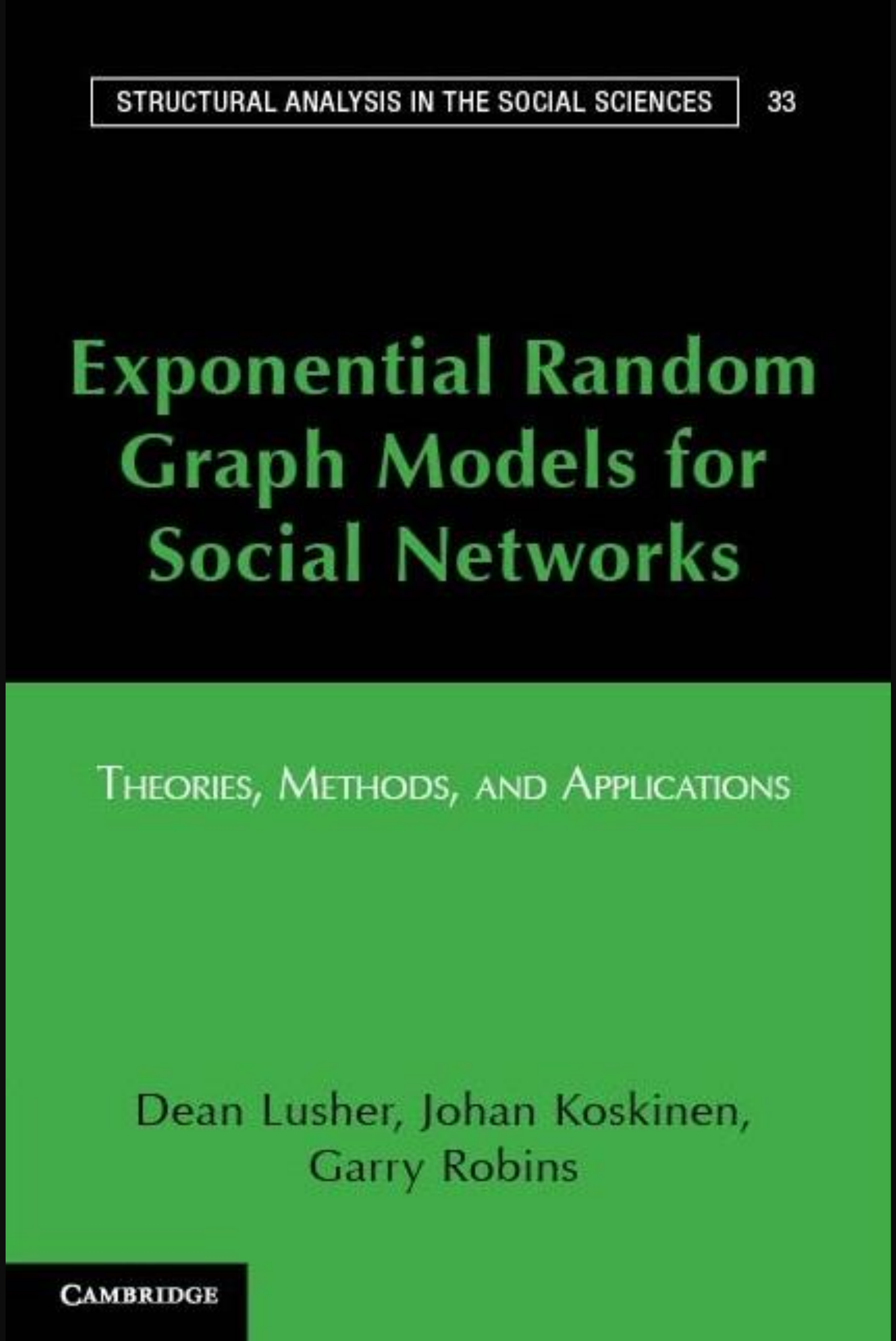
# Introduction to Exponential Random Graph Models

Peng Wang

# Reference material

- **Lusher, Koskinen & Robins (2013)** can provide more detail about the issues we present on ERGMs.



# Exponential random graph models (ERGMs)

What are they?

Why use them?

# Exponential random graph models

(Frank & Strauss, 1986; Wasserman & Pattison, 1999; Robins et al, 2009; Snijders et al, 2006)

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

Tie prediction tool

## ERGMs

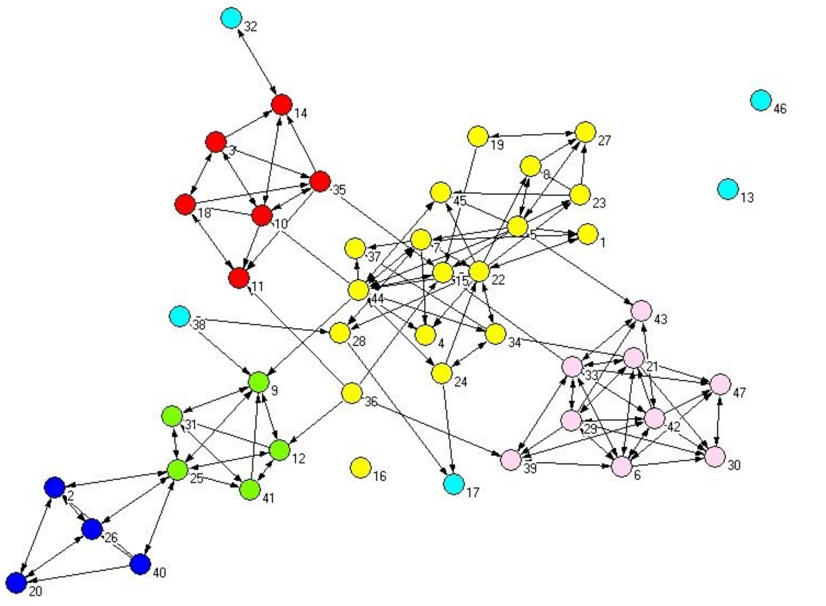
- ERGM is a tie-based model
- Predicting ties
- **Social selection**
- Number of data points is the number of possible ties:
    - $n(n\text{-}1)$

## ALAAMs

- ALAAM is a node-based model
- Predicting node outcomes
- **Social influence**
- Number of data points is the number of actors:
    - $n$

Snijders et al (2006) SAOM can do both simultaneously
But only with longitudinal data

# Elements of a “<u>social</u> network”



- Locally emergent structures
  - Local patterns form global structure
- Network ties self-organize
  - Through dependency of ties: it is because of one tie that another is formed (but also form due to actor attributes)
- Network patterns evidence of ongoing structural proc
  - Static trace of dynamic social processes
- Multiple processes can operate simultaneously
- Social networks are structured, yet stochastic
  - Structure and randomness

ERGM can take account of these important elements of

Local Structures

represent

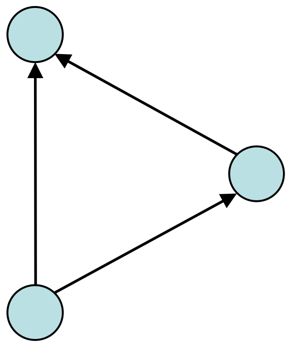
Social Processes

# Local everyday social “rules

Hypotheses about tie formation

- You scratch my back, I’ll scratch yours

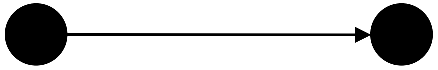
- A friend of a friend is a friend



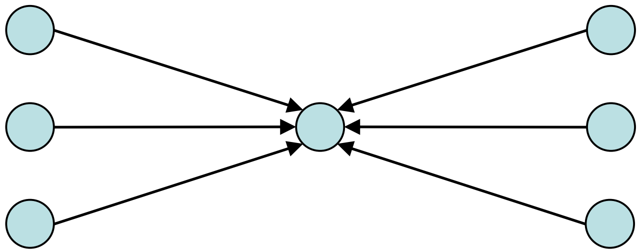
Social rules (processes) have related

- The broker

- Birds of a feather flock together



- Follow the crowd



# Multiple Processes

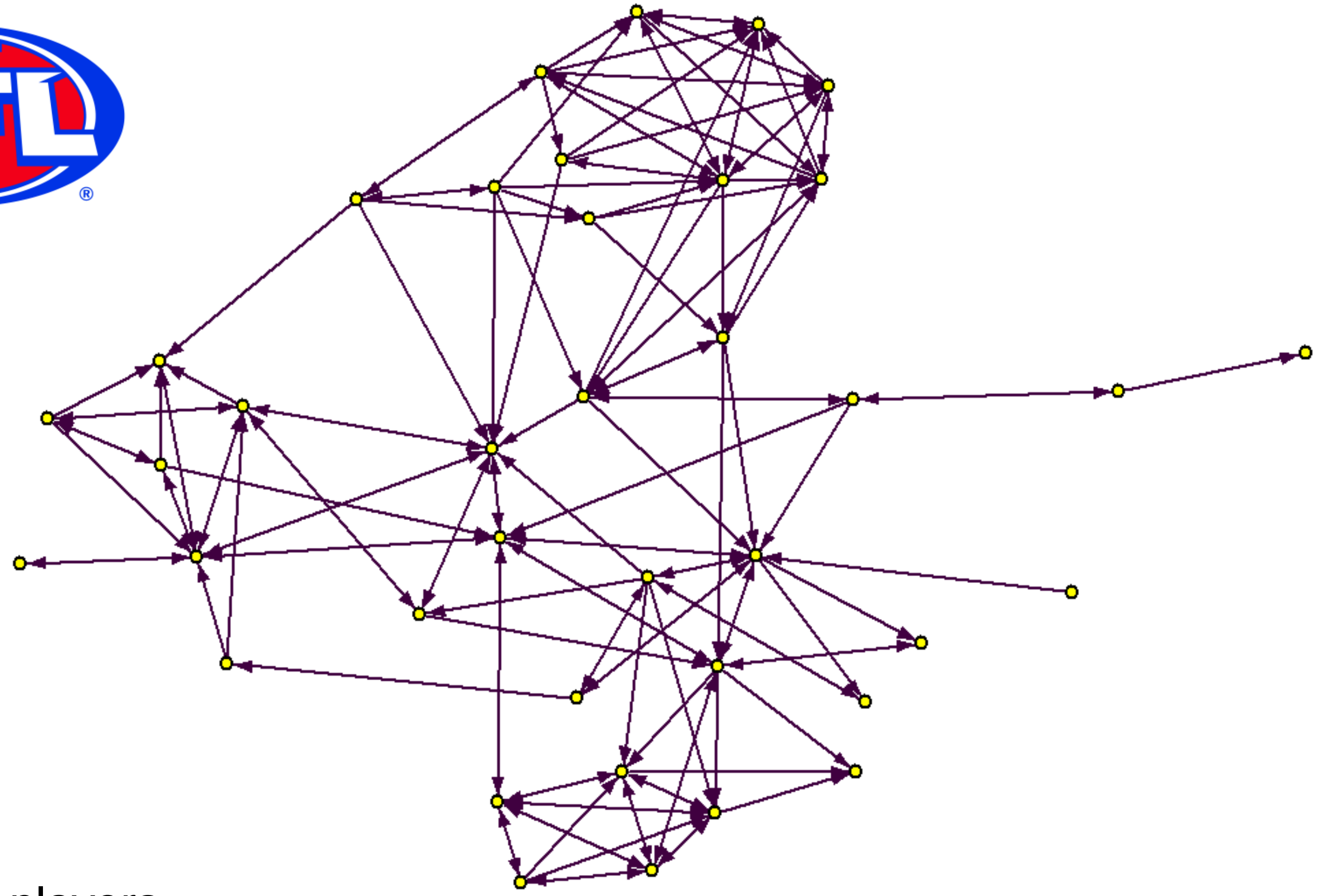
# Network concepts

- Reciprocity
- Transitivity
- Degrees
- Homophily
- Brokerage
- Clustering
- Popularity
- ……

- Until now, we have mostly analysed each of these separately

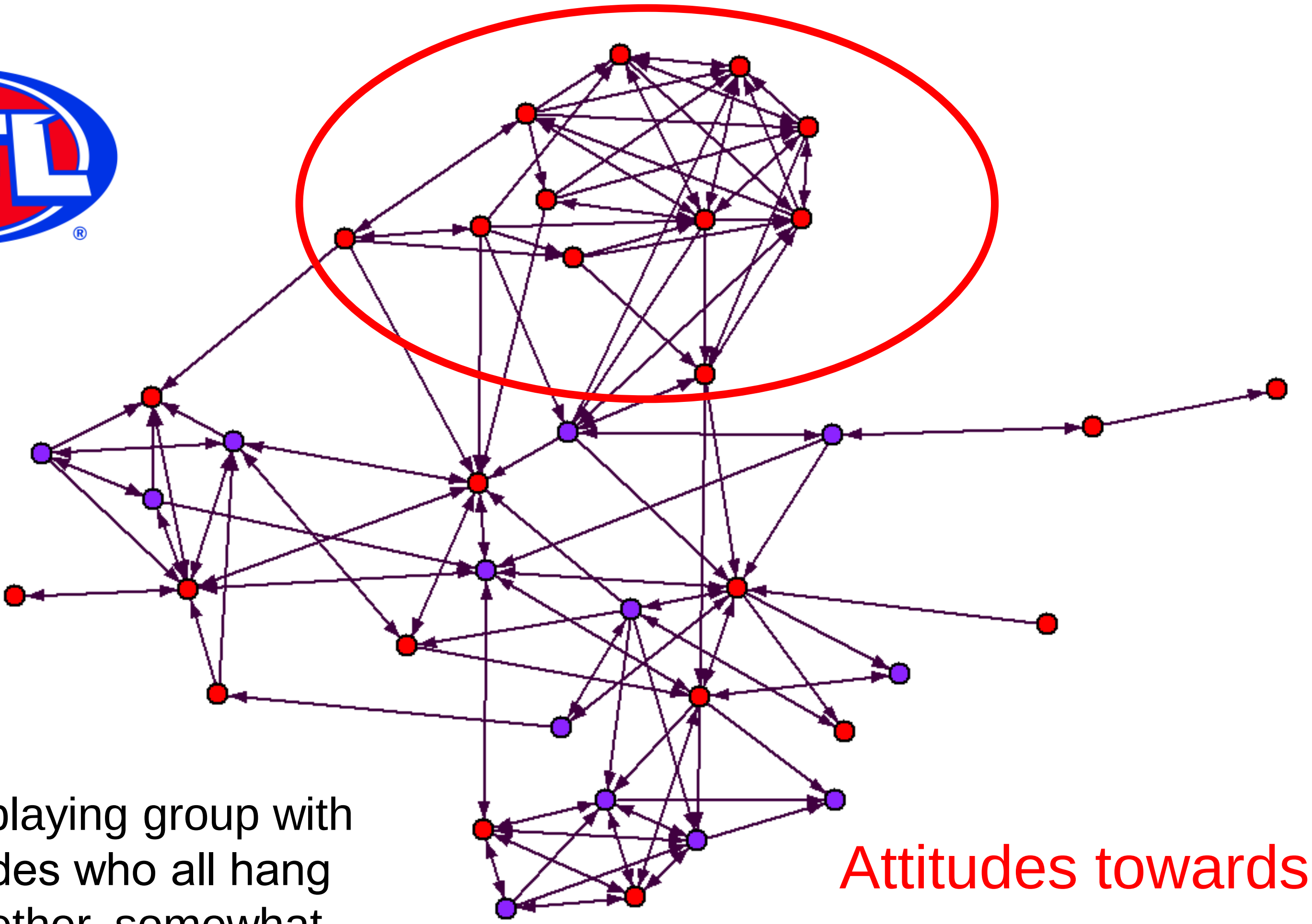
- Would you do a series of correlations, or would you use multiple regression?

