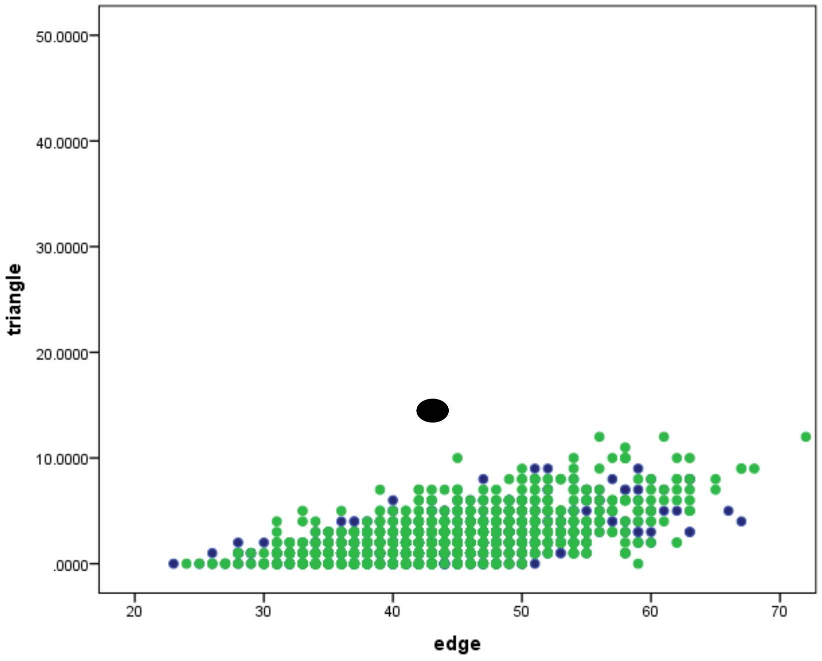


<u>Statistics from simulated samples</u>

Blue = Bernoulli

● Observed statistics

Simulated results from Markov graph model Edge, 2star parameters



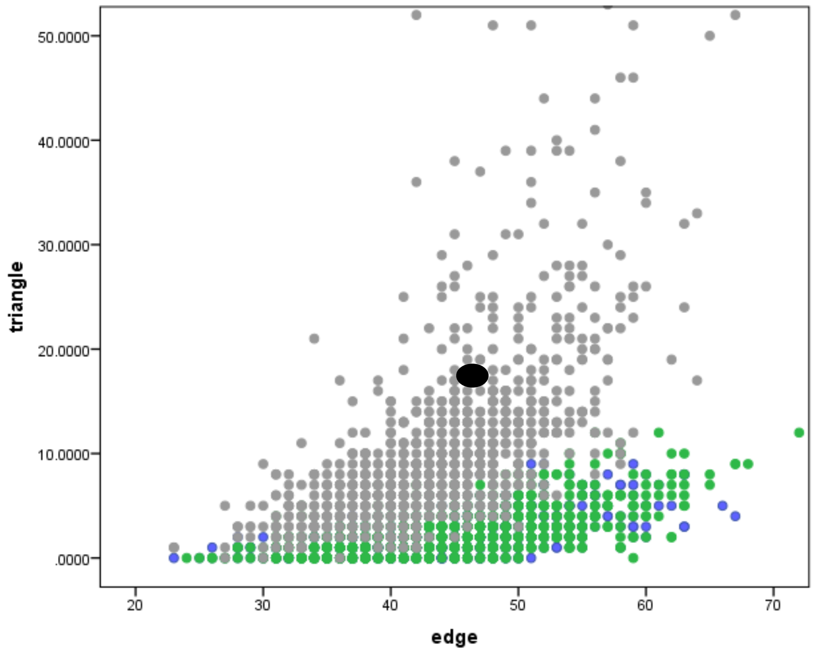
Statistics from simulated samples

Blue = Bernoulli

Green = L,2star

● Observed statistics

Simulated results from Markov graph model Edge, 2star, 3star, triangle parameters



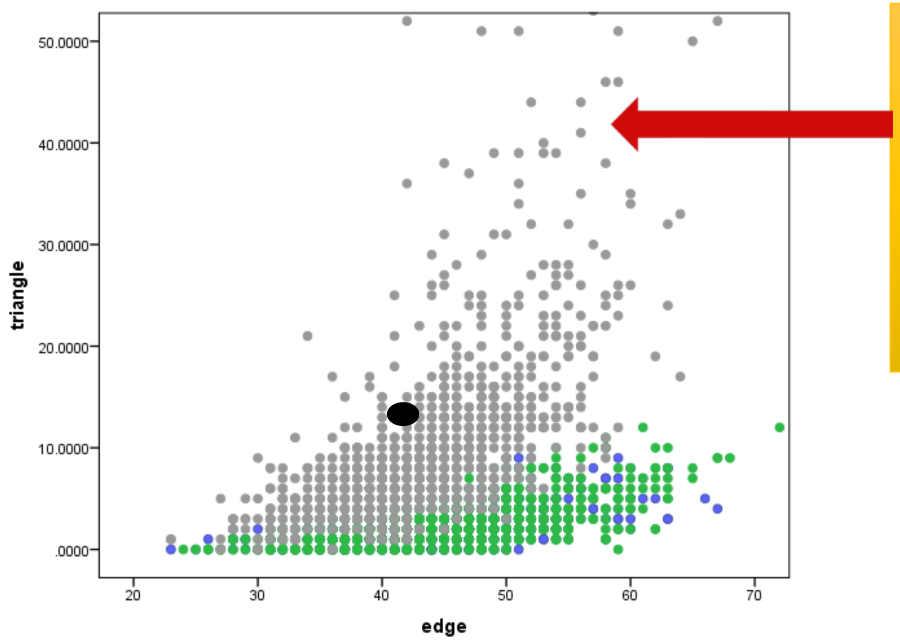
<u>Statistics from simulated samples</u>

Blue = Bernoulli

Green = L,2star

Observed statistics

Simulated results from Markov graph model
Edge, 2star, 3star, triangle parameters



This 'leakage' shows a common problem with Markov models – they are not always stable; and may be *degenerate.*

● Observed statistics

But, be careful!

- Markov random graph distributions provide statistical models for social networks based on plausible assumptions and importantly can represent clustering through the triangle parameter!

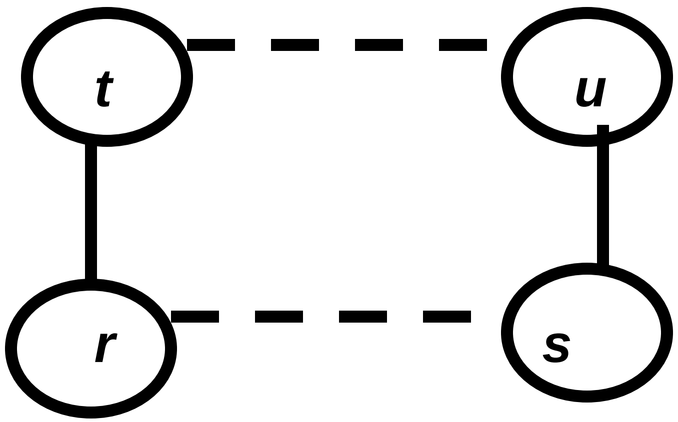
- THEY DON'T ALWAYS WORK!

Social circuit dependences

Network ties self-organize within 4-cycles.

