SNA assignment 3 by Yiqiu Wang

Yiqiu Wang

2023-02-22

1 Explain why the mechanism of gender homophily might create networks in which there is a substantial degree of reciprocity and transitivity, even if no explicit mechanisms of reciprocation and transitive closure operate.

Answers:

Gender homophily in a network means individuals in the network have the tendency to establish ties with other individuals who have the same gender as them. In the friendship network of a school, this means boys are more likely to befriend boys. The same for girls. The result could be gender segregation in terms of friendship in the school. Suppose there are 100 girls and 100 boys in the same school, gender homophily could cause the network of friendship to split into two components based on gender. Given this circumstance, a student has fewer options when making friends(99 students vs 199 students if gender homophily is extreme).

Reciprocity here refers to the situation that when a student nominate another student as his/her friend, the other student is very likely to nominate him/her. So that the two students recognize each other as "a friend of mine". Transitivity here means that if two students, A and B, are friends, A is likely to be a friend of student C, who is a friend of B. We have mentioned that gender homophily would reduce the number of potential friends. So gender homophily can induce transitivity as gender homophily increasing the likelihood of same-gender ties and in turn enhances the opportunity for completed triangles within the same-gender group, especially when groups are small(as in a school).

In the group consists of same gender individuals, group shared attributes could also be produced so that this group(girl) differentiate from the other group(boy). Based on each girl's unique attributes(music taste, sports, ect.), smaller groups with shared attributes are also likely to form within this girl group. The number of potential friends is again reduced and thus resulting in strong connection("we have so much in common") within a small group. This in turn increases the probability of A and B in the same small group nominating each other as a friend. Also, for A, B and C in the same small group, the probability of A befriending C after knowing C being a friend of B is increased. In this case, the effects of reciprocity and transitivity is less obvious than that of gender homophily.

2 Explain why the opposite is not generally true, i.e., why the mechanisms of reciprocation or transitive closure do not generally create networks which are gender segregated.

Answers:

In a school, same interests(music taste, sports, ect.) could be one of the main reasons why two students recognize each other as a friend and make friend with "a friend of my friend". In this case, if we expect

small groups in the network, these group should be different from each other based on the particular interest shared within this group. Even if one might prefer making friends with same gender students, the network is much more likely to be interest segregated than gender segregated. The mechanisms of reciprocity and transitivity closure do not emphasis a preference for same gender as gender homophily does so they should not result in gender segregated networks.

3 Identify three R-functions that can evaluate a given network in terms of (i) density, (ii) reciprocity, and (iii) transitivity. Identify a fourth function that evaluates a network and a group indicator variable in terms of (iv) homophily on that variable.

Answers:

We could first turn the original matrix(fri) into a network object(frinet) and an igraph object(graph) using network() and graph_from_adjacency_matrix() before we evaluate the network.

We would use functions from the sna and igraph package here.

We could use gden(frinet) or edge_density(graph) to get the index on the network's density. We could use grecip(frinet,measure='edgewise') or reciprocity(graph) to measure reciprocity. For transitivity, we could use transitivity(graph).

For homophily by gender, we need to first add additional attribute, sex, to the graph. Then we load the isnar package so that we could use isnar::assort(graph,'sex') to get the index.

4 Evaluate your personalised data set on these four dimensions by applying the functions from Exercise 3.

Table 1: Descriptive statistics of friendship network

Object	Index
Density	0.077
Reciprocity	0.520
Transitivity	0.282
Homophily by gender	0.711

Answers:

Table 1 shows the 4 indices on density, reciprocity, transitivity and homophily by gender of the friendship network in a school.

Density is the number of actual ties over number of potential ties. The density of the network is 0.077. So this network is not very well connected possibly because of certain attributes, such as gender, affecting the likelihood of two individuals establishing friendship.

The reciprocity index shows the likelihood of a friendship tie being mutual. In our network, this index equals to 0.520 which means the probability of A nominating B as a friend and B nominating A as well is about 52%. This means the friendship ties in the network are in general not very solid as only half of the times the friend one nominates would nominate that person back.

The transitivity index is calculated by dividing transitivity triads with potentially transitivity triads. The index here is 0.282. The probability of finding strongly connected groups is not very high.

The index for homophily by gender equals to 0.711. This suggests individuals in the network is very likely to befriend others who have the gender as them. We could expect gender segregated groups in our network.