# *After hours* socialising network



**Dots** are players
**Arrows** point to players who are selected as "*someone I socialize with*"

# *After hours* socialising network

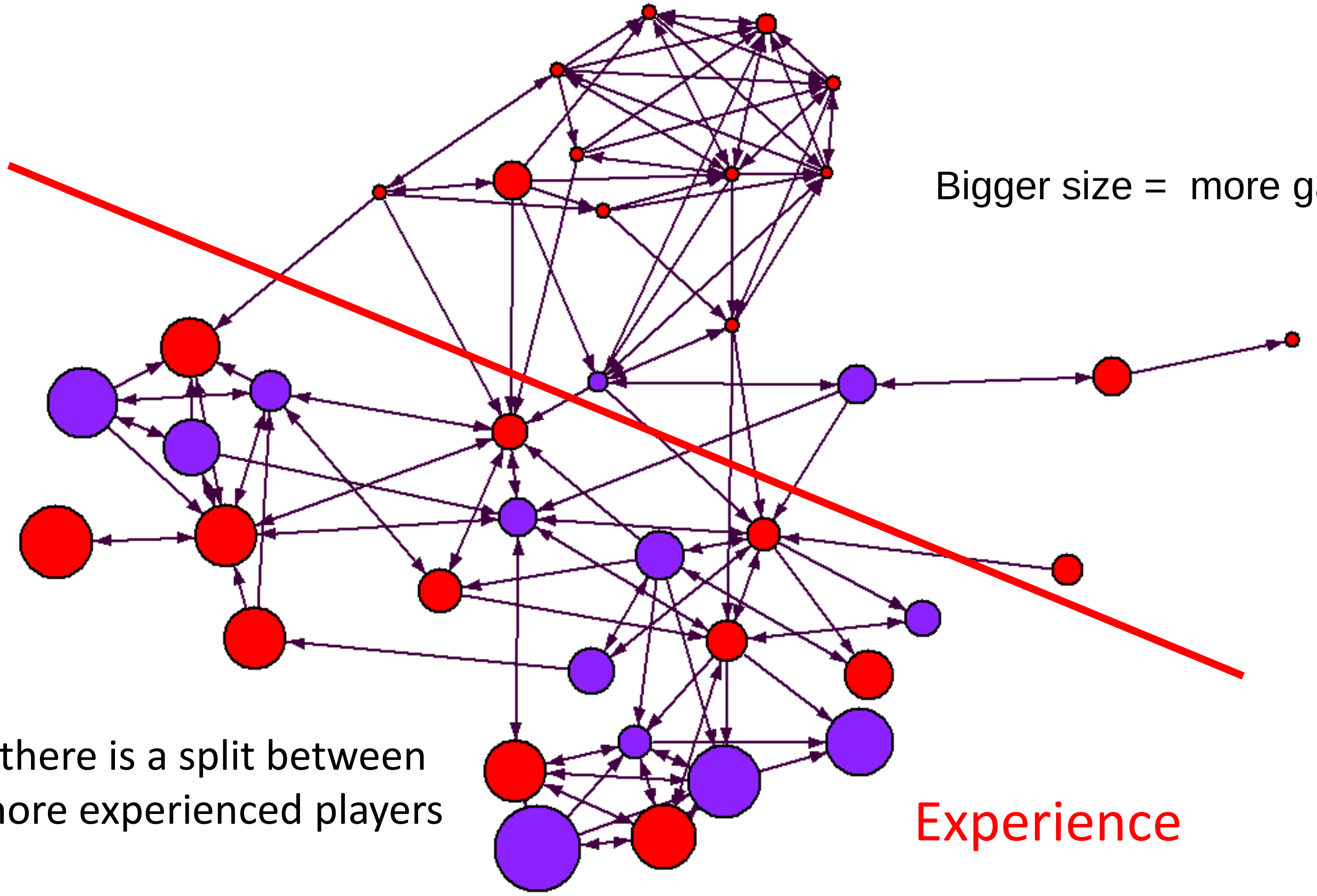


Section of playing group with “bad” attitudes who all hang around together, somewhat separate to the rest of the players

Attitudes towards women

**Red** = LOW on respect & responsibility; **Purple** = HIGH on respect

# *After hours* socialising network

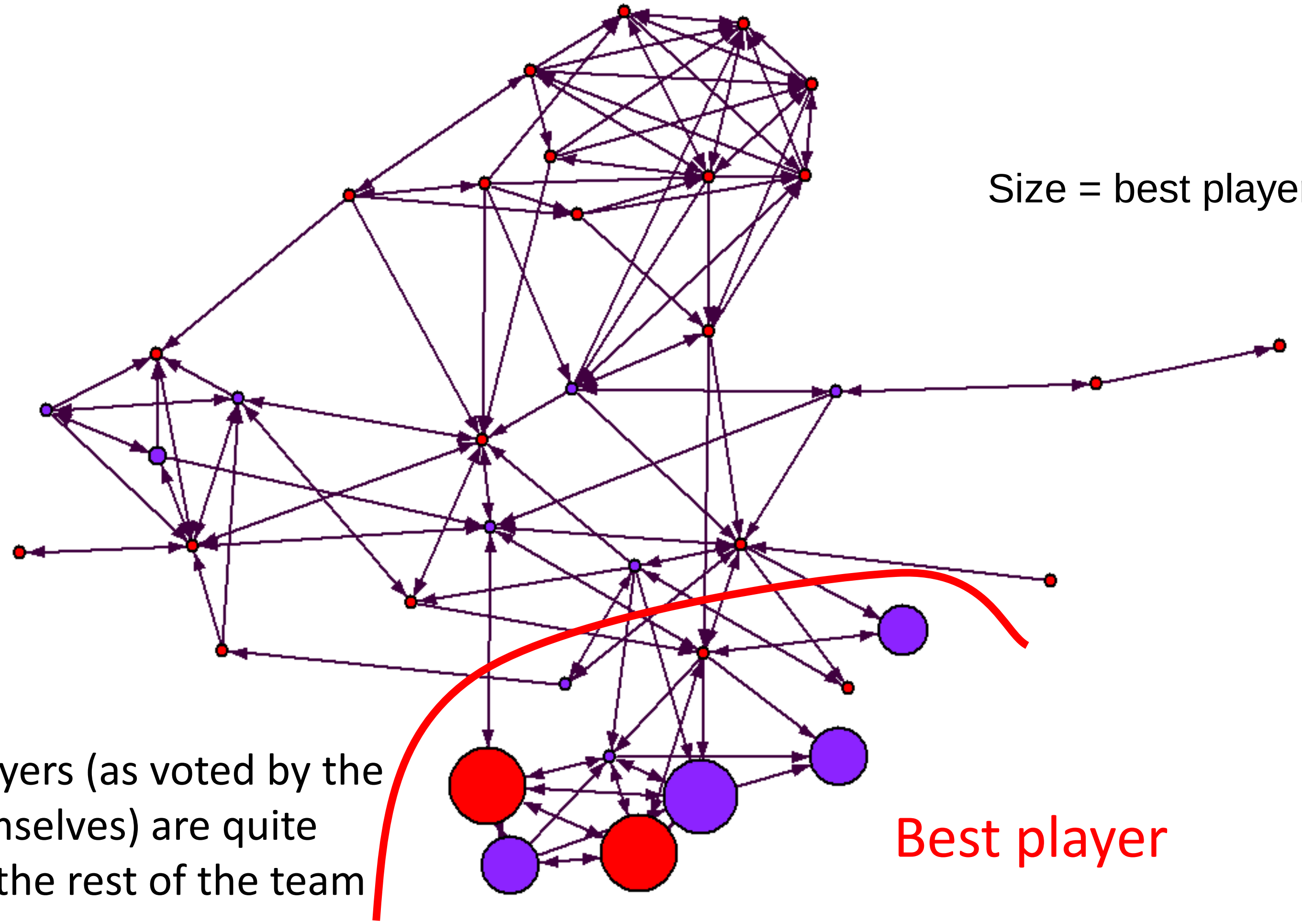



Bigger size = more games played

Note that there is a split between less and more experienced players

**Red** = LOW on respect; **Purple** = HIGH on respect

# *After hours* socialising network



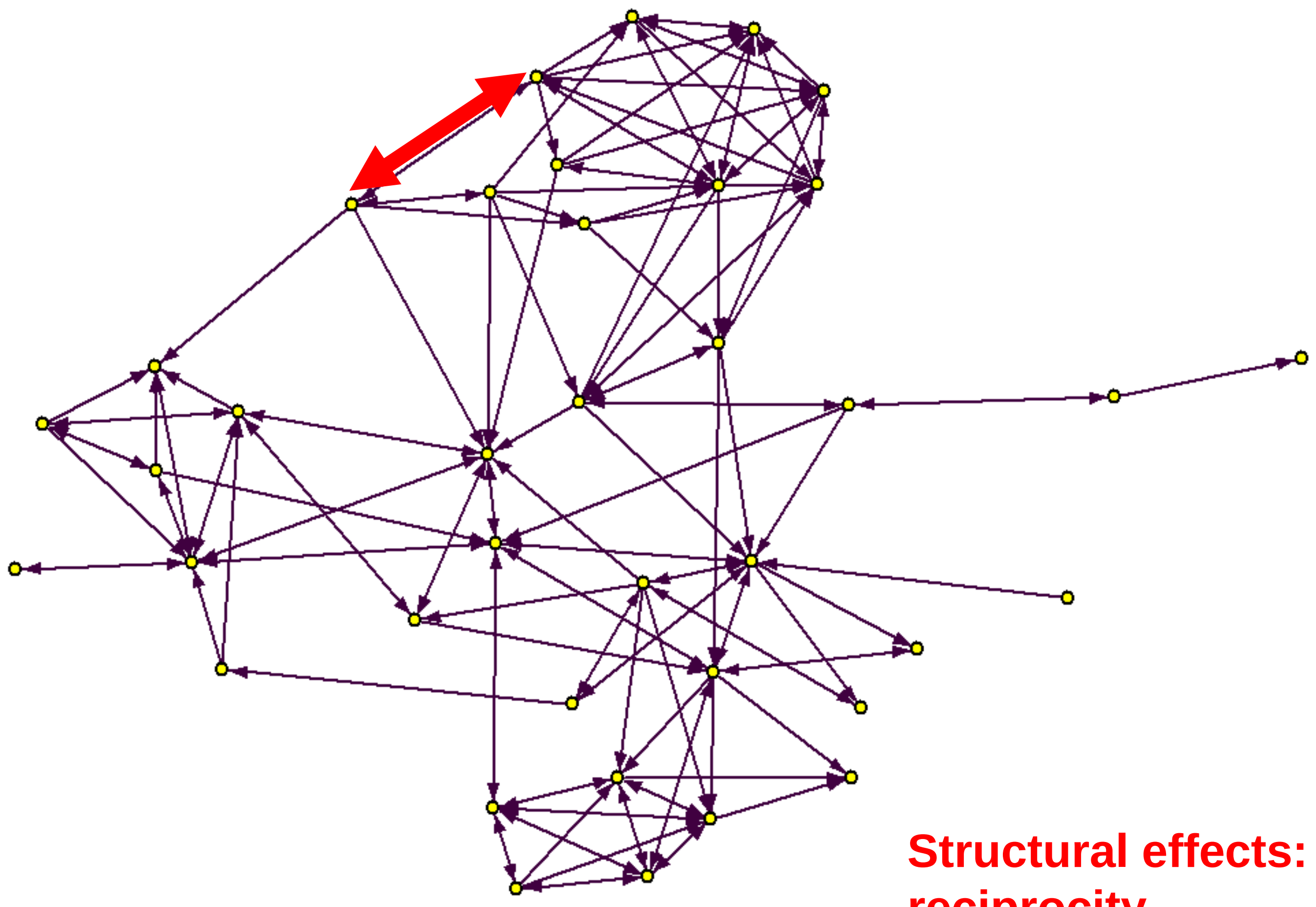
Size = best player

The best players (as voted by the players themselves) are quite separate to the rest of the team

