

Extended notes for the ERGM tutorial

Sasha Poquet (University of South Australia) & Nia Dowell (University of Michigan)

14/06/2017

INTRODUCTION

Exponential random graph modelling (ERGM) is used to analyse the properties describing network structures, i.e. does this local pattern occur in a given network beyond chance. ERGMs assume that individual characteristics influence the actors as to who they send ties to and receive ties from. ERGMs assume that the network structure emerges from network properties (endogeneous effects) and actor attributes (exogeneous effects). These micro-level interactions from social selection aggregate into group-level patterns that describe the entire network.

ERGMs also allow for the randomness in interactions, i.e. social networks are both structured and stochastic (Lusher et al., 2013, p.10). p^* /ERGM assumes that multiple local processes can take place simultaneously (Monge and Contractor, 2002, Lusher, et al., 2013). Monge and Contractor (2001) grouped various theories explaining local mechanisms for network formation into nine theoretical mechanisms: (1) theories of self-interest, (2) theories of mutual interest and collective action, (3) cognitive theories, (4) cognitive consistency theories, (5) contagion theories, (6) exchange and dependency theories, (7) homophily theories, (8) proximity theories, and (9) theories of network evolution and coevolution (Contractor, Wasserman, Faust, 2006, p. 682). These theories can help to form hypothesis if social relations are analysed. However, applications of network research in LA reaches beyond social relations; and hypothesing network formation in LA studies may be guided by prior empirical research of the phenomenon represented by the network.

ERGM is a probability model for network analysis that predicts the presence/absence of network ties. Lusher et al. (2013) explained that ERGM estimates the probability of a given network G given a sum of network statistics weighted, just as in regression, by parameters inside an exponential (p.10). The network statistics represent social processes by different forms expressed in sets of two actors (dyads), three actors (triads) or larger combinations, such as several triads sharing a tie. ERGM estimates the likelihood a selected configuration representing a theoretically formulated structural network characteristic will occur beyond chance. This is implemented by examining the likelihood of a studied element to occur in a randomly generated distribution of networks constrained by the selected configurations. The probability estimation is derived from comparing the ties/configurations in the simulated networks to the ties in the observed network.

The modelling of configurations depends on the sub-class of statistical models used by the ERGM for a particular configuration. Simply put, dyads are modelled by p^* models that do not need to predict a dependent third tie (Holland and Leinhardt, 1981). Triads could be modelled with the help of Markov models (Frank and Strauss, 1986, Wasserman and Pattison, 1996). However, their use can fail to produce converged networks due to the tendency of a network to cluster unevenly. Hence, network closure is often modelled with the help of the social circuit models that allow a specification of how the degree of clustering within the network (Snijders et al., 2006, Robins et al., 2007). Both dyadic, triadic and clustering configurations can be fitted within the same model. The geometric forms representative of the reciprocity, transitivity, network closure, and others, become building blocks from which the networks emerge.

Multi-level analysis within ERGM controls for the tendency of studied parameters against one another, due to their theoretical dependency. Estimation of whether new models produce a better fit are done in reference to the null model. A baseline model in ERGM examines only the network density to identify that the network does not occur at random. Further elements representing network structure and nodal and edge attributes are added consequently checking for the best model fit representing by AIC coefficients. The modelled output included a parameter estimate where zero indicates that the modelled effect was no different from random. These log-odds of a parameter can also be transformed into the odds of ties occurring.

Technical literature emphasizes several steps to check the goodness of fit. Statnet package embeds Monte Carlo

Markov Chain (MCMC) algorithm that stochastically approximates the Maximum Likelihood Estimation. These algorithms are used to check for network degeneracy and to examine the goodness of fit. Estimated models demonstrate reasonable goodness of fit if i) modelled number of networked configurations are similar to the configurations in the observed network, ii) MLE estimation produced non-degenerate networks, and iii) AIC parameter demonstrated better fit than the null model.

Describing the Data

Analysed network was built from the tweets of the participants in an open online course. The course was organized over a 12 week period from January 17th, 2011 to April 11th, 2011. The course was of interest to practitioners and researchers working in online education, and those facilitating online community development. Participation in the course was open, however those learners who wanted to receive a certificate, had to apply for university admission. For the analyses, we collected learner data and distributed course interactions to reconstruct the evolution of the course.

An edge between two nodes was created when @A mentioned @B in their tweet. The graph was directed; edge weights were calculated based on the count of links between two nodes.

Information about the participants was collected manually from publicly available sources such as Twitter profiles, social networking sites (e.g., LinkedIn, About.me, and Blogger profiles), and through manual Web searches. The following demographic data were found relevant for an overview of course participants, and are presented in Figure 1: i) domain of work (e.g., secondary education, higher education, and health) in 2011, ii) type of work (e.g., research or practice) in 2011, iii) demographic data (e.g., location, gender, and professional background) in 2011.

This dataset has been analysed in several ways as reported in these papers:

1. Skrypnik, O., Joksimović, S., Kovanović, V., Gašević, D., & Dawson, S. (2014). Roles of course facilitators, learners, and technology in the flow of information of a cMOOC. *The International Review of Research in Open and Distributed Learning*, 16(3).<http://dx.doi.org/10.19173/irrodl.v16i3.2170>
2. Joksimović, S., Kovanović, V., Jovanović, J., Zouaq, A., Gašević, D., & Hatala, M. (2015). What Do cMOOC Participants Talk About in Social Media?: A Topic Analysis of Discourse in a cMOOC. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 156–165). Poughkeepsie, New York: ACM. <https://doi.org/10.1145/2723576.2723609>

TUTORIAL

Prior to running this tutorial, make sure you installed the following packages: `igraph`, `ergm`, `statnet`, and `intergraph` Do not load the packages into the workspace yet, i.e. `igraph` and `statnet` need to be loaded in different order, and do not work well when loaded at the same time.

Set the working directory by running `'setwd("your path")'`.

Make sure that the network file `g2.gml` is in the same folder that you chose as your working directory.

