Lab 4 - Emergence of Network Dynamics (Stochastic Actor-Oriented Models)

CompSci 396-0: Social Networking Analysis ***Win 2022***

Student Name: Jiaqi Guo NetID: JGR9647

● Responses to Question

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.

■ Part I: Constructing Hypotheses

Relational Hypotheses:

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.
2. **Reciprocity/reciprocity**
   Hypothesis 2: The probability of having a reciprocal friendship between students will be higher over time than expected by random chance.
3. **Transitivity/gwespFF**
   Hypothesis 3: The probability of having a friend of friends, this kind of relationship between students will be higher over time than expected by random chance.
4. **Ego's drug behavior/ egoX**
   Hypothesis 4: The probability of drug users will have more friendship ties than non-drug users will be higher over time.
5. **Alter's drug behavior/ altX**
   Hypothesis 5: Over time, Drug users were more likely to be considered as friend than non-drug users.
6. **Homophily on the basis of drug behavior/sameX**
   Hypothesis 6: If two people have the similar individual attributes, they are more likely to have a friendship or have a relationship between each other over time.

Drug Behavior Hypothesis

7. **Assimilation of drug behavior/totSim**
   Hypothesis 7: Given that there is a relationship between two people, they will be more likely to have a similar individual attributes or have the same drug use behavior.

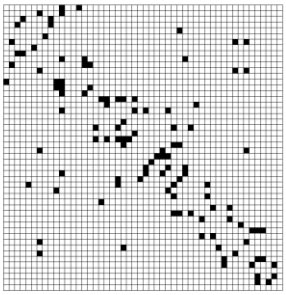
■ Part II: Hypothesis Testing

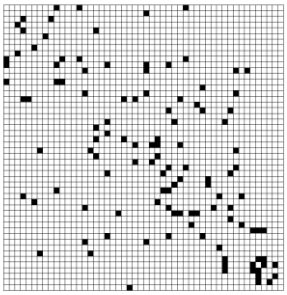
1. (10 points) A visual inspection of the adjacency matrices may help in highlighting how friendship changes at the three time. Include the sociometric plots in your report. Discuss what

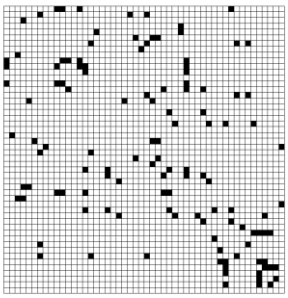
you observe from the plots (e.g., How does friendship change over time? Are the plots becoming denser over time? Is friendship between students mutual? Is there anyone who is nominated a lot by others? Is there anyone who nominates a lot of friends?)

Observation:

1. These plots are becoming denser over time, but the change is not very significant.
2. The friendships between students are usually mutual, as these adjacent matrixes are symmetric.
3. Over time, the number of reciprocal ties will increase.
4. There is nobody who is obviously nominated a lot by others.
5. There is nobody who obviously nominates a lot of friends.







2. (4 points) 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?

For the first question:

During Period 1 > 2: 59 relationships were created, while 56 relationships were dissolved.

During Period 2 > 3: 56 relationships were created, while 50 relationships were dissolved

Tie changes between subsequent observations:

periods	0 => 0	0 => 1	1 => 0	1 => 1	Distance	Jaccard	Missing
1 ==> 2	2278	59	56	57	115	0.331	0 (0%)
2 ==> 3	2278	56	50	66	106	0.384	0 (0%)

For the second question:

During Period 1 > 2: 5 students increased their use of drug, while 2 students decreased the use of drug

During Period 2 > 3: 13 students increased their use of drug, while 3 students decreased the use of drug

periods	actors: down	up	constant	missing ;	steps: down	up	total
1 => 2	2	5	43	0	2	6	8
2 => 3	3	13	34	0	3	17	20

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 several 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 and 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. (46 points) Type ans1 to view your results.
   a) Include a table including convergence t-ratios and overall maximum convergence ratio in your report.

```
Estimates, standard errors and convergence t-ratios

                                                Estimate    Standard    Convergence
                                                              Error       t-ratio
Network Dynamics
   1. rate constant friendship rate (period 1)   6.7419  ( 1.2624   )    0.0231
   2. rate constant friendship rate (period 2)   5.3156  ( 1.1201   )   -0.0816
   3. eval outdegree (density)                  -3.1052  ( 0.2059   )   -0.0227
   4. eval reciprocity                           2.4331  ( 0.1906   )   -0.0027
   5. eval GWESP I -> K -> J (69)                1.5133  ( 0.1701   )   -0.0348
   6. eval drugbeh alter                         0.0721  ( 0.1368   )    0.0418
   7. eval drugbeh ego                           0.1149  ( 0.1573   )    0.0333
   8. eval same drugbeh                          0.5879  ( 0.2581   )   -0.0150

Behavior Dynamics
   9. rate rate drugbeh (period 1)               0.5309  ( 0.2338   )   -0.0244
  10. rate rate drugbeh (period 2)               1.5698  ( 0.6885   )   -0.0198
  11. eval drugbeh linear shape                  1.5242  ( 2.7399   )   -0.0116
  12. eval drugbeh quadratic shape              -0.1767  ( 0.6459   )   -0.0068
  13. eval drugbeh total similarity              6.0974  ( 8.1633   )    0.0203

Overall maximum convergence ratio:    0.1229
```

   b) Explain whether your variables and model are converged based on your convergence ratios and overall maximum convergence ratio. See the hint (*)

Variables: The absolute value for this t-ratio ideally would be less than 0.1, which means that all of the model's variable converge perfectly.

