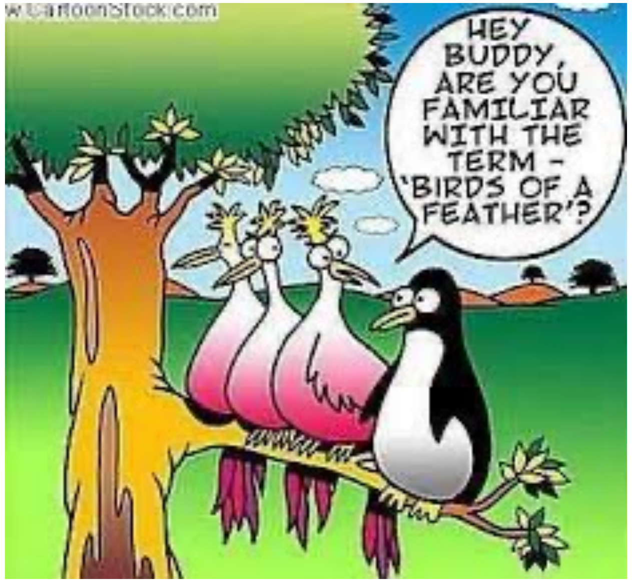
CHDH-SNA-2a

Social influence in cross-sectional data

Network regression

Homophily

"Birds of a feather flock together"

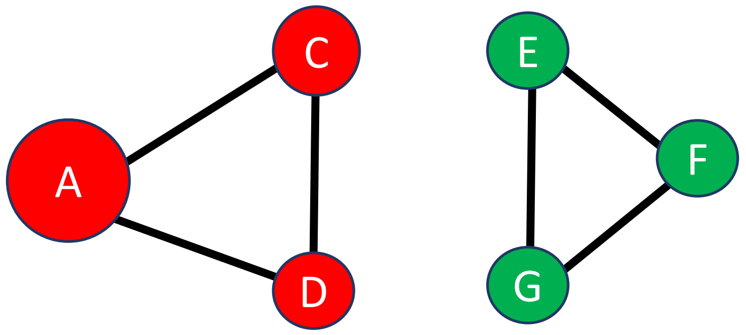



BEARDED
VILLAINS


NSYNC

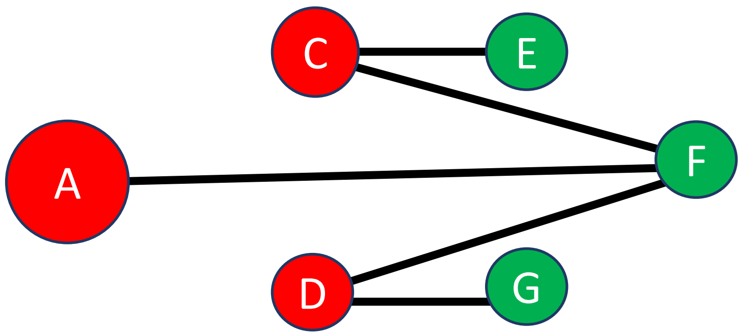
Homophilous/heterophilous dyads

Homophily



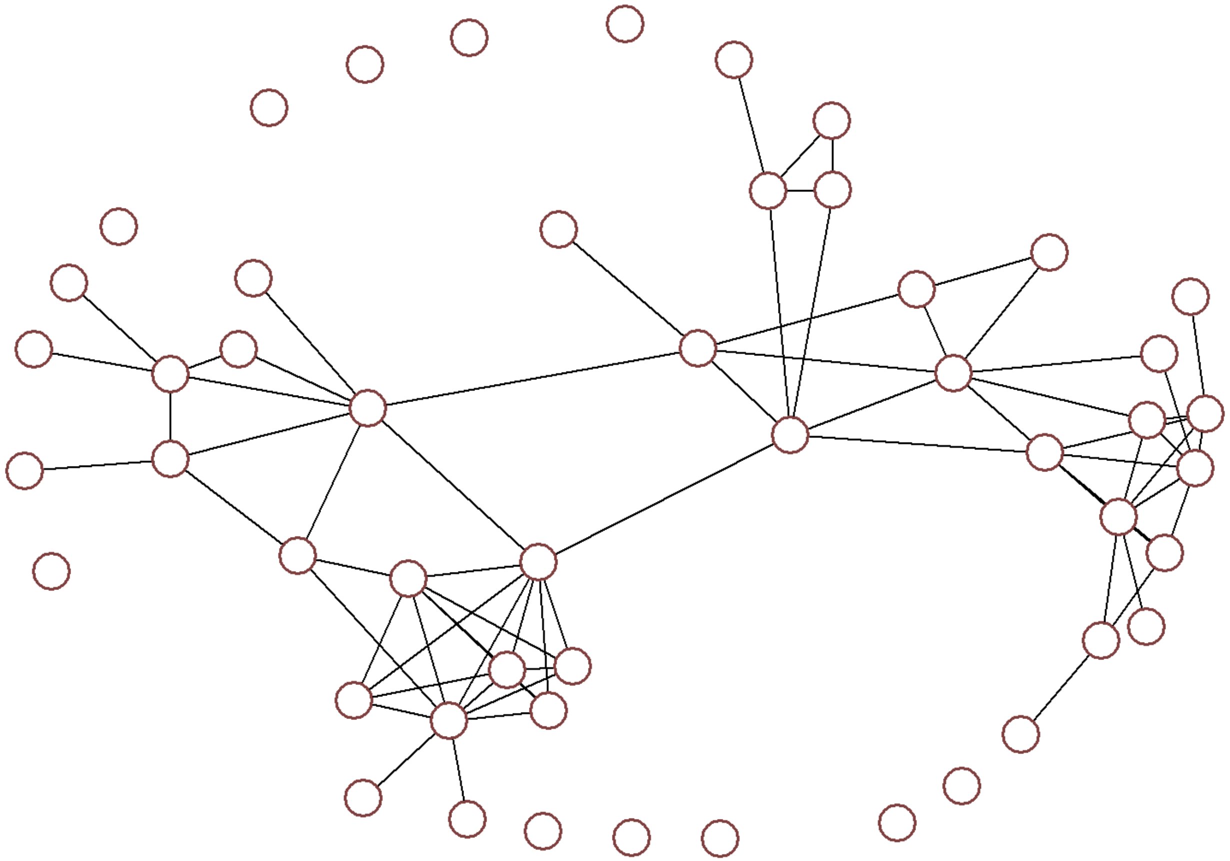
A, C, and D like sports *E, G, and F like the ballet*

Heterophily

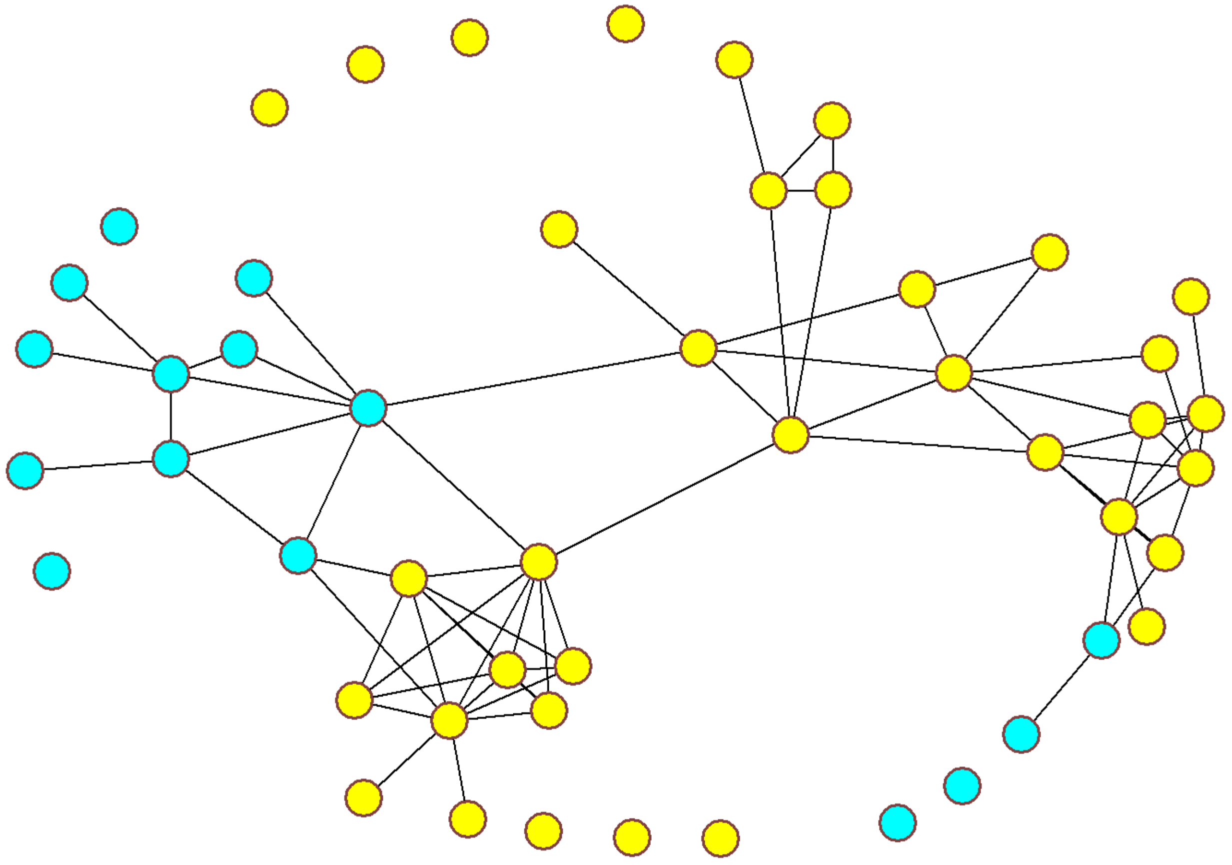


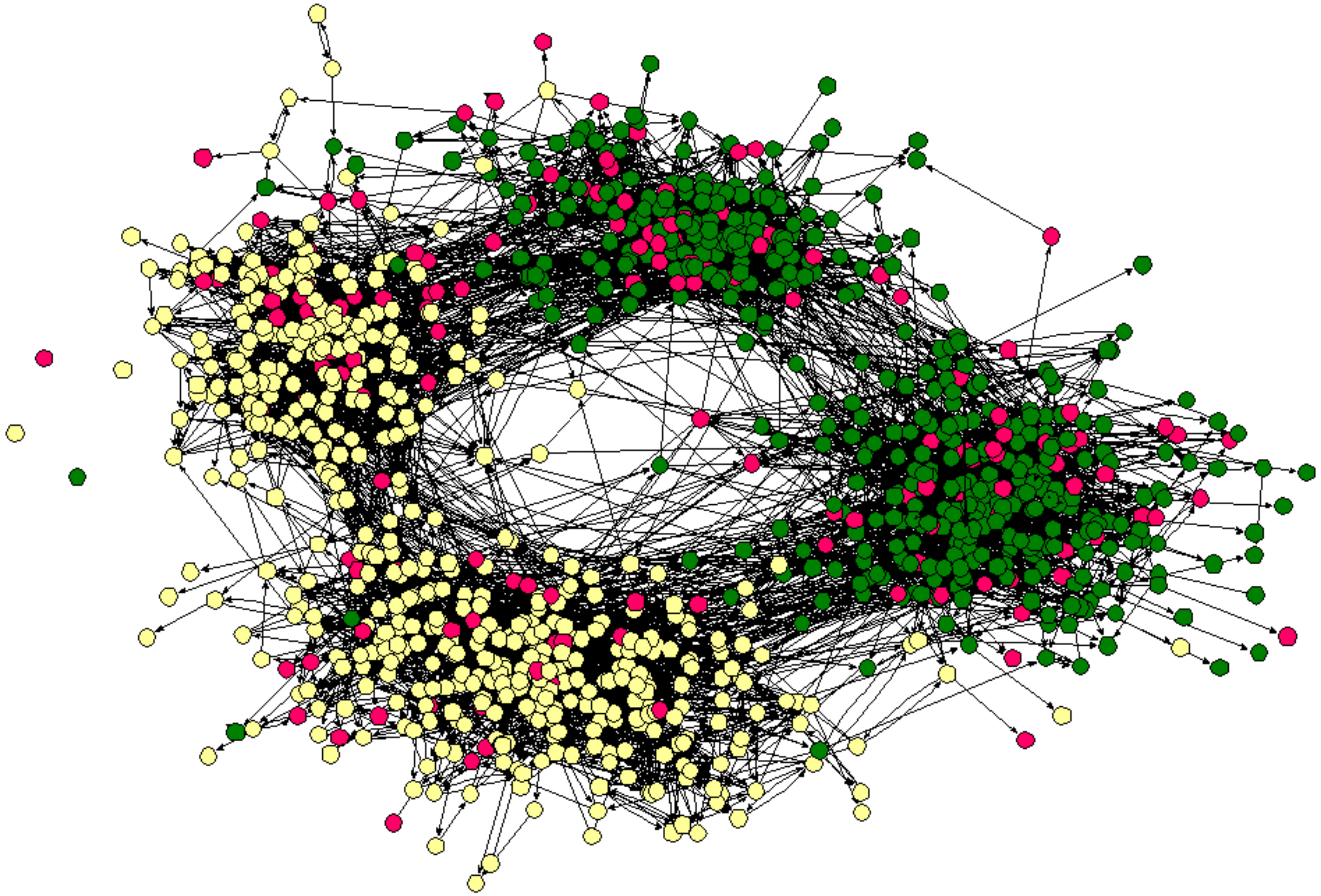
A, C, and D produce *E, G, and F consume*