Remember to load the library `library(igraph)` And set a seed for reproducibility `set.seed(234)`

EXPLORING THE NETWORK

```
g <- read_graph("g2.gml", format=c("gml")) #read graph from the working directory
summary(g)
```

```
## IGRAPH DNW- 767 1193 --
```

```
## + attr: id (v/n), name (v/c), Role (v/c), SocioTech (v/c), Gender
```

```
## | (v/c), Domain (v/c), WorkType (v/c), Continent (v/c), weight
## | (e/n)
```

This is a directed weighted network that has 767 nodes and 1193 edges. The network is rather sparse with the density of 0.0020306. The edges in the network have weights, as shown in the output. The nodes have attributes as well (marked by v/c) in the summary.

We can take a quick look at the descriptive summaries of the attributes, to get a better grasp of the dataset.

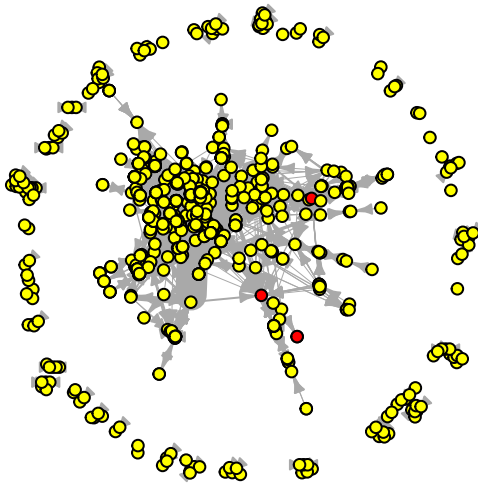
```
table(V(g)$Role)
```

```
##
## Course Instructor      Student
##                3                764
```

We know that there are three instructor nodes in the network: two representing personal Twitter accounts, and one representing course Twitter handle run by the instructors interchangeably. We can also take a quick look as to whether the instructors have a central position in this network - a dynamics that is often expected in an online course, especially its early stages. We will plot this network and colour the instructor nodes with red, and the remainder nodes in yellow. Sometimes visualising the network may help gain some insight, but sometimes it does not.

```
V(g)$RoleColour=V(g)$Role
V(g)$RoleColour=gsub("Student","yellow",V(g)$RoleColour)
V(g)$RoleColour=gsub("Course Instructor","red",V(g)$RoleColour)
#play with the layout that is better visually, remove the labels, adjust sizes of vertices and edge wei.

plot.igraph(g, layout=layout_with_mds, vertex.label=NA, vertex.size=5, vertex.color=V(g)$RoleColour, ed
```



It is evident that the instructors are embedded within the network, but not as prominent as some of the students. We observed a dense cluster in the left-side of the network indicating more intense activity levels. We could continue plotting the variables, or simply take a look at the descriptive summaries of the attributes.

```
table(V(g)$Domain)
```

```
##
##                Business                Community
##                50                    3
## Elementary and primary education      Entrepreneurship
##                10                    62
##                Government                Health
##                4                        9
```

```
##           Higher Education           Languages
##                190                37
##           Library           Organization
##                6                68
##           Other           Secondary education
##                9                78
##           Undergraduate           Unknown
##                4                216
##           Various
##                21
```

We can see that the domain of most nodes was unknown, followed by 190 nodes who were affiliated with the HigherEd sector, and 78 with K12, and 62 entrepreneurs.

```
table(V(g)$Continent)
```

```
##
##           Africa           Asia Australia and NZ           Europe
##                2           19           48           191
##   International   North America   South America           Unknown
##                2           232           56           217
```

Again, the majority of the nodes come from North America (232), some are from Europe (191), with 56 from South America and 48 from Australia and New Zealand. We could also see that most of the nodes come from English-speaking countries. If that is of interest to our research question, we can create an additional variable that describes this qualitative characteristic of the nodes.

```
table(V(g)$Gender)
```

```
##
##           F           M           Org Unknown
##           237           285           100           145
```

In relation to gender, we can see that there is a general balance between the number of men and women in the dataset. We also can see that there were 100 Twitter accounts representing organisations in this dataset.

```
table(V(g)$WorkType)
```

```
##
##   Mixed   Other Practice Research   Unknown
##        66         2         400         107         192
```

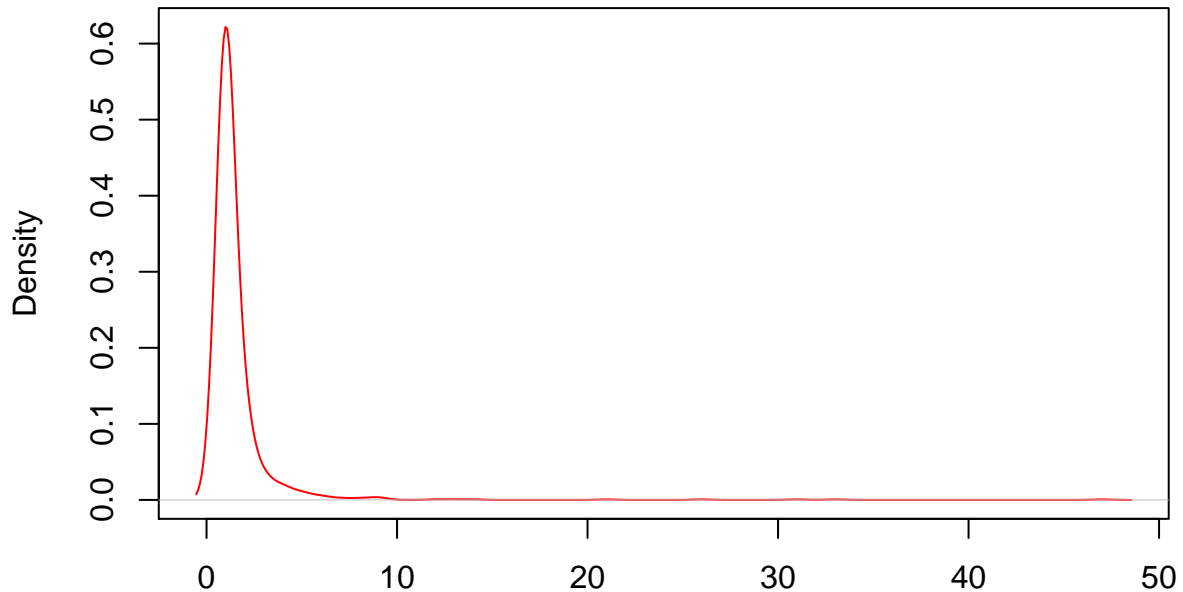
When we look at the worktype attribute, we observe that most of the nodes were related to practice, and 107 were researchers or research students.