5 Please fit three exponential random graph models to the data

```
# Model 1
# Assess overall tie creation tendencies, reciprocation tendencies
# gender homophily tendencies and transitive closure tendencies
model1 <- frinet~edges+mutual+nodematch("sex")+twopath+gwesp(0.5,fixed=TRUE)

# Model 2
# Drop reciprocation and transitive closure
# Retain gender homophily
model2 <- frinet~edges+nodematch("sex")

# Model 3
# Drop gender homophily
# Retain reciprocation and transitive closure
model3 <- frinet~edges+mutual+gwesp(0.5,fixed=TRUE)+twopath

# Estimate the model
set.seed(12345)
results1 <- ergm(model1)
results2 <- ergm(model2)
results3 <- ergm(model3)
```

6 Report the results in one table and interpret them in a brief text. What can a comparison of the coefficients from both analyses tell us?

Answers:

In Model 1, we assess overall tie creation tendencies, reciprocation tendencies, gender homophily tendencies and transitive closure tendencies. "edges" measures the number of edges in the friendship network. The coefficient equals to -2.974($p < 0.001$) which suggests that every additional edge in the friendship network lowers the log of the odds of establishing friendship with 2.974, holding all other variables as constant. So it is less likely for friendship to be established in a dense network. When an individual feels that he/she has made enough friends(or too many friends), he/she becomes less willingly to make more friends.

"mutual" measures the reciprocity of the network. That is, the likelihood that student A nominates student B as a friend and B nominates A as well. The coefficient here is 2.789($p < 0.001$) which suggests that every additional mutual tie increases the log of odds of establishing friendship with 2.789, holding all other variables as constant. If an individual witnesses many mutual friendships(solid/strong friendships) in the network, this individual would be encouraged to establish friendships with others in the network with the belief that such friendship ties should be strong.

"nodematch.sex" measures the proportion of ties between same-gender nodes. The coefficient is 1.228($p < 0.001$) which suggests every additional increase in the proportion of ties between same-gender nodes increases the log of odds of establishing friendship with 1.228, holding all other variables as constant. This also indicates gender homophily in the friendship network. People generally prefer making friends with other people of same gender.

Table 2: Exponential Random Graph Models on Friendship Network

	Model 1	Model 2	Model 3
edges	−2.974***	−3.832***	−2.229***
	(0.549)	(0.382)	(0.465)
mutual	2.789***		3.038***
	(0.489)		(0.490)
nodematch.sex	1.228***	1.969***	
	(0.346)	(0.416)	
twopath	−0.393**		−0.404**
	(0.144)		(0.153)
gwesp.fixed.0.5	0.640**		0.793***
	(0.219)		(0.232)
Num.Obs.	650	650	650
AIC	286.8	324.4	305.0
BIC	309.2	333.3	322.9

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

"twopath" measures the proportion of two ties that connect two students who do not nominate each other as a friend in our network. The coefficient is $-0.393(p < 0.01)$ which suggests that every additional increase in the proportion of such two ties increases the log of odds of establishing friendship with 0.393, holding all other variables as constant. It is less likely for friendship ties to be formed when there are many indirect ties in the friendship network. When there are many "two paths" to connect individuals in the friendship network, it means one does not have to establish friendship with another person to know that person so individuals could be disencouraged to form direct ties with others.

"gwesp.fixed.0.5" measures the proportion of ties that connects two nodes through shared node(s) in the network, with a threshold of 0.5. The coefficient is $0.640(p < 0.01)$ which suggests every addition increase in the proportion of such ties increases the log of odds of establishing friendship with 0.640, holding all other variables as constant. When there are many indirect connections between nodes, friendship is more likely to be formed. Students in the network could be encouraged to expand their friendship network through their friends.

In Model 2, we only assess overall tie creation tendencies and gender homophily tendencies. The overall tendencies are similar to those of Model 1. Dense network has a negative effect while same gender has a positive effect on the establishment of friendship. But the effects(whether negative or positive) have become greater when we dropped reciprocation tendencies and transitive closure tendencies in our model.The negative effect of dense network is the strongest in Model 2 among three models.

In Model 3, we assess overall tie creation tendencies, reciprocation tendencies and transitive closure tendencies. The coefficient of "edges" equals to $-2.229(p < 0.001)$ which suggests that every additional edge in the friendship network lowers the log of the odds of establishing friendship with 2.229, holding all other variables as constant. The negative effect of "edges" is the weakest in model 3 among three models. Apart from that, the overall tendencies are similar to those in Model 1. But the positive effect of "mutual", the negative effect of "twopath" and the positive effect of "gwesp.fixed.0.5" are greater than those in Model 1 when we dropped gender homophily tendencies in Model 3.

