10^{th} Winter School on Longitudinal Social Network Analysis

14-16 January 2019

University of Groningen

Academy Building

c.e.g.steglich@rug.nl

university of groningen

Network dynamics: basic issues

Why?

&

How?

Why model network dynamics?

- Research often starts as a question about associations between network features and *individual* features…
  - *Do popular students engage in risky behaviours?*
  - *Do South Europeans have more friends than North Europeans?*
- … or between network features and *dyadic* features.
  - *Do students ethnically segregate in school?*
  - *Do students with protective parents tend to be friends with risk-taking peers?*
- Any such association begs the question of *"Why?"* …

Explanatory mechanisms are of a dynamic nature

Competing explanations are the rule:

Do students ethnically segregate in friendship networks…

1. *… because they prefer their own ethnic group?*
2. *… because they prefer ties to students living in their neighbourhood, and there is* residential *segregation?*
3. *… because inter-ethnic ties break up more quickly?*

These explanations invoke dynamic mechanisms:

1. *Same ethnicity precedes tie formation.*
2. *Geographic proximity precedes tie formation.*
3. *Same ethnicity precludes tie dissolution.*

How to model network dynamics?

Ultimate criterion: *"Such that you can tell which of the competing mechanisms can account for the data."*

- Statistical approach needed to control different effects for each other
- Longitudinal approach needed to link antecedents to consequences
- Sociocentric network approach needed because conceptually, selection can only be studied when the complete pool of candidates is known

Statistics & non-independent data: specialised models!

Basic framework for stochastic network models

It is assumed that networks are random variables X with a (typically complex) probability distribution.

An observed network x is assumed to be drawn from the space $\mathfrak{X}$ of all possible networks according to this distribution.

This distribution ...

... can be formalized in a *parametric model*,

... can at least be *simulated* (e.g., by MCMC techniques),

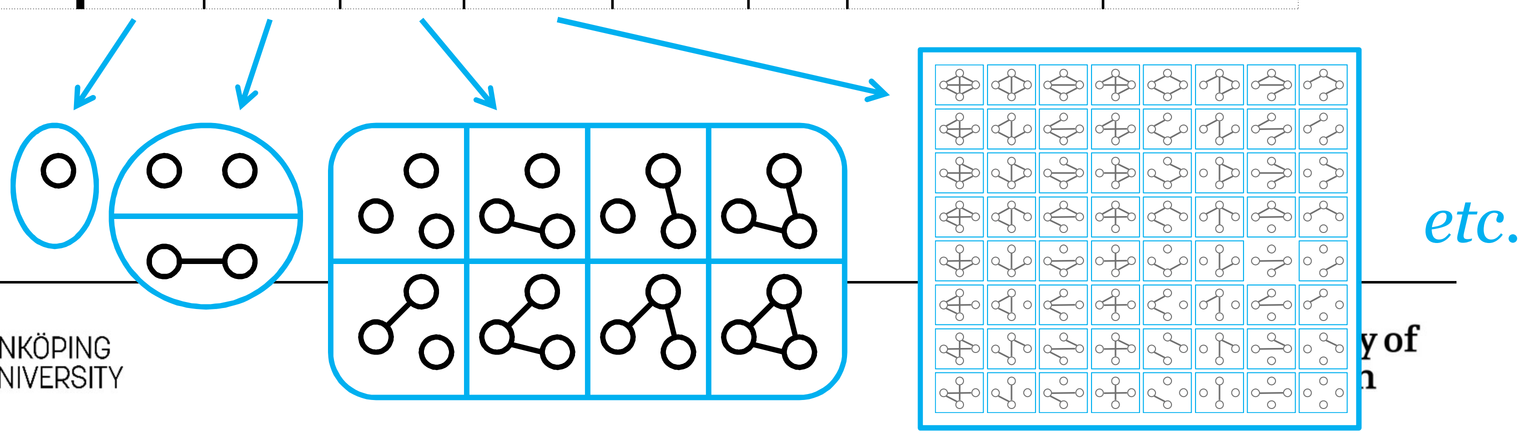
... and therefore be used for *hypothesis testing*.

The network space is huge!

For an undirected, binary network among *n* actors, how many networks are possible?

- For each dyad (i,j) there are two possibilities: $x_{ij}=0$ or $x_{ij}=1$.
- There are $\frac{n(n-1)}{2}$ dyads that can be combined in any way.
- So in total, there are $|\mathfrak{X}| = 2^{n(n-1)/2}$ networks possible.

n	1	2	3	4	5	…	10	…
$\lvert\mathfrak{X}\rvert$	1	2	8	64	1024	…	~35 trillion	…



The stochastic, actor-based framework

Model assumptions
Data format considerations
The network evolution algorithm
Model specification, selection of effects

LINKÖPING UNIVERSITY

university of groningen

Model assumption #1: Actor-basedness

Social actors are the locus of modelling

- change is due to actors' decisions*;
- actors control "their" (i.e., outgoing) network ties.

Two basic actor-level model components

- *When* can an actor make a decision? (***rate function***)
- *Which* decision does the actor make? (***objective function***)

[* assumption: Luce's (1959) "independence of irrelevant alternatives" axiom, or, equivalently, Yellot's (1977) weaker "invariance with uniform expansion" criterion.]

Model assumption #2: Decomposability

Subsequent *discrete-time* observations* are assumed to be related through an unobserved *continuous-time process* of change on $\mathfrak{X}$.

These changes are *as small as possible* ('mini steps').

Complex observed change patterns thus are assumed to be *decomposable into a mini step sequence* of many smallest-possible unobserved changes.

[*Fully observed continuous-time data can in principle be analysed with standard statistical software, but this requires considerable data organization skills.]

What are smallest possible changes?

- Changes between two networks that differ *by just one tie variable*, while all others are identical.
  - Example directed network:

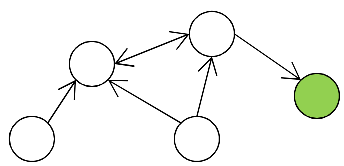
  - Example undirected network:

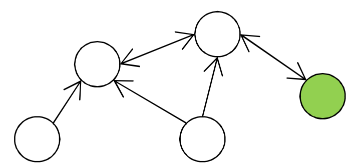
Terminology: these networks *differ by a "mini step"*.

Mini steps and locus of control

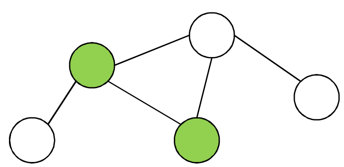
A mini step involves uniquely identified actors – these are assumed to ***control & decide about*** the tie variable:

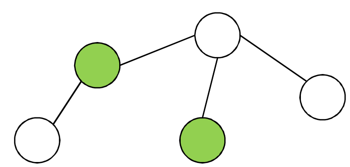
- Directed network: ONE actor





- Undirected network: TWO actors





A mini step is therefore simpler to model, in an actor-based way, in the ***directed case***. We here present only this simpler case.

Data requirements

Required are ***repeated measures of the same network***

- same definition of the network boundary

  (same group of actors, but some composition change is allowed)

- same relational variable.

(more requirements to follow)

<u>Example data:</u> *(Andrea Knecht, 2003/04)*

Networks among first grade pupils at Dutch secondary schools ("bridge class").

