Network Modeling

Viviana Amati Jürgen Lerner Bobo Nick

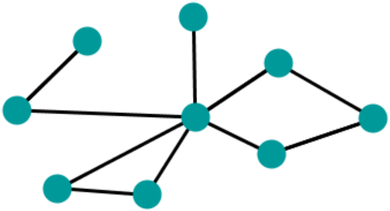
Dept. Computer & Information Science
University of Konstanz

Winter 2011/2012
(version 13 February 2012)

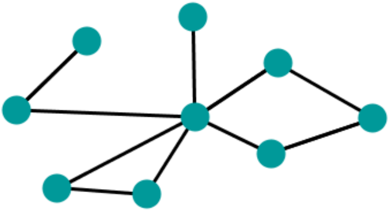
Outline

Outline

Where we are

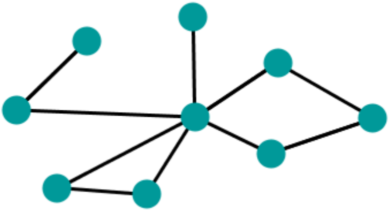


Where we are



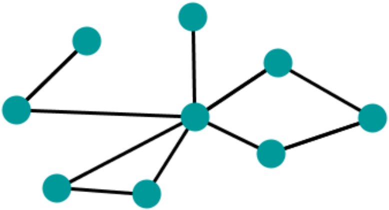
Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties are usually dependent

Where we are



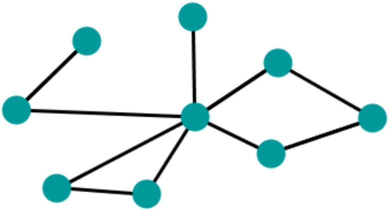
Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties are usually dependent
Preferential attachment	based on degree distribution	there are other structural properties

Where we are



Model	Main feature	Real data
$\mathcal{G}(n,p)$	ties are independent	ties are usually dependent
Preferential attachment	based on degree distribution	there are other structural properties
ERGM	class of models	reasonable representation of the data

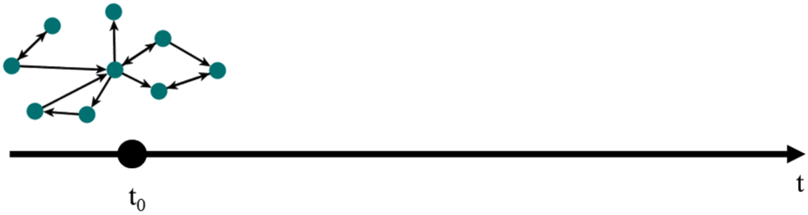
Where we are



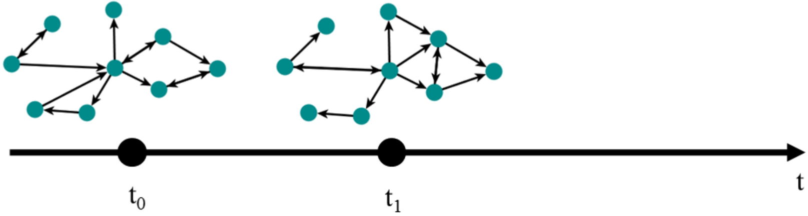
Model	Main feature	Real data
$\mathcal{G}(n,p)$	ties are independent	ties are usually dependent
Preferential attachment	based on degree distribution	there are other structural properties
ERGM	class of models	reasonable representation of the data

These are models for cross-sectional data

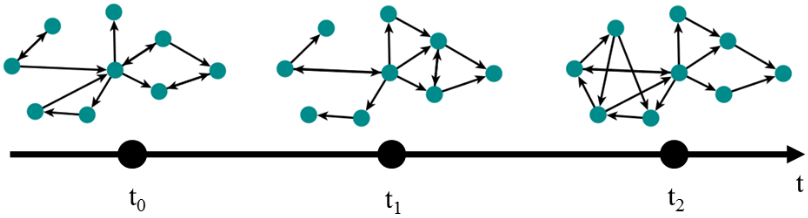
Where we are going



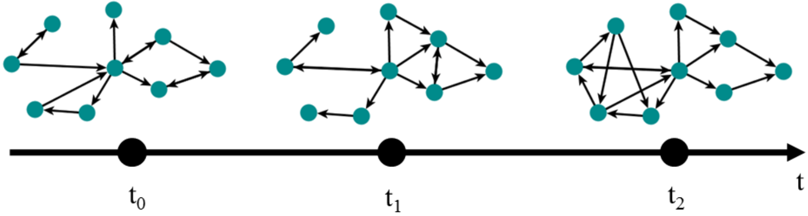
Where we are going



Where we are going

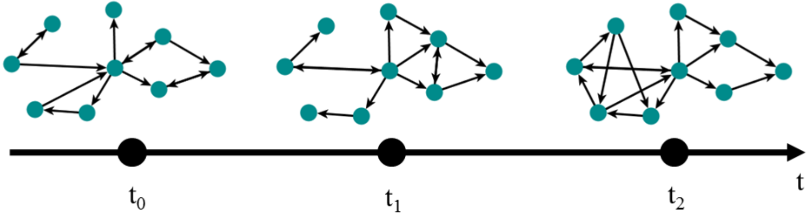


Where we are going



Network are dynamic by nature. How to model network evolution?

Where we are going



Network are dynamic by nature. How to model network evolution?

We need a model for longitudinal data

Networks are dynamic by nature: a real example

The *Teenage Friends and Lifestyle Study* analyzes smoking behavior and friendship

Data (available from www.stats.ox.ac.uk/$\sim$ snijders/siena/)

- One school year group from a Scottish secondary school starting at age 12-13 years, was monitored over 3 years;
- Data were collected using questionnaires at approximately one year interval
  1. Friendship relation: each pupil can name up to six friend
  2. Individual information and lifestyle elements: gender, age, substance use, smoking of parents and siblings etc.

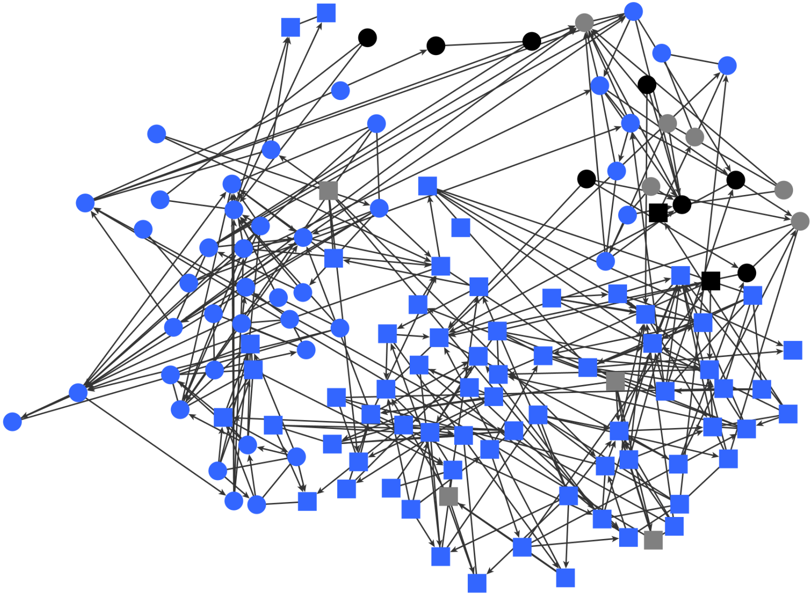
The dataset is depicted by the following graphs where:

arrows denote friendship relation

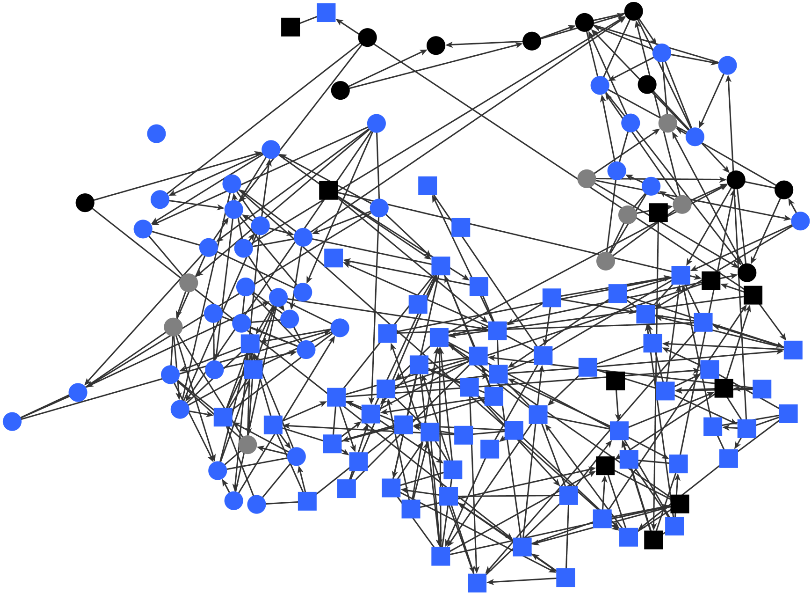
gender is represented by the shape of nodes (girls: circles, boys: squares)

smoking behavior is depicted by the color of nodes (non: blue, occasional: gray, regular: black)

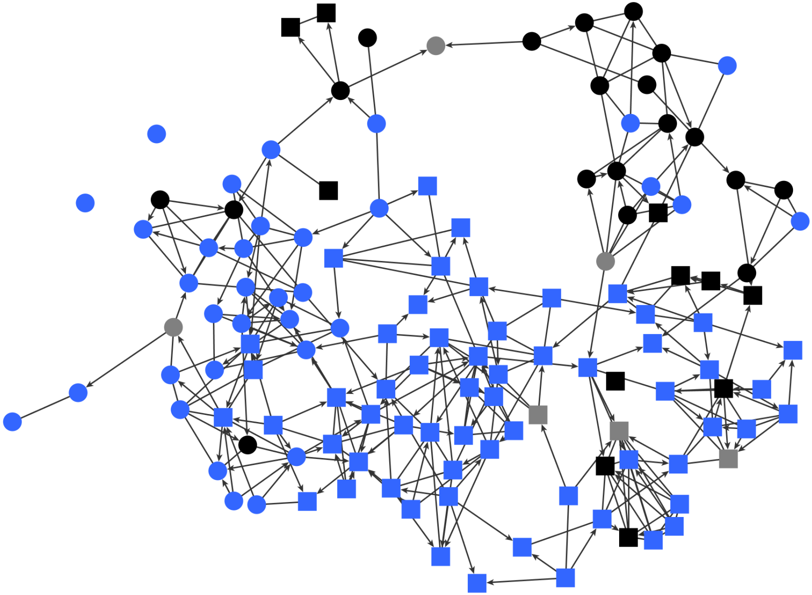
Networks are dynamic by nature: a real example



Networks are dynamic by nature: a real example



Networks are dynamic by nature: a real example



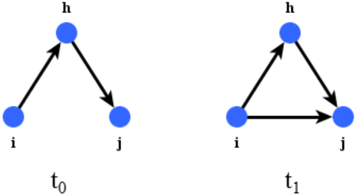
Questions

- Is there any tendency in friendship formation towards reciprocity?

Questions

- Is there any tendency in friendship formation towards reciprocity?

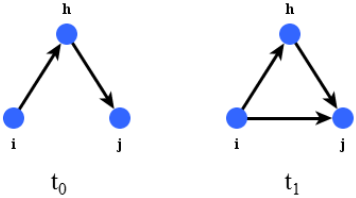
- Is there any tendency in friendship formation towards transitivity?



Questions

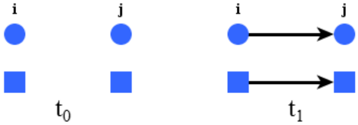
- Is there any tendency in friendship formation towards reciprocity?

- Is there any tendency in friendship formation towards transitivity?



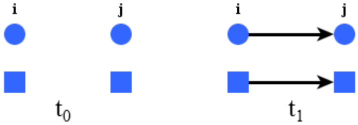
Questions

- Is there any homophily in friendship formation with respect to gender?

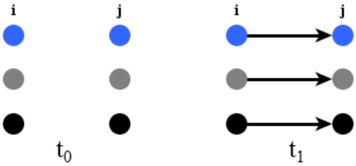


Questions

- Is there any homophily in friendship formation with respect to gender?



- Is there any homophily in friendship formation with respect to smoking behavior?



Solution

Stochastic actor-oriented model (SAOM)

Aim

Explain network evolution as a result of

- endogenous variables: structural effects (e.g. reciprocity, transitivity, etc.)
- exogenous variables: actor and dyadic covariates (gender, age, ethnicity, etc.)

simultaneously

Background: random variable

Definition

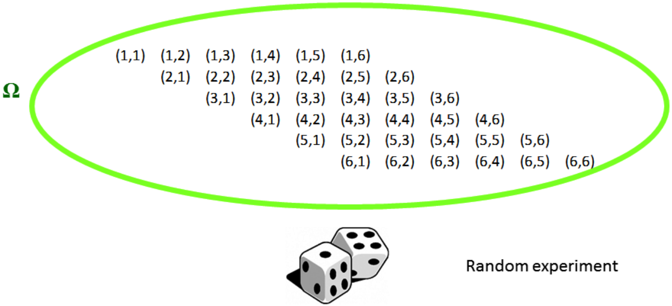
Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Background: random variable

Definition

Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example



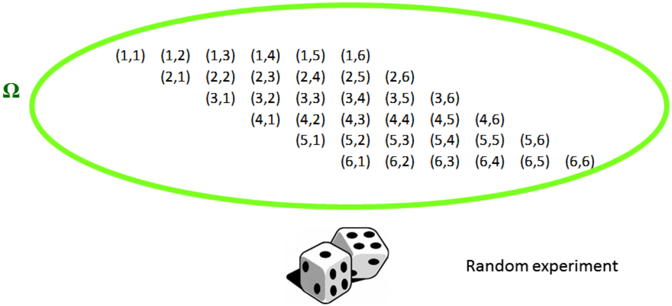
Background: random variable

Definition

Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

X: sum of the two dice



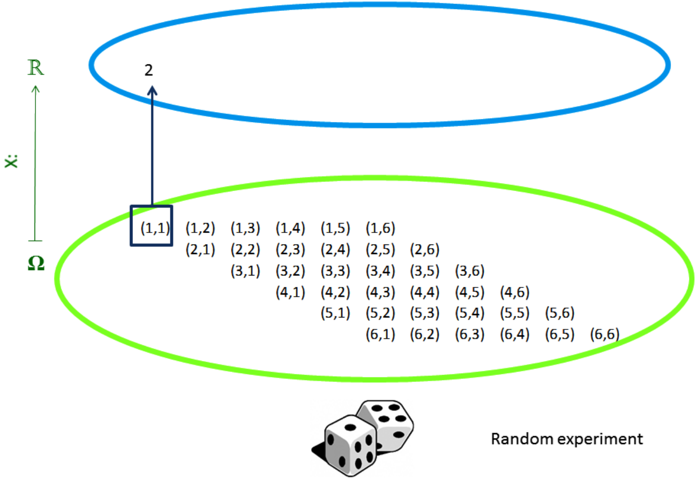
Random experiment

Background: random variable

Definition

Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

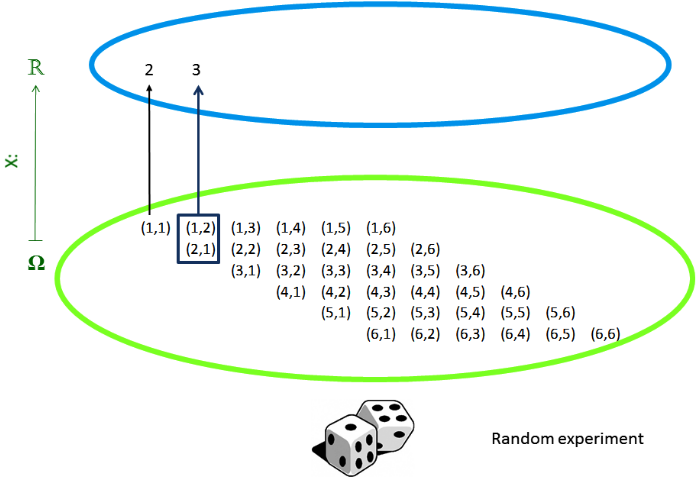


Background: random variable

Definition

Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

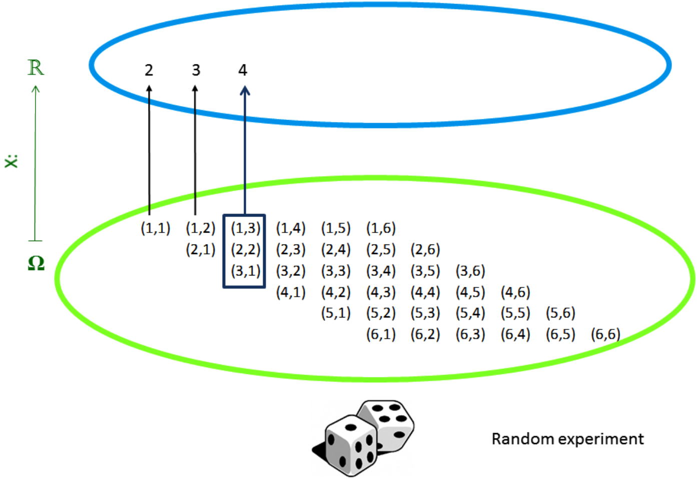


Background: random variable

Definition

Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

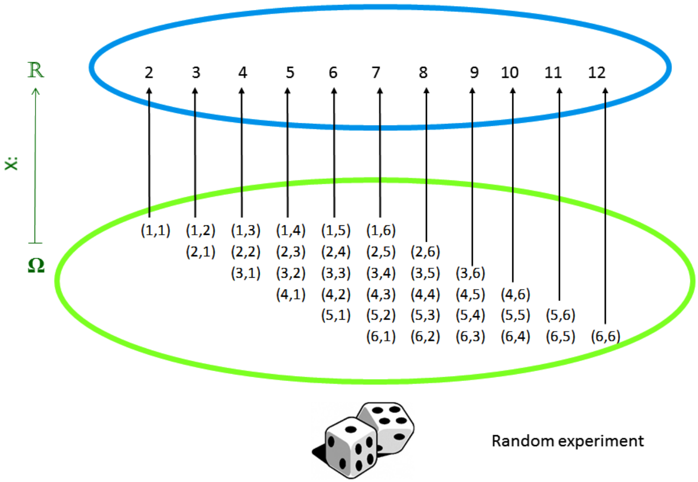


Background: random variable

Definition

Let (Ω, P) be a probability space. A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example



Background: continuous random variable

Definition

A random variable X is called *(absolutely) continuous* if there exists a function $f_X(x) : \mathbb{R} \rightarrow \mathbb{R}^+$ such that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)du \quad \forall x \in \mathbb{R}$$

$F_X(x)$ is the *cumulative distribution function* (c.d.f)

$f_X(x)$ is the *probability density function* (p.d.f)

- $f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$
- $P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f_X(x)dx = 1$

Background: continuous random variable

Definition

A random variable X is called *(absolutely) continuous* if there exists a function $f_X(x) : \mathbb{R} \to \mathbb{R}^+$ such that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)du \quad \forall x \in \mathbb{R}$$

$F_X(x)$ is the *cumulative distribution function* (c.d.f)

$f_X(x)$ is the *probability density function* (p.d.f)

- $f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$
- $P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f_X(x)dx = 1$

Be careful about the word *continuous*!!!

Background: continuous random variable

The p.d.f. $f_X(x)$ allows to compute all the probability statements about X. For instance, the probability that X takes values in $[a, b]$ is

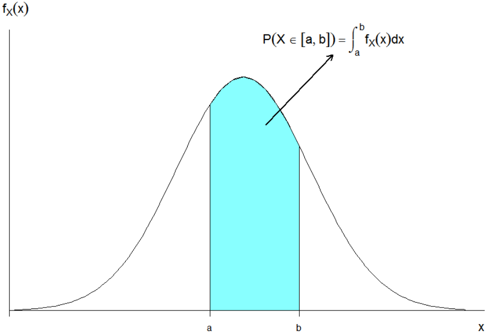
$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

Background: continuous random variable

The p.d.f. $f_X(x)$ allows to compute all the probability statements about X. For instance, the probability that X takes values in $[a, b]$ is

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

Geometrical interpretation



Intuition suggests that

$$P(X = x) = \int_x^x f_X(u)du = 0$$

Thus, we cannot determine a continuous random variable via its "mass function"

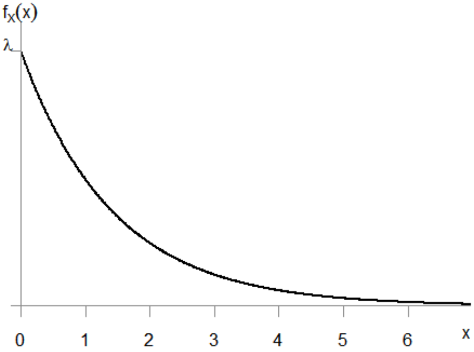
Background: Exponential random variable

Definition

A continuous random variable X whose p.d.f. is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is said to be an *Exponential* random variable with rate $\lambda > 0$.