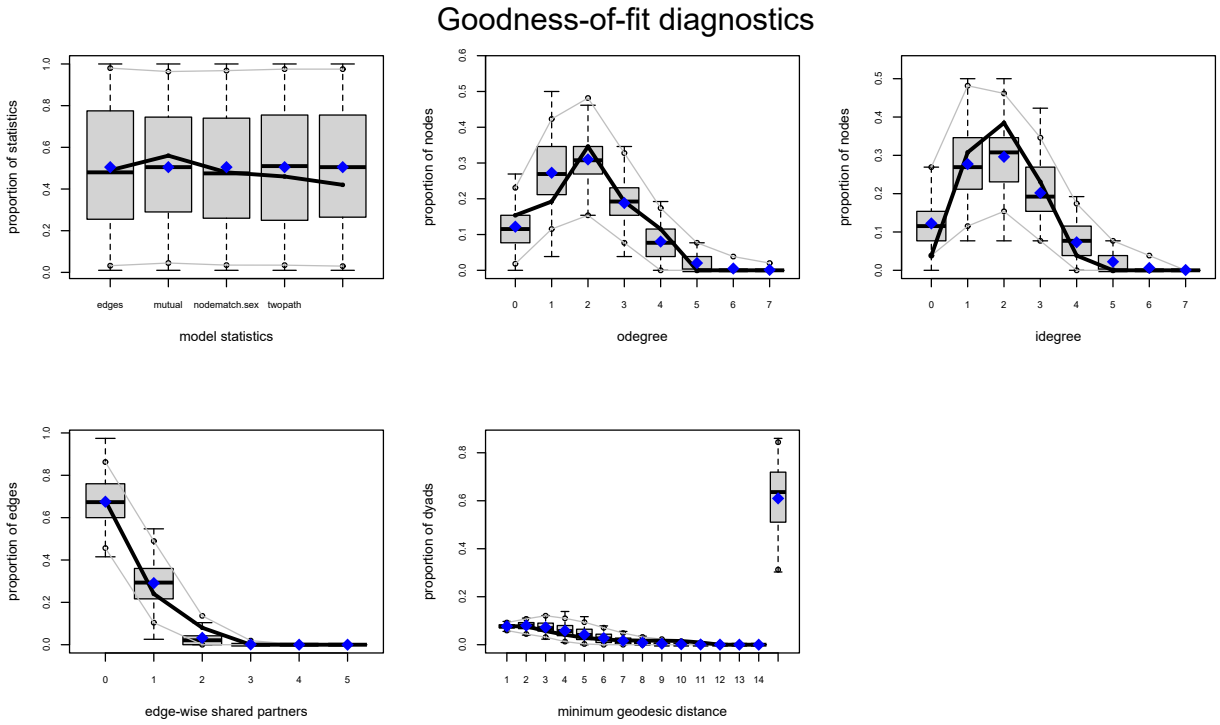
AIC is the Akaike's Information Criterion and BIC is the Bayesian Information Criterion. They are both used to compare model fit and the smaller the index the better. Model 1 has the smallest AIC and BIC among three models which suggests Model 1 has the best model fit. Model 2 has the greatest AIC and BIC among three models which suggests Model 2 has the worst model fit.