Homophilous/heterophilous dyad

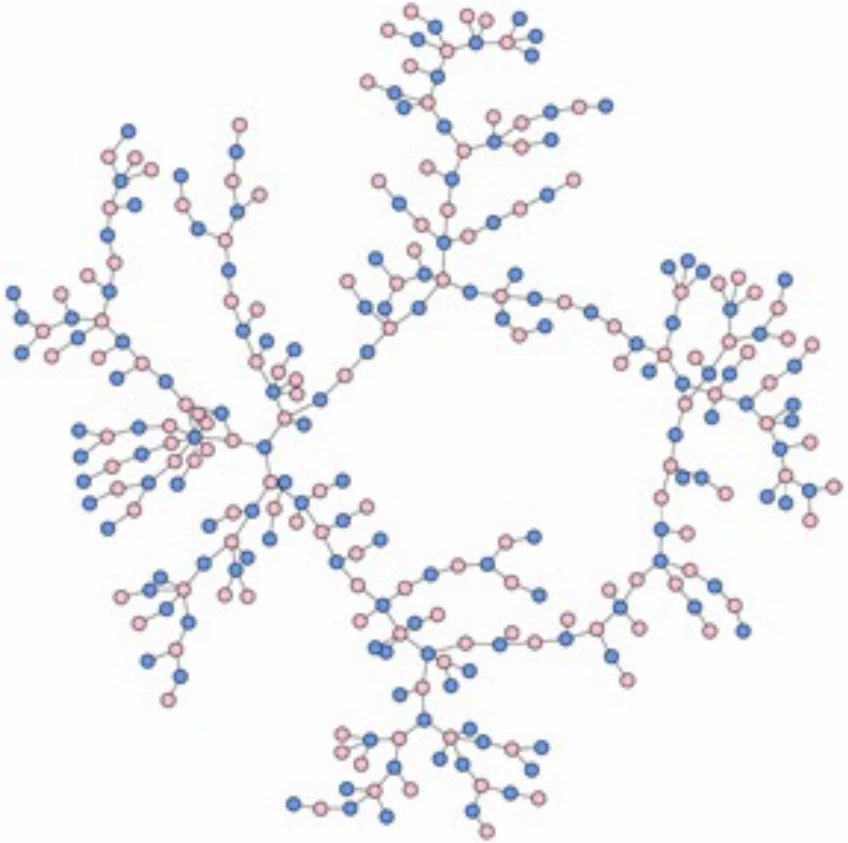


Homophilous/heterophilous dyad



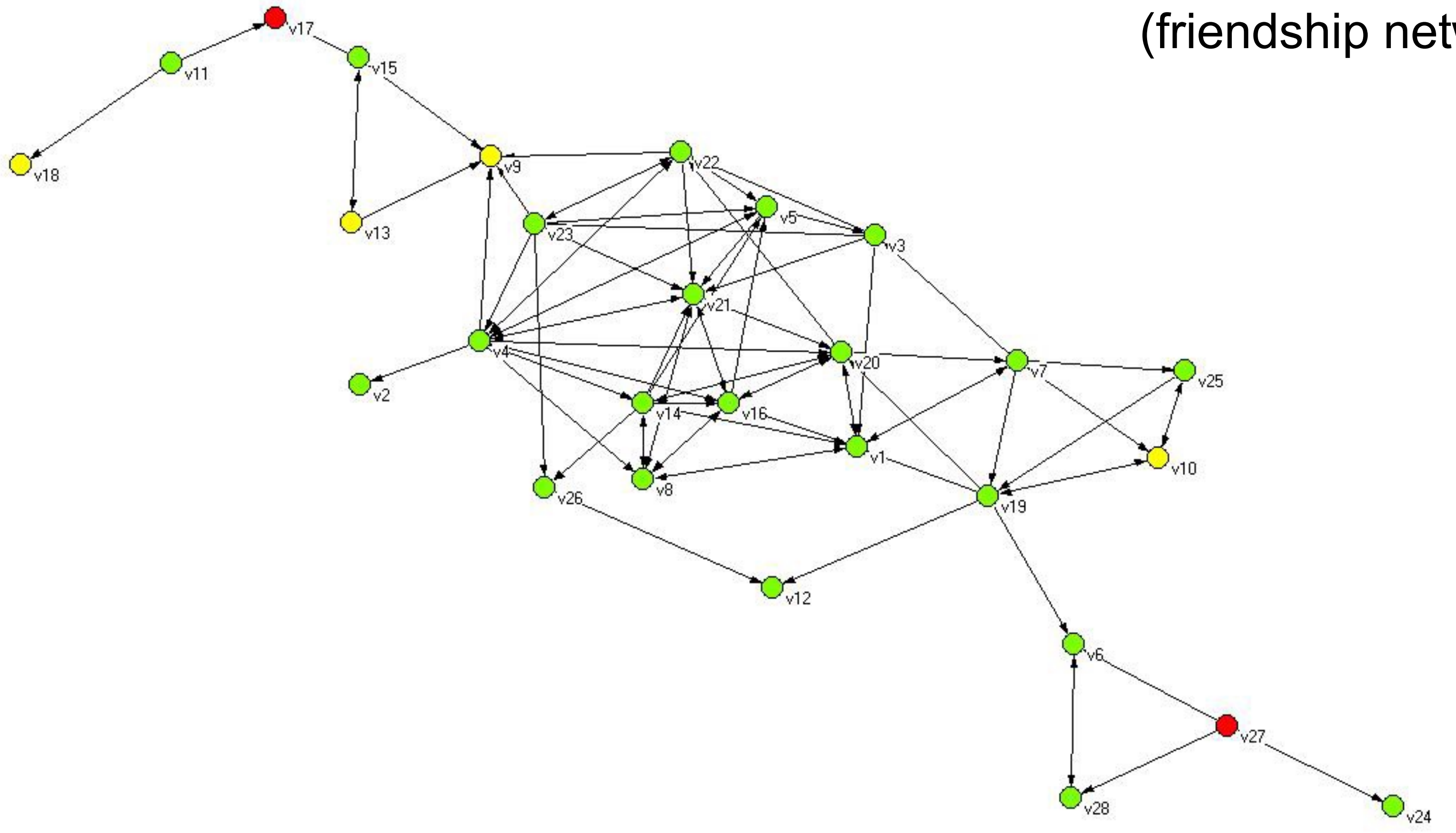


white black other

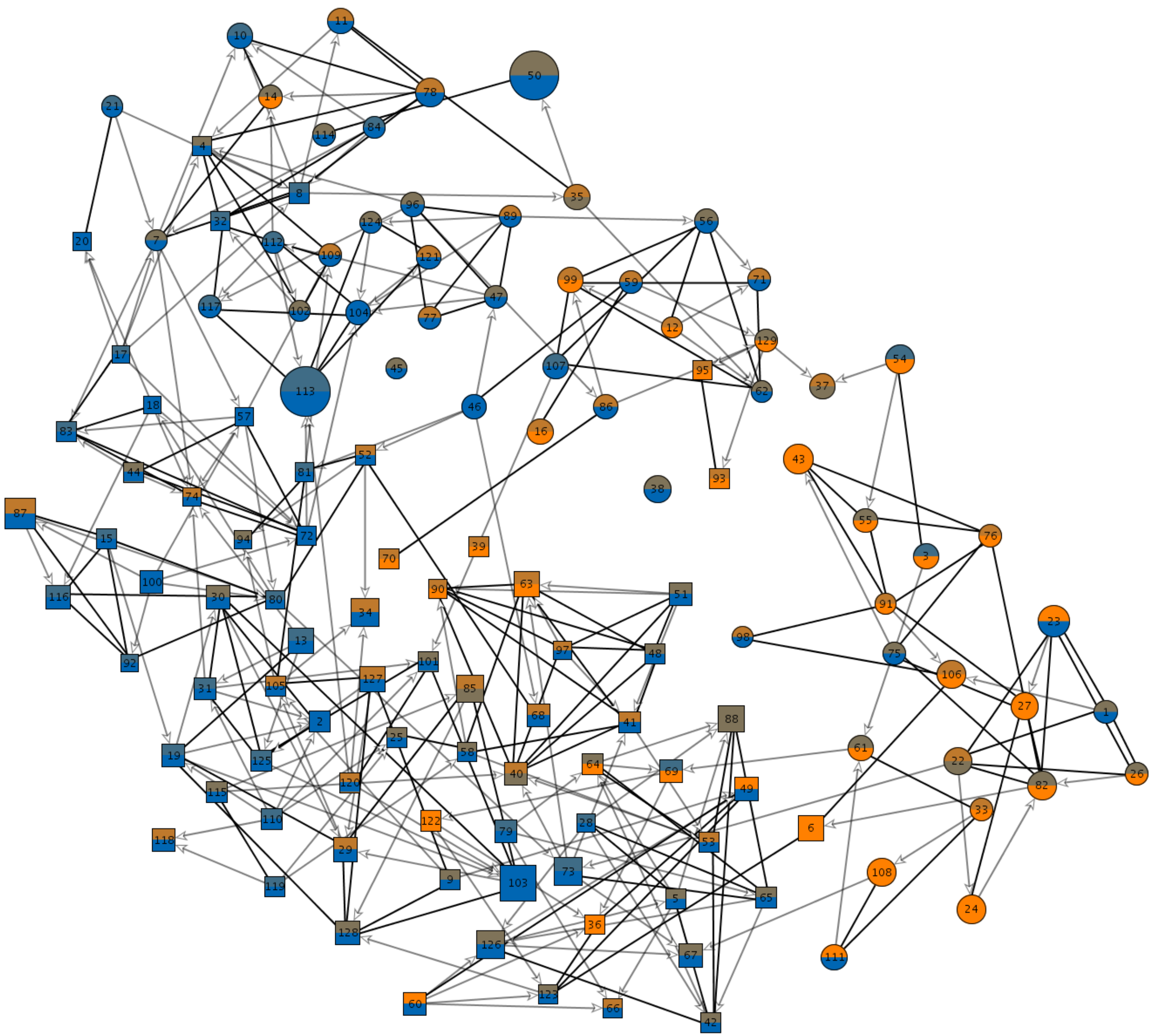


Team structures in training squads (Pane, 2003)

(friendship network in 12th week of training)



detached team oriented positive



example: friendship and substance abuse adolescents in Glasgow

EDGES

friend

reciprocal

NODES

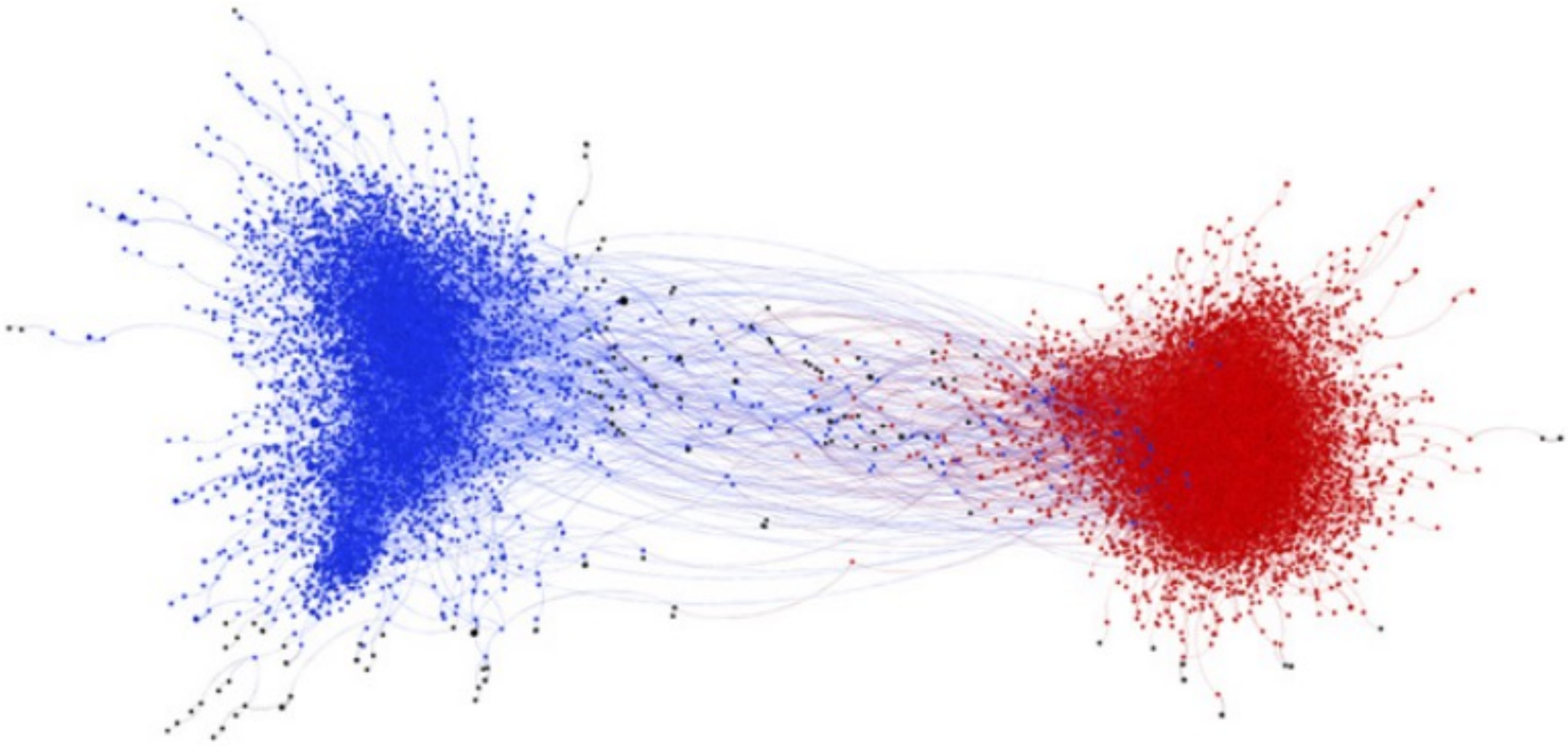
girl

Boy

Little Lot

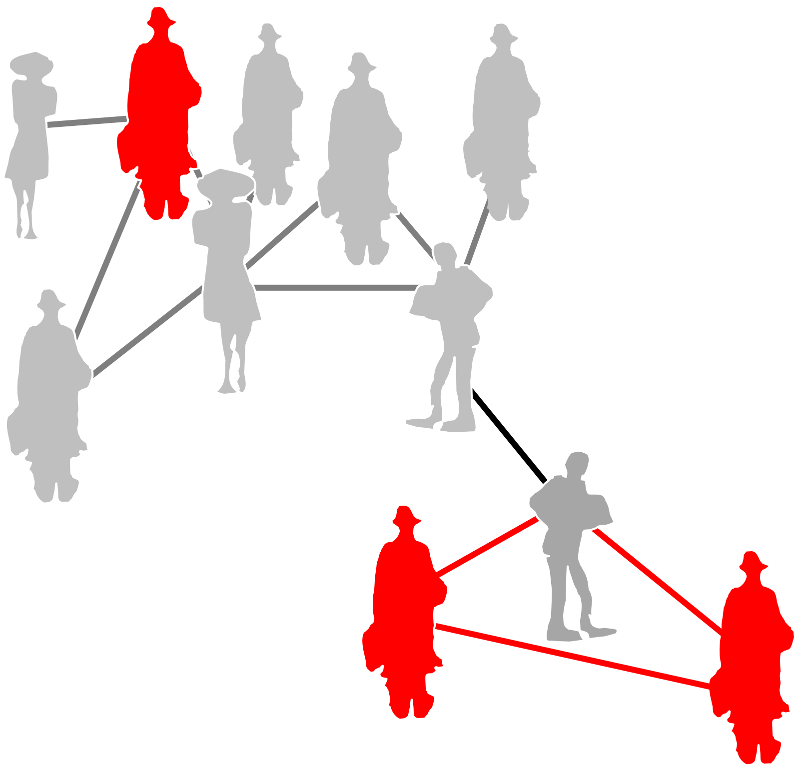
smoke

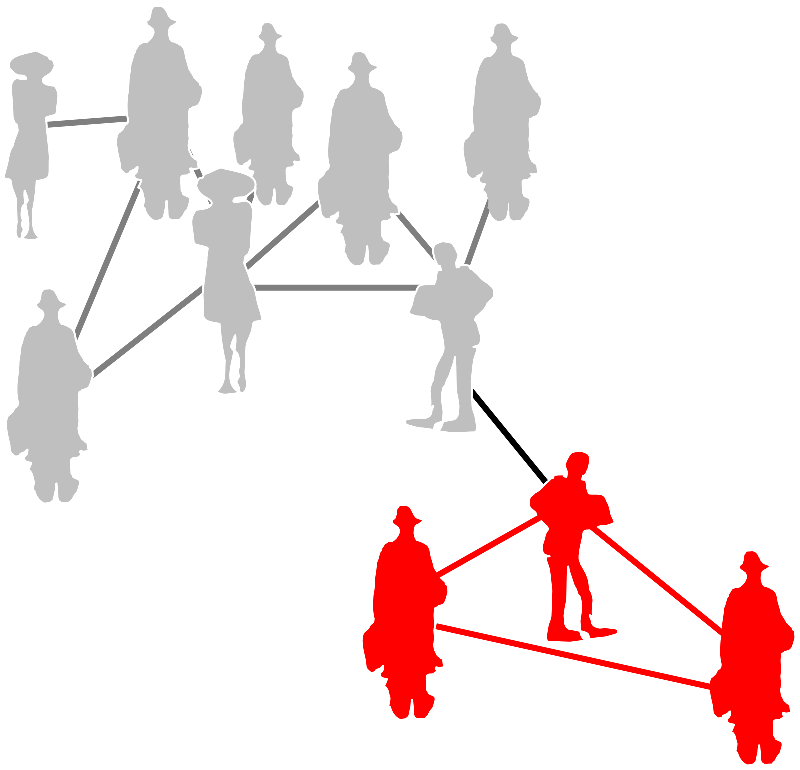
drink



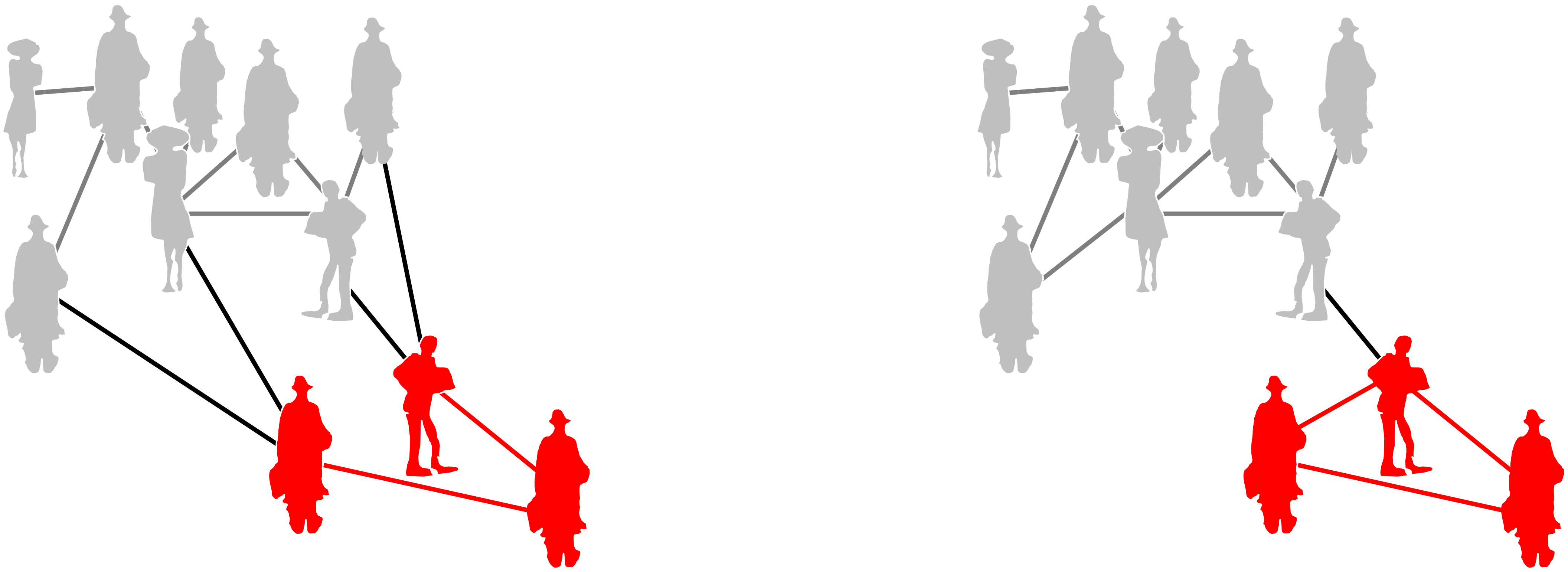
Brady et al.(2017), PNAS

Social influence





Social selection



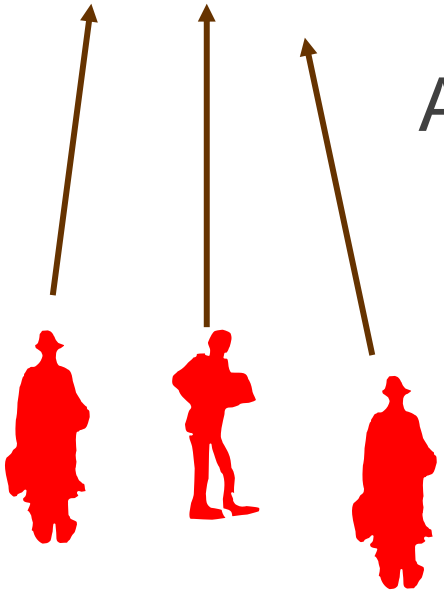
Propinquity, context, and foci

“Feld’s (1981, 1982, 1984) argument that focused activity puts people into contact with one another to foster the formation of personal relationships“
MCPHERSON et al (2001:431)

Affiliated with same things

Same attributes, same activities, form ties

Propinquity, context, and foci

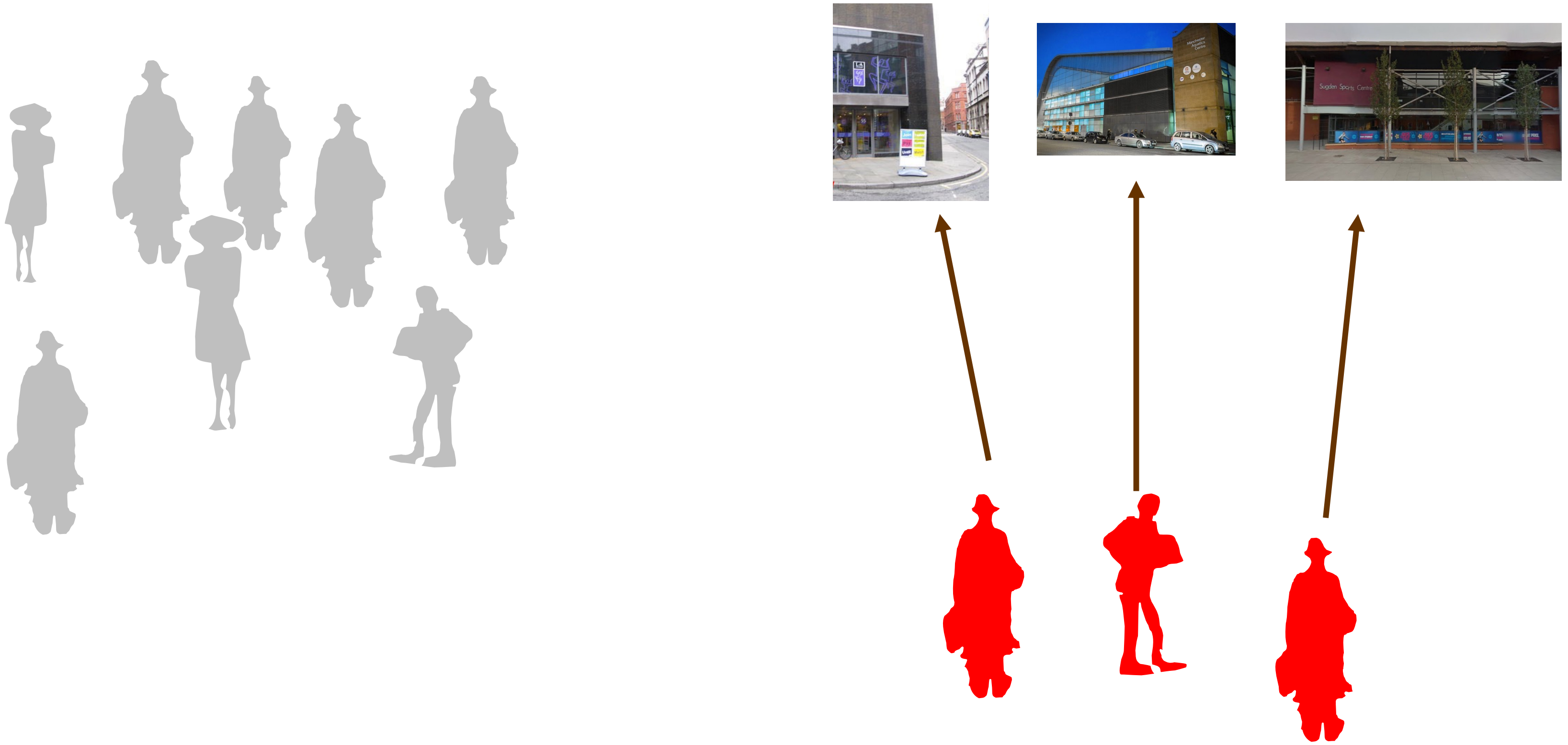


Affiliated with same things

Same attributes, same activities, form ties

Propinquity, context, and foci

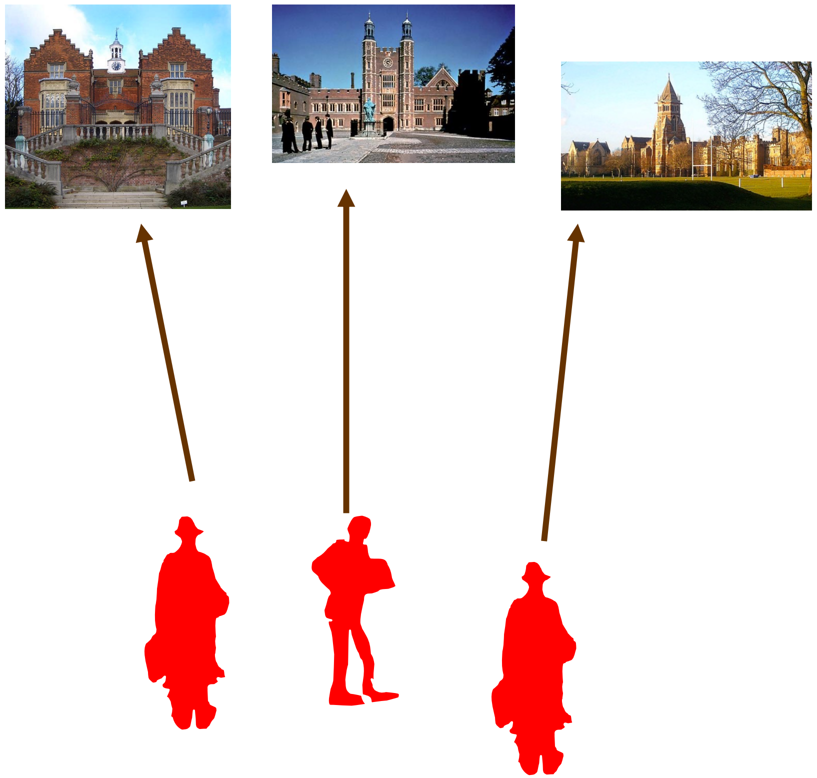
Affiliated with similar (not the same) things



