Selection and influence:

Exponential random graph models for social selection

- Social influence and social selection
- Review exponential random graph models
  - The conditional form of the model: interpretation
- Models incorporating social selection effects
  - Binary attributes
  - Continuous attributes
- Examples of estimation from pnet.
- Conclusions

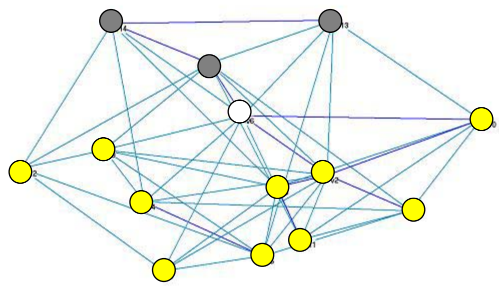
Social influence

- Given a social network, actors are "influenced" by their network partners.
  - a process of attribute change
  - applications in areas such as information spread, diffusion of innovations, disease
- Two possible mechanisms
  - social proximity
  - structural equivalence (Burt, 1987; Shah, 1998)
- Models for social influence
  - Network effects model: Friedkin (1998)
  - Exponential random graph version: Robins, Pattison & Elliott (2001)

Social influence in an organisational training group

Directed course interaction network

Grey = “disliked aspects of training course”



White = neutral

Yellow = “liked training course”

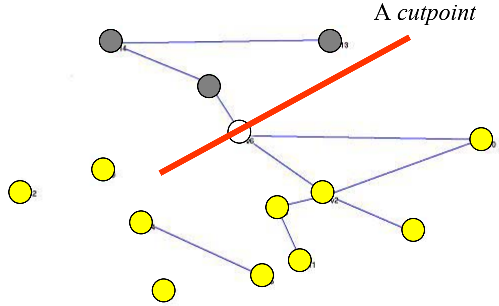
Robins, Pattison & Elliott, 2001

Social influence in an organisational training group

Mutual course interaction network

Grey = “disliked aspects of training course”

A *cutpoint*



White = neutral

Yellow = “liked training course”

Robins, Pattison & Elliott, 2001

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.

Review exponential random graph (p*) 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 from which the network is hypothesised to be built

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

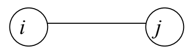
λ_Q parameter for the presence of Q

κ is a normalizing quantity.

Bernoulli models

Possible edges are independent of one another.

Configurations in this model relate to single possible edges (x_{ij}).



homogeneity assumption

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left(\sum \lambda_{ij} x_{ij}\right) = \frac{1}{\kappa}\exp\left(\theta \sum x_{ij}\right) = \frac{1}{\kappa}\exp\left(\theta L\right)$$

where θ is an edge parameter and L is the number of edges in the observed graph.

Bernoulli models: 2-blocks

If there are categorical attributes (e.g. sex), then relaxing the homogeneity assumption leads to a model akin to an a priori blockmodel.

Homogeneity assumptions (non-directed):

$\lambda_{ij} = \theta_{bb}$ if i,j are both male
$\lambda_{ij} = \theta_{bg}$ if one is male, the other female
$\lambda_{ij} = \theta_{gg}$ if i,j are both female

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left(\theta_{bb}L_{bb} + \theta_{bg}L_{bg} + \theta_{gg}L_{gg}\right)$$

where L_{bb} is the number of edges among males, etc.
See Robins, Pattison, Kalish & Lusher (2005) for an example.

The conditional form of the model e.g. for Bernoulli graphs

$$\Pr(\mathbf{X}=\mathbf{x})=\frac{1}{\kappa}\exp\left(\theta L\right)$$

$$\Pr(X_{ij}=1,"rest")=\frac{1}{\kappa}\exp\left(\theta\{1+L_{"rest"}\}\right)=\frac{1}{\kappa}\exp\left(\theta\right)\exp\left(\theta L_{"rest"}\right)$$

where "rest" signifies the rest of the graph.