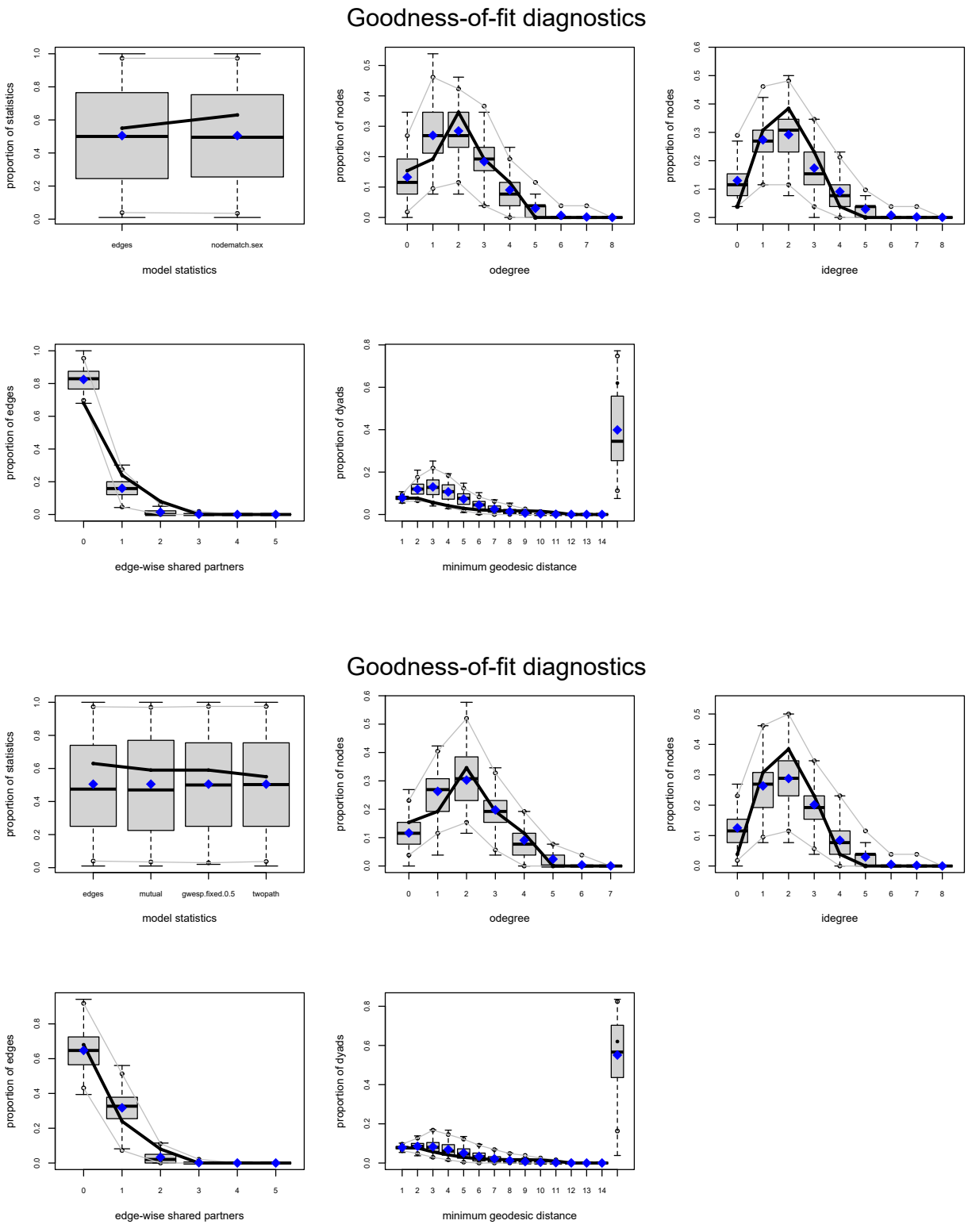
When we compare Model 2 to Model 1, we found that as we add reciprocity and transitivity to our model(from Model 2 to Model 1), the coefficient of "nodematch.sex" decreases. When we compare Model 1

to Model 3, we found that as we drop gender homophily tendencies(Model 1 to Model 3), the coefficients of "mutual" and "gwesp.fixed.0.5" increases.

The mechanisms of reciprocity and transitive closure could be masking the effect of gender homophily on the tendency of forming friendship. Then the coefficient of gender homophily decreases as we move from Model 2 to Model 1. Gender homophily in Model 1 could be suppressing the effect of reciprocity and transitivity on friendship forming tendency, so when we dropped gender homophily tendencies and moved from Model 1 to Model 3, the coefficients of reciprocity and transitivity increase.