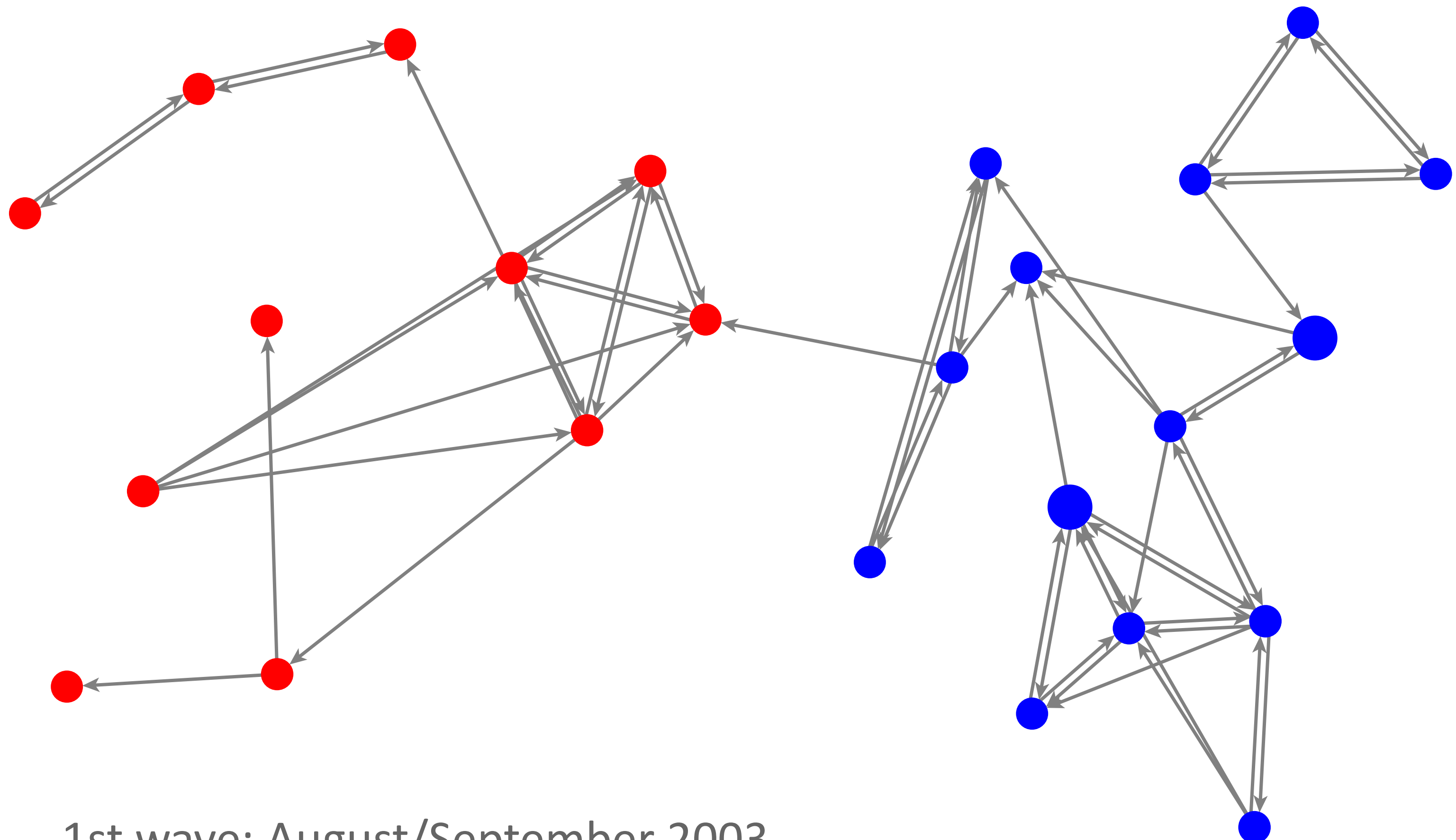
125 school classes

4 measurement points,

various network & individual measures.

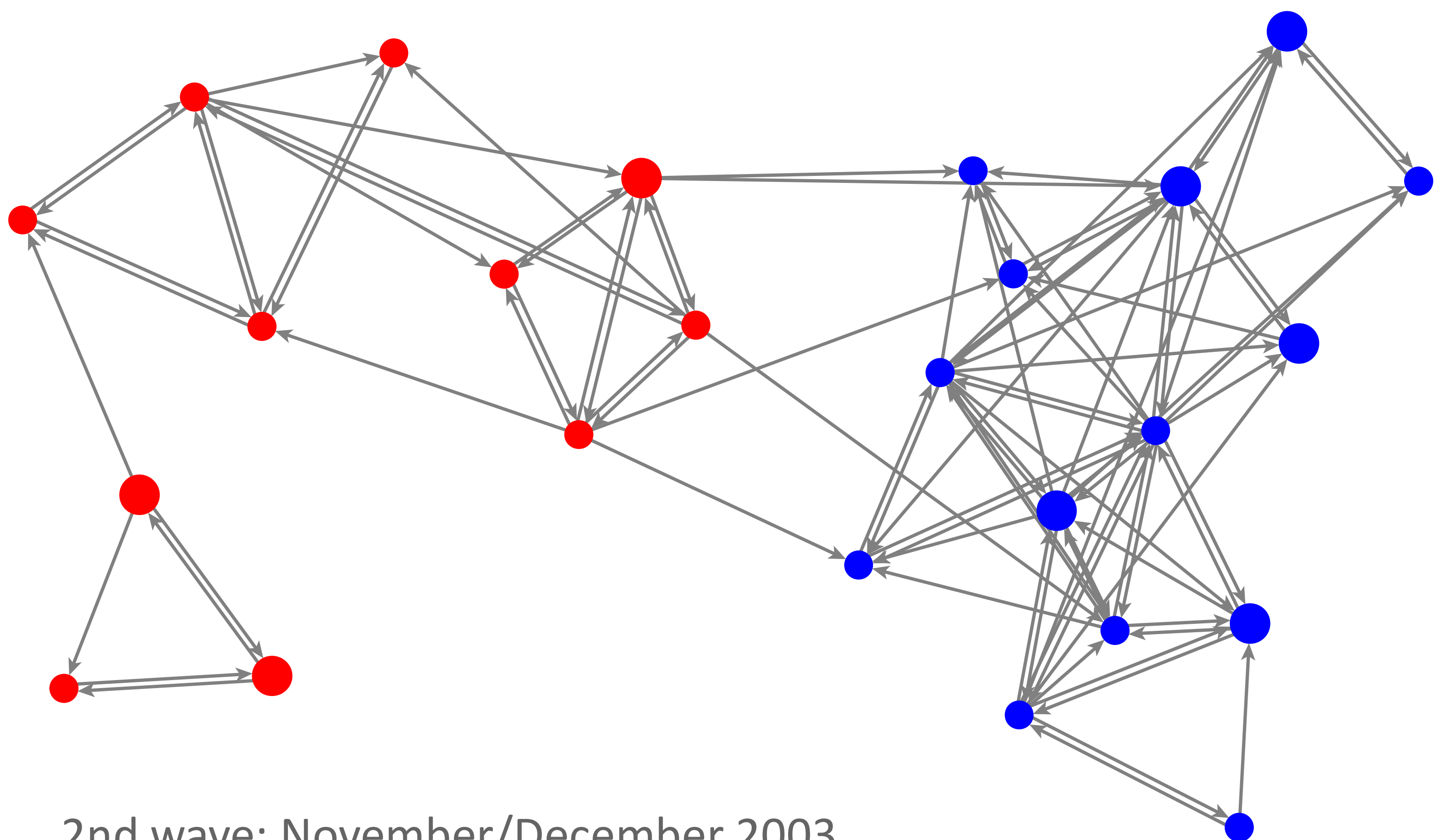
The following slides show the evolution of the friendship network in one classroom.

The graph layout is a bit messy for each observation alone, but optimal over time according to a stress minimisation algorithm.



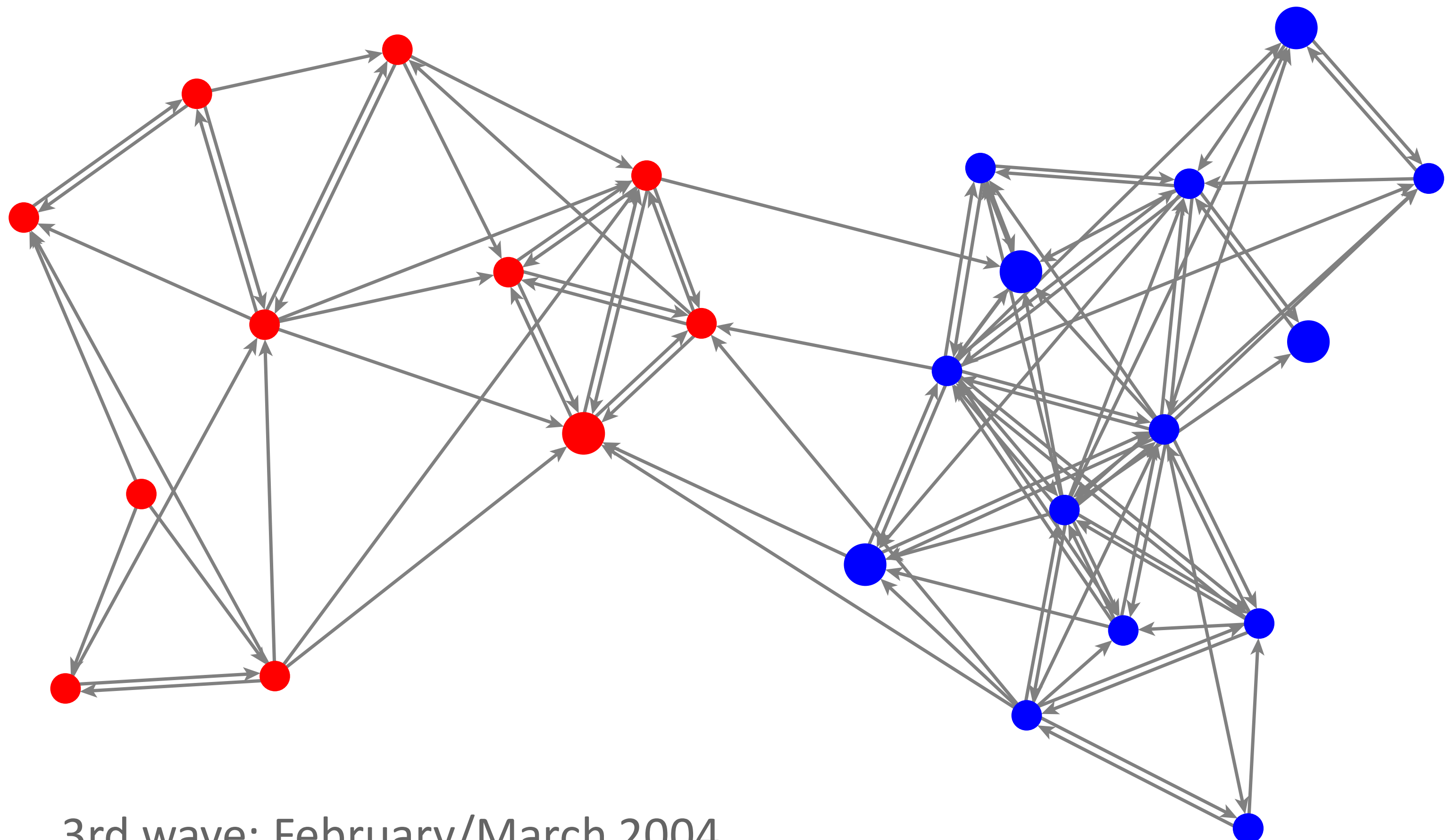
1st wave: August/September 2003

Node size indicates strength of delinquency...