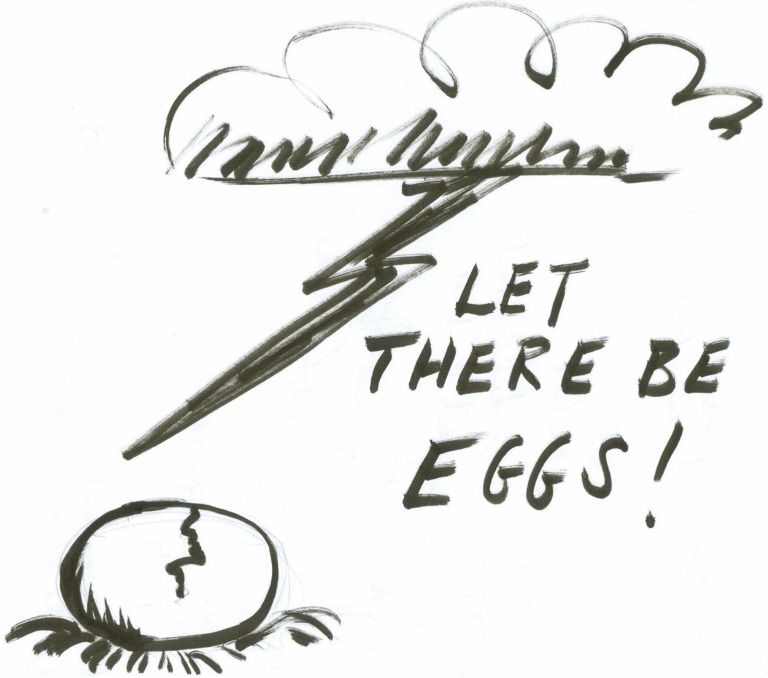
same type of activities leads to same attributes form ties

Propinquity, context, and foci

Affiliated with similar (not the same) things



same type of activities leads to same attributes form ties


LET
THERE BE
EGGS!


LET
THERE
BE
CHICKENS!

Mechanisms behind social influence

Social influence

- Persuasion
- Imitation
- Diffusion
- Adoption
- Contagion
- Learning
- etc

factors

- ✓ Active persuasion?
- ✓ Who is being influenced
- ✓ Is the attribute salient or hidden?
- ✓ Is it an absorbing state?
- ✓ Number of contacts or majority of contacts
- ✓ Is variable a single thing or many?
- ✓ do you seek it out or is it simply transmitted?

History

Jacob Moreno

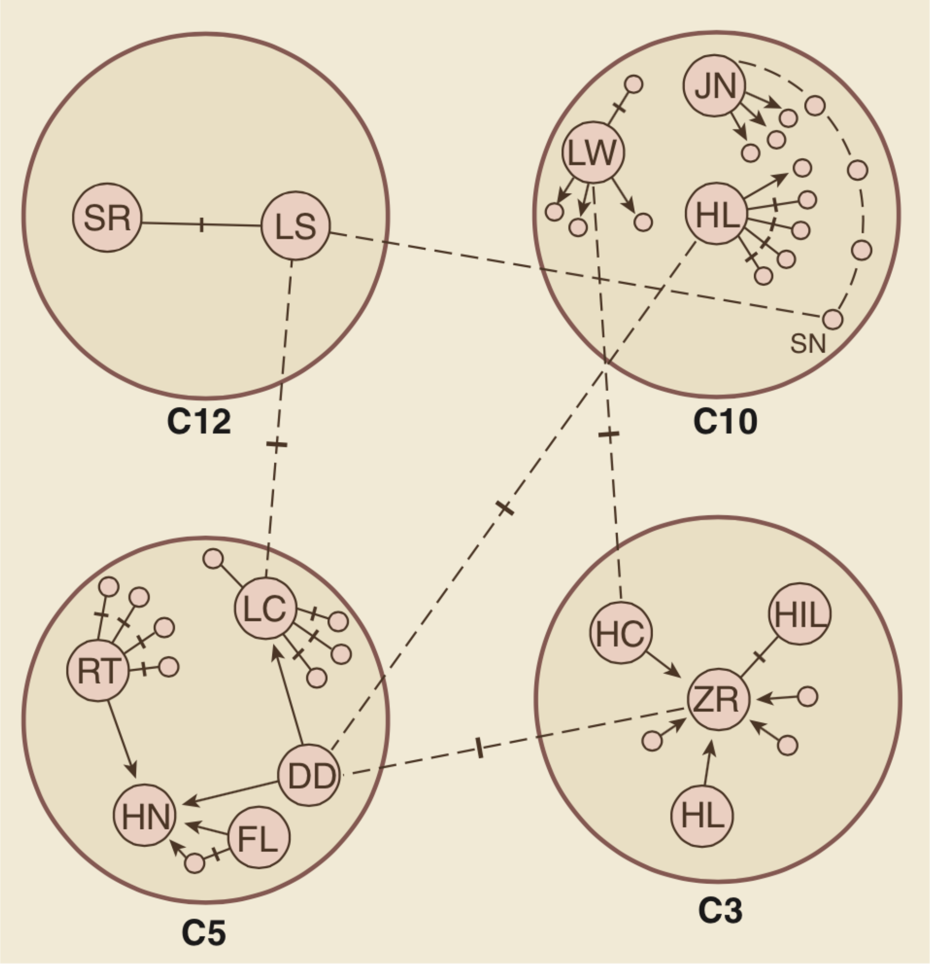
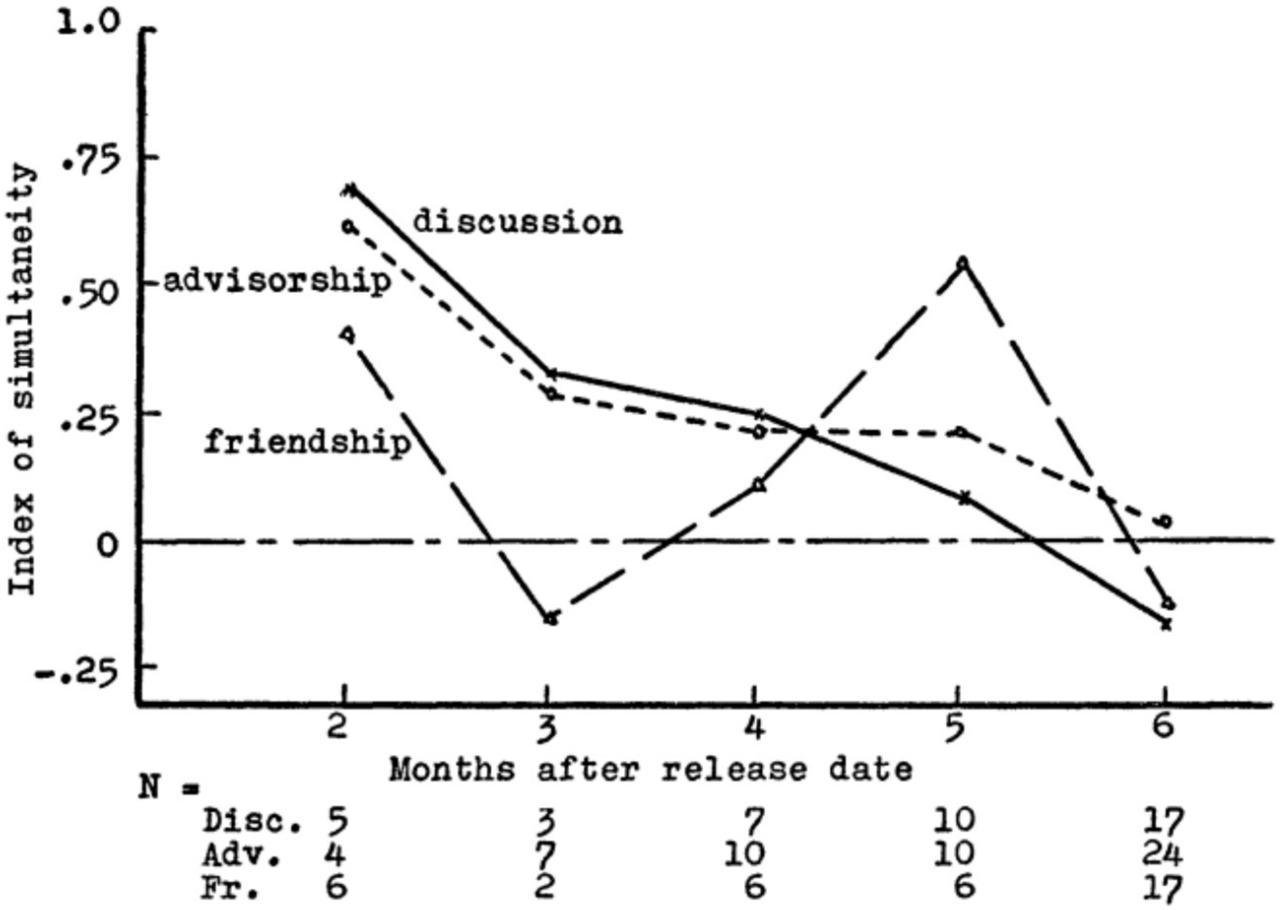


Fig. 1. Moreno's network of runaways. The four largest circles (C12, C10, C5, C3) represent cottages in which the girls lived. Each of the circles within the cottages represents an individual girl. The 14 runaways are identified by initials (e.g., SR). All nondirected lines between a pair of individuals represent feelings of mutual attraction. Directed lines represent one-way feelings of attraction.

feelings toward one another (Fig. 1). The links in this social network, Moreno argued, provided channels for the flow of social influence and ideas among the girls. In a way that even the girls themselves may not have been conscious of, it was their location in the social network that determined whether and when they ran away.

Coleman, Katz, and Menzel (1957)

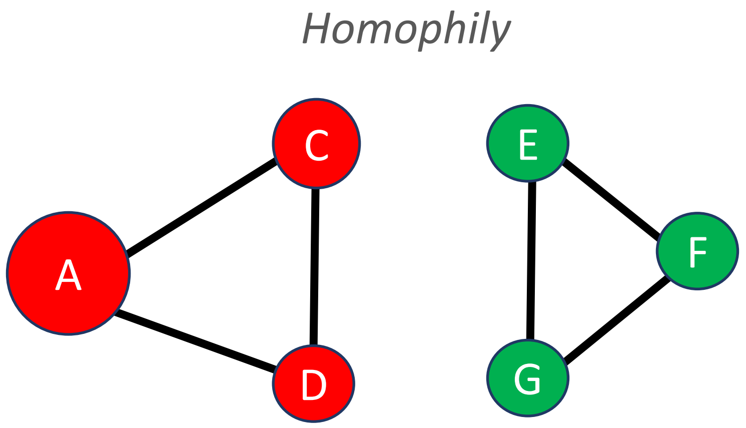


To analyse pairs of individuals instead of single individuals may seem like only a very modest step in the direction of the analysis of networks of social relations. And so it is; it would be more satisfactory, and truer to the complexity of actual events, if it were possible to use longer chains and more ramified systems of social relations as the units of analysis. But so little developed are the methods for the analysis of social processes, that it seemed best to be content with the analysis of pair relationships"
Coleman, Katz and Menzel 1966, p 114

The Diffusion of an Innovation Among Physicians, Sociometry

Investigating prevalence of influence

Statistic



$$Y_i = \begin{cases} 1 & \text{, if } i \text{ is a smoker} \\ 0 & \text{,otherwise} \end{cases}$$

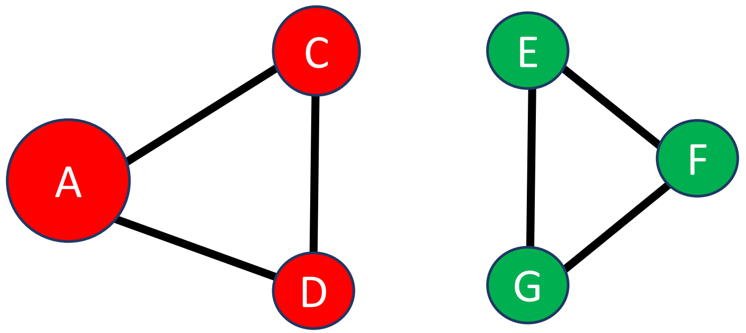
$$X_{ij} = \begin{cases} 1 & \text{, if } i \text{ has a tie to } j \\ 0 & \text{,otherwise} \end{cases}$$

Number of pairs of nodes that are connected and are smokers $\sum_j y_i y_j x_{ij}$

Number of pairs of nodes that are connected have same smoking status $\sum_j I\{y_i = y_j\} x_{ij}$

Statistic

Homophily



Y_i : *Trust in government (on a scale)*

$$X_{ij} = \begin{cases} 1 & \text{, if } i \text{ has a tie to } j \\ 0 & \text{,otherwise} \end{cases}$$

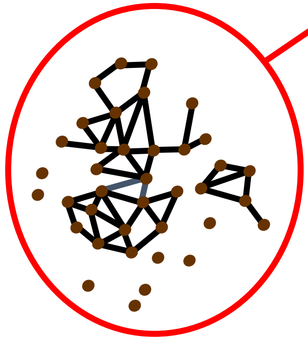
Tied partners with high values contribute more $\sum_j y_i y_j x_{ij}$

Distance between tied partners in trust $\sum_j |y_i - y_j| x_{ij}$

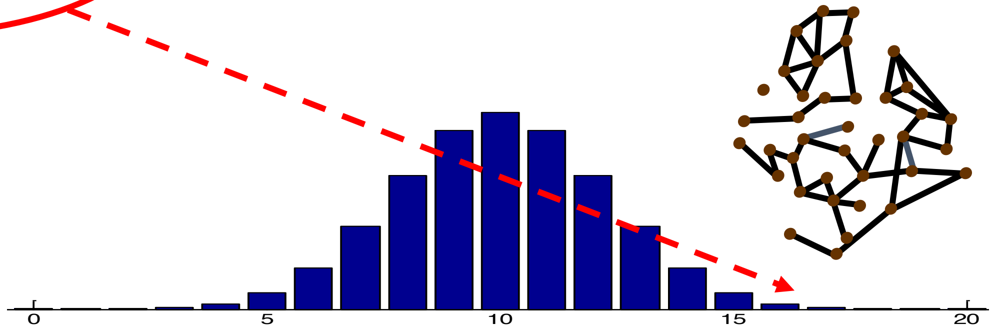
Similarity between tied partners in trust $\sum_j (1 - |y_i - y_j|/\mathrm{d}_{max}) x_{ij}$

How test?

$$\sum_j |y_i - y_j| x_{ij}$$



Observed network



Distribution of measure under null distribution

Randomise networks?

What are we testing

- What is the counterfactual?
- Drawing inference to network with same properties
- There are n(n-1)/2 dyads but only n outcomes

$X \sim$ Bern(p): the outdegrees are now independent of each other

$X \sim$ U | $L(X)$:

$X \sim$ U | d_{out} : we can fix outdegree only – what does activity explain

$X \sim$ U | d_{in} : we can fix indegree only – what does popularity explain

$X \sim$ U | d_{out}, d_{in} : we can fix outdegree AND indegree

$X \sim$ U | MAN: what can reciprocity explain

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{pmatrix}$$

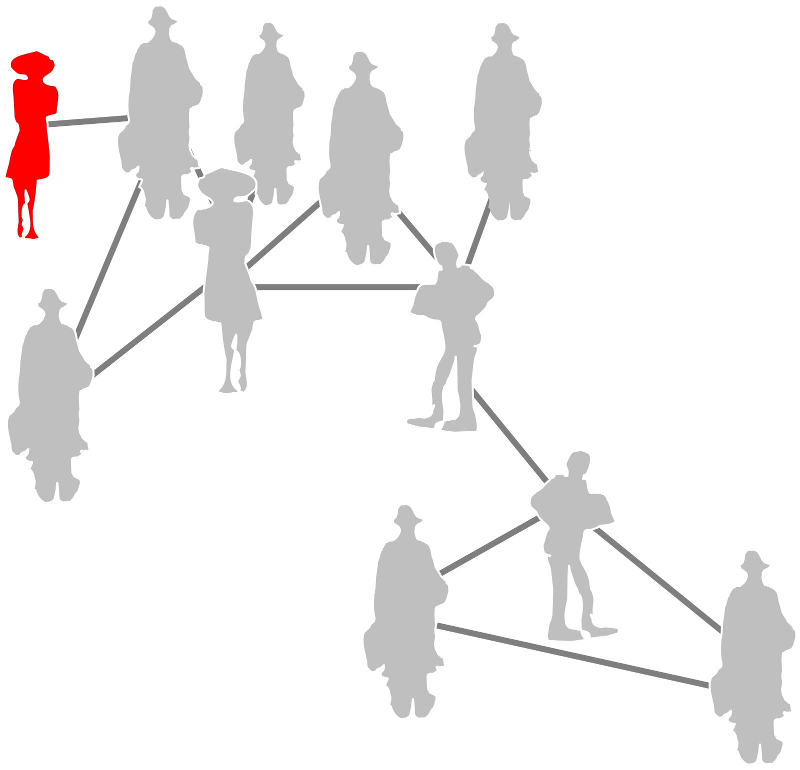
$$X = \begin{pmatrix} - & X_{12} & X_{13} & X_{14} & X_{15} \\ X_{21} & - & X_{23} & X_{24} & X_{25} \\ X_{31} & X_{32} & - & X_{34} & X_{35} \\ X_{41} & X_{42} & X_{43} & - & X_{45} \\ X_{51} & X_{52} & X_{53} & X_{54} & - \end{pmatrix}$$

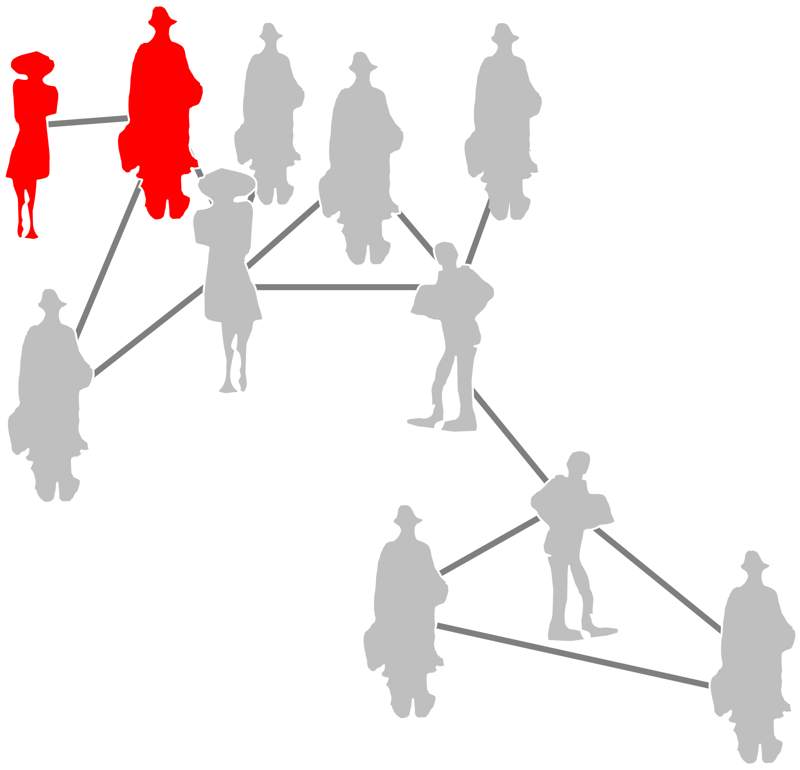
Randomise outcome?