Background: Exponential random variable

The c.d.f. of X is

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if} \quad x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

In fact

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f_X(x)dx = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}$$

Background: Exponential random variable

The Exponential r.v. has an important property: the memoryless property

Definition
A r.v. X is *memoryless* if

$$P(X > s + t | X > t) = P(X > s) \quad \forall s, t > 0$$

It is easy to prove the memoryless property for the Exponential r.v.

Background: Exponential random variable

The Exponential r.v. has an important property: the memoryless property

Definition

A r.v. X is *memoryless* if

$$P(X > s + t | X > t) = P(X > s) \quad \forall s, t > 0$$

It is easy to prove the memoryless property for the Exponential r.v.

Proof.

$$P(X > s + t | X > t) = \frac{P(X > t+s \cap X > t)}{P(X > t)} = \frac{P(X > t+s)}{P(X > t)} = \frac{1 - P(X \leq t+s)}{1 - P(X \leq t)} =$$

$$= \frac{1 - 1 + e^{-\lambda(t+s)}}{1 - 1 + e^{-\lambda t}} = e^{-\lambda s} = P(X > s)$$

$\square$

Background: stochastic process

Definition

A *stochastic process* is a collection $\{X(t), t \in T\}$ of random variables $X(t) : \Omega \longmapsto \mathbb{R}$

$$\forall t \in T \mapsto X(t) : \Omega \to \mathbb{R}$$

T = index set (usually interpreted as time)
$\mathcal{S}$ = state space

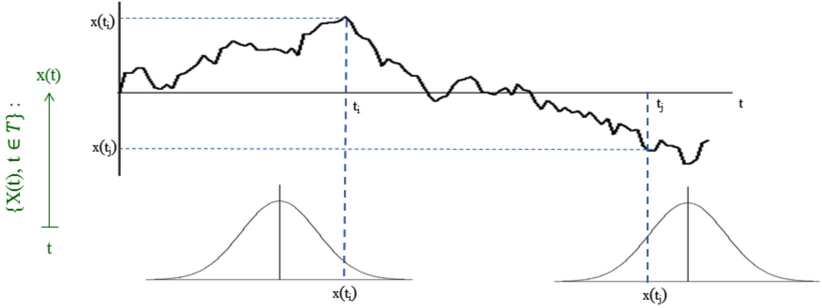
Background: stochastic process

Definition

A *stochastic process* is a collection $\{X(t), t \in T\}$ of random variables $X(t) : \Omega \longmapsto \mathbb{R}$

$$\forall t \in T \mapsto X(t) : \Omega \to \mathbb{R}$$

T = index set (usually interpreted as time)
$\mathcal{S}$ = state space



Background: stochastic process

Example

$X(t) =$ the outcome of flipping a coin

Background: stochastic process

Example

$X(t) =$ the outcome of flipping a coin

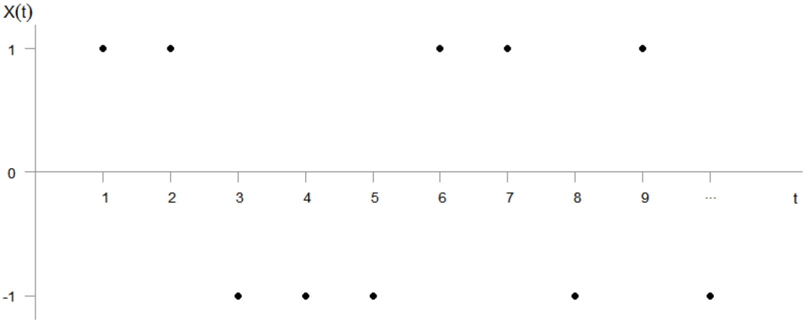
$\mathcal{S} = \{-1, 1\}$, where $-1 =$tail $1 =$head
$T = \{1, 2, \cdots\}$

Background: stochastic process

Example

$X(t) =$ the outcome of flipping a coin

$\mathcal{S} = \{-1, 1\}$, where $-1 =$tail $1 =$head
$T = \{1, 2, \cdots\}$

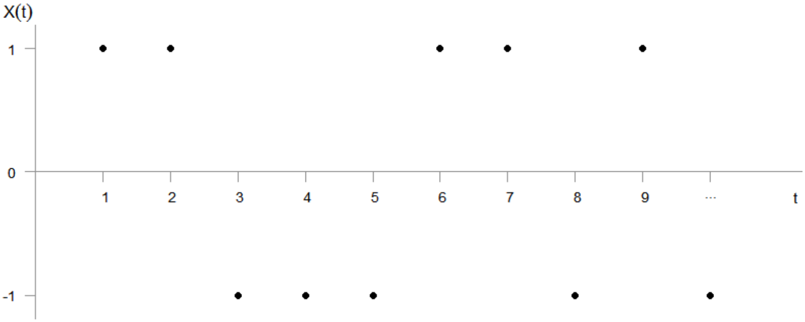


Background: stochastic process

Example

$X(t)$ = the outcome of flipping a coin

$\mathcal{S} = \{-1, 1\}$, where -1 =tail 1 =head
$T = \{1, 2, \cdots\}$



$\{X(t), t \in T\}$ is a discrete-time stochastic process with a discrete (or finite) state space

Background: stochastic process

Example

$X(t) =$ the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

Background: stochastic process

Example

$X(t) =$ the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

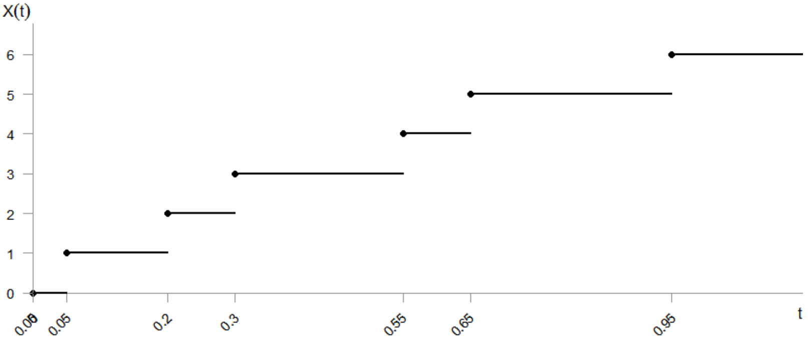
$\mathcal{S} = \{0, 1, 2, \cdots\}$
$T = [0, 12]$

Background: stochastic process

Example

$X(t)$ = the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

$\mathcal{S} = \{0, 1, 2, \cdots\}$
$T = [0, 12]$

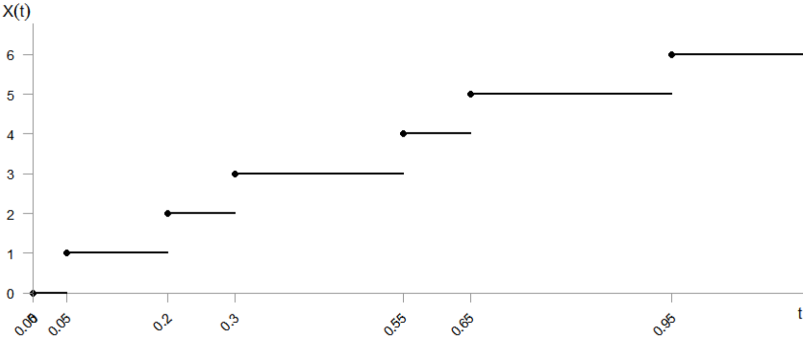


Background: stochastic process

Example

$X(t)$ = the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

$\mathcal{S} = \{0, 1, 2, \cdots\}$
$T = [0, 12]$



$\{X(t), t \in T\}$ is a continuous-time stochastic process with a discrete (or finite) state space

Background: stochastic process

Example

$X(t) =$ the amount of money that a Gambler has after each card games

Background: stochastic process

Example

$X(t) =$ the amount of money that a Gambler has after each card games

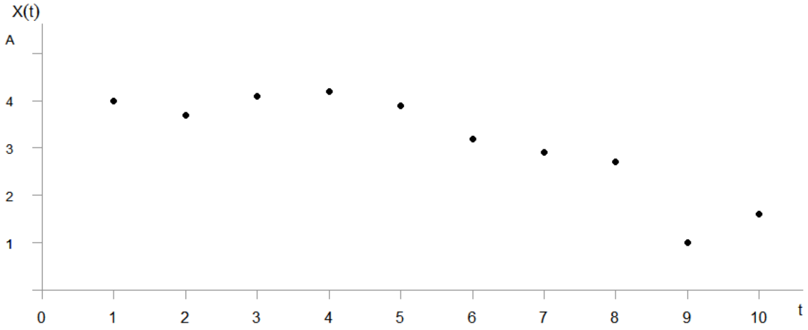
$\mathcal{S} = [0, A]$, $A =$maximum amount
$T = \{1, 2, \cdots\}$

Background: stochastic process

Example

$X(t)$ = the amount of money that a Gambler has after each card games

$\mathcal{S} = [0, A]$, A =maximum amount
$T = \{1, 2, \cdots\}$

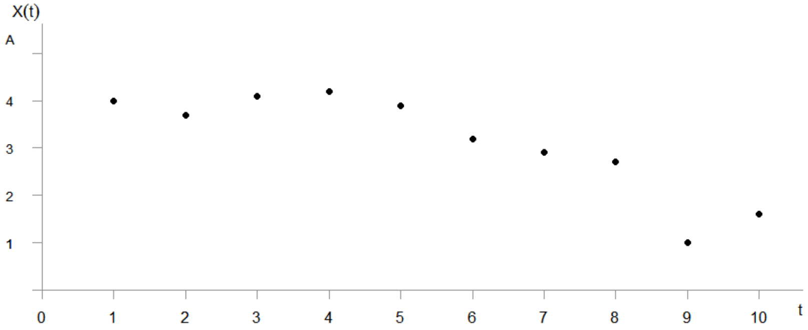


Background: stochastic process

Example

$X(t)$ = the amount of money that a Gambler has after each card games

$\mathcal{S} = [0, A]$, A =maximum amount
$T = \{1, 2, \cdots\}$



$\{X(t), t \in T\}$ is a discrete-time stochastic process with a continuous state space

Background: stochastic process

Example

$X(t) =$ the temperature in a room at each time moment

Background: stochastic process

Example

$X(t) =$ the temperature in a room at each time moment

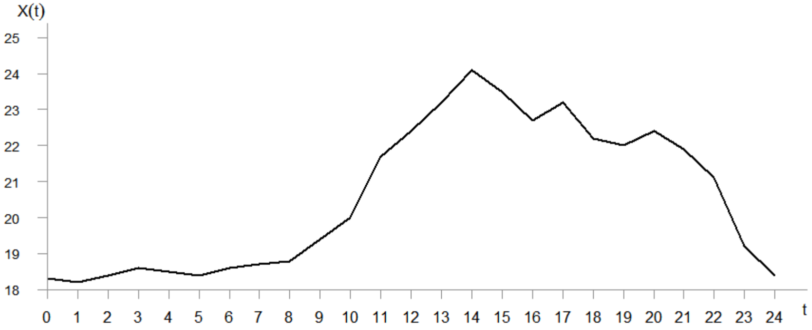
$\mathcal{S} = [18, 25]$
$T = [0, 24]$

Background: stochastic process

Example

$X(t)$ = the temperature in a room at each time moment

$\mathcal{S} = [18, 25]$
$T = [0, 24]$

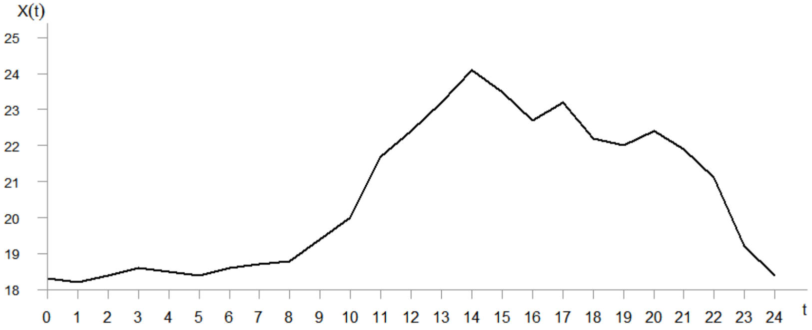


Background: stochastic process

Example

$X(t)$ = the temperature in a room at each time moment

$\mathcal{S} = [18, 25]$
$T = [0, 24]$



$\{X(t), t \in T\}$ is a continuous-time stochastic process with a continuous state space

Background: stochastic process

Different stochastic process can be defined according to the nature of the state space $\mathcal{S}$, the index set T and the dependence relations existing among the random variables $X(t)$.

We will consider **continuous-time Markov chains**

continuous-time = the process evolves in continuous time
Markov = the process have the Markov property
chains = the state space is a finite set

Definition

$\{X(t), t \in T\}$ has the *Markov property* if for any state $x \in \mathcal{S}$, and for any pair of time points $t_i < t_j$

$$P(X(t_j) = x(t_j) | X(t) = x(t) \ \text{ for all } \ t \leq t_i) = P(X(t_j) = x(t_j) | X(t_i) = x(t_i))$$

The future depends on the past and on the present only through the present

Background: continuous-time Markov Chains

Definition

A *continuous-time* Markov chain $\{X_t, t \geq 0\}$ is a **finite state**, **continuous-time** stochastic process having the **Markovian property**

Example

$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t), t \geq 0\} \text{ is a continuous-time Markov chains}$$

Why?

Background: continuous-time Markov Chains

Definition

A *continuous-time* Markov chain $\{X_t, t \geq 0\}$ is a **finite state**, **continuous-time** stochastic process having the **Markovian property**

Example

$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t), t \geq 0\} \text{ is a continuous-time Markov chains}$$

Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$, where B is the total number of goals scored by the player during his career

Background: continuous-time Markov Chains

Definition

A *continuous-time* Markov chain $\{X_t, t \geq 0\}$ is a **finite state**, **continuous-time** stochastic process having the **Markovian property**

Example

$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t), t \geq 0\} \text{ is a continuous-time Markov chains}$$

Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$, where B is the total number of goals scored by the player during his career
2. the time is continuous: time played in official games

Background: continuous-time Markov Chains

Definition

A *continuous-time* Markov chain $\{X_t, t \geq 0\}$ is a **finite state**, **continuous-time** stochastic process having the **Markovian property**

Example

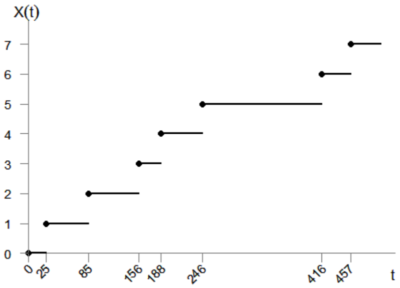
$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t), t \geq 0\} \text{ is a continuous-time Markov chains}$$

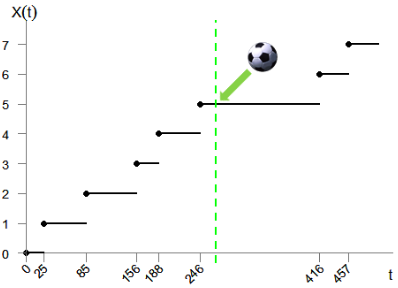
Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$, where B is the total number of goals scored by the player during his career
2. the time is continuous: time played in official games
3. the process $\{X(t), t \geq 0\}$ has the Markov property.

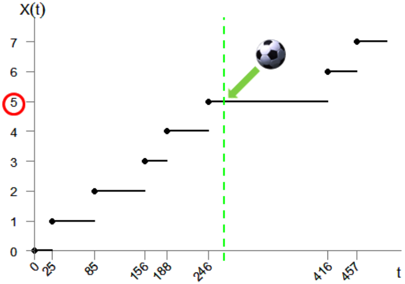
Background: Markov property

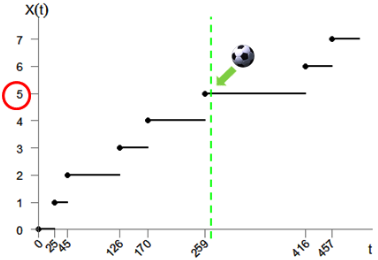


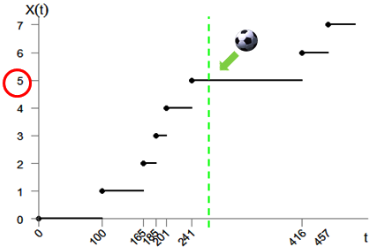
Background: Markov property



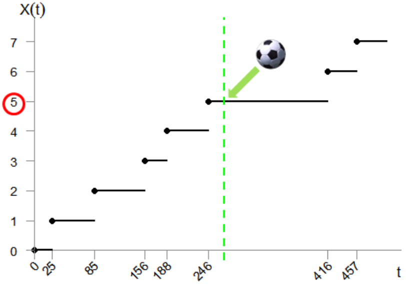
Background: Markov property

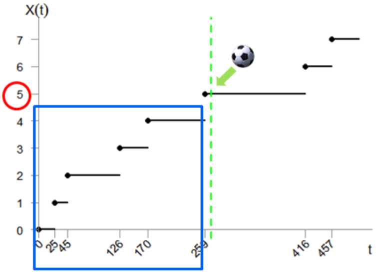


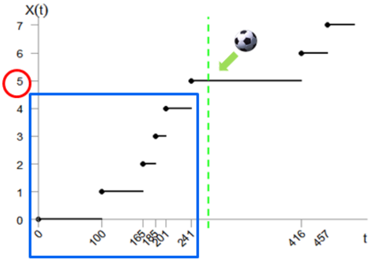




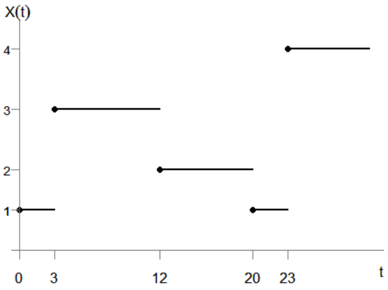
Background: Markov property



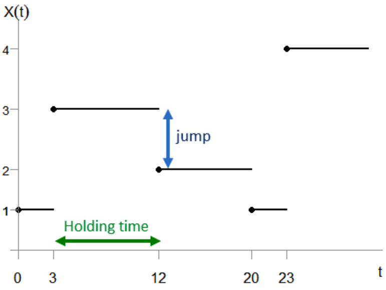




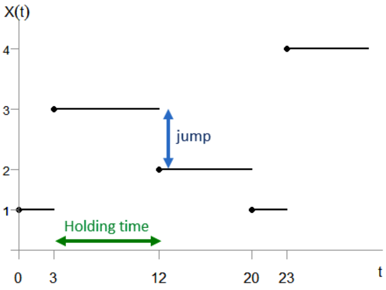
Background: describing a continuous-time Markov chain



Background: describing a continuous-time Markov chain



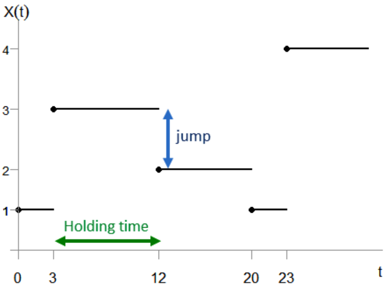
Background: describing a continuous-time Markov chain



Holding times and the jump chain

1. *Holding time*: for each state i, the amount of time we spend in that state is an exponentially distributed random variable, with parameter λ_i

Background: describing a continuous-time Markov chain

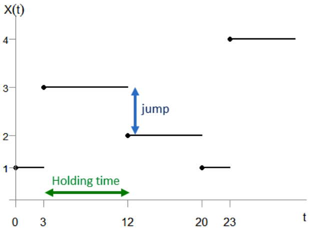


Holding times and the jump chain

1. *Holding time*: for each state i, the amount of time we spend in that state is an exponentially distributed random variable, with parameter λ_i
2. *Jump chain*: is described by a *jump matrix* $P = (p_{ij} : i, j \in \mathcal{S})$ which satisfies the following properties:

$$p_{ij} \geq 0 \qquad \sum_{j \in \mathcal{S}} p_{ij} = 1 \qquad \forall i, j \in S$$

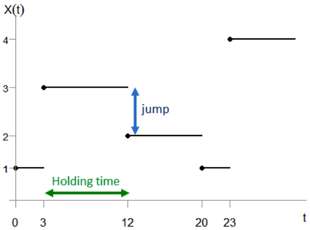
Background: describing a continuous-time Markov chain



P describes the probability of going to state j when we make a jump out a state i.

$$p_{ij} = P(X(t') = j | X(t) = i, \text{given the opportunity to leave state } i), \;\; t' > t$$

Background: describing a continuous-time Markov chain



P describes the probability of going to state j when we make a jump out a state i.

$$p_{ij} = P(X(t') = j | X(t) = i, \text{given the opportunity to leave state } i), \ t' > t$$

Example

$$P = \begin{bmatrix} 0.1 & 0 & 0.6 & 0.3 \\ 0.8 & 0.1 & 0.1 & 0 \\ 0.05 & 0.5 & 0.05 & 0.4 \\ 0.6 & 0.1 & 0.15 & 0.15 \end{bmatrix}$$

Background: describing a continuous-time Markov chain

The rate matrix

The rate matrix $Q = (q_{ij} : i,j \in \mathcal{S})$ defines the rate at which the process jump from the state i to the state j in a short time interval.
It satisfies the following statements:

1. $0 < -q_{ii} < \infty \quad \forall i \in S$
2. $q_{ij} > 0 \quad \forall i \neq j, \ i,j \in S$
3. $\sum_{j \in \mathcal{S}} q_{ij} = 0 \quad \forall i \in S$

The generic entry q_{ij} of this matrix gives the rate of transition from state i to state j is strictly related to the weighting times and the jump matrix. In particular:

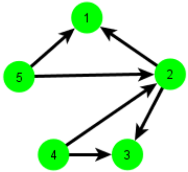
$$q_{ij} = \begin{cases} \lambda_i p_{ij} & \text{if } j \neq i \\ \\ -\lambda_i & \text{if} j = i \end{cases}$$

Outline

Background: adjacency matrix and directed relations

Social network consists of a set of actors $\mathcal{N}$ and the relation $\mathcal{R}$ existing among them

Graph $= G(\mathcal{N}, \mathcal{R})$



Adjacency matrix$=X$

-	0	0	0	0
1	-	1	0	0
0	0	-	0	0
0	1	1	-	0
1	1	0	0	-

In the following network will be represented as an adjacency matrix X

Background: adjacency matrix and directed relations

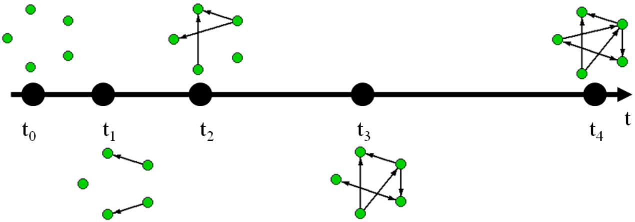
Adjacency matrix X (or digraph):

- $n = \#$ of actors
- the actor index the column and the rows of the $(n \times n)$ matrix X
- $$x_{ij} = \begin{cases} 1 & \text{if } i \text{ is related to } j \ (i \neq j) \\ 0 & \text{otherwise} \end{cases}$$
- Self-relations are not consider so that the diagonal values x_{ii} are meaningless

Directed relation: the existence of a tie from i to j does not imply the existence of a tie from j to i (and vice versa)

$$i \to j \qquad \not\Rightarrow \qquad j \to i$$

Data



Longitudinal (or panel) network data $=$ M (≥ 2) repeated observations on a network

$$x(t_0), x(t_1), \ldots, x(t_m), \ldots, x(t_{M-1}), x(t_M)$$

- set of actors $\mathcal{N} = \{1, 2, \ldots, n\}$
- a non reflexive and directed relation $\mathcal{R}$
- actor covariates (gender, age,social status, ...)