Besides exploring the node attributes, we should also look at the network structure, to be able to form some hypothesis about what may describe some of the processes driving the formation of this network.

We should also take a look at the edge weight distribution that indicates the number of times two people have had another Student's Twitter handle embedded within their Tweets.

```
plot(density(edge_attr(g)$weight), col='red') #explore the edge attributes
```

density.default(x = edge_attr(g)\$weight)



N = 1193 Bandwidth = 0.5144

```
summary(E(g)$weight)
```

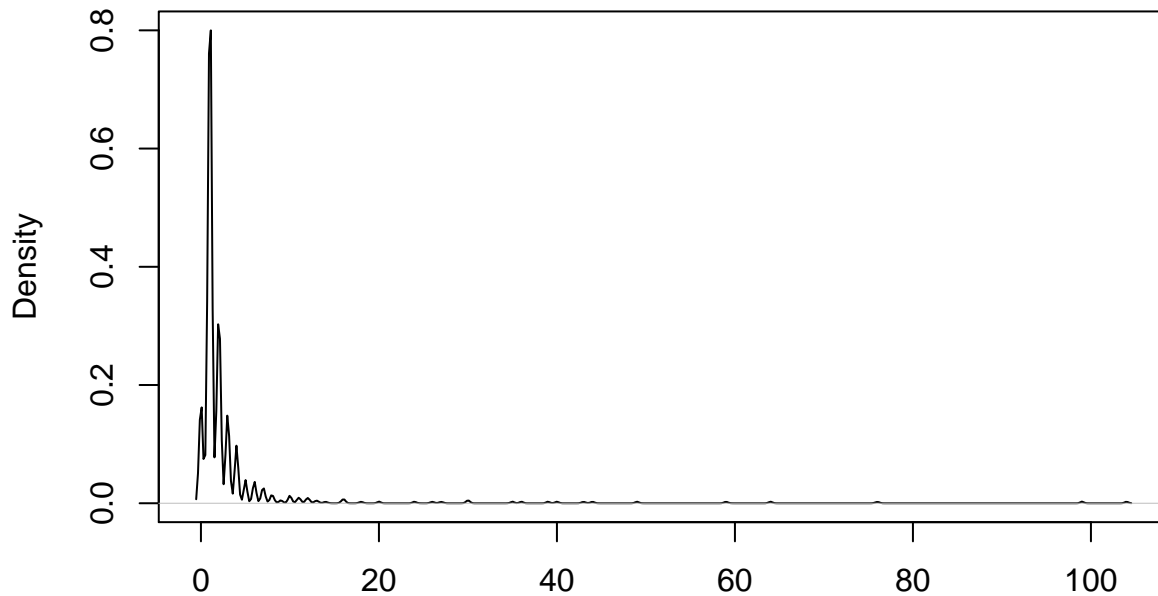
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.000  1.000  1.000  1.548  1.000  47.000
```

We observe that the edge weight has a skewed distribution, typical to online networks; with most edge weight being under 5, and a handful of edges over ten, reaching towards the maximum of 47.

Finally, at minimum, at the exploratory stage, we should take a look at the network's degree distribution.

```
deg <- degree(g, mode="all")
#deg.dist <- degree_distribution(g, cumulative=T, mode="all")
#plot( x=0:max(deg), y=1-deg.dist, pch=19, cex=1.2, col="orange",
#xlab="Degree", ylab="Cumulative Frequency")
plot(density(deg), main="Density, Degree Distribution, N=767")
```

Density, Degree Distribution, N=767



N = 767 Bandwidth = 0.1779

At this stage we have looked at the different qualitative aspects of the network, and may want to ask a few questions about the network structure.

Exploring Network for ERGM

First, some housekeeping. ERGM works within statnet package. In practice, igraph and statnet packages may have problems with one another. Hence, best to convert a network from an igraph object to a network object using intergraph package. Then, load statnet and ergm packages, as well as detach igraph package.

```
library(intergraph)
```

```
library(statnet)
```

```
library(network)
```

```
net <- asNetwork(g) #convert network from igraph object to sna object  
net #check network summary to make sure there was no bugs in conversion
```

```
## Network attributes:
```

```
## vertices = 767
```

```
## directed = TRUE
```

```
## hyper = FALSE
```

```
## loops = FALSE
```

```
## multiple = FALSE
```

```
## bipartite = FALSE
```

```
## total edges= 1193
```

```
## missing edges= 0
```

```
## non-missing edges= 1193
```

```
##
```

```
## Vertex attribute names:
```

```
## Continent Domain Gender id Role RoleColour SocioTech vertex.names WorkType
```

```
##
## Edge attribute names not shown
mixingmatrix(net, "Role") #look at the mixing matrices of the network
```

```
##           To
## From      Course Instructor Student Total
## Course Instructor      1      9    10
## Student                98    1085  1183
## Total                  99    1094  1193
```

So we can see that students mostly talked to one another, and instructors rarely did. There has been some activity between the students and instructors, but we can't know its intensity because the weight of the ties is not captured here.