<u>2nd wave:</u> November/December 2003

... and node node colour indicates sex.



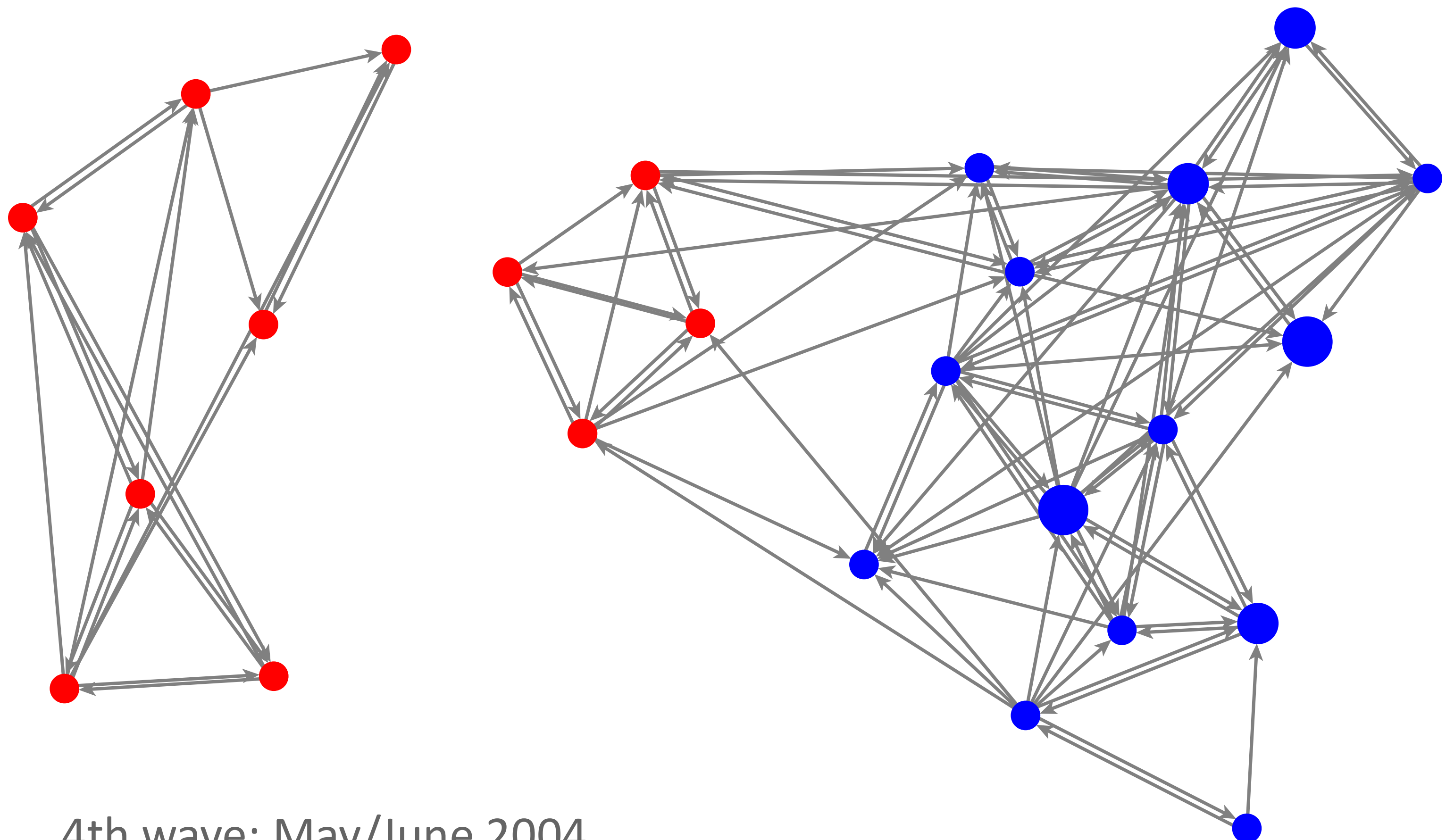
3rd wave: February/March 2004

Dynamic layout algorithms can be used to...

LINKÖPING UNIVERSITY

university of groningen

visone



4th wave: May/June 2004

... animate the data in a movie.

LINKÖPING UNIVERSITY

university of groningen

Points to consider before trying actor-based models

change ↔ stability

The networks should change ‘slowly’, contain a stable part.
Rules for structural change typically are about individual ties changing in response to surrounding ties (which remain stable, for that moment).

states ↔ events

NOT snapshots of e-mail traffic, BUT reliable measures of a slowly-changing social relation.
Event networks could be aggregated over (ideally: non-overlapping) time windows to obtain ‘state-type’ networks.

Quantifying stability & change between two observations

Hamming distance $H = n_{01} + n_{10}$

number of observed changes; indicates minimum number of mini steps needed to reach the second network from the first ($\rightarrow$ power).

Jaccard index $J = n_{11} / (n_{11} + n_{01} + n_{10})$

percentage of ties observed twice among ties that were observed at least once; indicates stability

(<0.1: very problematic; >0.3 safe)

	no tie at T2	tie at T2
no tie at T1	n_{00}	n_{01}
tie at T1	n_{10}	n_{11}

frequencies of observed change patterns

Data format issues to consider

binary	↔	**signed**	↔	**valued**
directed	↔	**undirected**		
tie loss possible	↔	**growth only networks**		
one-mode	↔	**two-mode / bipartite / affiliation-type**		
single dependent	↔	**multiplex**		

The standard model is developed for a single dependent, binary, directed, one-mode network that can both grow and shrink over time.

Everything else is a non-standard model extension, and not necessarily supported by the software implementation.

Modelling principles for such data sets

Random walk: Network evolution proceeds as a stochastic process on the space $\mathfrak{X}$ of all possible networks.

No contamination by the past: The first observation is not modelled but conditioned upon as the process' starting value.

Continuous-time model: <u>Change</u> is modelled as occurring in continuous time.

Mini steps: <u>Big</u> change from one observation to the next is assumed to accrue from a sequence of <u>smallest possible</u> changes.

The assumption of *temporal decomposability / separability* is quite a strong one, but crucial for statistical power!

The network evolution algorithm

Network evolution in observation period $t_0 \rightarrow t_1$ takes place as in this *straight simulation* algorithm:

1. Model time is set to $t = t_0$, and simulation starts out at the network observed at this time point.
2. For each actor, a *waiting time* is sampled according to the actor's *rate function*.
3. The actor with the shortest waiting time τ is identified.
4. If $t + \tau > t_1$, the simulation terminates.
5. Otherwise, the identified actor gets the opportunity to set a mini step. This step is determined by the actor's *objective function*.
6. Model time is updated and simulation proceeds at step 2.

The rate function $\lambda_i(x) = \sum_k \rho_k r_{ik}(x)$

- Models *speed* differences between actors i .
- Statistics r_{ik} of i's neighbourhood in x are weighted by model parameters ρ_k .
- These weights express whether the feature expressed in the statistic is related to more frequent ($\rho_k > 0$) or less frequent ($\rho_k < 0$) network changes by the actors.
- They are estimated from the data.

Technically, λ_i is parameter of an exponential distribution of waiting times – as in Poisson regression.

Typically, it is good to start an analysis under the assumption of a periodwise constant rate function.

The objective function $f_i(x) = \sum_k \beta_k s_{ik}(x)$

- Models attractiveness of network states x to actor i .
- Statistics s_{ik} of i's neighbourhood in x are weighted by model parameters β_k .
- These weights express whether the feature expressed in the statistic is desired ($\beta_k > 0$) or averted ($\beta_k < 0$).
- Also they are estimated from the data.

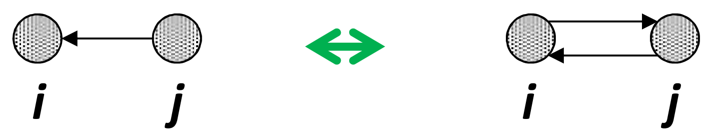
Technically, $f_i(x)$ *is parameter of a conditional logit model for discrete, probabilistic choice.*

The objective function is the main part of modelling.
Here, hypotheses typically are operationalised.

Some effect statistics

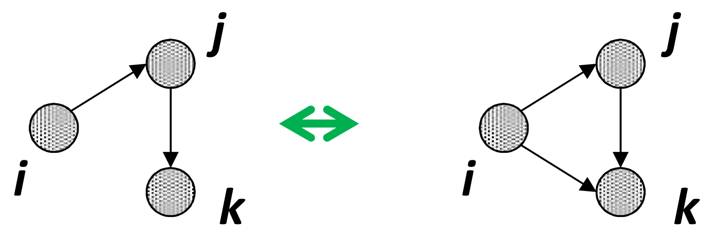
reciprocity effect

$$s_{i\ \mathrm{recip.}} = \sum_j x_{ij} x_{ji}$$



transitivity effect

$$s_{i\ \mathrm{tr.trip.}} = \sum_{jk} x_{ij} x_{jk} x_{ik}$$



Most often, effect statistics are subgraph counts.