How to model network evolution?

Stochastic actor-oriented model (SAOM)

Aim

Explain network evolution as a result of

- endogenous variables: structural effects (e.g. reciprocity, transitivity, etc.)
- exogenous variables: actor and dyadic covariates (gender, age, ethnicity, etc.)

simultaneously

Model definition: assumptions

Network evolution can be interpreted as the outcome of a **Continuous-time Markov-Chain**...but some assumptions are necessary

Model definition: assumptions

Network evolution can be interpreted as the outcome of a **Continuous-time Markov-Chain**...but some assumptions are necessary

1. *Ties are state*: network ties represent a state with a tendency to endure over time (e.g. friendship, trust, cooperation), rather than a brief event (e.g. telephone calls, e-mail exchanges).

Model definition: assumptions

Network evolution can be interpreted as the outcome of a **Continuous-time Markov-Chain**...but some assumptions are necessary

1. *Ties are state*: network ties represent a state with a tendency to endure over time (e.g. friendship, trust, cooperation), rather than a brief event (e.g. telephone calls, e-mail exchanges).

2. *Distribution of the process*: $\{X(t), t_0 \leq t \leq t_M\}$ is a continuous time Markov Chain defined on $\mathcal{X}$ and $\mathcal{N}$.

Model definition: assumptions

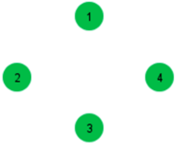
State space: $\mathcal{X}$ is the set of all possible adjacency matrix (digraphs) defined on the set of actors $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \text{ is a finite set}$$

Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrix (digraphs) defined on the set of actors $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \text{ is a finite set}$$

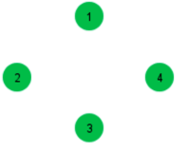


	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-

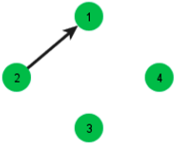
Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrix (digraphs) defined on the set of actors $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \text{ is a finite set}$$



	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-

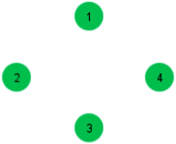


	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

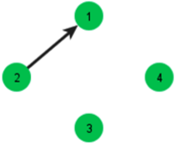
Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrix (digraphs) defined on the set of actors $\mathcal{N}$

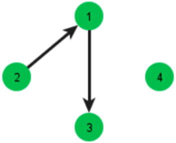
$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \text{ is a finite set}$$



	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

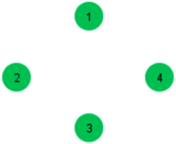


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

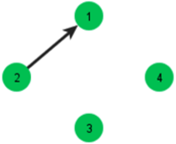
Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrix (digraphs) defined on the set of actors $\mathcal{N}$

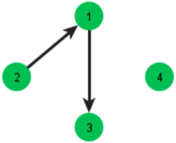
$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \text{ is a finite set}$$



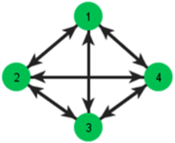
	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-



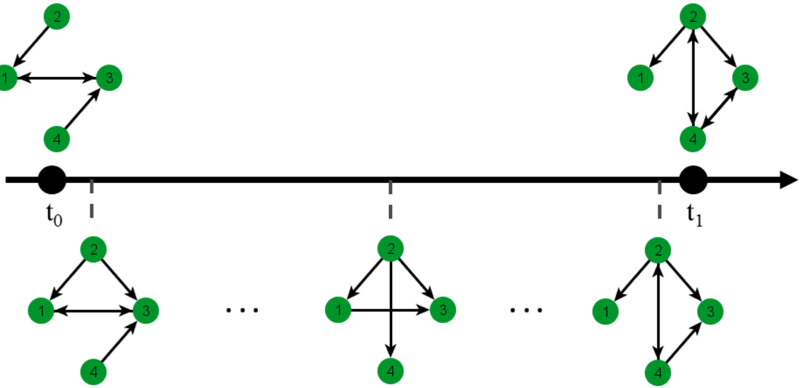
	1	2	3	4
1	-	1	1	1
2	1	-	1	1
3	1	1	-	1
4	1	1	1	-

Model definition: assumptions

Continuous-time process

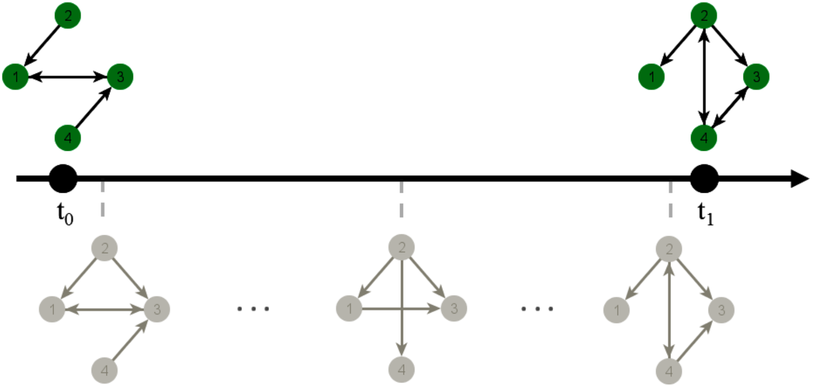
Model definition: assumptions

Continuous-time process



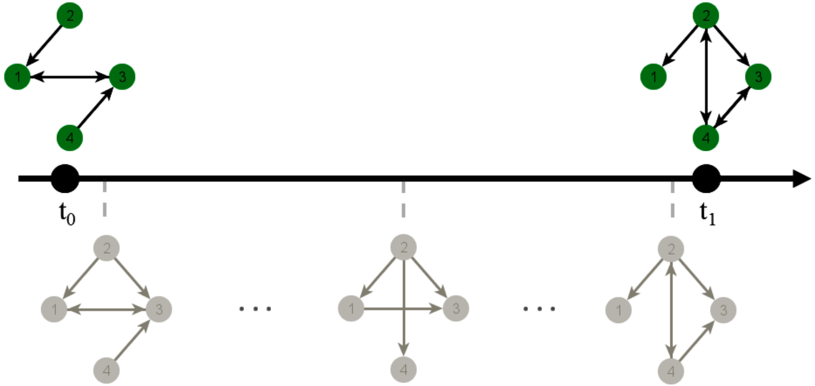
Model definition: assumptions

Continuous-time process



Model definition: assumptions

Continuous-time process



Latent process: the network evolves in continuous-time but we observed it only at discrete time points

Model definition: assumptions

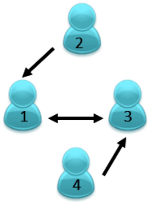
Markov property: the current state of the network determines probabilistically its further evolution

Model definition: assumptions

3. *Opportunity to change*: at a given moment one probabilistically selected actor has the opportunity to change

Model definition: assumptions

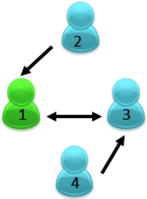
3. *Opportunity to change*: at a given moment one probabilistically selected actor has the opportunity to change



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

3. *Opportunity to change*: at a given moment one probabilistically selected actor has the opportunity to change



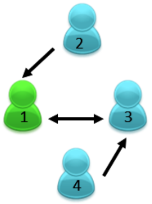
	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

4. *Absence of co-occuRrence*: no more than one tie can change at any given moment t

Model definition: assumptions

4. *Absence of co-occuRrence*: no more than one tie can change at any given moment t

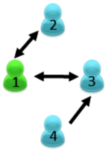


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

x = current state
of the network

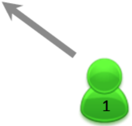
Model definition: assumptions

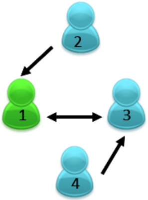
4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 2)$



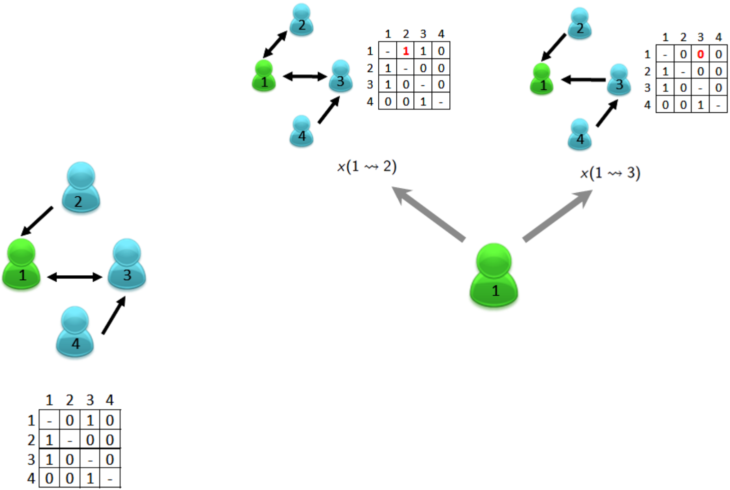


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

x = current state of the network

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

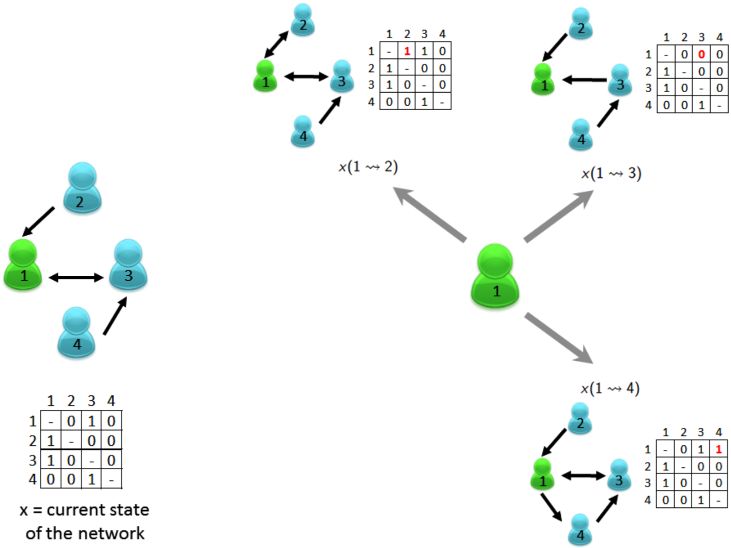
	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

x = current state of the network

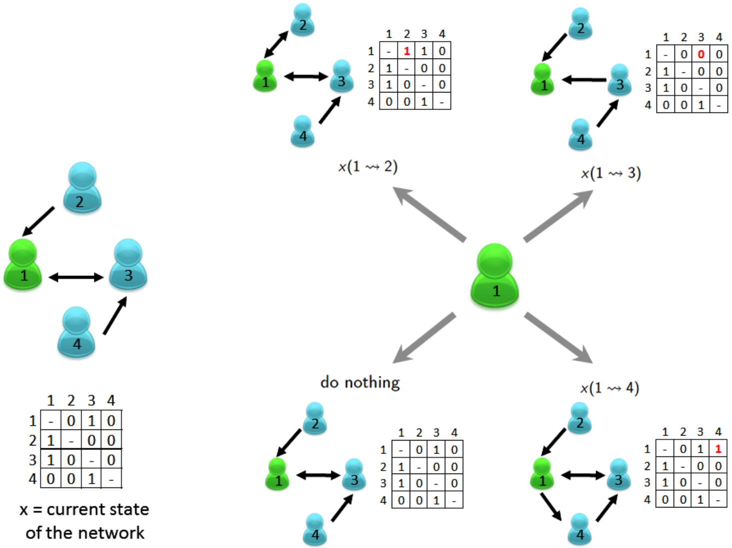
Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



Model definition: assumptions

5. *Actor-oriented perspective*: the actors control their outgoing ties
   - change in ties are made by the actors who send the ties
   - the actor decide to change one of his outgoing ties according to his position in the network, his attributes and the characteristics of the other actors

     Aim: maximize *a utility function*
   - actors have complete knowledge about the network and all the other actors
   - the maximization is based on immediate returns and not on long-run rewarding (myopic actors)

Model definition: assumptions (recap)

1. *Ties are state*

2. *The evolution process is a continuous-time Markov chain*

3. *At a given moment t one probabilistically selected actor has the opportunity to change*

4. *No more than one tie can change at any given moment t*

5. *Actor-oriented perspective*

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps: at one randomly determined moment t, one probabilistically selected actor i has the opportunity to change one of his outgoing ties x_{ij}

Model definition

Consequences of the assumptions
The evolution process can be decomposed into micro-steps: at one randomly determined moment t, one probabilistically selected actor i has the opportunity to change one of his outgoing ties x_{ij}

Micro-step	Continuous-time Markov chain
- the time at which i had the opportunity to change	- the waiting time until the next opportunity for a change made by an actor i (holding time)

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps: at one randomly determined moment t, one probabilistically selected actor i has the opportunity to change one of his outgoing ties x_{ij}

Micro-step	Continuous-time Markov chain
- the time at which i had the opportunity to change	- the waiting time until the next opportunity for a change made by an actor i (holding time)
- the precise change i made	- the probability of changing the link x_{ij} given the opportunity for changing (jump chain)

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps: at one randomly determined moment t, one probabilistically selected actor i has the opportunity to change one of his outgoing ties x_{ij}

Micro-step	Continuous-time Markov chain
- the time at which i had the opportunity to change	- the waiting time until the next opportunity for a change made by an actor i (holding time)
- the precise change i made	- the probability of changing the link x_{ij} given the opportunity for changing (jump chain)

The distribution of the waiting time and the transition matrix of the jump chain are modeled by the *rate function* and the *objective function* respectively.

Model definition: rate function

How fast is the opportunity for changing?

Waiting time between opportunities of change for actor $i \sim Exp(\lambda_i)$
$\Rightarrow$ λ_i is the expected frequency of changes by actor i between observations

Simple specification: all actors have the same rate of change λ

$$P(i \text{ has the opportunity of change}) = \frac{1}{n} \quad \forall i \in \mathcal{N}$$

More complex specification:

- actors may change their ties at different frequencies $\lambda_i(\alpha, x)$
- The parameter of the Exponential distribution is a function of the current state of the network x and the vector of parameter α

$$P(i \text{ has the opportunity of change}) = \frac{\lambda_i(\alpha, x)}{\lambda(\alpha, x)}$$

where

$$\lambda(\alpha, x) = \sum_{i=1}^{n} \lambda_i(\alpha, x)$$

Model definition: objective function

Which tie is changed?

Change a tie means turning it into its opposite:

$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

Model definition: objective function

Which tie is changed?

Change a tie means turning it into its opposite:

$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

Given that i has the opportunity to change:

Possible choices of i	Possible reachable states
$n-1$ changes	$n-1$ networks $x(i \rightsquigarrow j)$

Model definition: objective function

Which tie is changed?

Change a tie means turning it into its opposite:

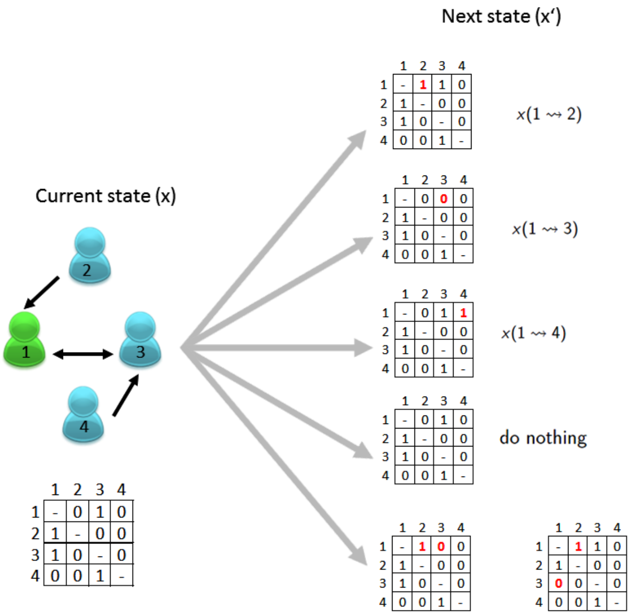
$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

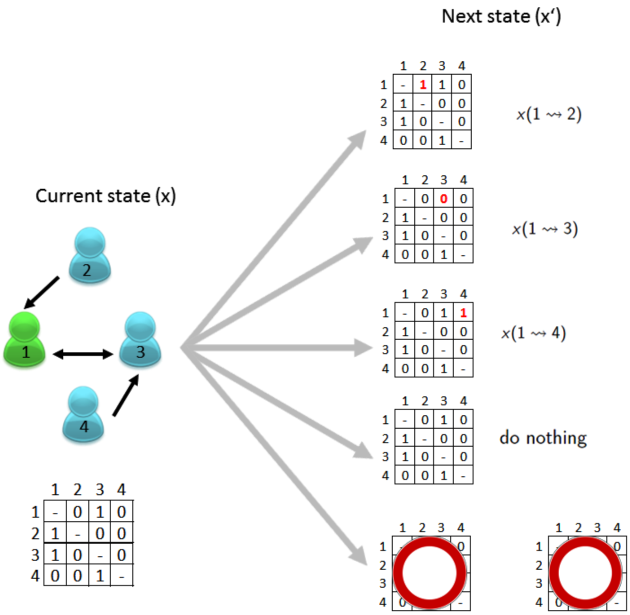
Given that i has the opportunity to change:

Possible choices of i	Possible reachable states
$n - 1$ changes	$n - 1$ networks $x(i \rightsquigarrow j)$
1 non-change	1 network equal to x

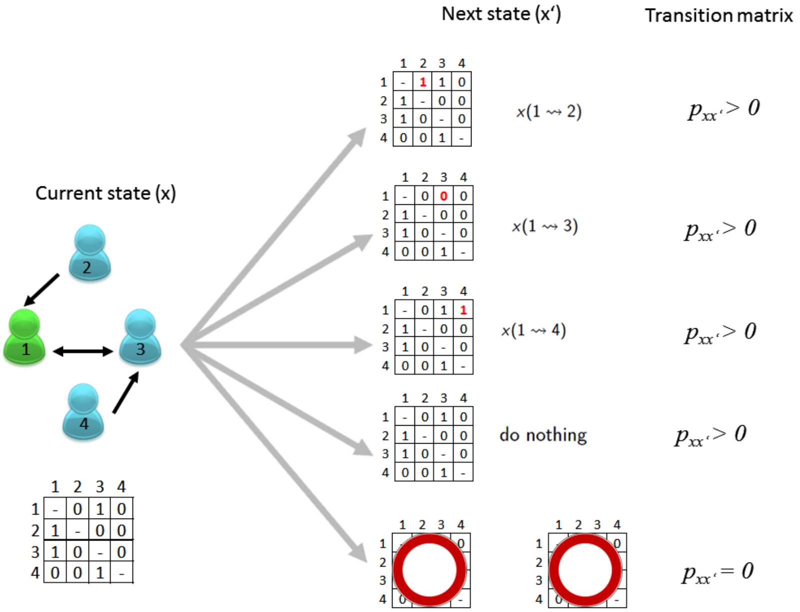
Model definition: objective function



Model definition: objective function



Model definition: objective function



Model definition: objective function

To define the non-zero transition probabilities we assume that actors change their ties in order to maximize a utility function.