$$\Pr(X_{ij}=0,"rest")=\frac{1}{\kappa}\exp\left(\theta L_{"rest"}\right)$$

$$\frac{\Pr(X_{ij}=1\,|\,"rest")}{\Pr(X_{ij}=0\,|\,"rest")}=\frac{\exp\left(\theta\right)\exp\left(\theta L_{"rest"}\right)}{\exp\left(\theta L_{"rest"}\right)}=\exp\left(\theta\right)$$

$$\log\frac{\Pr(X_{ij}=1\,|\,"rest")}{\Pr(X_{ij}=0\,|\,"rest")}=\theta$$ conditional log-odds

The conditional form of the model e.g. for Bernoulli graphs

$$\Pr(\mathbf{X}=\mathbf{x})=\frac{1}{\kappa}\exp\left(\theta L\right)$$

$$\log\frac{\Pr(X_{ij}=1\,|\,"rest")}{\Pr(X_{ij}=0\,|\,"rest")}=\theta$$ conditional log-odds

$$\log\frac{\Pr(X_{ij}=1\,|\,"rest")}{\Pr(X_{ij}=0\,|\,"rest")}=\theta$$ conditional log-odds

Conditional form for Markov graphs

e.g. edge/2-star/triangle model

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left(\theta L + \sigma_2 S_2 + \tau T\right)$$

conditional log-odds

$$\log \frac{\Pr(X_{ij}=1 \mid \text{"}rest\text{"})}{\Pr(X_{ij}=0 \mid \text{"}rest\text{"})} = \theta + \sigma_2 \Delta_{ij}(S_2) + \tau \Delta_{ij}(T)$$

where $\Delta_{ij}(S_2)$ is the number of extra 2-stars created if a tie on *ij* was created; and $\Delta_{ij}(T)$ the extra triangles (etc).
- ***change statistics***

$$\log \frac{\Pr(X_{ij}=1 \mid \text{"}rest\text{"})}{\Pr(X_{ij}=0 \mid \text{"}rest\text{"})} = \theta + \Delta_{ij}$$

where Δ_{ij} are the change statistics for the structural effects in the model.

The conditional form of the model e.g. for Bernoulli graph/2 blocks

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left(\theta_{bb} L_{bb} + \theta_{bg} L_{bg} + \theta_{gg} L_{gg}\right)$$

Conditional log-odds without the two blocks

$$\log \frac{\Pr(X_{ij}=1 \mid \text{"}rest\text{"})}{\Pr(X_{ij}=0 \mid \text{"}rest\text{"})} = \theta$$

So, analogously, for the 2-block version of the model

θ_{bb} is the log odds of a tie between males

θ_{bg} is the log odds of a tie between a male and a female

θ_{gg} is the log odds of a tie between females

The conditional form of the model e.g. for Bernoulli graph/2-blocks

Suppose that $y_i=1$ if i is male and $y_i=0$ if i is female

Then for a tie between males $y_i y_j=1$ (else 0)
between male/female $y_i(1-y_j)+(1-y_i)y_j=1$ (else 0)
between females $(1-y_i)(1-y_j)=1$ (else 0)

So
θ_{bb} is the log odds of a tie between males
θ_{bg} is the log odds of a tie between a male and a female
θ_{gg} is the log odds of a tie between females
Can be expressed as

$$\theta_{bb} y_i y_j + \theta_{bg} y_i (1 - y_j) + \theta_{bg} y_j (1 - y_i) + \theta_{gg} (1 - y_i)(1 - y_j)$$

The conditional form of the model e.g. for Bernoulli graph/2-blocks

$$\Pr(\mathbf{X} = \mathbf{x}) = \frac{1}{\kappa} \exp\left(\theta_{bb} L_{bb} + \theta_{bg} L_{bg} + \theta_{gg} L_{gg}\right)$$

is the same model as

$$\log \frac{\Pr(X_{ij} = 1 \,|\, "rest")}{\Pr(X_{ij} = 0 \,|\, "rest")} =$$

$$= \theta_{bb} y_i y_j + \theta_{bg} y_i (1 - y_j) + \theta_{bg} y_j (1 - y_i) + \theta_{gg} (1 - y_i)(1 - y_j)$$

$$= R_b y_i y_j + R(y_i + y_j) + \theta$$

The conditional form of the model e.g. for Bernoulli graph/2-blocks

$$\Pr(\mathbf{X}=\mathbf{x}) = \frac{1}{\kappa}\exp\left(\theta_{bb}L_{bb} + \theta_{bg}L_{bg} + \theta_{gg}L_{gg}\right)$$