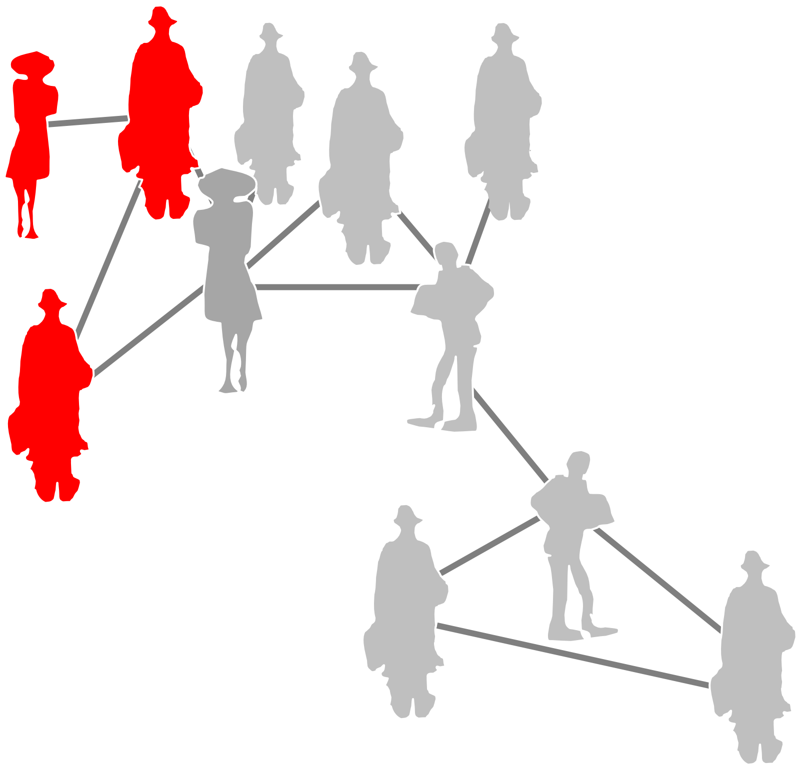
Interpretation: ‘if all actors swapped positions…’
Only values in the range of what we have observed
What about other covariates ‘what if this teen who loves tik-tok is all of sudden a 50-year-old man…’

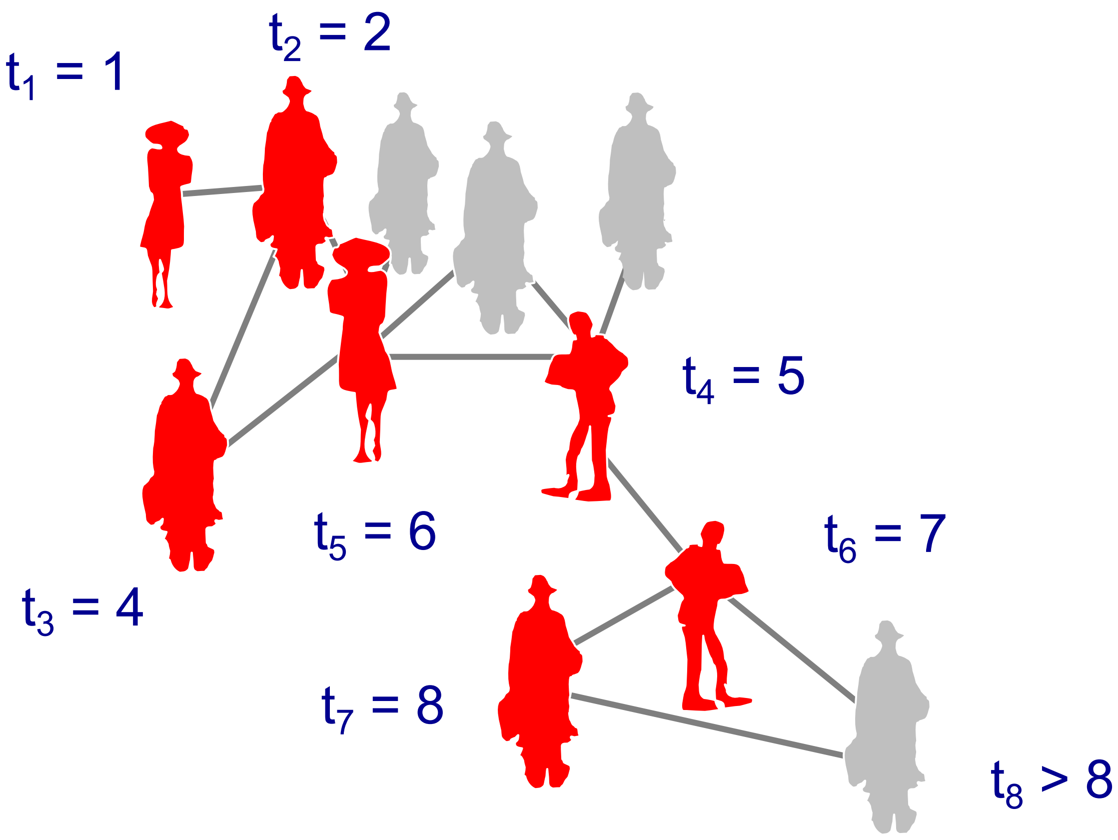
$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{pmatrix} \qquad X = \begin{pmatrix} - & X_{12} & X_{13} & X_{14} & X_{15} \\ X_{21} & - & X_{23} & X_{24} & X_{25} \\ X_{31} & X_{32} & - & X_{34} & X_{35} \\ X_{41} & X_{42} & X_{43} & - & X_{45} \\ X_{51} & X_{52} & X_{53} & X_{54} & - \end{pmatrix}$$

Event-history analysis



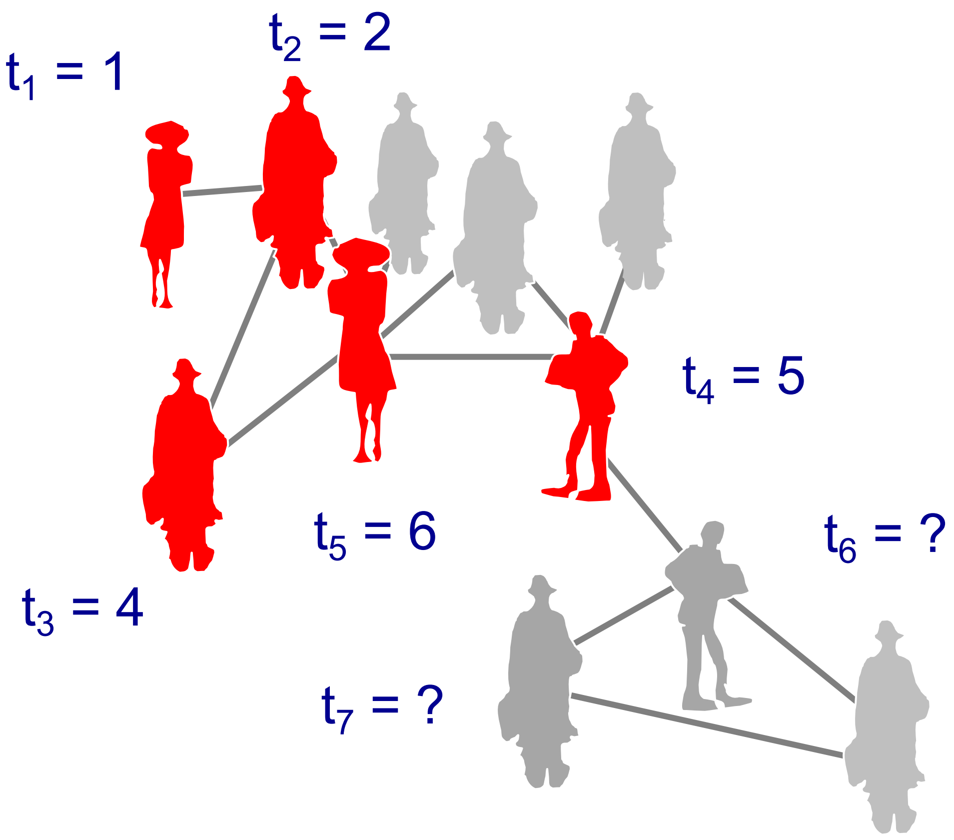






Empirical question:
Is the *time to adoption* shorter for individuals that have a tie to someone who has already adopted
Compared to people that do not have a tie to someone who has adopted?

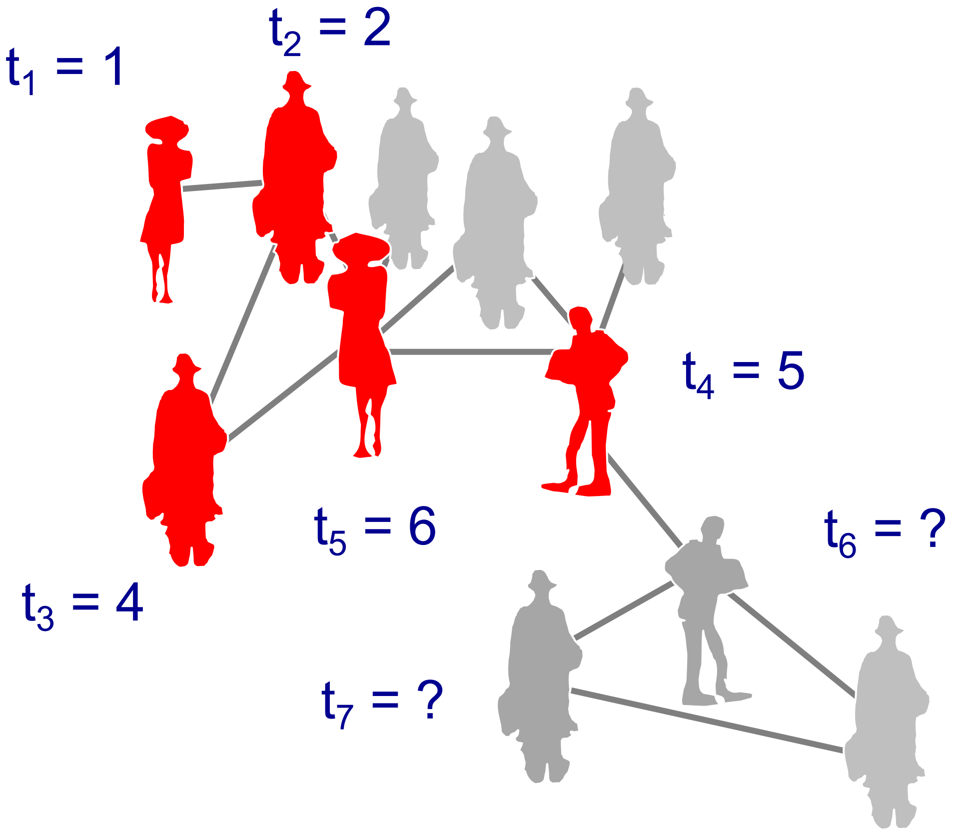
t_k = time of adoption of the k^{th} adopter or k^{th} ndividual



$\Pr(t_6 < t_7) = ?$

Empirical question:
Is the *time to adoption* shorter for individuals that have a tie to someone who has already adopted
Compared to people that do not have a tie to someone who has adopted?

Assume time t = 6:
who will adopt next?



$$\Pr(\ t_6 < \ t_7) = \Pr(\ (t_6 - 6) < \ (t_7 - 6)\ )$$

Assume time t = 6:
who will adopt next?

If $t_i \sim \mathrm{Exp}(\lambda_i)$:

$$\Pr(\ t_i < \ s+t \mid t > t) = \Pr(\ t_i < \ s\ )$$

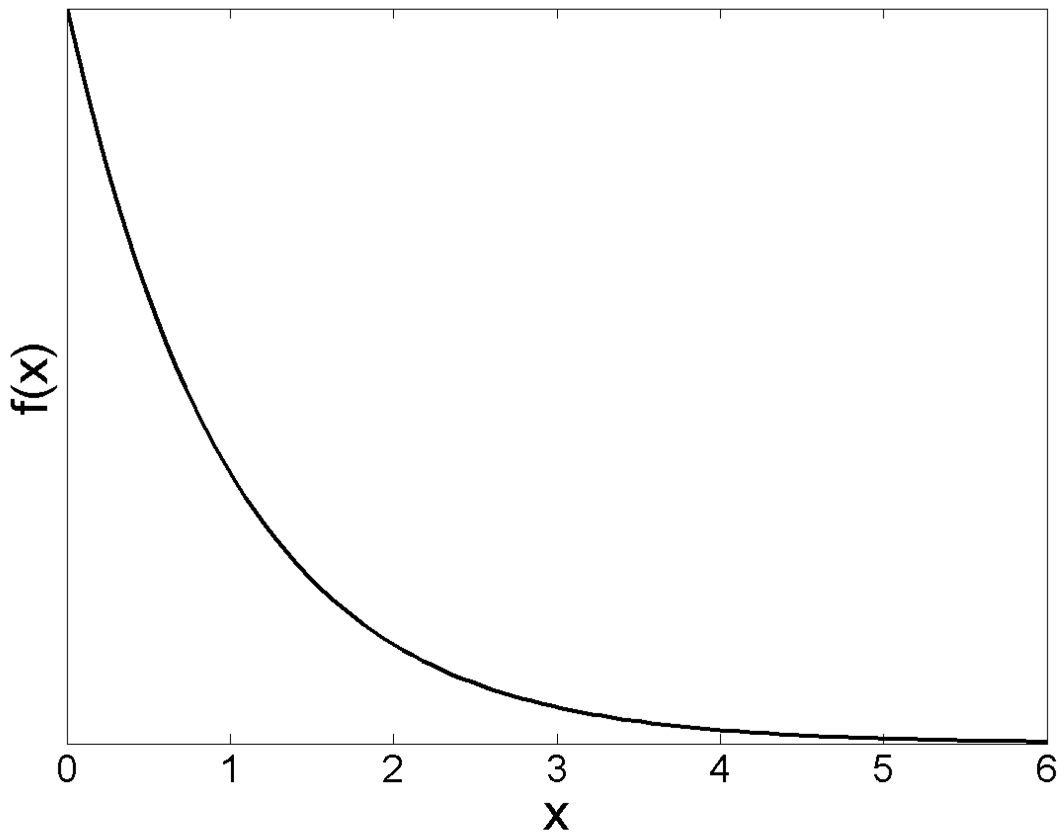
Exponential distribution memoryless

Assume that we want to model something positive that gets increasingly less probably the larger it gets

- The time till divorce
- The time spent as unemployed

$$f(x) = \lambda e^{-\lambda x}$$

probability density function (pdf)

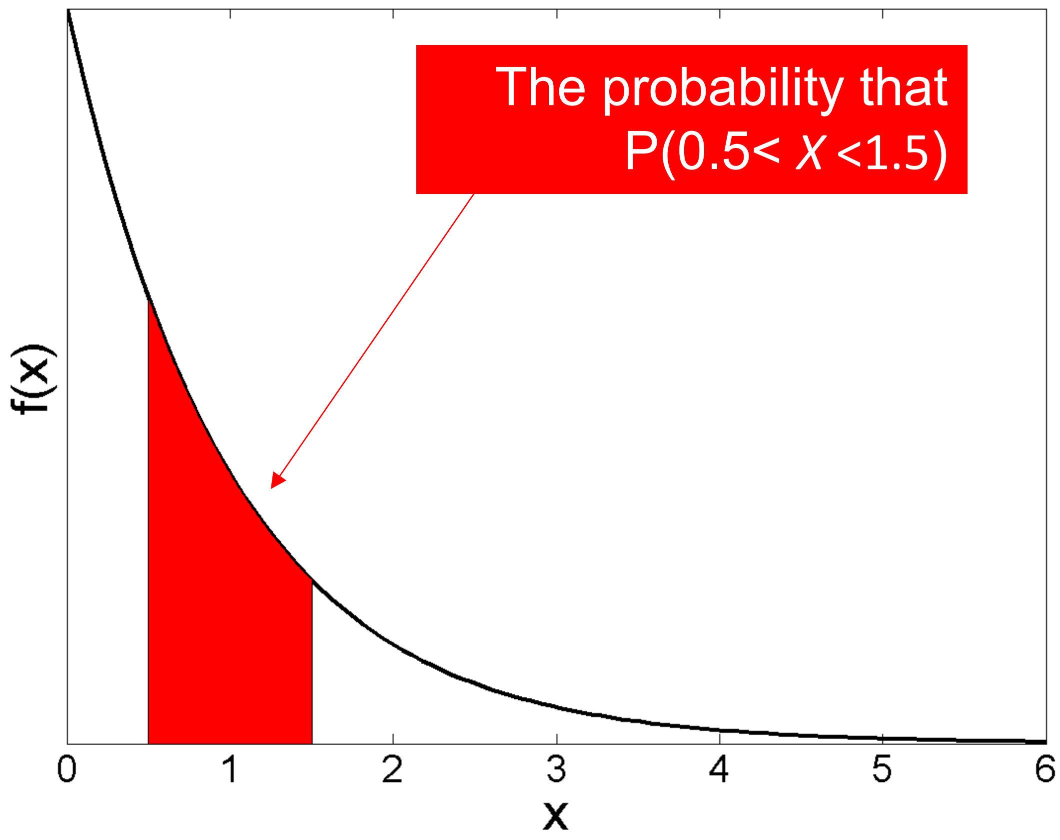


Assume that we want to model something positive that gets increasingly less probably the larger it gets

- The time till divorce
- The time spent as unemployed

$$f(x) = \lambda e^{-\lambda x}$$

probability density function (pdf)