(Pattison & Robins, 2002; Snijders et al, 2006)

- Two possible network ties are conditionally dependent if they would form a 4-cycle.
- Social circuit dependence



Configurations/model parameters for social circuit models

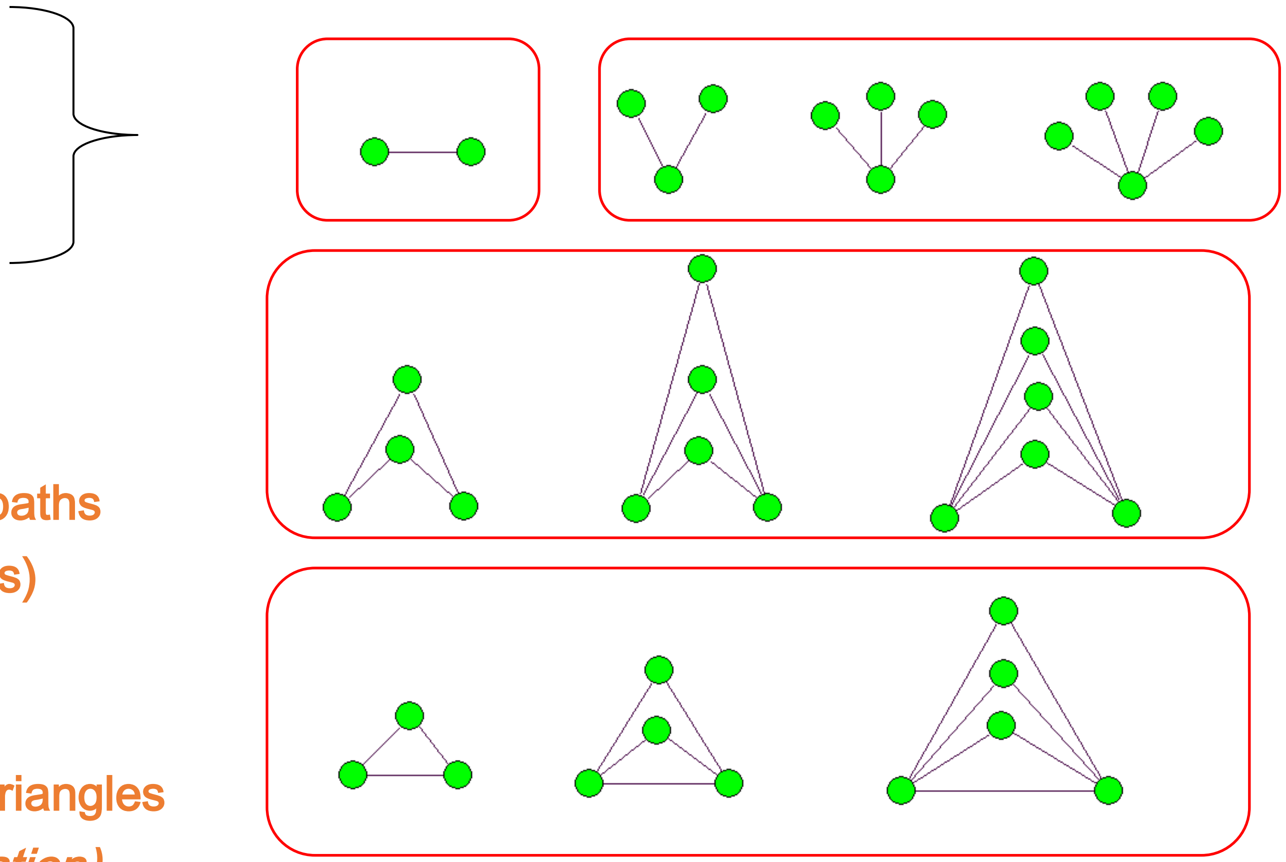
Parameters correspond to configurations of the following types:

Edges

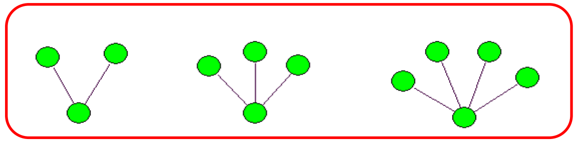
Stars
(*k*-stars)

Multiple paths
(*k*-2-paths)

Multiple triangles
(triangulation)



Centralization parameter: Alternating k-stars



$$z(\mathbf{x}) = S_2 - \frac{S_3}{\lambda} + \frac{S_4}{\lambda^2} - \ldots + (-1)^{n-2}\frac{S_{n-1}}{\lambda^{n-3}}$$

Usually we set λ = 2. Hunter & Handcock (2006) show how to estimate lambda

<u>Interpretation</u>:

Positive parameter indicates centralization through a small number of high degree nodes

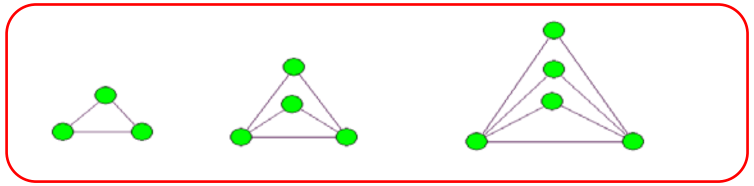
core-periphery based on popularity

More dispersed degree distribution

Negative parameter: “truncated” (less dispersed) degree distribution; nodes tend not to have particularly high degrees.

Equivalent to geometrically weighted degree distribution parameter (Hunter, 2007)

Network Closure parameter: Alternating k-triangles



$$u(\mathbf{y}) = T_1 - \frac{T_2}{\lambda} + \frac{T_3}{\lambda^2} - \ldots + (-1)^{n-2} \frac{T_{n-2}}{\lambda^{n-3}}$$

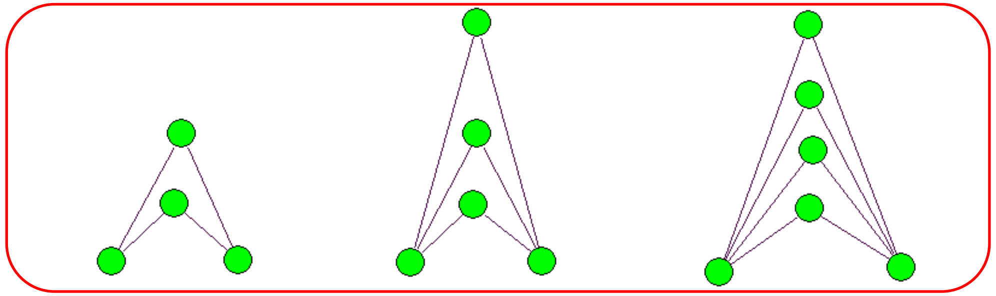
Interpretation:

a. Positive parameter suggests triangles tend to “clump” together in denser regions of the network.

b. Models the *edgewise shared partner* distribution:

For each pair of tied nodes, how many partners do they share? (Hunter, 2007)

Connectivity Parameter: Alternating k-2paths



Interpretation:

a. Localized structural equivalence

b. With the AKT parameter in the model, this indicates presence of structural holes.

c. Models the *pairwise shared partner* distribution:

For each pair of nodes (tied or not), how many partners do they share? (Hunter, 2007)

Social circuit dependences

Larger network configurations emerge:

Parameters for **degree sequences**, **denser regions of triangulation**, **multiple connectivity**.

These models fit much better.

A dependence assumption that captures emergence may be necessary to model real social networks.