The Objective function is defined as a linear combination

$$f_i(\beta, x(i \rightsquigarrow j)) = \sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow j)) + U_i(t, x, j)$$

- $s_{ik}(x(i \rightsquigarrow j))$ are effects
- β_k are statistical parameters
- $U_i(t, x, j)$ is a random utility term

For a suitable choice of the distribution of $U_i(t, x, j)$:

$$p_{ij} = \frac{exp\left(\sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow j))\right)}{\sum_{h=1}^{n} exp\left(\sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow h))\right)}$$

Observation: p_{ii} is interpreted as the probability of not changing

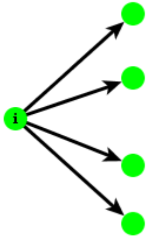
Objective function specification

Simplified notation: $x' = x(i \rightsquigarrow j)$.

Endogenous effects = *dependent on the network structures*

- Outdegree effect

$$s_{i\_out}(x') = \sum_j x'_{ij}$$



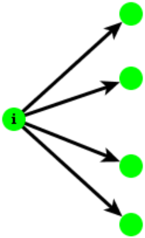
Objective function specification

Simplified notation: $x' = x(i \rightsquigarrow j)$.

Endogenous effects = *dependent on the network structures*

- Outdegree effect

$$s_{i\_out}(x') = \sum_j x'_{ij}$$



- Reciprocity effect

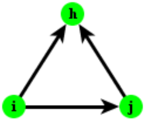
$$s_{i\_rec}(x') = \sum_j x'_{ij} x'_{ji}$$

Objective function specification

Endogenous effects = *dependent on the network structures*

- Transitive effect

$$s_{i\_trans}(x') = \sum_{j,h} x'_{ij} x'_{ih} x'_{jh}$$

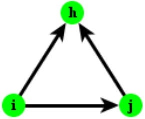


Objective function specification

Endogenous effects = *dependent on the network structures*

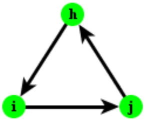
- Transitive effect

$$s_{i\_trans}(x') = \sum_{j,h} x'_{ij} x'_{ih} x'_{jh}$$



- three cycle-effect

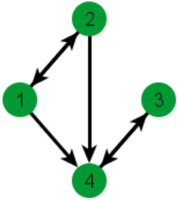
$$s_{i\_cyc}(x') = \sum_{j,h} x'_{ij} x'_{jh} x'_{hi}$$



Objective function specification

Example

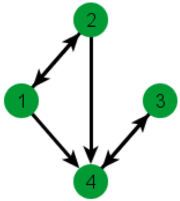
$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

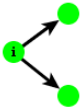
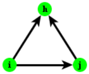


Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

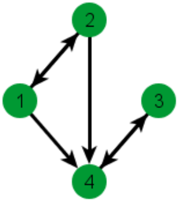


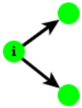
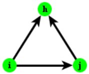
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$			
$1 \to 2$			
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

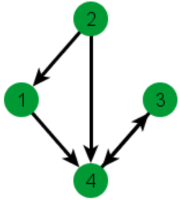


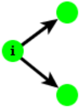
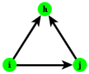
$i \rightarrow j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \rightarrow 1$	2	1	1
$1 \rightarrow 2$			
$1 \rightarrow 3$			
$1 \rightarrow 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

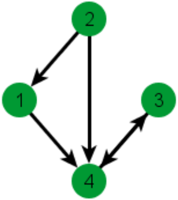


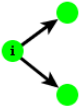
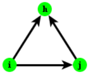
$i \rightarrow j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \rightarrow 1$	2	1	1
$1 \rightarrow 2$			
$1 \rightarrow 3$			
$1 \rightarrow 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

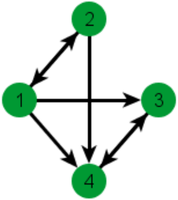


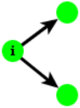
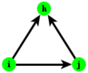
	s_{i_out}	s_{i_rec}	s_{i_trans}
$i \to j$			
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

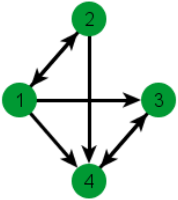


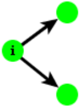
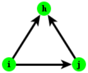
	s_{i_out}	s_{i_rec}	s_{i_trans}
$i \to j$			
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

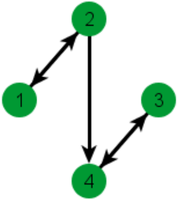


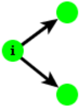
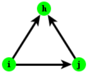
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$	3	1	3
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

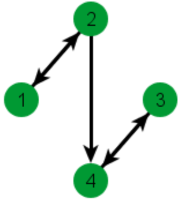


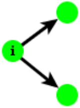
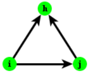
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$	3	1	3
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

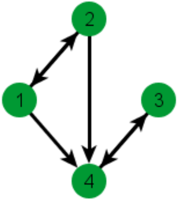


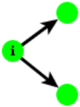
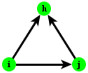
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$	3	1	3
$1 \to 4$	1	1	0

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

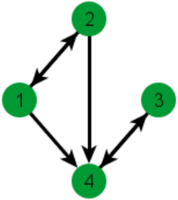


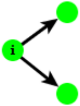
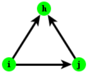
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 	f_i
$1 \to 1$	2	1	1	-1.75
$1 \to 2$	1	0	0	-1
$1 \to 3$	3	1	3	-3.25
$1 \to 4$	1	1	0	-0.5

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$



$i \rightarrow j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 	f_i
$1 \rightarrow 1$	2	1	1	-1.75
$1 \rightarrow 2$	1	0	0	-1
$1 \rightarrow 3$	3	1	3	-3.25
$1 \rightarrow 4$	1	1	0	-0.5

$$p_{11} = 0.146 \qquad p_{12} = 0.310 \qquad p_{13} = 0.033 \qquad p_{14} = 0.511$$

Objective function specification

Exogenous effects = *related to actor's attributes*

Example

- Friendship among pupils:

    Smoking: non, occasional, regular

    Gender: boys, girls

- Trade/Trust (Alliances) among countries:

    Geographical area: Europe, Asia, North-America,...

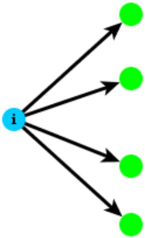
    Worlds: first, Second, Third, Fourth

Objective function specification

Exogenous effects = *related to actor's attributes*

- covariate-related activity

$$s_{i\_cact}(x) = \sum_j x_{ij} z_i$$



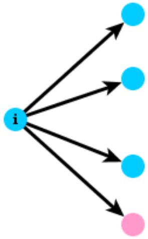
Objective function specification

Exogenous effects = *related to actor's attributes*

- covariate-related similarity

$$s_{i\_csim}(x) = \sum_j x_{ij}\left(1 - \frac{|z_i - z_j|}{R_Z}\right)$$

where R_Z is the range of Z.



Objective function specification

Which effects must be included in the objective function

Objective function specification

Which effects must be included in the objective function

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Objective function specification

Which effects must be included in the objective function

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Example

Friendship: sociological theory suggests that

Theory		Effect
the friend of my friend is also my friend	$\Rightarrow$	transitive effect

Objective function specification

Which effects must be included in the objective function

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Example

Friendship: sociological theory suggests that

Theory		Effect
the friend of my friend is also my friend	$\Rightarrow$	transitive effect
girls trust girls boys trust boys	$\Rightarrow$	covariate-related similarity

Parameter interpretation (part 1)

It is assumed that:

1. the frequencies at which actors have the opportunity to make a change depends on time $= \lambda$ is not constant over time

$$M \text{ time points} \Longrightarrow \text{we must specify } M-1 \text{ rate functions}$$

$$\lambda_1, \cdots, \lambda_{M-1}$$

Parameter interpretation (part 1)

It is assumed that:

1. the frequencies at which actors have the opportunity to make a change depends on time = λ is not constant over time

$$M \text{ time points} \Longrightarrow \text{we must specify } M-1 \text{ rate functions}$$

$$\lambda_1, \cdots, \lambda_{M-1}$$

2. the preferences that drive the choice of the actors have the same intensities over time

$$\beta_1, \cdots, \beta_K$$

are constant over time

Consequence: the number of parameters of the SAOM is equal to $M-1+K$

Parameter interpretation (part 1)

How to interpret the parameter of the SAOM?

Parameter interpretation (part 1)

How to interpret the parameter of the SAOM?

Parameter interpretation (part 1)

How to interpret the parameter of the SAOM?

- The parameter λ is the expected number of opportunities to change for each actor between two consecutive time points.

Parameter interpretation (part 1)

How to interpret the parameter of the SAOM?

- The parameter λ is the expected number of opportunities to change for each actor between two consecutive time points.

- The parameter β_k quantifies the role of each effect in the network evolution.

  $\beta_k = 0$ if the corresponding effect plays no role in the network dynamics

  $\beta_k > 0$ then there is higher probability of moving into networks where the corresponding effect is higher

  $\beta_k < 0$ there is higher probability of moving into networks where the corresponding effect is lower

Simulating network evolution

Reproducing a possible series of micro-steps between t_0 and t_1 according to fixed parameter value and the network $x(t_0)$.

t = the time
dt = the holding time between consecutive changes

Algorithm 1: Network evolution

Input: $x(t_0)$, λ,β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
    $dt \sim Exp(n\lambda)$
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
        $x = x(i \rightsquigarrow j)$
    else
        $x = x$
    $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

Simulating network evolution: simulation and unconditional estimation

There are two different stopping rules for the simulations of the network evolution:

1. *Unconditional* simulation:

   the simulations of the network evolution in each time period carry on until a predetermined time length has elapsed (usually until $t = 1$).

Simulating network evolution: simulation and unconditional estimation

There are two different stopping rules for the simulations of the network evolution:

1. *Unconditional* simulation:

   the simulations of the network evolution in each time period carry on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:

   simulations run on until the number of different entries between $x(t_0)$ and the simulated network $x^{sim}(t_1)$ is equal to the number of entries that differ between $x(t_0)$ and $x(t_1)$

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}^{obs}(t_1) - X_{ij}(t_0) \right| = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}^{sim}(t_1) - X_{ij}(t_0) \right|$$

   This criterion can be generalized conditioning on any other explanatory variable.

Background: population and sample moments

The formulation of the SAOM i depends on $M-1+K$ statistical parameters

$$\theta = (\lambda_1, \cdots, \lambda_{M-1}, \beta_1, \cdots, \beta_K)$$

Aim: estimate θ

Different estimation methods:

- the Method of Moments (MoM)
- the Maximum Likelihood Estimation (MLE)

Background: population and sample moments

Definition

$f_X(x;\theta)$ = probability distribution
$\theta \in \Theta \subset \mathbb{R}^p$ = p-dimensional parameter
$X_1, X_2, \cdots, X_n$ = random sample from the probability distribution $f_X(x;\theta)$

Background: population and sample moments

Definition

$f_X(x;\theta)$ = probability distribution
$\theta \in \Theta \subset \mathbb{R}^p$ = p-dimensional parameter
$X_1, X_2, \cdots, X_n$ = random sample from the probability distribution $f_X(x;\theta)$

The k-th *population moment* is:

$$E[X^k] = \sum_x x^k f_X(x) \qquad \text{(for the discrete case)}$$

$$E[X^k] = \int_{-\infty}^{+\infty} x^k f_X(x) \qquad \text{(for the continuous case)}$$

Background: population and sample moments

Definition

$f_X(x;\theta)$ = probability distribution
$\theta \in \Theta \subset \mathbb{R}^p$ = p-dimensional parameter
$X_1, X_2, \cdots, X_n$ = random sample from the probability distribution $f_X(x;\theta)$

The k-th *population moment* is:

$$E[X^k] = \sum_x x^k f_X(x) \qquad \text{(for the discrete case)}$$

$$E[X^k] = \int_{-\infty}^{+\infty} x^k f_X(x) \qquad \text{(for the continuous case)}$$

The corresponding k-th *sample moment* is

$$\mu_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Background: Method of Moments (MoM)

To estimate θ, one can observe that the theoretical moments of a certain distribution usually depend on the statistical parameters

Definition

The method of moment estimators are found by equating the first p population moments to the first p sample moments

$$E[X^1] = \mu_1$$

$$E[X^2] = \mu_2$$

$$\dots$$

$$E[X^p] = \mu_p$$

and solving the resulting equations for the unknown parameters.

Background: Method of Moments (MoM)

Example

The time to failure of an electronic module used in an automobile engine controller is tested at an elevated temperature to accelerate the failure mechanism.
The time to failure is exponentially distributed with parameter λ.

To estimate the rate parameter λ, eight units are randomly selected and tested, resulting in the following failure time (in hours):

$$x_1 = 12.1 \qquad x_2 = 5.7 \qquad x_3 = 17.8 \qquad x_4 = 16.5$$

$$x_5 = 31.6 \qquad x_6 = 7.7 \qquad x_7 = 11.9 \qquad x_8 = 22.7$$

What is the estimate for λ according to the observed data and the the MoM?

Background: Method of Moments (MoM)

Example

The first population moment of the Exponential random variable is

$$E[X] = \frac{1}{\lambda}$$

and the corresponding sample moment is

$$\mu_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Background: Method of Moments (MoM)

Example

The first population moment of the Exponential random variable is

$$E[X] = \frac{1}{\lambda}$$

and the corresponding sample moment is

$$\mu_1 = \frac{1}{n} \sum_{i=1}^{n} X_i$$

According to the MoM, the estimator for the parameter λ is:

$$\frac{1}{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad \Leftrightarrow \qquad \lambda = \frac{n}{\sum_{i=1}^{n} X_i}$$

and the corresponding estimate is

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{8}{126} = 0.063$$

Background: Method of Moments (MoM)

The principle of the MoM can be easily generalized to any set of p functions $s_k(X), \quad k = 1, \ldots, p$

Population moment

$$E[s_k(X)] = \sum_x s_k(x) f_X(x) \qquad \text{(for the discrete case)}$$

$$E[s_k(X)] = \int_{-\infty}^{+\infty} s_k(x) f_X(x) \qquad \text{(for the continuous case)}$$

Sample moment

$$\gamma_k = \frac{1}{n} \sum_{i=1}^{n} s_k(X_i)$$

$s_k(X)$ are called statistics

They **must be sensitive** to the parameter θ, i.e. higher values of θ lead to higher values of $s(X)$.

Background: Method of Moments (MoM)

Remark

- An *estimator* is a function of the sample, e.g. $\frac{n}{\displaystyle\sum_{i=1}^{n} X_i}$
- An *estimate* is the realized value of an estimator, e.g. $\frac{n}{\displaystyle\sum_{i=1}^{n} x_i}$
- The estimate of a parameter varies according to the selected sample. Thus, we usually associate to an estimator its standard error.

Background: Method of Moments (MoM)

Example

We assume to randomly select and test other eight electronic modules, resulting in the following failure time (in hours):

$$x_1 = 9.5 \qquad x_2 = 7.2 \qquad x_3 = 13.4 \qquad x_4 = 10.2$$

$$x_5 = 15.0 \qquad x_6 = 16.3 \qquad x_7 = 13.9 \qquad x_8 = 34.5$$

The new estimate for the parameter λ is

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{8}{120} = 0.067$$

This value is close to the previous but it is not the same!

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1, \cdots, \lambda_{M-1}, \beta_1, \cdots, \beta_K)$$

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1, \cdots, \lambda_{M-1}, \beta_1, \cdots, \beta_K)$$

In practice:

- find $M - 1 + K$ statistics
- set the theoretical expected value of each statistic equal to its sample counterpart
- solve the resulting system of equations with respect to θ.

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1, \cdots, \lambda_{M-1}, \beta_1, \cdots, \beta_K)$$

In practice:

- find $M - 1 + K$ statistics
- set the theoretical expected value of each statistic equal to its sample counterpart
- solve the resulting system of equations with respect to θ.

Which statistics are suitable for estimating θ?

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1, \cdots, \lambda_{M-1}, \beta_1, \cdots, \beta_K)$$

In practice:

- find $M-1+K$ statistics
- set the theoretical expected value of each statistic equal to its sample counterpart
- solve the resulting system of equations with respect to θ.

Which statistics are suitable for estimating θ?

For simplicity, let us assume to have observed a network at two time points t_0 and t_1 and to condition the estimation on the first observation $x(t_0)$

Estimating the parameter of the SAOM using the MoM

- The rate parameter λ describes the frequencies at which changes happen.

$$s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

Estimating the parameter of the SAOM using the MoM

- The rate parameter λ describes the frequencies at which changes happen.

$$s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
s_λ	94	135	171

$\Rightarrow$ a high value of λ leads to a high value of $s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0))$

Estimating the parameter of the SAOM using the MoM

- The parameter β_k quantifies the role played by each effect in the network evolution.

$$s_k(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{ik}(X(t_1))$$

Estimating the parameter of the SAOM using the MoM

- The parameter β_k quantifies the role played by each effect in the network evolution.

$$s_k(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{ik}(X(t_1))$$

E.g.: let us consider the parameter β_{out}.
The corresponding statistic is

$$s_{out}(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{i\_out}(X(t_1)) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}(t_1)$$

	$\beta_{out} = -2.5$	$\beta_{out} = -2$	$\beta_{out} = -1.5$
s_{out}	195	214	234

$\Rightarrow$ a high value of β_{out} leads to a high value of $s_{out}(X(t_1)|X(t_0) = x(t_0))$

Estimating the parameter of the SAOM using the MoM

Generalizing to $M-1$ periods:

- Statistics for the rate function parameters

$$s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

...

$$s_{\lambda_{M-1}}(X(t_M), X(t_{M-1})|X(t_{M-1}) = x(t_{M-1})) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_M) - X_{ij}(t_{M-1}) \right|$$

Estimating the parameter of the SAOM using the MoM

Generalizing to $M-1$ periods:

- Statistics for the rate function parameters

$$s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

$\ldots$

$$s_{\lambda_{M-1}}(X(t_M), X(t_{M-1})|X(t_{M-1}) = x(t_{M-1})) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_M) - X_{ij}(t_{M-1}) \right|$$

- Statistics for the objective function parameters:

$$\sum_{m=1}^{M-1} s_{mk}\,(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{M-1} s_{mk}(X(t_m), X(t_{m-1}))$$

Estimating the parameter of the SAOM using the MoM

The MoM estimator for θ is defined as the solution of the system of $M+K-1$ equations

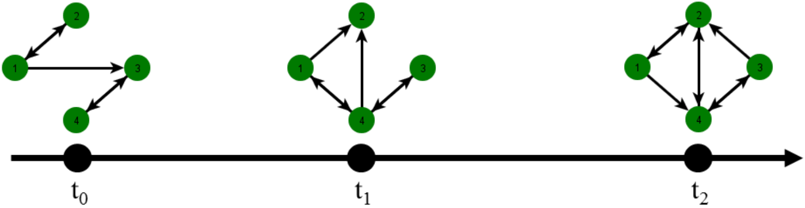
$$\begin{cases} E_\theta\left[s_{\lambda_m}(X(t_m), X(t_{m+1}) | X(t_m) = x(t_m))\right] = s_{\lambda_m}(x(t_1), x(t_0)) \\ E_\theta\left[\sum_{m=1}^{M-1} s_{mk}(X(t_{m+1}) | X(t_m) = x(t_m))\right] = \sum_{m=1}^{M-1} s_{mk}(x(t_{m+1}), x(t_m)) \end{cases}$$

with $m = 1, \ldots, M-1$ and $k = 1, \cdots, K$

Estimating the parameter of the SAOM using the MoM

Example

Let us assume to have observed a network at $M = 3$ time points



	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

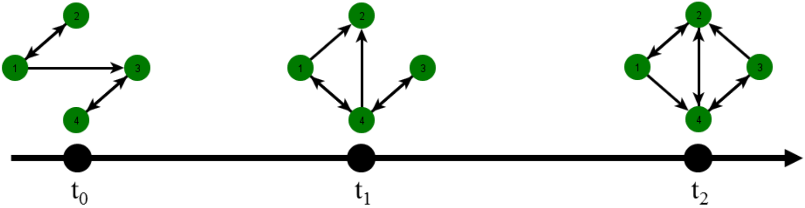
	1	2	3	4
1	-	1	0	1
2	1	-	0	1
3	0	1	-	1
4	0	1	1	-

and to model network evolution considering the outdegree and the reciprocity effects.

