university of groningen / behavioural and social sciences / sociology

Statistical Analysis of Complete Social Networks
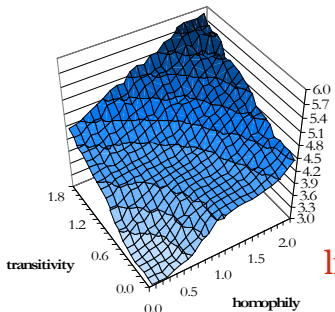
Model-based inference;
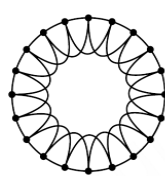*social relations model, p2 model*
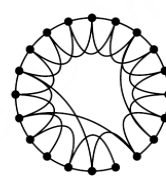
Christian Steglich   c.e.g.steglich@rug.nl
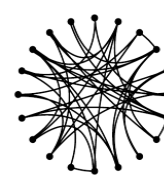
$$\ln(\frac{\Pr(x^c \rightarrow_i x^b)}{\Pr(x^c \rightarrow_i x^a)}) = \sum_{k=1}^{K} \beta_k \left(s_{ik}(x^b) - s_{ik}(x^a) \right)$$

**university of
groningen** / **behavioural and
social sciences** / **sociology**

Overview

› Stochastic modelling & model-based inference
› Dyad dependence models
- The p1 model
- The p2 model
- The social relations model

Basic framework for stochastic network models:

› It is assumed that <u>*networks are random variables*</u> (called X) with a (complex) probability distribution.

› An observed network (called x) is assumed to be drawn from the space of all possible networks according to this distribution.

The distribution...

› ...can be formulated in a model,

› ...can (at least) be simulated ("Markov Chain Monte Carlo"),

› ...can be used for hypothesis testing.

The network space is <u>huge</u>...

› For an undirected, binary ("zero-one") network among
n actors, how many networks are possible?

- For each dyad (i, j), there are **2** possibilities:
$x_{ij}=0$ or $x_{ij}=1$,
- There are $n \times (n-1)/2$ dyads ,
- Dyad outcomes can be combined in any way:
totality of $2^{n \times (n-1)/2}$.

n	1	2	3	4	5	...	10	...
# of networks	1	2	8	64	1024	...	~**35** trillion	...

State space for undirected networks with $n=4$ actors
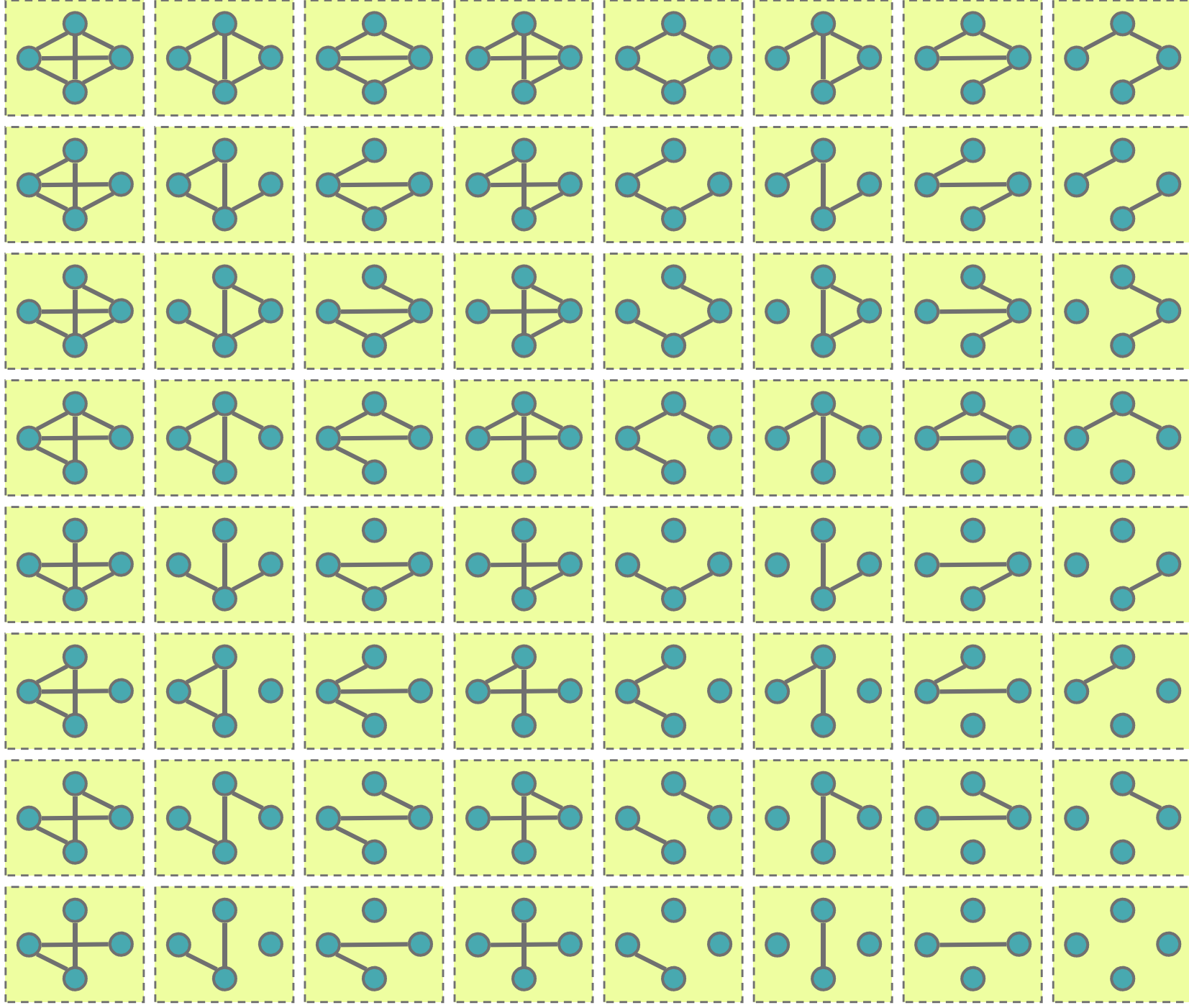
Once more independence:
The Erdős-Rényi (Bernoulli graph) model:

› Suppose all dyads are <u>*independent*</u>, and that a dyad (i, j) is connected with the probability p.

› Then the probability of any network x can be written as the product of the dyad probabilities (simple product rule holds for independent events).

› *Formally, we have* $\Pr(X = x) = p^{\#\text{ties}} \times (1 - p)^{\#\text{non-ties}}$, *where* $\#\text{non-ties} = (n \times (n - 1)/2) - \#\text{ties}$

The probability distribution on the network space

› …depends not on "structure" but only on tie counts! (see following slide)

› Now suppose that in a data collection, we observed the following particular network:
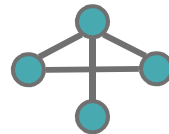
› Then the empirical tie probability is:

$$p = \#\text{ties}/(n{\times}(n{-}1)/2) = 2/3$$

The 'best-fitting' probability distribution on the network space is given on the following slide ... and has some problems:

- *Observed network is "lumped together" with other, non-equivalent networks,*

- *Highest probability has the full network, not the observed one...*

Probabilities under independence model with $p=2/3$

8.8%	4.4%	4.4%	4.4%	2.2%	2.2%	2.2%	0.5%
4.4%	2.2%	2.2%	2.2%	0.5%	0.5%	0.5%	0.3%
4.4%	2.2%	2.2%	2.2%	0.5%	0.5%	0.5%	0.3%
4.4%	2.2%	2.2%	2.2%	0.5%	0.5%	0.5%	0.3%
2.2%	0.5%	0.5%	0.5%	0.3%	0.3%	0.3%	0.1%
2.2%	0.5%	0.5%	0.5%	0.3%	0.3%	0.3%	0.1%
2.2%	0.5%	0.5%	0.5%	0.3%	0.3%	0.3%	0.1%
0.5%	0.3%	0.3%	0.3%	0.1%	0.1%	0.1%	0.1%

What about permutation-based distributions?

› Suppose again that in a data collection, we observed the same network:
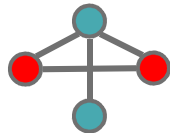
› For *n*=**4** actors, the number of permutations of these actors is **4!**=**4**×**3**×**2**×**1** = **24** , so there are **24** permuted networks

...of which each has a *structurally indistinguishable twin* because the actors marked **red** above are in fully equivalent positions,

...so **12** networks remain, they all have the same probability

Pr(*X*=*x*) =**1/12** ≈ **8.3%** while all other networks have Pr(*X*=*x*)=**0** .