(Robins, Snijders, Wang, Handcock & Pattison, 2007)

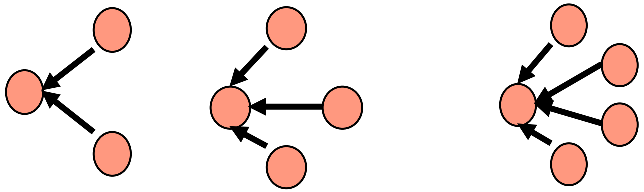
ERGMs for directed networks

Dyadic parameters

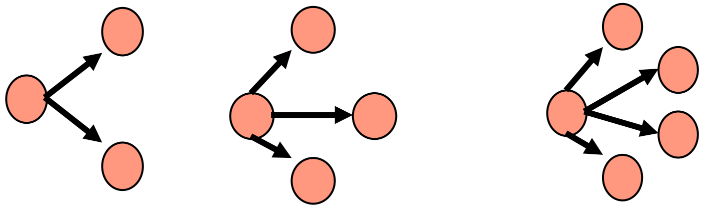
Arc

Reciprocity

Degree-based parameters



Popularity spread (centralization)
Alternating in-stars



Activity spread (centralization)
Alternating out-stars

Also:

Sources: nodes with 0 in-degree

Sinks: nodes with 0 out-degree

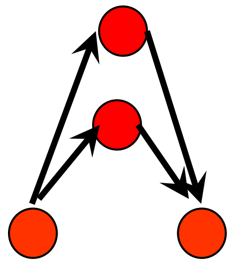
Isolates: nodes with 0 in- and 0 out-degree

Connectivity parameters

(alternating 2-paths)

Multiple 2paths (A2P-T)

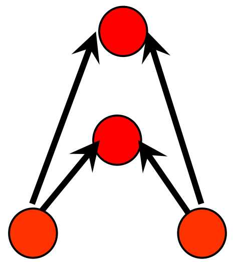
(T = "transitive")



Interpretation: Localized connectivity

Shared activity (A2P-U)

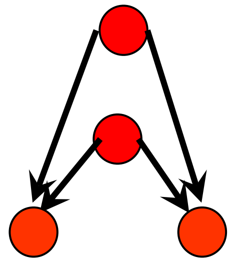
(U = "up")



Interpretation: Activity-based structural equivalence

Shared popularity (A2P-D)

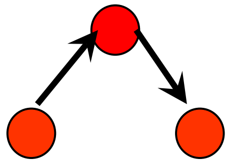
(D= "down")



Interpretation: Popularity-based structural equivalence

Also, ***Simple connectivity***

(Markov 2path parameter)

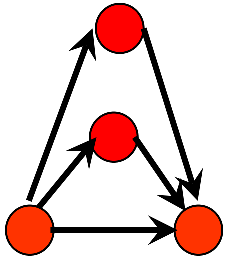


Interpretation: Correlation between in- and out-degrees

Closure parameters

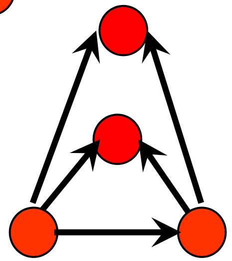
(alternating triads)

Path closure (AT-T)



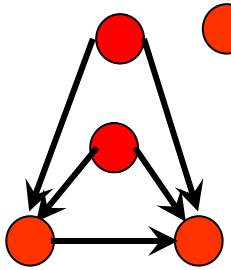
Interpretation: Closure of 2paths

Activity closure (AT-U)



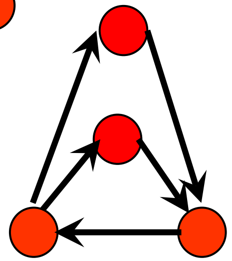
Interpretation: Activity-based structural homophily

Popularity closure (AT-D)



Interpretation: Popularity-based structural homophily

Cyclic closure(AT-C)



Interpretation: Generalized exchange

Generalized transitivity (AT-TDU) includes all of first three

Actor attributes and dyadic covariates

Social selection

- Actors select network partners based on actor attributes.
  - a process of tie formation

- Possible mechanisms
  - Homophily: actors of similar attributes tend to form ties (McPherson et al, 2001).
  - homophily in itself cannot explain the emergence of hierarchy in relations (so difference may also be important)
  - more generalized selection: individuals select social positions for themselves.

Social selection

- Also actor *main effects*
  - Nondirected - activity: Actors with certain attributes might be more active (involved in more ties)
  - Directed
    - Sender effects: Actors with certain attributes may send out more ties (more active or expansive)
    - Receiver effects: Actors with certain attributes may received more ties (more popular)

Terminology and notation: Network variables

For node set N

Let $X_{ij} = 1$ if there is a tie from node i to node j

$= 0$ if there is no tie from i to j.

A binary network

Let $X_{ii} = 0$ for all nodes i.

Define $\mathbf{X}$ as the matrix of variables $[X_{ij}]$

Define $\mathbf{x}$ as the *adjacency matrix*, the matrix of observed network ties

Terminology and notation: Attribute variables

For node set N

Let $Y_i = 1$ if node i has attribute Y

$= 0$ otherwise.

A binary attribute (e.g. gender)

Alternatively Y_i can represent categories (e.g. political party)

Or can be continuous (e.g. age)

Define **Y** as the vector of variables $[Y_i]$

Define **y** as the *attribute vector*, the vector of observed attributes.

Three types of attribute variables

1. Binary – eg male/female

2. Categorical – eg Workteams within a company

3. Continuous – eg Age, attitudes

ERGM Social selection models

Probability of observing graph **x** GIVEN observed attribute vector **y**

$$\Pr(\mathbf{X}=\mathbf{x} \mid \mathbf{Y}=\mathbf{y}) = \frac{1}{\kappa}\exp\left\{\sum_{Q}\lambda_Q z_Q(\mathbf{x}) + \sum_{R}\lambda_R z_R(\mathbf{x},\mathbf{y})\right\}$$

structural part – just as before

selection part – interaction of ties and attributes

Second summation is over all selection configurations *R*

Possible binary attribute configurations (non-directed graphs)

Activity

Positive parameter indicates node with attribute has many ties

<u>Statistic</u>: For each tie, count the number of attributed nodes

$$X_{ij}(Y_i + Y_j)$$

Then sum across all ties:

Interaction

Positive parameter indicates nodes with attribute tend to share ties

<u>Statistic</u>: For each tie, count those where both nodes are attributed

$$X_{ij}Y_iY_j$$

Then sum across all ties.

Possible binary attribute configurations (directed graphs)

Sender

Positive parameter indicates nodes with attribute tend to be more expansive (active)

Receiver

Positive parameter indicates nodes with attribute tend to be more popular

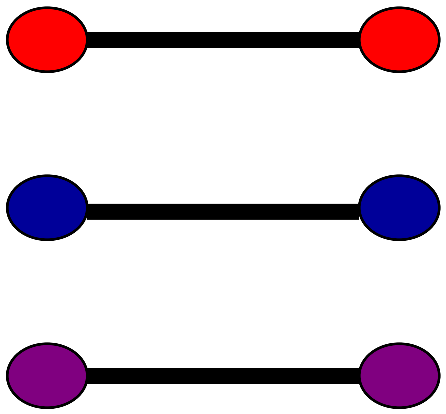
Interaction

Homophily: Positive parameter indicates nodes with attribute tend to have ties with each other (over and above sender and receiver effects)

Possible categorical attribute configurations (non-directed graphs)

Match

category



Positive parameter indicates ties within categories are more likely

Possible continuous attribute configurations (non-directed graphs)

Difference

For each pair of tied nodes the statistic is the absolute difference between the attribute values – (then summed over all pairs)

$$X_{ij} | Y_i - Y_j |$$

Negative parameter indicates that a smaller absolute difference is associated with the presence of a tie:

Ties are more likely when nodes have similar attribute values - HOMOPHILY

Possible continuous attribute configurations (non-directed graphs)

Activity (sum)

For each pair of tied nodes the statistic is the sum of the attribute values – (then summed over all pairs)

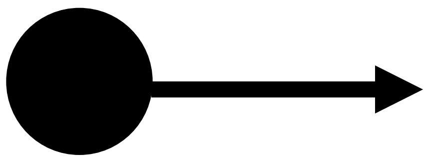
$$X_{ij}(Y_i + Y_j)$$

Positive parameter indicates that pairs of nodes with large (average) attribute values tend to be tied.

Possible continuous attribute configurations (directed graphs)

Sender

(Activity)

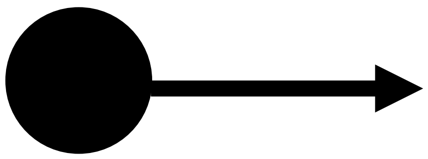


For each node, the statistic is attribute value × outdegree.

(Then summed over all nodes)

Receiver

(Popularity)



For each node, the statistic is attribute value × indegree.

(Then summed over all nodes)

Positive parameter for activity indicates that nodes with large attribute values tend to be more active. Positive parameter for popularity indicates that nodes with large attribute values tend to be more popular.