Best player

Red = LOW on respect; Purple = HIGH on respect

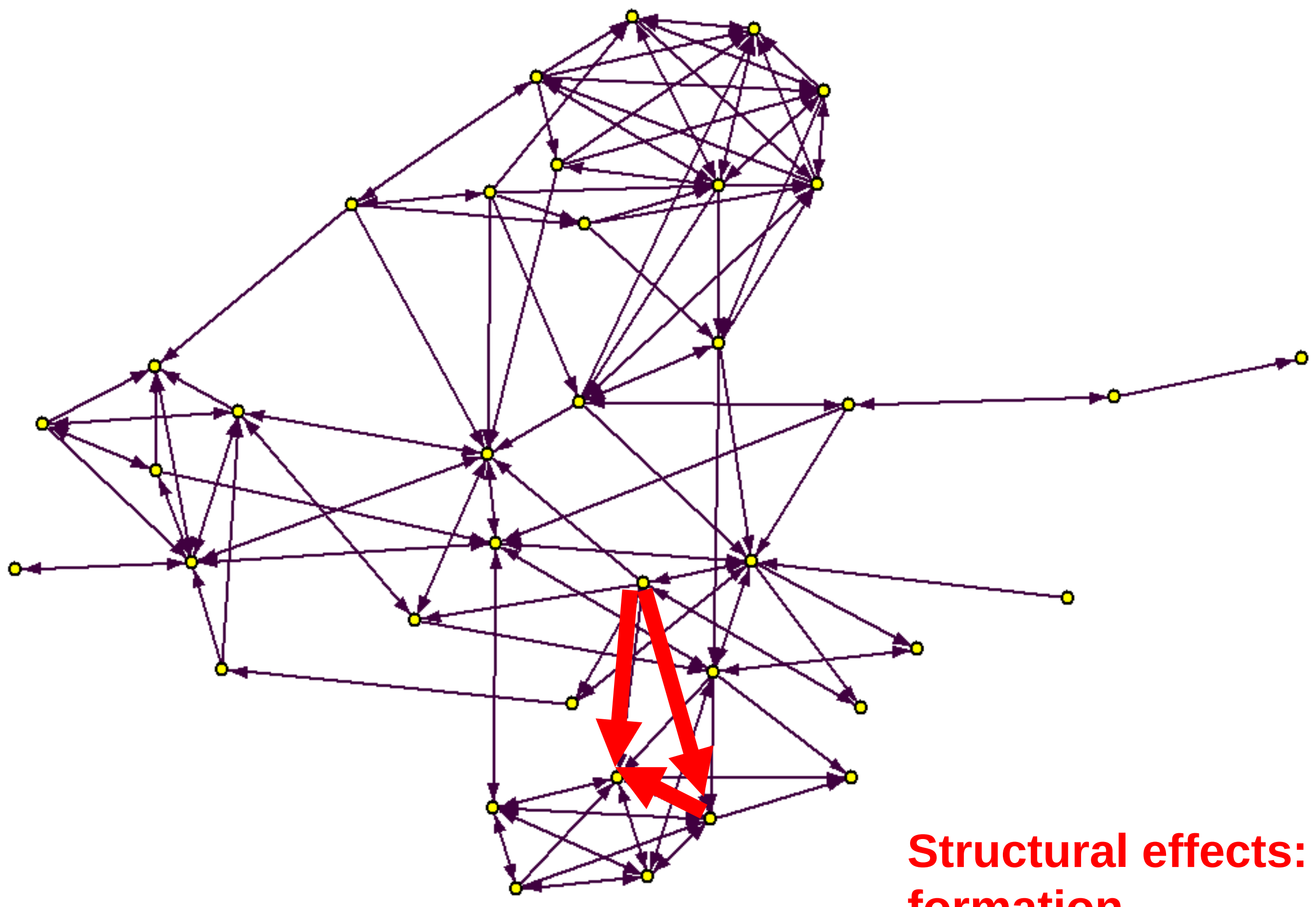
# *After hours* socialising network



**Structural effects: reciprocity**

**Red** = LOW on respect; **Purple** = HIGH on respect

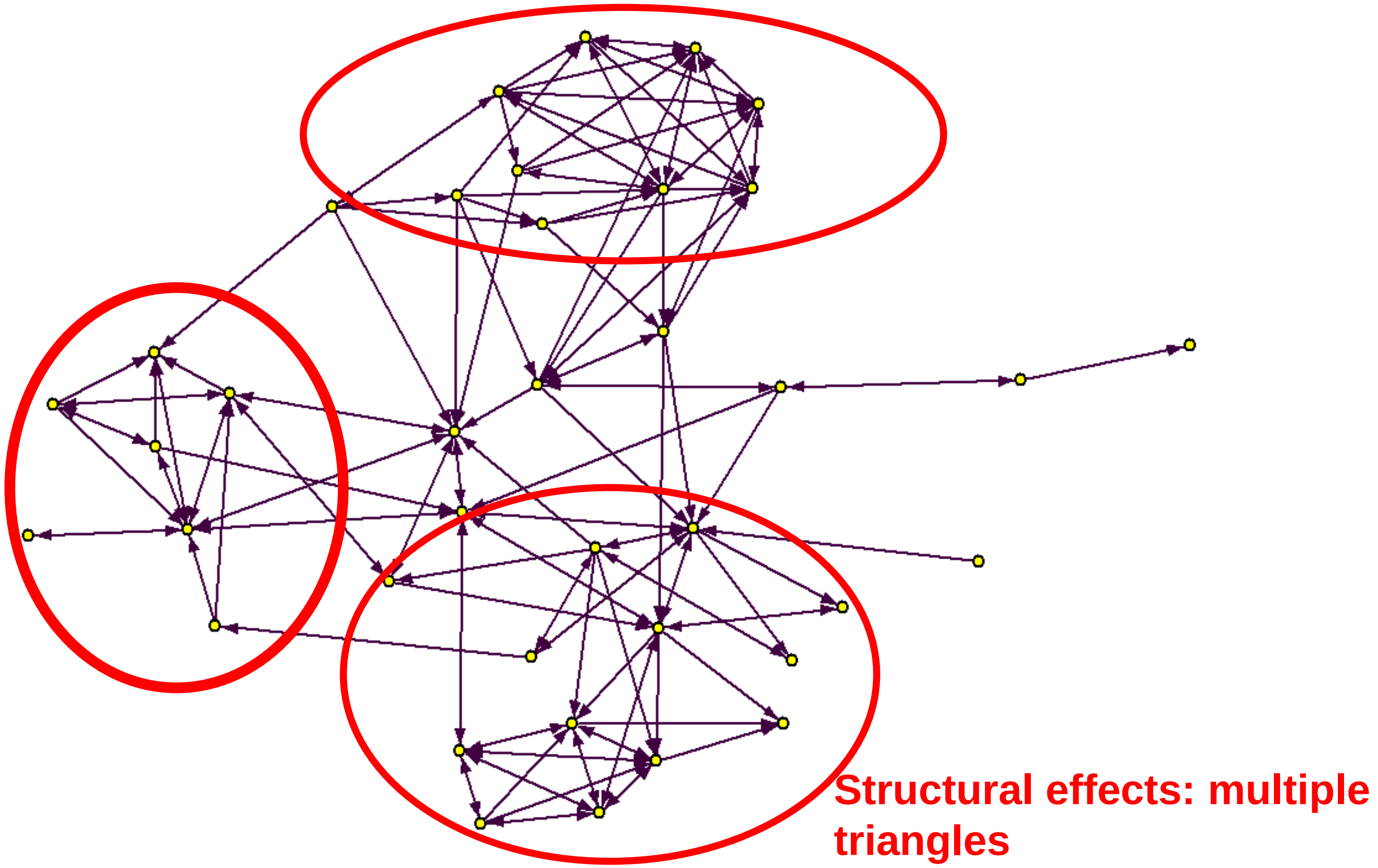
# *After hours* socialising network



**Structural effects: triangle formation**

**Red** = LOW on respect; **Purple** = HIGH on respect

# *After hours* socialising network




**Red** = LOW on respect; **Purple** = HIGH on respect

# *After hours* socialising network

**Competing explanations for tie formation**

Attitudes towards women

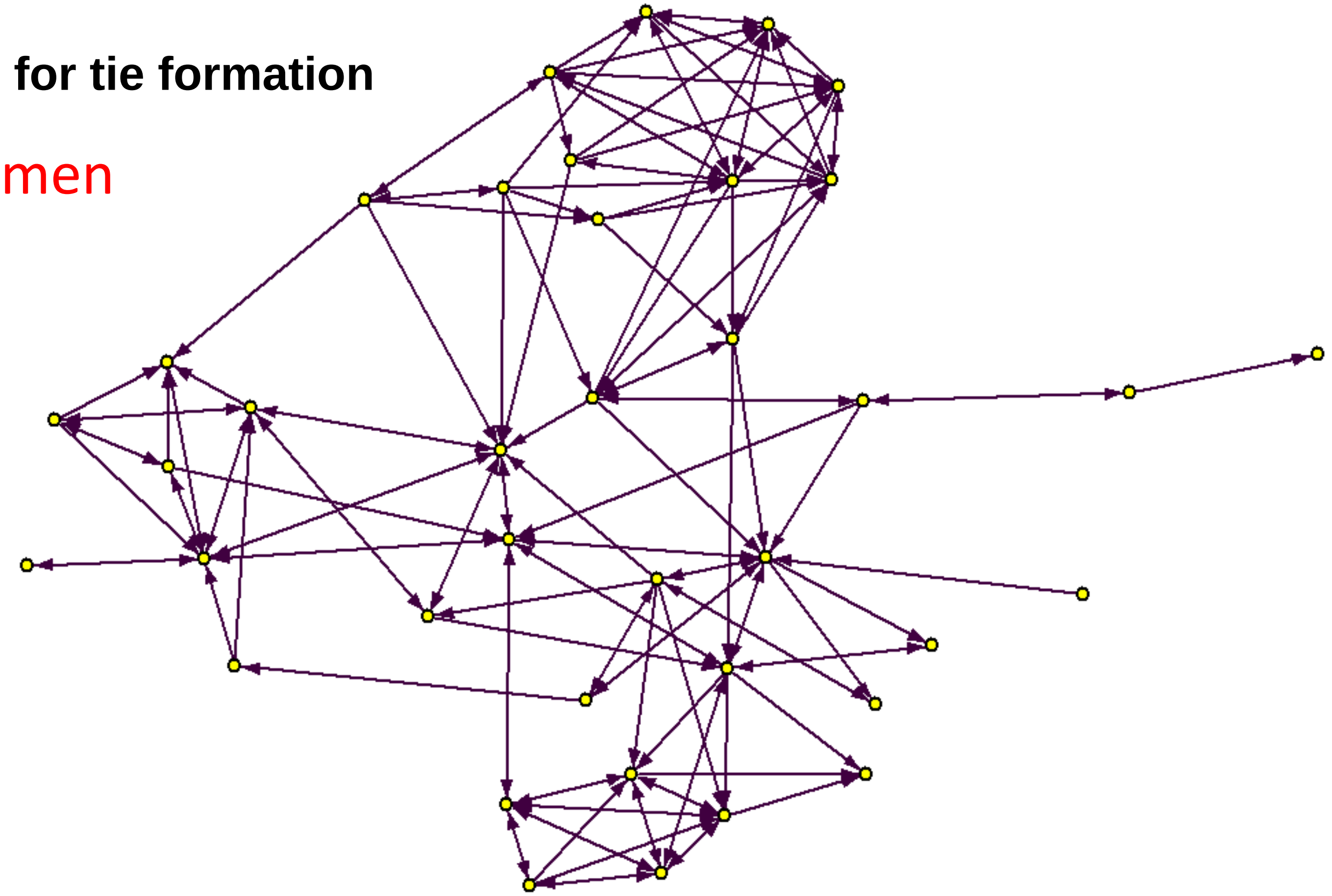
Experience

Best player

Reciprocity

Triangles

Multiple triangles



**Red** = LOW on respect; **Purple** = HIGH on respect

# Multiple social processes

Tie

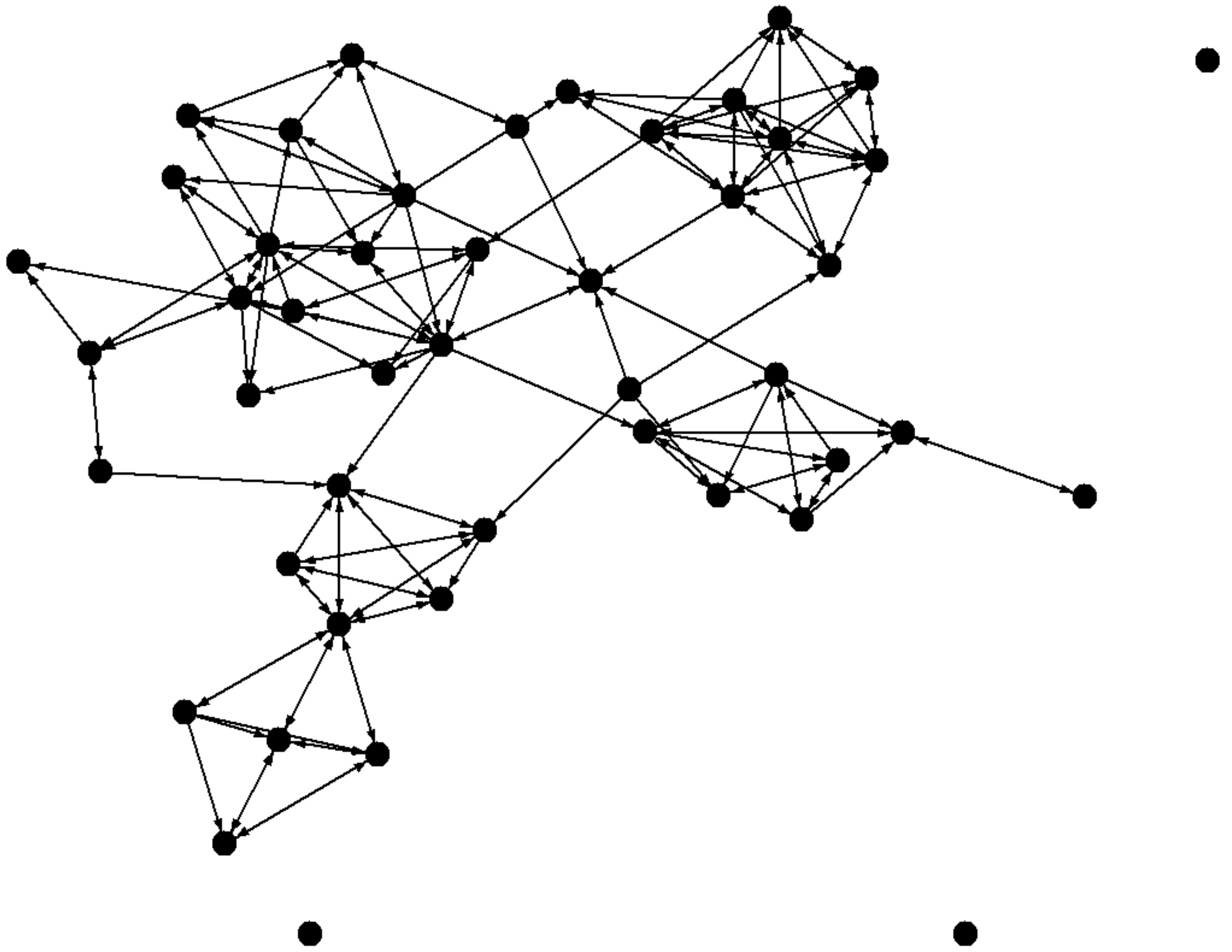
Reciprocity

Activity

Popularity

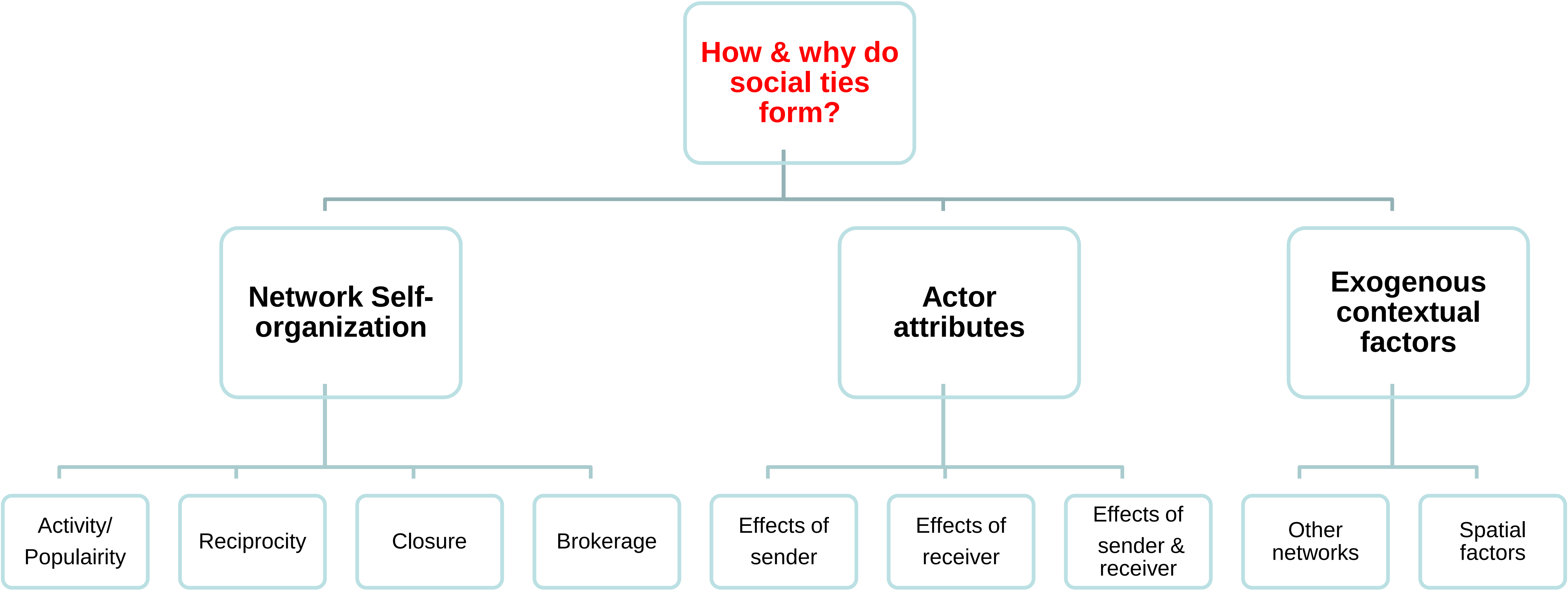
Triads

Brokerage



# Why do network ties form?

## Typology of network structures and social processes

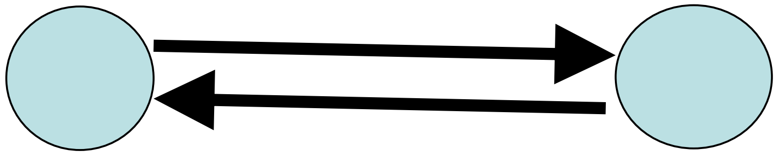


# Dependency

# Local, everyday social rules

….. are underpinned by a very important commonality