› See next slide for how the best-fitting permutation-based distribution for this network looks like.

Probabilities under permutations of the actors

0	0	0	0	0	1/12	1/12	0
0	1/12	1/12	0	0	0	0	0
0	1/12	1/12	1/12	0	0	0	0
0	1/12	1/12	1/12	0	0	0	0
1/12	0	0	0	0	0	0	0
1/12	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

What to conclude for permutation-based distributions?

› They distinguish optimally between equivalent and non-equivalent structures ("isomorphic networks"),

› and do so better than the Bernoulli graph model (4-cycles are not treated identically to the example network),

› but do *only* this and nothing else – probabilities are zero for *all* non-isomorphic networks!

› This is a bad approach when considering *measurement error:*

  • small deviations between two networks are treated the same as huge differences! Error is *inflated* this way.

Better would be a model where <u>similar</u> networks have <u>similar</u> probabilities...

Dyad dependence models

Recall: A dyad is a pair of actors **‹i,j›** in the network, together with the configuration of tie variables **‹x_{ij}, x_{ji}›** between them.

Dependence *within* dyads:

- tie variable x_{ij} depends on tie variable x_{ji} (reciprocity)

Dependence *between* dyads:

- tie variable x_{ij} depends on tie variables x_{ik} (sender/ego effects)
- tie variable x_{ij} depends on tie variables x_{kj} (receiver/alter effects)

The p1 model *(Holland & Leinhardt, 1981, JASA)*

According to the probability distribution of the p1 model, the probability to observe a network x is this:

$$p_1(x) = Pr(X = x)$$

$$= \frac{1}{\text{const.}} \exp\left(\theta \sum_{ij} x_{ij} + \rho \sum_{ij} x_{ij} x_{ji} + \sum_{ij} \alpha_i x_{ij} + \sum_{ij} \beta_j x_{ij}\right)$$

A constant guaranteeing that the probability over all possible networks sums up to one.

Total number of ties. θ models the density of the network.

Number of reciprocal ties. ρ models the degree of reciprocity of the network.

...

Sender and receiver (fixed) effects for each actor

$$\frac{1}{\text{const.}}\exp\left(\theta\sum_{ij}x_{ij} + \rho\sum_{ij}x_{ij}x_{ji} + \sum_{ij}\alpha_i x_{ij} + \sum_{ij}\beta_j x_{ij}\right)$$

$$\sum_i \alpha_i \sum_j x_{ij} = \sum_i \alpha_i x_{i+}$$

$$\sum_j \beta_j \sum_i x_{ij} = \sum_j \beta_j x_{+j}$$

...

Number of outgoing ties of actor i. α_i models the outdegree of actor i (*expansiveness*).

Number of incoming ties of actor j. β_j models the indegree of actor j (*attractiveness*).

The α and β parameters should each sum up to zero.

They are many! For n actors, 2n such parameters in the model ...

The p1 model is a conditional dyad independence model

To see this, reformulate probability function:

$$p_1(x) = \frac{1}{\text{const.}} \exp\left(\theta \sum_{ij} x_{ij} + \rho \sum_{ij} x_{ij} x_{ji} + \sum_{ij} \alpha_i x_{ij} + \sum_{ij} \beta_j x_{ij}\right)$$

$$= \frac{1}{\text{const.}} \exp\left(\sum_{ij} x_{ij}\left(\theta + \rho x_{ji} + \alpha_i + \beta_j\right)\right)$$

$$= \frac{1}{\text{const.}} \prod_{ij} \exp\left(x_{ij}\left(\theta + \rho x_{ji} + \alpha_i + \beta_j\right)\right)$$

$$= \prod_{i<j} \Pr\left(\left\langle X_{ij}, X_{ji}\right\rangle = \left\langle x_{ij}, x_{ji}\right\rangle\right)$$

Total probability is product of dyad probabilities…

The p1 model in dyadic notation

The p1 probability for a dyad ‹x_{ij}, x_{ji}› is this:

$$\frac{1}{\text{const.}}\exp\left(x_{ij}(\theta+\alpha_i+\beta_j)+x_{ji}(\theta+\alpha_j+\beta_i)+x_{ij}x_{ji}\rho\right)$$

with the norming constant now being a (dyad-specific) different one than before.