Estimating the parameter of the SAOM using the MoM

Example

Let us assume to have observed a network at $M = 3$ time points



	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

	1	2	3	4
1	-	1	0	1
2	1	-	0	1
3	0	1	-	1
4	0	1	1	-

and to model network evolution considering the outdegree and the reciprocity effects.

$$\theta = (\lambda_1, \lambda_2, \beta_{out}, \beta_{rec})$$

Estimating the parameter of the SAOM using the MoM

Example

Statistics:

$$s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{\substack{i,j=1 \\ i \neq j}}^{4} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

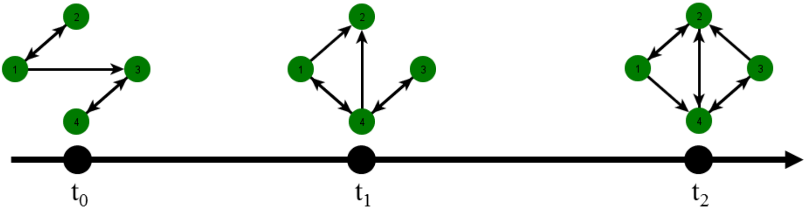
$$s_{\lambda_2}(X(t_2), X(t_1)|X(t_1) = x(t_1)) = \sum_{\substack{i,j=1 \\ i \neq j}}^{4} \left| X_{ij}(t_2) - X_{ij}(t_1) \right|$$

$$\sum_{m=1}^{2} s_{m\_out}(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{2} \sum_{i=1}^{4} \sum_{j=1}^{4} X_{ij}(t_m)$$

$$\sum_{m=1}^{2} s_{m\_rec}(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{2} \sum_{i=1}^{4} \sum_{j=1}^{4} X_{ij}(t_m) X_{ji}(t_m)$$

Estimating the parameter of the SAOM using the MoM

Example



	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

	1	2	3	4
1	-	1	0	1
2	1	-	0	1
3	0	1	-	1
4	0	1	1	-

Observed values of the statistics:

$$s_{\lambda_1} = 5 \qquad s_{\lambda_2} = 4$$

$$\sum_{m=1}^{2} s_{m\_out} = 6 + 8 = 14 \qquad \sum_{m=1}^{2} s_{m\_rec} = 2 + 3 = 5$$

Estimating the parameter of the SAOM using the MoM

Example

We look for the value of θ that satisfies the system:

$$\begin{cases} E_\theta \left[s_{\lambda_1}(X(t_1), X(t_0) | X(t_0) = x(t_0)) \right] = 5 \\ E_\theta \left[s_{\lambda_2}(X(t_2), X(t_1) | X(t_1) = x(t_1)) \right] = 4 \\ E_\theta \left[\sum_{m=1}^{2} s_{m\_out} \left(X(t_m) | X(t_{m-1}) = x(t_{m-1}) \right) \right] = 14 \\ E_\theta \left[\sum_{m=1}^{2} s_{m\_rec} \left(X(t_m) | X(t_{m-1}) = x(t_{m-1}) \right) \right] = 5 \end{cases}$$

Stochastic approximation method for the SAOM

Simplified notation:

- S: $(M-1+K)$-dimensional vector of statistics
- s: $(M-1+K)$-dimensional vector of the observed values of the statistics

Consequently, the system of moment equations can be written as

$$E_\theta[S] = s$$

or equivalently as

$$E_\theta[S-s] = 0$$

Stochastic approximation method for the SAOM

Simplified notation:

- S: $(M-1+K)$-dimensional vector of statistics
- s: $(M-1+K)$-dimensional vector of the observed values of the statistics

Consequently, the system of moment equations can be written as

$$E_\theta[S] = s$$

or equivalently as

$$E_\theta[S - s] = 0$$

Problem: analytical and usual numerical procedures cannot be applied to solve this system, but we can use a stochastic approximation method

Stochastic approximation method for the SAOM

Simplified notation:

- S: $(M-1+K)$-dimensional vector of statistics
- s: $(M-1+K)$-dimensional vector of the observed values of the statistics

Consequently, the system of moment equations can be written as

$$E_\theta[S] = s$$

or equivalently as

$$E_\theta[S-s] = 0$$

Problem: analytical and usual numerical procedures cannot be applied to solve this system, but we can use a stochastic approximation method

Definition

Stochastic approximation methods are a family of iterative stochastic optimization algorithms that attempt to find zeros or extrema of functions which cannot be computed in an analytical way.

Stochastic approximation method for the SAOM

1. The expected values of the statistics are approximated via Monte Carlo methods $\Rightarrow$ stochastic
2. The value of θ is iteratively updated according to the “distance” between the approximated expected values and the observed values.

Stochastic approximation method for the SAOM

1. The expected values of the statistics are approximated via Monte Carlo methods $\Rightarrow$ stochastic
2. The value of θ is iteratively updated according to the "distance" between the approximated expected values and the observed values.

1. Approximation of the expected value $E_\theta[S]$

1. Given $x(t_0)$ and θ, simulate the sequence of the observed networks at time $t_1, \ldots, t_M$ q times. Denote these sequences by

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\cdots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ x^{(q)}(t_M)$$

2. For each sequence compute the value $S^{(l)}$ assumed by S
3. Approximate the expected value by

$$\overline{S} = \frac{1}{q}\sum_{l=1}^{q} S^{(l)}$$

Stochastic approximation method for the SAOM

1. The expected values of the statistics are approximated via Monte Carlo methods $\Rightarrow$ stochastic
2. The value of θ is iteratively updated according to the "distance" between the approximated expected values and the observed values.

1. Approximation of the expected value $E_\theta[S]$

1. Given $x(t_0)$ and θ, simulate the sequence of the observed networks at time $t_1, \ldots, t_M$ q times. Denote these sequences by

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ x^{(q)}(t_M)$$

2. For each sequence compute the value $S^{(l)}$ assumed by S
3. Approximate the expected value by

$$\overline{S} = \frac{1}{q} \sum_{l=1}^{q} S^{(l)} \rightarrow E_\theta[S]$$

Stochastic approximation method for the SAOM

1. Approximation of the expected value $E_\theta[S]$

Example

1. Given $x(t_0)$ and θ, simulate the sequence of the observed networks at time $t_1, \ldots, t_M$ q times. Denote these sequences by

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ x^{(q)}(t_M)$$

Stochastic approximation method for the SAOM

1. Approximation of the expected value $E_\theta[S]$

Example

1. Given $x(t_0)$ and θ, simulate the sequence of the observed networks at time $t_1, \ldots, t_M$ q times. Denote these sequences by

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ x^{(q)}(t_M)$$

2. For each sequence compute the value assumed by the statistics

$$S^{(l)}_{out} = \sum_{m=1}^{M-1} \sum_{i=1}^{n} \sum_{j=1}^{n} x^{(l)}_{ij}(t_m)$$

Stochastic approximation method for the SAOM

1. Approximation of the expected value $E_\theta[S]$

Example

1. Given $x(t_0)$ and θ, simulate the sequence of the observed networks at time $t_1, \dots, t_M$ q times. Denote these sequences by

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \dots,\ x^{(1)}(t_M)$$

$$\dots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \dots,\ x^{(q)}(t_M)$$

2. For each sequence compute the value assumed by the statistics

$$S_{out}^{(l)} = \sum_{m=1}^{M-1} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}^{(l)}(t_m)$$

3. Approximate the expected value by

$$\overline{S}_{out} = \frac{1}{q} \sum_{i=1}^{q} S_{out}^{(l)} \to E_\theta[S_{out}]$$

Stochastic approximation method for the SAOM

2. Updating the value of θ

The value of θ is updated by the Robbins-Monro step:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1}(\overline{S}_i - s)$$

where

- a_i is a sequence of positive numbers such that:

$$\lim_{(i \to \infty)} a_i = 0 \qquad \sum_{i=1}^{\infty} a_i = \infty \qquad \sum_{i=1}^{\infty} a_i^2 < \infty$$

- D denotes the diagonal matrix of the first order derivative matrix of S with respect to θ:

$$D = \frac{\partial}{\partial \theta} E_\theta[S | X(t_0) = x(t_0)]$$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

We look for the value of θ that satisfies the system:

$$\begin{cases} E_\theta\left[s_{\lambda_1}(X(t_1), X(t_0) | X(t_0) = x(t_0))\right] = 5 \\ E_\theta\left[s_{\lambda_2}(X(t_2), X(t_1) | X(t_1) = x(t_1))\right] = 4 \\ E_\theta\left[\sum_{m=1}^{2} s_{m\_out}\left(X(t_m) | X(t_{m-1}) = x(t_{m-1})\right)\right] = 14 \\ E_\theta\left[\sum_{m=1}^{2} s_{m\_rec}\left(X(t_m) | X(t_{m-1}) = x(t_{m-1})\right)\right] = 5 \end{cases}$$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

- Initial guess $\theta_0 = (4.5, 3.2, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to θ_0
- Approximation of the expected values $E_{\theta_0}[\overline{S}]$

$\overline{S}_{\lambda_1} = 6.211$ $\qquad$ $\overline{S}_{\lambda_2} = 4.567$

$\overline{S}_{\beta_{out}} = 13.806$ $\qquad$ $\overline{S}_{\beta_{rec}} = 4.702$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

- Initial guess $\theta_0 = (4.5, 3.2, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to θ_0
- Approximation of the expected values $E_{\theta_0}[\overline{S}]$

$\overline{S}_{\lambda_1} = 6.211$ $\quad\quad$ $\overline{S}_{\lambda_2} = 4.567$

$\overline{S}_{\beta_{out}} = 13.806$ $\quad\quad$ $\overline{S}_{\beta_{rec}} = 4.702$

- Approximation of the moment equation

$\overline{S}_{\lambda_1} - 5 = 1.211$ $\quad\quad$ $\overline{S}_{\lambda_2} - 4 = 0.567$

$\overline{S}_{\beta_{out}} - 14 = -0.194$ $\quad\quad$ $\overline{S}_{\beta_{rec}} - 5 = -0.298$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

- Initial guess $\theta_0 = (4.5, 3.2, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to θ_0
- Approximation of the expected values $E_{\theta_0}[\overline{S}]$

$\overline{S}_{\lambda_1} = 6.211$ $\qquad$ $\overline{S}_{\lambda_2} = 4.567$

$\overline{S}_{\beta_{out}} = 13.806$ $\qquad$ $\overline{S}_{\beta_{rec}} = 4.702$

- Approximation of the moment equation

$\overline{S}_{\lambda_1} - 5 = 1.211$ $\qquad$ $\overline{S}_{\lambda_2} - 4 = 0.567$

$\overline{S}_{\beta_{out}} - 14 = -0.194$ $\qquad$ $\overline{S}_{\beta_{rec}} - 5 = -0.298$

The Robbins-Monro step

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1}(\overline{S}_i - s)$$

suggests how to modify the parameter value to satisfy the moment equation through the difference $(\overline{S}_i - s)$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

- Guess $\theta_1 = (4.1, 2.9, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to θ_1
- Approximation of the expected values $E_{\theta_1}[\overline{S}]$

$\overline{S}_{\lambda_1} = 5.345$ $\qquad$ $\overline{S}_{\lambda_2} = 4.215$

$\overline{S}_{\beta_{out}} = 13.813$ $\qquad$ $\overline{S}_{\beta_{rec}} = 4.759$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

- Guess $\theta_1 = (4.1, 2.9, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to θ_1
- Approximation of the expected values $E_{\theta_1}[\overline{S}]$

$\overline{S}_{\lambda_1} = 5.345$ $\qquad$ $\overline{S}_{\lambda_2} = 4.215$

$\overline{S}_{\beta_{out}} = 13.813$ $\qquad$ $\overline{S}_{\beta_{rec}} = 4.759$

- Approximation of the moment equation

$\overline{S}_{\lambda_1} - 5 = 0.345$ $\qquad$ $\overline{S}_{\lambda_2} - 4 = 0.215$

$\overline{S}_{\beta_{out}} - 14 = -0.187$ $\qquad$ $\overline{S}_{\beta_{rec}} - 5 = -0.241$

2. Updating the value of θ

Example

- Guess $\theta_i = (3.87, 2.56, -0.11, 0.87)$
- Simulate the network evolution 1000 times according to θ_1
- Approximation of the expected values $E_{\theta_i}[\overline{S}]$

$\overline{S}_{\lambda_1} = 5.055$ $\qquad$ $\overline{S}_{\lambda_2} = 4.018$

$\overline{S}_{\beta_{out}} = 14.012$ $\qquad$ $\overline{S}_{\beta_{rec}} = 4.974$

Stochastic approximation method for the SAOM

2. Updating the value of θ

Example

- Guess $\theta_i = (3.87, 2.56, -0.11, 0.87)$
- Simulate the network evolution 1000 times according to θ_1
- Approximation of the expected values $E_{\theta_i}[\overline{S}]$

$\overline{S}_{\lambda_1} = 5.055$ $\qquad$ $\overline{S}_{\lambda_2} = 4.018$

$\overline{S}_{\beta_{out}} = 14.012$ $\qquad$ $\overline{S}_{\beta_{rec}} = 4.974$

- Approximation of the moment equation

$\overline{S}_{\lambda_1} - 5 = 0.055$ $\qquad$ $\overline{S}_{\lambda_2} - 4 = 0.018$

$\overline{S}_{\beta_{out}} - 14 = 0.012$ $\qquad$ $\overline{S}_{\beta_{rec}} - 5 = -0.026$

The Robbins-Monro algorithm

Phase 1: given the network at time t_0 and an initial guess θ_0 for θ a small number q_1 of steps are made to estimate D.

1. Network evolution is simulated from θ_0 and the values S_{i0} are computed

The Robbins-Monro algorithm

Phase 1: given the network at time t_0 and an initial guess θ_0 for θ a small number q_1 of steps are made to estimate D.

1. Network evolution is simulated from θ_0 and the values S_{i0} are computed
2. Network evolution is simulated from $\theta_0 + \epsilon_j e_j$ and the values S_{ij} are computed where
   - e_j is the j-th unit vector
   - $0.1 < \epsilon_j < 1$

The Robbins-Monro algorithm

Phase 1: given the network at time t_0 and an initial guess θ_0 for θ a small number q_1 of steps are made to estimate D.

1. Network evolution is simulated from θ_0 and the values S_{i0} are computed
2. Network evolution is simulated from $\theta_0 + \epsilon_j e_j$ and the values S_{ij} are computed where
   - e_j is the j-th unit vector
   - $0.1 < \epsilon_j < 1$
3. Compute $d_{ij} = \epsilon_j^{-1}(S_{ij} - S_{i0})$

The Robbins-Monro algorithm

Phase 1: given the network at time t_0 and an initial guess θ_0 for θ a small number q_1 of steps are made to estimate D.

1. Network evolution is simulated from θ_0 and the values S_{i0} are computed
2. Network evolution is simulated from $\theta_0 + \epsilon_j e_j$ and the values S_{ij} are computed where
   - e_j is the j-th unit vector
   - $0.1 < \epsilon_j < 1$
3. Compute $d_{ij} = \epsilon_j^{-1}(S_{ij} - S_{i0})$
4. Repeat steps 1. to 3. until $i = q_1$

The Robbins-Monro algorithm

Phase 1: given the network at time t_0 and an initial guess θ_0 for θ a small number q_1 of steps are made to estimate D.

1. Network evolution is simulated from θ_0 and the values S_{i0} are computed
2. Network evolution is simulated from $\theta_0 + \epsilon_j e_j$ and the values S_{ij} are computed where
   - e_j is the j-th unit vector
   - $0.1 < \epsilon_j < 1$
3. Compute $d_{ij} = \epsilon_j^{-1}(S_{ij} - S_{i0})$
4. Repeat steps 1. to 3. until $i = q_1$
5. Estimate $E_\theta[S]$ and D by the Monte Carlo method

$$\overline{S} = \frac{1}{q_1}\sum_{i=1}^{q_1} S_{i0} \qquad \widehat{d}_j = \frac{1}{q_1}\sum_{i=1}^{q_1} d_{ij}$$

The Robbins-Monro algorithm

Phase 1: given the network at time t_0 and an initial guess θ_0 for θ a small number q_1 of steps are made to estimate D.

1. Network evolution is simulated from θ_0 and the values S_{i0} are computed
2. Network evolution is simulated from $\theta_0 + \epsilon_j e_j$ and the values S_{ij} are computed where
   - e_j is the j-th unit vector
   - $0.1 < \epsilon_j < 1$
3. Compute $d_{ij} = \epsilon_j^{-1}(S_{ij} - S_{i0})$
4. Repeat steps 1. to 3. until $i = q_1$
5. Estimate $E_\theta[S]$ and D by the Monte Carlo method

$$\overline{S} = \frac{1}{q_1}\sum_{i=1}^{q_1} S_{i0} \qquad \widehat{d_j} = \frac{1}{q_1}\sum_{i=1}^{q_1} d_{ij}$$

6. A new value of θ is estimated via the Robbins-Monro step with $a_i = 1$:

$$\widehat{\theta}_{q_1} = \theta_0 - \widehat{D}^{-1}(\overline{S} - s)$$

The Robbins-Monro algorithm

Phase 2: the main phase, consisting of c sub-phases. Each sub-phase h iterate the Robbins-Monro step for at most q_h step.

1. Generate the network evolution according to $\widehat{\theta}_i$ and compute S_{ih}

The Robbins-Monro algorithm

Phase 2: the main phase, consisting of c sub-phases. Each sub-phase h iterate the Robbins-Monro step for at most q_h step.

1. Generate the network evolution according to $\widehat{\theta}_i$ and compute S_{ih}
2. Updated θ according to the Robbins-Monro step

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_h \widehat{D}^{-1}(\overline{S}_i - s)$$

The Robbins-Monro algorithm

Phase 2: the main phase, consisting of c sub-phases. Each sub-phase h iterate the Robbins-Monro step for at most q_h step.

1. Generate the network evolution according to $\widehat{\theta}_i$ and compute S_{ih}
2. Updated θ according to the Robbins-Monro step

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_h \widehat{D}^{-1}(\overline{S}_i - s)$$

3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$

The Robbins-Monro algorithm

Phase 2: the main phase, consisting of c sub-phases. Each sub-phase h iterate the Robbins-Monro step for at most q_h step.

1. Generate the network evolution according to $\widehat{\theta}_i$ and compute S_{ih}
2. Updated θ according to the Robbins-Monro step

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_h \widehat{D}^{-1}(\overline{S}_i - s)$$

3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$
4. Compute

$$\widehat{\theta}_h = \frac{1}{i}\sum_i \widehat{\theta}_i$$

The Robbins-Monro algorithm

Phase 2: the main phase, consisting of c sub-phases. Each sub-phase h iterate the Robbins-Monro step for at most q_h step.

1. Generate the network evolution according to $\widehat{\theta}_i$ and compute S_{ih}
2. Updated θ according to the Robbins-Monro step

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_h \widehat{D}^{-1}(\overline{S}_i - s)$$

3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$
4. Compute

$$\widehat{\theta}_h = \frac{1}{i} \sum_i \widehat{\theta}_i$$

5. $a_{h+1} = \frac{a_h}{2}$ is the eventual estimate for θ

$\widehat{\theta}_c$ is the eventual estimate for θ

The Robbins-Monro algorithm

Phase 2: the main phase, consisting of c sub-phases. Each sub-phase h iterate the Robbins-Monro step for at most q_h step.

1. Generate the network evolution according to $\widehat{\theta}_i$ and compute S_{ih}
2. Updated θ according to the Robbins-Monro step

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_h \widehat{D}^{-1}(\overline{S}_i - s)$$

3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$
4. Compute

$$\widehat{\theta}_h = \frac{1}{i} \sum_i \widehat{\theta}_i$$

5. $a_{h+1} = \frac{a_h}{2}$ is the eventual estimate for θ

$\widehat{\theta}_c$ is the eventual estimate for θ

Phase 3: a number of q_3 simulations is used to evaluate the convergence of the algorithm and the accuracy of the estimator.

- Phase 3 requires the computation of D, thus it is similar to Phase 1.

The Maximum-likelihood estimation (MLE)

The model assumptions allow to decompose the process in a series of micro-steps:

$$\{(T_r, i_r, j_r), r = 1, \ldots, R\}$$

where

- T_r is the time point for an opportunity for change
- i_r denotes the actor who has the opportunity to change
- j_r is the actor towards whom the tie is changed

We denote by R the total number of micro-steps between t_0 and t_1 and we assume that the time point T_r are ordered increasingly:

$$t_0 = T_0 < T_1 < \ldots < T_R < t_1$$

The Maximum-likelihood estimation (MLE)

Definition

Given the sequence of micro-steps $\{(T_r, i_r, j_r), r = 0, \ldots, R\}$, the likelihood function of the network evolution process is defined by:

$$L(\theta) = \prod_{r=1}^{R} P((T_r, i_r, j_r))$$

Then, the estimate for θ is the vector of values $\widehat{\theta}$ such that:

$$\widehat{\theta} = arg \max_{\theta \in \Theta} L(\theta)$$

or equivalently, the vector of values $\widehat{\theta}$ such that:

$$U(\theta) = \frac{\partial}{\partial \theta} log(L(\theta)) = 0$$

where $U(\theta)$ is the score function.