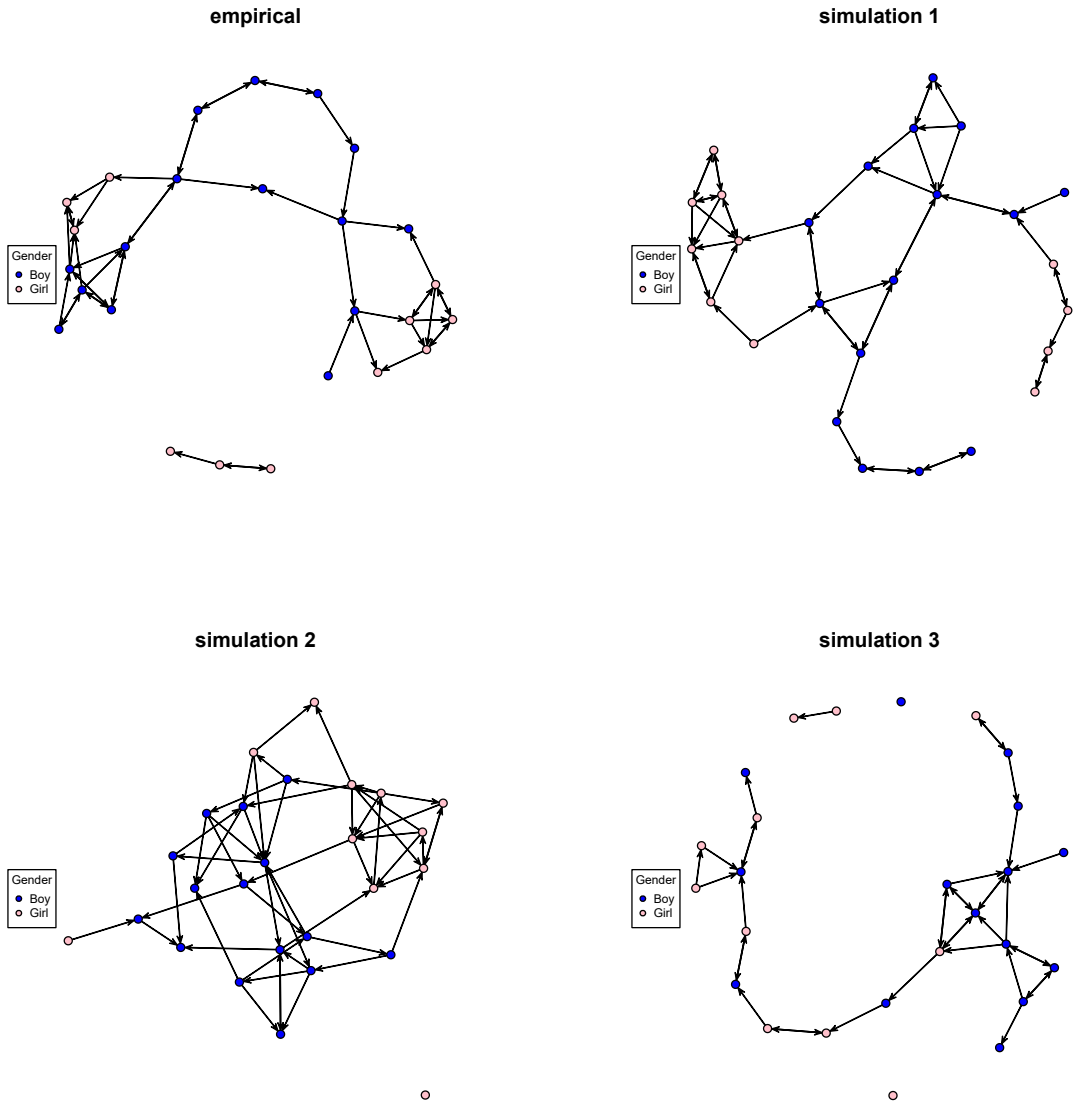
So as long as there is explicit mechanism of gender homophily in the network, it is hard to find explicit mechanisms of reciprocity and transitivity as their effect could be masked by gender homophily, but we could still find a substantial degree of reciprocity and transitivity. Without explicit mechanism of gender homophily, the effects of reciprocity and transitivity is no longer suppressed so we see explicit mechanisms of reciprocity and transitivity which do not resulting in gender segregated networks.