```
mixingmatrix(net, "Gender")
```

```
##           To
## From      F    M Org Unknown Total
## F          143 244 67     30   484
## M          149 240 59     31   479
## Org         47  69 20     20   156
## Unknown     9  17  3     45    74
## Total      348 570 149    126  1193
```

We can see that females sent more ties to the males, and males sent more ties to the males.

```
mixingmatrix(net, "Domain")
```

```
##           To
## From      Business Community
## Business      4      0
## Community     5      0
## Elementary and primary education  1      0
## Entrepreneurship  9      1
## Government     0      0
## Health         2      0
## Higher Education 11      1
## Languages      4      1
## Library        0      0
## Organization   7      1
## Other          0      0
## Secondary education  3      0
## Undergraduate   0      0
## Unknown        5      0
## Various        1      0
## Total         52      4

##           To
## From      Elementary and primary education
## Business      0
## Community     0
## Elementary and primary education  0
## Entrepreneurship  1
## Government     0
## Health         0
## Higher Education  2
## Languages      0
```

##	Library	0
##	Organization	1
##	Other	0
##	Secondary education	1
##	Undergraduate	0
##	Unknown	0
##	Various	1
##	Total	6

##		To			
##	From	Entrepreneurship	Government	Health	
##	Business	6	0	2	
##	Community	6	0	2	
##	Elementary and primary education	1	0	0	
##	Entrepreneurship	28	0	3	
##	Government	1	0	0	
##	Health	3	0	1	
##	Higher Education	36	1	7	
##	Languages	16	0	3	
##	Library	0	0	0	
##	Organization	13	1	2	
##	Other	0	1	0	
##	Secondary education	22	1	0	
##	Undergraduate	0	0	0	
##	Unknown	15	1	4	
##	Various	3	0	2	
##	Total	150	5	26	

##		To			
##	From	Higher Education	Languages	Library	
##	Business	18	8	2	
##	Community	19	3	0	
##	Elementary and primary education	2	0	0	
##	Entrepreneurship	55	8	0	
##	Government	0	0	0	
##	Health	8	1	0	
##	Higher Education	164	20	1	
##	Languages	28	26	1	
##	Library	5	0	0	
##	Organization	39	4	1	
##	Other	1	0	0	
##	Secondary education	19	16	0	
##	Undergraduate	0	1	0	
##	Unknown	30	11	1	
##	Various	17	2	0	
##	Total	405	100	6	

##		To			
##	From	Organization	Other	Secondary education	
##	Business	10	0	4	
##	Community	5	2	3	
##	Elementary and primary education	3	0	5	
##	Entrepreneurship	16	1	16	
##	Government	0	0	0	
##	Health	2	0	0	
##	Higher Education	28	3	8	
##	Languages	2	0	4	


```

## Library 0 0 0
## Organization 17 1 11
## Other 0 0 0
## Secondary education 18 2 15
## Undergraduate 0 0 0
## Unknown 16 0 2
## Various 1 0 2
## Total 118 9 70
##
## To
## From Undergraduate Unknown Various Total
## Business 0 10 3 67
## Community 0 10 2 57
## Elementary and primary education 0 0 0 12
## Entrepreneurship 0 21 5 164
## Government 0 0 0 1
## Health 0 2 1 20
## Higher Education 0 39 17 338
## Languages 0 12 3 100
## Library 0 0 0 5
## Organization 0 30 3 131
## Other 0 1 0 3
## Secondary education 0 11 1 109
## Undergraduate 0 0 0 1
## Unknown 0 58 8 151
## Various 0 4 1 34
## Total 0 198 44 1193

```

We can see that there are many zeros in these matrix, i.e. there were no ties between the categories. In practice that means that these relations will be difficult to model, and they need to be less granular. We will show what the output would look like, and suggest a solution to that during modelling.

```

mixingmatrix(net, "WorkType")

```

```

## To
## From Mixed Other Practice Research Unknown Total
## Mixed 51 0 60 21 11 143
## Other 1 0 1 0 2 4
## Practice 137 0 471 80 98 786
## Research 43 0 45 30 15 133
## Unknown 17 0 50 4 56 127
## Total 249 0 627 135 182 1193

```

```

mixingmatrix(net, "Continent")

```

```

## To
## From Africa Asia Australia and NZ Europe International
## Africa 0 1 0 1 0
## Asia 0 3 2 3 0
## Australia and NZ 1 3 17 52 1
## Europe 2 13 20 99 3
## International 0 0 0 0 0
## North America 1 6 13 51 3
## South America 0 2 0 20 0
## Unknown 0 3 14 50 3
## Total 4 31 66 276 10
## To

```

```
## From          North America South America Unknown Total
## Africa                0           0           0      2
## Asia                   4           1          11     24
## Australia and NZ      46          11          26    157
## Europe                 110          17          79    343
## International          0           0           1      1
## North America         141          15          81    311
## South America          29          34          18    103
## Unknown                73          19          90    252
## Total                  403          97         306   1193
```

Having looked at the basic counts of network attributes, we may want to look at the counts of structural features prior to modelling the network. Basic structural properties include controlling for density (edges terms in statnet), degree (a range of terms), closure (a range of terms associated with different types of triads within the triadic census).

For instance, we can look at the number of nodes with in-degree and out-degree from 0 to ten.

```
summary(net ~ idegree(0:10)) # what does the in-degree summary
```

```
## idegree0 idegree1 idegree2 idegree3 idegree4 idegree5 idegree6
##      348      259       64       27        18         8         7
## idegree7 idegree8 idegree9 idegree10
##         8         3         2         6
```

```
#for the range of 1 to ten look like in this network
```

```
summary(net ~ odegree(0:10)) # what does the out-degree summary
```

```
## odegree0 odegree1 odegree2 odegree3 odegree4 odegree5 odegree6
##      349      218       99       45        19        12         4
## odegree7 odegree8 odegree9 odegree10
##         3         2         1         0
```

```
#for the range of 1 to ten look like in this network
```

It appears that the network has highest number of nodes with degrees of 1 and 2. This may become a part of the hypothesis about the structure of the network at a later stage.

Alternatively, we can look at the other structural elements that theoretically appear meaningful in describing the network. The full description of all possible terms can be found within the ergm-terms documentation (<http://svitsrv25.epfl.ch/R-doc/library/ergm/html/ergm-terms.html>)

```
summary(net ~ edges + mutual
+ triangles + simmelianties
+ intransitive + transitive
+ cyclicalities + twopath)
```

```
##          edges          mutual      triangle simmelianties  intransitive
##          1193             55         1020             18           7415
##   transitive cyclicalities      twopath
##           964             108         8379
```

```
# you can also check the counts for the selected structural features. Minimum structural features could
```

Finally, we can also look at the counts of some of the attributes, commonly used are nodefactor, nodemix, nodematch, nodecov, dyadcov.