Effects' parameters β indicate the attractiveness difference for the focal actor i between the configurations depicted on the right and on the left.

Many other effects are possible to include in the objective function (consult RSiena manual for what is currently possible)…

TABLE 2
SELECTION OF POSSIBLE EFFECTS FOR MODELING NETWORK EVOLUTION

effect	network statistic	effective transitions in network*	verbal description
1. outdegree	x_{ij}		preference for ties to arbitrary others
2. reciprocity	$x_{ij}x_{ji}$		preference for reciprocated ties
3. transitive triplets	$x_{ij}\sum_h x_{ih}x_{hj}$		preference for being friend of the friends' friends
4. balance	x_{ij} strsim$_{ij}$		preference for ties to structurally similar others
5. actors at distance two	$\begin{cases}1 & \text{if between}(h;ij)=1 \text{ for some } h\\ 0 & \text{else}\end{cases}$	(the *number* of intermediaries is irrelevant)	preference for keeping others at social distance two
6. popularity alter	$x_{ij}\sum_h x_{hj}$		preference for attaching to popular others, i.e., others who are often named as friend ('preferential attachment')
7. activity alter	$x_{ij}\sum_h x_{jh}$		preference for attaching to active others, i.e., others who name many friends
8. 3-cycles	$x_{ij}\sum_h x_{jh}x_{hi}$		preference for forming relationship cycles (negative indicator for hierarchical relations)
9. betweenness	$\sum_h \text{between}(i;hj)$	(*no direct link* from the left to the right actor)	preference for being be in an intermediary position between unrelated others
10. dense triads	$\sum_h \text{group}(ijh)$		preference for being part of cohesive subgroups
11. peripheral	$\sum_{hk} \text{peripheral}(i;jhk)$		preference for unilaterally attaching to cohesive subgroups
12. similarity	x_{ij} sim$_{ij}$		preference for ties to similar others (selection)
13. behavior alter	$x_{ij}z_j$		main effect of alter's behavior on tie preference
14. behavior ego	$x_{ij}z_i$		main effect of ego's behavior on tie preference
15. similarity × reciprocity	$x_{ij}x_{ji}$ sim$_{ij}$		preference for reciprocated ties to similar others
16. between dis-similar alters	$\sum_h (1-\text{sim}_{jh})\text{between}(i;jh)$		preference for being in an intermediary position between unrelated, dissimilar others (brokerage potential)
17. similarity × dense triads	$\sum_h \text{group}(ijh)\left(\text{sim}_{ij}+\text{sim}_{ih}\right)$		preference for being part of behaviorally similar cohesive subgroups
18. behavior × peripheral	$z_i\sum_{hk} \text{peripheral}(i;jhk)$		behavior-specific preference for unilaterally attaching to cohesive subgroups
19. similarity × peripheral	$\sum_{hk}\left(\text{peripheral}(i;jhk) \times(\text{sim}_{ij}+\text{sim}_{ih}+\text{sim}_{ik})\right)$		preference for unilaterally attaching to behaviorally similar cohesive subgroups

* In the *effective transitions* illustrations, it is assumed that the behavioral dependent variable is dichotomous and centered at zero; the color coding is ○ = low score (negative), ● =high score (positive), ○ =arbitrary score. The tie x_{ij} from actor i to actor j is the one that changes in the transition indicated by the double arrow. Illustrations are not exhaustive.

LINKÖPING UNIVERSITY

groningen

Choice probabilities $\Pr(x \to_i x') \propto \exp(f_i(x'))$

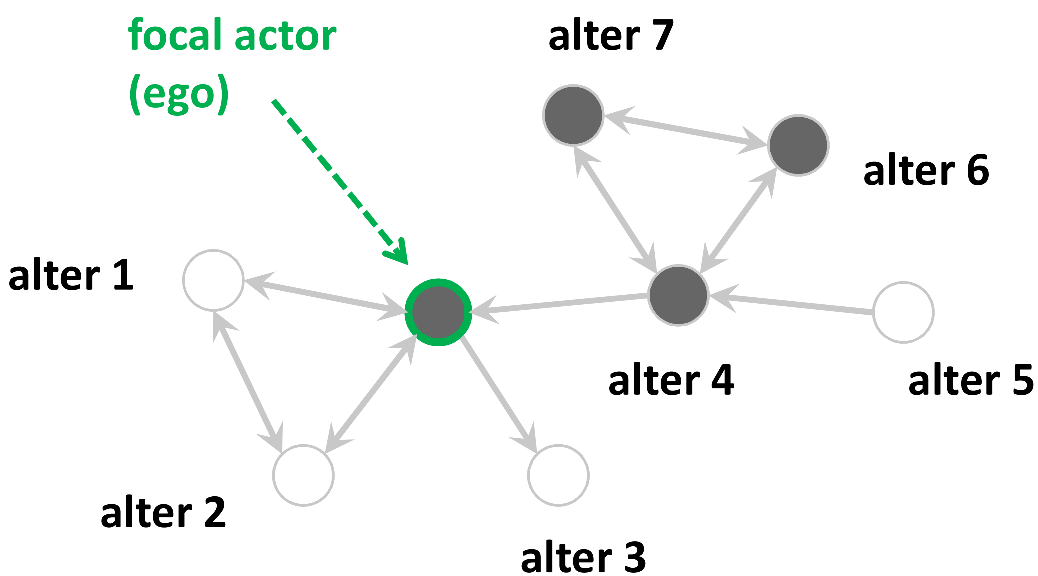
- Choice probabilities for mini steps are proportional to the exponential function of the objective function.
- Valid options are all possible mini steps, plus the option not to change the status quo.
- This probability distribution can be interpreted as optimisation of a ***random utility function***, namely the objective function f_i plus a Gumbel-distributed error term.
- Note that the probabilities only depend on x' and not on past states, not even x . This can be relaxed while keeping the Markov property (keyword: ***creation & maintenance functions***).

Illustration of "how a mini step works"

Assume that a model specification with the following objective function parameters was estimated on a classroom friendship network:

- outdegree $\beta_{\mathbf{outdg.}} = -2.6$ *friendship is rare*
- reciprocity $\beta_{\mathbf{recip.}} = 1.8$ *friendship is reciprocal*
- transitivity $\beta_{\mathbf{tr.trip.}} = 0.4$ *friendship shows clustering*
- three-cycles $\beta_{\mathbf{3\text{-}cycl.}} = -0.7$ *friendship shows hierarchy*
- same gender $\beta_{\mathbf{same}} = 0.8$ *friendship is sex segregated*

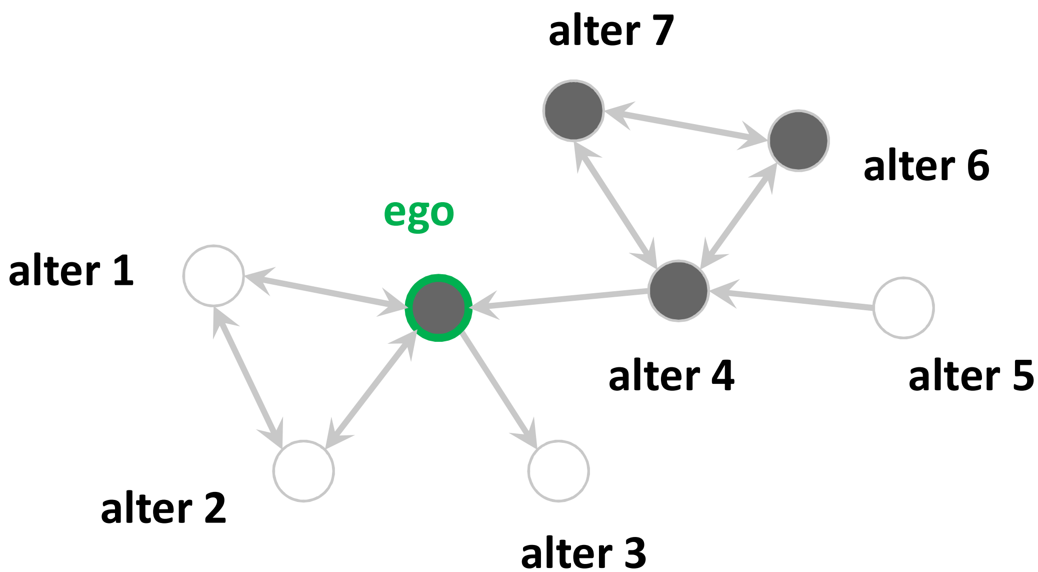
Example of an actor's decision



Ego's choice options:

- drop tie to alter 1
- drop tie to alter 2
- drop tie to alter 3
- create tie to alter 4
- create tie to alter 5
- create tie to alter 6
- create tie to alter 7
- keep status quo