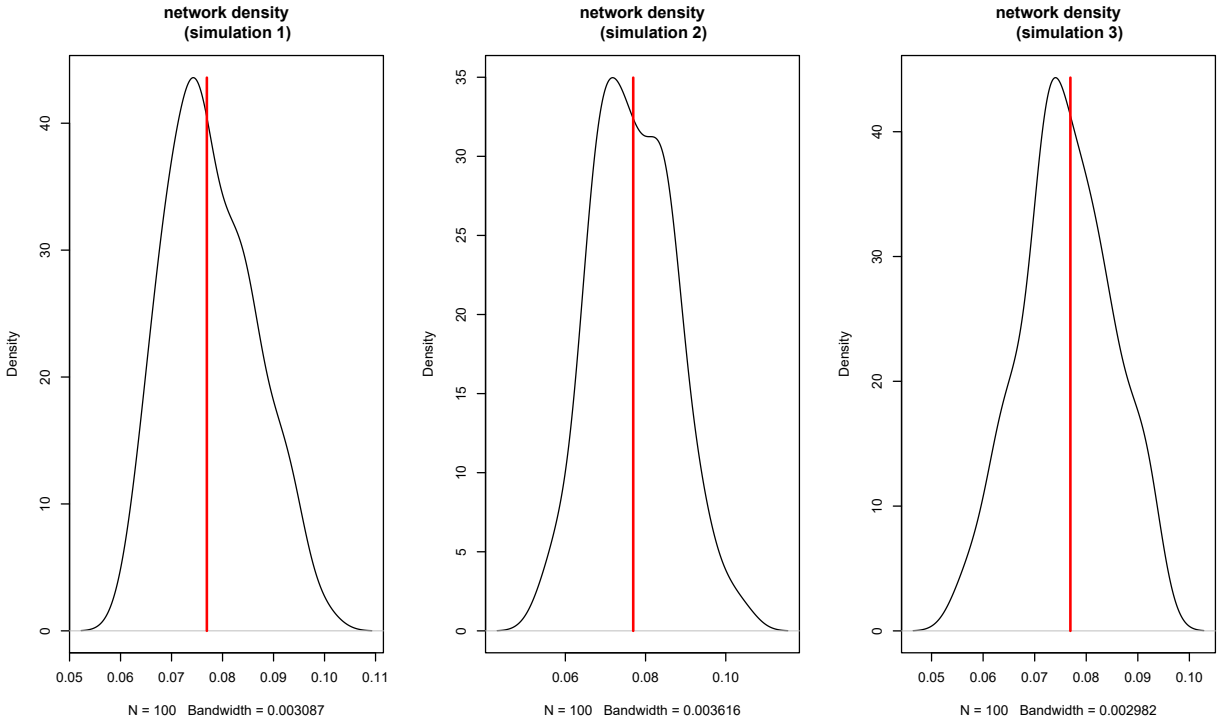


The three graphs above show the results goodness-of-fit diagnostics of Model 1, Model 2 and Model 3. It appears that in the graphs of Model 1, the black lines of empirical data are located well within the box plots of simulated data on these dimensions while it is not the case for Model 2 and Model 3. Model 1 has the best model fit.

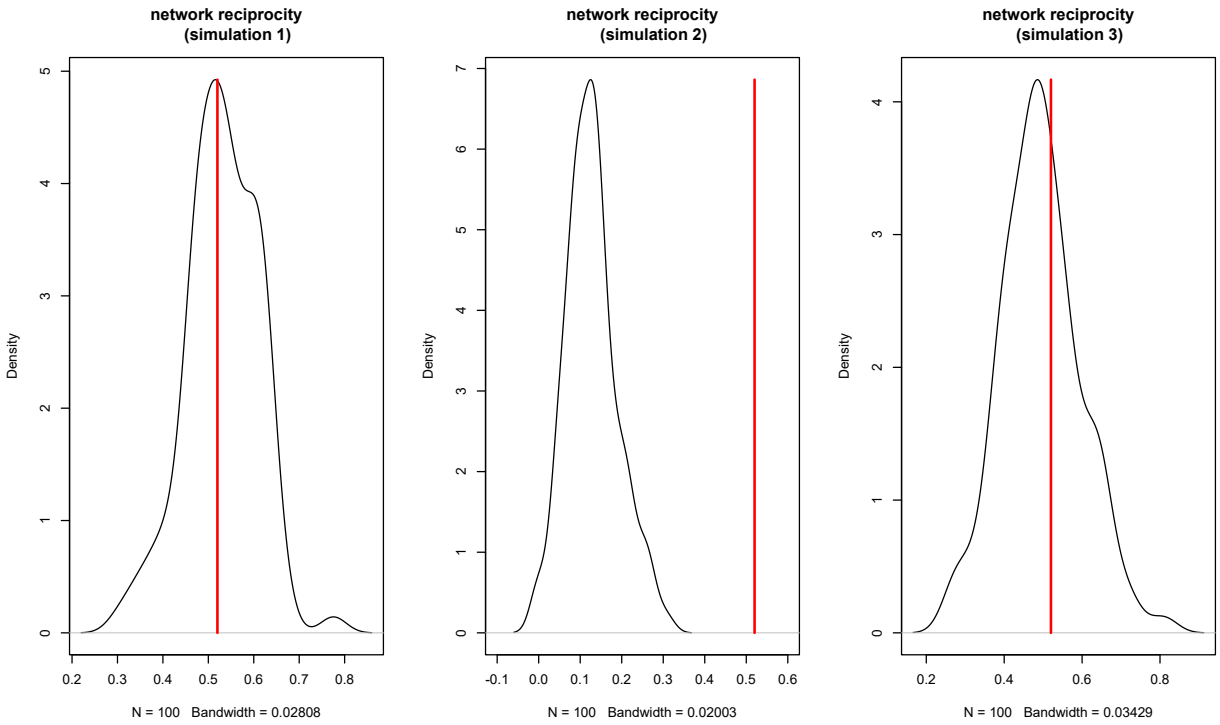
7.1 Make four network visualisations: one example graph from each set of simulations, and the empirical data set, with nodes coloured by gender.