```
summary(net ~ nodefactor('Continent') + nodematch ('Continent', diff=T) + nodemix('Continent'))
```

```

##           nodefactor.Continent.Asia
##                               55
##   nodefactor.Continent.Australia and NZ
##                               223
##           nodefactor.Continent.Europe
##                               619
##   nodefactor.Continent.International
##                               11
##   nodefactor.Continent.North America
##                               714
##   nodefactor.Continent.South America
##                               200
##   nodefactor.Continent.Unknown
##                               558
##   nodematch.Continent.Africa
##                               0
##   nodematch.Continent.Asia
##                               3
##   nodematch.Continent.Australia and NZ
##                               17
##   nodematch.Continent.Europe
##                               99
##   nodematch.Continent.International
##                               0
##   nodematch.Continent.North America
##                               141
##   nodematch.Continent.South America
##                               34
##   nodematch.Continent.Unknown
##                               90
##   mix.Continent.Africa.Africa
##                               0
##   mix.Continent.Asia.Africa
##                               0
##   mix.Continent.Australia and NZ.Africa
##                               1
##   mix.Continent.Europe.Africa
##                               2
##   mix.Continent.International.Africa
##                               0
##   mix.Continent.North America.Africa
##                               1
##   mix.Continent.South America.Africa
##                               0
##   mix.Continent.Unknown.Africa
##                               0
##   mix.Continent.Africa.Asia
##                               1
##   mix.Continent.Asia.Asia
##                               3
##   mix.Continent.Australia and NZ.Asia
##                               3
##   mix.Continent.Europe.Asia
##                               13

```

```

##          mix.Continent.International.Asia
##                                     0
##          mix.Continent.North America.Asia
##                                     6
##          mix.Continent.South America.Asia
##                                     2
##          mix.Continent.Unknown.Asia
##                                     3
##          mix.Continent.Africa.Australia and NZ
##                                     0
##          mix.Continent.Asia.Australia and NZ
##                                     2
## mix.Continent.Australia and NZ.Australia and NZ
##                                     17
##          mix.Continent.Europe.Australia and NZ
##                                     20
##          mix.Continent.International.Australia and NZ
##                                     0
##          mix.Continent.North America.Australia and NZ
##                                     13
##          mix.Continent.South America.Australia and NZ
##                                     0
##          mix.Continent.Unknown.Australia and NZ
##                                     14
##          mix.Continent.Africa.Europe
##                                     1
##          mix.Continent.Asia.Europe
##                                     3
##          mix.Continent.Australia and NZ.Europe
##                                     52
##          mix.Continent.Europe.Europe
##                                     99
##          mix.Continent.International.Europe
##                                     0
##          mix.Continent.North America.Europe
##                                     51
##          mix.Continent.South America.Europe
##                                     20
##          mix.Continent.Unknown.Europe
##                                     50
##          mix.Continent.Africa.International
##                                     0
##          mix.Continent.Asia.International
##                                     0
##          mix.Continent.Australia and NZ.International
##                                     1
##          mix.Continent.Europe.International
##                                     3
##          mix.Continent.International.International
##                                     0
##          mix.Continent.North America.International
##                                     3
##          mix.Continent.South America.International
##                                     0
##

```

```

##          mix.Continent.Unknown.International
##                                     3
##          mix.Continent.Africa.North America
##                                     0
##          mix.Continent.Asia.North America
##                                     4
## mix.Continent.Australia and NZ.North America
##                                     46
##          mix.Continent.Europe.North America
##                                     110
##          mix.Continent.International.North America
##                                     0
##          mix.Continent.North America.North America
##                                     141
##          mix.Continent.South America.North America
##                                     29
##          mix.Continent.Unknown.North America
##                                     73
##          mix.Continent.Africa.South America
##                                     0
##          mix.Continent.Asia.South America
##                                     1
## mix.Continent.Australia and NZ.South America
##                                     11
##          mix.Continent.Europe.South America
##                                     17
##          mix.Continent.International.South America
##                                     0
##          mix.Continent.North America.South America
##                                     15
##          mix.Continent.South America.South America
##                                     34
##          mix.Continent.Unknown.South America
##                                     19
##          mix.Continent.Africa.Unknown
##                                     0
##          mix.Continent.Asia.Unknown
##                                     11
##          mix.Continent.Australia and NZ.Unknown
##                                     26
##          mix.Continent.Europe.Unknown
##                                     79
##          mix.Continent.International.Unknown
##                                     1
##          mix.Continent.North America.Unknown
##                                     81
##          mix.Continent.South America.Unknown
##                                     18
##          mix.Continent.Unknown.Unknown
##                                     90

```

Modelling the Networks

First, we run the null model to see if the density is negative, i.e. network does not exist on its own beyond chance.

```
null <- ergm(net ~edges)
```

```
## Evaluating log-likelihood at the estimate.
```

```
summary(null)
```

```
##
## =====
## Summary of model fit
## =====
##
## Formula:    net ~ edges
##
## Iterations: 8 out of 20
##
## Monte Carlo MLE Results:
##      Estimate Std. Error MCMC % p-value
## edges -6.19741    0.02898      0 <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 814478  on 587522  degrees of freedom
##      Residual Deviance: 17175  on 587521  degrees of freedom
##
## AIC: 17177    BIC: 17189    (Smaller is better.)
```

We note the values for AIC and BIC. As we add more elements to the model, these values should become smaller, if the model is describing the data better.

The next model adds some basic structural elements, i.e. degree and reciprocity.