Possible continuous attribute configurations (non-directed graphs)

Difference

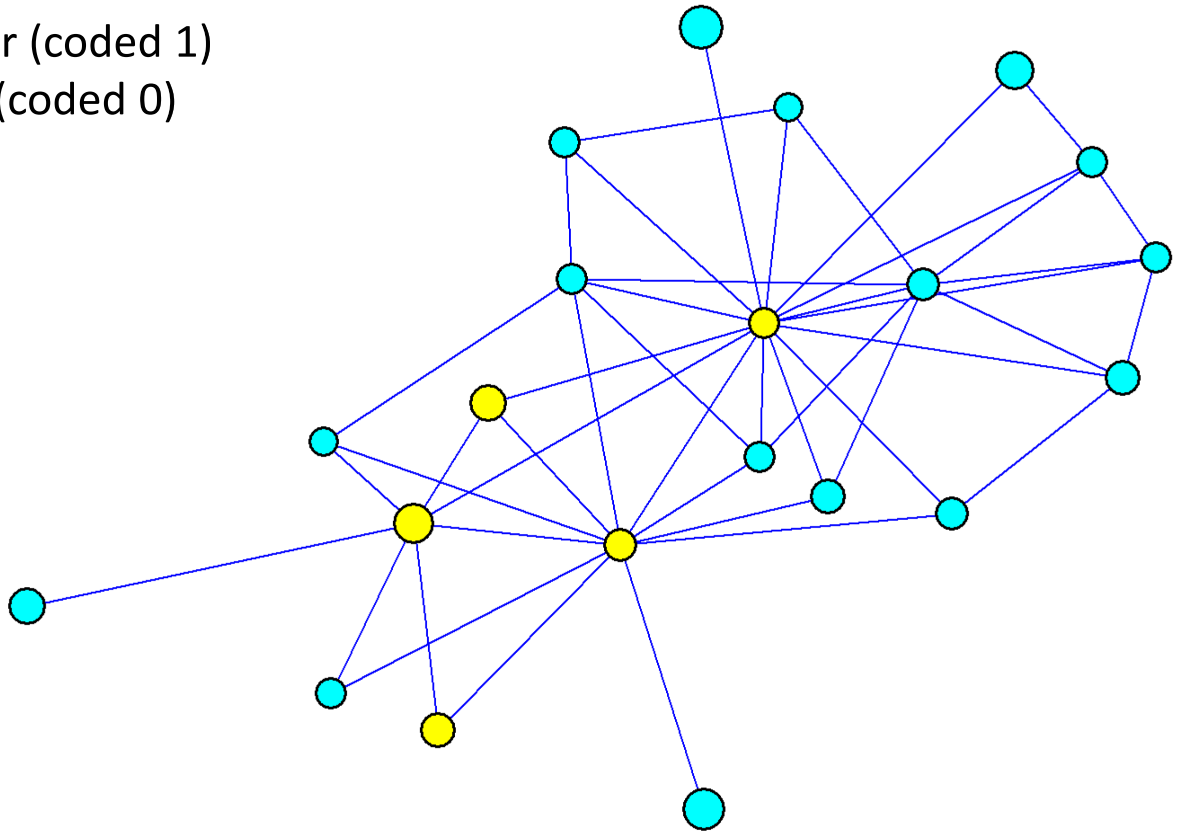
For each pair of tied nodes the statistic is the absolute difference between the attribute values – (then summed over all pairs)

Negative parameter indicates that a smaller absolute difference is associated with the presence of a tie:

Ties are more likely between nodes with similar attribute values - HOMOPHILY

Krackhardt hi-tech managers: Mutual advice network

Binary attribute: *Level*
yellow = senior (coded 1)
blue = junior (coded 0)



Krackhardt hi-tech managers: Mutual advice network

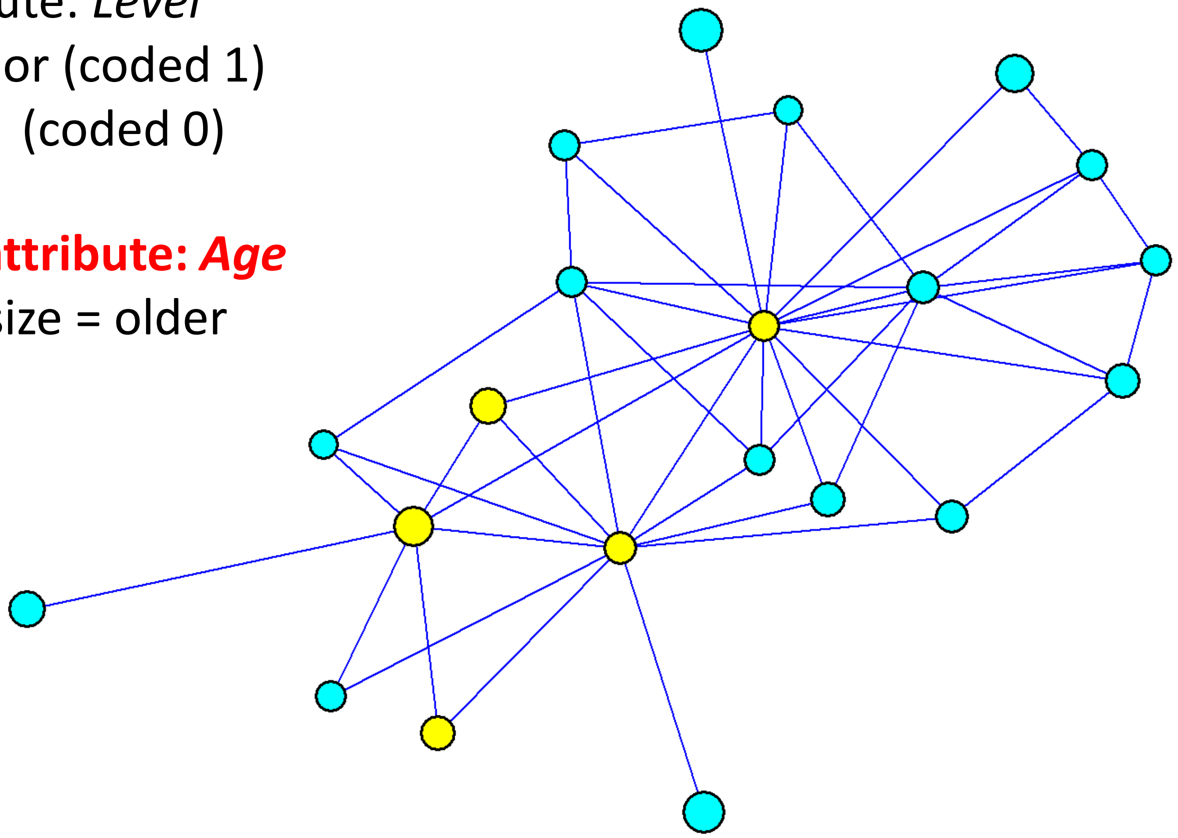
Bernoulli model for *level*

Parameter	Estimate	Standard error	Convergence
Edge	-1.96*	0.29	0.07
Interaction (Homophily)	1.40	1.00	0.002
Activity	0.97*	0.40	0.02

Krackhardt hi-tech managers: Mutual advice network – Continuous attributes

Binary attribute: *Level*
yellow = senior (coded 1)
blue = junior (coded 0)

Continuous attribute: *Age*
Larger node size = older



REMEMBER: Continuous attributes

Difference

For each pair of tied nodes the statistic is the absolute difference between the attribute values – (then summed over all pairs)

$$X_{ij} \, | Y_i - Y_j |$$

REMEMBER: Continuous attributes

Activity (sum)

For each pair of tied nodes the statistic is the sum of the attribute values – (then summed over all pairs)

$$X_{ij}(Y_i + Y_j)$$

Krackhardt hi-tech managers: Mutual advice network

Bernoulli model for *age*

Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence
Edge:	0.85	1.31	-0.01
sum *age*	-0.022	0.018	-0.003
difference *age*	-0.045	0.029	0.01

Small values of parameter estimates for sum and difference reflect the scale of age – these estimates may get multiplied by 60 or more in calculating log-odds.