The p1 model can be estimated in UCINET or S†OCNET.
See following slides for illustrative results.

```
P1 /…excerpts of output…/
----------------------------------------------------------------
Input dataset:  campnet (C:\...\Ucinet 6\DataFiles\campnet)

Theta = -3.3380
Rho   =  5.4418
```

Estimated coefficients …

```
Expansiveness and Popularity Parameters


                   1      2                             1      2
                Alpha   Beta                         Alpha   Beta
                ------ ------                         ------ ------
   1    HOLLY  -0.623  0.607         10    BILL      1.602
   2  BRAZEY   1.219 -2.405          11     LEE      0.036 -0.218
   3   CAROL   0.703 -1.217          12     DON     -0.623  0.607
   4     PAM  -1.140  1.230          13    JOHN      1.602
   5     PAT  -0.623  0.607          14   HARRY      0.036 -0.218
   6  JENNIE   0.036 -0.218          15    GERY      0.703 -1.217
   7 PAULINE  -0.623  0.607          16   STEVE     -1.140  1.230
   8     ANN   0.703 -1.217          17    BERT     -0.623  0.607
   9 MICHAEL  -0.623  0.607          18    RUSS     -0.623  0.607
```

```
P1 Expected Values

              1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18
           HOLL BRAZ CARO  PAM  PAT JENN PAUL  ANN MICH BILL  LEE  DON JOHN HARR GERY STEV BERT RUSS
           ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
  1  HOLLY  0.00 0.07 0.14 0.27 0.23 0.19 0.23 0.14 0.23 0.00 0.19 0.23 0.00 0.19 0.14 0.27 0.23 0.23
  2  BRAZEY 0.24 0.00 0.08 0.34 0.24 0.15 0.24 0.08 0.24 0.00 0.15 0.24 0.00 0.15 0.08 0.34 0.24 0.24
  3  CAROL  0.23 0.05 0.00 0.30 0.23 0.17 0.23 0.11 0.23 0.00 0.17 0.23 0.00 0.17 0.11 0.30 0.23 0.23
  4  PAM    0.24 0.07 0.13 0.00 0.24 0.20 0.24 0.13 0.24 0.00 0.20 0.24 0.00 0.20 0.13 0.27 0.24 0.24
  5  PAT    0.23 0.07 0.14 0.27 0.00 0.19 0.23 0.14 0.23 0.00 0.19 0.23 0.00 0.19 0.14 0.27 0.23 0.23
  6  JENNIE 0.23 0.07 0.13 0.28 0.23 0.00 0.23 0.13 0.23 0.00 0.18 0.23 0.00 0.18 0.13 0.28 0.23 0.23
  7  PAULINE 0.23 0.07 0.14 0.27 0.23 0.19 0.00 0.14 0.23 0.00 0.19 0.23 0.00 0.19 0.14 0.27 0.23 0.23
  8  ANN    0.23 0.05 0.11 0.30 0.23 0.17 0.23 0.00 0.23 0.00 0.17 0.23 0.00 0.17 0.11 0.30 0.23 0.23
  9  MICHAEL 0.23 0.07 0.14 0.27 0.23 0.19 0.23 0.14 0.00 0.00 0.19 0.23 0.00 0.19 0.14 0.27 0.23 0.23
 10  BILL   0.24 0.02 0.05 0.38 0.24 0.12 0.24 0.05 0.24 0.00 0.12 0.24 0.00 0.12 0.05 0.38 0.24 0.24
 11  LEE    0.23 0.07 0.13 0.28 0.23 0.18 0.23 0.13 0.23 0.00 0.00 0.23 0.00 0.18 0.13 0.28 0.23 0.23
 12  DON    0.23 0.07 0.14 0.27 0.23 0.19 0.23 0.14 0.23 0.00 0.19 0.00 0.00 0.19 0.14 0.27 0.23 0.23
 13  JOHN   0.24 0.02 0.05 0.38 0.24 0.12 0.24 0.05 0.24 0.00 0.12 0.24 0.00 0.12 0.05 0.38 0.24 0.24
 14  HARRY  0.23 0.07 0.13 0.28 0.23 0.18 0.23 0.13 0.23 0.00 0.18 0.23 0.00 0.00 0.13 0.28 0.23 0.23
 15  GERY   0.23 0.05 0.11 0.30 0.23 0.17 0.23 0.11 0.23 0.00 0.17 0.23 0.00 0.17 0.00 0.30 0.23 0.23
 16  STEVE  0.24 0.07 0.13 0.27 0.24 0.20 0.24 0.13 0.24 0.00 0.20 0.24 0.00 0.20 0.13 0.00 0.24 0.24
 17  BERT   0.23 0.07 0.14 0.27 0.23 0.19 0.23 0.14 0.23 0.00 0.19 0.23 0.00 0.19 0.14 0.27 0.00 0.23
 18  RUSS   0.23 0.07 0.14 0.27 0.23 0.19 0.23 0.14 0.23 0.00 0.19 0.23 0.00 0.19 0.14 0.27 0.23 0.00
```

... model-derived (predicted) tie probabilities ...