Overall Model: The value of the overall maximum convergence ratio ideally would be less than 0.25, which means our model also converge perfectly.

   c) Include another table including the estimates, standard errors, and p-values (or estimate/standard error) in your report.

	Model 1
constant friendship rate (period 1)	6.70 ***
	(1.16)
constant friendship rate (period 2)	5.36 ***
	(0.93)
outdegree (density)	-3.08 ***
	(0.19)
reciprocity	2.43 ***
	(0.18)
GWESP I -> K -> J (69)	1.51 ***
	(0.16)
drugbeh alter	0.08
	(0.13)
drugbeh ego	0.11
	(0.14)
same drugbeh	0.57 *
	(0.25)
rate drugbeh (period 1)	0.52 *
	(0.23)
rate drugbeh (period 2)	1.52 **
	(0.49)
drugbeh linear shape	1.53
	(2.69)
drugbeh quadratic shape	-0.17
	(0.55)
drugbeh total similarity	5.99
	(7.91)
Iterations	2905

*** p < 0.001; ** p < 0.01; * p < 0.05

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.

Hypothesis 1: According to the parameter *outdegree*, the p-value is less than 0.001(statistically significant), and it is negative estimate. The probablity of individual reporting a friendship tie is lower than or 0.04 times the random chance over time, which support this hypothesis.

Hypothesis 2: According to the parameter *reciprocity*, the p-value is less than 0.001(statistically significant), and it is positive estimate. Students are 4.52 times more likely to have a friend of friends than random chance over time, which support this hypothesis.

Hypothesis 3: According to the parameter *GWESP*, the p-value is less than 0.001(statistically significant), and it is positive estimate. The probablity of having a reciproical friendship between students will be 11.35 times higher than the random chance over time, which support this hypothesis.

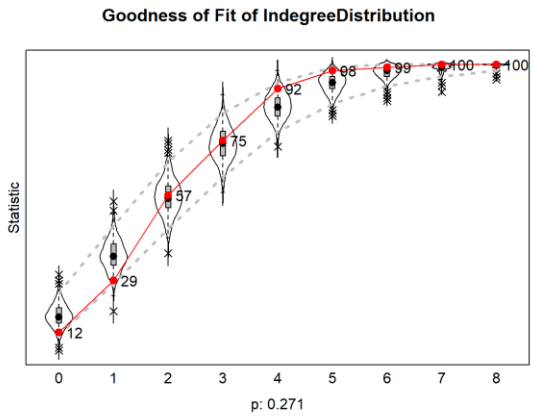
Hypothesis 4: According to the parameter *drugbeh ego*, the p-value is bigger than 0.05(not statistically significant), so, the hypothesis is rejected.

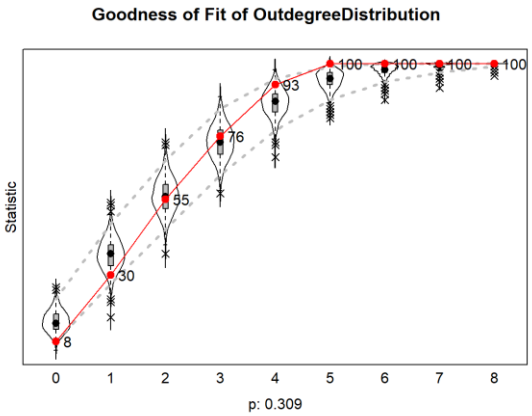
Hypothesis 5: According to the parameter *drugbeh alter*, the p-value is bigger than 0.05(not statistically significant), so, the hypothesis is rejected.

Hypothesis 6: According to the parameter *same drugbeh*, the p-value is less than 0.05(statistically significant), and it is positive estimate. If two people have the same individual attributes, they are 1.68 times more likely to be friend or have certain relationship, which support this hypothesis.

Hypothesis 7: According to the parameter *drugbeh total similarity*, the p-value is bigger than 0.05(not statistically significant), so, the hypothesis is rejected.

6. **(10 points) Report the goodness of fit for your model regarding in-degree and out-degree distributions. Include the plots and interpret the results of each plot.**





It is not difficult to observe that the red lines in both figures are contained by the gray dot lines, which suggests that our model is a good fit for both in-degree and out-degree distribution.