Dyadic covariate (or dyadic attribute)

- Some other relationship among nodes that could influence the network structure:

  Examples:

  - Formal organisation structure
  - Geography
  - Another network

The machinery of simulation and estimation

Markov Chain Monte Carlo Maximum Likelihood Estimation (MCMCMLE)

- Simulate a distribution of random graphs from a starting set of parameter values
- Change the parameter values by comparing the distribution of graphs against the observed graph
- Repeat until the parameter estimates stabilize. (convergence)

Markov Chain Monte Carlo (MCMC): important points

- To simulate a distribution of graphs we generate graphs using a random walk, changing one edge at a time
- You have to make sure
  - Process has forgotten about the past (burn-in)
  - Chain is “moving freely”
- The “multiplication factor” controls this
  - Large -> more iterations for exploring
  - Large -> more iterations means longer time

How to simulate a graph distribution:
An intuitive description of the Metropolis algorithm

Fixed number of nodes; fixed parameter values

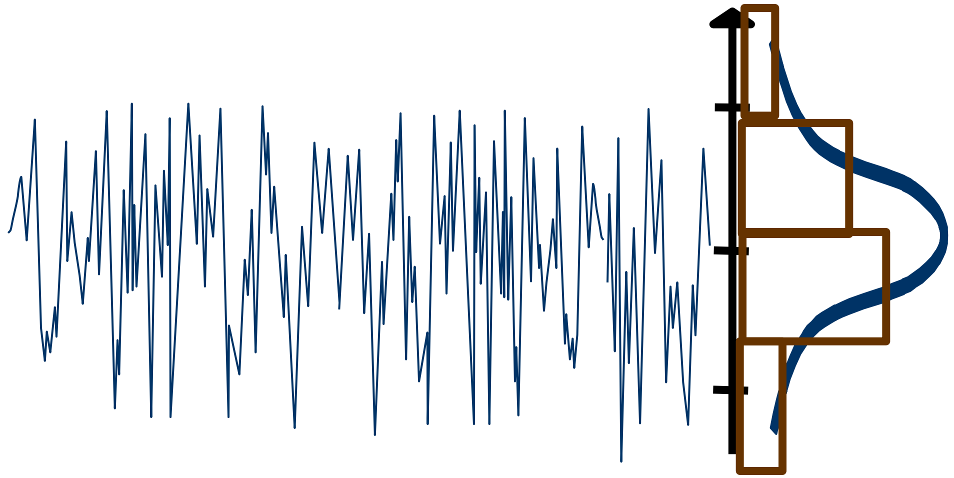
1. Start from a random graph
2. For each step, propose to change one edge at a time
   If the probability of the graph increases, make the change
   If the probability decreases, don’t make the change (EXCEPT SOMETIMES – this makes it a proper statistical distribution)
3. Throw away the early iterations so the starting graph has no effect on the distribution – “burn-in”
4. Sample as many graphs as needed (e.g. every 1000th)
5. Stop after a suitable number of iterations

Random walk for drawing graphs

Make sure you are drawing from target distribution

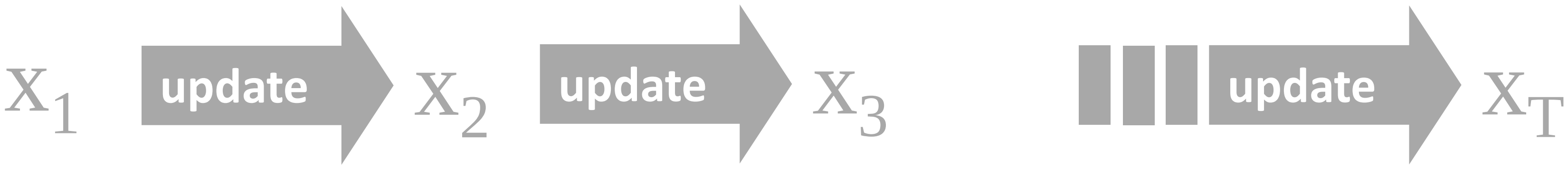
x_1 → update → x_2 → update → x_3 … update → x_T

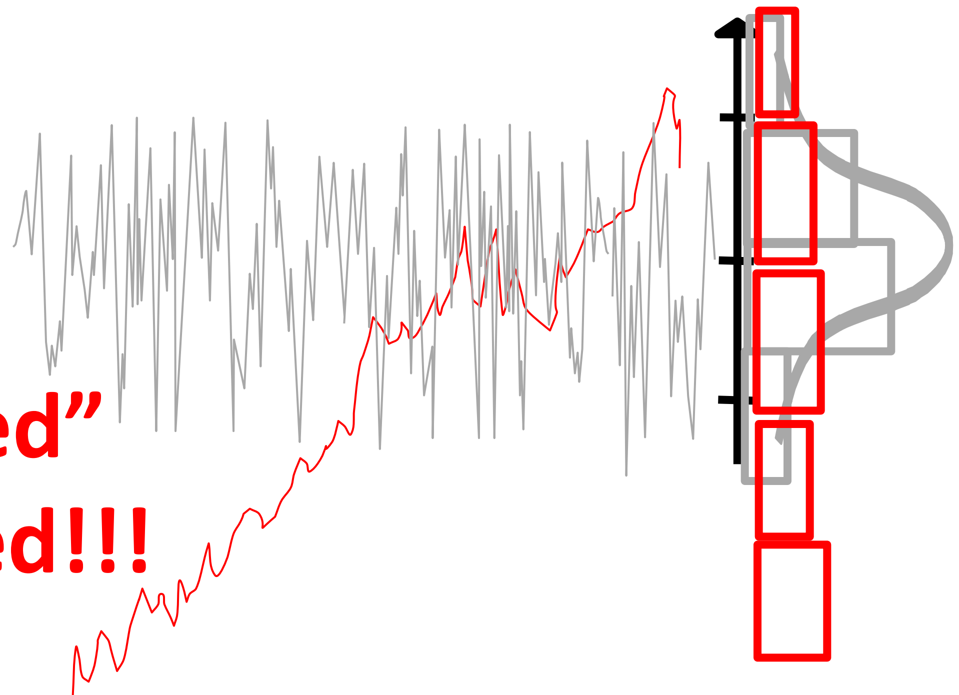
Simulate **graphs**



Random walk for drawing graphs

Make sure you are drawing from target distribution





Not “converged”

Not settled!!!

