Network Modeling

Viviana Amati    Jürgen Lerner    Bobo Nick

Dept. Computer & Information Science
University of Konstanz

Winter 2011/2012
(version 13 February 2012)
Outline

Introduction
   Where are we going?

The Stochastic actor-oriented model
   Data and model definition
   Model specification
   Simulating network evolution
   Parameter estimation: MoM and MLE
   Parameter interpretation

Extending the model: analyzing the co-evolution of networks and behavior
   Motivation
   Selection and influence
   Model definition and specification
   Simulating the co-evolution of networks and behavior
   Parameter estimation
   Parameter interpretation
Outline

Introduction
  Where are we going?

The Stochastic actor-oriented model

Extending the model: analyzing the co-evolution of networks and behavior
Where we are

Model | Main feature | Real data
--- | --- | ---
G(n, p) | ties are independent | ties are usually dependent
Preferential attachment | based on there are other degree distribution structural properties | ERGM class of models reasonable representation of the data
These are models for cross-sectional data
Where we are

<table>
<thead>
<tr>
<th>Model</th>
<th>Main feature</th>
<th>Real data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(n, p)$</td>
<td>ties are independent</td>
<td>ties are usually dependent</td>
</tr>
</tbody>
</table>
Where we are

<table>
<thead>
<tr>
<th>Model</th>
<th>Main feature</th>
<th>Real data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(n,p)$</td>
<td>ties are independent</td>
<td>ties are usually dependent</td>
</tr>
<tr>
<td>Preferential attachment</td>
<td>based on degree distribution</td>
<td>there are other structural properties</td>
</tr>
</tbody>
</table>
Where we are

<table>
<thead>
<tr>
<th>Model</th>
<th>Main feature</th>
<th>Real data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(n, p)$</td>
<td>ties are independent</td>
<td>ties are usually dependent</td>
</tr>
<tr>
<td>Preferential attachment</td>
<td>based on degree distribution</td>
<td>there are other structural properties</td>
</tr>
<tr>
<td>ERGM</td>
<td>class of models</td>
<td>reasonable representation of the data</td>
</tr>
</tbody>
</table>
Where we are

<table>
<thead>
<tr>
<th>Model</th>
<th>Main feature</th>
<th>Real data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(n, p)$</td>
<td>ties are independent</td>
<td>ties are usually dependent</td>
</tr>
<tr>
<td>Preferential attachment</td>
<td>based on degree distribution</td>
<td>there are other structural properties</td>
</tr>
<tr>
<td>ERGM</td>
<td>class of models</td>
<td>reasonable representation of the data</td>
</tr>
</tbody>
</table>

These are models for cross-sectional data
Where we are going
Where we are going
Where we are going

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

We need a model for longitudinal data
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/~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 friends
  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 reci-procity?
Questions

- Is there any tendency in friendship formation towards reciprocity?

  ![Diagram](image1)

- Is there any tendency in friendship formation towards transitivity?

  ![Diagram](image2)
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 \((\Omega, P)\) be a probability space. A (real-valued) random variable (r.v.) is a function \(X : \Omega \rightarrow \mathbb{R}\).
Background: random variable

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

Example
Background: random variable

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

Example

\(X: \text{sum of the two dice}\)
Background: random variable

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

Example
Background: random variable

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

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

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

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

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

![Graph of the exponential distribution](image)
The c.d.f. of \( X \) is

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

In fact

\[
F_X(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f_X(x)dx = \int_{0}^{x} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}
\]
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.
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.

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

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

**Definition**

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

\[ \forall t \in T \mapsto X(t) : \Omega \rightarrow \mathbb{R} \]

\( T = \) index set (usually interpreted as time)
\( S = \) state space
Background: stochastic process

Definition

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

\[
\forall t \in T \mapsto X(t) : \Omega \rightarrow \mathbb{R}
\]

\( T = \text{index set (usually interpreted as time)} \)

\( S = \text{state space} \)
Background: stochastic process

Example

\[ X(t) = \text{the outcome of flipping a coin} \]
Background: stochastic process

Example

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

\( S = \{-1, 1\}, \) where \(-1 =\) tail \(1 =\) head

\( T = \{1, 2, \ldots\} \)
Background: stochastic process

Example

$X(t) = \text{the outcome of flipping a coin}$

$S = \{-1, 1\}$, where $-1 = \text{tail}$, $1 = \text{head}$

$T = \{1, 2, \cdots\}$
Background: stochastic process

Example

\[ X(t) = \text{the outcome of flipping a coin} \]

\[ S = \{-1, 1\}, \text{ where } -1 = \text{tail} \ 1 = \text{head} \]

\[ T = \{1, 2, \cdots\} \]

\{X(t), t \in T\} \text{ 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

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

$S = \{0, 1, 2, \ldots\}$

$T = [0, 12]$
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.

$S = \{0, 1, 2, \cdots \}$

$T = [0, 12]$
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.

\( S = \{0, 1, 2, \cdots\} \)

\( T = [0, 12] \)

\( \{X(t), t \in T\} \) 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

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

$S = [0, A], A = \text{maximum amount}$

$T = \{1, 2, \ldots\}$
Background: stochastic process

Example

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

\( S = [0, A], \ A = \) maximum amount

\( T = \{1, 2, \ldots\} \)
Background: stochastic process

Example

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

$S = [0, A], \ A =$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

\( S = [18, 25] \)

\( T = [0, 24] \)
Background: stochastic process

Example

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

\( S = [18, 25] \)
\( T = [0, 24] \)
Background: stochastic process

Example

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

\[ S = [18, 25] \]
\[ T = [0, 24] \]

\( \{ X(t), t \in T \} \) 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 $S$, 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 $x \in S$, 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)

\[ \{X(t), t \geq 0\} \] is a continuous-time Markov chain

Why?
1. state space: \( S = \{0,1,2,\ldots,B\} \), where \( B \) is the total number of goals scored by the player during his career
Background: continuous-time Markov Chains

Definition
A \textit{continuous-time} Markov chain \( \{X_t, t \geq 0\} \) is a \textbf{finite state, continuous-time} stochastic process having the \textbf{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: \( S = \{0, 1, 2, \ldots, B\} \), 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: \( S = \{0, 1, 2, \ldots, B\} \), 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

\[ X(t) \]

- X(t) = 1 at t = 0
- X(t) = 2 at t = 3
- X(t) = 3 at t = 12
- X(t) = 4 at t = 23
Background: describing a continuous-time Markov 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 $\lambda_i$.

2. Jump chain: is described by a jump matrix $P = (p_{ij}: i, j \in S)$ which satisfies the following properties:
   - $p_{ij} \geq 0$
   - $p_{ij} = 1 \quad \forall i, j \in S$
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 $\lambda_i$;
Background: describing a continuous-time Markov 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 $\lambda_i$.

2. **Jump chain**: is described by a jump matrix $P = (p_{ij} : i, j \in S)$ which satisfies the following properties:

   
   $$
   p_{ij} \geq 0 \quad \sum_{j \in S} p_{ij} = 1 \quad \forall i, j \in S
   $$
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 $Q = (q_{ij} : i, j \in S)$ 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. $0 < -q_{ii} < \infty \; \forall i \in S$
2. $q_{ij} > 0 \; \forall i \neq j, \; i, j \in S$
3. $\sum_{j \in S} q_{ij} = 0 \; \forall i \in S$

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

Introduction

The Stochastic actor-oriented model
  Data and model definition
  Model specification
  Simulating network evolution
  Parameter estimation: MoM and MLE
  Parameter interpretation

Extending the model: analyzing the co-evolution of networks and behavior
Social network consists of a set of actors $\mathcal{N}$ and the relation $\mathcal{R}$ existing among them

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

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

$$X = \begin{pmatrix}
- & 0 & 0 & 0 & 0 \\
1 & - & 1 & 0 & 0 \\
0 & 0 & - & 0 & 0 \\
0 & 1 & 1 & - & 0 \\
1 & 1 & 0 & 0 & -
\end{pmatrix}$$
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)
Longitudinal (or panel) network data = $M \geq 2$ 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
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).
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 $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \text{ is a finite set}$$
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**: \( X \) is the set of all possible adjacency matrix (digraphs) defined on the set of actors \( \mathcal{N} \)

\[
X = 2^{n(n-1)} \Rightarrow X \text{ is a finite set}
\]
Model definition: assumptions

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

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

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

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

$x = \text{current state of the network}$
Model definition: assumptions

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

$x = \text{current state of the network}$
Model definition: assumptions

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

$x = \text{current state of the network}$
4. *Absence of co-occurrence*: no more than one tie can change at any given moment $t$

$x = \text{current state of the network}$
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*
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 $x_{ij}$

<table>
<thead>
<tr>
<th>Micro-step</th>
<th>Continuous-time Markov chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>- the time at which $i$ had the opportunity to change</td>
<td>- the waiting time until the next opportunity for a change made by an actor $i$ (holding time)</td>
</tr>
</tbody>
</table>
**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 $x_{ij}$

<table>
<thead>
<tr>
<th>Micro-step</th>
<th>Continuous-time Markov chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>- the time at which $i$ had the opportunity to change</td>
<td>- the waiting time until the next opportunity for a change made by an actor $i$ (holding time)</td>
</tr>
<tr>
<td>- the precise change $i$ made</td>
<td>- the probability of changing the link $x_{ij}$ given the opportunity for changing (jump chain)</td>
</tr>
</tbody>
</table>
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 $x_{ij}$.

<table>
<thead>
<tr>
<th>Micro-step</th>
<th>Continuous-time Markov chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>- the time at which $i$ had the opportunity to change</td>
<td>- the waiting time until the next opportunity for a change made by an actor $i$ (holding time)</td>
</tr>
<tr>
<td>- the precise change $i$ made</td>
<td>- the probability of changing the link $x_{ij}$ given the opportunity for changing (jump chain)</td>
</tr>
</tbody>
</table>

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 $i \sim \text{Exp}(\lambda_i)$
$\Rightarrow \lambda_i$ is the expected frequency of changes by actor $i$ between observations

Simple specification: all actors have the same rate of change $\lambda$

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

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 $\alpha$

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

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:

\[ x_{ij} = 0 \text{ is changed into } x_{ij} = 1 \quad \text{tie creation} \]

\[ x_{ij} = 1 \text{ is changed into } x_{ij} = 0 \quad \text{tie deletion} \]
Model definition: objective function

Which tie is changed?