RESIDUALS

```
              1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18
           HOLL BRAZ CARO  PAM  PAT JENN PAUL  ANN MICH BILL  LEE  DON JOHN HARR GERY STEV BERT RUSS
           ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
  1   HOLLY  0.00-0.07-0.14 0.73 0.77-0.19-0.23-0.14-0.23 0.00-0.19 0.77 0.00-0.19-0.14-0.27-0.23-0.23
  2  BRAZEY -0.24 0.00-0.08-0.34-0.24-0.15-0.24-0.08-0.24 0.00 0.85-0.24 0.00-0.15-0.08 0.66 0.76-0.24
  3   CAROL -0.23-0.05 0.00 0.70 0.77-0.17 0.77-0.11-0.23 0.00-0.17-0.23 0.00-0.17-0.11-0.30-0.23-0.23
  4     PAM -0.24-0.07-0.13 0.00-0.24 0.80 0.76 0.87-0.24 0.00-0.20-0.24 0.00-0.20-0.13-0.27-0.24-0.24
  5     PAT  0.77-0.07 0.86-0.27 0.00 0.81-0.23-0.14-0.23 0.00-0.19-0.23 0.00-0.19-0.14-0.27-0.23-0.23
  6  JENNIE -0.23-0.07-0.13 0.72 0.77 0.00-0.23 0.87-0.23 0.00-0.18-0.23 0.00-0.18-0.13-0.28-0.23-0.23
  7 PAULINE -0.23-0.07 0.86 0.73 0.77-0.19 0.00-0.14-0.23 0.00-0.19-0.23 0.00-0.19-0.14-0.27-0.23-0.23
  8     ANN -0.23-0.05-0.11 0.70-0.23 0.83 0.77 0.00-0.23 0.00-0.17-0.23 0.00-0.17-0.11-0.30-0.23-0.23
  9 MICHAEL  0.77-0.07-0.14-0.27-0.23-0.19-0.23-0.14 0.00 0.00-0.19 0.77 0.00 0.81-0.14-0.27-0.23-0.23
 10    BILL -0.24-0.02-0.05-0.38-0.24-0.12-0.24-0.05 0.76 0.00-0.12 0.76 0.00 0.88-0.05-0.38-0.24-0.24
 11     LEE -0.23 0.93-0.13-0.28-0.23-0.18-0.23-0.13-0.23 0.00 0.00-0.23 0.00-0.18-0.13 0.72 0.77-0.23
 12     DON  0.77-0.07-0.14-0.27-0.23-0.19-0.23-0.14 0.77 0.00-0.19 0.00 0.00 0.81-0.14-0.27-0.23-0.23
 13    JOHN -0.24-0.02-0.05-0.38-0.24-0.12 0.76-0.05-0.24 0.00-0.12-0.24 0.00-0.12 0.95-0.38-0.24 0.76
 14   HARRY  0.77-0.07-0.13-0.28-0.23-0.18-0.23-0.13 0.77 0.00-0.18 0.77 0.00 0.00-0.13-0.28-0.23-0.23
 15    GERY -0.23-0.05-0.11-0.30-0.23-0.17-0.23-0.11 0.77 0.00-0.17-0.23 0.00-0.17 0.00 0.70-0.23 0.77
 16   STEVE -0.24-0.07-0.13-0.27-0.24-0.20-0.24-0.13-0.24 0.00 0.80-0.24 0.00-0.20-0.13 0.00 0.76 0.76
 17    BERT -0.23-0.07-0.14-0.27-0.23-0.19-0.23-0.14-0.23 0.00 0.81-0.23 0.00-0.19-0.14 0.73 0.00 0.77
 18    RUSS -0.23-0.07-0.14-0.27-0.23-0.19-0.23-0.14-0.23 0.00-0.19-0.23 0.00-0.19 0.86 0.73 0.77 0.00
```

… & residuals (observed minus predicted).