```
m1.1 <- ergm(net ~edges + mutual + idegree(1) + odegree(1))
```

```
## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 2.646
## Step length converged once. Increasing MCMC sample size.
## Iteration 2 of at most 20:
## The log-likelihood improved by 0.1279
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.
```

```
summary(m1.1)
```

```
##
## =====
## Summary of model fit
## =====
##
## Formula:    net ~ edges + mutual + idegree(1) + odegree(1)
##
```

```

## Iterations: 2 out of 20
##
## Monte Carlo MLE Results:
##      Estimate Std. Error MCMC % p-value
## edges    -6.31085    0.03214    0 < 1e-04 ***
## mutual     4.00520    0.15815    1 < 1e-04 ***
## idegree1   0.04638    0.08048    0 0.56443
## odegree1  -0.22703    0.08422    0 0.00703 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 814478 on 587522 degrees of freedom
## Residual Deviance: 16862 on 587518 degrees of freedom
##
## AIC: 16870    BIC: 16915    (Smaller is better.)
m1.2 <- ergm(net ~edges + mutual + idegree(1) + odegree(1))

## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 2.638
## Step length converged once. Increasing MCMC sample size.
## Iteration 2 of at most 20:
## The log-likelihood improved by 0.2537
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.

summary(m1.2)

##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + mutual + idegree(1) + odegree(1)
##
## Iterations: 2 out of 20
##
## Monte Carlo MLE Results:
##      Estimate Std. Error MCMC % p-value
## edges    -6.31226    0.03332    0 < 1e-04 ***
## mutual     3.99514    0.14059    1 < 1e-04 ***
## idegree1   0.04450    0.07991    0 0.57765
## odegree1  -0.22925    0.08504    0 0.00702 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 814478 on 587522 degrees of freedom
## Residual Deviance: 16862 on 587518 degrees of freedom
##
## AIC: 16870    BIC: 16915    (Smaller is better.)
m2 <- ergm(net ~edges + mutual + istar(2) + ostar(2))

```

However, although different features model the same structure, they may have different types of statistical models behind them. If you run `m2`, you will discover that the network did not converge using `istar` and `outstar` features, whereas modelling was successful when we used `idegree` and `odegree`. We can also try adding triangles to control for the closure in the network.

```
m3 <- ergm(net ~edges + mutual + idegree(2) + odegree(2) + triangles)
```

Again, modelling closure with the triangles term was not successful, giving an error message. We can control for both degree and closure using `gwesp` and `gwdegree` terms, and adjusting the lamda.

```
m4 <- ergm(net ~ edges + mutual + gwdegree(0.6, fixed=T))
```

```
## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 5.989
## Iteration 2 of at most 20:
## The log-likelihood improved by 2.036
## Iteration 3 of at most 20:
## The log-likelihood improved by 3.246
## Iteration 4 of at most 20:
## The log-likelihood improved by 2.86
## Iteration 5 of at most 20:
## The log-likelihood improved by 2.58
## Iteration 6 of at most 20:
## The log-likelihood improved by 1.342
## Iteration 7 of at most 20:
## The log-likelihood improved by 0.2233
## Step length converged once. Increasing MCMC sample size.
## Iteration 8 of at most 20:
## The log-likelihood improved by 0.3189
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.
```

```
summary(m4)
```

```
##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + mutual + gwdegree(0.6, fixed = T)
##
## Iterations: 8 out of 20
##
## Monte Carlo MLE Results:
##      Estimate Std. Error MCMC % p-value
## edges      -5.16230    0.03242     0 <1e-04 ***
## mutual       4.01544    0.15353     1 <1e-04 ***
## gwdegree    -3.06112    0.09100     0 <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Null Deviance: 814478 on 587522 degrees of freedom
## Residual Deviance: 15945 on 587519 degrees of freedom
```



```

##
## AIC: 15951    BIC: 15985    (Smaller is better.)

The network converged, negative idegree at lambda 0.6 (which we have experimented with prior to the
tutorial) indicates that the network has areas of ‘clustered closure’ with many triangles around one or few
nodes. We can also add other features that could describe the network structure, or we can start adding
attributes to model ‘who talks to whom’ and other qualitative patterns that describe the network formation.

m_effects <- ergm(net ~ edges + nodefactor('Role') + nodefactor( 'Domain') + nodefactor( 'Continent')+

## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 2.83
## Iteration 2 of at most 20:
## The log-likelihood improved by 1.351
## Iteration 3 of at most 20:
## The log-likelihood improved by 1.717
## Iteration 4 of at most 20:
## The log-likelihood improved by 1.945
## Iteration 5 of at most 20:
## The log-likelihood improved by 1.576
## Iteration 6 of at most 20:
## The log-likelihood improved by 2.212
## Iteration 7 of at most 20:
## The log-likelihood improved by 1.671
## Iteration 8 of at most 20:
## The log-likelihood improved by 2.024
## Iteration 9 of at most 20:
## The log-likelihood improved by 1.719
## Iteration 10 of at most 20:
## The log-likelihood improved by 2.009
## Iteration 11 of at most 20:
## The log-likelihood improved by 2.28
## Iteration 12 of at most 20:
## The log-likelihood improved by 2.105
## Iteration 13 of at most 20:
## The log-likelihood improved by 2.426
## Iteration 14 of at most 20:
## The log-likelihood improved by 1.8
## Step length converged once. Increasing MCMC sample size.
## Iteration 15 of at most 20:
## The log-likelihood improved by 1.834
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.

summary(m_effects)

##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + nodefactor("Role") + nodefactor("Domain") + nodefactor("Continent") +
## mutual + gwidegree(0.6, fixed = T)

```