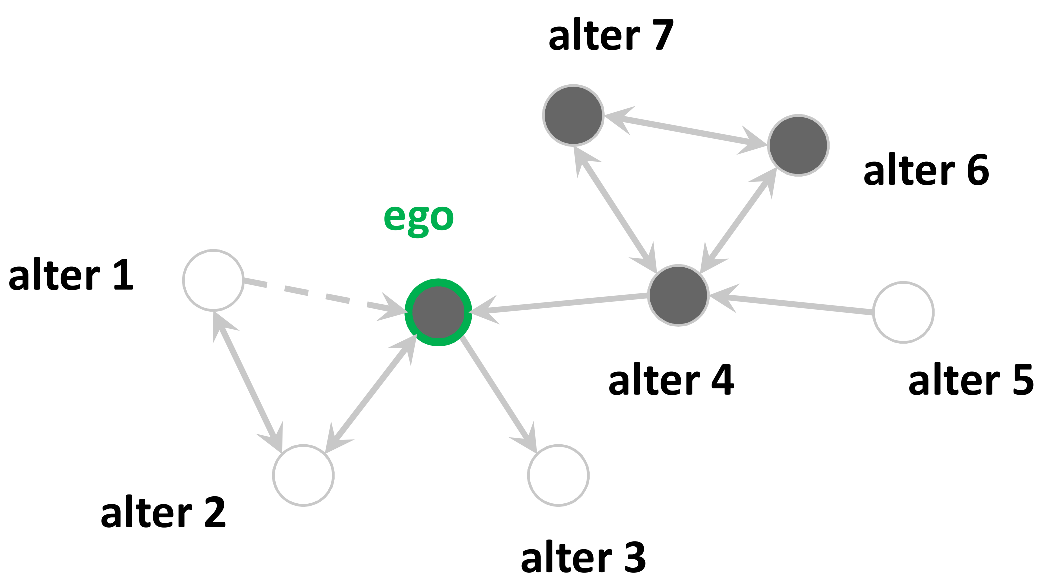
Count model-relevant subgraphs for all options



Keep status quo

- 3 outgoing ties
- 2 reciprocated ties
- 2 transitive triplets
- 2 three-cycles
- 0 same gender ties

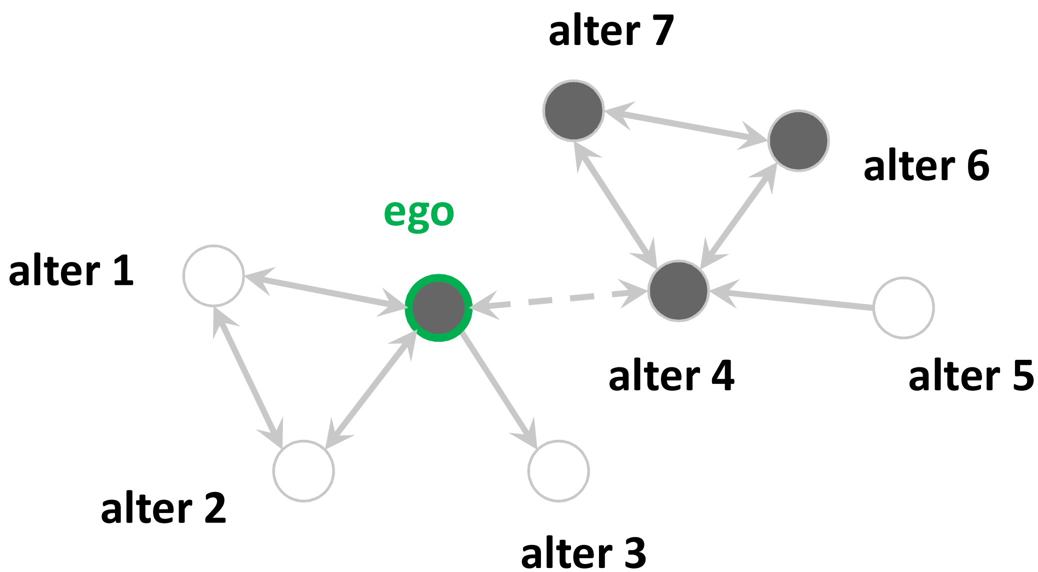
Count model-relevant subgraphs for all options



Drop tie to alter 1

- 2 outgoing ties
- 1 reciprocated tie
- 0 transitive triplets
- 1 three-cycles
- 0 same gender ties

Count model-relevant subgraphs for all options



Create tie to alter 4

- 4 outgoing ties
- 3 reciprocated ties
- 2 transitive triplets
- 2 three-cycles
- 1 same gender tie

…these calculations are done for all the eligible options.

Result: a decision matrix of options by attributes

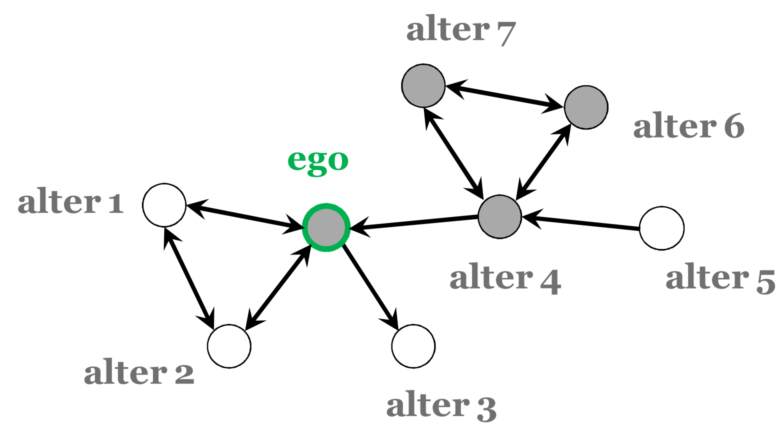
Option	# out-ties	# recip. ties	# tr.triplets	# 3-cycles	# same sex
drop tie to alter 1	2	1	0	1	0
drop tie to alter 2	2	1	0	1	0
drop tie to alter 3	2	2	2	2	0
create tie to alter 4	4	3	2	2	1
create tie to alter 5	4	2	2	3	0
create tie to alter 6	4	2	2	3	1
create tie to alter 7	4	2	2	3	1
keep status quo	3	2	2	2	0

Calculation of objective function values:

$$
\begin{pmatrix} f_{\text{drop-1}} \\ f_{\text{drop-2}} \\ f_{\text{drop-3}} \\ f_{\text{create-4}} \\ f_{\text{create-5}} \\ f_{\text{create-6}} \\ f_{\text{create-7}} \\ f_{\text{stat.quo}} \end{pmatrix} = \textbf{Matrix } S_{ego} \begin{pmatrix} \beta_{\text{outdg.}} \\ \beta_{\text{recip.}} \\ \beta_{\text{tr.trip.}} \\ \beta_{\text{3-cycles}} \\ \beta_{\text{same}} \end{pmatrix}
$$

Option	# out-ties	# recip. ties	# tr.triplets	# 3-cycles	# same col.
drop tie to alter 1	2	1	0	1	0
drop tie to alter 2	2	1	0	1	0
drop tie to alter 3	2	2	2	2	0
create tie to alter 4	3	2	2	2	1
create tie to alter 5	4	2	2	3	0
create tie to alter 6	4	2	2	3	1
create tie to alter 7	4	2	2	3	1
keep status quo	3	2	2	2	0

Option	objective function	exponential transform	probability
drop tie to alter 1	-4.1	0.017	10%
drop tie to alter 2	-4.1	0.017	10%
drop tie to alter 3	-2.2	0.111	68%
create tie to alter 4	-4.8	0.008	5%
create tie to alter 5	-8.1	0.000	0%
create tie to alter 6	-7.3	0.001	0%
create tie to alter 7	-7.3	0.001	0%
keep status quo	-4.8	0.008	5%



Dropping the tie to alter 3 clearly dominates this decision situation.

Note: RSiena internally centers many variables – this does not affect the choice probabilities.

LINKÖPING UNIVERSITY

university of groningen

Homogeneity assumptions

Unless otherwise specified (by including interaction terms with nuancing variables), the model assumes…

Actor homogeneity:

All actors follow the same behavioural rules in their networking activities.

Time homogeneity:

These behavioural rules do not change over time.

This can be problematic whenever actors or time periods are heterogeneous but there are no predictors for differences in the data. *So: check this in your models!*

Model specification / effect selection

When investigating social network dynamics, researchers ususally do not come empty-handed but have *theories* or (at least) *hypotheses* about the mechanisms that might operate.

- *These mechanisms [hopefully] can be expressed in terms of SIENA parameters, and the hypotheses can be restated in terms of the corresponding model parameters.*
- *By estimating the parameters and calculating significance tests for them, the theories / hypotheses can be tested empirically.*

But... how do parameters & hypotheses relate to each other?

Local characterisation of choice probabilities

- For two networks that could be obtained in competing mini steps from the same network of origin, the ratio of choice probabilities is this ("odds"):

$$\frac{\Pr(x^c \rightarrow_i x^a)}{\Pr(x^c \rightarrow_i x^b)} = \exp\left(\sum_{k=1}^{K} \beta_k \underbrace{\left(s_{ik}(x^a) - s_{ik}(x^b)\right)}\right)$$

compared are two moves ('mini steps') made by actor ***i*** from a network $\boldsymbol{x^c}$ to two "neighbouring networks" $\boldsymbol{x^a}$ and $\boldsymbol{x^b}$

model parameters

difference in model statistics of actor ***i*** between the two compared moves

The main part of the formula in detail:

The sum $\sum_{k=1}^{K} \beta_k \left(s_{ik}(x^a) - s_{ik}(x^b) \right)$ determines whether x^a or x^b is more likely to succeed x^c in the network evolution process.