The log-odds of a tie

$$\log\frac{\Pr(X_{ij}=1\,|\,"rest")}{\Pr(X_{ij}=0\,|\,"rest")} = R_b y_i y_j + R(y_i + y_j) + \theta$$

Exponential random graph social selection models with binary attributes

The log-odds of a tie

$$\log\frac{\Pr(X_{ij}=1\,|\,"rest")}{\Pr(X_{ij}=0\,|\,"rest")} = R_b y_i y_j + R(y_i + y_j) + \theta + \Delta_{ij}$$

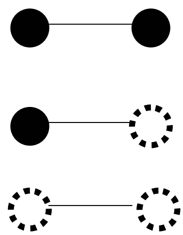
where Δ_{ij} is a change statistic for the structural effects in the model.

If y_i=y_j=1, log odds= R_b+2R+θ+Δ_{ij}
If i and j have different attributes
log odds= R+θ+Δ_{ij}
If y_i=y_j=0, log odds= θ+Δ_{ij}

©Department of Psychology, University of Melbourne

Exponential random graph social selection models with binary attributes

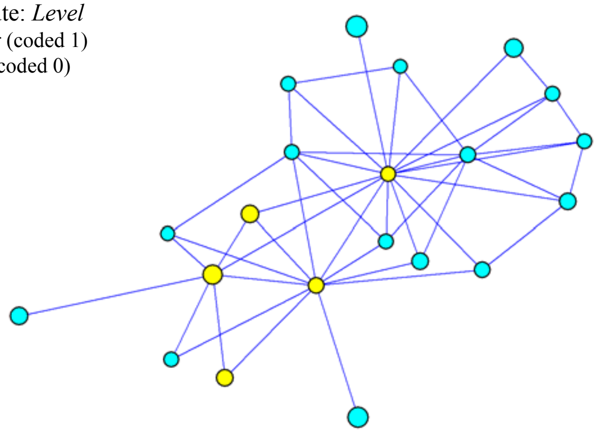
Edge Configurations



R_b		$x_{ij}y_iy_j$
R		$x_{ij}y_i$
θ		x_{ij}

Example: Krackhardt hi-tech managers: Mutual advice network

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



Example: Krackhardt hi-tech managers: Mutual advice network

Edge-based model for *level*
Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence (t-statistic)
Edge:	-1.96*	0.29	0.07
Rb	1.40	1.00	0.002
R	0.97*	0.40	0.02

Log-odds of mutual trust between juniors = – 1.96
between junior/senior = 0.97 – 1.96 = – 0.99
between seniors = 1.40 + 2×0.97 – 1.96 = 1.38

Exponential random graph social selection models with continuous attributes

y_i is a continuous scale, not binary
Sum and difference model

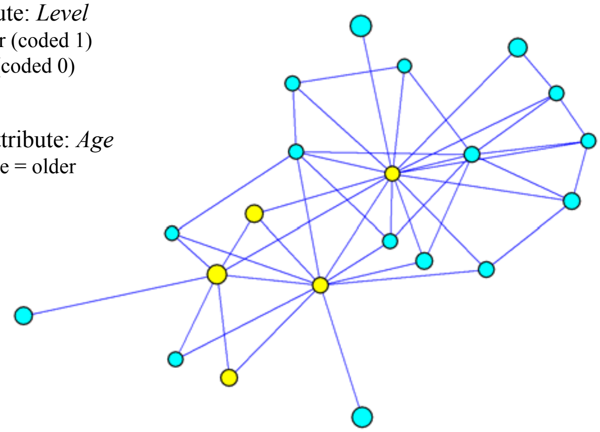
$$\log \frac{\Pr(X_{ij} = 1 \mid "rest")}{\Pr(X_{ij} = 0 \mid "rest")} = C_d \mid y_i - y_j \mid + C_s(y_i + y_j) + \theta + \Delta_{ij}$$

For positive Cs
log-odds increases as $y_i + y_j$ gets larger
For negative Cd
log-odds increases as $y_i - y_j$ gets closer to 0

Example: Krackhardt hi-tech managers: Mutual advice network

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

Continuous attribute: *Age*
Larger node size = older



Example: Krackhardt hi-tech managers: Mutual advice network

Edge-based model for *age*
Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence (t-statistic)
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.