Augmented data

Problem: we cannot observe the complete data and the likelihood of the observed data $(x(t_1), \ldots, x(t_M))$ conditional on $x(t_0)$.

Augmented data

Problem: we cannot observe the complete data and the likelihood of the observed data $(x(t_1), \ldots, x(t_M))$ conditional on $x(t_0)$.

$$\Downarrow$$

the parameter estimation requires finding the root of a system of equations in which the functions cannot be computed analytically $\Rightarrow$ a stochastic approximation method must be applied.

Augmented data

Problem: we cannot observe the complete data and the likelihood of the observed data $(x(t_1), \ldots, x(t_M))$ conditional on $x(t_0)$.

$$\Downarrow$$

the parameter estimation requires finding the root of a system of equations in which the functions cannot be computed analytically $\Rightarrow$ a stochastic approximation method must be applied.

The idea is to augment the observed data so that an easily computable likelihood is obtained.

Augmented data

Problem: we cannot observe the complete data and the likelihood of the observed data $(x(t_1), \dots, x(t_M))$ conditional on $x(t_0)$.

$$\Downarrow$$

the parameter estimation requires finding the root of a system of equations in which the functions cannot be computed analytically $\Rightarrow$ a stochastic approximation method must be applied.

The idea is to augment the observed data so that an easily computable likelihood is obtained.

Note: the data augmentation can be done separately for each time period (t_{m-1}, t_m). It is not restrictive to describe it only for two observations $x(t_0)$ and $x(t_1)$.

Augmented data

Definition

The *augmented data* (or *sample path*) consist of R and of the sequence of tie changes that brings the network from $x(t_0)$ to $x(t_1)$

$$(i_1, j_1), \ldots, (i_R, j_R)$$

Formally:

$$\underline{v} = \{(i_1, j_1), \ldots, (i_R, j_R)\} \in \mathcal{V}$$

where $\mathcal{V}$ is the set of all sample paths connecting $x(t_0)$ and $x(t_1)$.

Augmented data

Definition

The *augmented data* (or *sample path*) consist of R and of the sequence of tie changes that brings the network from $x(t_0)$ to $x(t_1)$

$$(i_1, j_1), \ldots, (i_R, j_R)$$

Formally:

$$\underline{v} = \{(i_1, j_1), \ldots, (i_R, j_R)\} \in \mathcal{V}$$

where $\mathcal{V}$ is the set of all sample paths connecting $x(t_0)$ and $x(t_1)$.

We can approximate the likelihood function (and then the score function) of the observed data using the probability of $\underline{v}$

$$P(\underline{v}|x(t_0), x(t_1)) \propto \frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{\lambda} p_{i_r j_r}(\beta, x(T_{r-1}))$$

Sampling the augmented data

The augmented data are sampled through a Markov chain simulation defined on$\mathcal{V}$

Sampling the augmented data

The augmented data are sampled through a Markov chain simulation defined on$\mathcal{V}$

Definition

The *Metropolis-Hastings algorithm* is defined by the following transition probabilities:

1. Given $\underline{v}_i = \underline{v}$, generate $\widetilde{\underline{v}}$ form the proposal distribution $u(\widetilde{\underline{v}}|\underline{v}_i)$

Sampling the augmented data

The augmented data are sampled through a Markov chain simulation defined on$\mathcal{V}$

Definition

The *Metropolis-Hastings algorithm* is defined by the following transition probabilities:

1. Given $\underline{v}_i = \underline{v}$, generate $\widetilde{\underline{v}}$ form the proposal distribution $u(\widetilde{\underline{v}}|\underline{v}_i)$

2. Take

$$\underline{v}_{i+1} = \begin{cases} \widetilde{\underline{v}} & \text{with probability} \quad \rho(\widetilde{\underline{v}}, \underline{v}) \\ \underline{v} & \text{with probability} \quad 1 - \rho(\widetilde{\underline{v}}, \underline{v}) \end{cases}$$

   where

$$\rho(\widetilde{\underline{v}}, \underline{v}) = \min \left\{ \frac{P(\widetilde{\underline{v}}) u(\underline{v}|\widetilde{\underline{v}})}{P(\underline{v}) u(\widetilde{\underline{v}}|\underline{v})}, 1 \right\}$$

The transition probabilities of the chain generate by the Metropolis-Hastings algorithm are given by $\rho(\widetilde{\underline{v}}, \underline{v}) u(\widetilde{\underline{v}}|\underline{v})$

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

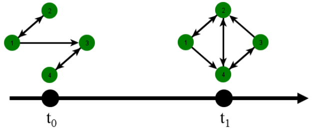
1. *Pairwise deletions*: one pair of indices r_1 and r_2 such that $(i_{r_1}, j_{r_1}) = (i_{r_2}, j_{r_2})$ is selected and the corresponding pairs (i_{r_1}, j_{r_1}) and (i_{r_2}, j_{r_2}) are deleted from $\underline{v}$

Sampling the augmented data

The proposal distribution $u(\underset{\sim}{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

1. *Pairwise deletions*: one pair of indices r_1 and r_2 such that $(i_{r_1}, j_{r_1}) = (i_{r_2}, j_{r_2})$ is selected and the corresponding pairs (i_{r_1}, j_{r_1}) and (i_{r_2}, j_{r_2}) are deleted from $\underline{v}$

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random (r_1, r_2) in $\{(1,7)(1,10)(2,8)\}$, e.g. $(r_1, r_2) = (1,7)$
- Delete the elements $(2,4)$

$$\underset{\sim}{\widetilde{v}} = (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

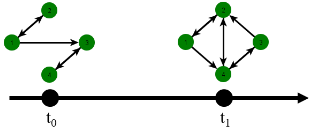
2. *Pairwise insertions*: one pair of $(i,j) \in \mathcal{N}^2$ and two indices r_1 and r_2 are randomly chosen. The element (i,j) is inserted in $\underline{v}$ immediately before r_1 and r_2.

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

2. *Pairwise insertions*: one pair of $(i,j) \in \mathcal{N}^2$ and two indices r_1 and r_2 are randomly chosen. The element (i,j) is inserted in $\underline{v}$ immediately before r_1 and r_2.

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random (i,j) and (r_1, r_2), e.g. $i=4,\ j=1,\ r_1=5,\ r_2=7$
- Insert the elements $(4,1)$ before $r_1 = 5$ and $r_2 = 7$

$$\underline{\widetilde{v}} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (\mathbf{4,1})\ (3,2)\ (1,4)\ (\mathbf{4,1})\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

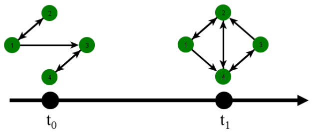
3. *Single deletion*: one pair (i_r, j_r) satisfying $i_r = j_r$ is randomly selected and deleted from $\underline{v}$

Sampling the augmented data

The proposal distribution $u(\underset{\sim}{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

3. *Single deletion*: one pair (i_r, j_r) satisfying $i_r = j_r$ is randomly selected and deleted from $\underline{v}$

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random r in $\{3, 11\}$, e.g. $r = 11$
- Delet the elements $(3,3)$

$$\underline{\widetilde{v}} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)$$

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}} \,|\, \underline{v})$ assigns non-null probabilities to the following changes:

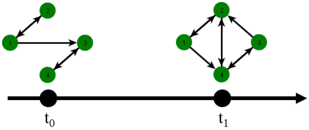
4. *Single insertion*: one actor $i \in \mathcal{N}$ and an index r are selected. The element (i, i) is inserted immediately before r

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}\,|\,\underline{v})$ assigns non-null probabilities to the following changes:

4. *Single insertion*: one actor $i \in \mathcal{N}$ and an index r are selected. The element (i, i) is inserted immediately before r

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random $i \in \mathcal{N}$ and r, e.g. $i = 4$ $r = 6$
- Insert the elements $(4,4)$ before $r = 6$

$$\underline{\widetilde{v}} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ \mathbf{(4,4)}\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

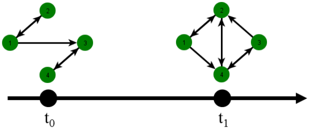
5. *Permutations*: for randomly chosen indices $r_1 < r_2$, the sequence $(i_{r_1}, j_{r_1}), \ldots, ((i_{r_2}, j_{r_2}))$ is randomly permuted

Sampling the augmented data

The proposal distribution $u(\underline{\widetilde{v}}|\underline{v})$ assigns non-null probabilities to the following changes:

5. *Permutations*: for randomly chosen indices $r_1 < r_2$, the sequence $(i_{r_1}, j_{r_1}), \ldots, ((i_{r_2}, j_{r_2}))$ is randomly permuted

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random (r_1, r_2) and r, e.g. $r_1 = 2, r_2 = 5$
- Permute the sequence $(i_2, j_2), \ldots, (i_5, j_5)$

$$\underline{v} = (2,4)\ \mathbf{(1,1)}\ \mathbf{(2,3)}\ \mathbf{(3,2)}\ \mathbf{(4,2)}\ (1,4)\ (4,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

Sampling the augmented data

Theorem

The Metropolis-Hastings algorithm leads to an irreducible, aperiodic and reversible Markov-chain.

Proof

- *The Markov chain is irreducible.*
  Pairwise deletions and insertions and single deletion and insertion are sufficient for all $\underline{v} \in$ to communicate.

- *The Markov chain is aperiodic.*
  The graph associated to the resulting Markov-chain contains all the loops and thus the greatest common divisor of all cycles is one.

Sampling the augmented data

- *The Markov chain is reversible.* The detailed balance condition:

$$\rho(\underline{\widetilde{v}}, \underline{v})u(\underline{\widetilde{v}}\,|\,\underline{v})P(\underline{v}) = \rho(\underline{v}, \underline{\widetilde{v}})u(\underline{v}\,|\,\underline{\widetilde{v}})P(\underline{\widetilde{v}})$$

is satisfied.

$$\rho(\underline{\widetilde{v}}, \underline{v})u(\underline{\widetilde{v}}\,|\,\underline{v})P(\underline{v}) = \min\left\{\frac{P(\underline{\widetilde{v}})u(\underline{v}\,\big|\,\underline{\widetilde{v}})}{P(\underline{v})u(\underline{\widetilde{v}}\,|\,\underline{v})}, 1\right\} u(\underline{\widetilde{v}}\,|\,\underline{v})P(\underline{v}) =$$

$$= \min\left\{\frac{P(\underline{\widetilde{v}})u(\underline{v}\,\big|\,\underline{\widetilde{v}})}{u(\underline{\widetilde{v}}\,|\,\underline{v})}, P(\underline{v})\right\} u(\underline{\widetilde{v}}\,|\,\underline{v}) =$$

$$= \min\left\{\frac{u(\underline{v}\,\big|\,\underline{\widetilde{v}})}{u(\underline{\widetilde{v}}\,|\,\underline{v})}, \frac{P(\underline{v})}{P(\underline{\widetilde{v}})}\right\} u(\underline{\widetilde{v}}\,|\,\underline{v})P(\underline{\widetilde{v}}) =$$

$$= \min\left\{1, \frac{P(\underline{v})u(\underline{\widetilde{v}}\,|\,\underline{v})}{P(\underline{\widetilde{v}})u(\underline{v}\,\big|\,\underline{\widetilde{v}})}\right\} u(\underline{v}\,|\,\underline{\widetilde{v}})P(\underline{\widetilde{v}}) =$$

$$= \rho(\underline{v}, \underline{\widetilde{v}})u(\underline{v}\,|\,\underline{\widetilde{v}})P(\underline{\widetilde{v}})$$

Sampling the augmented data

The ML estimation algorithm can be sketched in the following way:

1. For each $m = 1, \ldots, M-1$ makes a large number of Metropolis-Hastings steps yielding $v^{(i)} = (v_1^{(i)}, \ldots, v_{M-1}^{(i)})$
2. Compute the score function:

$$\frac{\partial}{\partial\theta} log(L(\widehat{\theta}_i; x; v_m^{(i)}))$$

3. Update the parameter estimate using the Robbins-Monro step

$$\theta_{i+1} = \theta_i + a_i D^{-1} U(L(\widehat{\theta}_i; x; v_m^{(i)}))$$

The estimate $\widehat{\theta}$ is calculated as the average of the θ_{i+1} values generated by this algorithm.

Parameter estimation

The Robbins-Monro algorithm and the ML estimation are implemented in the R library *RSiena* (Simulation Investigation for Empirical Network Analysis)

You need to load the following libraries:

1. library(snow)
2. library(rlecuyer)
3. library(RSiena)

The R script "estimation.R" contains the R commands to implement the estimation procedure in R and the folder "tfls.zip" includes the data files.

Example data: an excerpt from the "Teenage Friends and Lifestyle Study" data set:

- Networks: relation = friendship
  actors = 129 pupils present at all three measurement points
- Covariates: gender (1 = Male, 2 = Female)
  smoking behavior (1 = no, 2= occasional, 3 = regular)

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Rate parameter: expected frequency, between two consecutive network observations, with which actors get the opportunity to change a network tie

- about 9 opportunities for change in the first period
- about 7 opportunities for change in the second period

The estimated rate parameters will be higher than the observed number of changes per actor (why?)

Parameter interpretation: a very simple model

Interpreting the objective function parameters:

The parameter β_k quantifies the role of the effect s_k in the network evolution.

$\beta_k = 0$ s_k plays no role in the network dynamics

$\beta_k > 0$ higher probability of moving into networks where s_k is higher

$\beta_k < 0$ higher probability of moving into networks where s_k is lower

Which β_k are "significantly" different from 0?

E.g. $\beta_{rec} = 0.13$ is "significantly" different from 0?

Parameter interpretation: a very simple model

Hypothesis test:

1. State the hypotheses.
   - The *null hypothesis* (H_0) states that the observed increase or decrease in the number of network configurations related to a certain effect results purely from chance.

$$H_0 : \beta_k = 0$$

Parameter interpretation: a very simple model

Hypothesis test:

1. State the hypotheses.
   - The *null hypothesis* (H_0) states that the observed increase or decrease in the number of network configurations related to a certain effect results purely from chance.

$$H_0 : \beta_k = 0$$

   - The *alternative hypothesis* (H_1) states that the observed increase or decrease in the number of network configurations related to a certain effect is influenced by some non-random cause.

$$H_1 : \beta_k \neq 0$$

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left|\frac{\beta_k}{s.e.(\beta_k)}\right| \geq 2 & \text{reject } H_0 \\ \\ \left|\frac{\beta_k}{s.e.(\beta_k)}\right| < 2 & \text{fail to reject } H_0 \end{cases}$$

The logic behind this decision rule is based on the standard error concept.

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left|\frac{\beta_k}{s.e.(\beta_k)}\right| \geq 2 & \text{reject } H_0 \\ \\ \left|\frac{\beta_k}{s.e.(\beta_k)}\right| < 2 & \text{fail to reject } H_0 \end{cases}$$

The logic behind this decision rule is based on the standard error concept.

Example

Is the value $\beta_{rec} = 0.13$ far enough from 0?

If $s.e.(\beta_{rec}) = 0.9$, a more or less plausible set of values that the parameter can assume is approximately

$$[0.04, 0.22]$$

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left|\frac{\beta_k}{s.e.(\beta_k)}\right| \geq 2 & \text{reject } H_0 \\ \left|\frac{\beta_k}{s.e.(\beta_k)}\right| < 2 & \text{fail to reject } H_0 \end{cases}$$

The logic behind this decision rule is based on the standard error concept.

Example

Is the value $\beta_{rec} = 0.13$ far enough from 0?

If $s.e.(\beta_{rec}) = 0.9$, a more or less plausible set of values that the parameter can assume is approximately

$$[0.04, 0.22]$$

$$\left|\frac{\beta_{rec}}{s.e.(\beta_{rec})}\right| = \left|\frac{0.13}{0.9}\right| = 0.14 < 2$$

β_{rec} is not significantly different from 0

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Objective function parameters:

- outdegree parameter: the observed networks have low density

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Objective function parameters:

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties

Parameter interpretation: a very simple model

In more detail

$$\beta_{out} \sum_{j=1}^{n} x_{ij} + \beta_{rec} \sum_{j=1}^{n} x_{ij} x_{ji} = -2.4147 \sum_{j=1}^{n} x_{ij} + 2.7106 \sum_{j=1}^{n} x_{ij} x_{ji}$$

Adding a reciprocated tie (i.e., for which $x_{ji} = 1$) gives

$$-2.4147 + 2.7106 = 0.2959$$

while adding a non-reciprocated tie (i.e., for which $x_{ji} = 0$) gives

$$-2.4147$$

Conclusion: reciprocated ties are valued positively and non-reciprocated ties are valued negatively by actors

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

- the statement "the friend of my friend is also my friend" is almost always true

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

- the statement “the friend of my friend is also my friend” is almost always true $\rightarrow$ transitive triplets effect

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- pupils prefer to establish friendship relations with others that are similar to themselves

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- pupils prefer to establish friendship relations with others that are similar to themselves → covariate similarity

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

This effect must be controlled for the sender and receiver effects of the covariate.

- Covariate ego effect

- Covariate alter effect

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties
- transitivity parameter: preference for being friends with friends'friends

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- sex alter: gender does not affect actor popularity
- sex ego: gender does not affect actor activity
- sex similarity: tendency to choose friends with the same gender

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.

- All actor covariates are centered: $\overline{z} = 1.434$ is the mean of the covariate

$$z_i - \overline{z} = \begin{cases} -0.434 & \text{for boys} \\ \\ 0.566 & \text{for girls} \end{cases}$$

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.
- All actor covariates are centered: $\bar{z} = 1.434$ is the mean of the covariate

$$z_i - \bar{z} = \begin{cases} -0.434 & \text{for boys} \\ 0.566 & \text{for girls} \end{cases}$$

- The contribution of x_{ij} to the objective function is

$$\beta_{ego}(z_i - \bar{z}) + \beta_{alter}(z_j - \bar{z}) + \beta_{same}\mathbb{I}\{z_i = z_j\} =$$

$$= 0.1571(z_i - \bar{z}) - 0.1513(z_j - \bar{z}) + 0.9191\mathbb{I}\{z_i = z_j\}$$

where $\mathbb{I}\{z_i = z_j\}$ is the indicator function

$$\mathbb{I}\{z_i = z_j\} \begin{cases} 1 & z_i = z_j \\ 0 & \text{otherwise} \end{cases}$$

Parameter interpretation: a more complex model

	Male	Female
Male	0.9166	0.1546
Female	-0.1538	0.9224

Table: Gender-related contributions to the objective function

Conclusions:

- preference for similar alters
- the negative value associated to the the single tie form a girl to a boy, suggests that girls seem not to like male friends.

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- smoke alter: smoking behavior does not affect actor popularity
- smoke ego: smoking behavior not affect actor activity
- smoke similarity: tendency to choose friends with the same smoking behavior

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for “no”, 2 for “occasional”, and 3 for “regular” smokers.