The p2 model

A problem with the p1 model is its inflationary use of parameters. The p2 model...

- keeps a conditional dyad independence assumption,
- replaces the fixed effects for expansiveness and attractiveness for each actor by a regression equation plus random effects.

The probability for a dyad in the p2 model is still this:

$$\frac{1}{\text{const.}} \exp\left(x_{ij}(\theta + \alpha_i + \beta_j) + x_{ji}(\theta + \alpha_j + \beta_i) + x_{ij}x_{ji}\rho \right)$$

...but now the α and β parameters are regressed on a set of predictor variables! (& θ and ρ can be, too).

Regression equations for the p2 model (1)

"Sender effects" $$\alpha_i = \sum_k \gamma_k^S v_{ik}^S + \varepsilon_i^S$$

- Parameters γ^S measure effect of actor variables v^S on the *activity / expansiveness* of actors;
- Random parameters ε^S (expected value zero) model unexplained activity.

"Receiver effects" $$\beta_i = \sum_k \gamma_k^R v_{ik}^R + \varepsilon_i^R$$

- Parameters γ^R measure effect of actor variables v^R on the *popularity / attractiveness* of actors;
- Random parameters ε^R (expected value zero) model unexplained popularity.

Regression equations for the p2 model (2)

"Density effects"
$$\theta_{ij} = \theta + \sum_k \gamma_k^D w_{ijk}^D$$

- Parameters γ^D measure effect of dyadic variables w^D on the tie x_{ij} (here – unlike in the p1 model – allowed to differ between dyads);
- Intercept parameter θ models density.

"Reciprocity effects"
$$\rho_{ij} = \rho + \sum_k \gamma_k^M w_{ijk}^M$$

- Parameters γ^M measure effect of variables w^M on *reciprocity*;
- Intercept parameter ρ models reciprocity.

Assumptions about random terms (basic model)

$$\text{var}(\varepsilon_i^S) = \sigma_S^2$$

$$\text{var}(\varepsilon_i^R) = \sigma_R^2$$

$$\text{cov}(\varepsilon_i^S, \varepsilon_i^R) = \sigma_{SR}$$

$$\left.\begin{array}{l} \text{cov}(\varepsilon_i^S, \varepsilon_j^R) = 0 \\[2mm] \text{cov}(\varepsilon_i^S, \varepsilon_j^S) = 0 \\[2mm] \text{cov}(\varepsilon_i^R, \varepsilon_j^R) = 0 \end{array}\right\} \text{for } i \neq j$$

› Sender and receiver terms are *correlated* within actors, assumption: bivariate normal;

› Sender and receiver terms are *uncorrelated* between actors;

Hypothesis testing with p2 (1)

› Sender / ego parameters γ^S have (pos. or neg., depending on sign) effect on *outgoing* ties' probabilities;

These parameters can be used to test hypotheses about which type of actors nominate many / few others.

› Receiver / alter parameters γ^R have (pos. or neg.) effect on *incoming* ties' probabilities;

These parameters can be used to test hypotheses about which type of actors are nominated by many / few others.

Hypothesis testing with p2 (2)

› Density parameters γ^D have (pos. or neg.) *main effect* on tie probabilities;

These parameters can be used to test hypotheses about which types of actors select which other types of actors as network partners.

› Reciprocity parameters γ^R have (pos. or neg.) *interaction effect with reciprocity* on tie probabilities.

These parameters can be used to test hypotheses about which types of actors reciprocate ties to which other types of actors.

Hypothesis testing with p2 (3)

Single parameters γ can be tested for departure from zero by a Wald type test, making use of the parameter estimate divided by its standard error.

Assuming the null hypothesis H_o: γ=0 , the test statistic $t = \dfrac{\gamma}{s.e.(\gamma)}$ follows an approximate standard normal distribution. P-values can be calculated for alternative hypotheses H_A: γ<0 , H_A: γ>0 , H_A: γ≠0 .

This way of significance testing is quite common (linear regression, logistic regression, ERGM, Siena models, …).

Example for a p2 analysis: Lazega & van Duijn (1997)

Setting:

An intra-organizational advice network among
lawyers in a New England law firm.