Example: Krackhardt hi-tech managers: Mutual advice network

Edge-based model for *age*
Parameter estimates from pnet:

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

Log-odds of mutual trust = 0.85 – 0.022 ×(sum of ages) – 0.045 ×(difference in ages)

Average age = 39.7; Range is from 27 to 62

Log-odds of mutual trust between two managers of average age:
$= 0.85 - 0.022 \times 79.4 - 0.045 \times 0 = -0.90$

Log-odds of mutual trust between oldest and youngest managers:
$= 0.85 - 0.022 \times 89 - 0.045 \times 35 = -2.68$

Example: Krackhardt hi-tech managers: Mutual advice network

Edge-based model for both *age* and *level*
Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence (t-statistic)
Edge:	2.32	1.57	-0.004
Rb *level*	1.56	1.02	-0.02
R *level*	1.21*	0.42	-0.01
Sum *age*	-0.054*	0.023	0.002
Difference *age*	-0.031	0.035	0.03

Example: Krackhardt hi-tech managers: Mutual advice network

Edge-based model for both *age* and *level*
retaining only significant parameter effects from previous model

Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence (t-statistic)
Edge:	2.74*	1.29	0.02
R *level*	1.62*	0.32	0.06
Sum *age*	-0.065*	0.018	0.03

Log-odds of mutual trust
between juniors = 2.74 – 0.065 ×(sum of ages)
between junior/senior = 2.74 + 1.62 – 0.065 ×(sum of ages) = 4.36 – 0.065 ×(sum of ages)
between seniors = 2.74 + 2 × 1.62 – 0.065 ×(sum of ages) = 5.98 – 0.065 ×(sum of ages)
Average age of juniors = 39; average age of seniors = 43

Estimated log-odds of mutual trust between managers with average ages
between juniors = 2.74 – 0.065 ×78 = – 2.33
between junior/senior = 4.36 – 0.065 × 82 = – 0.97
between seniors = 5.98 – 0.065 × 86 = – 0.39

Example: Krackhardt hi-tech managers: Mutual advice network

Structural model without attributes

Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence (t-statistic)
Edge:	- 1.26	1.23	- 0.01
k-star	- 0.66	0.43	- 0.01
k-triangle	1.11*	0.31	-0.02

Example: Krackhardt hi-tech managers: Mutual advice network

Structural model with attribute effects

Parameter estimates from pnet:

Parameter	Estimate	Standard error	Convergence (t-statistic)
Edge:	6.28*	3.33	0.02
k-star	- 1.59*	0.71	0.02
k-triangle	0.94*	0.32	0.03
R *level*	1.41*	0.43	0.04
Sum *age*	-0.064*	0.023	0.03

References

Burt, R.S. (1987). Social contagion and innovation: cohesion versus structural equivalence. *American Journal of Sociology*, *92*, 205-211.

Friedkin, N. (1998). *A structural theory of social influence.* NY: Cambridge University Press.

McPherson M, Smith-Lovin, L., & Cook, J.M. (2001). Birds of a feather: homophily in social networks. *Annual Reiew of Sociology* , *27*, 415-444.

Robins, G.L., Elliott, P., & Pattison, P. (2001). Network models for social selection processes. *Social Networks, 23,* 1-30.

Robins, G.L., Pattison, P, & Elliott, P. (2001). Network models for social influence processes. *Psychometrika, 66,* 161-190.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2005). A workshop on exponential random graph (*p**) models for social networks. University of Melbourne.

Shah, P.P. (1998). Who are employees' social referents: Using a network perspective to determine referent others. *Academy of Management Journal, 41*, 249-268.