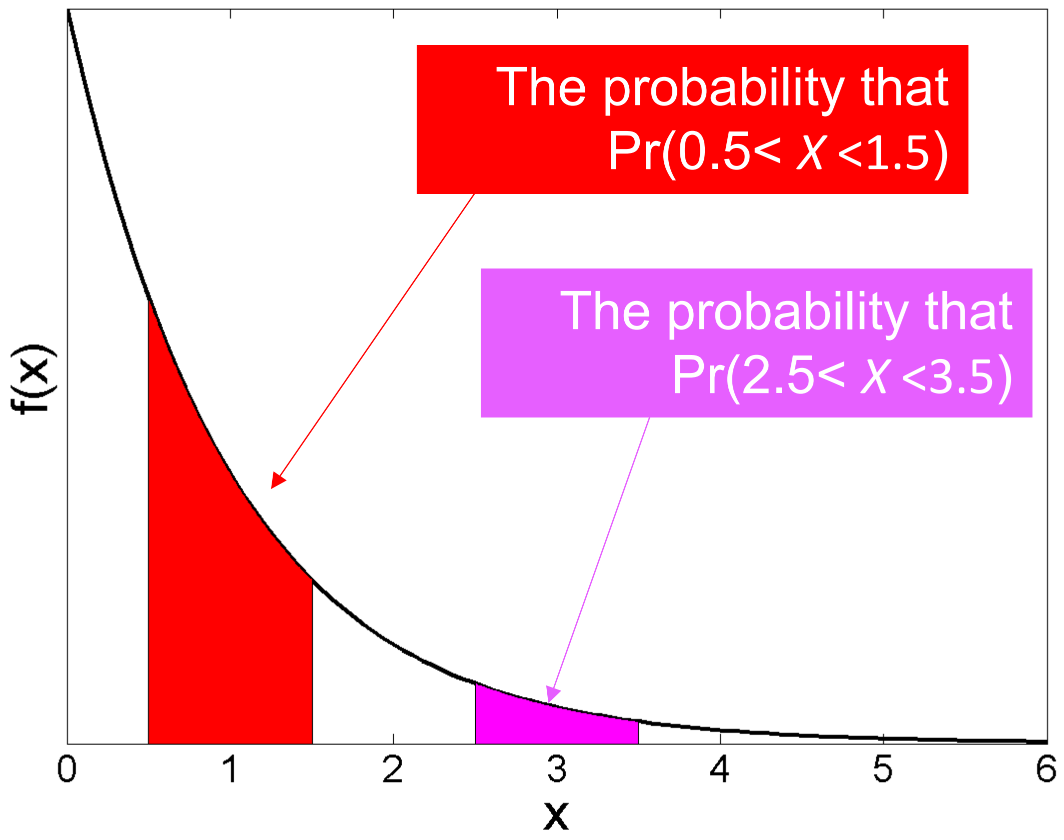


Assume that we want to model something positive that gets increasingly less probably the larger it gets

- The time till divorce
- The time spent as unemployed

$$f(x) = \lambda e^{-\lambda x}$$

probability density function (pdf)



The higher the rate

λ

The quicker the change

The parameter

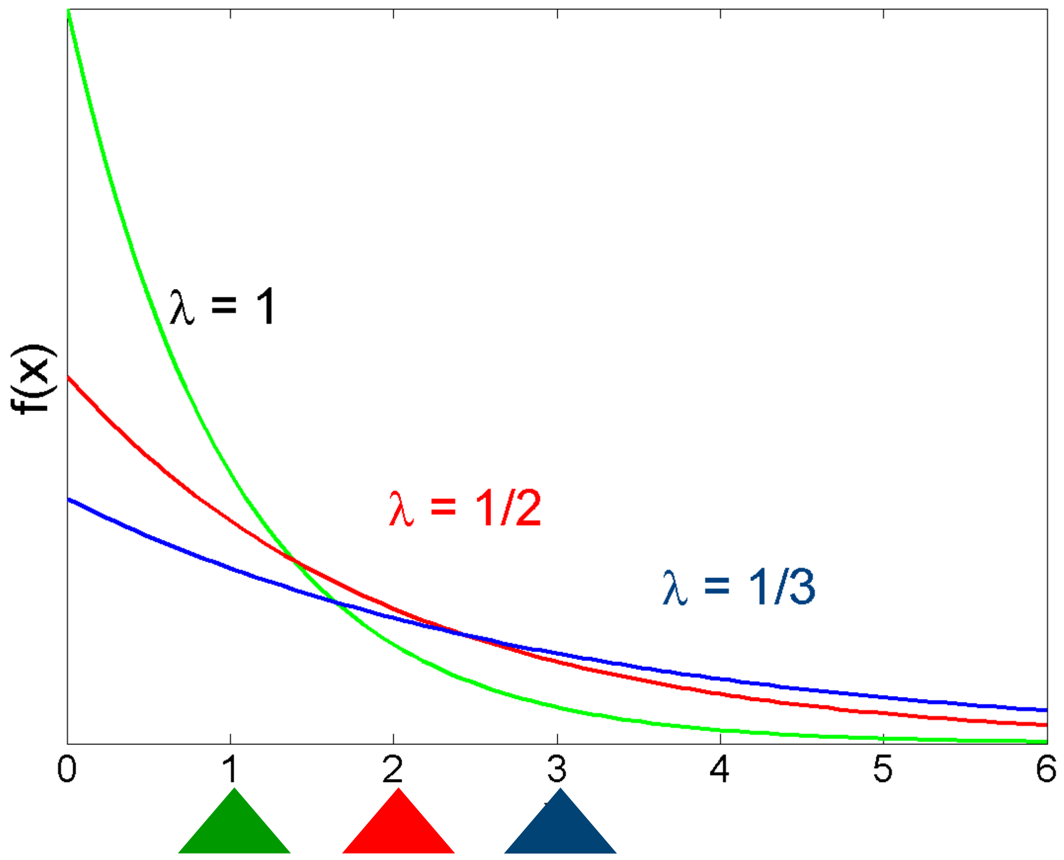
λ

Describes the location:

$$E(X) = 1/\lambda$$

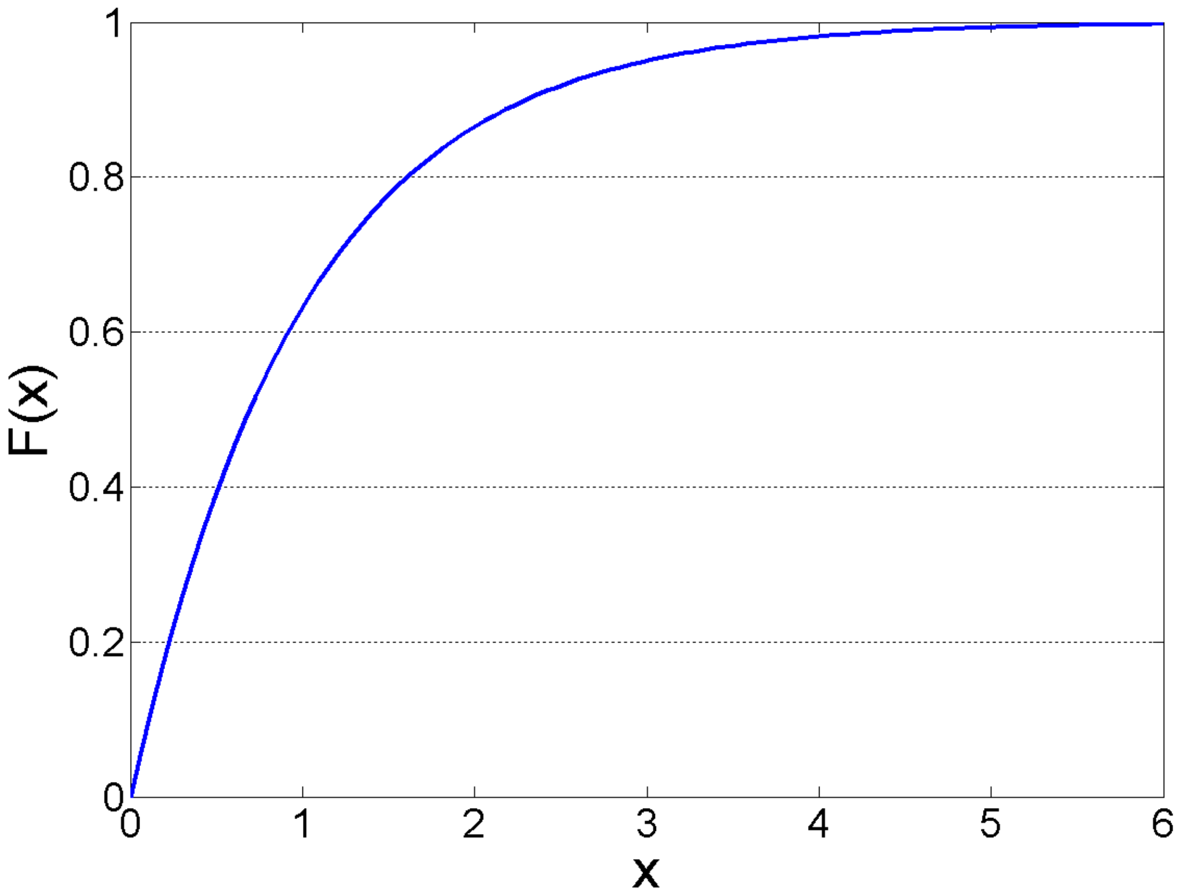
And the spread

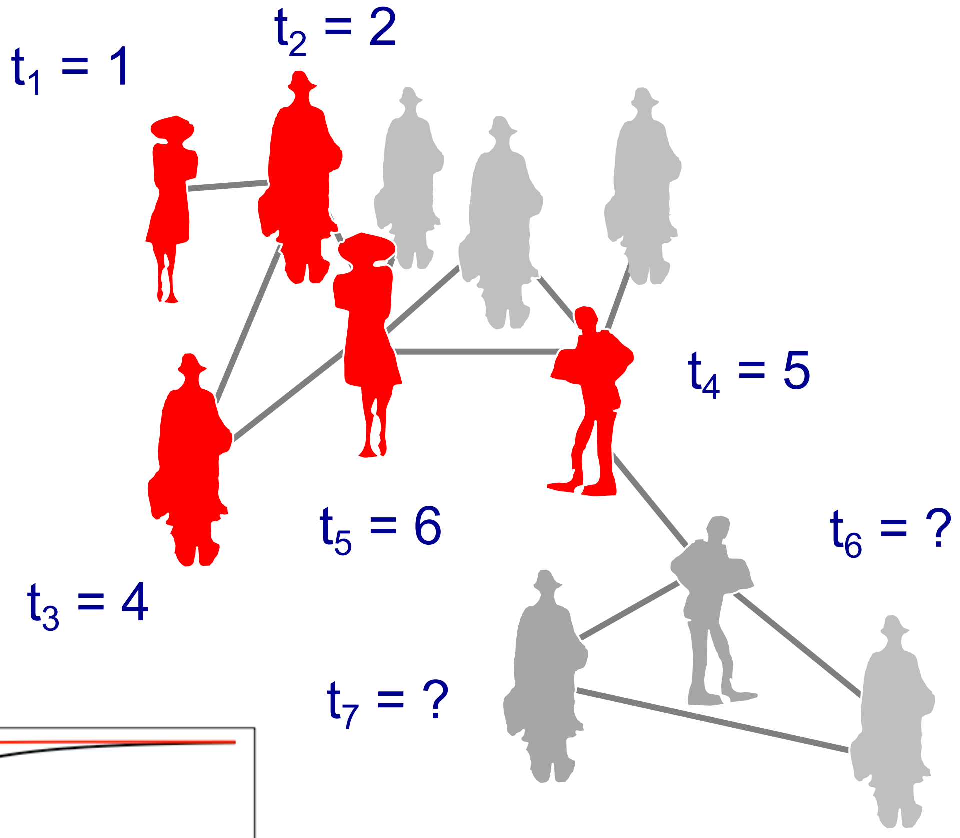
$$V(X) = 1/\lambda^2$$

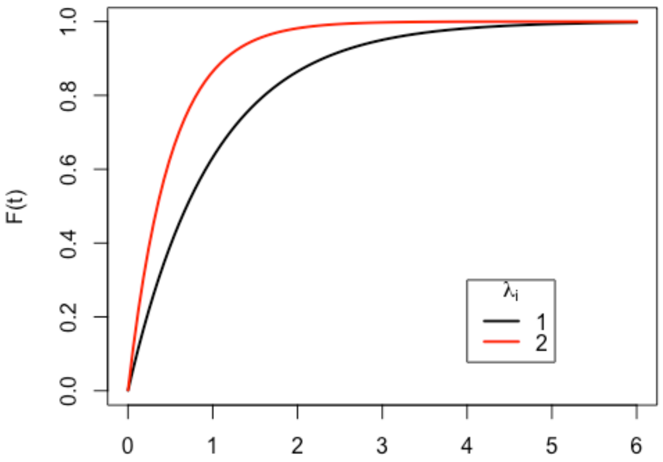


A useful property is that we can write down the cdf

$$\Pr(X < x) = F(x) = 1 - e^{-\lambda x}$$





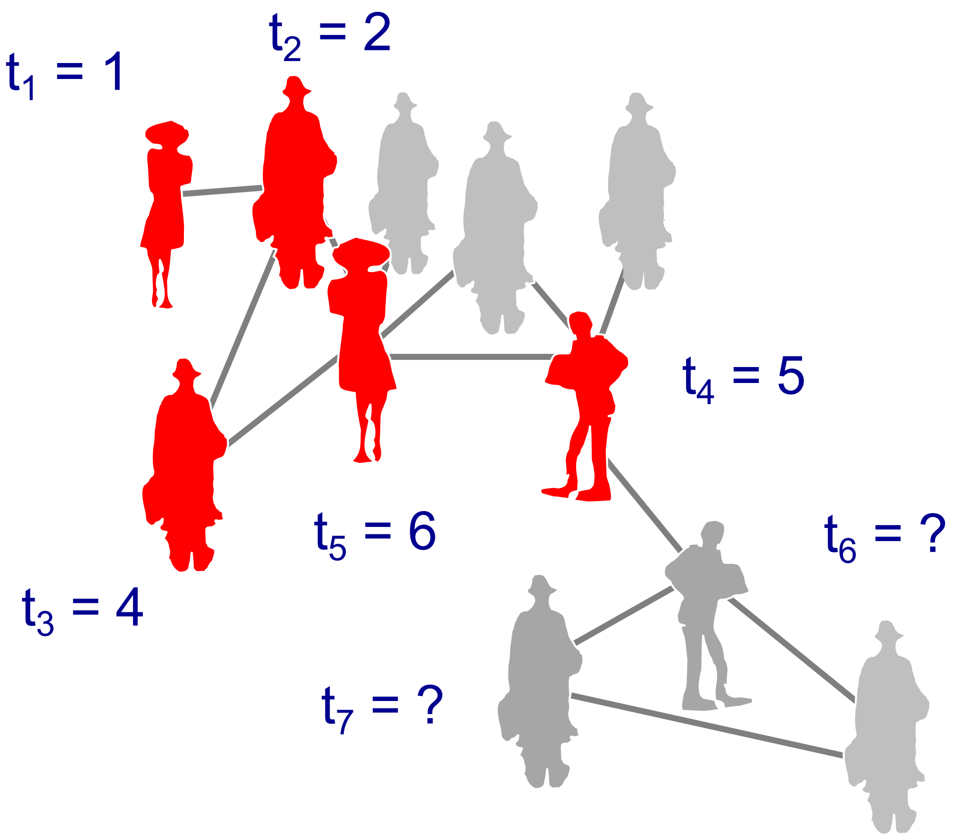


For diffusion we want

$t_6 \sim \mathrm{Exp}(\lambda_6)$
$t_7 \sim \mathrm{Exp}(\lambda_7)$

So that $\Pr(t_6 < t_7)$ is large

If we set $\lambda_6 > \lambda_7$, then
$E(t_6) = 1/\lambda_6 < 1/\lambda_7 = E(t_7)$
and
$\Pr(t_6 < t) > \Pr(t_7 < t)$



Define for each i

$$Y_i(t) = \begin{cases} 1 & \text{,if } i \text{ has adopted at time } t \\ 0 & \text{,otherwise} \end{cases}$$

For any point in time t_0

$$\Pr(t_i < t_0 + s \mid Y_i(t_0) = 0)$$
$$= \Pr(t_i < s \mid Y_i(t_0) = 0)$$

$$t_i - t_0 \sim \mathrm{Exp}\{ \lambda_i(Y(t_0), X) \}$$

A rate that depends on all $Y_j(t)$ and network X

Piecewise constant rate functions

rate must be positive

function linking current state to i's rate of change

$$\lambda_i(Y(t), X(t)) = \exp\{\alpha + \beta a_i(Y(t), X(t))\}$$

rate specific to *i*

A rate that depends on all $Y_j(t)$ and network $X(t)$

Statistical parameter for influence strength

Note: if a change occur to $Y(t)$, rates are updated

Piecewise constant rate functions

$$\lambda_i(Y(t), X(t)) = \exp\{\alpha+\beta a_i(Y(t), X(t))\}$$

Positive and large β – stronger dependence on current state

If β>0 the greater a_i, the quicker the change

Note: if a change occur to $Y(t)$, rates are updated

Meyers (2000): rates depend on important adopters

$$a_i(y, x) = \sum_j y_j s_j(x)$$

y_j = 1 if other person j has adopted

The adoption y_j = 1 carries more weight if importance $s_j(x)$ of j is great

The more 'important' nodes that have adopted, the more other people will adopt

Valente (2005): rates depend on important adopters that you are connected to

The adoption $y_j = 1$ carries more weight if importance $s_j(x)$ of j is great

$$a_i(y, x) = \sum_j y_j x_{ij} s_j(x)$$

$y_j = 1$ if other person j has adopted

$x_{ij} = 1$ if i has a tie to j

The more 'important' nodes that have adopted, the more other people will adopt

Total exposure

Count of how many people that i is connected to have already adopted

$$a_i(y, x) = \underbrace{\sum_j y_j x_{ij}}$$

y_j = 1 if other person j has adopted

x_{ij} = 1 if i has a tie to j

The more people you know that have adopted, the quicker you will adopt yourself

Average exposure

proportion of how many people that i is connected to have already adopted

$$a_i(y, x) = \frac{\overbrace{\sum_j y_j x_{ij}}}{\sum_j x_{ij}}$$

degree of i

The greater the proportion of people you know that have adopted, the quicker you will adopt yourself

Average v total exposure

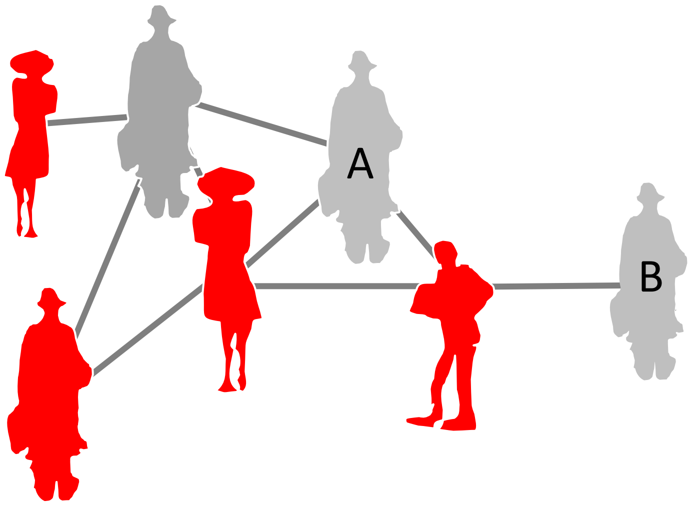
$$a_i(y, x) = \frac{\sum_j y_j x_{ij}}{\sum_j x_{ij}}$$

$$a_A(y, x) = \frac{2}{3}$$

$$a_B(y, x) = \frac{1}{1} = 1$$

RED: $y_j = 1$

GREY: $y_j = 0$



$$a_i(y, x) = \sum_j y_j x_{ij}$$

$$a_B(y, x) = 1$$

$$a_A(y, x) = 2$$

When is the total number important and when is it the proportion – examples?