Change a tie means turning it into its opposite:

\[ x_{ij} = 0 \text{ is changed into } x_{ij} = 1 \quad \text{tie creation} \]

\[ x_{ij} = 1 \text{ is changed into } x_{ij} = 0 \quad \text{tie deletion} \]

Given that \( i \) has the opportunity to change:

<table>
<thead>
<tr>
<th>Possible choices of ( i )</th>
<th>Possible reachable states</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n - 1 ) changes</td>
<td>( n - 1 ) networks ( x(i \sim j) )</td>
</tr>
</tbody>
</table>
Model definition: objective function

Which tie is changed?

Change a tie means turning it into its opposite:

\[ x_{ij} = 0 \text{ is changed into } x_{ij} = 1 \text{ tie creation} \]

\[ x_{ij} = 1 \text{ is changed into } x_{ij} = 0 \text{ tie deletion} \]

Given that \( i \) has the opportunity to change:

<table>
<thead>
<tr>
<th>Possible choices of ( i )</th>
<th>Possible reachable states</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n - 1 ) changes</td>
<td>( n - 1 ) networks ( x(i \sim j) )</td>
</tr>
<tr>
<td>1 non-change</td>
<td>1 network equal to ( x )</td>
</tr>
</tbody>
</table>
Model definition: objective function

Current state (x)

Next state (x')

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & - & 1 & 1 & 0 \\
2 & 1 & - & 0 & 0 \\
3 & 1 & 0 & - & 0 \\
4 & 0 & 0 & 1 & - \\
\end{array}
\]

x(1 \rightarrow 2)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & & 0 & 0 & 0 \\
2 & 1 & - & 0 & 0 \\
3 & 1 & 0 & - & 0 \\
4 & 0 & 0 & 1 & - \\
\end{array}
\]

x(1 \rightarrow 3)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & & 0 & 1 & 1 \\
2 & 1 & - & 0 & 0 \\
3 & 1 & 0 & - & 0 \\
4 & 0 & 0 & 1 & - \\
\end{array}
\]

x(1 \rightarrow 4)

do nothing

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & & 0 & 1 & 0 \\
2 & 1 & - & 0 & 0 \\
3 & 1 & 0 & - & 0 \\
4 & 0 & 0 & 1 & - \\
\end{array}
\]
Model definition: objective function

Next state ($x'$)

Current state ($x$)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

$x(1 \rightarrow 2)$

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

$x(1 \rightarrow 3)$

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

$x(1 \rightarrow 4)$

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

do nothing
Model definition: objective function

Current state (x)

Next state (x')

Transition matrix

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & - & 1 & 0 \\
2 & 1 & - & 0 \\
3 & 1 & 0 & - \\
4 & 0 & 0 & 1 \\
\end{array}
\]

\(p_{xx'} > 0\)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & - & 0 & 0 \\
2 & 1 & - & 0 \\
3 & 1 & 0 & - \\
4 & 0 & 0 & 1 \\
\end{array}
\]

\(p_{xx'} > 0\)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & - & 0 & 1 \\
2 & 1 & - & 0 \\
3 & 1 & 0 & - \\
4 & 0 & 0 & 1 \\
\end{array}
\]

\(p_{xx'} > 0\)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & - & 0 & 1 \\
2 & 1 & - & 0 \\
3 & 1 & 0 & - \\
4 & 0 & 0 & 1 \\
\end{array}
\]

do nothing

\(p_{xx'} > 0\)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & 1 & 0 & 0 \\
2 & 1 & 0 & - \\
3 & 0 & 0 & 1 \\
4 & 0 & 1 & - \\
\end{array}
\]

\(p_{xx'} = 0\)
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

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

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

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

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

Observation: $p_{ii}$ is interpreted as the probability of not changing
Objective function specification

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

**Endogenous effects** = dependent on the network structures

- Outdegree effect

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

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

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

\[ s_{i\_trans}(x') = \sum_{j,h} x'_i x'_i x'_j \]

- three cycle-effect

\[ s_{i\_cyc}(x') = \sum_{j,h} x'_i x'_j x'_h \]
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]
Objective function specification

Example

\[
\beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25
\]

<table>
<thead>
<tr>
<th>(i \rightarrow j)</th>
<th>(S_{i\text{-out}})</th>
<th>(S_{i\text{-rec}})</th>
<th>(S_{i\text{-trans}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (\rightarrow) 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 (\rightarrow) 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (\rightarrow) 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (\rightarrow) 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Objective function specification

Example

\[ \beta_{\text{out}} = -1 \quad \beta_{\text{rec}} = +0.5 \quad \beta_{\text{trans}} = -0.25 \]
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]

<table>
<thead>
<tr>
<th>( i \rightarrow j )</th>
<th>( S_{i_out} )</th>
<th>( S_{i_rec} )</th>
<th>( S_{i_trans} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 \rightarrow 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 \rightarrow 2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 \rightarrow 3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 \rightarrow 4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]

<table>
<thead>
<tr>
<th>( i \rightarrow j )</th>
<th>( S_{i_out} )</th>
<th>( S_{i_rec} )</th>
<th>( S_{i_trans} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 \rightarrow 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 \rightarrow 2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 \rightarrow 3</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1 \rightarrow 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]

<table>
<thead>
<tr>
<th>i \rightarrow j</th>
<th>S_{i, out}</th>
<th>S_{i, rec}</th>
<th>S_{i, trans}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 \rightarrow 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 \rightarrow 2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 \rightarrow 3</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1 \rightarrow 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Objective function specification

Example

\[ \beta_{\text{out}} = -1 \quad \beta_{\text{rec}} = +0.5 \quad \beta_{\text{trans}} = -0.25 \]

<table>
<thead>
<tr>
<th>( i \rightarrow j )</th>
<th>( S_{i,\text{out}} )</th>
<th>( S_{i,\text{rec}} )</th>
<th>( S_{i,\text{trans}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 \rightarrow 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 \rightarrow 2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 \rightarrow 3</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1 \rightarrow 4</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Objective function specification

Example

\[ \beta_{out} = -1 \quad \beta_{rec} = +0.5 \quad \beta_{trans} = -0.25 \]

\begin{tabular}{|c|c|c|c|c|}
\hline
\[ i \rightarrow j \] & \[ S_{i\_out} \] & \[ S_{i\_rec} \] & \[ S_{i\_trans} \] & \[ f_i \] \\
\hline
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 \\
\hline
\end{tabular}
Objective function specification

Example

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

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

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 the friend of my friend is also my friend.
- Girls trust girls.
- Boys trust boys.
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

<table>
<thead>
<tr>
<th>Theory</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>the friend of my friend is also my friend</td>
<td>transitive effect</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Theory</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>the friend of my friend is also my friend</td>
<td>⇒ transitive effect</td>
</tr>
<tr>
<td>girls trust girls boys trust boys</td>
<td>⇒ covariate-related similarity</td>
</tr>
</tbody>
</table>
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

   $M$ time points $\implies$ we must specify $M-1$ rate functions

   $\lambda_1, \ldots, \lambda_{M-1}$
Parameter interpretation (part 1)

It is assumed that:

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

   \[ M \text{ time points} \implies \text{we must specify} \ M - 1 \text{ rate functions} \]
   
   $\lambda_1, \ldots, \lambda_{M-1}$

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

   \[ \beta_1, \ldots, \beta_K \]

   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?

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

- The parameter \( \beta_k \) 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
Parameter interpretation (part 1)

How to interpret the parameter of the SAOM?

- The parameter $\lambda$ is the expected number of opportunities to change for each actor between two consecutive time points.
- The parameter $\beta_k$ 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.
Parameter interpretation (part 1)

How to interpret the parameter of the SAOM?

- The parameter $\lambda$ is the expected number of opportunities to change for each actor between two consecutive time points.
Parameter interpretation (part 1)

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

- The parameter $\beta_k$ 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:* $x(t_0)$, $\lambda, \beta, n$

*Output:* $x^{sim}(t_1)$

$t \leftarrow 0$

$x \leftarrow x(t_0)$

**while** condition = TRUE **do**

$$dt \sim \text{Exp}(n\lambda)$$

$$i \sim \text{Uniform}(1, \ldots, n)$$

$$j \sim \text{Multinomial}(p_{i1}, \ldots, p_{in})$$

**if** $i \neq j$ **then**

$$x = x(i \sim j)$$

**else**

$$x = x$$

$$t \leftarrow t + dt$$

$x^{sim}(t_1) \leftarrow x$

**return** $x^{sim}(t_1)$
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 \( x(t_1) \)

\[
\sum_{i,j=1}^{n} |X_{ij}^{obs}(t_1) - X_{ij}(t_0)| = \sum_{i,j=1}^{n} |X_{ij}^{sim}(t_1) - X_{ij}(t_0)|
\]

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 $\theta$

Different estimation methods:
- the Method of Moments (MoM)
- the Maximum Likelihood Estimation (MLE)
Background: population and sample moments

**Definition**

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

Definition

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

The \( k \)-th population moment is:

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

\[
E[X^k] = \int_{-\infty}^{+\infty} x^k f_X(x) \quad \text{(for the continuous case)}
\]
Background: population and sample moments

Definition

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

The \textit{k-th population moment} is:

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

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

The corresponding \textit{k-th sample moment} is

\[
\mu_k = \frac{1}{n} \sum_{i=1}^{n} X_i^k
\]
To estimate $\theta$, 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

\[
E[X^1] = \mu_1 \\
E[X^2] = \mu_2 \\
\vdots \\
E[X^p] = \mu_p
\]

and solving the resulting equations for the unknown parameters.
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 $\lambda$.