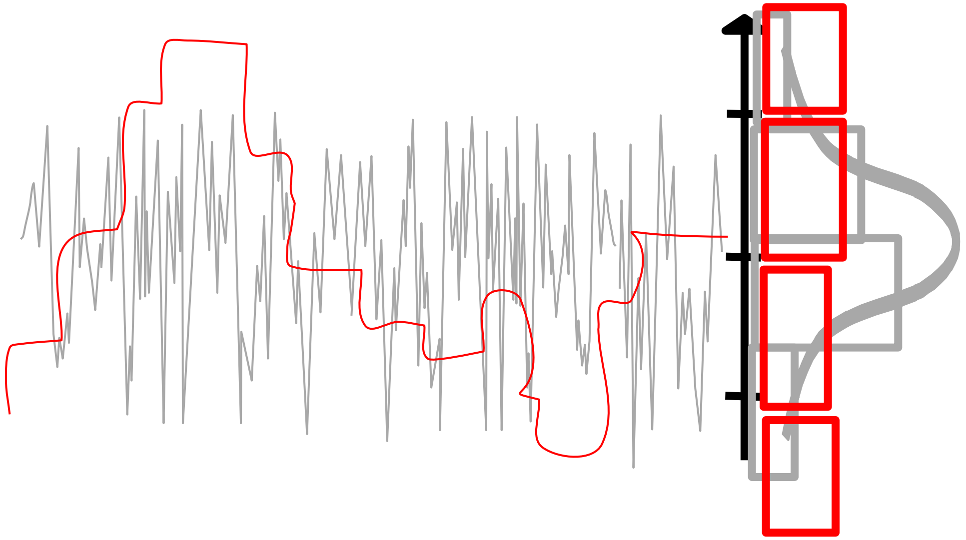
Random walk for drawing graphs

Make sure you are drawing from target distribution

x_1 update → x_2 update → x_3 ... update → x_T

Bad mixing

Too large moves



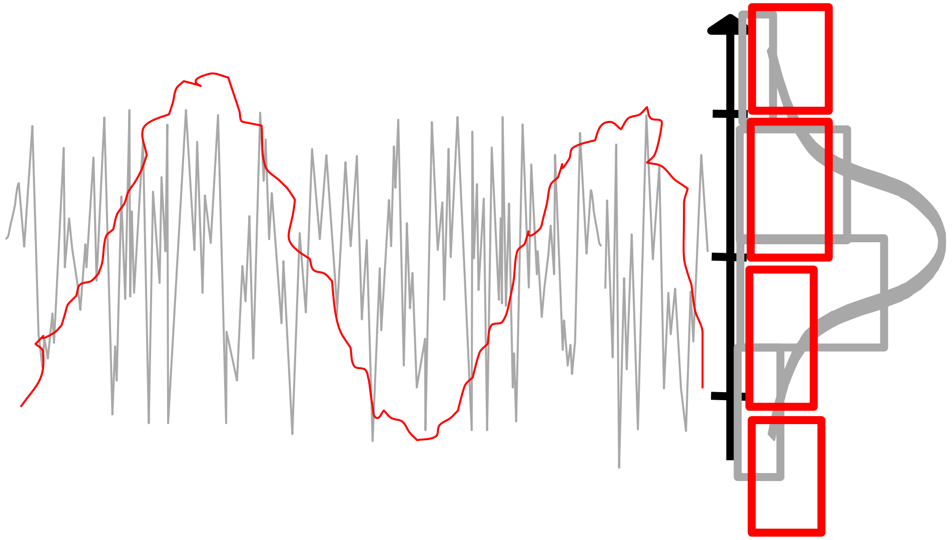
Random walk for drawing graphs

Make sure you are drawing from target distribution

x_1 update x_2 update x_3 ... update x_T

Bad mixing

Too small moves



The technicalities:
Maximum Likelihood- “Method of moments”

- Find those parameter values such that the average number of configurations in the distribution equals the observed values

$$E_{\hat{\theta}}\{z(X)\} = z(x_{obs})$$

...as we have observed

We get the same number of configurations (in expectation – i.e. average)

The technicalities:
Maximum Likelihood- “Method of moments”

Snijders (2002):

Algorithm for solving equation

$E_{\theta}\{z(X)\} = z(x_{obs})$ for θ:

- *Phase 1*: Initialize
- *Phase 2*: Estimate
- *Phase 3*: Check!

The technicalities:
Maximum Likelihood- "Method of moments"

- *Phase 1*: Gets the process started to reach some very rough and ready parameter "guesses"

- *Phase 2*: Several subphases:
    - In each subphase we simulate distributions of graphs.
    - Check the observed graph against the simulation
    - Change the parameter estimate
    - Stop at the end of the specified number of subphases

- *Phase 3*: Simulates from the final parameter estimates, checks convergence and estimates standard errors

NOTE: **CONVERGENCE IS NOT GUARANTEED!!**

- So:
  - If hard to get convergence, try with bigger multiplication factor
  - If close to convergence, can use a smaller gaining factor (already quite precise)
  - Results differ slightly from run to run – this is a stochastic algorithm

Parameter interpretation

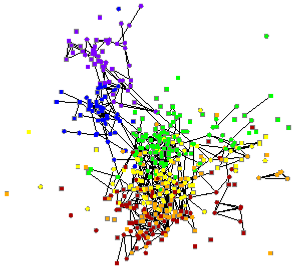
- Note that for purely structural (i.e., endogenous) network effects:
  - Negative parameter = less of such substructures
  - Positive parameter = more of such substructures

This can differ for actor-relation (attribute) effects

Goodness of fit

Goodness of fit

DATA



MODEL

MODEL PARAMETER ESTIMATES

SIMULATE A DISTRIBUTION OF GRAPHS

INSPECT THE GRAPH STATISTIC FROM DATA

TAKE A SAMPLE OF SIMULATED GRAPHS AND INSPECT DISTRIBUTION OF ANY GRAPH STATISTIC OF INTEREST

IS THE DATA EXTREME IN THE DISTRIBUTION?

YES: THIS MODEL CANNOT EXPLAIN THIS STATISTIC

NO: AS FAR AS THIS GRAPH FEATURE IS CONCERNED THE DATA COULD HAVE COME THIS MODEL

GOODNESS OF FIT

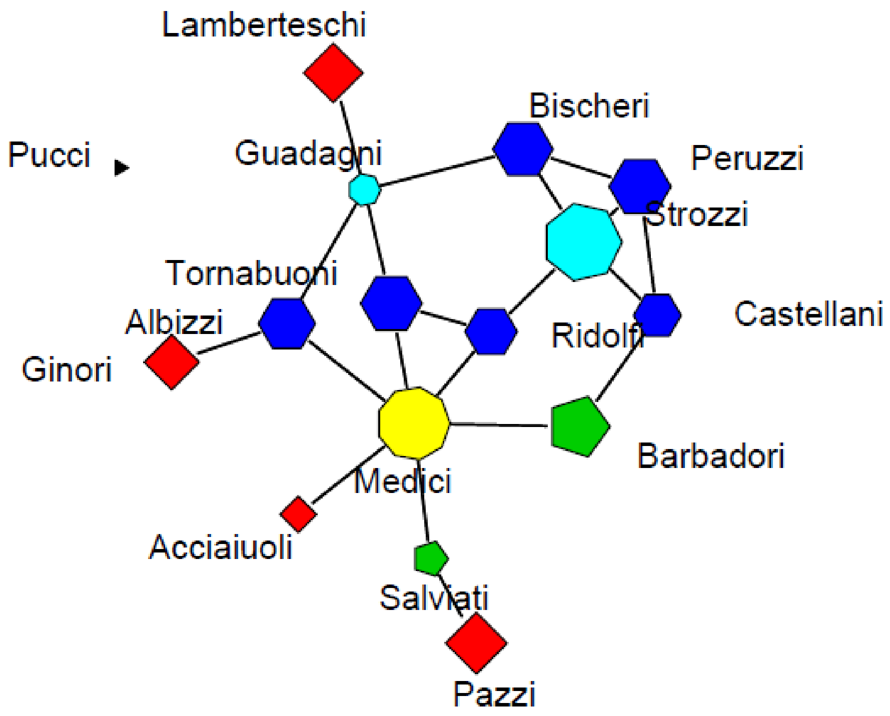
With thanks to Steve Goodreau

Goodness of fit (GOF)

- Estimate parameters
- Simulate a distribution of graphs using these parameters
- From the simulation, collect graph statistics of any sort
- Compare the observed data with the collected statistics:
  - If the data is not extreme (e.g. $|t| < 2.0$), then the model plausibly explains that feature of the data
  - For parameters in the model, we want the data to be central in the distribution (say, $|t| < 0.2$), else model may not have converged.