Further implications of exponential times

In the process, all (susceptible) nodes draw a time $T_i \sim \text{Exp}\{\, \lambda_i(Y(t), X(t)) \,\}$
The node with the shortest time, gets to change – we call that the winner

What is the probability that *i* is the winner?

$$\Pr(T_i = \text{argmin}\{T_1, \ldots, T_n\}) = \frac{\lambda_i(Y(t), X(t))}{\sum_j \lambda_j(Y(t), X(t))}$$

A property of the exponential distribution

Further implications of exponential times

In the process, all (susceptible) nodes draw a time $T_i \sim \mathrm{Exp}\{ \lambda_i(Y(t),X(t)) \}$
The node with the shortest time, gets to change – we call that the winner

What is distribution of the winning time?

$$\min\{T_1 ,\dots, T_n\} \sim \mathrm{Exp}\left(\textstyle\sum_j \lambda_j(Y(t),X(t)) \right)$$

A property of the exponential distribution

Simulation of the diffusion process

At time t

Pick a winner i with probability

$$\frac{\lambda_i(Y(t),X(t))}{\sum_j \lambda_j(Y(t),X(t))}$$

Draw winning time

$$s \sim \mathrm{Exp}\left(\sum_j \lambda_j(Y(t),X(t)) \right)$$

Update status of winner

$$Y(t+s) := Y(t)+1$$

Increment time

$$t := t+s$$

Estimation of the diffusion process

Based on the network X, and the times of adoption

$$t_1, \ldots, t_k$$

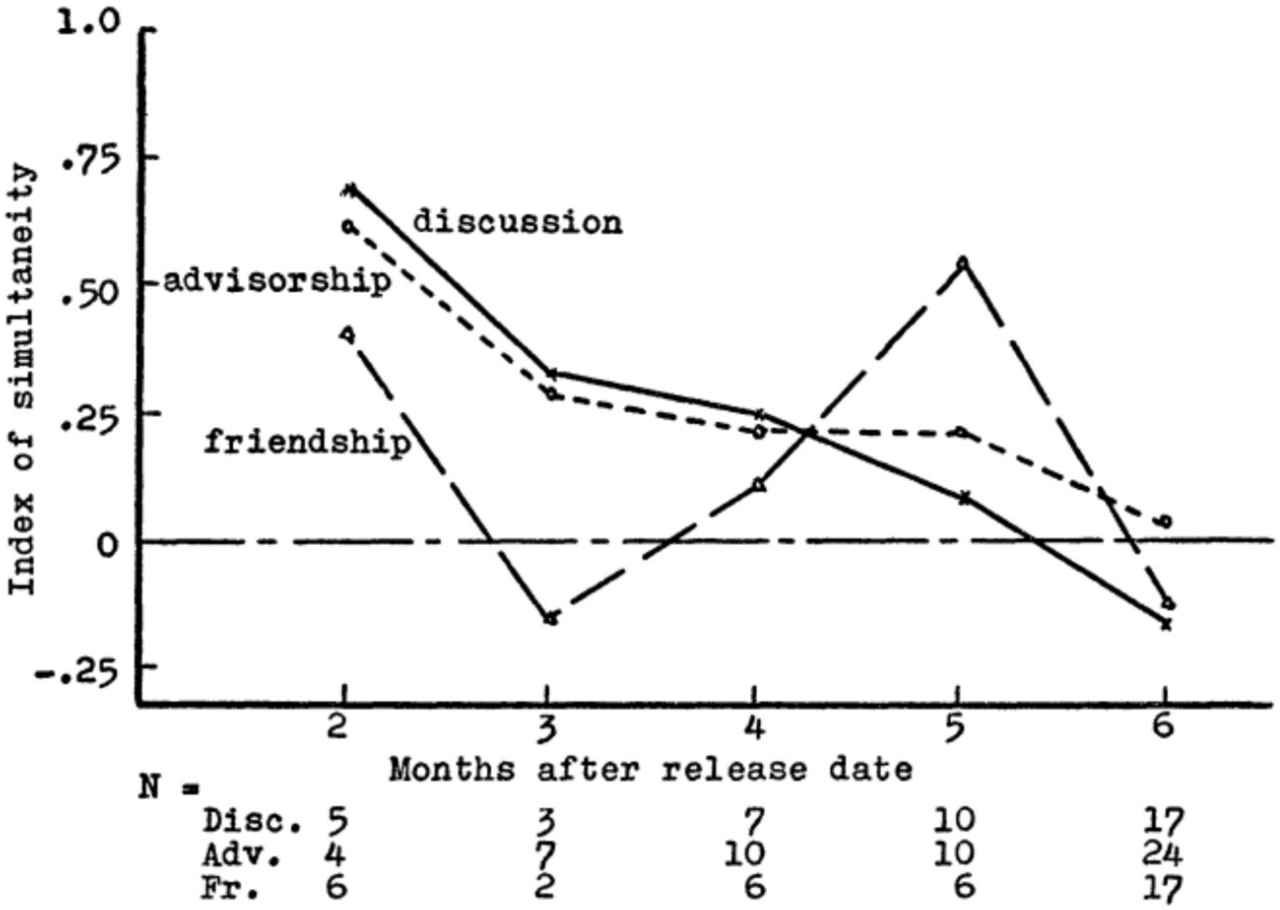
The likelihood is just a product of exponential pdf's and cdf's (for non-adopters)

$$\prod \lambda_j(Y(t_j), X(t_j)) \exp\left\{ -\sum_j (t_j - t_{j-1}) \lambda_j(Y(t_j), X(t_j)) \right\}$$

$$\times \underbrace{\prod \exp\{ -(t_j - t_{j-1}) \lambda_h(Y(t_j), X(t_j)) \}}_{\boxed{\textit{Pr(T>t): censoring for those who didn't adopt}}}$$

Parameters can be estimated using Maximum likelihood or Bayes

EXAMPLE: Coleman, Katz, and Menzel (1957)

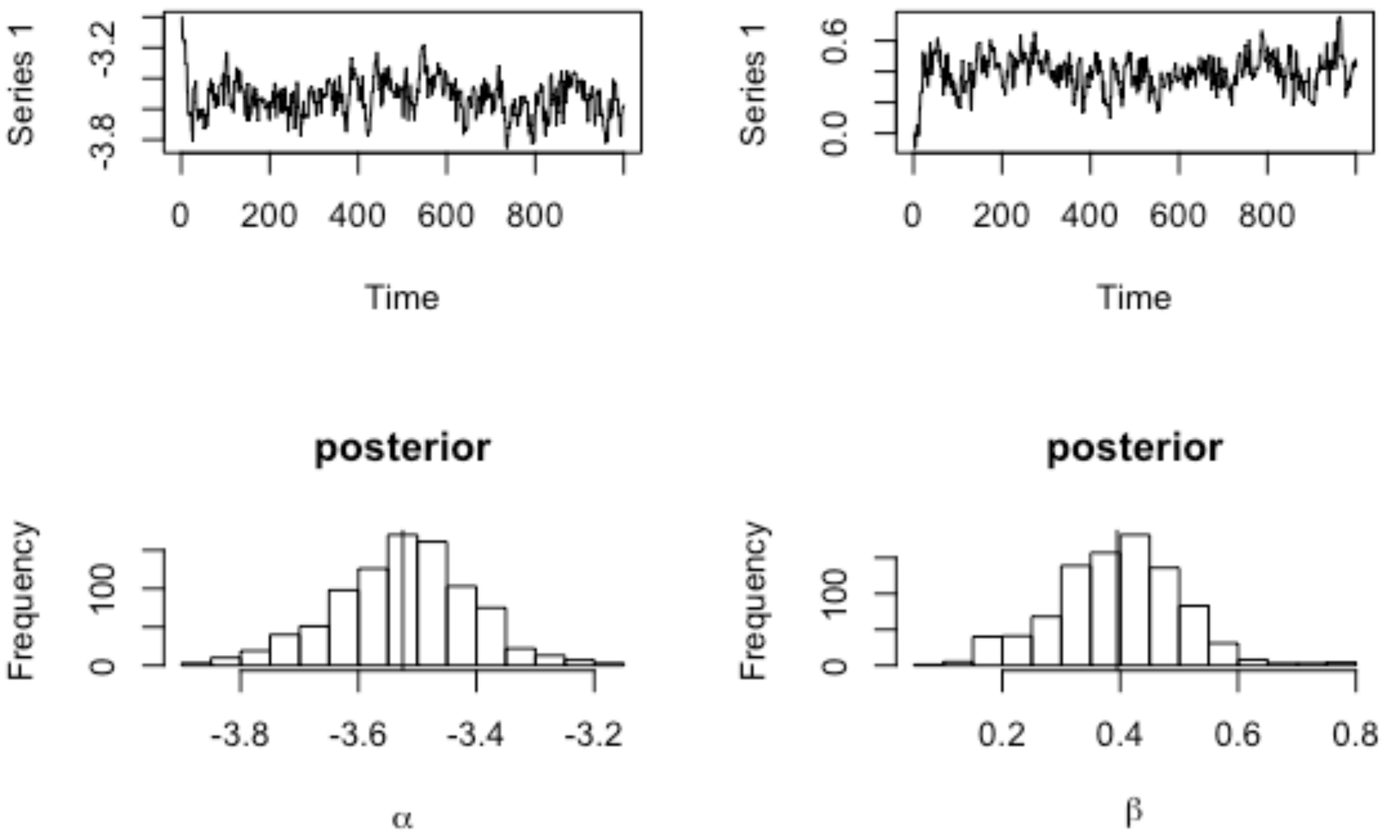


To analyse pairs of individuals instead of single individuals may seem like only a very modest step in the direction of the analysis of networks of social relations. And so it is; it would be more satisfactory, and truer to the complexity of actual events, if it were possible to use longer chains and more ramified systems of social relations as the units of analysis. But so little developed are the methods for the analysis of social processes, that it seemed best to be content with the analysis of pair relationships"
Coleman, Katz and Menzel 1966, p 114

The Diffusion of an Innovation Among Physicians, Sociometry

EXAMPLE: Coleman, Katz, and Menzel (1957)

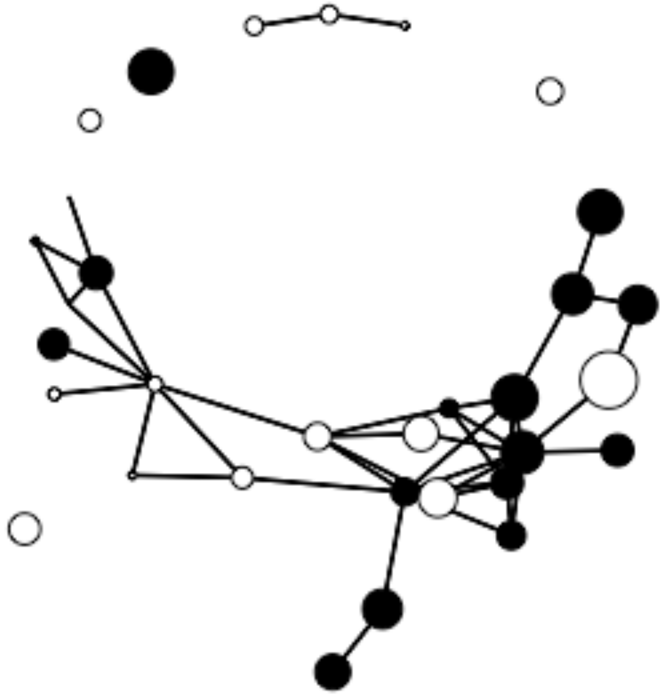
difusion.html



The Diffusion of an Innovation Among Physicians, Sociometry

Social influence in cross-sectional data

A dataset of academics

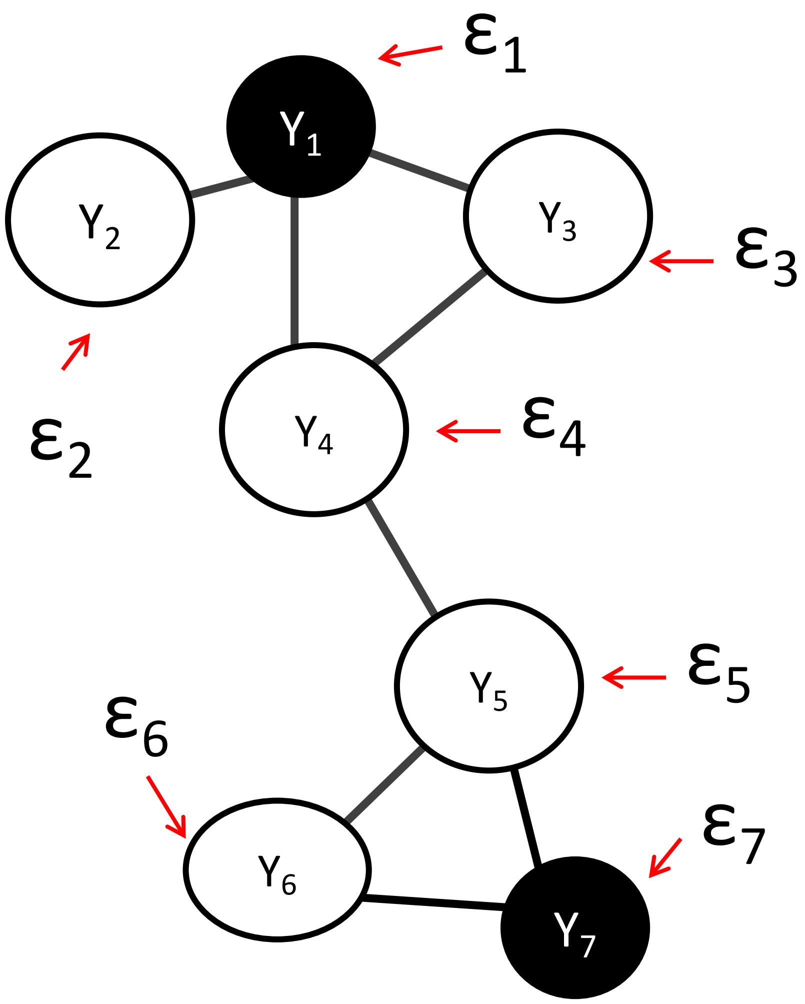


Y_i : *Log citation of researcher i*

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

$$M_i : \begin{cases} 1 & \text{, if } i \text{ sociologist} \\ 0 & \text{,otherwise} \end{cases}$$

Ordinary least squares regression



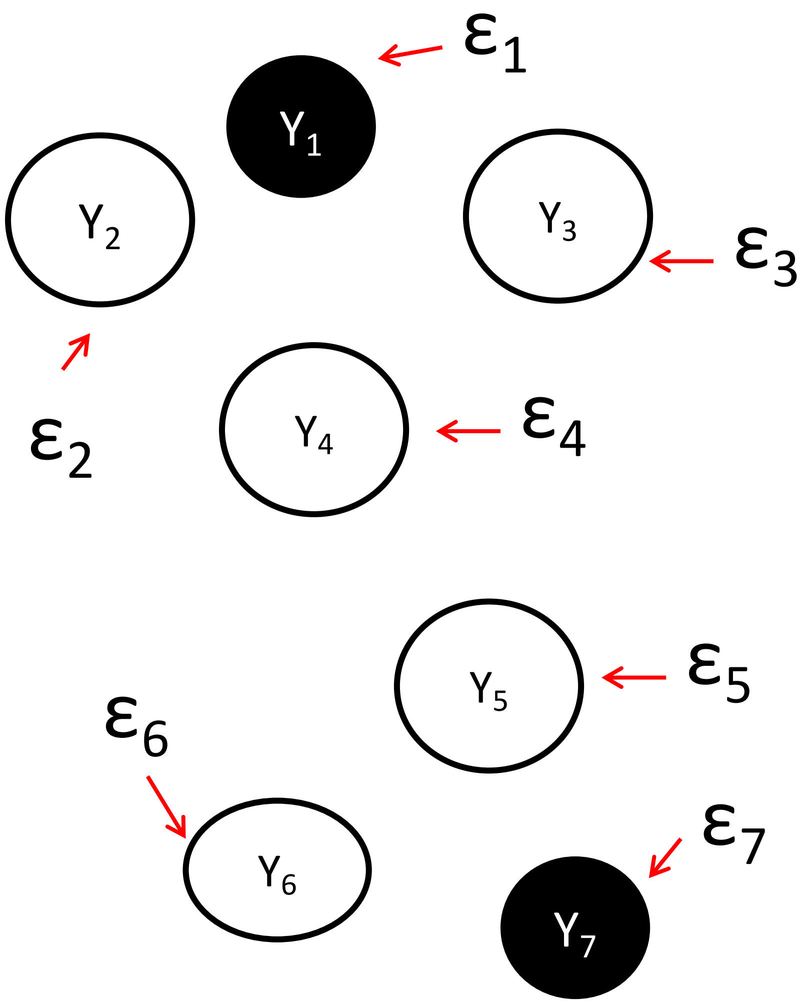
$$Y_i = \alpha + \beta M_i + \varepsilon_i$$