E.g., You scratch my back, I'll scratch yours



- It is because you have scratched my back that I will scratch yours

One tie follows the other in time

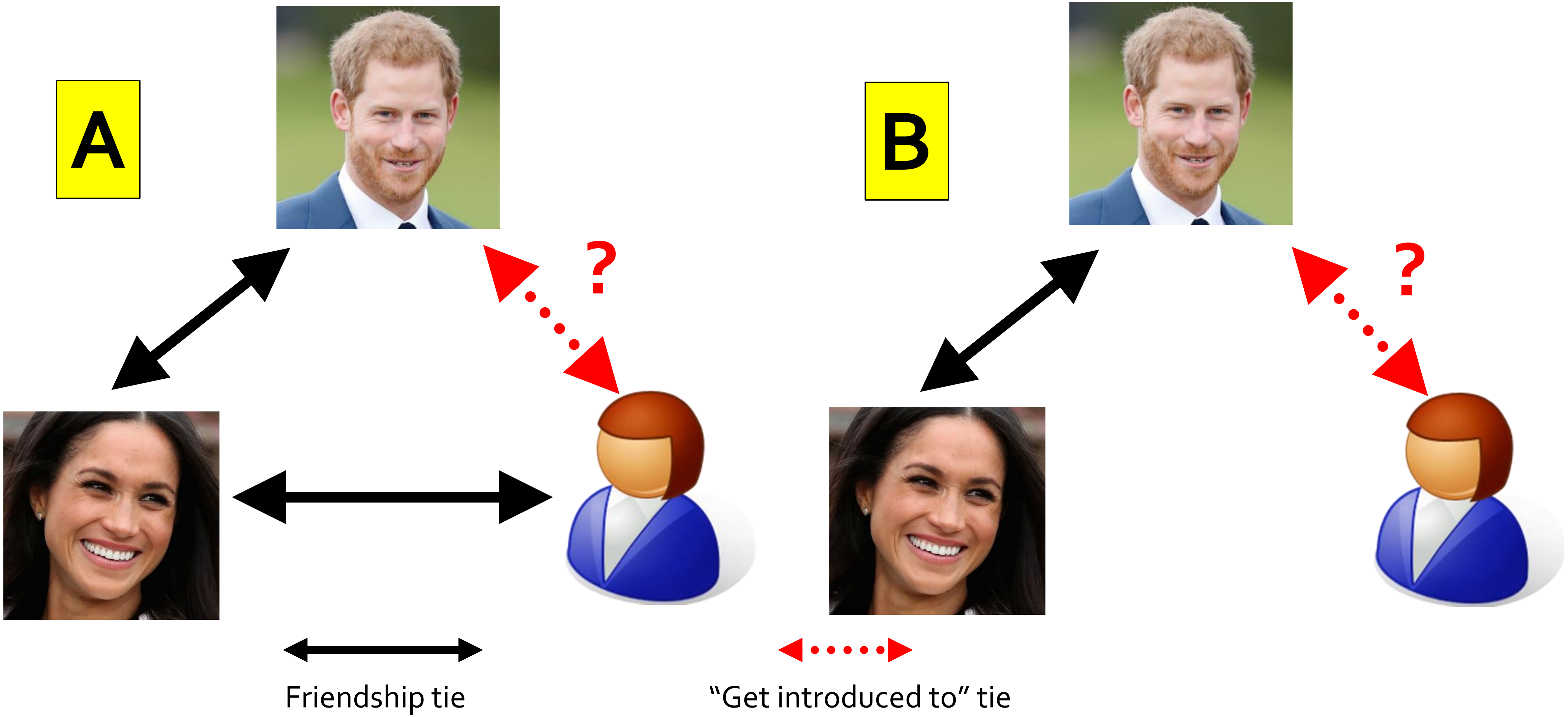
- One tie is *dependent* on the other


# Dependency

**The presence of one tie affects the presence/absence of another**

**Network self-organization**

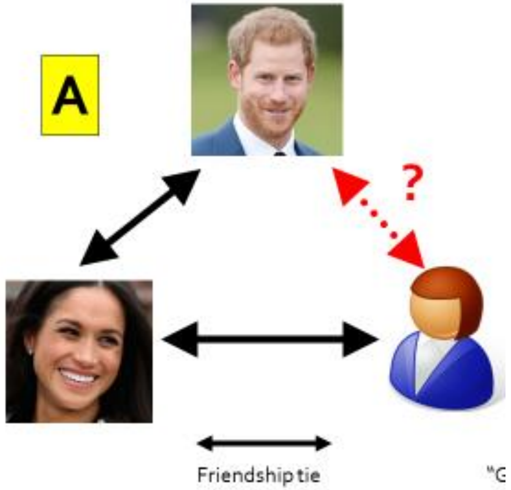
# In which scenario are we more likely to meet Prince Harry?



# ERGM vs Standard Stats

- ERGM uses conditional dependence

i.e., Nodes are considered independent UNLESS they share a tie

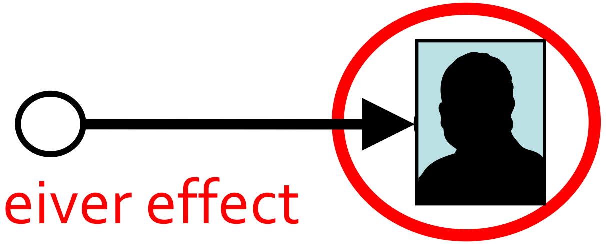



- Standard statistics assumes independent observations

**ERGM brings method closer to theory**

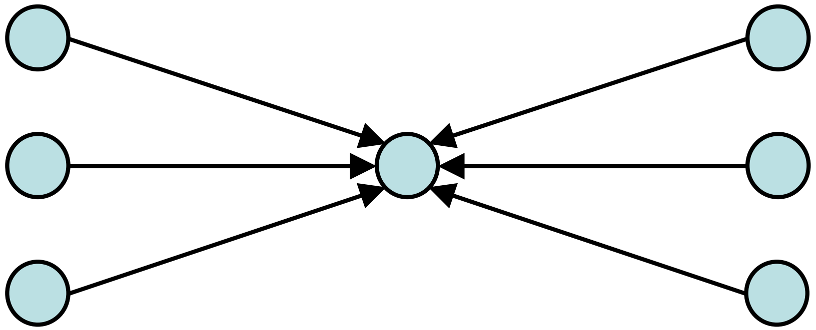
# So what?

- Centrality in networks due to attributes



Receiver effect

- Centrality in networks due to existing ties



Preferential attachment, Matthew effect, 'rich get richer'

ERGM can delineate these two competing explanations regarding why people form ties from one another

Standard statistics **CANNOT** delineate these two competing explanations

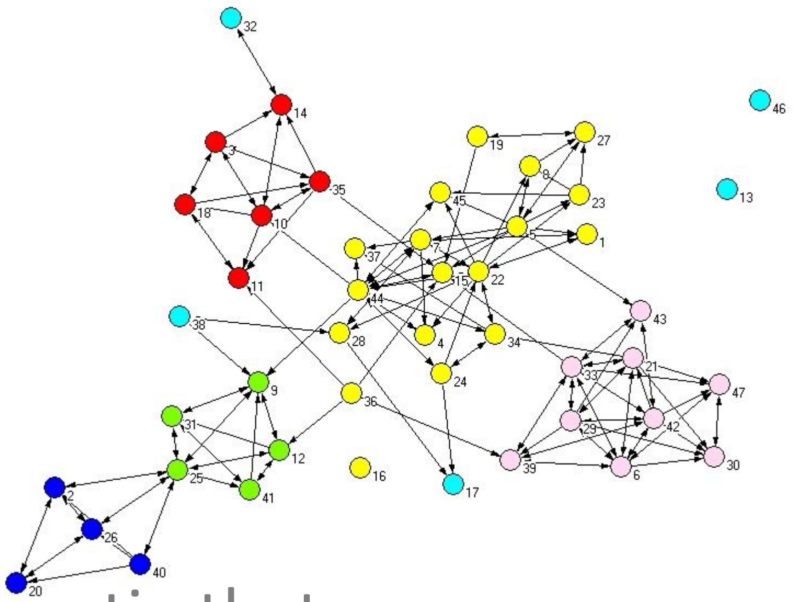
# ....and?