```

##
## Iterations: 15 out of 20
##
## Monte Carlo MLE Results:
##
## Estimate Std. Error
## edges -2.13231 0.62187
## nodefactor.Role.Student -1.70811 0.11053
## nodefactor.Domain.Community 1.38084 0.17064
## nodefactor.Domain.Elementary and primary education -0.10233 0.22045
## nodefactor.Domain.Entrepreneurship 0.40678 0.08362
## nodefactor.Domain.Government -0.32507 0.34753
## nodefactor.Domain.Health 0.36416 0.13293
## nodefactor.Domain.Higher Education 0.20886 0.07770
## nodefactor.Domain.Languages 0.39752 0.09125
## nodefactor.Domain.Library -0.25147 0.24796
## nodefactor.Domain.Organization 0.02147 0.09510
## nodefactor.Domain.Other -0.43643 0.22309
## nodefactor.Domain.Secondary education -0.01329 0.09102
## nodefactor.Domain.Undergraduate -1.78933 1.00944
## nodefactor.Domain.Unknown -0.52903 0.09743
## nodefactor.Domain.Various 0.21597 0.10178
## nodefactor.Continent.Asia -0.03583 0.30758
## nodefactor.Continent.Australia and NZ 0.21406 0.28633
## nodefactor.Continent.Europe -0.03621 0.28589
## nodefactor.Continent.International 0.62003 0.35209
## nodefactor.Continent.North America -0.11421 0.28781
## nodefactor.Continent.South America 0.02688 0.29331
## nodefactor.Continent.Unknown 0.28563 0.28969
## mutual 3.33258 0.18765
## gwidegree -2.45995 0.10430
##
## MCMC % p-value
## edges 0 0.000606 ***
## nodefactor.Role.Student 1 < 1e-04 ***
## nodefactor.Domain.Community 0 < 1e-04 ***
## nodefactor.Domain.Elementary and primary education 0 0.642516
## nodefactor.Domain.Entrepreneurship 0 < 1e-04 ***
## nodefactor.Domain.Government 0 0.349596
## nodefactor.Domain.Health 0 0.006153 **
## nodefactor.Domain.Higher Education 0 0.007186 **
## nodefactor.Domain.Languages 0 < 1e-04 ***
## nodefactor.Domain.Library 0 0.310506
## nodefactor.Domain.Organization 0 0.821364
## nodefactor.Domain.Other 0 0.050426 .
## nodefactor.Domain.Secondary education 0 0.883866
## nodefactor.Domain.Undergraduate 0 0.076296 .
## nodefactor.Domain.Unknown 0 < 1e-04 ***
## nodefactor.Domain.Various 0 0.033841 *
## nodefactor.Continent.Asia 0 0.907270
## nodefactor.Continent.Australia and NZ 0 0.454717
## nodefactor.Continent.Europe 0 0.899224
## nodefactor.Continent.International 0 0.078234 .
## nodefactor.Continent.North America 0 0.691505
## nodefactor.Continent.South America 0 0.926973
## nodefactor.Continent.Unknown 0 0.324139

```

```

## mutual 1 < 1e-04 ***
## gwidegree 0 < 1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Null Deviance: 814478 on 587522 degrees of freedom
## Residual Deviance: 15409 on 587497 degrees of freedom
##
## AIC: 15459 BIC: 15741 (Smaller is better.)
#check if the order makes a difference here
m_homophily <- ergm (net ~ edges + mutual + gwidegree(0.4, fixed=T)+ nodematch('Role', diff=T) + nodema

## Observed statistic(s) nodematch.Domain.Community, nodematch.Domain.Elementary and primary education,
## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 2.946
## Iteration 2 of at most 20:
## The log-likelihood improved by 2.214
## Iteration 3 of at most 20:
## The log-likelihood improved by 1.565
## Iteration 4 of at most 20:
## The log-likelihood improved by 2.542
## Iteration 5 of at most 20:
## The log-likelihood improved by 1.894
## Iteration 6 of at most 20:
## The log-likelihood improved by 1.468
## Iteration 7 of at most 20:
## The log-likelihood improved by 2.2
## Iteration 8 of at most 20:
## The log-likelihood improved by 2.335
## Iteration 9 of at most 20:
## The log-likelihood improved by 1.832
## Iteration 10 of at most 20:
## The log-likelihood improved by 1.828
## Iteration 11 of at most 20:
## The log-likelihood improved by 2.064
## Iteration 12 of at most 20:
## The log-likelihood improved by 3.091
## Iteration 13 of at most 20:
## The log-likelihood improved by 1.861
## Iteration 14 of at most 20:
## The log-likelihood improved by 1.915
## Iteration 15 of at most 20:
## The log-likelihood improved by 0.5632
## Iteration 16 of at most 20:
## The log-likelihood improved by 1.805
## Iteration 17 of at most 20:
## The log-likelihood improved by 1.853
## Step length converged once. Increasing MCMC sample size.
## Iteration 18 of at most 20:
## The log-likelihood improved by 3.142
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##

```

This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.