β_k positive: states with *higher* scores s_{ik} are more likely than states with *lower* scores;

β_k negative: states with *lower* scores s_{ik} are more likely than states with *higher* scores.

This way, parameter values β_k express dynamic tendencies of network evolution: *"actors are moving towards a high [low] score on the corresponding network statistic $s_{.k}$"*

Examples

Advice seeking among MBA students

Friendship of adolescents in a classroom

First example (Torlò, Steglich, Lomi & Snijders, 2007)

- **75 students** enrolled in an MBA program;
- **4 network variables**: advice-seeking, communication, friendship, acknowledge-contribution-to-learning;
- **co-evolving behavioural dimension**: performance in examinations;
- **several other actor variables**: gender, age, experience, nationality;
- **3 waves** in yearly intervals.

We focus here on the analysis of the evolution of the *advice network* only.

Which hypotheses are investigated? *[just 3 of them…]*

1. You seek advice from your friends.

Mechanism: presence of a friendship tie between two actors increases the likelihood that an advice tie is present between the same actors.

If x_{ij} stands for i seeking advice from j and w_{ij} stands for i naming j as a friend, then the effect

$$s_{i\ \mathrm{friend}}(x) = \sum_j x_{ij} w_{ij}$$

operationalises the above mechanism, and the corresponding parameter β_{friend} can be used to test it.

- The effect statistic $s_{i\,\text{friend}}$ counts the degree to which advice seeking and friendship 'overlap'.
- The parameter β_{friend} expresses whether by changing the advice network, such an overlap is sought or avoided, i.e., whether friendship enhances or weakens advice seeking:

  β_{friend} positive: advice seeking is more likely when it coincides with friendship;

  β_{friend} negative: advice seeking is less likely when it coincides with friendship.
- In SIENA, the effect can be included as main effect of a dyadic covariate (friendship) on network evolution.

Hypothesis 1: $\beta_{\text{friend}} > 0$; test the null hypothesis $\beta_{\text{friend}} = 0$.

2. ***The lower your performance, the more advice you need*** *[and the more you will seek]***.**

Mechanism: actors with low performance scores are likely to have more outgoing advice ties than actors with high performance scores.

If z_i stands for performance of actor i, then the effect

$$s_{i\ \text{own-performance}}(x) = z_i \sum_j x_{ij}$$

operationalises the above mechanism, and the parameter $\beta_{\text{own-performance}}$ can be used to test it.

- The effect statistic $s_{i\ \text{own-performance}}$ counts the degree to which active advice seeking and performance coincide.
- The parameter $\beta_{\text{own-performance}}$ expresses whether by changing the advice network, such an coincidence is sought or avoided, i.e., whether own performance enhances or weakens advice seeking:

  $\beta_{\text{own-performance}}$ positive: high performers seek more advice than low performers;

  $\beta_{\text{own-performance}}$ negative: high performers seek less advice than low performers.
- In SIENA, the effect can be included as an ego-effect of an actor variable (performance) on network evolution.

Hypothesis 2: $\beta_{\text{own-p.}} < 0$; test the null hypothesis $\beta_{\text{own-p.}} = 0$.

3. The higher your performance, the better the advice you can give *[and the more you will be asked for advice]**.***

Mechanism: actors with high performance scores are likely to attract more incoming advice ties than actors with low performance scores.

Let z_j now stand for performance of actor j, then effect

$$s_{i\ \text{others-performance}}(x) = \sum_j z_j x_{ij}$$

operationalises the above mechanism, and the parameter $\beta_{\text{others-performance}}$ can be used to test it.

- The effect statistic $s_{i\ \text{others-performance}}$ counts the degree to which passive advice seeking (‘being asked’) and performance coincide.
- The parameter $\beta_{\text{others-performance}}$ expresses whether by changing the advice network, such a coincidence is sought or avoided, i.e., whether others’ performance makes them more or less attractive as sources of advice:

  $\beta_{\text{others-perf.}}$ positive: high performers are more often asked for advice than low p’fs.; ☑

  $\beta_{\text{others-perf.}}$ negative: high performers are less often asked for advice that low p’fs. ☒

- In SIENA, this is the alter-effect of an actor variable.

Hypothesis 3: $\beta_{\text{oth-p.}} > 0$; test the null hypothesis $\beta_{\text{oth-p.}} = 0$.

Significance testing of parameters

- The RSiena software estimates parameters β_k and their standard errors st.err.(β_k) .
- By calculating the *standard score* of those, parameter significance can be tested:

  β_k /st.err.(β_k)

  - is approximately normally distributed*
  - under the assumption (null hypothesis) that *actual network evolution* follows a model in which the parameter is constrained to zero (**H_0: β_k =0**).

* Thus far, this claim largely rests on extensive simulation studies.

Results for these particular hypotheses

Dependent network: Advice	Estimate	St.Error	conv.	pred.?
1. eval outdegree (density)	-2.6541	(0.0890)	0.0131	
2. eval reciprocity	0.9973	(0.1231)	-0.0307	
3. eval transitive triplets	0.2781	(0.0291)	-0.0246	
4. eval 3-cycles	-0.1199	(0.0534)	-0.0362	
5. eval indegree - popularity	0.0410	(0.0055)	-0.0368	
6. eval friendship	**1.0230**	**(0.0823)**	-0.0324	✓
7. eval same background	0.1661	(0.0726)	-0.0311	
8. eval same experience	0.1174	(0.0757)	-0.0704	
9. eval performance alter	**0.1035**	**(0.0272)**	-0.0643	✓
10. eval performance ego	**-0.0840**	**(0.0256)**	-0.0564	✓
11. eval performance similarity	0.8371	(0.3044)	-0.0358	

More specifically: tests & p-values

Hypothesis 1: "You seek advice from your friends."

standard score = 1.0230 / 0.0823 = 12.4 ; $p < 0.001$

Hypothesis 2: "The lower your performance, the more advice you seek."

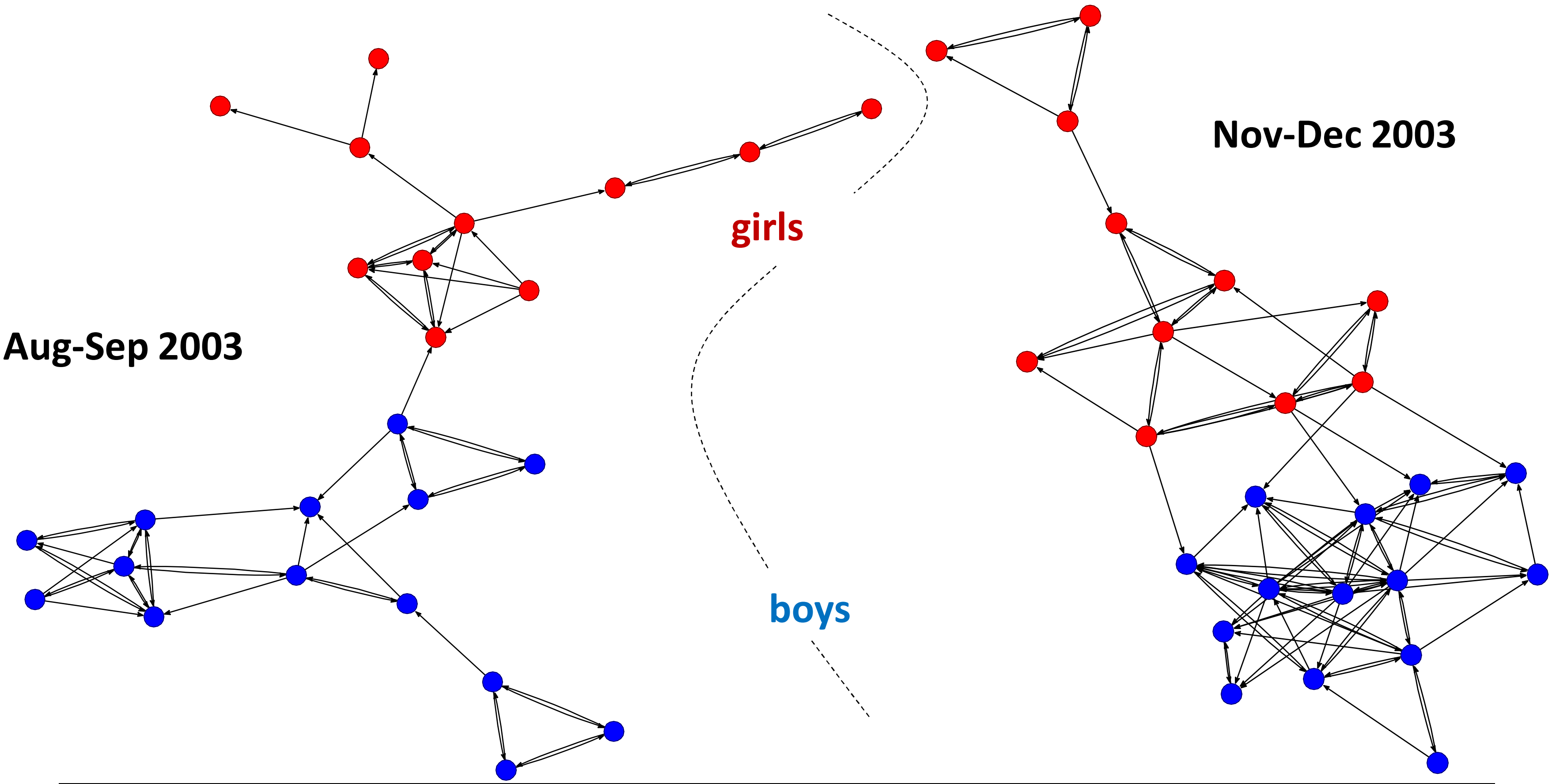
standard score = -0.0840 / 0.0256 = -3.28 ; $p = 0.001$

Hypothesis 3: "The higher your performance, the more others ask you for advice."

standard score = 0.1035 / 0.0272 = 3.81 ; $p < 0.001$

All three hypotheses are confirmed!