Goodness of fit (GOF)

A well-known example in social networks:



Florentine marriage data

- Edge indicates marriage tie between families
- Sides=degree + 3
- Color=degree
- Size=log(wealth)

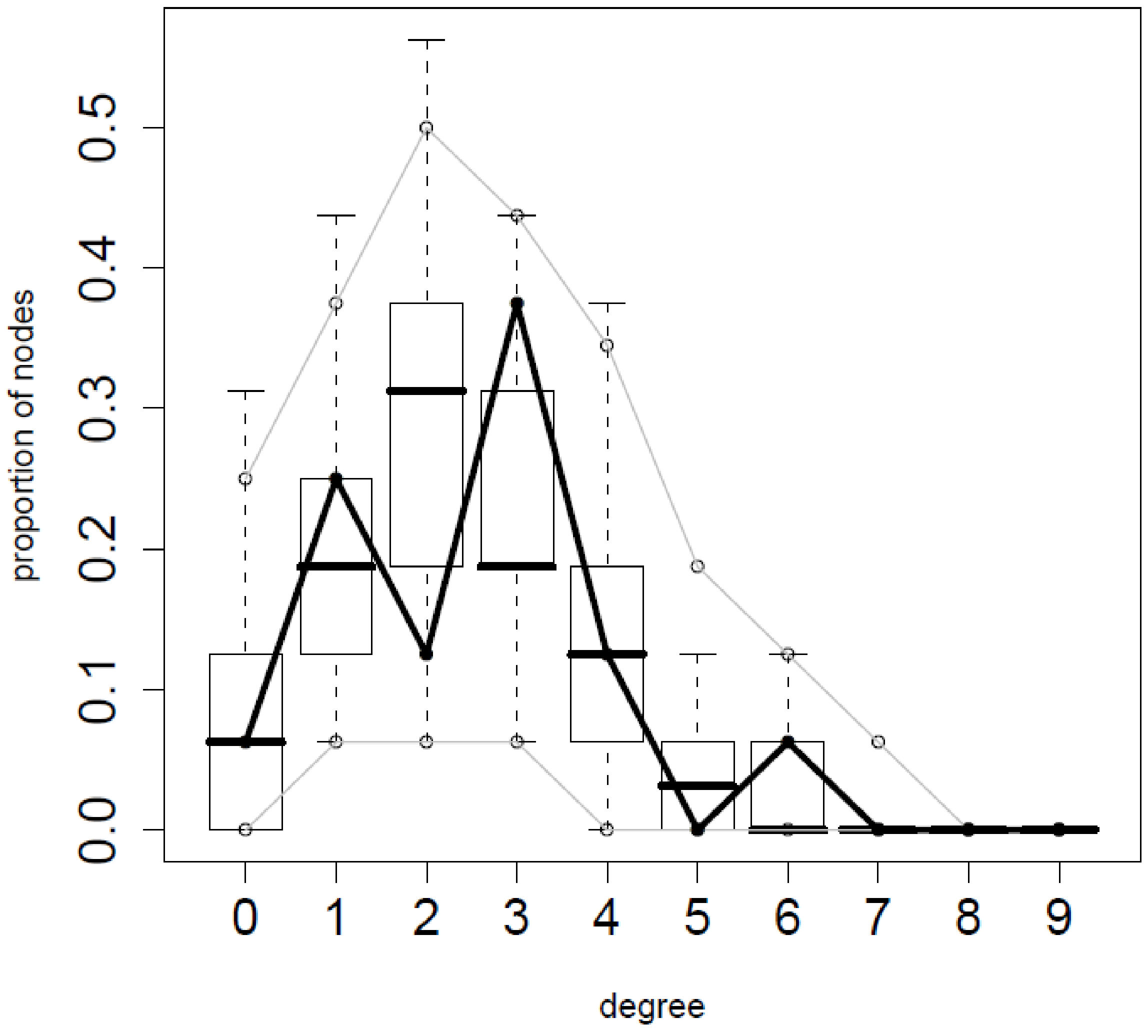
```
model1 <- ergm(flomarriage ~ edges + kstar(2))
```

Model fit

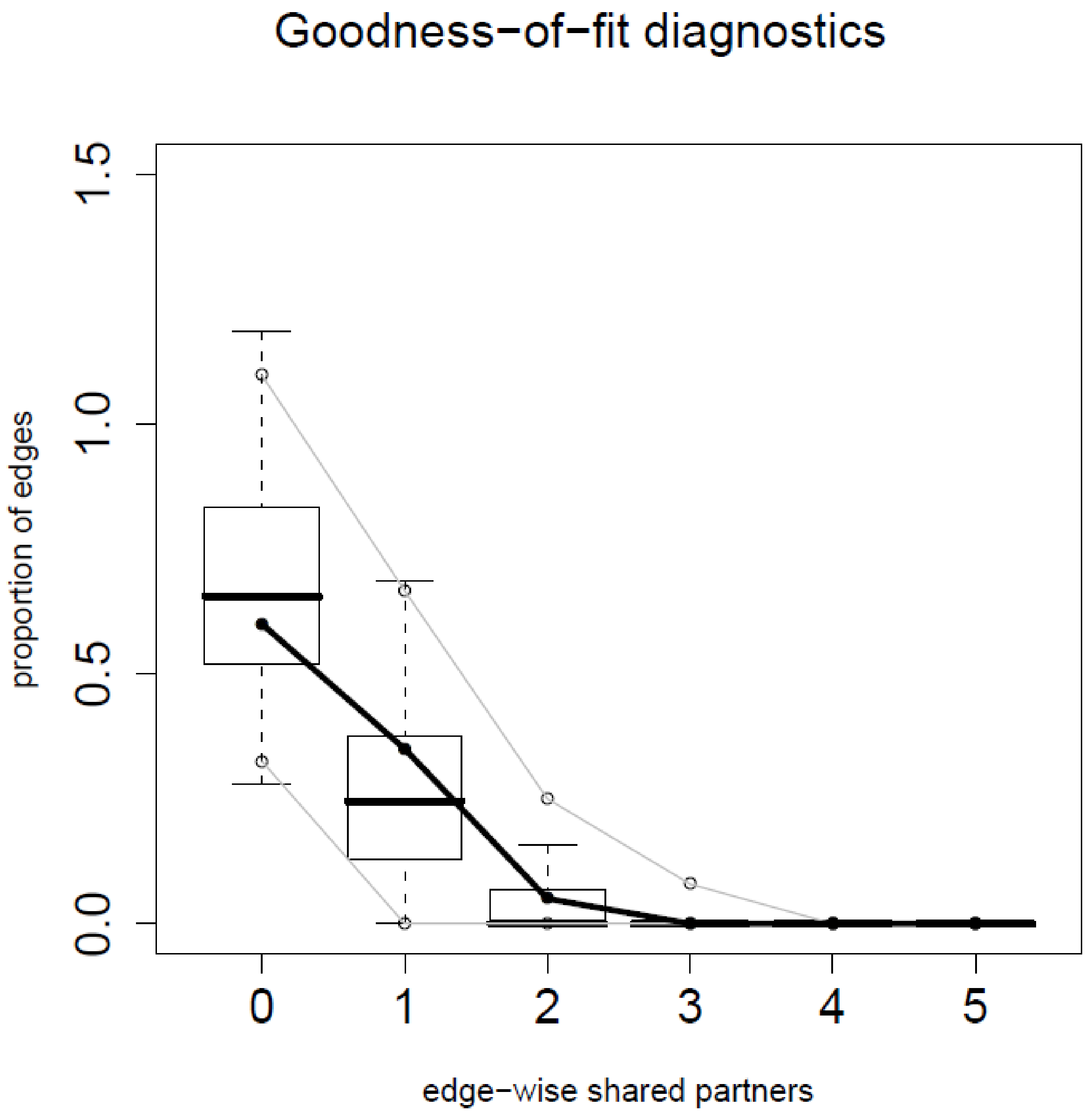
```
R> model1 <- ergm(flomarriage ~ edges + kstar(2))
R> summary(model)
...
Monte Carlo MLE Results:
       estimate  s.e.      p-value MCMC s.e.
edges  -1.651664 0.8342    0.050   0.008394
kstar2  0.009063 0.1705    0.958   0.001703

Null Deviance:  166.355 on 120 degrees of freedom
Residual Deviance:  108.126 on 118 deg of freedom
Deviance:  58.229 on 2 degrees of freedom
```

```
model1 <- ergm(flomarriage ~ edges + kstar(2))
```