- If we do not control for preferential attachment, then we may overestimate (or underestimate) the impact of actor attributes

- Conclusions for the AFL study may have been wrong

- Lusher & Ackland, 2011

# Elements of a "social network"

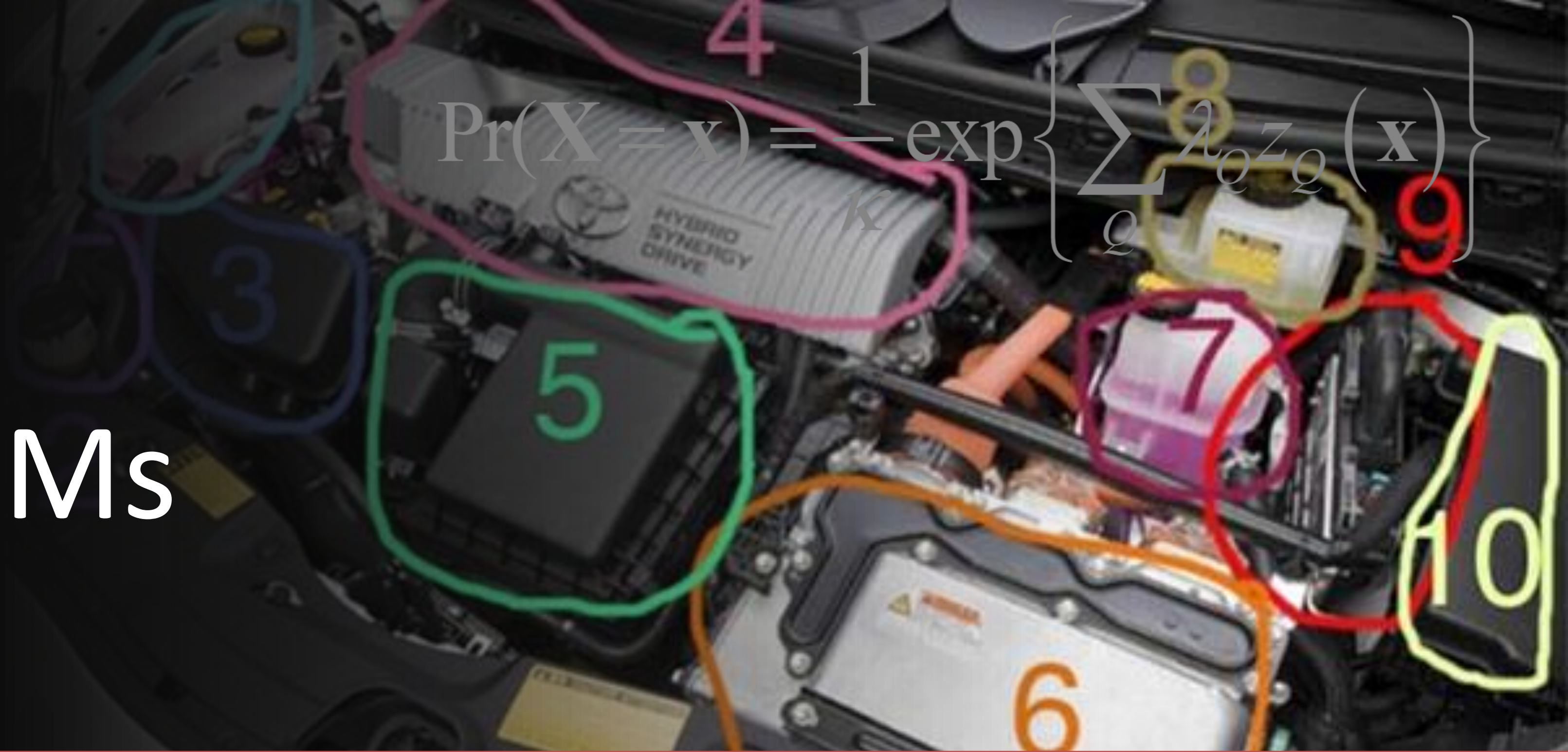


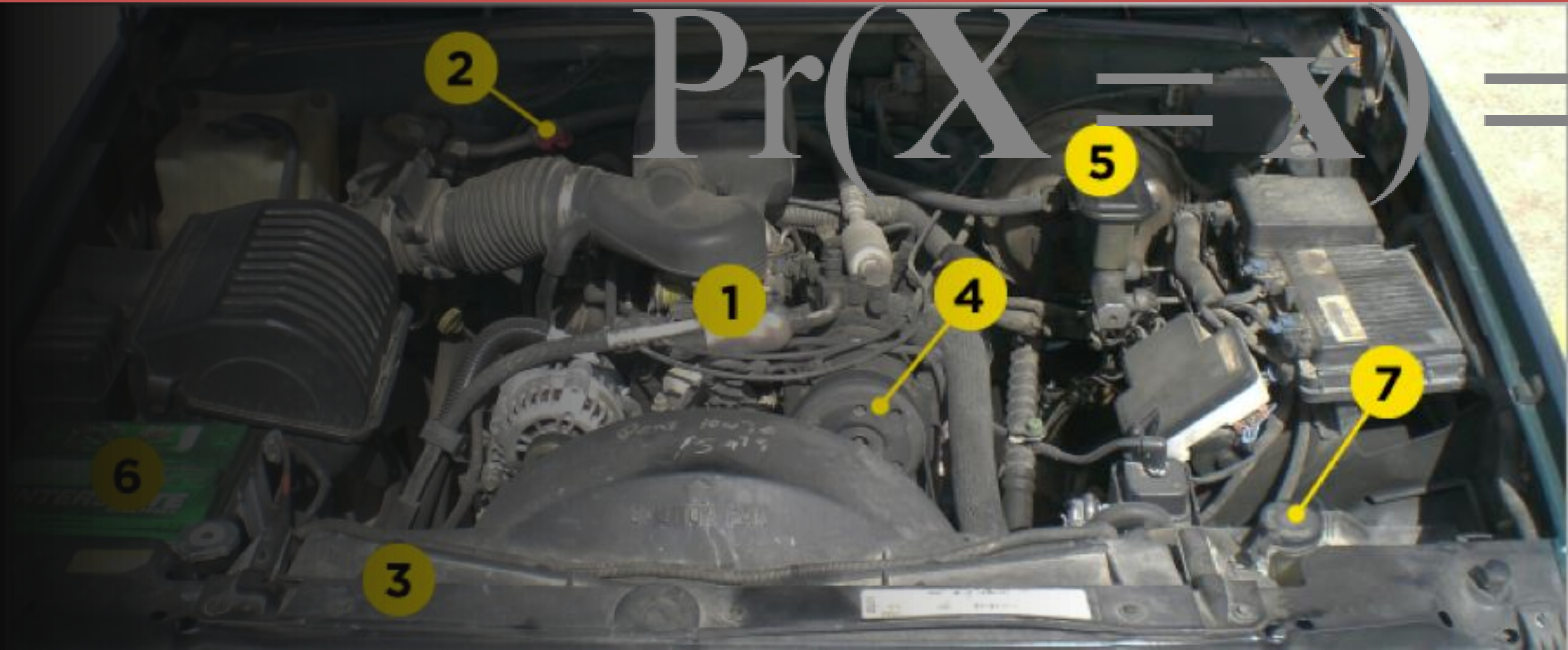
- Locally emergent
  - Local patterns form global structure
- Network ties self-organize
  - Through dependency of ties: it is because of one tie that another is formed (but also form due to actor attributes)
- Network patterns evidence of ongoing structural processes
  - Static trace of dynamic social processes
- Multiple processes can operate simultaneously
- Social networks are structured, yet stochastic
  - Structure and randomness

ERGM can take account of these important elements of social networks

# How do ERGMs work?

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



# Why do we **model** a network?

- Models represent theories we have about our observed data,
  - Is that theory valid? A model can test this
- A desirable goal of a model is to best represent our observed data
  - Reproduce the structures we see in our observed network
- In ERGM, our model represents the combination of structures of which our observed network is composed
  - Permits inferences about the social processes of network tie formation.
- When run an ERGM, we are in fact estimating a model for (or modelling) that data
  - Just like a regression model to non-network data.

**Multiple**

# How does ERGM work?

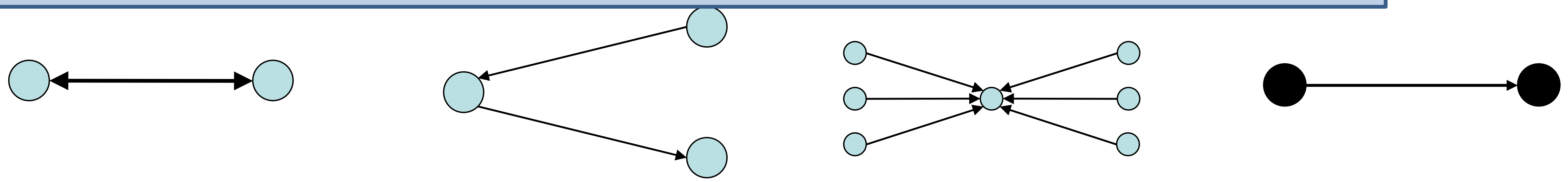
- Like logistic regression
  - …… but with more complex dependence assumpti

**Network effects**
Reciprocity
Transitivity
Degrees
Homophily
Brokerage
Clustering
Popularity

**Predicting ties**

| | *Coefficient* | [illegible] | *t Stat* | *P-value* | |
|---|---|---|---|---|---|
| Intercept | [illegible] | 0.0731 | 4.6624 | 0.0000 | |
| projects | -0.0064 | 0.0082 | -0.7849 | 0.4378 | |
| seniority | **0.2491** | 0.1104 | 2.2576 | 0.0303 | * |

**Multiple predictors**



# Some important ERGM concepts

- Random graphs
  - Randomness and structure
  - Do random graphs have some structure to them?
  - When is there <u>enough</u> structure?

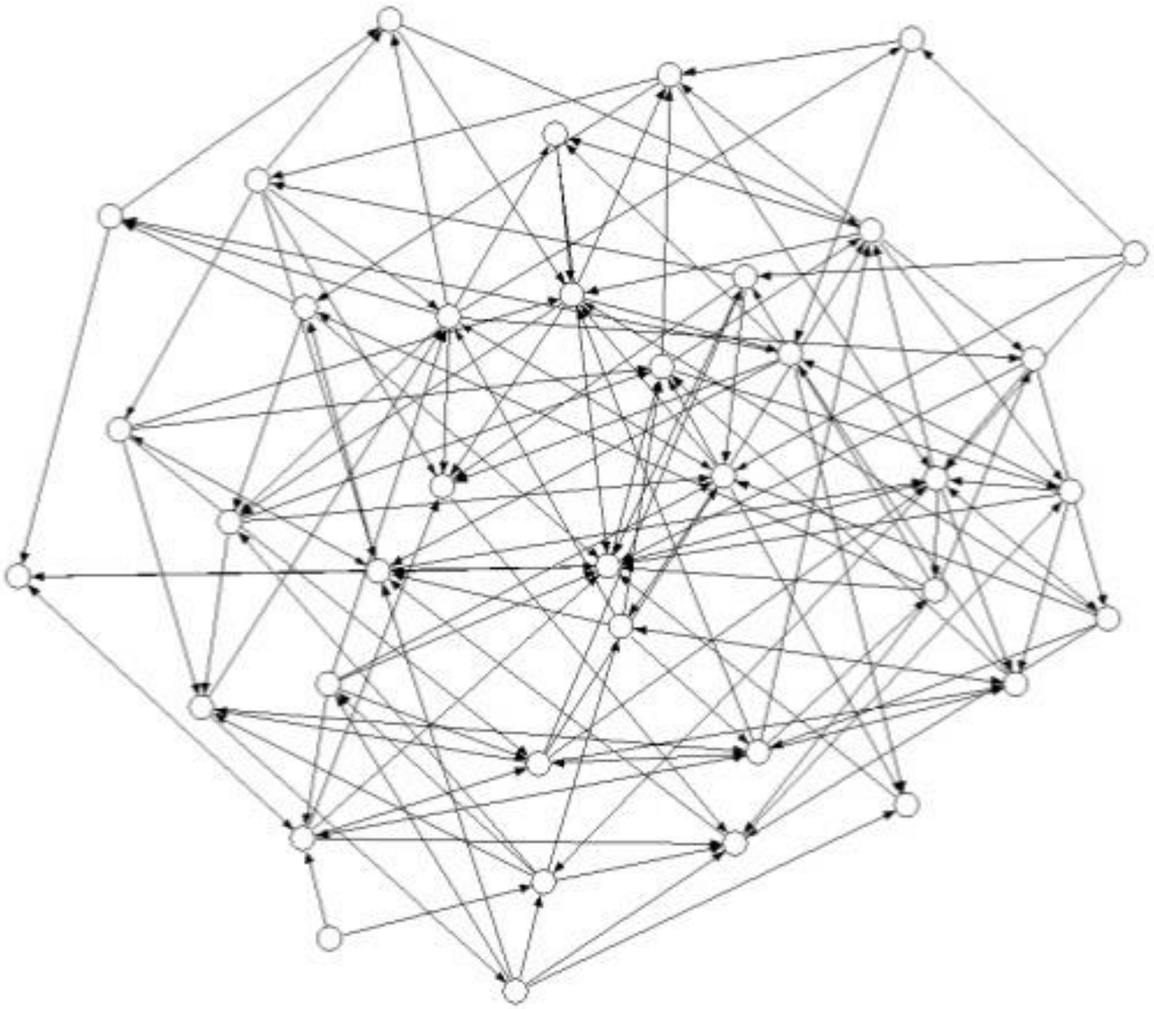
**Dependency**

- Distribution of graphs
  - How many different ways could a network with the same number of nodes and same number of ties be arranged?