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

\[
\begin{align*}
  x_1 &= 12.1 & x_2 &= 5.7 & x_3 &= 17.8 & x_4 &= 16.5 \\
  x_5 &= 31.6 & x_6 &= 7.7 & x_7 &= 11.9 & x_8 &= 22.7
\end{align*}
\]

What is the estimate for $\lambda$ 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 \( \lambda \) is:
\[ \frac{1}{\hat{\lambda}} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \Leftrightarrow \quad \lambda = \frac{n}{\sum_{i=1}^{n} X_i} \]
and the corresponding estimate is
\[ \hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{8}{126} = 0.063 \]
Background: Method of Moments (MoM)

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

Population moment

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

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

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 $\theta$, i.e. higher values of $\theta$ lead to higher values of $s(X)$. 
Remark

- An estimator is a function of the sample, e.g. \[ \frac{n}{\sum_{i=1}^{n} X_i} \]

- An estimate is the realized value of an estimator, e.g. \[ \frac{n}{\sum_{i=1}^{n} x_i} \]

- 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):

\[
x_1 = 9.5 \quad x_2 = 7.2 \quad x_3 = 13.4 \quad x_4 = 10.2
\]

\[
x_5 = 15.0 \quad x_6 = 16.3 \quad x_7 = 13.9 \quad x_8 = 34.5
\]

The new estimate for the parameter \( \lambda \) is

\[
\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{8}{120} = 0.067
\]

This value is close to the previous but it is not the same!
Estimating the parameter of the SAOM using MoM

**Aim:** estimate $\theta$ using the MoM

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

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 $\theta$.

Which statistics are suitable for estimating $\theta$?

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 MoM

**Aim:** estimate $\theta$ 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 $\theta$. 
Estimating the parameter of the SAOM using MoM

**Aim:** estimate $\theta$ 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 $\theta$.

Which statistics are suitable for estimating $\theta$?
Estimating the parameter of the SAOM using MoM

**Aim:** estimate \( \theta \) using the MoM

\[
\theta = (\lambda_1, \ldots, \lambda_{M-1}, \beta_1, \ldots, \beta_K)
\]

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 \( \theta \).

Which statistics are suitable for estimating \( \theta \)?

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 $\lambda$ describes the frequencies at which changes happen.

\[ s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{i,j=1 \atop i \neq j}^n \left| X_{ij}(t_1) - X_{ij}(t_0) \right| \]
Estimating the parameter of the SAOM using the MoM

- The rate parameter $\lambda$ describes the frequencies at which changes happen.

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

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{\lambda}$</td>
<td>94</td>
<td>135</td>
<td>171</td>
</tr>
</tbody>
</table>

$\Rightarrow$ a high value of $\lambda$ 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 $\beta_k$ 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 $\beta_k$ 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 $\beta_{out}$.

The corresponding statistic is

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

<table>
<thead>
<tr>
<th>$\beta_{out}$</th>
<th>$\beta_{out} = -2.5$</th>
<th>$\beta_{out} = -2$</th>
<th>$\beta_{out} = -1.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{out}$</td>
<td>195</td>
<td>214</td>
<td>234</td>
</tr>
</tbody>
</table>

$\Rightarrow$ a high value of $\beta_{out}$ 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

\[
s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1 \atop i \neq j}^n |X_{ij}(t_1) - X_{ij}(t_0)|
\]

\[
\ldots
\]

\[
s_{\lambda_{M-1}}(X(t_M), X(t_{M-1})|X(t_{M-1}) = x(t_{M-1})) = \sum_{i,j=1 \atop i \neq j}^n |X_{ij}(t_M) - X_{ij}(t_{M-1})|
\]
Estimating the parameter of the SAOM using the MoM

Generalizing to $M - 1$ periods:

- Statistics for the rate function parameters

\[ s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1 \atop i \neq j}^{n} |X_{ij}(t_1) - X_{ij}(t_0)| \]

\[ \ldots \]

\[ s_{\lambda_{M-1}}(X(t_M), X(t_{M-1})|X(t_{M-1}) = x(t_{M-1})) = \sum_{i,j=1 \atop i \neq j}^{n} |X_{ij}(t_M) - X_{ij}(t_{M-1})| \]

- 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 $\theta$ is defined as the solution of the system of $M + K - 1$ equations

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

with $m = 1, \ldots, M - 1$ and $k = 1, \ldots, K$
Estimating the parameter of the SAOM using the MoM

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

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

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

and to model network evolution considering the outdegree and the reciprocity effects.
Estimating the parameter of the SAOM using the MoM

**Example**

Statistics:

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

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

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

\[
\sum_{m=1}^2 s_{m\_rec}(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^2 \sum_{i=1}^4 \sum_{j=1}^4 X_{ij}(t_m)X_{ji}(t_m)
\]
Estimating the parameter of the SAOM using the MoM

Example

Observed values of the statistics:

\[
\begin{align*}
    s_{\lambda_1} &= 5 \\
    s_{\lambda_2} &= 4 \\
    \sum_{m=1}^{2} s_{m_{\text{out}}} &= 6 + 8 = 14 \\
    \sum_{m=1}^{2} s_{m_{\text{rec}}} &= 2 + 3 = 5
\end{align*}
\]
Estimating the parameter of the SAOM using the MoM

Example
We look for the value of $\theta$ that satisfies the system:

\[
\left\{
\begin{align*}
E_{\theta} \left[ s_{\lambda_1} (X(t_1), X(t_0) \mid X(t_0) = x(t_0)) \right] &= 5 \\
E_{\theta} \left[ s_{\lambda_2} (X(t_2), X(t_1) \mid X(t_1) = x(t_1)) \right] &= 4 \\
E_{\theta} \left[ \sum_{m=1}^{2} s_{\text{m-out}} (X(t_m) \mid X(t_{m-1}) = x(t_{m-1})) \right] &= 14 \\
E_{\theta} \left[ \sum_{m=1}^{2} s_{\text{m-rec}} (X(t_m) \mid X(t_{m-1}) = x(t_{m-1})) \right] &= 5
\end{align*}
\right.
\]
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 ⇒ stochastic

2. The value of \( \theta \) 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 ⇒ stochastic

2. The value of $\theta$ 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 $\theta$, 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

$$\overline{S} = \frac{1}{q} \sum_{l=1}^{q} S^{(l)}$$
Stochastic approximation method for the SAOM

1. The expected values of the statistics are approximated via Monte Carlo methods ⇒ stochastic
2. The value of \( \theta \) 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 \( \theta \), 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
      \[
      \bar{S} = \frac{1}{q} \sum_{l=1}^{q} S^{(l)} \rightarrow E_\theta [S]
      \]
Stochastic approximation method for the SAOM

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

Example

1. Given $x(t_0)$ and $\theta$, 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 $\theta$, 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 $\theta$, 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

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

3. Approximate the expected value by

   $$\bar{S}_{out} = \frac{1}{q} \sum_{i=1}^{q} S^{(l)}_{out} \to E_\theta[S_{out}]$$
2. Updating the value of $\theta$

The value of $\theta$ is updated by the Robbins-Monro step:

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

where

- $a_i$ is a sequence of positive numbers such that:

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

- $D$ denotes the diagonal matrix of the first order derivative matrix of $S$ with respect to $\theta$:

$$D = \frac{\partial}{\partial \theta} E_{\theta}[S|X(t_0) = x(t_0)]$$
2. Updating the value of $\theta$

**Example**

We look for the value of $\theta$ that satisfies the system:

\[
\begin{align*}
E_\theta \left[ s_{\lambda_1}(X(t_1), X(t_0) | X(t_0) = x(t_0)) \right] &= 5 \\
E_\theta \left[ s_{\lambda_2}(X(t_2), X(t_1) | X(t_1) = x(t_1)) \right] &= 4 \\
E_\theta \left[ \sum_{m=1}^{2} s_{m\_out}(X(t_m) | X(t_{m-1}) = x(t_{m-1})) \right] &= 14 \\
E_\theta \left[ \sum_{m=1}^{2} s_{m\_rec}(X(t_m) | X(t_{m-1}) = x(t_{m-1})) \right] &= 5
\end{align*}
\]
Stochastic approximation method for the SAOM

2. Updating the value of $\theta$

Example

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

\[
\bar{S}_{\lambda_1} = 6.211 \quad \bar{S}_{\lambda_2} = 4.567 \\
\bar{S}_{\beta_{out}} = 13.806 \quad \bar{S}_{\beta_{rec}} = 4.702
\]
Stochastic approximation method for the SAOM

2. Updating the value of $\theta$

Example

- Initial guess $\theta_0 = (4.5, 3.2, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to $\theta_0$
- Approximation of the expected values $E_{\theta_0}[\bar{S}]$
  \[
  \bar{S}_{\lambda_1} = 6.211 \quad \bar{S}_{\lambda_2} = 4.567 \\
  \bar{S}_{\beta_{out}} = 13.806 \quad \bar{S}_{\beta_{rec}} = 4.702 
  \]
- Approximation of the moment equation
  \[
  \bar{S}_{\lambda_1} - 5 = 1.211 \quad \bar{S}_{\lambda_2} - 4 = 0.567 \\
  \bar{S}_{\beta_{out}} - 14 = -0.194 \quad \bar{S}_{\beta_{rec}} - 5 = -0.298 
  \]
Stochastic approximation method for the SAOM

2. Updating the value of $\theta$

Example

- Initial guess $\theta_0 = (4.5, 3.2, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to $\theta_0$
- Approximation of the expected values $E_{\theta_0}[S]$
  \[
  \bar{S}_{\lambda_1} = 6.211 \quad \bar{S}_{\lambda_2} = 4.567 \\
  \bar{S}_{\beta_{out}} = 13.806 \quad \bar{S}_{\beta_{rec}} = 4.702 
  \]
- Approximation of the moment equation
  \[
  \bar{S}_{\lambda_1} - 5 = 1.211 \quad \bar{S}_{\lambda_2} - 4 = 0.567 \\
  \bar{S}_{\beta_{out}} - 14 = -0.194 \quad \bar{S}_{\beta_{rec}} - 5 = -0.298 
  \]