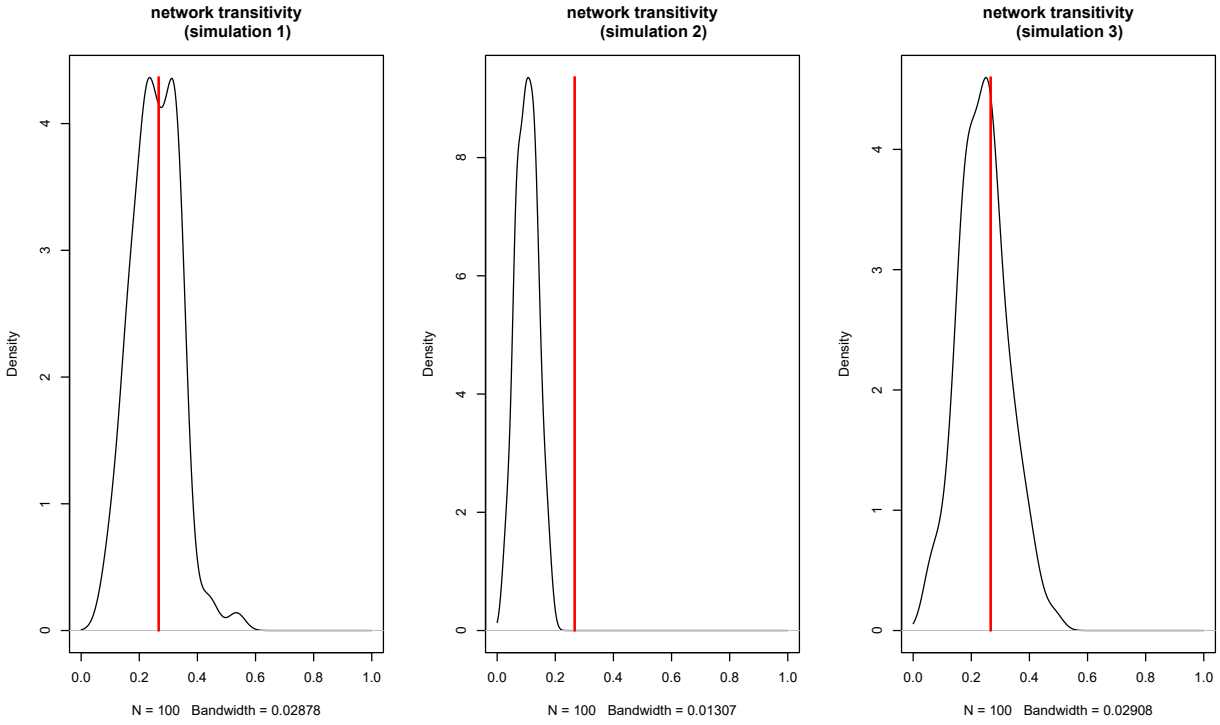


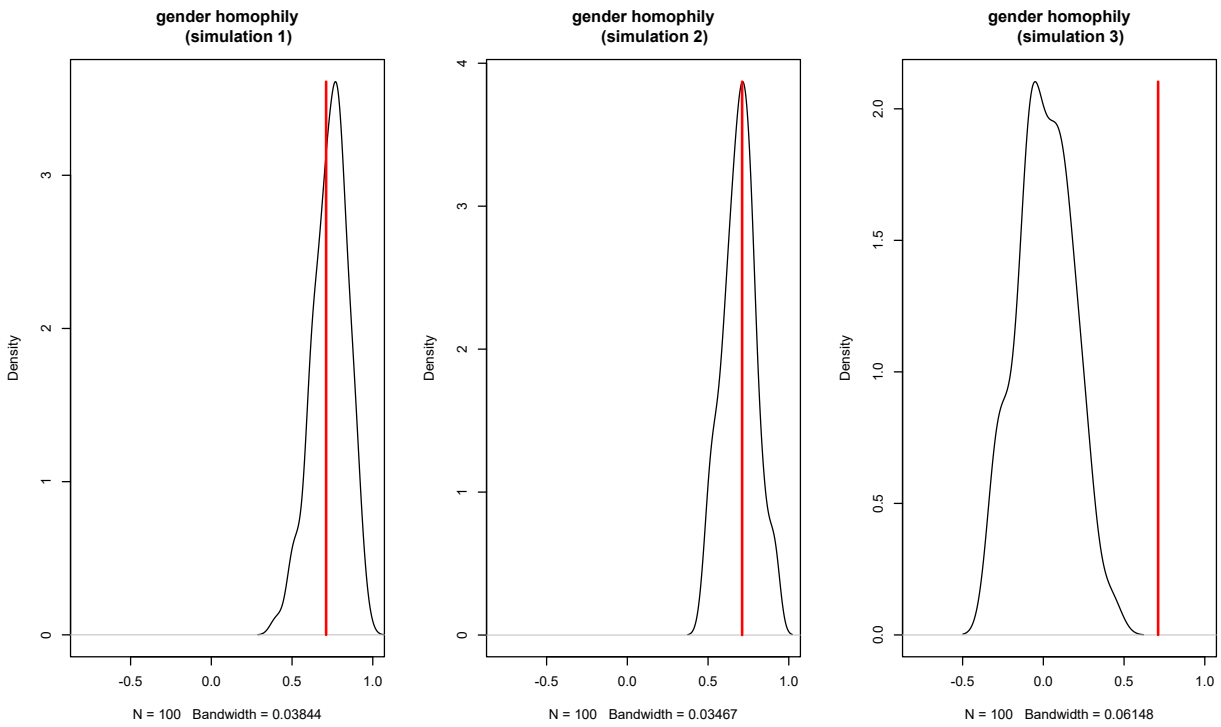
7.2 Make four plots, one for each index, each of which contains the three simulated distributions and the empirical value of the index as reference (i.e., the value you calculated in Exercise 4).



Since we included edges in all three models, the distributions of density indices of our simulated networks


network reciprocity
(simulation 1)
Density
N = 100 Bandwidth = 0.02808
network reciprocity
(simulation 2)
Density
N = 100 Bandwidth = 0.02003
network reciprocity
(simulation 3)
Density
N = 100 Bandwidth = 0.03429


network transitivity
(simulation 1)
Density
N = 100 Bandwidth = 0.02878
network transitivity
(simulation 2)
Density
N = 100 Bandwidth = 0.01307
network transitivity
(simulation 3)
Density
N = 100 Bandwidth = 0.02908



8 Try to substantiate your reasoning in Exercises 1) and 2) above based on these results.

Answers:

We found it when answering question 6 that the mechanism of gender homophily could mask as well suppress the effect of reciprocity and transitivity on the tendency of friendship establishment in the friendship network.

By looking at the 4 plots in question 7.1, we find that the plot based on the full model(Model 1) is very much similar to the plot of the empirical network. The plot based on Model 2 shows significant gender homophily tendency. Although we did not include reciprocity and transitivity in Model 2, we still find a degree of reciprocity and transitivity in the plot. The plot based on Model 3 shows great degree of reciprocity and transitivity, but no obvious tendency of gender homophily.

When we look at the distribution plots of simulation 2(based on Model 2) about reciprocity indices and transitivity indices in question 7.2, even when we did not include reciprocity and transitivity in Model 2, we can still find a certain degree of reciprocity and transitivity though the indices are lower than those of the empirical network. When we look at the distribution plot about gender homophily of simulation 3(based on Model 3), if we include reciprocity and transtivity but no gender homophily in the model, the gender homophily indices of our simulated networks in general equal to or are extremely close to 0. This means there is no obvious preference for same gender or different gender and gender is hardly of any importance when individuals are making friends in our network.

So when we find gender homophily in the network, we also find reciprocity and transitivity, even if no explicit mechanisms of reciprocation and transitive closure operate. If we find the mechanisms of reciprocity and transitivity in the network, the effects of reciprocity and transitivity(not masked) are stronger than that of gender homophily, meaning we cannot find explicit mechanism of gender homophily, so the network would not be gender segregated.