Second example: A classroom friendship network



Analyse this network the lab

- … making use of the following effects:
  - outdegree (density),
  - reciprocity,
  - transitive triplets,
  - an interaction effect of the two preceding ones,
  - gender effects of sender and receiver,
  - a gender homophily effect.
- (see exercise on segregation & homophily).

Results of classroom data lab exercise

	Model 1		Model 2		Model 3	
Rate function friendship						
Rate of change $t_1 \rightarrow t_2$	7,54 (0,97)		8,81 (1,31)		10,87 (2,63)	
Rate of change $t_2 \rightarrow t_3$	2,73 (0,45)		2,92 (0,50)		3,04 (0,52)	
Rate of change $t_3 \rightarrow t_4$	3,29 (0,49)		3,56 (0,54)		3,80 (0,65)	
Objective function friendship						
Outdegree	-1,92 (0,17)	***	-2,03 (0,16)	***	-2,19 (0,16)	***
Reciprocity	—		1,09 (0,16)	***	0,84 (0,17)	***
Transitive triplets	—		—		0,18 (0,03)	***
primary school friendship	0,54 (0,21)	*	0,30 (0,21)		0,40 (0,20)	*
Male alter	0,30 (0,18)		0,28 (0,18)		0,05 (0,17)	
Male ego	0,11 (0,19)		0,07 (0,19)		-0,17 (0,18)	
Same sex	1,70 (0,18)	***	1,39 (0,18)	***	0,93 (0,18)	***

'significantly biased'

LINKÖPING UNIVERSITY

university of groningen

On estimation & goodness of fit

Model estimation by estimating equations

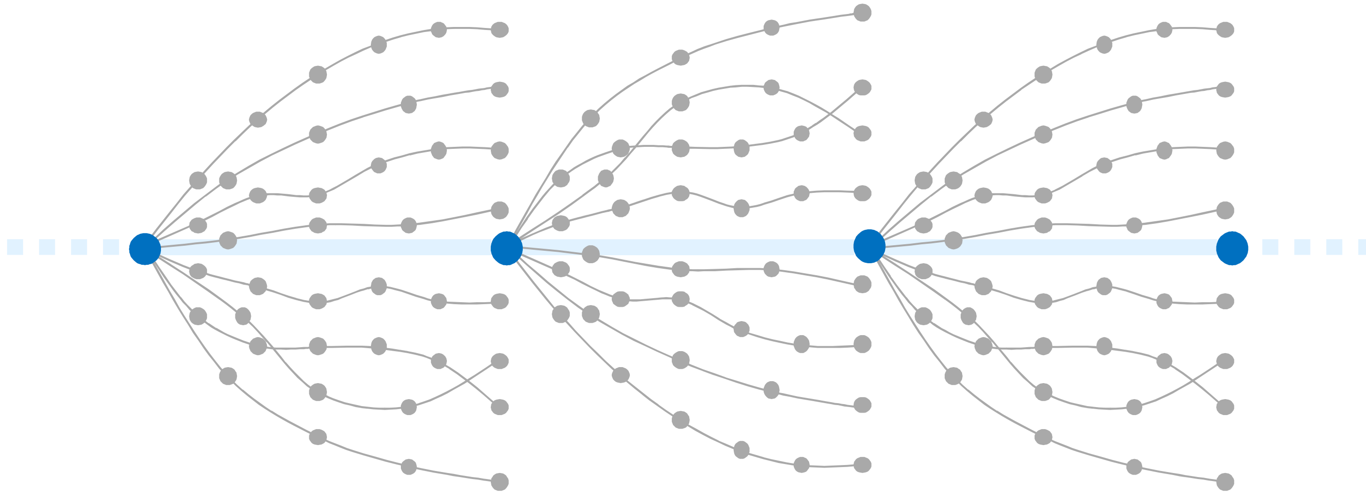
For each model parameter θ. …

whether part of the rate or of the objective function doesn't matter

- a statistic *S.* is identified that can be evaluated on a data set *y* (e.g. the ***observed data*** *x* or draws from the probability model *X* of these data, i.e., ***simulated by straight simulation***)
- this statistic defines an ***estimating equation*** for its parameter: *"Expected value over simulations must equal the observed value."*

The vector θ of parameter estimates is obtained as the joint solution to the corresponding system of equations.

Simulations under Estimating Equations algorithm



- Observed data are met in expectation over simulations, on a vector of *k* statistics corresponding to the *k* model parameters (Snijders, 1996, *Journal of Mathematical Sociology*).
- Trajectories are sampled by straight simulations.

Estimating statistics used

- Basic rate parameter period *m*: $\theta_{\bullet} = \lambda_{\mathbf{m}}$

  Estimating statistic: $\mathbf{S}_{\bullet}(\mathbf{y}) = \sum_{\mathbf{ij}} \left| \mathbf{y}_{\mathbf{ij}}(\mathbf{t}_{\mathbf{m}+1}) - \mathbf{x}_{\mathbf{ij}}(\mathbf{t}_{\mathbf{m}}) \right|$

- Objective function parameters: $\theta_{\bullet} = \beta_{\mathbf{h}}$

  Estimating statistic: $\mathbf{S}_{\bullet}(\mathbf{y}) = \sum_{\mathbf{k}} \sum_{\mathbf{i}} \mathbf{s}_{\mathbf{ih}}(\mathbf{y}(\mathbf{t}_{\mathbf{k}+1}))$

Estimating equations (for all parameters):

$$\mathbf{E}(\mathbf{S}_{\bullet}(\mathbf{X})) = \mathbf{S}_{\bullet}(\mathbf{x})$$

Conditional estimation with estimating equations

- It can be useful to not estimate the rate parameter (modelling the observed amount of network change):
  - Can improve model convergence
  - Focus often is not on rate of change anyway
- Then, the straight simulation algorithm needs to be slightly modified:
  - Stopping rule is not any more "model time exceeds period end time"…
  - …but becomes *"Hamming distance in simulations reaches observed Hamming distance"*

Parameter updating based on simulations

- Parameters are iteratively updated according to the rule

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - a_{k+1}\,\mathbf{D}_0^{-1}\left(\mathbf{S}_k^{\text{sim}} - \mathbf{S}^{\text{obs}}\right)$$

  where…
  - $\mathbf{D}_0$ is the approximation of the derivative matrix of statistics **S** by parameters θ, evaluated at the parameter's starting value θ_0
  - $\boldsymbol{a_k}$ is a sequence of numbers that approach zero at rate $\boldsymbol{k^{-c}}$
  - **c** is chosen (0.5 < **c** < 1) so as to obtain good convergence properties.
- The final parameter estimate $\hat{\boldsymbol{\theta}}$ is the tail average $\frac{1}{N}\sum_{r=1}^{N}\hat{\boldsymbol{\theta}}_{k+r}$

Estimation of covariance structure

- The approximative covariance matrix of the estimator function, evaluated at the estimate, is given by

$$\mathrm{cov}_{\hat{\theta}}(\tilde{\theta}) = \mathbf{D}_{\hat{\theta}}^{-1} \Sigma_{\hat{\theta}}(\mathbf{S}) \mathbf{D}_{\hat{\theta}}$$

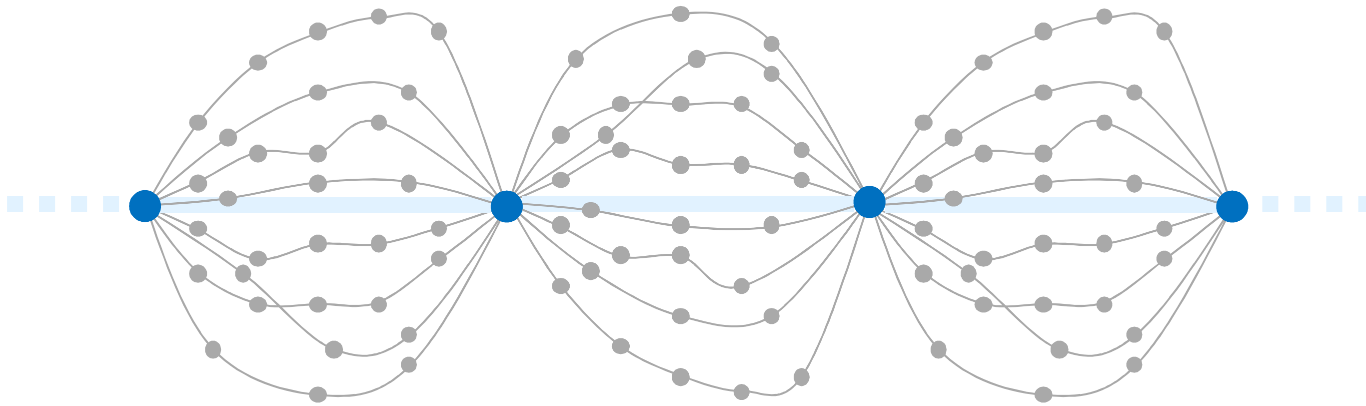
  where…
  - $\mathbf{D}_{\hat{\theta}}$ again is an approximation of the derivative matrix of statistics **S** by parameters θ, now evaluated at the estimate $\hat{\theta}$,
  - $\Sigma_{\hat{\theta}}(\mathbf{S})$ is the matrix of covariance of the simulated vector **S** of estimation statistics, also evaluated at the estimate.
- *Standard errors* of the estimates are calculated as the square roots of the diagonal elements of this matrix.