The Robbins-Monro step

\[
\hat{\theta}_{i+1} = \hat{\theta}_i - a_i D^{-1}(\bar{S}_i - s) 
\]

suggests how to modify the parameter value to satisfy the moment equation through the difference $(\bar{S}_i - s)$
2. Updating the value of $\theta$

Example

- Guess $\theta_1 = (4.1, 2.9, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to $\theta_1$
- Approximation of the expected values $E_{\theta_1}[\bar{S}]$
  
  \[
  \begin{align*}
  \bar{S}_{\lambda_1} &= 5.345 \\
  \bar{S}_{\lambda_2} &= 4.215 \\
  \bar{S}_{\beta_{out}} &= 13.813 \\
  \bar{S}_{\beta_{rec}} &= 4.759
  \end{align*}
  \]
2. Updating the value of $\theta$

Example

- Guess $\theta_1 = (4.1, 2.9, -0.2, 0.9)$
- Simulate the network evolution 1000 times according to $\theta_1$
- Approximation of the expected values $E_{\theta_1}[S]$
  \[
  \bar{S}_{\lambda_1} = 5.345 \quad \bar{S}_{\lambda_2} = 4.215 \\
  \bar{S}_{\beta_{out}} = 13.813 \quad \bar{S}_{\beta_{rec}} = 4.759 
  \]
- Approximation of the moment equation
  \[
  \bar{S}_{\lambda_1} - 5 = 0.345 \quad \bar{S}_{\lambda_2} - 4 = 0.215 \\
  \bar{S}_{\beta_{out}} - 14 = -0.187 \quad \bar{S}_{\beta_{rec}} - 5 = -0.241 
  \]
Stochastic approximation method for the SAOM

2. Updating the value of $\theta$

Example

- Guess $\theta_i = (3.87, 2.56, -0.11, 0.87)$
- Simulate the network evolution 1000 times according to $\theta_1$
- Approximation of the expected values $E_{\theta_i}[\bar{S}]$

\[
\begin{align*}
\bar{S}_{\lambda_1} & = 5.055 & \bar{S}_{\lambda_2} & = 4.018 \\
\bar{S}_{\beta_{out}} & = 14.012 & \bar{S}_{\beta_{rec}} & = 4.974
\end{align*}
\]
2. Updating the value of $\theta$

Example

- Guess $\theta_i = (3.87, 2.56, -0.11, 0.87)$
- Simulate the network evolution 1000 times according to $\theta_1$
- Approximation of the expected values $E_{\theta_i}[\bar{S}]$
  
  $\bar{S}_{\lambda_1} = 5.055$ 
  $\bar{S}_{\lambda_2} = 4.018$

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

- Approximation of the moment equation
  
  $\bar{S}_{\lambda_1} - 5 = 0.055$ 
  $\bar{S}_{\lambda_2} - 4 = 0.018$

  $\bar{S}_{\beta_{out}} - 14 = 0.012$ 
  $\bar{S}_{\beta_{rec}} - 5 = -0.026$
The Robbins-Monro algorithm

**Phase 1**: given the network at time $t_0$ and an initial guess $\theta_0$ for $\theta$ a small number $q_1$ of steps are made to estimate $D$.

1. Network evolution is simulated from $\theta_0$ 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 $d_{ij} = \epsilon - 1_j (S_{ij} - S_{i0})$

4. Repeat steps 1. to 3. until $i = q_1$

5. Estimate $E_{\theta_0}[S]$ and $D$ by the Monte Carlo method

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

$$\hat{\theta}_{q_1} = \theta_0 - \hat{D} - 1 (S - s)$$
The Robbins-Monro algorithm

**Phase 1**: given the network at time $t_0$ and an initial guess $\theta_0$ for $\theta$ a small number $q_1$ of steps are made to estimate $D$.

1. Network evolution is simulated from $\theta_0$ 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 $d_{ij} = \epsilon_j - 1 \left( S_{ij} - S_{i0} \right)$
4. Repeat steps 1. to 3. until $i = q_1$
5. Estimate $E_{\theta} \left[ S \right]$ and $D$ by the Monte Carlo method $S = \frac{1}{q_1} \sum_{i=1}^{q_1} S_{i0}$  $\hat{d}_{ij} = \frac{1}{q_1} \sum_{i=1}^{q_1} d_{ij}$
6. A new value of $\theta$ is estimated via the Robbins-Monro step with $a_i = 1$: $\hat{\theta}_{q_1} = \theta_0 - \hat{D} - \left( S - s \right)$
The Robbins-Monro algorithm

**Phase 1**: given the network at time $t_0$ and an initial guess $\theta_0$ for $\theta$ a small number $q_1$ of steps are made to estimate $D$.

1. Network evolution is simulated from $\theta_0$ 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 $\theta_0$ for $\theta$ a small number $q_1$ of steps are made to estimate $D$.

1. Network evolution is simulated from $\theta_0$ 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 $E_{\theta} \{ S \}$ by the Monte Carlo method $S = \frac{1}{q_1} \sum_{i=1}^{q_1} S_{i0}$

6. A new value of $\hat{\theta}$ is estimated via the Robbins-Monro step with $a_i = 1$:
   $$\hat{\theta}_{q_1} = \theta_0 - \hat{D} - 1 (S - s)$$
The Robbins-Monro algorithm

**Phase 1:** given the network at time $t_0$ and an initial guess $\theta_0$ for $\theta$ a small number $q_1$ of steps are made to estimate $D$.

1. Network evolution is simulated from $\theta_0$ 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

\[
\bar{S} = \frac{1}{q_1} \sum_{i=1}^{q_1} S_{i0}, \quad \hat{d}_j = \frac{1}{q_1} \sum_{i=1}^{q_1} d_{ij}
\]
**The Robbins-Monro algorithm**

**Phase 1**: given the network at time $t_0$ and an initial guess $\theta_0$ for $\theta$ a small number $q_1$ of steps are made to estimate $D$.

1. Network evolution is simulated from $\theta_0$ 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

\[
\overline{S} = \frac{1}{q_1} \sum_{i=1}^{q_1} S_{i0} \quad \hat{d}_j = \frac{1}{q_1} \sum_{i=1}^{q_1} d_{ij}
\]

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

\[
\hat{\theta}_{q_1} = \theta_0 - \hat{D}^{-1}(\overline{S} - s)
\]
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 \(\hat{\theta}_i\) and compute \(S_{ih}\)

**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 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 \( \hat{\theta}_i \) and compute \( S_{ih} \)
2. Update \( \theta \) according to the Robbins-Monro step
   \[
   \hat{\theta}_{i+1} = \hat{\theta}_i - a_h \hat{D}^{-1}(S_i - s)
   \]
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 \( \hat{\theta}_i \) and compute \( S_{ih} \)
2. Updated \( \theta \) according to the Robbins-Monro step
   \[
   \hat{\theta}_{i+1} = \hat{\theta}_i - a_h \hat{D}^{-1}(\bar{S}_i - s)
   \]
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 $\hat{\theta}_i$ and compute $S_{ih}$
2. Updated $\theta$ according to the Robbins-Monro step
   \[ \hat{\theta}_{i+1} = \hat{\theta}_i - a_h D^{-1}(\bar{S}_i - s) \]

3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$
4. Compute
   \[ \hat{\theta}_h = \frac{1}{i} \sum_{i} \hat{\theta}_i \]
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 $\hat{\theta}_i$ and compute $S_{ih}$
2. Updated $\theta$ according to the Robbins-Monro step
   
   $$\hat{\theta}_{i+1} = \hat{\theta}_i - a_h D^{-1}(\bar{S}_i - s)$$

3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$
4. Compute
   
   $$\hat{\theta}_h = \frac{1}{i} \sum_i \hat{\theta}_i$$

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

$\hat{\theta}_c$ is the eventual estimate for $\theta$
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 $\hat{\theta}_i$ and compute $S_{ih}$
2. Updated $\theta$ according to the Robbins-Monro step
   \[
   \hat{\theta}_{i+1} = \hat{\theta}_i - a_h \hat{D}^{-1}(\overline{S}_i - s)
   \]
3. Repeat steps 1. and 2. until $i > q_h$ or $(S_{ih} - s)(S_{(i-1)h} - s) < 0$
4. Compute
   \[
   \hat{\theta}_h = \frac{1}{i} \sum_{i} \hat{\theta}_i
   \]
5. $a_{h+1} = \frac{a_h}{2}$ is the eventual estimate for $\theta$

$\hat{\theta}_c$ is the eventual estimate for $\theta$

**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 \( \theta \) is the vector of values \( \hat{\theta} \) such that:

\[
\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta)
\]

or equivalently, the vector of values \( \hat{\theta} \) 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 \((x(t_1), \ldots, x(t_M))\) conditional on \(x(t_0)\).

\[
\text{⇓}
\]

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 \(x(t_1)\).
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:

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

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 $\mathcal{V}$
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:

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

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 $v$

$$P(v|x(t_0),x(t_1)) \propto \frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{\lambda} p_{ir,jr}(\beta, x(T_{r-1}))$$
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 $v_i = v$, generate $\tilde{v}$ from the proposal distribution $u(\tilde{v} | v_i)$.
2. Take $v_{i+1} = \begin{cases} \tilde{v} & \text{with probability } \rho(\tilde{v}, v) \\ v & \text{with probability } 1 - \rho(\tilde{v}, v) \end{cases}$

where

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

The transition probabilities of the chain generated by the Metropolis-Hastings algorithm are given by

$$\rho(\tilde{v}, v) u(\tilde{v} | v)$$
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 $v_i = v$, generate $\tilde{v}$ from the proposal distribution $u(\tilde{v} | v_i)$

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

The transition probabilities of the chain generated by the Metropolis-Hastings algorithm are given by $\rho(\tilde{v}, v) u(\tilde{v} | v_i)$
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 \( v_i = v \), generate \( \tilde{v} \) from the proposal distribution \( u(\tilde{v} | v_i) \)

2. Take

\[
\begin{align*}
    v_{i+1} &= \begin{cases} 
        \tilde{v} & \text{with probability } \rho(\tilde{v}, v) \\
        v & \text{with probability } 1 - \rho(\tilde{v}, v)
    \end{cases}
\end{align*}
\]

where

\[
\rho(\tilde{v}, v) = \min \left\{ \frac{P(\tilde{v}) u(\tilde{v} | \tilde{v})}{P(v) u(v | \tilde{v})}, 1 \right\}
\]

The transition probabilities of the chain generated by the Metropolis-Hastings algorithm are given by \( \rho(\tilde{v}, v) u(\tilde{v} | v) \)
Sampling the augmented data

The proposal distribution $u(\tilde{v} | v)$ 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 $v$. 