# Random graphs

# Random Graphs

- A random network is a theory about how network ties form
  - i.e. ties form <u>independently</u>
  - i.e. ties form <u>randomly</u>

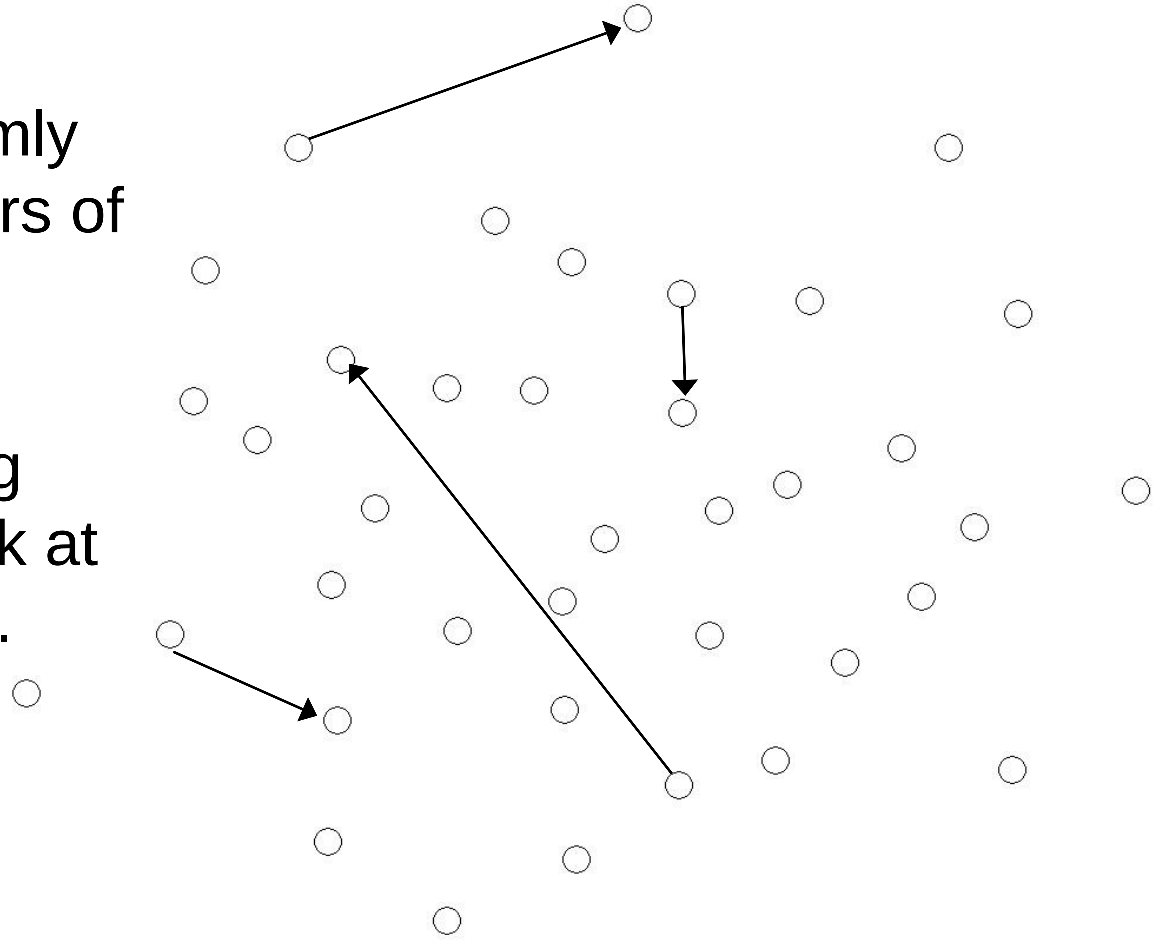


**Even random graphs may have some structure**

# Creating a random graph

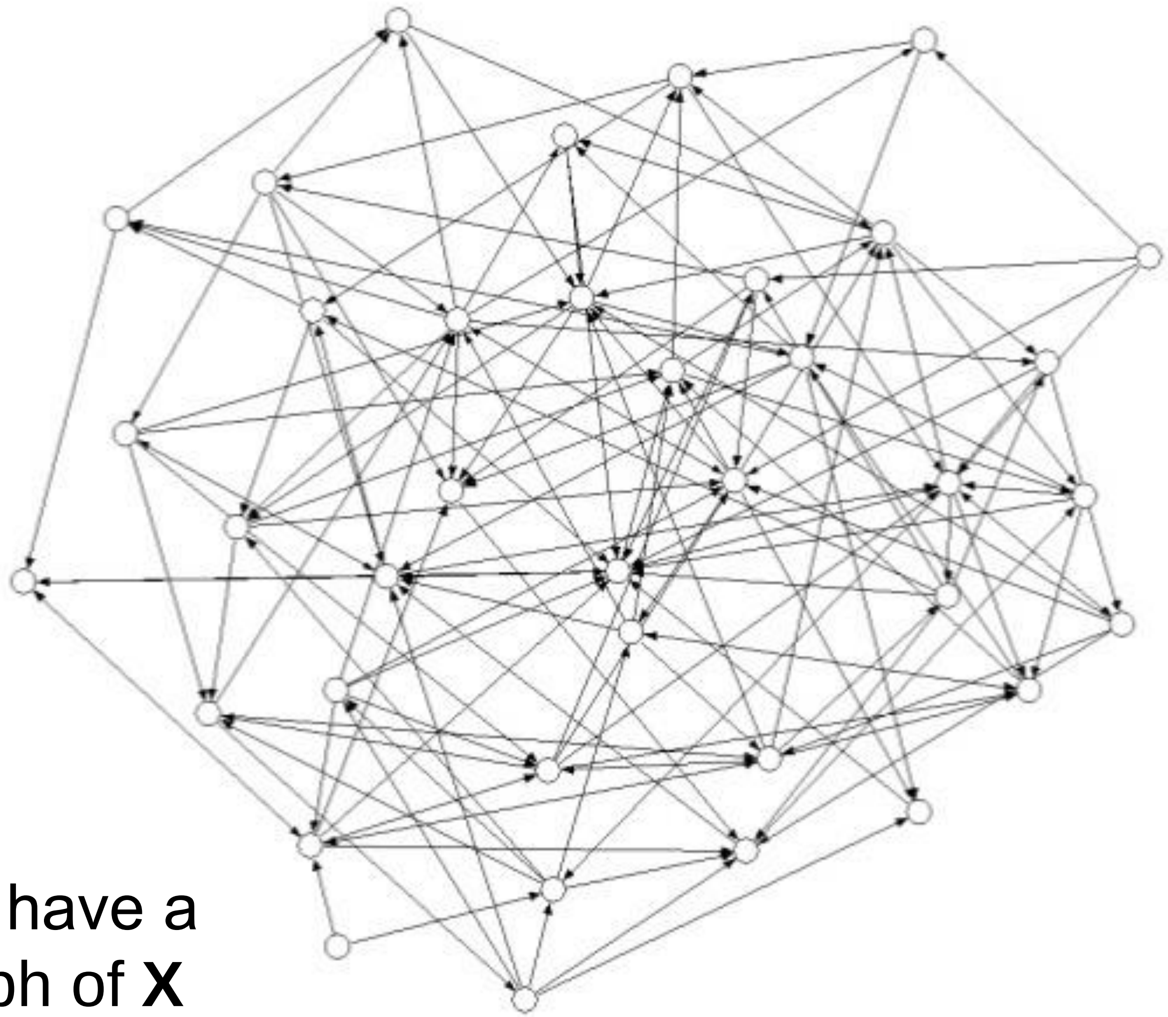
Ties are randomly assigned to pairs of actors

We keep adding ties to a network at random until…..



Simulation = creation

# Creating a random graph



….. until we have a random graph of **X** nodes and **Y** ties

We specify X and Y ourselves

Simulation = creation

# Random graphs

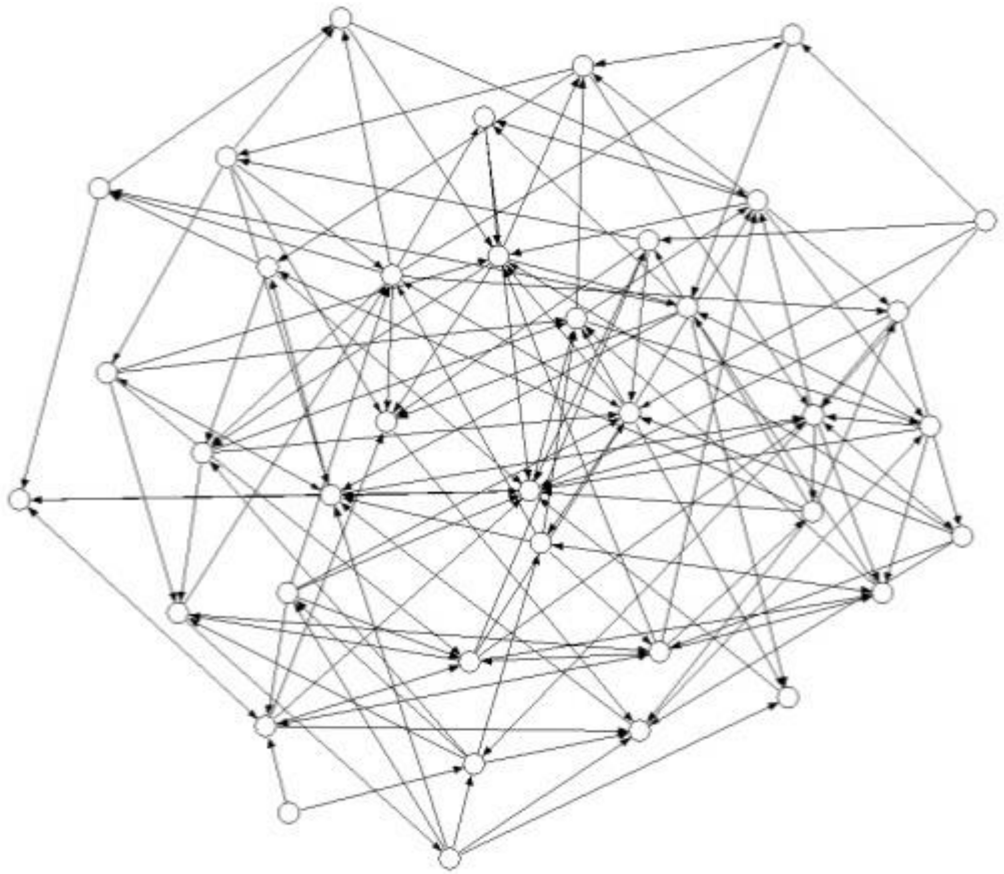
- So is the point of ERGM to see if our data differs from a random graph?
  – Sort of, but not precisely

- Random graphs are a starting point
  – We can start with the assumption of random connections
  – But we introduce different dependencies to weight some network distributions
  – Over the next few days, we will learn such dependence assumptions

# Why use random graphs?

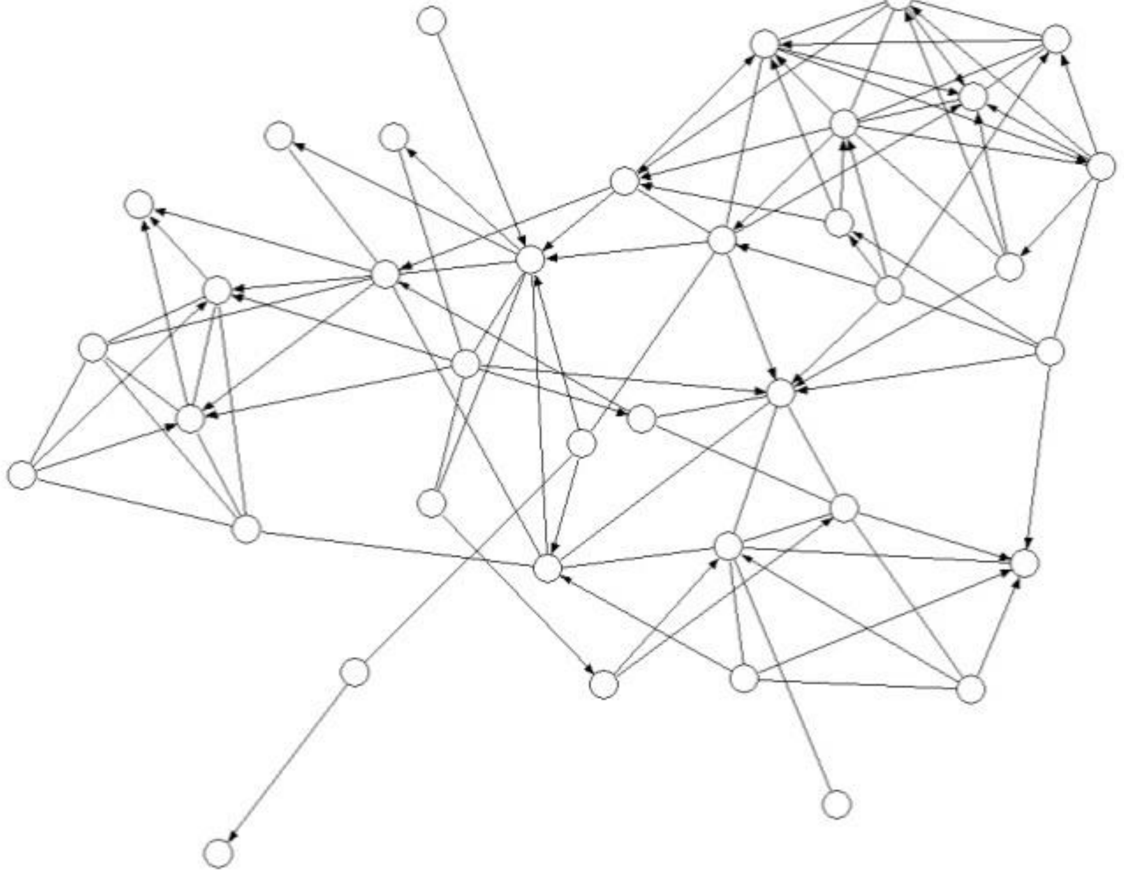
(Both graphs below have same number on nodes, same number of ties)

**Random graph**



Simulated random graph

**Observed graph**



Network we have collected data on

# Random vs Observed

But even random graphs have some structure

There are still some triads, just not as many as the observed network

Random Network

Communication Network

| | Random Network | Communication Network |
|---|---|---|
| Actors | 38 | 38 |
| Arcs | 146 | 146 |
| Reciprocated arcs | 6 | 44 |
| Transitive triads | 53 | 212 |
| in-2stars | 292 | 313 |
| out-2stars | 254 | 283 |

# Random vs Observed

But even random graphs have some structure

So if our network follows non-random processes, which ones are important?