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{z} = 1.310$ is the mean of the covariate

$$z_i - \overline{z} = \begin{cases} -0.310 & \text{for no smokers} \\ 0.690 & \text{for occasional smokers} \\ 1.690 & \text{for regular smokers} \end{cases}$$

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{z} = 1.310$ is the mean of the covariate

$$z_i - \overline{z} = \begin{cases} -0.310 & \text{for no smokers} \\ 0.690 & \text{for occasional smokers} \\ 1.690 & \text{for regular smokers} \end{cases}$$

- The contribution of x_{ij} to the objective function is

$$\beta_{ego}(z_i - \overline{z}) + \beta_{alter}(z_j - \overline{z}) + \beta_{same}\left(1 - \frac{|z_i - z_j|}{R_z} - sim_z\right) =$$

$$= 0.0714(z_i - \overline{z}) + 0.1055(z_j - \overline{z}) + 0.3724\left(1 - \frac{|z_i - z_j|}{2} - 0.7415\right)$$

where $R_z = z_{max} - z_{min}$

Parameter interpretation: a more complex model

	no	occasional	regular
no	0.0414	-0.0734	-0.1882
occasional	-0.0393	0.2183	0.1035
regular	-0.1200	0.1376	0.3952

Table: Smoking-related contributions to the objective function

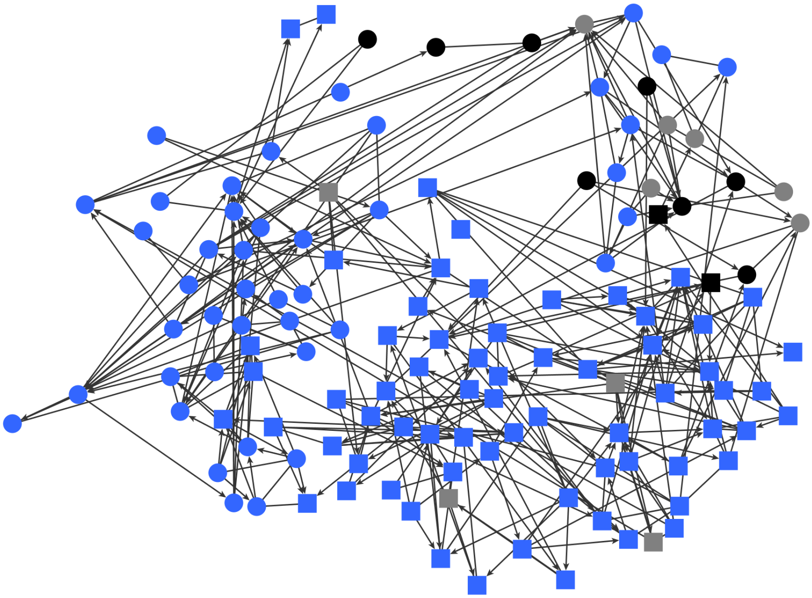
Conclusions:

- preference for similar alters
- this tendency is strongest for high values on smoking behavior

Outline

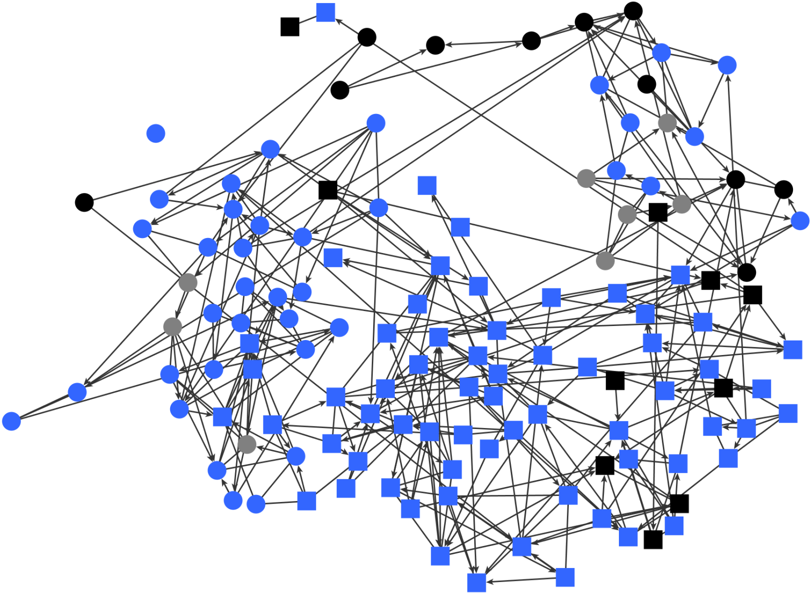
Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



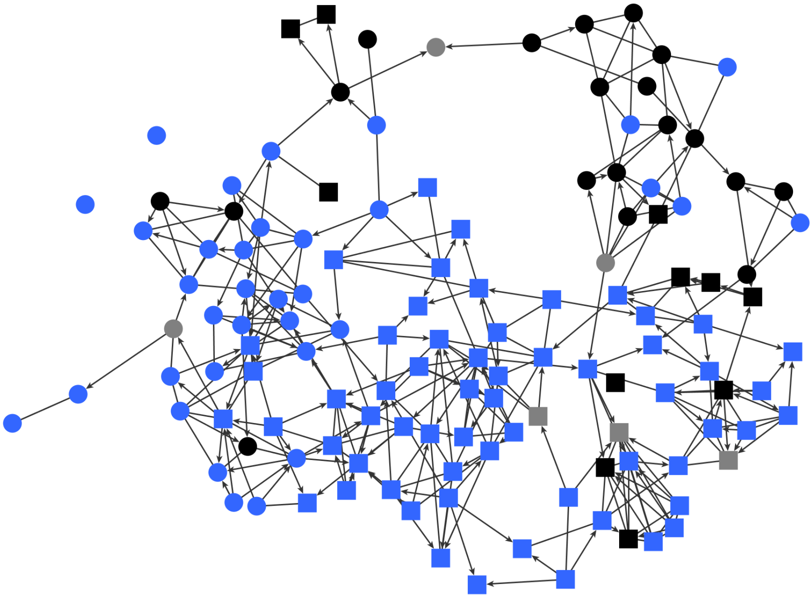
Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



Motivation

1. Social network dynamics can depend on actors'characteristics.

Selection process: relationship *partners* are selected according to their characteristics

Motivation

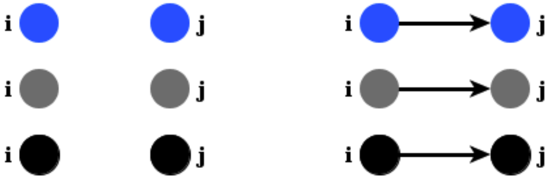
1. Social network dynamics can depend on actors'characteristics.

Selection process: relationship *partners* are selected according to their characteristics

Example

Homophily: the formation of relations based on the similarity of two actors

E.g. smoking behavior



Motivation

2. Changeable actors'characteristics can depend on the social network

E.g.: opinions, attitudes, intentions, etc. - we use the word behavior for all of these!

Influence process: actors adjust their characteristics according to the characteristics of other actors to whom they are tied

Motivation

2. Changeable actors'characteristics can depend on the social network

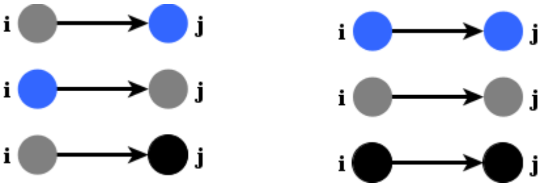
E.g.: opinions, attitudes, intentions, etc. - we use the word behavior for all of these!

Influence process: actors adjust their characteristics according to the characteristics of other actors to whom they are tied

Example

Assimilation/contagion: connected actors become increasingly similar over time

E.g. smoking behavior



Competing explanatory stories

Homophily and assimilation give rise to the same outcome (similarity of connected individuals)

$$\Downarrow$$

study of influence requires the consideration of selection and vice versa.

Fundamental question: is this similarity caused mainly by influence or mainly by selection?

Competing explanatory stories

Homophily and assimilation give rise to the same outcome (similarity of connected individuals)

$$\Downarrow$$

study of influence requires the consideration of selection and vice versa.

Fundamental question: is this similarity caused mainly by influence or mainly by selection?

Extending the SAOM for the co-evolution of networks and behaviors

Competing explanatory stories

Example

Similarity in smoking:

Selection: “a smoker may tend to have smoking friends because, once somebody is a smoker, he or she is likely to meet other smokers in smoking areas and thus has more opportunities to form friendship ties with them”

Competing explanatory stories

Example

Similarity in smoking:

Selection: “a smoker may tend to have smoking friends because, once somebody is a smoker, he or she is likely to meet other smokers in smoking areas and thus has more opportunities to form friendship ties with them”

Influence: “a smoker may have been the friendship with a smoker that made him or her start smoking in the first place”

Longitudinal network-behavior panel data

1. a network x represented by its adjacency matrix
2. a series of actors' attributes:
   - H constant covariates $V_1, \cdots, V_H$
   - L behavior covariates $Z_1(t), \cdots, Z_L(t)$
     Behavior variables are ordinal categorical variables.

Longitudinal network-behavior panel data

1. a network x represented by its adjacency matrix
2. a series of actors' attributes:
   - H constant covariates $V_1, \cdots, V_H$
   - L behavior covariates $Z_1(t), \cdots, Z_L(t)$
     Behavior variables are ordinal categorical variables.

Longitudinal network-behavior panel data: networks and behaviors observed at $M \geq 2$ time points $t_1, \cdots, t_M$

$$(x, z)(t_1),\ (x, z)(t_2), \cdots,\ (x, z)(t_M)$$

and the constant covariates $V_1, \cdots, V_H$.

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

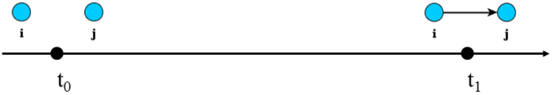
     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*

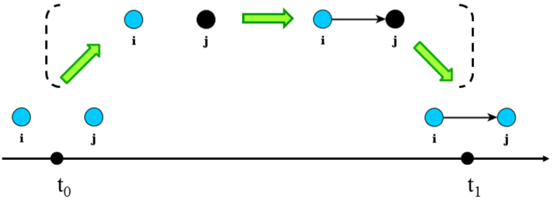


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^n(n-1) \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*

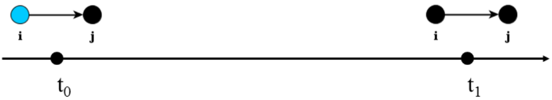


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^n(n-1) \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network and behavior
   - *Continuous-time:*

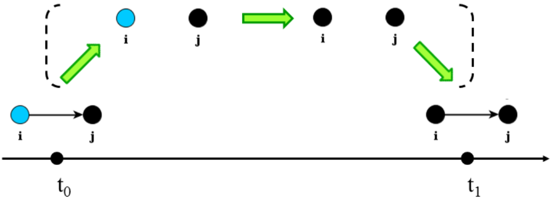


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

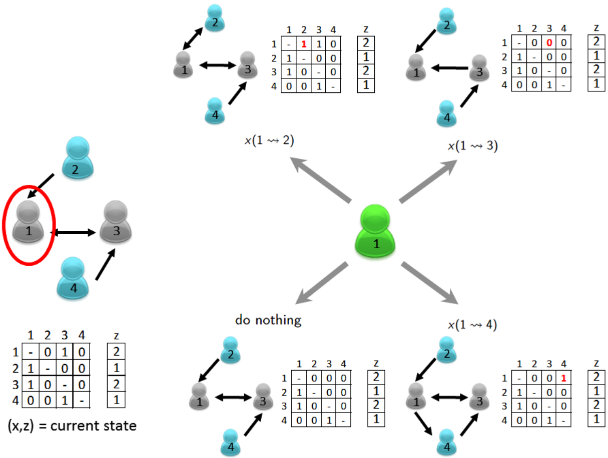
     $$|C| = 2^n(n-1) \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*



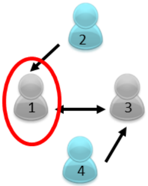
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing tie or his behavior.



Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing tie or his behavior.

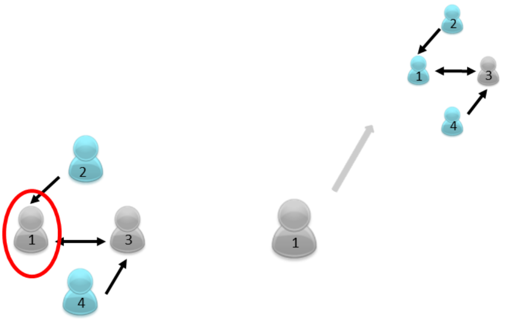


	1	2	3	4		z
1	-	0	1	0		2
2	1	-	0	0		1
3	1	0	-	0		2
4	0	0	1	-		1

(x,z) = current state

Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing tie or his behavior.



	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

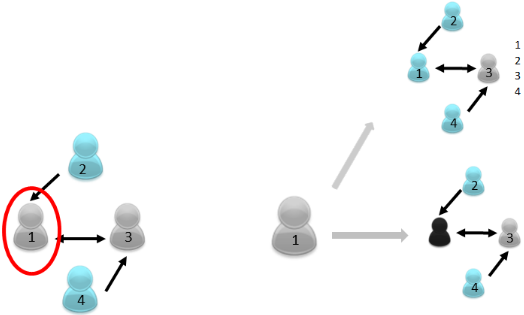
(x,z) = current state

	1	2	3	4	z
1	-	0	1	0	1
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

z(2⇝1)

Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing tie or his behavior.



	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

(x,z) = current state

	1	2	3	4	z
1	-	0	1	0	1
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

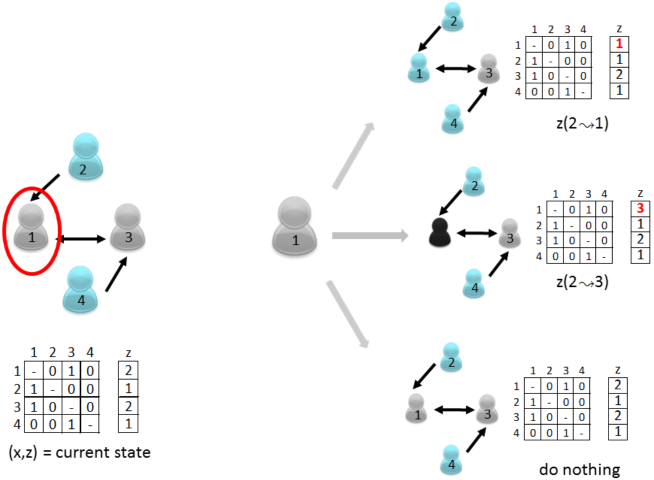
z(2⤳1)

	1	2	3	4	z
1	-	0	1	0	3
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

z(2⤳3)

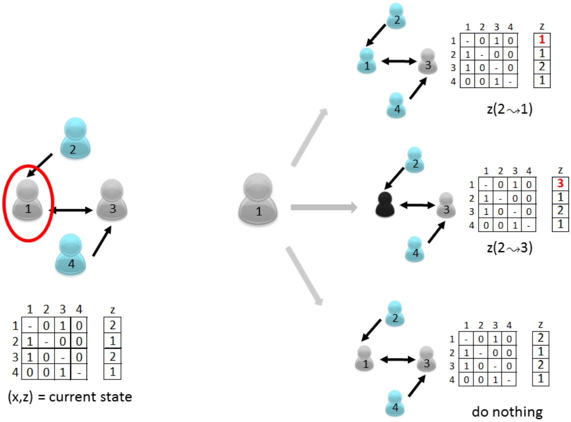
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing tie or his behavior.



Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing tie or his behavior.



The moments at which an actor has the opportunity for a tie change or a behavior change are modeled by two distinct rate functions.

Assumptions

3. *Absence of co-occurrence.*
   At each instant t, only one actor has the opportunity to change (one of his outgoing ties or his behavior)

4. *Actor-oriented perspective.*
   Actors control their outgoing ties as well as their own behavior.
   - the actor decide to change one of his outgoing ties or his behavior according to his position in the network, his attributes and the characteristics of the other actors

     Aim: maximize *a utility function*
   - two distinct objective functions: one for the network and one for the behavior change
   - actors have complete knowledge about the network and the behaviors of all the the other actors
   - the maximization is based on immediate returns and not on long-run rewarding (myopic actors)

Model definition

According to the previous assumptions, the network-behavior co-evolution process is decomposed into a series of micro-steps:

- the opportunity of changing one network tie and the corresponding tie changed
- the opportunity of changing a behavior and the corresponding unit changed in behavior

Model definition

According to the previous assumptions, the network-behavior co-evolution process is decomposed into a series of micro-steps:

- the opportunity of changing one network tie and the corresponding tie changed
- the opportunity of changing a behavior and the corresponding unit changed in behavior

$$\Downarrow$$

every micro-step requires the identification of a focal actor who gets the opportunity to make a change and the identification of the change outcome

	Occurrence	*Preference*
Network changes	Network rate function	Network objective function
Behavioral changes	Behavioral rate function	Behavioral objective function

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

Practically always, one type of decision will be made more frequently than the other

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

Practically always, one type of decision will be made more frequently than the other

Example

In the joint study of friendship and smoking behavior at high school, we would expect more frequent changes in the network than in behavior

The rate functions

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda_i^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda_i^{beh})$$

The rate functions

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda_i^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda_i^{beh})$$

The waiting time until the occurrence of the next micro step of either kind by any actor is exponentially distributed with parameter:

$$\lambda_{tot} = \sum_i (\lambda_i^{net} + \lambda_i^{beh})$$

The rate functions

Network rate function

T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda_i^{net})$$

Behavior rate function

T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda_i^{beh})$$

The waiting time until the occurrence of the next micro step of either kind by any actor is exponentially distributed with parameter:

$$\lambda_{tot} = \sum_i (\lambda_i^{net} + \lambda_i^{beh})$$

Probabilities

$$P(i \text{ has the opportunity to change one of his tie}) = \frac{\lambda_i^{net}}{\lambda_{tot}}$$

$$P(i \text{ has the opportunity to change his behavior}) = \frac{\lambda_i^{beh}}{\lambda_{tot}}$$

The rate functions (simplest specification)

Network rate function

T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda^{net})$$

Behavior rate function

T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda^{beh})$$

The rate functions (simplest specification)

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda^{beh})$$

The waiting time until the occurrence of the next micro step of either kind by any actor is exponentially distributed with parameter:

$$\lambda_{tot} = n(\lambda^{net} + \lambda^{beh})$$

The rate functions (simplest specification)

Network rate function

T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda^{net})$$

Behavior rate function

T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda^{beh})$$

The waiting time until the occurrence of the next micro step of either kind by any actor is exponentially distributed with parameter:

$$\lambda_{tot} = n(\lambda^{net} + \lambda^{beh})$$

Probabilities

$$P(\text{network micro-step}) = \frac{n\lambda^{net}}{\lambda_{tot}}$$

$$P(\text{behavioral micro-step}) = \frac{n\lambda^{beh}}{\lambda_{tot}}$$

The objective functions

The probability of going from one state to another state of the co-evolution Markov chain is defined by the objective functions.

Why must we specify two different objective functions?

The objective functions

The probability of going from one state to another state of the co-evolution Markov chain is defined by the objective functions.

Why must we specify two different objective functions?

- The network objective function represents how likely it is for *i* to change one of its outgoing tie
- The behavioral objective function represents how likely it is for the actor *i* the current level of his behavior

The objective functions

The probability of going from one state to another state of the co-evolution Markov chain is defined by the objective functions.

Why must we specify two different objective functions?

- The network objective function represents how likely it is for i to change one of its outgoing tie
- The behavioral objective function represents how likely it is for the actor i the current level of his behavior

Network objective function

$$f_i^{net}(\beta, x(i \rightsquigarrow j), z) = \sum_{k=1}^{K} \beta_k s_{ik}^{net}(x, z) + U_i(t, x, j)$$

already described for the SAOM

The objective functions

Behavioral objective function

$$f_i^{beh}(\gamma, z(l \rightsquigarrow l'), x) = \sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l')) + \epsilon_i(t, z, l, l')$$

where

- $s_w^{beh}(x, z)$ are effects
- γ_w are statistical parameters
- $\epsilon_i(t, z, l, l')$ is a random term

The probability that an actor i changes his own behavior by one unit is:

$$p_{ll'}(\gamma; z(l \rightsquigarrow l'); x) = \frac{exp\left(\sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l'))\right)}{\sum_{l'' \in \{l+1, l-1, l\}} exp\left(\sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l''))\right)}$$

p_{ll} is the probability that i does not change his behavior

The objective functions

The specification of the behavioral objective function

- *Basic shape effects*
  The linear shape effect $s^{beh}_{i\_linear}(x, z)$ and the quadratic shape effect $s^{beh}_{i\_quadratic}(x, z)$ are defined by

$$s^{beh}_{i\_linear}(x, z) = z_i \qquad\qquad s^{beh}_{i\_quadratic}(x, z) = z_i^2$$

  The basic shape effects must be always included in the model specification.

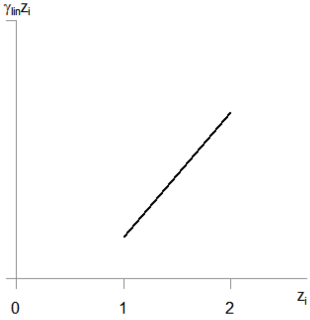
The objective functions

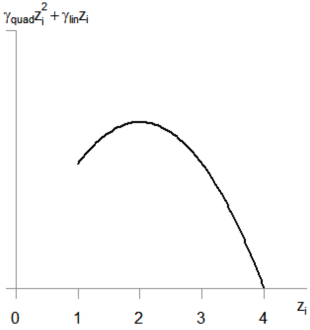
The specification of the behavioral objective function

- *Basic shape effects*
  The linear shape effect $s^{beh}_{i\_linear}(x,z)$ and the quadratic shape effect $s^{beh}_{i\_quadratic}(x,z)$ are defined by

$$s^{beh}_{i\_linear}(x,z) = z_i \qquad\qquad s^{beh}_{i\_quadratic}(x,z) = z_i^2$$

The basic shape effects must be always included in the model specification.





The specification of the behavioral objective function

- *Classical influence effects*

1. The *average similarity effect* $s^{beh}_{i\_avsim}(x,z)$ expressing the preference of actors to be similar in behavior to their alters, in such a way that the total influence of the alters is the same regardless ego's outdegree

$$s^{beh}_{i\_avsim}(x,z) = \frac{1}{x_{i+}} \sum_{j=1}^{n} x_{ij}(sim_z(ij) - \widehat{sim_z})$$

where

$$sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_z}$$

R_z is the range of the behavior z and $\widehat{sim_z}$ is the mean similarity value.

The objective functions

The specification of the behavioral objective function

- *Classical influence effects*

1. The *average similarity effect* $s^{beh}_{i\_avsim}(x,z)$ expressing the preference of actors to be similar in behavior to their alters, in such a way that the total influence of the alters is the same regardless ego's outdegree

$$s^{beh}_{i\_avsim}(x,z) = \frac{1}{x_{i+}} \sum_{j=1}^{n} x_{ij}(sim_z(ij) - \widehat{sim}_z)$$

   where

$$sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_z}$$

   R_z is the range of the behavior z and $\widehat{sim}_z$ is the mean similarity value.