Example: $v = (2, 4) (2, 3) (1, 1) (4, 2) (3, 2) (1, 4) (2, 4) (2, 3) (1, 3) (2, 4) (3, 3)$

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

$v = (2, 3) (1, 1) (4, 2) (3, 2) (1, 4) (2, 3) (1, 3) (2, 4) (3, 3)$
Sampling the augmented data

The proposal distribution \( u(\tilde{v} | v) \) 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 \( v \)

Example

\[
v = (2,4) \hspace{1em} (2,3) \hspace{1em} (1,1) \hspace{1em} (4,2) \hspace{1em} (3,2) \hspace{1em} (1,4) \hspace{1em} (2,4) \hspace{1em} (2,3) \hspace{1em} (1,3) \hspace{1em} (2,4) \hspace{1em} (3,3)
\]

- 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 \((2,4)\)

\[
\tilde{v} = (2,3) \hspace{1em} (1,1) \hspace{1em} (4,2) \hspace{1em} (3,2) \hspace{1em} (1,4) \hspace{1em} (2,3) \hspace{1em} (1,3) \hspace{1em} (2,4) \hspace{1em} (3,3)
\]
Sampling the augmented data

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

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

Example: $v = (2, 4) (2, 3) (1, 1) (4, 2) (3, 2) (1, 4) (2, 4) (2, 3) (1, 3) (2, 4) (3, 3)$

- 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 $r_2 = 7$
Sampling the augmented data

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

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

Example

$v = (2, 4) (2, 3) (1, 1) (4, 2) (3, 2) (1, 4) (2, 4) (2, 3) (1, 3) (2, 4) (3, 3)$

- 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 $r_2 = 7$

$\tilde{v} = (2, 4) (2, 3) (1, 1) (4, 2) (4, 1) (3, 2) (1, 4) (4, 1) (2, 4) (2, 3) (1, 3) (2, 4) (3, 3)$
Sampling the augmented data

The proposal distribution $u(\widetilde{v} | v)$ 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 $v$.

Example:
$v = (2, 4) (2, 3) (1, 1) (4, 2) (3, 2) (1, 4) (2, 4) (2, 3) (1, 3) (2, 4)$
- Select at random $r$ in $\{3, 11\}$, e.g. $r = 11$
- Delete the elements $(3, 3)$

$\widetilde{v} = (2, 4) (2, 3) (1, 1) (4, 2) (3, 2) (1, 4) (2, 4) (2, 3) (1, 3) (2, 4)$
Sampling the augmented data

The proposal distribution \( u(\tilde{v} | v) \) 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 \( v \)

Example

\[
v = (2,4) (2,3) (1,1) (4,2) (3,2) (1,4) (2,4) (2,3) (1,3) (2,4) (3,3)
\]

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

\[
\tilde{v} = (2,4) (2,3) (1,1) (4,2) (3,2) (1,4) (2,4) (2,3) (1,3) (2,4)
\]
Sampling the augmented data

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

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

The proposal distribution $u(\tilde{v} | v)$ 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**

\[ v = (2,4) (2,3) (1,1) (4,2) (3,2) (1,4) (2,4) (2,3) (1,3) (2,4) (3,3) \]

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

\[ \tilde{v} = (2,4) (2,3) (1,1) (4,2) (3,2) (1,4) (4,4) (2,4) (2,3) (1,3) (2,4) (3,3) \]
Sampling the augmented data

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

5. \textit{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

\textbf{Example}

\( v = (2,4) \ (2,3) \ (1,1) \ (4,2) \ (3,2) \ (1,4) \ (2,4) \ (2,3) \ (1,3) \ (2,4) \ (3,3) \)

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

\( v = (2,4) \ (1,1) \ (2,3) \ (3,2) \ (4,2) \ (1,4) \ (4,4) \ (2,4) \ (2,3) \ (1,3) \ (2,4) \ (3,3) \)
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 \( v \in \) 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:

\[
\rho(\tilde{v}, v)u(\tilde{v} | v)P(v) = \rho(v, \tilde{v})u(v | \tilde{v})P(\tilde{v})
\]

is satisfied.

\[
\rho(\tilde{v}, v)u(\tilde{v} | v)P(v) = \min \left\{ \frac{P(\tilde{v})u(v | \tilde{v})}{P(v)u(v | v)}, 1 \right\} u(\tilde{v} | v)P(v) =
\]

\[
= \min \left\{ \frac{P(\tilde{v})u(v | \tilde{v})}{u(\tilde{v} | v)}, P(v) \right\} u(\tilde{v} | v) =
\]

\[
= \min \left\{ \frac{u(\tilde{v} | \tilde{v})}{u(\tilde{v} | v)}, \frac{P(v)}{P(\tilde{v})} \right\} u(\tilde{v} | v)P(\tilde{v}) =
\]

\[
= \min \left\{ 1, \frac{P(v)u(\tilde{v} | v)}{P(\tilde{v})u(\tilde{v} | \tilde{v})} \right\} u(\tilde{v} | v)P(\tilde{v}) =
\]

\[
= \rho(v, \tilde{v})u(v | \tilde{v})P(\tilde{v})
\]
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 $\nu^{(i)} = (\nu_1^{(i)}, \ldots, \nu_{M-1}^{(i)})$

2. Compute the score function:

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

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

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

The estimate $\hat{\theta}$ is calculated as the average of the $\theta_{i+1}$ 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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rate parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 1</td>
<td>8.5948</td>
<td>(0.7091)</td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 2</td>
<td>7.2115</td>
<td>(0.5751)</td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4147</td>
<td>(0.0387)</td>
<td>-62.3875</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7106</td>
<td>(0.0811)</td>
<td>33.4061</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rate parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter</td>
<td>8.5948</td>
<td>(0.7091)</td>
<td></td>
</tr>
<tr>
<td>period 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter</td>
<td>7.2115</td>
<td>(0.5751)</td>
<td></td>
</tr>
<tr>
<td>period 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4147</td>
<td>(0.0387)</td>
<td>-62.3875</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7106</td>
<td>(0.0811)</td>
<td>33.4061</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rate parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 1</td>
<td>8.5948</td>
<td>(0.7091)</td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 2</td>
<td>7.2115</td>
<td>(0.5751)</td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4147</td>
<td>(0.0387)</td>
<td>-62.3875</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7106</td>
<td>(0.0811)</td>
<td>33.4061</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate parameters:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 1</td>
<td>8.5948</td>
<td>( 0.7091 )</td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 2</td>
<td>7.2115</td>
<td>( 0.5751 )</td>
<td></td>
</tr>
<tr>
<td>Other parameters:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4147</td>
<td>( 0.0387 )</td>
<td>-62.3875</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7106</td>
<td>( 0.0811 )</td>
<td>33.4061</td>
</tr>
</tbody>
</table>

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?)
Interpreting the objective function parameters:

The parameter $\beta_k$ 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 $\beta_k$ 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.

\[
H_0 : \beta_k = 0
\]

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

\[
\begin{align*}
\left| \frac{\beta_k}{s.e.(\beta_k)} \right| & \geq 2 \quad \text{reject } H_0 \\
\left| \frac{\beta_k}{s.e.(\beta_k)} \right| & < 2 \quad \text{fail to reject } H_0
\end{align*}
\]

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

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

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

\[
\begin{align*}
\left| \frac{\beta_k}{s.e.(\beta_k)} \right| & \geq 2 \quad \text{reject } H_0 \\
\left| \frac{\beta_k}{s.e.(\beta_k)} \right| & < 2 \quad \text{fail to reject } H_0
\end{align*}
\]

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

\[
[0.04, 0.22]
\]

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

\( \beta_{rec} \) is not significantly different from 0
**Parameter interpretation: a very simple model**

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rate parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 1</td>
<td>8.5948</td>
<td>(0.7091)</td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 2</td>
<td>7.2115</td>
<td>(0.5751)</td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4147</td>
<td>(0.0387)</td>
<td>-62.3875</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7106</td>
<td>(0.0811)</td>
<td>33.4061</td>
</tr>
</tbody>
</table>

**Objective function parameters:**
- **outdegree parameter:** the observed networks have low density
Parameter interpretation: a very simple model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rate parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 1</td>
<td>8.5948</td>
<td>(0.7091)</td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 2</td>
<td>7.2115</td>
<td>(0.5751)</td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4147</td>
<td>(0.0387)</td>
<td>-62.3875</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7106</td>
<td>(0.0811)</td>
<td>33.4061</td>
</tr>
</tbody>
</table>

**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

\[ \text{Diagram: Reciprocal relationship} \]

\[ \text{Diagram: Transitive triplets effect} \]
Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated → 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

![Diagram showing similarity and direction of influence]
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

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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rate parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 1</td>
<td>10.6809</td>
<td>(1.0425)</td>
<td></td>
</tr>
<tr>
<td>Rate parameter period 2</td>
<td>9.0116</td>
<td>(0.8386)</td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.8597</td>
<td>(0.0608)</td>
<td>-47.0288</td>
</tr>
<tr>
<td>reciprocity</td>
<td>1.9855</td>
<td>(0.0876)</td>
<td>22.6765</td>
</tr>
<tr>
<td>transitive triplets</td>
<td>0.4480</td>
<td>(0.0257)</td>
<td>17.4558</td>
</tr>
<tr>
<td>sex alter</td>
<td>-0.1513</td>
<td>(0.0980)</td>
<td>-1.5445</td>
</tr>
<tr>
<td>sex ego</td>
<td>0.1571</td>
<td>(0.1072)</td>
<td>1.4659</td>
</tr>
<tr>
<td>sex similarity</td>
<td>0.9191</td>
<td>(0.1076)</td>
<td>8.5440</td>
</tr>
<tr>
<td>smoke alter</td>
<td>0.1055</td>
<td>(0.0577)</td>
<td>1.8272</td>
</tr>
<tr>
<td>smoke ego</td>
<td>0.0714</td>
<td>(0.0623)</td>
<td>1.1469</td>
</tr>
<tr>
<td>smoke similarity</td>
<td>0.3724</td>
<td>(0.1177)</td>
<td>3.1647</td>
</tr>
</tbody>
</table>
Parameter interpretation: a more complex model