Model estimation by MCMC maximum likelihood

ML estimation requires approximation of the likelihood of the data, therefore…

- Straight simulation inappropriate (typically doesn't end up in the observed data set)
- Needed: construction of model-consistent distribution of simulated *network evolution trajectories that connect observed data points*
- Is achieved by MCMC techniques (Snijders, Koskinen & Schweinberger, 2010, *Annals of Applied Statistics*)

Simulations under Likelihood-based algorithms



- Observed data are met exactly.
- Connecting trajectories are sampled from the model by MCMC technique (Snijders, Koskinen & Schweinberger, 2010, *Annals of Applied Statistics*).

When to use ML estimation?

Pro:

- Makes more efficient use of available information
- Therefore has higher statistical power
- Relevant for datasets that have low information content already (many missings, little change, small size,…)

Contra:

- Takes considerably more time

Recommendation:

- Use it only when *estimating equations* gives problems

Convergence and goodness of fit checking

What is "goodness of fit" for stochastic network models?

- Simulate many networks from the estimated model, and see how well these simulated networks replicate statistics of the data that were not part of the model.

Since long (since StOCNET): use of convergence indicator

$$\frac{\text{E(simulated values)} - \text{observed value}}{\text{st.dev.(simulated values)}}$$

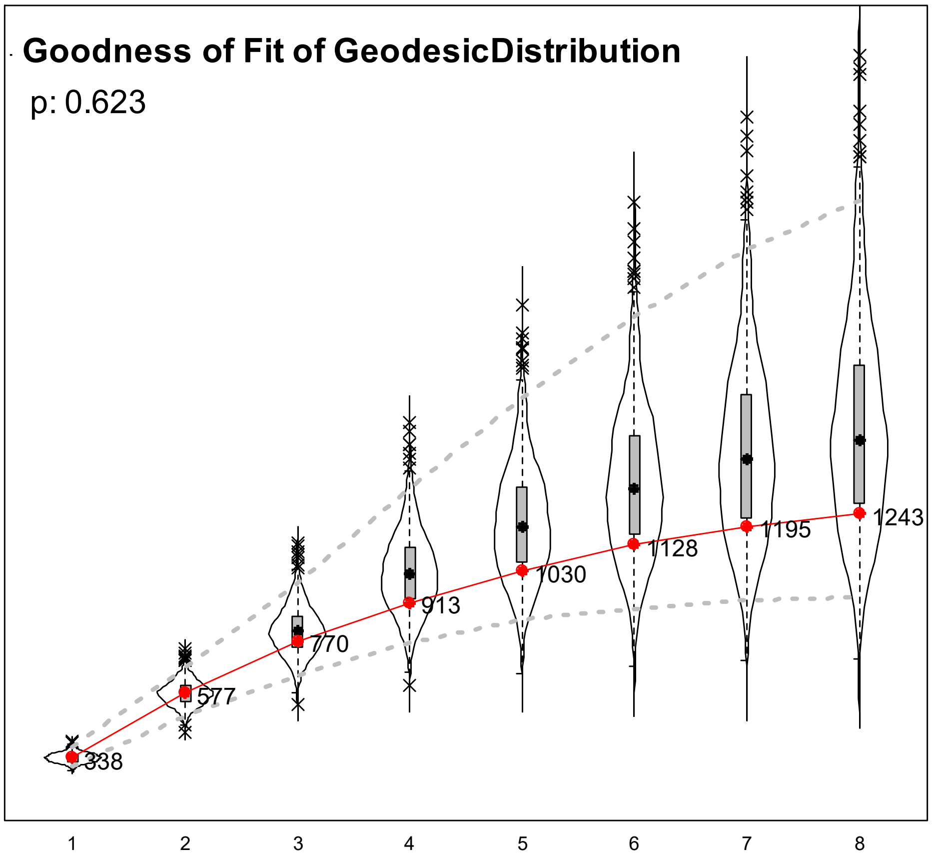
These standard scores are practically zero for estimating statistics (by definition of the estimating algorithm). They should not be too different from zero (ideally n.s.) for other important fit statistics.

Since a while available in RSiena: ‘violin plots’

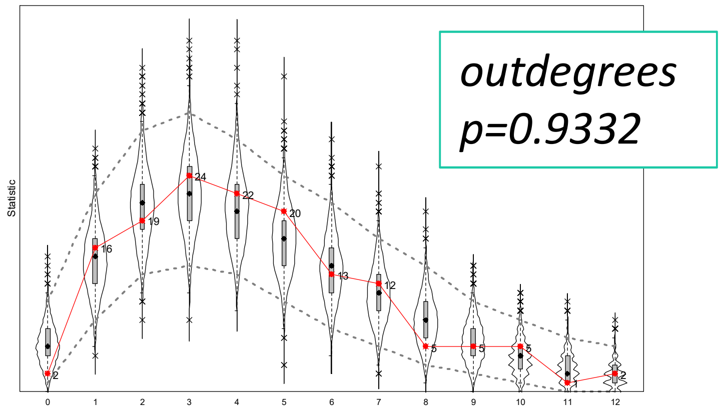
Shows more detail than just a collection of standard scores

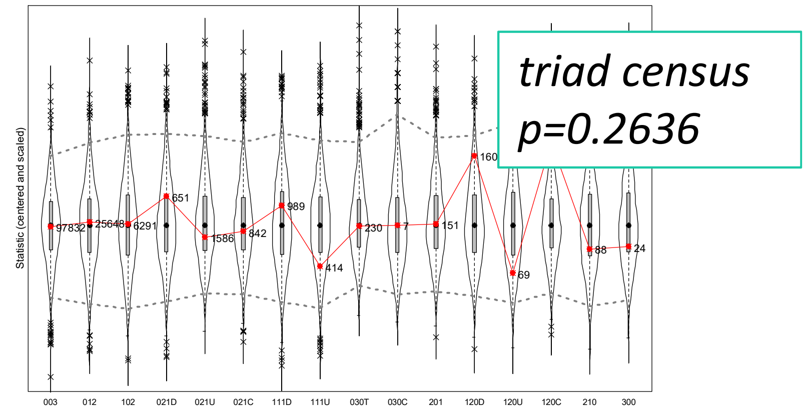
- red solid line shows observed values,
- boxplots & violins show distribution of simulated values,
- p-value is based on Mahalanobis’ distance from centre of simulations.

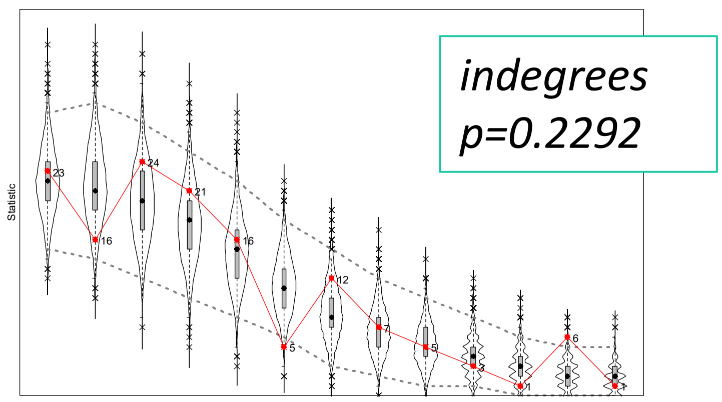
Examples will be elaborated in a lab exercise.

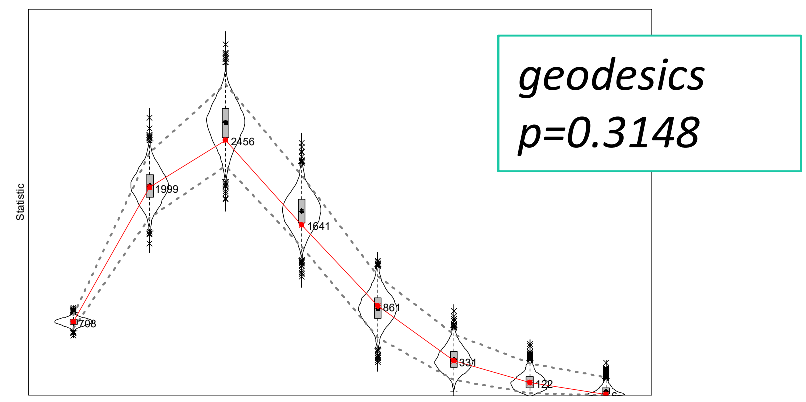


Goodness of fit plots & Mahalanobis distance p-value









Plots belong to the MBA advice seeking example above.

c.e.g.steglich@rug.nl

https://steglich.gmw.rug.nl

LINKÖPING UNIVERSITY

university of groningen