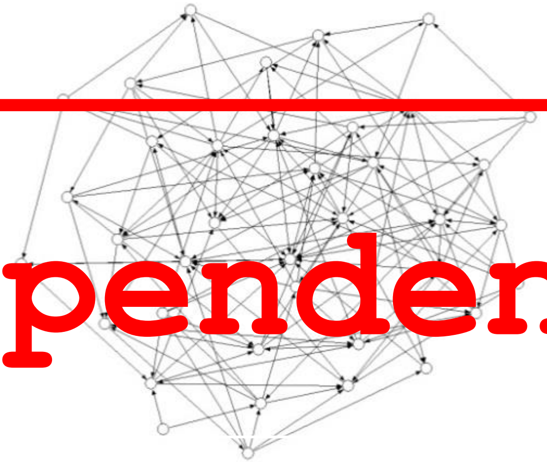
Random Network

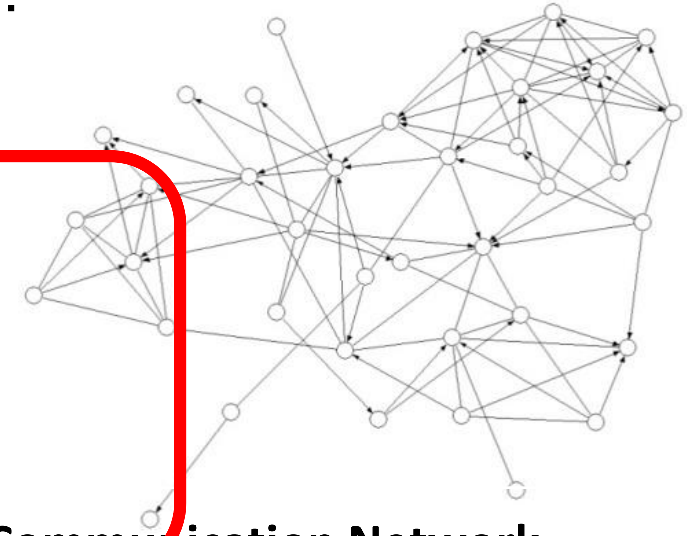
Communication Network

| | Random Network | Communication Network |
|---|---|---|
| Actors | 38 | 38 |
| Arcs | 146 | 146 |
| Reciprocated arcs | 6 | 44 |
| Transitive triads | 53 | 212 |
| in-2stars | 292 | 313 |
| out-2stars | 254 | 283 |


# Why are there differences?

Between random and observed data?



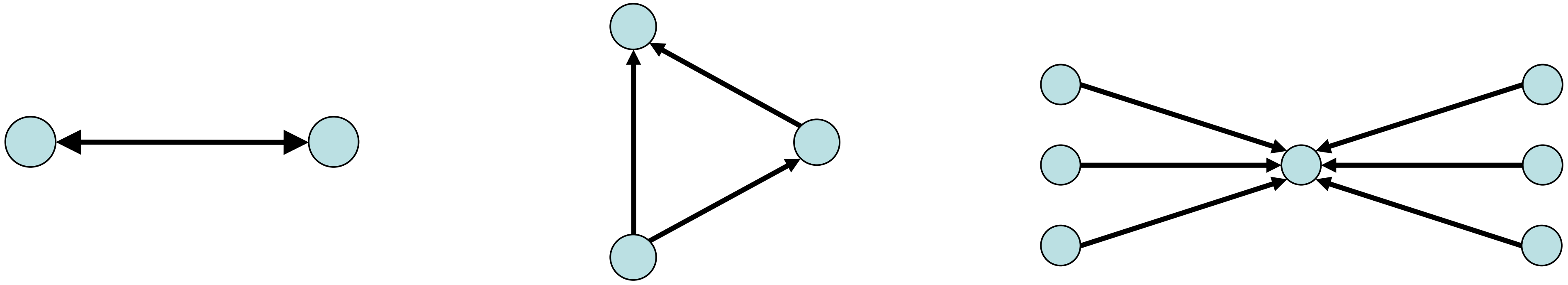


**Dependency**

| | Random Network | Communication Network |
|---|---|---|
| Actors | 38 | 38 |
| Arcs | 146 | 146 |
| Reciprocated arcs | 6 | 44 |
| Transitive triads | 53 | 212 |
| in-2stars | 292 | 313 |
| out-2stars | 254 | 283 |

# Dependency makes some configurations more likely

- The dependency inherent in social processes make some network patterns more likely, and others less likely

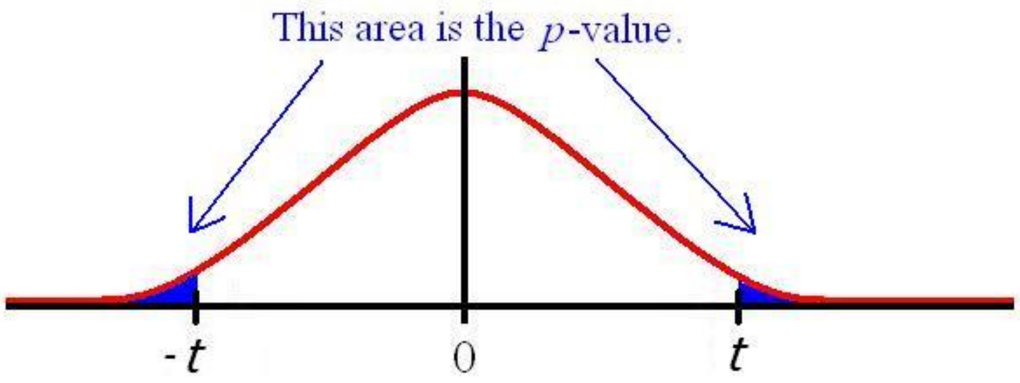


*Dependency is the ‘social glue’ of network structures*

# Distribution of graphs

# Comparing Data to Distributions

- We often compare our <u>observed data</u> to a <u>distribution</u> to see if our data are ‘extreme’
- A *t*-test is precisely this, and let’s us know if something is significant



# Distribution of graphs

- We can compare our observed network data to a distribution of graphs

- That is, how does the data we collected (observed) compare with all of the other ways that this network could be configured?
  – Assuming same number of nodes
  – Assuming same number of ties

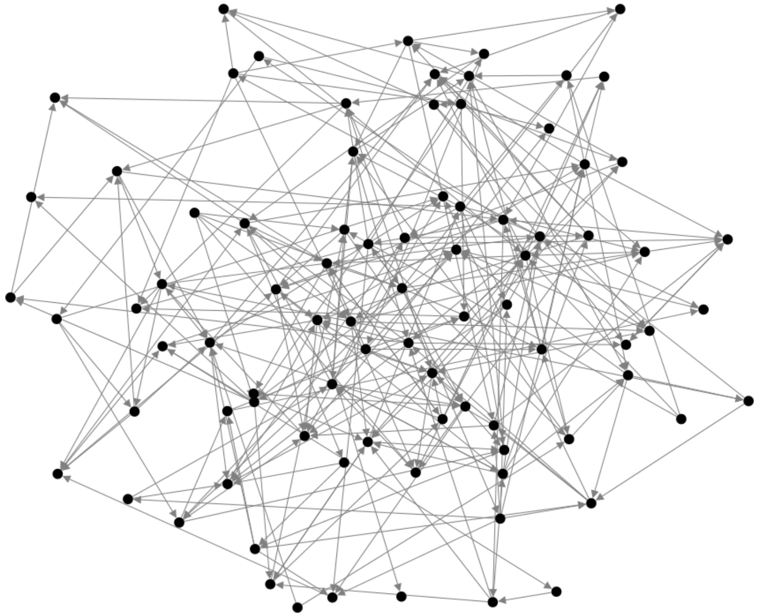
# Distribution of graphs

Ah, we've done this already…

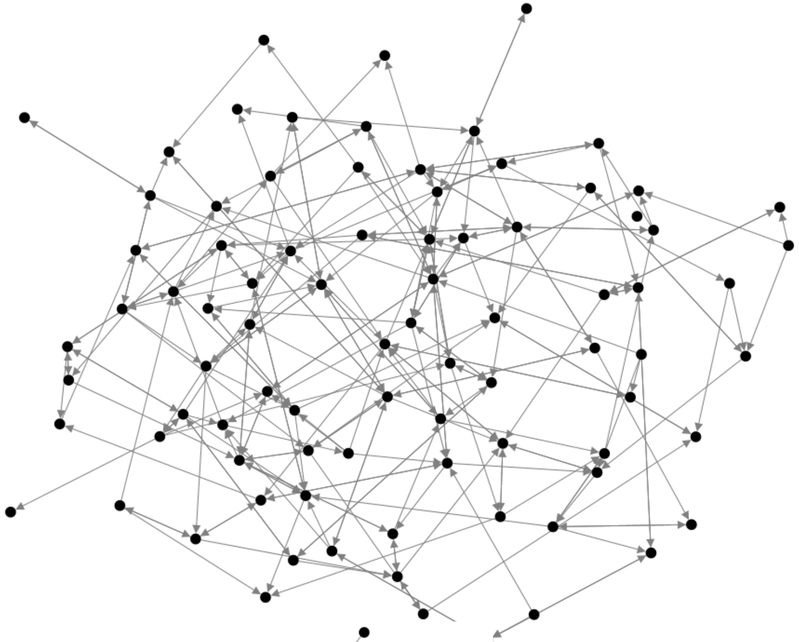
- net1, net2, net3 & net4
  - All 85 nodes
  - All same density, but different structure
  - Each represents a different possible way that a graph can be arranged

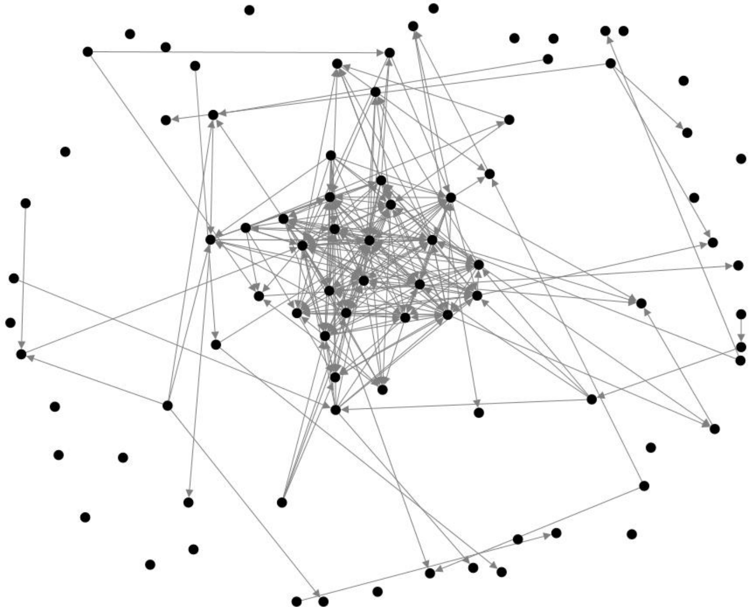
- Distribution of graphs
  - All possible arrangements of a graph (with 85 nodes and 256 ties)
  - Sample space is the finite number of ways a graph can be arranged
  - Some graphs are more or less likely (more or less probable)

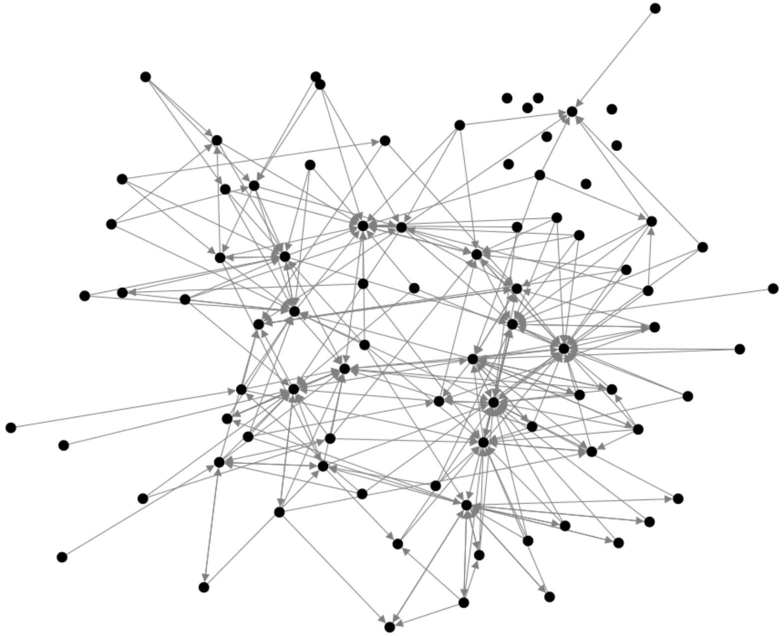
# Networks of **85** nodes and **256** ties



Do these 4 networks have the same structure?