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

- **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

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

- **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: $\bar{z} = 1.434$ is the mean of the covariate

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

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.9166</td>
<td>0.1546</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1538</td>
<td>0.9224</td>
</tr>
</tbody>
</table>

Table: Gender-related contributions to the objective function

Conclusions:

- preference for similar alters
- the negative value associated to the single tie from a girl to a boy, suggests that girls seem not to like male friends.
Parameter interpretation: a more complex model

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

- **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: $\bar{z} = 1.310$ is the mean of the covariate

$$z_i - \bar{z} = \begin{cases} 
-0.310 & \text{for no smokers} \\
0.690 & \text{for occasional smokers} \\
1.690 & \text{for regular smokers}
\end{cases}$$
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: $\bar{z} = 1.310$ is the mean of the covariate

$$z_i - \bar{z} = \left\{ \begin{array}{ll}
-0.310 & \text{for no smokers} \\
0.690 & \text{for occasional smokers} \\
1.690 & \text{for regular smokers}
\end{array} \right.$$  

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

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

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

where $R_z = z_{max} - z_{min}$
Parameter interpretation: a more complex model

<table>
<thead>
<tr>
<th></th>
<th>no</th>
<th>occasional</th>
<th>regular</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>0.0414</td>
<td>-0.0734</td>
<td>-0.1882</td>
</tr>
<tr>
<td>occasional</td>
<td>-0.0393</td>
<td>0.2183</td>
<td>0.1035</td>
</tr>
<tr>
<td>regular</td>
<td>-0.1200</td>
<td>0.1376</td>
<td>0.3952</td>
</tr>
</tbody>
</table>

**Table:** Smoking-related contributions to the objective function

**Conclusions:**
- preference for similar alters
- this tendency is strongest for high values on smoking behavior
Introduction

The Stochastic actor-oriented model

Extending the model: analyzing the co-evolution of networks and behavior
  Motivation
  Selection and influence
  Model definition and specification
  Simulating the co-evolution of networks and behavior
  Parameter estimation
  Parameter interpretation
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 \textit{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)

⇓

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)

⇓

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

     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
     \[ |\mathcal{C}| = 2^{n(n-1)} \times B^n \]
     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
     $$|\mathcal{C}| = 2^{n(n-1)} \times B^n$$
     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
     
     $$|\mathcal{C}| = 2^n(n-1) \times B^n$$
     
     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*: 

![Diagram](image)
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
     \[ |\mathcal{C}| = 2^n(n-1) \times B^n \]
     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*:

```plaintext
i → j

| t₀ | t₁ |
```

- $i$: initial state
- $j$: final state
Assumptions

1. **Distribution of the process.**
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space* $C$: 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:*
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.
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.
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.

$z(2 \sim 1)$

$z(2 \sim 3)$

$z(2 \sim 3)$

$z(2 \sim 3)$

$(x,z) = \text{current state}$

do nothing
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

<table>
<thead>
<tr>
<th>Occurrence</th>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network changes</td>
<td>Network rate function</td>
</tr>
<tr>
<td>Behavioral changes</td>
<td>Behavioral rate function</td>
</tr>
<tr>
<td></td>
<td>Network objective function</td>
</tr>
<tr>
<td></td>
<td>Behavioral objective function</td>
</tr>
</tbody>
</table>
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 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**

\[ T_{i}^{\text{net}} \] = the waiting time until \( i \) gets the opportunity to make a network change

\[ T_{i}^{\text{net}} \sim \text{Exp}(\lambda_{i}^{\text{net}}) \]

**Behavior rate function**

\[ T_{i}^{\text{beh}} \] = the waiting time until \( i \) gets the opportunity to make a behavior change

\[ T_{i}^{\text{beh}} \sim \text{Exp}(\lambda_{i}^{\text{beh}}) \]
The rate functions

**Network rate function**

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

\[ T_i^{net} \sim \text{Exp}(\lambda_i^{net}) \]

**Behavior rate function**

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

\[ T_i^{beh} \sim \text{Exp}(\lambda_i^{beh}) \]

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

\[ T_{i}^{net} \sim \text{Exp}(\lambda_{i}^{net}) \]

**Behavior rate function**

\( T_{i}^{beh} = \) the waiting time until \( i \) gets the opportunity to make a behavior change

\[ T_{i}^{beh} \sim \text{Exp}(\lambda_{i}^{beh}) \]

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

\( P( \text{\( i \) has the opportunity to change his behavior}) = \frac{\lambda_{i}^{beh}}{\lambda_{tot}} \)
The rate functions (simplest specification)

Network rate function
\[ T_{i}^{net} = \text{the waiting time until } i \text{ gets the opportunity to make a network change} \]
\[ T_{i}^{net} \sim \text{Exp}(\lambda^{net}) \]

Behavior rate function
\[ T_{i}^{beh} = \text{the waiting time until } i \text{ gets the opportunity to make a behavior change} \]
\[ T_{i}^{beh} \sim \text{Exp}(\lambda^{beh}) \]
The rate functions (simplest specification)

**Network rate function**

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

\[ T_i^{\text{net}} \sim \text{Exp}(\lambda^{\text{net}}) \]

**Behavior rate function**

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

\[ T_i^{\text{beh}} \sim \text{Exp}(\lambda^{\text{beh}}) \]

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

\[ \lambda_{tot} = n(\lambda^{\text{net}} + \lambda^{\text{beh}}) \]
The rate functions (simplest specification)

**Network rate function**

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

\[ T_i^{\text{net}} \sim \text{Exp}(\lambda^{\text{net}}) \]

**Behavior rate function**

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

\[ T_i^{\text{beh}} \sim \text{Exp}(\lambda^{\text{beh}}) \]

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

\[ \lambda_{\text{tot}} = n(\lambda^{\text{net}} + \lambda^{\text{beh}}) \]

**Probabilities**

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

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

\[
\begin{align*}
    f_i^{\text{net}}(\beta, x(i \sim j), z) &= \sum_{k=1}^{K} \beta_k s_{ik}^{\text{net}}(x, z) + U_i(t, x, j) \\
\end{align*}
\]

already described for the SAOM
The objective functions

**Behavioral objective function**

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

where
- \( s_{w}^{beh}(x, z) \) are effects
- \( \gamma_w \) 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:

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

\( p_{ll} \) 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 $s_{i, \text{linear}}^{\text{beh}}(x, z)$ and the quadratic shape effect $s_{i, \text{quadratic}}^{\text{beh}}(x, z)$ are defined by

$$s_{i, \text{linear}}^{\text{beh}}(x, z) = z_i \quad \quad s_{i, \text{quadratic}}^{\text{beh}}(x, z) = z_i^2$$

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 $s_{i\_linear}^{\text{beh}}(x, z)$ and the quadratic shape effect $s_{i\_quadratic}^{\text{beh}}(x, z)$ are defined by

$$
s_{i\_linear}^{\text{beh}}(x, z) = z_i
$$

$$
s_{i\_quadratic}^{\text{beh}}(x, z) = z_i^2
$$

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

The specification of the behavioral objective function

- Classical influence effects

1. The average similarity effect \( s_{i\_avsim}^{beh}(x, z) \) 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

\[
s_{i\_avsim}^{beh}(x, z) = \frac{1}{x_i + \sum_{j=1}^{n} x_{ij} (\text{sim}_z(ij) - \text{sim}_z)}
\]

where

\[
\text{sim}_z(ij) = 1 - \frac{|z_i - z_j|}{R_z}
\]

\( R_z \) is the range of the behavior \( z \) and \( \text{sim}_z \) is the mean similarity value.
The objective functions

The specification of the behavioral objective function

- Classical influence effects

1. The average similarity effect $s_{i_{avsim}}^{beh}(x, z)$ 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

$$s_{i_{avsim}}^{beh}(x, z) = \frac{1}{x_i + \sum_{j=1}^{n} x_{ij}} \sum_{j=1}^{n} x_{ij} (sim_z(ij) - sim_z)$$

where

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

$R_z$ is the range of the behavior $z$ and $sim_z$ is the mean similarity value.

