

## Social Network Analysis – Winter 2022

### Lab 4 Option A:

#### Emergence of Network Dynamics (Stochastic Actor-Oriented Models, SAOM)

#### Either Lab 4 Option A OR Option B Due:

March 4th, at 11:59 PM CT (upload to Canvas)

**Deliverables: a single PDF file. Please include all the plots and needed tables in this file. If you write the code from scratch and do not use the provided R code, please also submit your code as a separate R file.**

In this lab, you'll be building, estimating, and interpreting actor-based longitudinal network models using RSiena. RSiena is used to model stochastic actor-oriented models (SAOM) in order to examine the effects of network ties over time on a certain behavior, or the effect of a certain behavior on tie formation over time.

You will be using data from the *Teenage Friends and Lifestyle Study*<sup>1</sup>. The dataset includes 3 network files containing friendship relationships between 50 teenage girls recorded at three consecutive points in time. (Note that there is a tie from A to B, if A mentions B as a friend).

1. s50-network1.dat
2. s50-network2.dat
3. s50-network3.dat

The data also includes information about:

1. The *drug* use of the 50 female students (s50-drugs.dat). The drug behavior variable has four levels: 1 (non), 2 (tried once), 3 (occasional) and 4 (regular).
2. The *smoking behavior* of the 50 female students (s50-smoke.dat): 1 (does not smoke), 2 (smokes occasionally) and 3 (smokes regularly).
3. The *alcohol consumption* (s50-alcohol.dat). It has five levels: 1 (does not drink alcohol), 2 (once or twice a year), 3 (once a month), 4 (once a week) and 5 (more than once a week).

---

<sup>1</sup>For more details, check this website [https://www.stats.ox.ac.uk/~snijders/siena/s50\\_data.htm](https://www.stats.ox.ac.uk/~snijders/siena/s50_data.htm)

## **PART I: Constructing Hypotheses (30 points)**

Formulate hypotheses using network terminology based on the following SIENA terms (indicated after the forward slash). Please check [the RSiena Manual](#) (p. 120-163) for the terms. You'll test the hypotheses that you'll construct here in the Part II. The first one is done for you:

1. Low Outdegree Density/ outdegree

*Hypothesis 1: The probability of having friendship relation between students will be lower over time than expected by random chance.*

### **Relational Hypotheses:**

2. Reciprocity/reciprocity (5 points)
3. Transitivity/gwespFF (5 points)
4. Ego's drug behavior/ egoX (5 points)
5. Alter's drug behavior/ altX (5 points)
6. Homophily on the basis of drug behavior/sameX (5 points)

### **Drug Behavior Hypothesis:**

7. Assimilation of drug behavior/totSim (5 points)

## **PART II: Hypothesis Testing (70 points)**

In this section, you'll test the hypotheses you formulated using RSiena.

1. A visual inspection of the adjacency matrices may help in highlighting how friendship changes at the three time. Include the sociomatrix plots in your report. Discuss what you observe from the plots (e.g., How does friendship change over time? Are the plots becoming denser over time? Is friendship between students mutual? Is there anyone who is nominated a lot by others? Is there anyone who nominates a lot of friends?) (10 points)
2. Create a *siena* data object including the longitudinal friendship networks and the drug behavioral variable. Then run `print01report` function which creates an output file in your working directory. Using your text editor, open the output file (if you use the provided script, `s50_3_init.out`) where you can see data descriptions. In the output file, how many friendship relations were created and dissolved between period 1 and 2? How many students increased their use of drug or decreased the use of drug between the same periods? (4 points)
3. Using your hypotheses, you can begin to construct a list of parameters (effects) to test using your Siena model. Create a data frame of effects using the `getEffects` function. The created data frame will include a number of extra properties for use with RSiena. Include the effects of triadic closure, the effects of drug use on friendship formation (i.e. the effects of the ego drug behavior, the alter drug behavior, and if ego and alter are the same in their drug behavior –both drug user or non-drug user).

4. Include the effects of all of a node's friends' drug behavior on the node's own drug behavior, the specified model and effects to the data using the function `siena07`. A new window labeled "Siena07" with a picture of an old building should pop up, showing the iterations of simulations R goes through. The window should close after the simulations complete.
5. Type `ans1` to view your results.
  - (a) Include a table including convergence t-ratios and overall maximum convergence ratio in your report.
  - (b) Explain whether your variables and model are converged based on your convergence t-ratios and overall maximum convergence ratio. See the hint (\*).
  - (c) Include another table including the estimates, standard errors and p-values (or estimate/standard error) in your report.
  - (d) Use the estimates and p-value (or estimate/standard error) to explain whether your hypotheses (7 hypotheses in total) are supported or not. Provide interpretations of the estimates and discuss if the results make sense. **When you interpret the results, you should convert log-odds ratios (estimates) into either log-odds or probabilities.**

\*[Remember, a  $|t\text{-ratio}| < 0.1$  means that the parameter converged to a reliable value. An overall maximum convergence ratio  $< 0.25$  indicates the model convergence. A parameter is significant if the  $\frac{|estimate|}{standard\ error} > 1.96$  (i.e.,  $p < 0.05$ )]. **(46 points)**
6. Report the goodness of fit for your model with regard to in-degree and out-degree distributions. Include the plots and interpret the results of each plot. **(10 points)**