2. The *total similarity effect* $s^{beh}_{i\_totsim}(x,z)$, expressing the preference of actors to be similar in behavior to their alters, in such a way that the total influence of the alters is proportional to the number of alters:

$$s^{beh}_{i\_totsim}(x,z) = \sum_{j=1}^{n} x_{ij}(sim_z(ij) - \widehat{sim}_z)$$

The objective functions

The specification of the behavioral objective function

- *Position-dependent influence effects*
  Network position could also have an effect on the behavior of dynamics.
  1. *outdegree effect*
  $$s^{beh}_{i\_out}(x,z) = z_i \sum_{j=1}^{n} x_{ij}$$
  2. *indegree effect*
  $$s^{beh}_{i\_ind}(x,z) = z_i \sum_{j=1}^{n} x_{ji}$$
- *Effects of other actor variables.*
  For each actor's attribute a main effect on the behavior can be included in the model.

Simulating the co-evolution of networks and behavior

The algorithm consists in reproducing a possible series of micro-steps between two observation moments t_0 and t_1 according to fixed parameter value.

1. Set the time $t = 0$, $x = x(t_0)$ and $z = z(t_0)$
2. Generate T^{net} according to an exponential distribution with parameter λ^{net}
3. Generate T^{beh} according to an exponential distribution with parameter λ^{beh}

Simulating the co-evolution of networks and behavior

4. If $\min\left\{T^{net}, T^{beh}\right\} = T^{net}$ a **network micro-step** is implemented:

   - Select the actor $i \in \mathcal{N}$, who makes the changes, with probability

   $$P(i \text{ has the opportunity to change} \mid \text{network micro-steps}) = \frac{\lambda^{net}}{\lambda_{tot}}$$

   - Select the actor $j \in \mathcal{N}$, to whom i changes his outgoing tie, with probability:

   $$p_{ij} = \frac{exp\left(\sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow j), z)\right)}{\sum_{h=1}^{n} exp\left(\sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow h), z)\right)}$$

   - If $i \neq j$ then $x = x(i \rightsquigarrow j)$. If $i = j$ then $x = x$
   - Set $t = t + T^{net}$

Simulating the co-evolution of networks and behavior

4. If $\min\left\{T^{net}, T^{beh}\right\} = T^{beh}$ a **behavioral micro-step** is implemented:

   - Select the actor $i \in \mathcal{N}$, who makes the changes, with probability

$$P(i \text{ has the opportunity to change}|\text{behavioral micro-steps}) = \frac{\lambda^{beh}}{\lambda_{tot}}$$

   - Determine the behavioral change $l' \in \{l+1, l-1, l\}$ with probability:

$$p_{ll'}(\gamma; z(l \rightsquigarrow l'); x) = \frac{exp\left(\sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l'))\right)}{\sum_{l'' \in \{l+1, l-1, l\}} exp\left(\sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l''))\right)}$$

   - If $l \neq l'$ then $z = z(l \rightsquigarrow l')$. If $l = l'$ then $z = z$
   - Set $t = t + T^{beh}$

5. Repeat step 2. to 4. until the stopping criterion is satisfied.

Simulating the co-evolution of networks and behavior

1. *Unconditional* simulation:

   the simulations in each time period carry on until a predetermined time length has elapsed (usually until $t = 1$).

Simulating the co-evolution of networks and behavior

1. *Unconditional* simulation:

   the simulations in each time period carry on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:

   - simulations run on until the number of different entries between $x(t_0)$ and the simulated network $x^{sim}(t_1)$ is equal to the number of entries that differ between $x(t_0)$ and $x(t_1)$

   $$\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}^{obs}(t_1) - X_{ij}(t_0) \right| = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}^{sim}(t_1) - X_{ij}(t_0) \right|$$

   - simulations run on until the number of different entries between $z(t_0)$ and the simulated behavior $z^{sim}(t_1)$ is equal to the number of entries that differ between $z(t_0)$ and $z(t_1)$

   $$\sum_{i=1}^{n} \left| z_i^{obs}(t_1) - z_i(t_0) \right| = \sum^{n} \left| z_i^{sim}(t_1) - z_i(t_0) \right|$$

The parameter estimation (MoM)

Aim: estimate θ, the 2(M-1)+K+W dimensional vector of parameters of the co-evolution model

Statistics:

- Network rate parameters for the period m

$$s^{net}_{\lambda_m}(X(t_m), X(t_{m-1}) | X(t_{m-1}) = x(t_{m-1})) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}(t_m) - X_{ij}(t_{m-1}) \right|$$

- Behavior rate parameters for the period m

$$s^{beh}_{\lambda_m}(Z(t_m), Z(t_{m-1}) | Z(t_{m-1}) = z(t_{m-1})) = \sum_{i=1}^{n} |Z_i(t_m) - Z_i(t_{m-1})|$$

The parameter estimation (MoM)

Aim: estimate θ, the 2(M-1)+K+W dimensional vector of parameters of the co-evolution model

Statistics:

- Network objective function effects

$$\sum_{m=1}^{M-1} s_{mk}^{net}(X(t_m)|X(t_{m-1})=x(t_{m-1}))=\sum_{m=1}^{M-1} s_{mk}^{net}(X(t_m),X(t_{m-1}))$$

- Behavior objective function effects

$$\sum_{m=1}^{M-1} s_{mw}^{beh}(X(t_m)|X(t_{m-1})=x(t_{m-1}))=\sum_{m=1}^{M-1} s_{mw}^{beh}(X(t_m),X(t_{m-1}))$$

The parameter estimation (MoM)

Consequently the MoM estimator for θ is provided by the solution of the system of equations:

$$\begin{cases} E_\theta\left[s_{\lambda_m}(X(t_m), X(t_{m-1})|X(t_{m-1}) = x(t_{m-1}))\right] = s_{\lambda_m}(x(t_m), x(t_{m-1})) \\ E_\theta\left[s_{\lambda_m}(Z(t_m), Z(t_{m-1})|Z(t_{m-1}) = z(t_{m-1}))\right] = s_{\lambda_m}(z(t_m), z(t_{m-1})) \\ E_\theta\left[\sum_{m=1}^{M-1} s_{mk}^{net}(X(t_m)|X(t_{m-1}) = x(t_{m-1}))\right] = \sum_{m=1}^{M-1} s_{mk}^{net}(x(t_m), x(t_{m-1})) \\ E_\theta\left[\sum_{m=1}^{M-1} s_{mw}^{beh}(X(t_m)|X(t_{m-1}) = x(t_{m-1}))\right] = \sum_{m=1}^{M-1} s_{mw}^{beh}(x(t_m), x(t_{m-1})) \end{cases}$$

The parameter estimation (MoM)

Remarks

- The system of equation cannot be solved analytically $\Longrightarrow$ Robbins-Monro algorithm

- The Maximum likelihood estimation is under construction. At the moment is too slow

Example

Example data: excerpt from the "Teenage Friends and Lifestyle Study" data set

We will use the SAOM for the co-evolution of networks and behaviors to disentangle influence and selection processes.

1. Do pupils select friends based on similar smoking behavior?
2. Are pupils influenced by friends to adjust to their smoking behavior?

Dependent variables: friendship networks and smoking behavior

Covariate: gender

Precondition of the analysis

To find out whether it makes sense to analyze the data with a co-evolution model one should check whether:

1. the data are sufficiently informative to allow for identification of effects

$$J = \frac{N_{11}}{N_{11} + N_{01} + N_{10}} > 0.3 \qquad \textit{Jaccard index}$$

```
Tie changes between subsequent observations:
 periods        0 =>  0   0 =>  1   1 =>  0   1 =>  1   Distance Jaccard   Missing
  1 ==>  2      15827       237       240       208     477      0.304     0 (0%)
  2 ==>  3      15839       228       209       236     437      0.351     0 (0%)
```

Precondition of the analysis

2. there is interdependence between networks and behavioral variables

$$I = -\frac{\overline{d}(\delta)}{(n-1)\overline{d}} \qquad \text{Moran index}$$

where

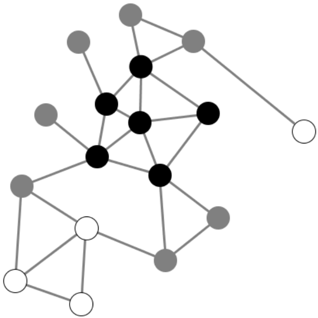
-
$$\overline{d}(\delta) = \frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}}$$

is the mean of the cross products of the centered behavioral variable for connected actors

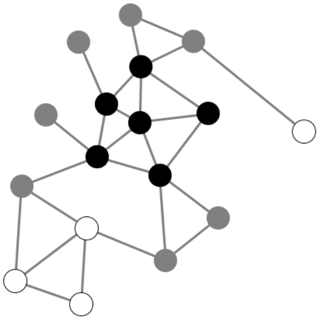
-
$$\overline{d} = \frac{\sum_{ij}(z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

is the overall mean of the cross products of the centered behavioral variable for all the possible pairs of actors in the network

Precondition of the analysis: understanding the Moran index



Precondition of the analysis: understanding the Moran index

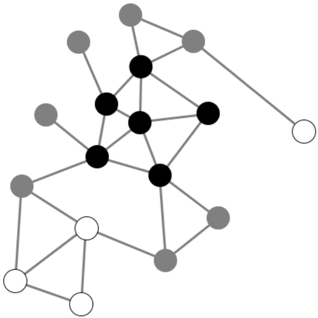


$$\frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}} \qquad \frac{\sum_{ij} (z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

If two connected actors i and j are similar in their behaviors we expect that

- $z_i > \overline{z}$ and $z_j > \overline{z}$
- $z_i < \overline{z}$ and $z_j < \overline{z}$

Precondition of the analysis: understanding the Moran index

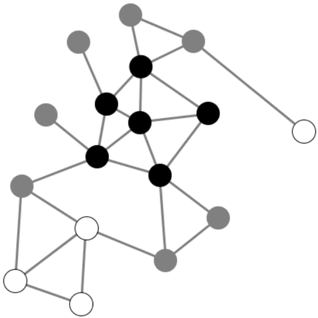


$$\frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}} > \frac{\sum_{ij}(z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

If two connected actors i and j are similar in their behaviors we expect that

- $z_i > \overline{z}$ and $z_j > \overline{z}$
- $z_i < \overline{z}$ and $z_j < \overline{z}$

Precondition of the analysis: understanding the Moran index



$$\frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}} > \frac{\sum_{ij}(z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

If two connected actors i and j are similar in their behaviors we expect that

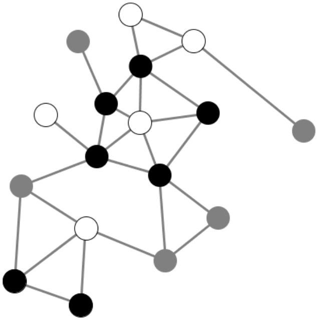
- $z_i > \overline{z}$ and $z_j > \overline{z}$
- $z_i < \overline{z}$ and $z_j < \overline{z}$

$\Downarrow$

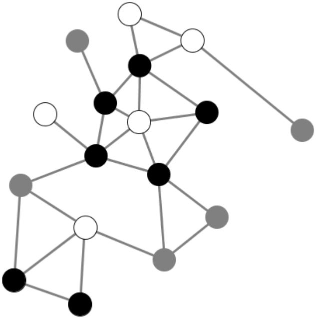
Positive interdependence

$$0 < I < 1$$

Precondition of the analysis: understanding the Moran index



Precondition of the analysis: understanding the Moran index

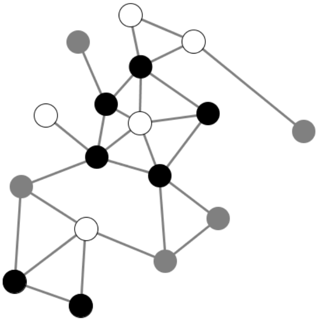


$$\frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}} \qquad \frac{\sum_{ij} (z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

If two connected actors i and j are extremely different in their behaviors we expect that

- $z_i > \overline{z}$ and $z_j < \overline{z}$
- $z_i < \overline{z}$ and $z_j > \overline{z}$

Precondition of the analysis: understanding the Moran index

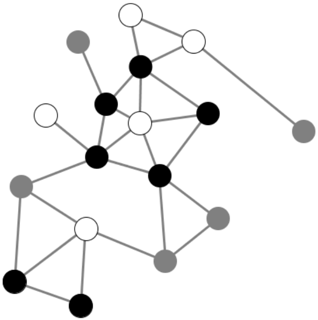


$$\frac{\sum_{ij} x_{ij}(z_i - \bar{z})(z_j - \bar{z})}{\sum_{ij} x_{ij}} < \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)}$$

If two connected actors i and j are extremely different in their behaviors we expect that

- $z_i > \bar{z}$ and $z_j < \bar{z}$
- $z_i < \bar{z}$ and $z_j > \bar{z}$

Precondition of the analysis: understanding the Moran index



$$\frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}} < \frac{\sum_{ij}(z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

If two connected actors i and j are extremely different in their behaviors we expect that

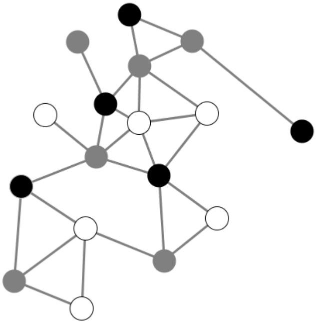
- $z_i > \overline{z}$ and $z_j < \overline{z}$
- $z_i < \overline{z}$ and $z_j > \overline{z}$

$$\Downarrow$$

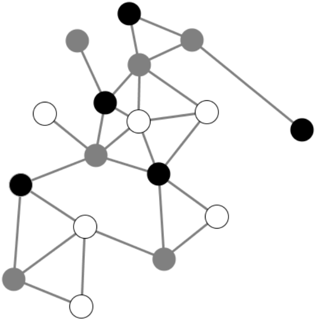
Negative interdependence

$$-1 < I < 0$$

Precondition of the analysis: understanding the Moran index



Precondition of the analysis: understanding the Moran index



If:

$$\frac{\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_{ij} x_{ij}} \approx \frac{\sum_{ij}(z_i - \overline{z})(z_j - \overline{z})}{n(n-1)}$$

the *local* distribution of the behavioral variable follows the *global* distribution of the behavioral variable

$$\Downarrow$$

Absence of interdependence

$$I \approx -\frac{1}{n-1}$$

Precondition of the analysis

In theory $-1 \leq I \leq +1$

- values close to zero indicates independence between networks and behaviors (i.e.absence of interdependence)
- value $+1$ indicates perfect identity of the behaviors of two friends (i.e. very strong positive interdependence)
- value -1 indicates perfect complementarity of the behaviors of two friends (i.e. very strong negative interdependence).

Precondition of the analysis

In theory $-1 \leq I \leq +1$

- values close to zero indicates independence between networks and behaviors (i.e.absence of interdependence)
- value $+1$ indicates perfect identity of the behaviors of two friends (i.e. very strong positive interdependence)
- value -1 indicates perfect complementarity of the behaviors of two friends (i.e. very strong negative interdependence).

The computation of the Moran index for the friendship networks and smoking behaviors leads to

0.244 0.258 0.341

Conclusion: there is considerable dependence between networks and behaviors.

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Network rate parameters:

- about 9 opportunities for a network change in the first period
- about 7 opportunities for a network change in the second period

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Network objective function parameters:

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Behavioral rate parameters:

- about 4 opportunities for a behavioral change in the first period
- about 4 opportunities for a behavioral change in the second period

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.12676
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Behavioral objective function parameters: express the attractiveness of different behavioral levels taking into account the current structure of the network and the behavior of the other actors

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{z} = 1.377$ is the mean of the covariate

$$z_i - \overline{z} = \begin{cases} -0.377 & \text{for no smokers} \\ 0.623 & \text{for occasional smokers} \\ 1.623 & \text{for regular smokers} \end{cases}$$

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{z} = 1.377$ is the mean of the covariate

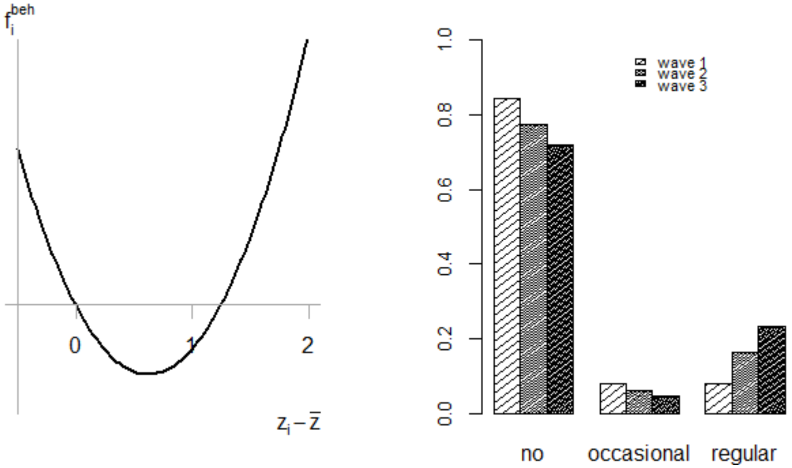
$$z_i - \overline{z} = \begin{cases} -0.377 & \text{for no smokers} \\ 0.623 & \text{for occasional smokers} \\ 1.623 & \text{for regular smokers} \end{cases}$$

- The contribution to the behavioral objective function is

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 =$$

$$= -3.5464_{linear}(z_i - \overline{z}) + 2.8464_{quadratic}(z_i - \overline{z})^2$$

Parameter interpretation: a baseline model



U-shaped changes in the behavior are drawn to the extreme of the range

A more complex model

The baseline model does not provide any information about selection and influence processes:

- the network dynamics are explained by the preference towards creating and reciprocating ties
- the behavior dynamic are described only by the distribution of the behavior in the population

A more complex model

The baseline model does not provide any information about selection and influence processes:

- the network dynamics are explained by the preference towards creating and reciprocating ties
- the behavior dynamic are described only by the distribution of the behavior in the population

If we want to disentangle the selection and influence effects we should include in the objective functions specification:

- the effects that capture the dependence of social network dynamics on actor's characteristic
- the effects that capture the dependence of behavior dynamics on social network

A more complex model

Effects that capture the dependence of social network dynamics on actor's characteristic

- pupils prefer to establish friendship relations with others that are similar to themselves

A more complex model

Effects that capture the dependence of social network dynamics on actor's characteristic

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

A more complex model

Effects that capture the dependence of social network dynamics on actor's characteristic

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

This effect must be controlled for the sender and receiver effects of the covariate.

- Covariate ego effect

- Covariate alter effect

A more complex model

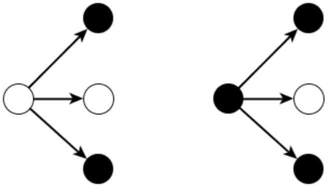
Effects that capture the dependence of behavior dynamics on social network

- pupils tend to adjust their smoking behavior according to the behaviors of their friends

A more complex model

Effects that capture the dependence of behavior dynamics on social network

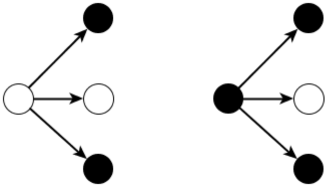
- pupils tend to adjust their smoking behavior according to the behaviors of their friends $\rightarrow$ average similarity effect



A more complex model

Effects that capture the dependence of behavior dynamics on social network

- pupils tend to adjust their smoking behavior according to the behaviors of their friends $\rightarrow$ average similarity effect



This effect must be controlled for the indegree and the outdegree effects

- Indegree effect

- Outdegree effect

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

Network objective function parameters: tendency towards reciprocity, transitivity and homophily with respect to gender

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

Network objective function parameters: pupils selected others with similar smoking behavior as friends $\rightarrow$ evidence for selection process

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shap	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters: U-shaped distribution of the smoking behavior

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shap	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters: indegree and outdegree effects are not significant

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shap	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters: pupils are influenced by the smoking behavior of the others → evidence for influence process

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573_{linear}(z_i - \overline{z}) + 2.8406_{quadratic}(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

Since the behavioral objective function depends not generally on the average behavior of the actor's friends, here we present a table only for the special case of actors all whose friend have the same behavior z_j.

z_j / z_i	no	occasional	regular
1	2.56	-1.82	-0.51
2	0.84	-0.10	1.20
3	-0.88	-1.82	2.92

- The row maximum is assumed at the diagonal for the non-smokers and for the regular smokers $\rightarrow$ the focal actor prefers to have the same behavior as all these friends.
- In the case where the friends do not smoke at all, the preference toward imitating their behavior is less strong than in the case where all the friends smoke a lot.

Recent, current and near future

- Distinction among creating and deleting ties
- Estimation procedures (MLE and bayesian estimation)
- Goodness of fit of the model
- Model selection
- Non-directed relations
- Time-heterogeneity tests
- ...