2. The total similarity effect $s_{i_{totsim}}^{beh}(x, z)$, 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:

$$s_{i_{totsim}}^{beh}(x, z) = \sum_{j=1}^{n} x_{ij} (sim_z(ij) - sim_z)$$
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*

  \[
  s_{i\_out}(x, z) = z_i \sum_{j=1}^{n} x_{ij}
  \]

  2. *indegree effect*

  \[
  s_{i\_ind}(x, z) = z_i \sum_{j=1}^{n} x_{ji}
  \]

- **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 \( \lambda^{net} \)

3. Generate \( T^{beh} \) according to an exponential distribution with parameter \( \lambda^{beh} \)
Simulating the co-evolution of networks and behavior

4. If \( \min \{ T^{net}, T^{beh} \} = T^{net} \) a **network micro-step** is implemented:

   - Select the actor \( i \in \mathbb{N} \), who makes the changes, with probability
     \[
     P(i \text{ has the opportunity to change } | \text{ network micro-steps}) = \frac{\lambda^{net}}{\lambda_{tot}}
     \]

   - Select the actor \( j \in \mathbb{N} \), to whom \( i \) changes his outgoing tie, with probability:
     \[
     p_{ij} = \frac{\exp \left( \sum_{k=1}^{K} \beta_k s_{ik}(x(i \sim j), z) \right)}{\sum_{h=1}^{n} \exp \left( \sum_{k=1}^{K} \beta_k s_{ik}(x(i \sim h), z) \right)}
     \]

   - If \( i \neq j \) then \( x = x(i \sim j) \). If \( i = j \) then \( x = x \)

   - Set \( t = t + T^{net} \)
Simulating the co-evolution of networks and behavior

4. If \( \min \{ T^{\text{net}}, T^{\text{beh}} \} = T^{\text{beh}} \) a behavioral micro-step is implemented:

   - Select the actor \( i \in \mathcal{N} \), who makes the changes, with probability

   \[
   P(i \text{ has the opportunity to change}|\text{behavioral micro-steps}) = \frac{\lambda^{\text{beh}}}{\lambda^{\text{tot}}}
   \]

   - Determine the behavioral change \( l' \in \{l+1, l-1, l\} \) with probability:

   \[
   p_{ll'}(\gamma; z(l \sim l'); x) = \frac{\exp \left( \sum_{w=1}^{W} \gamma \omega s_{iw}^{\text{beh}}(x, z(l \sim l')) \right)}{\sum_{l'' \in \{l+1, l-1, l\}} \exp \left( \sum_{w=1}^{W} \gamma \omega s_{iw}^{\text{beh}}(x, z(l \sim l'')) \right)}
   \]

   - If \( l \neq l' \) then \( z = z(l \sim l') \). If \( l = l' \) then \( z = z \)

   - Set \( t = t + T^{\text{beh}} \)

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 $x^{\text{sim}}(t_1)$ is equal to the number of entries that differ between $x(t_0)$ and $x(t_1)$

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

   - simulations run on until the number of different entries between $z(t_0)$ and the simulated behavior $z^{\text{sim}}(t_1)$ is equal to the number of entries that differ between $z(t_0)$ and $z(t_1)$

   $$\sum_{i=1}^{n} |z_{i}^{\text{obs}}(t_1) - z_{i}(t_0)| = \sum_{i=1}^{n} |z_{i}^{\text{sim}}(t_1) - z_{i}(t_0)|$$
The parameter estimation (MoM)

**Aim:** estimate $\theta$, the $2(M-1)+K+W$ dimensional vector of parameters of the co-evolution model

**Statistics:**

- Network rate parameters for the period $m$

$$s_{\lambda_m}^{net}(X(t_m), X(t_{m-1})|X(t_{m-1}) = x(t_{m-1})) = \sum_{i,j=1 \atop i \neq j}^{n} |X_{ij}(t_m) - X_{ij}(t_{m-1})|$$

- Behavior rate parameters for the period $m$

$$s_{\lambda_m}^{beh}(Z(t_m), Z(t_{m-1})|Z(t_{m-1}) = z(t_{m-1})) = \sum_{i=1}^{n} |Z_i(t_m) - Z_i(t_{m-1})|$$
The parameter estimation (MoM)

**Aim:** estimate $\theta$, 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
  \[
  \sum_{m=1}^{M-1} s_{mw}^{beh}(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{M-1} s_{mw}^{beh}(X(t_m), X(t_{m-1}))
  \]
Consequently the MoM estimator for $\theta$ is provided by the solution of the system of equations:

\[
\begin{align*}
E_\theta \left[ s_{\lambda_m} (X(t_m), X(t_{m-1}) | X(t_{m-1}) = x(t_{m-1})) \right] &= s_{\lambda_m} (x(t_m), x(t_{m-1})) \\
E_\theta \left[ s_{\lambda_m} (Z(t_m), Z(t_{m-1}) | Z(t_{m-1}) = z(t_{m-1})) \right] &= s_{\lambda_m} (z(t_m), z(t_{m-1})) \\
E_\theta \left[ \sum_{m=1}^{M-1} s^{\text{net}}_{mk} (X(t_m) | X(t_{m-1}) = x(t_{m-1})) \right] &= \sum_{m=1}^{M-1} s^{\text{net}}_{mk} (x(t_m), x(t_{m-1})) \\
E_\theta \left[ \sum_{m=1}^{M-1} s^{\text{beh}}_{mw} (X(t_m) | X(t_{m-1}) = x(t_{m-1})) \right] &= \sum_{m=1}^{M-1} s^{\text{beh}}_{mw} (x(t_m), x(t_{m-1}))
\end{align*}
\]
The parameter estimation (MoM)

Remarks

- The system of equation cannot be solved analytically $\implies$ Robbins-Monro algorithm

- The Maximum likelihood estimation is under construction. At the moment is too slow.
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

\[ J = \frac{N_{11}}{N_{11} + N_{01} + N_{10}} > 0.3 \]

*Jaccard index*
Precondition of the analysis

2. there is interdependence between networks and behavioral variables

\[ I = -\frac{d(\delta)}{(n-1)d} \]

Moran index

where

\[ d(\delta) = \sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z}) \]

is the mean of the cross products of the centered behavioral variable for connected actors

\[ d = \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)} \]

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

If two connected actors $i$ and $j$ are similar in their behaviors we expect that $-z_i > z_i$ and $z_j > z_j$.

$\Rightarrow$ Positive interdependence

$0 < I < 1$
Precondition of the analysis: understanding the Moran index

\[ \frac{\sum_{ij} x_{ij}(z_i - \bar{z})(z_j - \bar{z})}{\sum_{ij} x_{ij}} \cdot \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)} \]

If two connected actors \( i \) and \( j \) are similar 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

\[
\sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z}) > \frac{\sum (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)}
\]

If two connected actors \( i \) and \( j \) are similar 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

\[
\frac{\sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{ij} x_{ij}} > \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)}
\]

If two connected actors \(i\) and \(j\) are similar in their behaviors we expect that

- \(z_i > \bar{z}\) and \(z_j > \bar{z}\)
- \(z_i < \bar{z}\) and \(z_j < \bar{z}\)

\[\Downarrow\]

Positive interdependence

\(0 < I < 1\)
Precondition of the analysis: understanding the Moran index
Precondition of the analysis: understanding the Moran index

\[ \sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z}) \] \[ \sum_{ij} x_{ij} \] \[ \sum_{ij} (z_i - \bar{z})(z_j - \bar{z}) \] \[ n(n-1) \]

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

\[
\frac{\sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{ij} x_{ij}} < \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)}
\]

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

\[
\frac{\sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{ij} x_{ij}} < \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)}
\]

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 \(z_j > \bar{z}\)

\[
\downarrow
\]

Negative interdependence

\(-1 < I < 0\)
Precondition of the analysis: understanding the Moran index
Precondition of the analysis: understanding the Moran index

If:

\[
\frac{\sum_{ij} x_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{ij} x_{ij}} \approx \frac{\sum_{ij} (z_i - \bar{z})(z_j - \bar{z})}{n(n-1)}
\]

the local distribution of the behavioral variable follows the global distribution of the behavioral variable

\[\downarrow\]

Absence of interdependence

\[ I \approx -\frac{1}{n-1} \]
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).

Conclusion: there is considerable dependence between networks and behaviors.
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

\[
\begin{array}{ccc}
0.244 & 0.258 & 0.341 \\
\end{array}
\]

**Conclusion**: there is considerable dependence between networks and behaviors.
Parameter interpretation: a baseline model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
<tr>
<td><strong>Behavior Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>(1.9689)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>(2.3679)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>(0.4394)</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>(0.3628)</td>
<td>7.8447</td>
</tr>
</tbody>
</table>

Network rate parameters:
- about 9 opportunities for change in the first period
- about 7 opportunities for change in the second period
Parameter interpretation: a baseline model

<table>
<thead>
<tr>
<th>Network Dynamics</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>( 0.6666 )</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>( 0.5466 )</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>( 0.0407 )</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>( 0.0823 )</td>
<td>32.8337</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior Dynamics</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>( 1.9689 )</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>( 2.3679 )</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>( 0.4394 )</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>( 0.3628 )</td>
<td>7.8447</td>
</tr>
</tbody>
</table>
Parameter interpretation: a baseline model

<table>
<thead>
<tr>
<th>Network Dynamics</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior Dynamics</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>(1.9689)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>(2.3679)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>(0.4394)</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>(0.3628)</td>
<td>7.8447</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
<tr>
<td><strong>Behavior Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>(1.9689)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>(2.3679)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>(0.4394)</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>(0.3628)</td>
<td>7.8447</td>
</tr>
</tbody>
</table>
Parameter interpretation: a baseline model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
</tbody>
</table>

**Behavior Dynamics**

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>(1.9689)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>(2.3679)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>(0.4394)</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>(0.3628)</td>
<td>7.8447</td>
</tr>
</tbody>
</table>

Network objective function parameters:

- **outdegree parameter**: the observed networks have low density
- **reciprocity parameter**: strong tendency towards reciprocated ties
Parameter interpretation: a baseline model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>( 0.6666 )</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>( 0.5466 )</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>( 0.0407 )</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>( 0.0823 )</td>
<td>32.8337</td>
</tr>
<tr>
<td><strong>Behavior Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>( 1.9689 )</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>( 2.3679 )</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>( 0.4394 )</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>( 0.3628 )</td>
<td>7.8447</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
<tr>
<td><strong>Behavior Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>(1.9689)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>(2.3679)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>(0.4394)</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>(0.3628)</td>
<td>7.8447</td>
</tr>
</tbody>
</table>

**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

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
</tbody>
</table>

| **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

<table>
<thead>
<tr>
<th><strong>Network Dynamics</strong></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant friendship rate (period 1)</td>
<td>8.6287</td>
<td>(0.6666)</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>7.2489</td>
<td>(0.5466)</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.4084</td>
<td>(0.0407)</td>
<td>-59.12676</td>
</tr>
<tr>
<td>reciprocity</td>
<td>2.7024</td>
<td>(0.0823)</td>
<td>32.8337</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Behavior Dynamics</strong></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.8922</td>
<td>(1.9689)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>4.4813</td>
<td>(2.3679)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.5464</td>
<td>(0.4394)</td>
<td>-8.0712</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8464</td>
<td>(0.3628)</td>
<td>7.8447</td>
</tr>
</tbody>
</table>