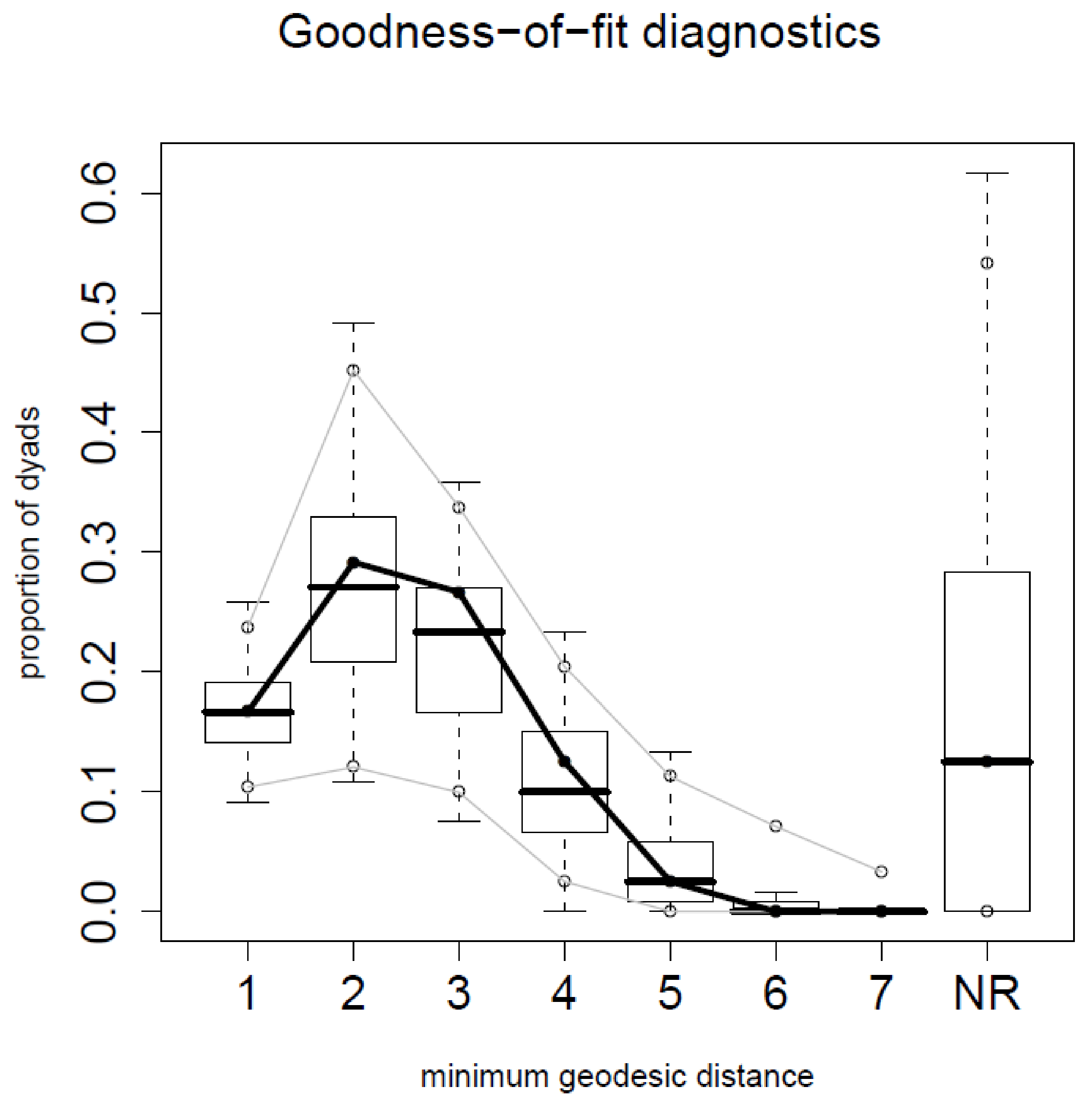
Goodness−of−fit diagnostics

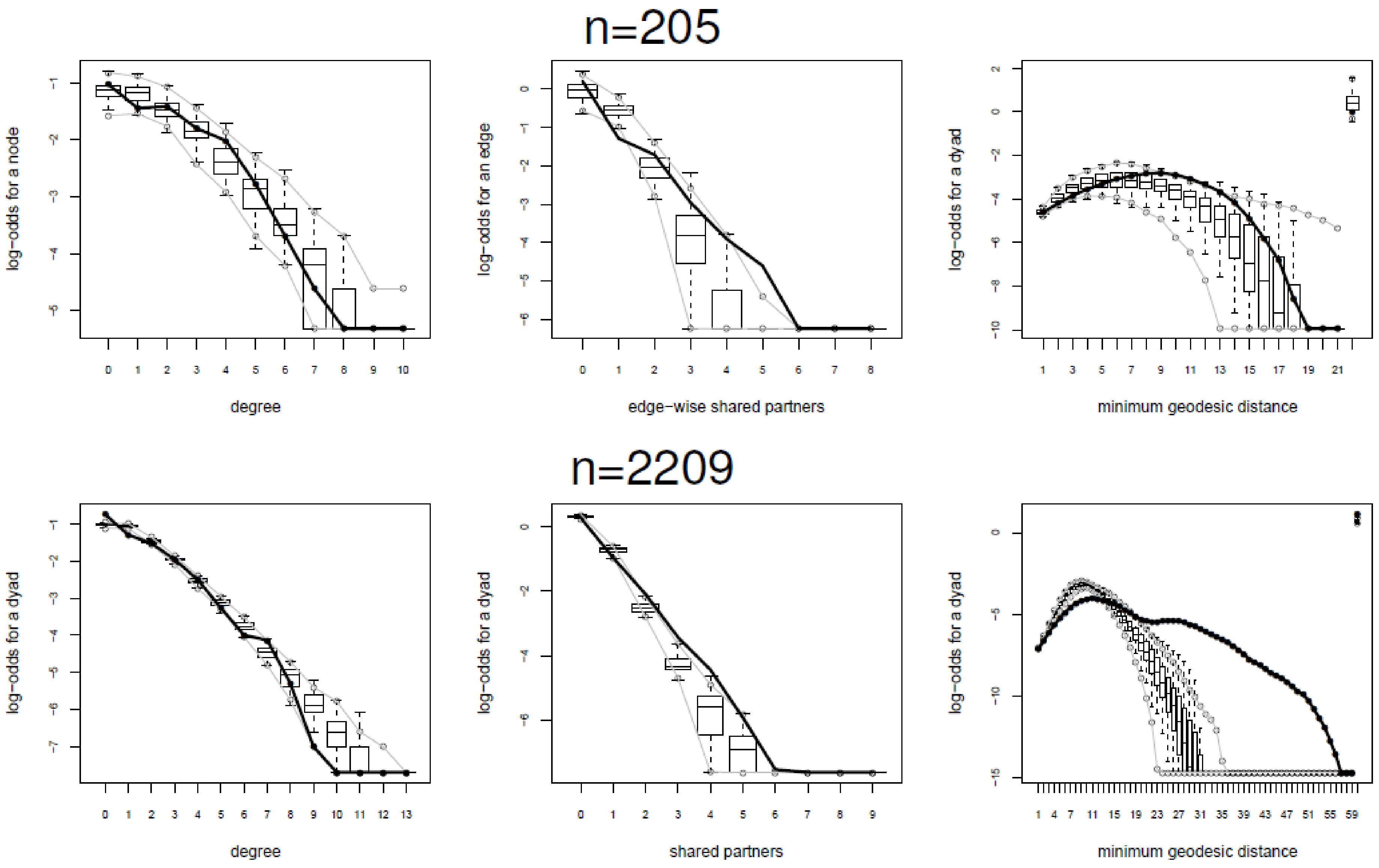


```
model1 <- ergm(flomarriage ~ edges + kstar(2))
```



```
model1 <- ergm(flomarriage ~ edges + kstar(2))
```





Hunter, Goodreau, Handcock (2008, *JASA*), to appear.