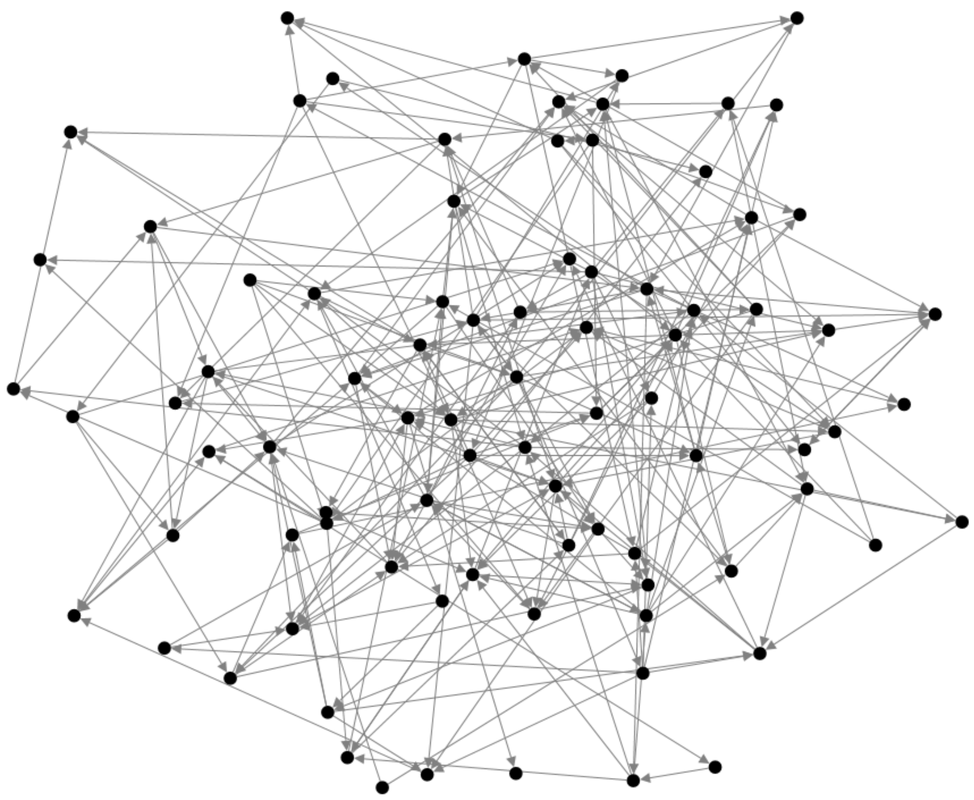


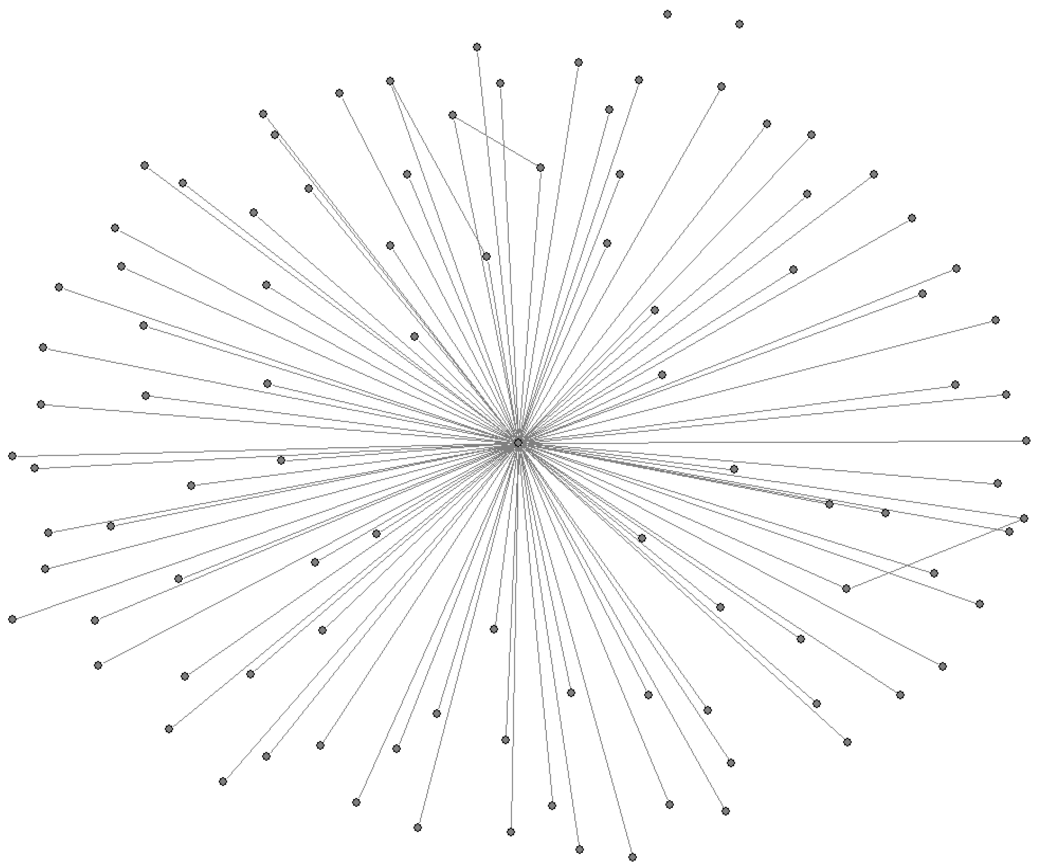




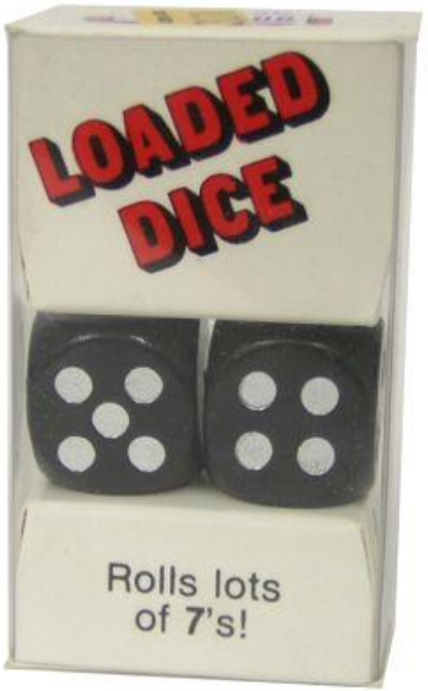
# Distribution of Graphs

Some graphs are more or less likely (more or less probable)





# Dependencies

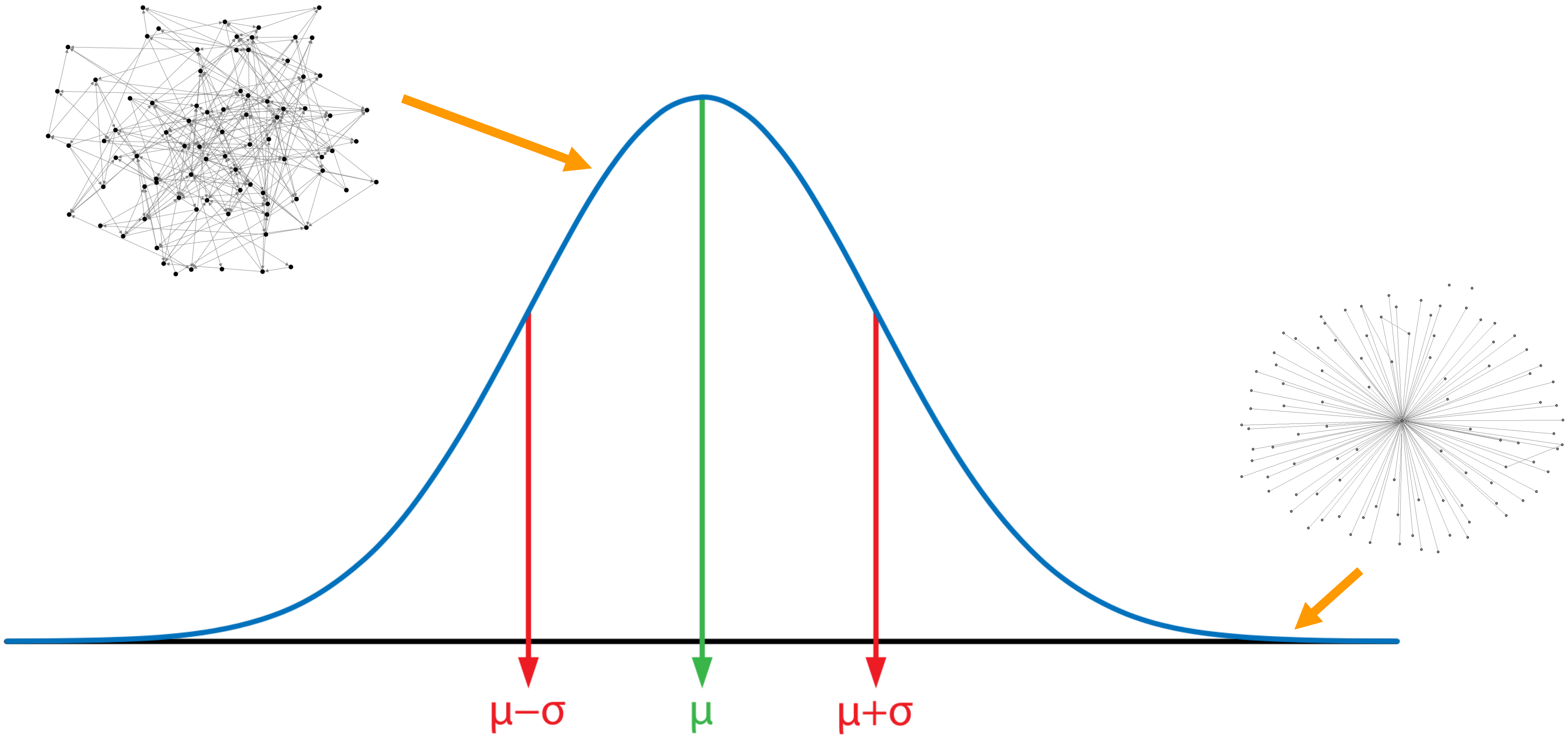



- With weighted/loaded dice (dependencies), then we would expect to see **Sevens (7s)** more than chance.
- If we compare to normal dice (random), we will have more sevens than we expect.

- Similarly, given we know there are **dependencies** in network data (e.g., you scratch my back, I'll scratch yours)  how do we know there are more reciprocated ties that expected?

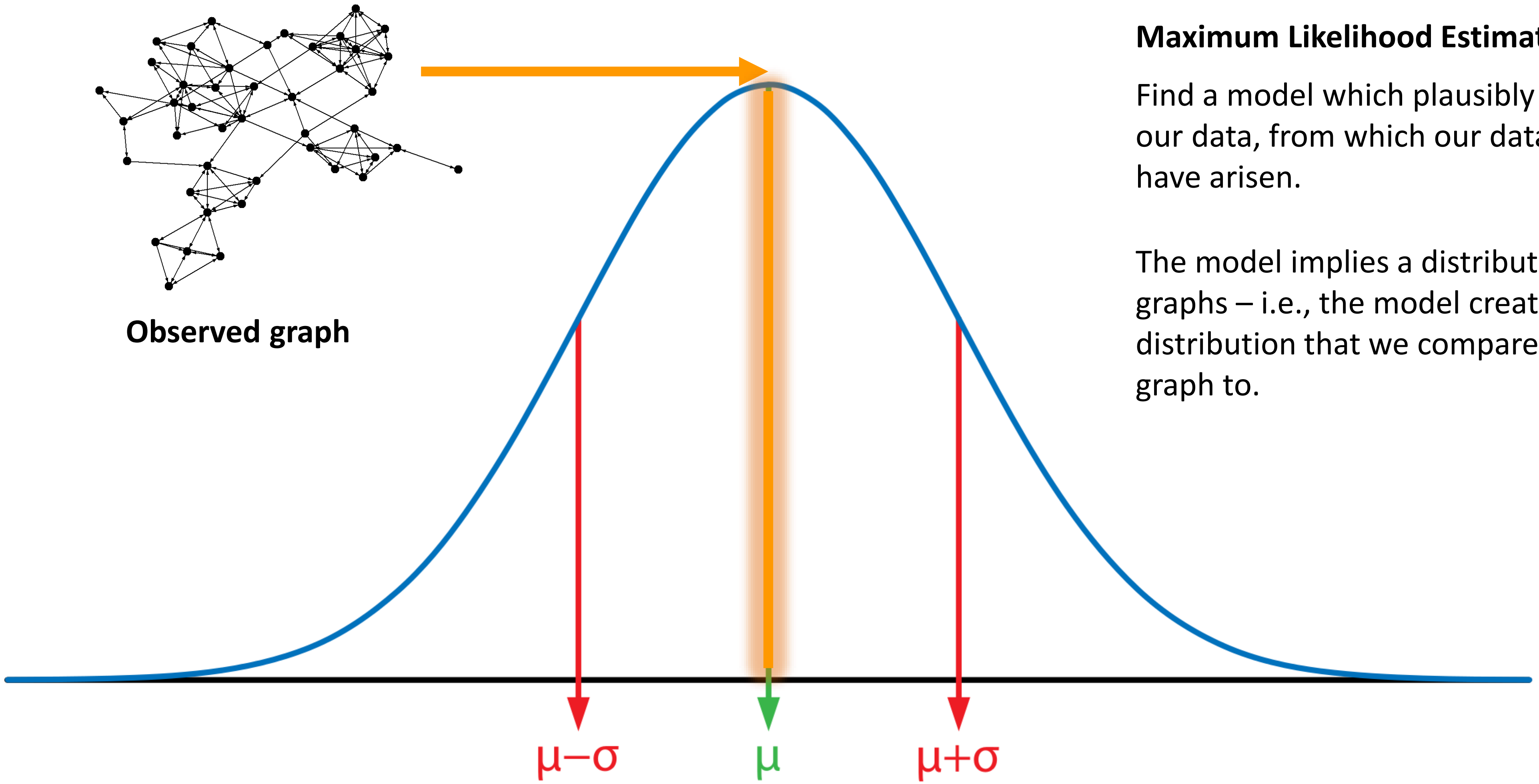
Dependency

# Distribution of graphs




We can constrain our distribution of graphs by introducing dependencies, making some graphs more likely and some less likely

# Distribution of graphs




**Observed graph**

**Maximum Likelihood Estimate (MLE)**

Find a model which plausibly explains our data, from which our data could have arisen.

The model implies a distribution of graphs – i.e., the model creates the distribution that we compare our graph to.

We can constrain our distribution of graphs by introducing dependencies, making some graphs more likely and some less likely

# Distribution of Graphs

1. Could our network have arisen from the combination of multiple, local processes?
   - Is our network central in the distribution of graphs? We want it to be.

**Maximum Likelihood Estimate (MLE)**

2. Are any features of the network extreme?
   - i.e., non-random
   - Do we see more *reciprocity* in our network than we might expect by chance?

**Which network parameters are significant?**