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.
- 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: $\bar{z} = 1.377$ is the mean of the covariate

$$z_i - \bar{z} = \begin{cases} 
-0.377 & \text{for no smokers} \\
0.623 & \text{for occasional smokers} \\
1.623 & \text{for regular smokers}
\end{cases}$$
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: $\bar{z} = 1.377$ is the mean of the covariate $z_i - \bar{z} = \begin{cases} 
-0.377 & \text{for no smokers} \\
0.623 & \text{for occasional smokers} \\
1.623 & \text{for regular smokers} 
\end{cases}$

- The contribution to the behavioral objective function is

$$\gamma_{linear}(z_i - \bar{z}) + \gamma_{quadratic}(z_i - \bar{z})^2 =$$

$$= -3.5464_{linear}(z_i - \bar{z}) + 2.8464_{quadratic}(z_i - \bar{z})^2$$
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 → covariate similarity
  
  ![Diagram](image1)

  ![Diagram](image2)

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

  - Covariate ego effect
    
    ![Diagram](image3)

    ![Diagram](image4)

  - Covariate alter effect
    
    ![Diagram](image5)

    ![Diagram](image6)
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 → 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 → average similarity effect

This effect must be controlled for the indegree and the outdegree effects

- Indegree effect

- Outdegree effect
A more complex model

<table>
<thead>
<tr>
<th>Network Dynamics</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant friendship rate (period 1)</td>
<td>10.7166</td>
<td>1.4036</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>9.0005</td>
<td>0.7709</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.8435</td>
<td>0.0572</td>
<td>-49.6776</td>
</tr>
<tr>
<td>reciprocity</td>
<td>1.9683</td>
<td>0.0933</td>
<td>21.1077</td>
</tr>
<tr>
<td>transitive triplets</td>
<td>0.4447</td>
<td>0.0322</td>
<td>13.7964</td>
</tr>
<tr>
<td>sex ego</td>
<td>0.1612</td>
<td>0.1206</td>
<td>1.3368</td>
</tr>
<tr>
<td>sex alter</td>
<td>-0.1476</td>
<td>0.1064</td>
<td>-1.3871</td>
</tr>
<tr>
<td>sex similarity</td>
<td>0.9104</td>
<td>0.0882</td>
<td>10.3244</td>
</tr>
<tr>
<td>smoke ego</td>
<td>0.0665</td>
<td>0.0846</td>
<td>0.7857</td>
</tr>
<tr>
<td>smoke alter</td>
<td>0.1121</td>
<td>0.0761</td>
<td>1.4719</td>
</tr>
<tr>
<td>smokebeh similarity</td>
<td>0.5114</td>
<td>0.1735</td>
<td>2.9479</td>
</tr>
</tbody>
</table>

Rate parameters: the speed at which tie change occurs is higher than the speed at which behavioral change occurs.
A more complex model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>10.7166</td>
<td>( 1.4036 )</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>9.0005</td>
<td>( 0.7709 )</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.8435</td>
<td>( 0.0572 )</td>
<td>-49.6776</td>
</tr>
<tr>
<td>reciprocity</td>
<td>1.9683</td>
<td>( 0.0933 )</td>
<td>21.1077</td>
</tr>
<tr>
<td>transitive triplets</td>
<td>0.4447</td>
<td>( 0.0322 )</td>
<td>13.7964</td>
</tr>
<tr>
<td>sex ego</td>
<td>0.1612</td>
<td>( 0.1206 )</td>
<td>1.3368</td>
</tr>
<tr>
<td>sex alter</td>
<td>-0.1476</td>
<td>( 0.1064 )</td>
<td>-1.3871</td>
</tr>
<tr>
<td>sex similarity</td>
<td>0.9104</td>
<td>( 0.0882 )</td>
<td>10.3244</td>
</tr>
<tr>
<td>smoke ego</td>
<td>0.0665</td>
<td>( 0.0846 )</td>
<td>0.7857</td>
</tr>
<tr>
<td>smoke alter</td>
<td>0.1121</td>
<td>( 0.0761 )</td>
<td>1.4719</td>
</tr>
<tr>
<td>smokebeh similarity</td>
<td>0.5114</td>
<td>( 0.1735 )</td>
<td>2.9479</td>
</tr>
</tbody>
</table>

**Network objective function parameters:** tendency towards reciprocity, transitivity and homophily with respect to gender
A more complex model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 1)</td>
<td>10.7166</td>
<td>( 1.4036 )</td>
<td></td>
</tr>
<tr>
<td>constant friendship rate (period 2)</td>
<td>9.0005</td>
<td>( 0.7709 )</td>
<td></td>
</tr>
<tr>
<td>outdegree (density)</td>
<td>-2.8435</td>
<td>( 0.0572 )</td>
<td>-49.6776</td>
</tr>
<tr>
<td>reciprocity</td>
<td>1.9683</td>
<td>( 0.0933 )</td>
<td>21.1077</td>
</tr>
<tr>
<td>transitive triplets</td>
<td>0.4447</td>
<td>( 0.0322 )</td>
<td>13.7964</td>
</tr>
<tr>
<td>sex ego</td>
<td>0.1612</td>
<td>( 0.1206 )</td>
<td>1.3368</td>
</tr>
<tr>
<td>sex alter</td>
<td>-0.1476</td>
<td>( 0.1064 )</td>
<td>-1.3871</td>
</tr>
<tr>
<td>sex similarity</td>
<td>0.9104</td>
<td>( 0.0882 )</td>
<td>10.3244</td>
</tr>
<tr>
<td>smoke ego</td>
<td>0.0665</td>
<td>( 0.0846 )</td>
<td>0.7857</td>
</tr>
<tr>
<td>smoke alter</td>
<td>0.1121</td>
<td>( 0.0761 )</td>
<td>1.4719</td>
</tr>
<tr>
<td>smokebeh similarity</td>
<td>0.5114</td>
<td>( 0.1735 )</td>
<td>2.9479</td>
</tr>
</tbody>
</table>

**Network objective function parameters:** pupils selected others with similar smoking behavior as friends → evidence for selection process
## Behavior Dynamics

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.9041</td>
<td>1.7402</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>3.8059</td>
<td>1.4323</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.3573</td>
<td>0.5678</td>
<td>-5.9129</td>
</tr>
<tr>
<td>behavior smokebeh quadratic</td>
<td>2.8406</td>
<td>0.4125</td>
<td>6.8864</td>
</tr>
<tr>
<td>behavior smokebeh indegree</td>
<td>0.1711</td>
<td>0.1812</td>
<td>0.9444</td>
</tr>
<tr>
<td>behavior smokebeh outdegree</td>
<td>0.0128</td>
<td>0.1926</td>
<td>0.0662</td>
</tr>
<tr>
<td>behavior smokebeh average</td>
<td>3.4361</td>
<td>1.4170</td>
<td>2.4250</td>
</tr>
<tr>
<td>similarity</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Behavioral objective function parameters:** U-shaped distribution of the smoking behavior
A more complex model

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavior Dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.9041</td>
<td>(1.7402)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>3.8059</td>
<td>(1.4323)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.3573</td>
<td>(0.5678)</td>
<td>-5.9129</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8406</td>
<td>(0.4125)</td>
<td>6.8864</td>
</tr>
<tr>
<td>behavior smokebeh indegree</td>
<td>0.1711</td>
<td>(0.1812)</td>
<td>0.9444</td>
</tr>
<tr>
<td>behavior smokebeh outdegree</td>
<td>0.0128</td>
<td>(0.1926)</td>
<td>0.0662</td>
</tr>
<tr>
<td>behavior smokebeh average similarity</td>
<td>3.4361</td>
<td>(1.4170)</td>
<td>2.4250</td>
</tr>
</tbody>
</table>

**Behavioral objective function parameters:** indegree and outdegree effects are not significant
A more complex model

<table>
<thead>
<tr>
<th>Behavior Dynamics</th>
<th>Estimates</th>
<th>s.e.</th>
<th>t-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate smokebeh (period 1)</td>
<td>3.9041</td>
<td>(1.7402)</td>
<td></td>
</tr>
<tr>
<td>rate smokebeh (period 2)</td>
<td>3.8059</td>
<td>(1.4323)</td>
<td></td>
</tr>
<tr>
<td>behavior smokebeh linear shap</td>
<td>-3.3573</td>
<td>(0.5678)</td>
<td>-5.9129</td>
</tr>
<tr>
<td>behavior smokebeh quadratic shape</td>
<td>2.8406</td>
<td>(0.4125)</td>
<td>6.8864</td>
</tr>
<tr>
<td>behavior smokebeh indegree</td>
<td>0.1711</td>
<td>(0.1812)</td>
<td>0.9444</td>
</tr>
<tr>
<td>behavior smokebeh outdegree</td>
<td>0.0128</td>
<td>(0.1926)</td>
<td>0.0662</td>
</tr>
<tr>
<td>behavior smokebeh average similarity</td>
<td>3.4361</td>
<td>(1.4170)</td>
<td>2.4250</td>
</tr>
</tbody>
</table>

**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:

$$
\gamma_{linear}(z_i - \bar{z}) + \gamma_{quadratic}(z_i - \bar{z})^2 + \gamma_{avsim} \frac{1}{x_{ij+}} \sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) = 
$$

$$
= -3.3573_{linear}(z_i - \bar{z}) + 2.8406_{quadratic}(z_i - \bar{z})^2 + 3.4361 \frac{1}{x_{ij+}} \sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)
$$

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 $z_j$.

<table>
<thead>
<tr>
<th>$z_j$ / $z_i$</th>
<th>no</th>
<th>occasional</th>
<th>regular</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.56</td>
<td>-1.82</td>
<td>-0.51</td>
</tr>
<tr>
<td>2</td>
<td>0.84</td>
<td>-0.10</td>
<td>1.20</td>
</tr>
<tr>
<td>3</td>
<td>-0.88</td>
<td>-1.82</td>
<td>2.92</td>
</tr>
</tbody>
</table>

- The row maximum is assumed at the diagonal for the non-smokers and for the regular smokers → 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
- ...