Theory:

› *The advice getting network expresses 'partial equilibrium' of a
status competition process; advice is the resource competed for.*

› *Status (partner vs. associate) and seniority (levels 1 through 5)
signal resource possession and therefore should be the
dominant predictors of getting advice.*

› *Other predictors (office, specialty, sex, school) are secondary.*

Data description for Lazega & van Duijn (1997) study

Distribution of lawyers per variable

	Partner	Associate	Total
Seniority Level 1	14	7	
Seniority Level 2	13	10	
Seniority Level 3	9	5	
Seniority Level 4		7	
Seniority Level 5		6	
Total	36	35	71
Office 1	22	26	48
Office 2	13	6	19
Office 3	1	3	4
Specialty litigation	20	21	41
Specialty corporate	16	14	30
Men	33	20	53
Women	3	15	18
Lawschool Ivy League	12	3	15
Lawschool New-England non-Ivy League	11	17	28
Lawschool other	13	15	28

Example analysis (Lazega & van Duijn, 1997, Soc.Netw.)

P_2 estimates of all lawyers' choices (standard errors in brackets)

	Parameter	Empty model	Final model
Sender	Variance σ_A^2	0.58 (0.08)	0.75 (0.11)
	Partner seniority Level 1		−0.92 (0.30)
Receiver	Variance σ_B^2	0.76 (0.10)	0.49 (0.08)
	Associate seniority level		−0.50 (0.06)
Sender–receiver	Covariance σ_{AB}	−0.25 (0.07)	−0.05 (0.06)
Density	μ	−1.87 (0.12)	−3.98 (0.22)
	Similarity status		0.89 (0.22)
	Similarity seniority associate		0.98 (0.19)
	Superiority seniority		−0.29 (0.11)
	Similarity office		1.79 (0.11)
	Similarity specialty		1.60 (0.12)
	Similarity gender		0.29 (0.11)
	Similarity lawschool		0.20 (0.09)
Reciprocity	ρ	1.42 (0.13)	1.46 (0.25)
	Similarity specialty		−0.81 (0.28)

Exemplary hypothesis test (Lazega & van Duijn, 1997)

› The *superiority seniority* effect measures whether the potential advice seeker is more senior than the potential advice giver.

› The estimate is $\gamma = -0.29$ and the estimated standard error is $s.e.(\gamma) = 0.11$, indicating that *superiority seniority* makes actual advice giving less likely.

› The t-ratio for the Wald test becomes
$t = -0.29/0.11 = -2.64$.

› The corresponding left-sided p-value is $p = 0.004$.

› The null hypothesis $H_0: \gamma=0$ can (at conventional signifi-cance levels) be rejected against the alternative $H_A: \gamma<0$.

Estimation of the p2 model

› The norming constant in the p2 probability model cannot be calculated anymore (for p1, it could – see Holland & Leinhardt, 1981).

› MCMC techniques allow (after burn-in phase) to draw simulated networks from the distribution that a (preliminary) p2 model parameterisation implies.

› These draws are used to iteratively (by comparison to the actual data) improve estimates until convergence.

› Maximum likelihood & Bayesian procedures available.

More on p2

Multilevel extensions of the p2 model were developed by van Duijn, Snijders & Zijlstra (2004, Methodology):

› *Multiplex networks (several relations in the same group)*

› *Multiple groups (same relation for several groups)*

The p2 model can be estimated with StOCNET.

The p2 model is – like many network analysis methods – confined to binary data. A "valued ties" counterpart is David Kenny's *Social Relations Model…*

The social relations model

The social relations model assumes the dependent tie variables to be normally distributed. Their values are modelled directly, in a linear regression / ANOVA type framework:
$$x_{ij} = m + a_i + b_j + g_{ij}$$

m is the relational <u>mean</u> / intercept parameter

a models sender/ ego (here called <u>actor</u>) effects

b models receiver/ alter (here called <u>partner</u>) effects

g models dyad-specific (here <u>relationship</u>) effects

Random effects in the social relations model (1)

›   The a, b and g parameters can in principle be regressed
   on a set of predictor variables. They always are modelled
   with a random component.