where

$$\varepsilon_i \sim N(0, \sigma^2)$$

INDEPENDENT for all *i*

Ordinary least squares regression



$$Y_i = \alpha+\beta M_i+\varepsilon_i$$

where

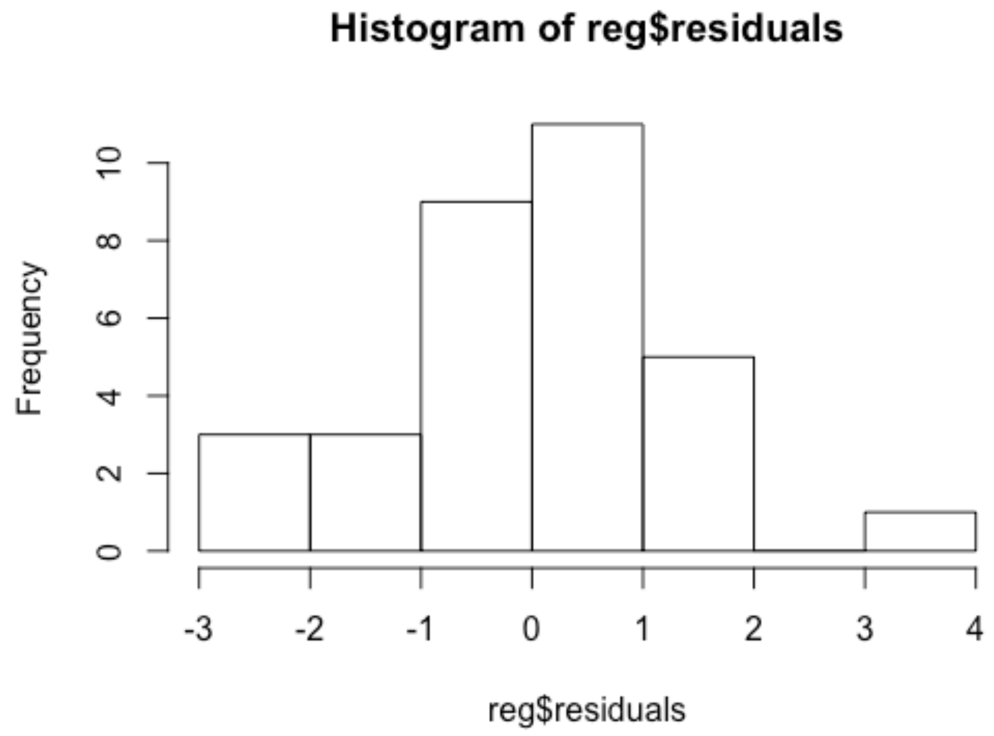
$$\varepsilon_i \sim N(0,\sigma^2)$$

INDEPENDENT for all *i*

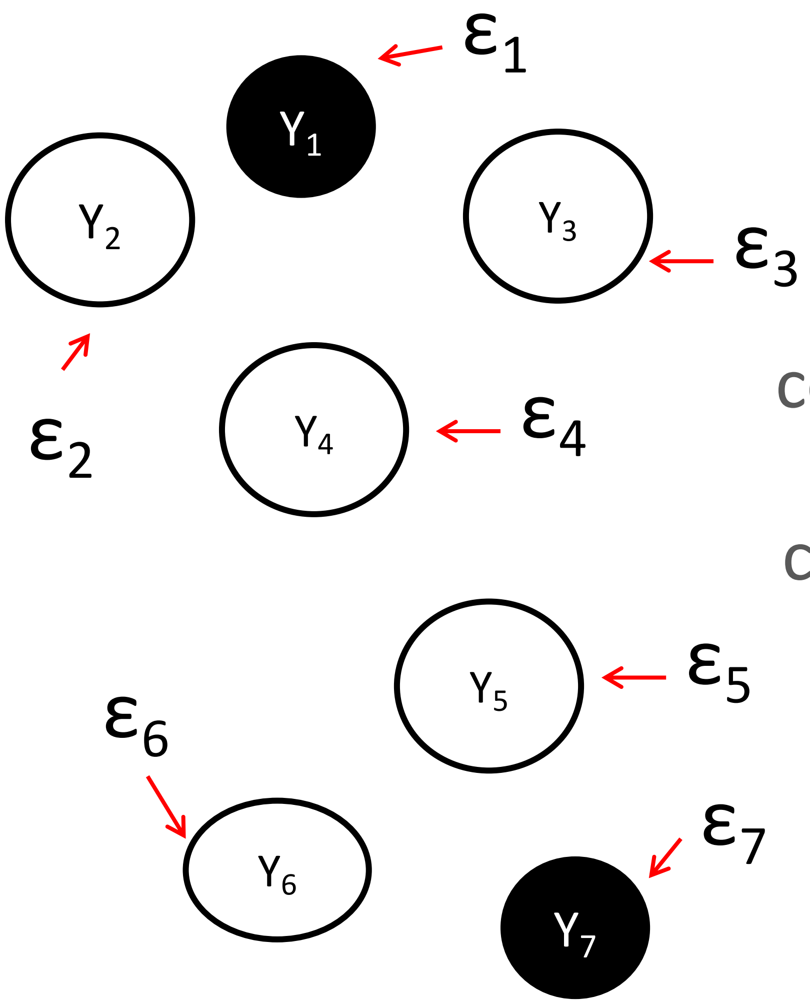
OLS

$\varepsilon_i \sim N(0,\sigma^2)$?

```
Call:
lm(formula = logcite ~ covariates[, 2])

Residuals:
     Min       1Q   Median       3Q      Max 
-2.87901 -0.47427  0.03839  0.81261  3.11374 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)       2.0279     0.3170   6.397 4.62e-07 ***
covariates[, 2]   0.8511     0.4349   1.957   0.0597 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.228 on 30 degrees of freedom
Multiple R-squared:  0.1132,	Adjusted R-squared:  0.08364 
F-statistic: 3.829 on 1 and 30 DF,  p-value: 0.05973
```

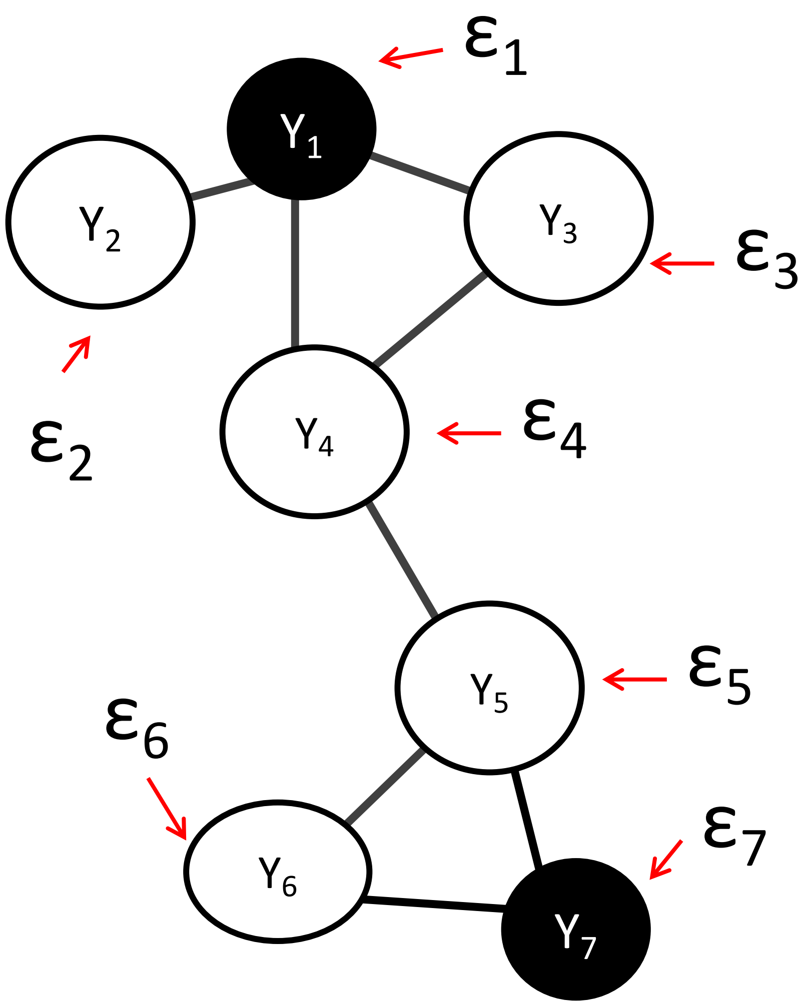


Independent error terms?



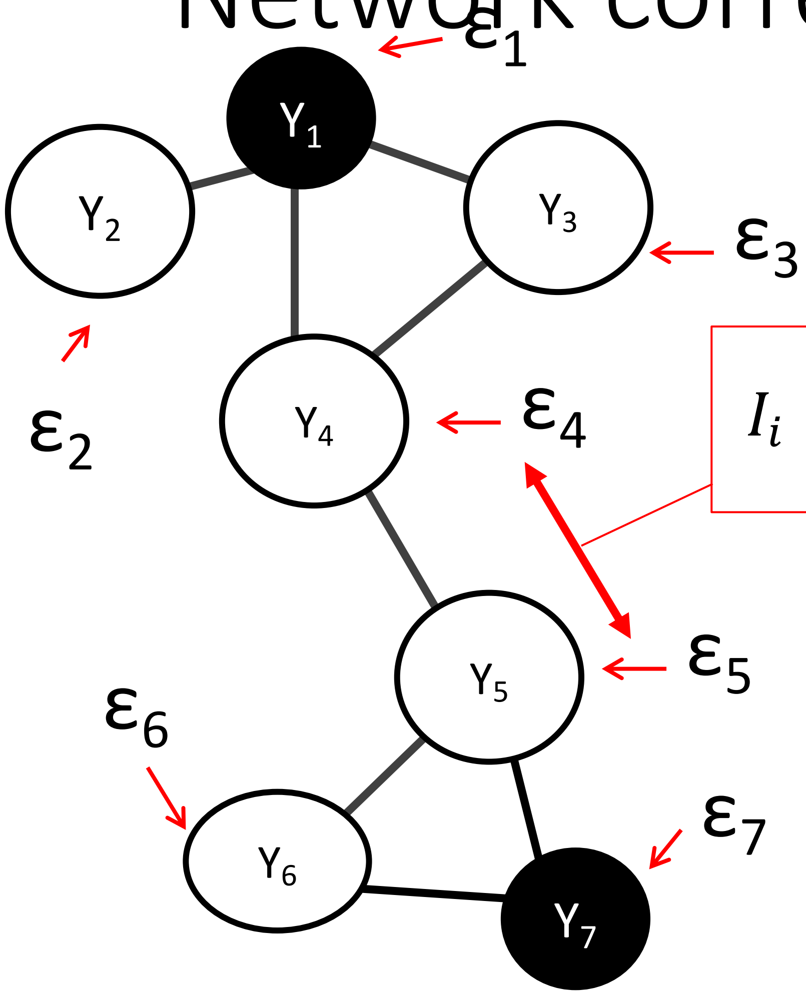
$\mathrm{cor}(\varepsilon_5 , \varepsilon_2) =?$

$\mathrm{cor}(\varepsilon_5 , \varepsilon_7) =?$



ε_i INDEPENDENT for all i ?

Network correlation – Moran’s I



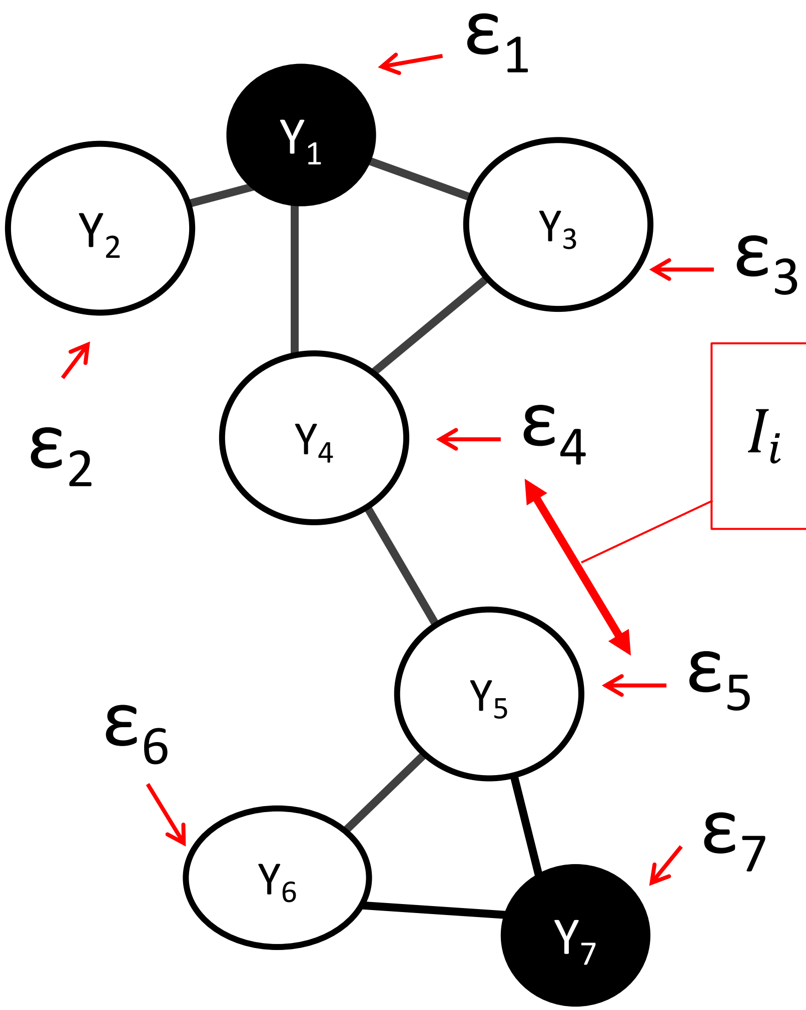
cov(ε_j , ε_k) but only for X_{jk} = 1

$$I_i = \frac{n\sum_{j=1}^{n}\sum_{k=1}^{n}(y_j - \bar{y})(y_k - \bar{y})A_{ijk}}{\sum_{j,k} A_{ijk}\sum_{j=1}^{n} y_j^2}$$

We test directly if there are departures from independence

Alternative:

$$C_i = \frac{(n - 10\sum_{j=1}^{n}\sum_{k=1}^{n}(y_j - y_k)^2 A_{ijk}}{2\sum_{j,k} A_{ijk}\sum_{j=1}^{n}(y_j - \bar{y})^2}$$



$\boxed{cov(\varepsilon_j\,, \varepsilon_k\,) \text{ but only for } X_{jk} = 1}$

$$\boxed{I_i = \frac{n\sum_{j=1}^{n}\sum_{k=1}^{n}(y_j - \bar{y})(y_k - \bar{y})A_{ijk}}{\sum_{j,k} A_{ijk}\sum_{j=1}^{n} y_j^2}}$$

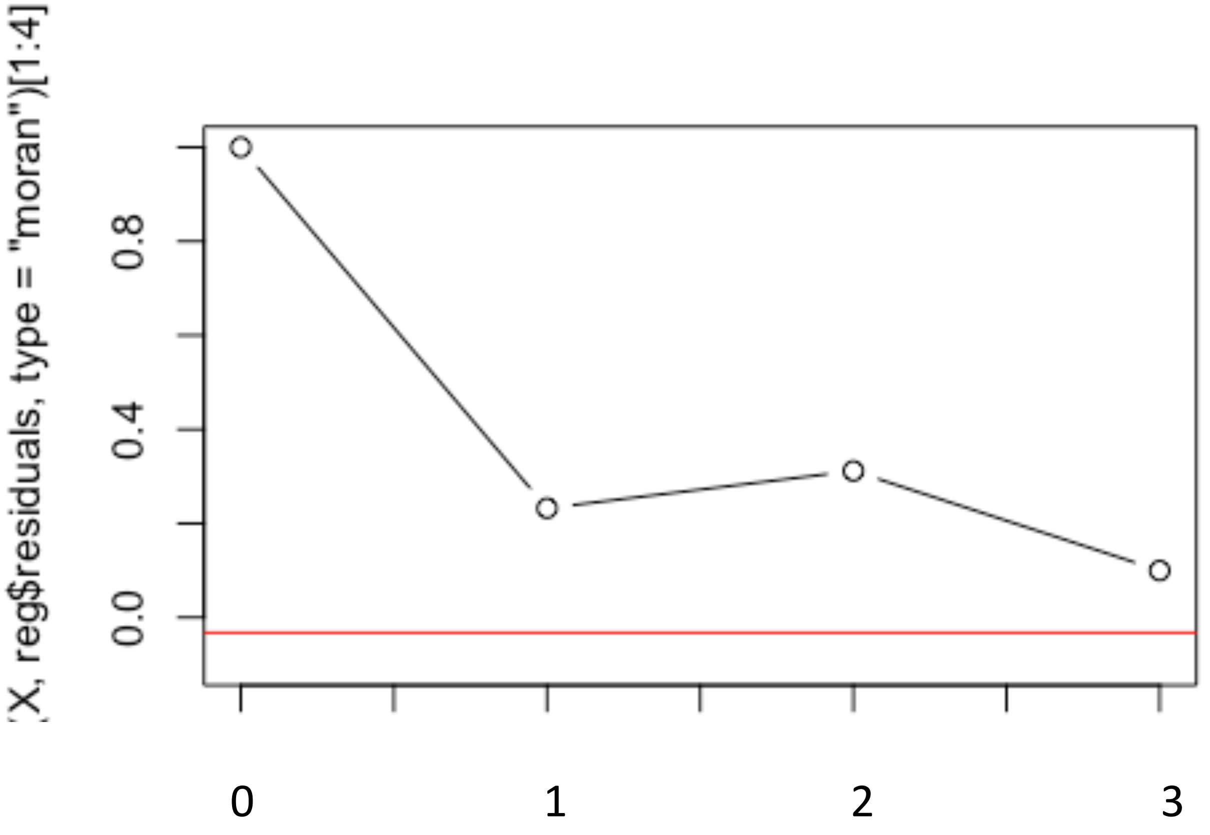
If I_i >0 what does it mean?

Residual for connected nodes

$$y_j - \hat{y}_j > 0$$

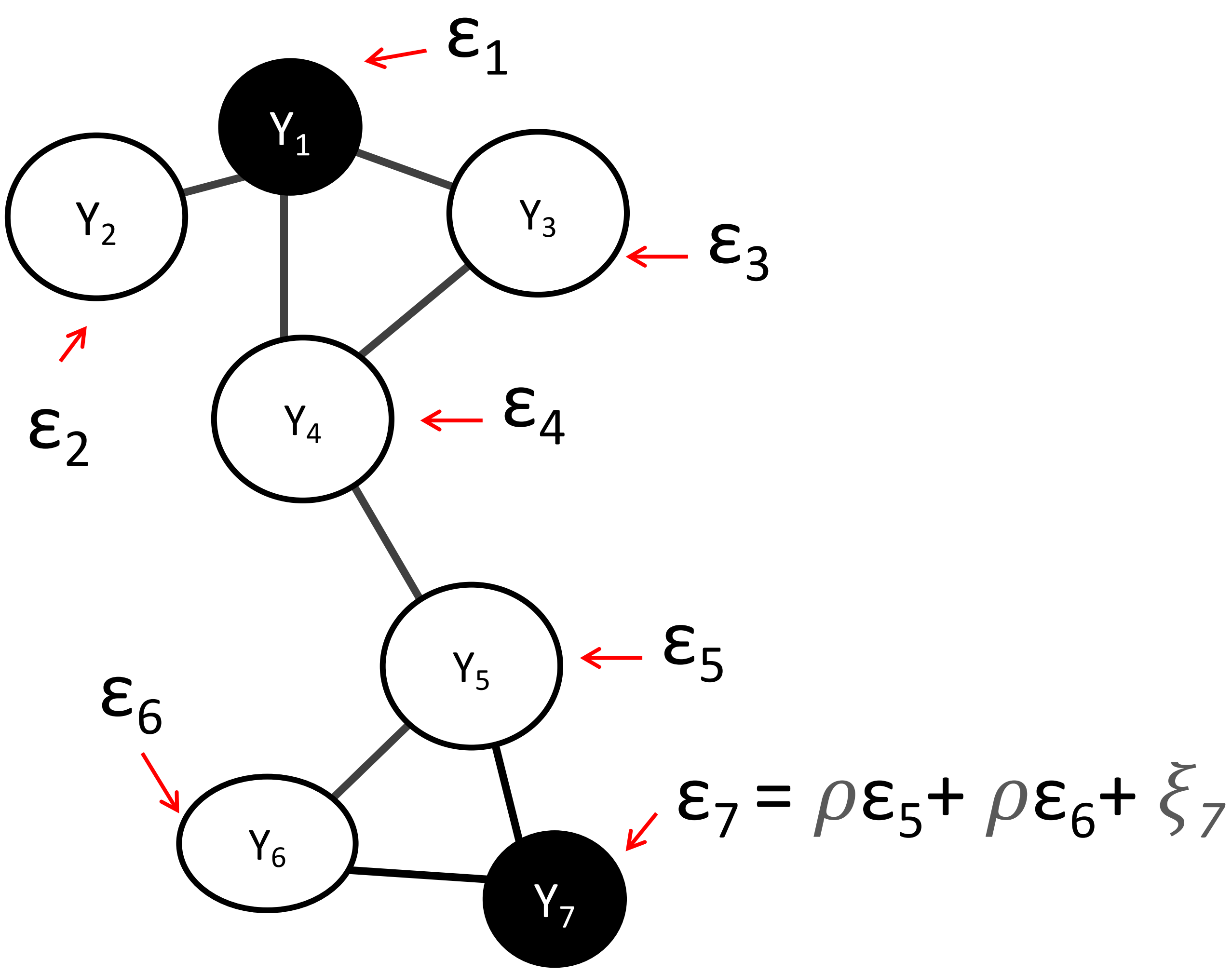
$$y_k - \hat{y}_k > 0$$

Are errors uncorrelated?




```r
plot(nacf(X,reg$residuals,type="moran")[1:4],
     type='b',ylim=c(-0.1,1))
abline(h=-1/(n-1),col='red')
```