- *Actor effects* $a_i = \sum_k \gamma_k^{act} v_{ik}^{act} + \varepsilon_i^{act}$
- *Partner effects* $b_i = \sum_k \gamma_k^{par} v_{ik}^{par} + \varepsilon_i^{par}$
- *Relationship effects* $g_{ij} = \sum_k \gamma_k^{rel} w_{ik}^{rel} + \varepsilon_{ij}^{rel}$

›   The basic model always includes six parameters: next to
   the intercept m the 5 (co)variance components…

Assumptions about random terms (basic model)

$$\text{var}(\varepsilon_i^{\text{act}}) = \sigma_{\text{act}}^2$$

$$\text{var}(\varepsilon_i^{\text{par}}) = \sigma_{\text{par}}^2$$

$$\text{var}(\varepsilon_{ij}^{\text{rel}}) = \sigma_{\text{rel}}^2$$

$$\text{cov}(\varepsilon_i^{\text{act}}, \varepsilon_i^{\text{par}}) = \sigma_{\text{act,par}}$$

$$\text{cov}(\varepsilon_{ij}^{\text{rel}}, \varepsilon_{ji}^{\text{rel}}) = \sigma_{\text{rel,rel}}$$

everything else is zero

› Actor and partner effects can be correlated within (not between) individuals

› Relationship effects can be correlated within (not between) dyads

› Reciprocity is – unlike in the p2 model – not modelled as separate regression equation, but as covariance.

Typical analysis (Snijders & Kenny, 1999, Pers. Rel.)

Table 1. *Parameter estimates with standard errors (SE) for the Warner et al. data set*

Parameter	Interpretation	Estimate	SE
μ	Constant term	50.8	2.7
Variance (A_i)	Actor variance	92.0	53.9
Variance (B_j)	Partner variance	40.9	27.8
Covariance (A_i, B_j)	Actor–partner covariance	−40.4	32.0
Variance (E_{ij})	Dyadic variance	78.4	18.2
Covariance (E_{ij}, E_{ji})	Within-dyad covariance	−27.8	18.2

In most applications, variance partitioning is reported,
not the actual estimates.

More complicated analysis from same paper

The social relations model is very often applied to data sets with many very small groups (e.g., families)

⇒ Multilevel extension, adding another layer of random effects

Table 2. Estimated effects and standard errors (SE) for recalled affection in families, for models without (Model 1) and with (Model 2) reciprocity effects

Effect	Model 1 Estimate	Model 1 SE	Model 2 Estimate	Model 2 SE
Fixed Effects				
Constant term (fixed effect) μ	27.62	0.54	27.60	0.54
Father as actor (fixed effect) β_{FA}	1.72	0.86	1.75	0.86
Mother as actor (fixed effect) β_{MA}	1.60	0.63	1.60	0.57
Father–mother difference as partners $\beta_{FP}-\beta_{MP}$	−1.69	0.37	−1.70	0.37
Random Effects				
Family variance (F_k)	0	*	0	*
Father as actor variance (A^F_{sk})	16.77	4.30	16.43	4.38
Mother as actor variance (A^M_{sk})	6.63	1.91	7.05	1.91
Child as actor variance (A^C_{sk})	13.56	2.57	13.80	3.15
Father–Mother actor covariance (A^F_{sk}, A^M_{sk})	−0.22	2.02	−0.51	2.03
Father partner variance (B^F_{sk})	20.27	5.32	19.78	5.45
Mother partner variance (B^M_{sk})	0	*	0.30	2.45
Child partner variance (B^C_{sk})	0	*	0	*
Father actor-partner covariance (A^F_{sk}, B^F_{sk})	9.05	3.68	8.00	3.76
Mother actor-partner covariance (A^M_{sk}, B^M_{sk})	0	*	1.66	1.43
Child actor-partner covariance (A^C_{sk}, B^C_{sk})	0	*	0	*
Dyad variance relationship with father ($R^F_{(ij)k}$)	#	#	2.02	1.25
Dyad variance relationship with mother ($R^M_{(ij)k}$)	#	#	0.24	0.85
Residual variance (E_{ijk})	5.27	0.69	4.19	0.89
Deviance	1994.92		1989.18	

[#]Not estimated. *No standard error.

Software to estimate the social relations model

› *The social relations model is typically estimated with the software SOREMO by David Kenny, available at the website http://davidakenny.net/srm/srm.htm*

  Downside: SOREMO cannot handle missing data.

› *It can be estimated as well with software that allows to fit random effects models, such as (Win)BUGS, available at the website http://www.mrc-bsu.cam.ac.uk/bugs/ , or MLwiN (see Tom Snijders' website for macros)*

› *Simple specifications (with actor-partner covariance assumed to be zero) can be estimated with standard software like SPSS. Consult David Kenny's website (see above) for how to do this.*