Network auto-correlation model



$$Y_i = \alpha + \beta M_i + \varepsilon_i$$

where

$$\varepsilon_i = \rho \sum_j X_{ij}\varepsilon_j + \xi_i$$

and

$$\xi_i \sim N(0,\sigma^2)$$

INDEPENDENT for all *i*

Network auto-correlation model

```
Call:
lnam(y = logcite, x = covariates, W2 = X)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6925 -0.5811  0.1828  0.9024  3.0069

Coefficients:
            Estimate Std. Error Z value Pr(>|z|)
cons         2.13474    0.33909   6.296 3.06e-10 ***
sociologist  0.55778    0.47269   1.180   0.2380
rho2.1       0.10050    0.05868   1.713   0.0868 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      Estimate Std. Error
Sigma     1.13      0.021
```

$$Y_i = \alpha + \beta M_i + \varepsilon_i$$

where

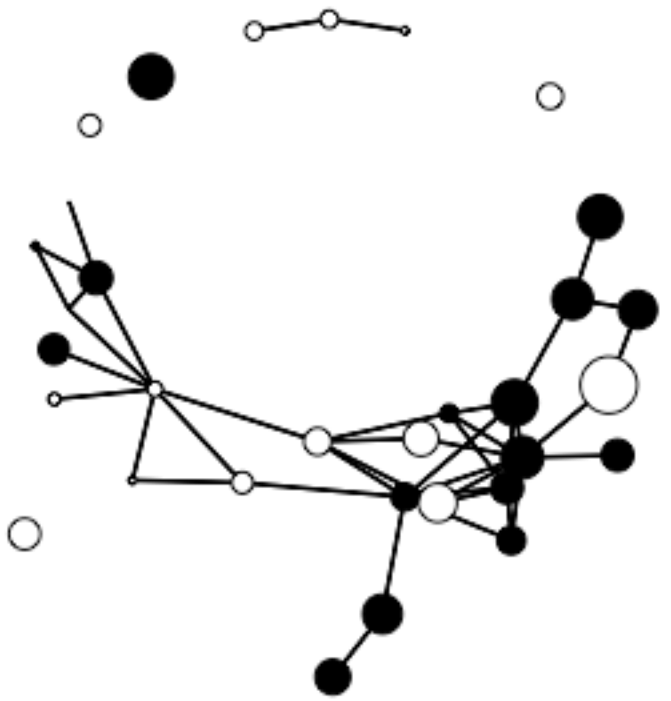
$$\varepsilon_i = \rho \sum_j X_{ij} \varepsilon_j + \xi_i$$

and

$$\xi_i \sim \mathrm{N}(0, \sigma^2)$$

INDEPENDENT for all *i*

Weight matrix



Since

$$\varepsilon_i = \rho \sum_j X_{ij} \varepsilon_j + \xi_i$$

Higher degree nodes
'get more errors'
What about scaling ties

$$\varepsilon_i = \rho \frac{\sum_j X_{ij} \varepsilon_j}{\sum_j X_{ij}} + \xi_i$$

Weight matrix

```
Call:
lnam(y = logcite, x = covariates, W2 = W)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6855 -0.7442  0.1371  0.8724  2.8438

Coefficients:
            Estimate Std. Error Z value Pr(>|z|)
cons          2.2978     0.3810   6.032 1.62e-09 ***
sociologist   0.3877     0.4640   0.836  0.40341
rho2.1        0.9271     0.3596   2.578  0.00993 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

      Estimate Std. Error
Sigma    1.066      0.019
```

$$\sum_j W_{ij} = 1 \text{ (or 0)}$$

Since

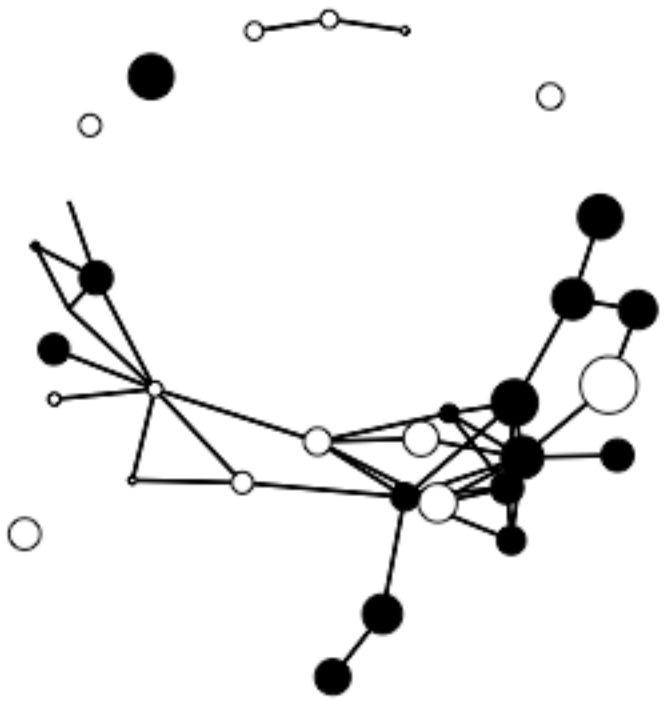
$$\varepsilon_i = \rho \sum_j X_{ij}\varepsilon_j + \xi_i$$

Higher degree nodes
‘get more errors’
What about scaling ties

$$\varepsilon_i = \rho \frac{\sum_j X_{ij}\varepsilon_j}{\sum_j X_{ij}} + \xi_i$$

$$= \rho \sum_j W_{ij}\varepsilon_j + \xi_i$$

Interpretation



Researchers tend to be cited more than average if their collaborators are cited more than average

Correlation through errors mean that we do not predict with the network only explaining unobserved heterogenity

Network effects model

What if there was a process of social influence

$$Y_{i,t} = \rho \sum_j X_{ij} \, Y_{j,t-1} \, \alpha + \beta M_i + \varepsilon_{i,t}$$

Citation at time *t*

Citation of all others at time *t*-1

Network effects model

What if there was a process of social influence

$$Y_{i,t} = \rho \sum_j X_{ij} Y_{j,t-1} + \alpha + \beta M_i + \varepsilon_{i,t}$$

Citation at time *t*

Citation of all others at time *t*-1

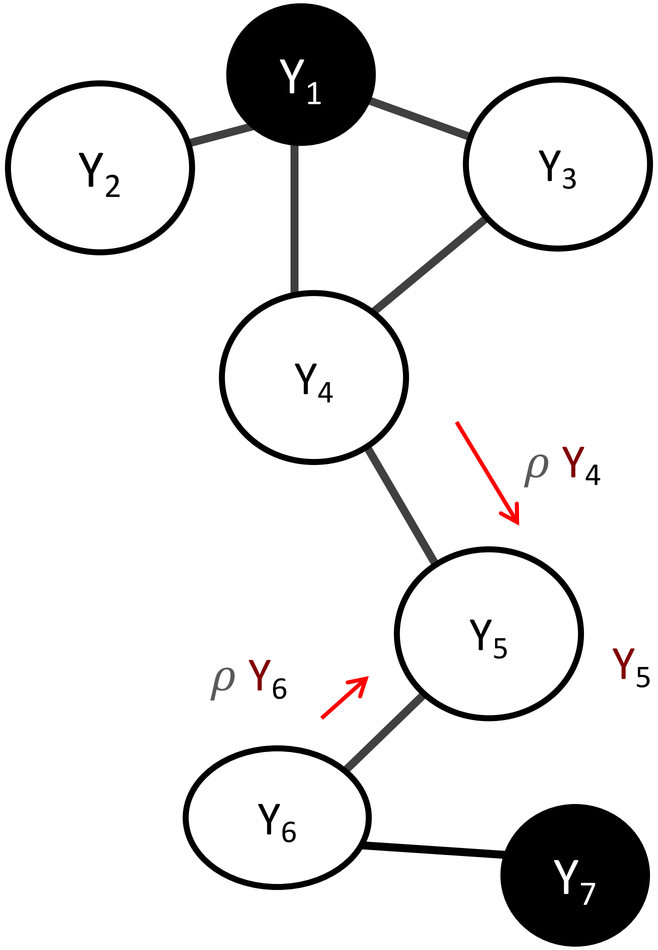
The equilibrium (when t tends to infinity):

$$Y_i = \rho \sum_j X_{ij} Y_j + \alpha + \beta M_i + \varepsilon_i$$

Citations

Citation of all others

Network effects model



$Y_5 = \rho\ Y_4 + \rho\ Y_6 + \beta M_5 + \varepsilon_5$

The network effect model (element-wise):

$$Y_i = \rho \sum_j X_{ij}\, Y_j \ + \alpha + \beta M_i + \varepsilon_i$$

Citations

Citation of all others

The network effect model (simulataneous):

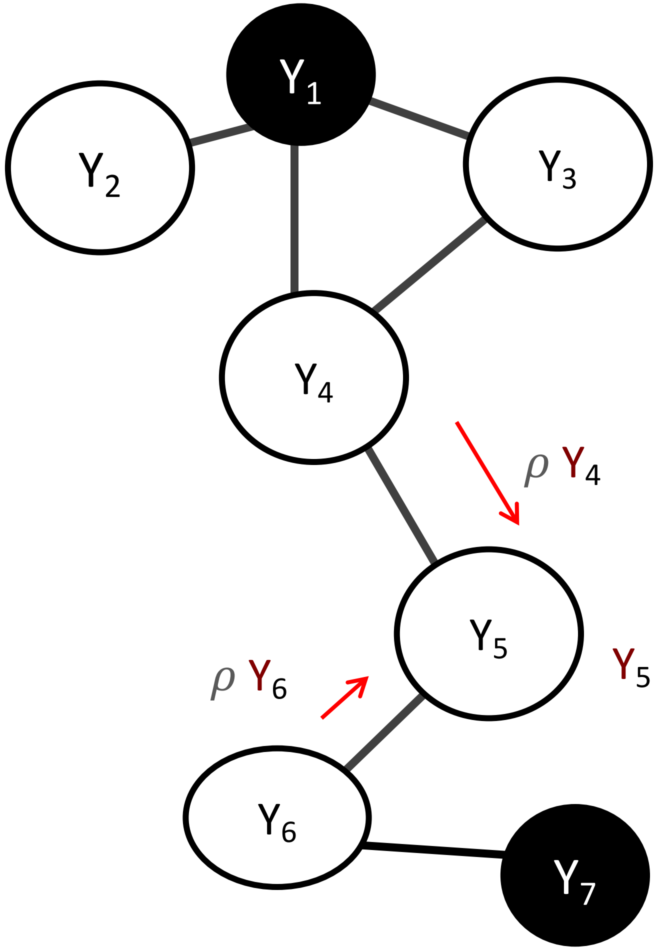
$$Y = \rho X Y \ + \alpha + M\beta + \varepsilon$$

Vector of citations

Citation of all others

Network effects model

The network effect model (simultaneous):



$Y_5 = \rho\ Y_4 + \rho\ Y_6 + \beta M_5 + \varepsilon_5$

$$Y = \rho XY + \alpha + M\beta + \varepsilon$$

```
Call:
lnam(y = logcite, x = covariates, W1 = X)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.6665 -0.8386  0.1060  0.8026  3.0625 

Coefficients:
            Estimate Std. Error Z value Pr(>|z|)    
cons         1.87066    0.33036   5.663 1.49e-08 ***
sociologist  0.69711    0.43357   1.608    0.108    
rho1.1       0.03397    0.02969   1.144    0.253    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

      Estimate Std. Error
Sigma    1.163      0.021
```

Weight matrix

```
Call:
lnam(y = logcite, x = covariates, W1 = W)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.8873 -0.5815  0.1188  0.8561  2.8750 

Coefficients:
            Estimate Std. Error Z value Pr(>|z|)    
cons          1.7051     0.3926   4.343  1.4e-05 ***
sociologist   0.6244     0.4465   1.398    0.162    
rho1.1        0.3945     0.3124   1.263    0.207    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

      Estimate Std. Error
Sigma    1.154      0.021
```

Is everyone equally important?

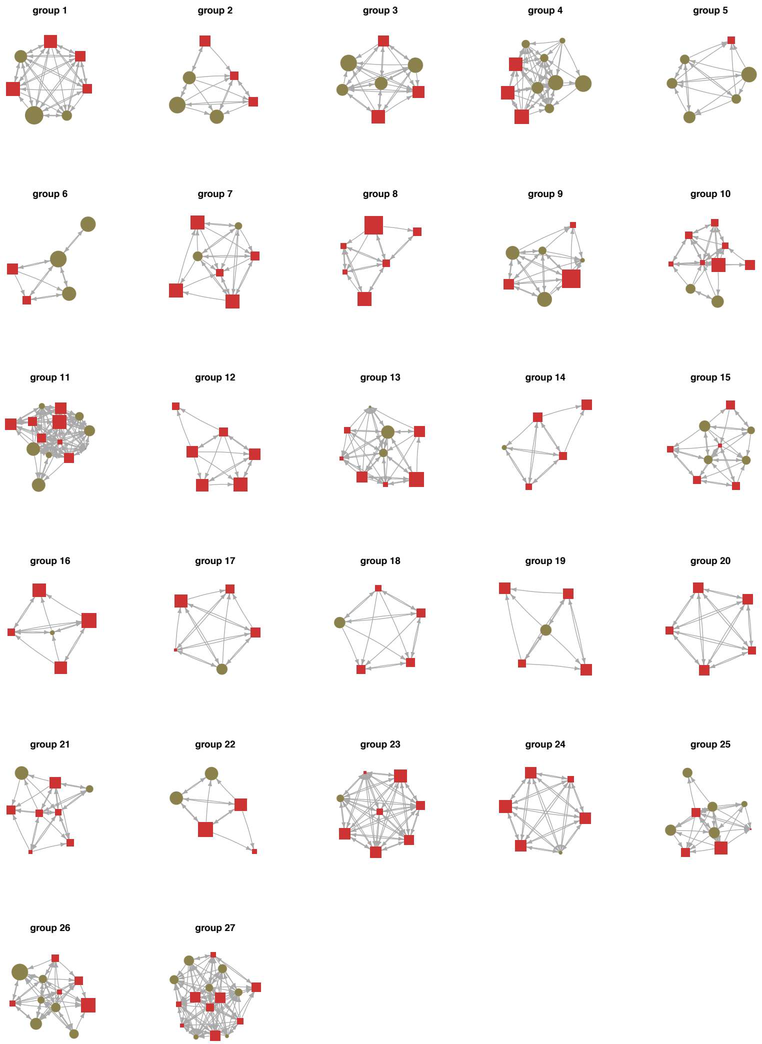
“I have two friends and they both like the government”

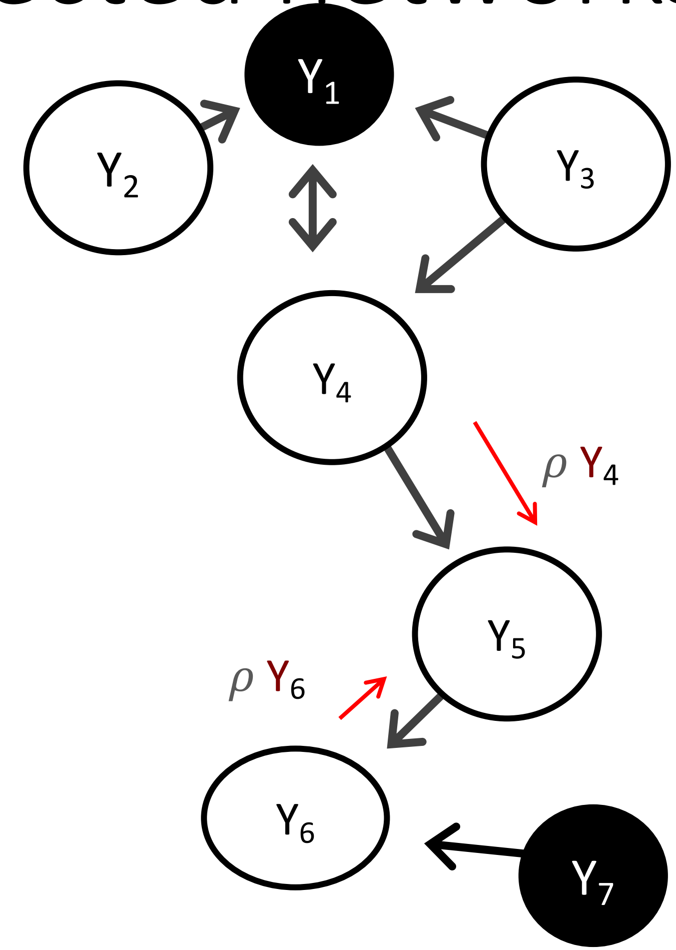
“I have two fifty friends and two like the government”

$$Y_i = \rho \frac{\sum_j X_{ij} Y_j}{\sum_j X_{ij}} + \varepsilon_i$$

$$Y_i = \rho \sum_j W_{ij} Y_j + \varepsilon_i$$

Network Effects model for directed networks





In-ties $Y_5 = \rho\ Y_4 + \beta M_5 + \varepsilon_5$

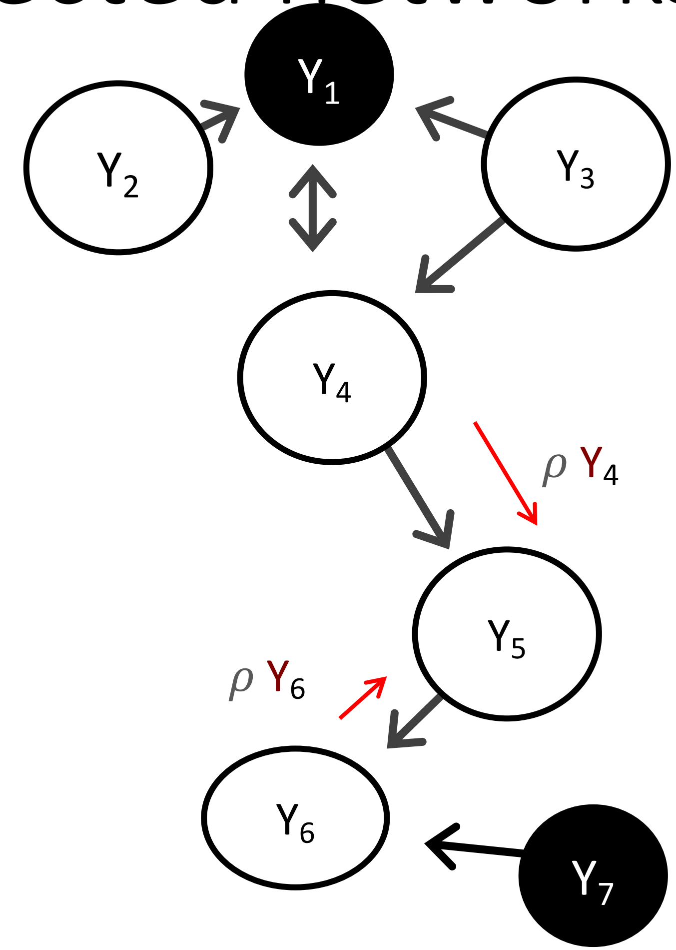
Out-ties $Y_5 = \rho\ Y_6 + \beta M_5 + \varepsilon_5$

Network Effects model for directed networks

Consider the differences between:

'I become similar to the people I like'

'I influence the people that I chose'



In-ties $Y_5 = \rho\ Y_4 + \beta M_5 + \varepsilon_5$

Out-ties $Y_5 = \rho\ Y_6 + \beta M_5 + \varepsilon_5$

… but what about selection