```
summary(m_homophily)
```

```
##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + mutual + gwidegree(0.4, fixed = T) + nodematch("Role",
## diff = T) + nodematch("Domain", diff = T) + nodematch("Gender",
## diff = T)
##
## Iterations: 18 out of 20
##
## Monte Carlo MLE Results:
##
## Estimate Std. Error
## edges -3.96516 0.11390
## mutual 3.44622 0.17734
## gwidegree -2.64893 0.10148
## nodematch.Role.Course Instructor 5.85789 NA
## nodematch.Role.Student -1.63267 0.11206
## nodematch.Domain.Business -0.05569 0.44184
## nodematch.Domain.Community -Inf 0.00000
## nodematch.Domain.Elementary and primary education -Inf 0.00000
## nodematch.Domain.Entrepreneurship 1.09072 0.15805
## nodematch.Domain.Government -Inf 0.00000
## nodematch.Domain.Health 2.37777 1.05100
## nodematch.Domain.Higher Education 0.50060 0.07986
## nodematch.Domain.Languages 1.82175 0.18526
## nodematch.Domain.Library -Inf 0.00000
## nodematch.Domain.Organization 1.69152 0.60511
## nodematch.Domain.Other -Inf 0.00000
## nodematch.Domain.Secondary education 0.35483 0.24677
## nodematch.Domain.Undergraduate -Inf 0.00000
## nodematch.Domain.Unknown -0.76446 0.24136
## nodematch.Domain.Various 0.45656 1.02991
## nodematch.Gender.F 0.26946 0.08138
## nodematch.Gender.M 0.22473 0.06791
## nodematch.Gender.Org -0.88015 0.56513
## nodematch.Gender.Unknown 1.14886 0.27583
##
## MCMC % p-value
## edges 1 < 1e-04 ***
## mutual 1 < 1e-04 ***
## gwidegree 0 < 1e-04 ***
## nodematch.Role.Course Instructor NA NA
## nodematch.Role.Student 1 < 1e-04 ***
## nodematch.Domain.Business 0 0.899696
## nodematch.Domain.Community 0 < 1e-04 ***
## nodematch.Domain.Elementary and primary education 0 < 1e-04 ***
## nodematch.Domain.Entrepreneurship 0 < 1e-04 ***
## nodematch.Domain.Government 0 < 1e-04 ***
## nodematch.Domain.Health 0 0.023674 *
## nodematch.Domain.Higher Education 0 < 1e-04 ***
## nodematch.Domain.Languages 1 < 1e-04 ***
```

```

## nodematch.Domain.Library          0 < 1e-04 ***
## nodematch.Domain.Organization     0 0.005183 **
## nodematch.Domain.Other            0 < 1e-04 ***
## nodematch.Domain.Secondary education 0 0.150473
## nodematch.Domain.Undergraduate    0 < 1e-04 ***
## nodematch.Domain.Unknown          0 0.001539 **
## nodematch.Domain.Various          0 0.657549
## nodematch.Gender.F                0 0.000929 ***
## nodematch.Gender.M                0 0.000936 ***
## nodematch.Gender.Org              0 0.119371
## nodematch.Gender.Unknown          0 < 1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Null Deviance: 814478 on 587522 degrees of freedom
## Residual Deviance: NaN on 587498 degrees of freedom
##
## AIC: NaN BIC: NaN (Smaller is better.)
##
## Warning: The following terms have infinite coefficient estimates:
## nodematch.Domain.Community nodematch.Domain.Elementary and primary education nodematch.Domain.Gover
#check if the order makes a difference here
m_homophily2 <- ergm (net ~ edges + mutual + nodefactor( 'Domain' ) + nodefactor( 'Continent' )+ nodema

## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 1.539
## Step length converged once. Increasing MCMC sample size.
## Iteration 2 of at most 20:
## The log-likelihood improved by 2.355
## Iteration 3 of at most 20:
## The log-likelihood improved by 1.194
## Step length converged once. Increasing MCMC sample size.
## Iteration 4 of at most 20:
## The log-likelihood improved by 0.9378
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.

summary(m_homophily2)

##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + mutual + nodefactor("Domain") + nodefactor("Continent") +
## nodematch("Gender") + nodematch("Domain") + nodematch("Continent")
##
## Iterations: 4 out of 20
##
## Monte Carlo MLE Results:
##
## Estimate Std. Error

```

## edges	-6.76123	0.84021
## mutual	3.36167	0.15542
## nodefactor.Domain.Community	2.07824	0.15882
## nodefactor.Domain.Elementary and primary education	-0.14822	0.26379
## nodefactor.Domain.Entrepreneurship	0.71366	0.10163
## nodefactor.Domain.Government	-0.28348	0.42301
## nodefactor.Domain.Health	0.68182	0.17364
## nodefactor.Domain.Higher Education	0.35828	0.09641
## nodefactor.Domain.Languages	0.76109	0.11731
## nodefactor.Domain.Library	-0.21655	0.31433
## nodefactor.Domain.Organization	0.06825	0.11447
## nodefactor.Domain.Other	-0.55964	0.31272
## nodefactor.Domain.Secondary education	-0.01842	0.11289
## nodefactor.Domain.Undergraduate	-2.07859	0.98113
## nodefactor.Domain.Unknown	-0.80120	0.11994
## nodefactor.Domain.Various	0.42135	0.14281
## nodefactor.Continent.Asia	-0.07084	0.43516
## nodefactor.Continent.Australia and NZ	0.30906	0.42283
## nodefactor.Continent.Europe	-0.20567	0.42116
## nodefactor.Continent.International	1.11228	0.51326
## nodefactor.Continent.North America	-0.21037	0.42483
## nodefactor.Continent.South America	0.01807	0.42758
## nodefactor.Continent.Unknown	0.40737	0.42734
## nodematch.Gender	0.14296	0.05837
## nodematch.Domain	0.55434	0.07587
## nodematch.Continent	0.37141	0.06645
##	MCMC %	p-value
## edges	0	< 1e-04 ***
## mutual	1	< 1e-04 ***
## nodefactor.Domain.Community	0	< 1e-04 ***
## nodefactor.Domain.Elementary and primary education	0	0.574191
## nodefactor.Domain.Entrepreneurship	0	< 1e-04 ***
## nodefactor.Domain.Government	0	0.502751
## nodefactor.Domain.Health	0	< 1e-04 ***
## nodefactor.Domain.Higher Education	0	0.000202 ***
## nodefactor.Domain.Languages	0	< 1e-04 ***
## nodefactor.Domain.Library	0	0.490864
## nodefactor.Domain.Organization	0	0.551055
## nodefactor.Domain.Other	0	0.073519 .
## nodefactor.Domain.Secondary education	0	0.870397
## nodefactor.Domain.Undergraduate	0	0.034127 *
## nodefactor.Domain.Unknown	0	< 1e-04 ***
## nodefactor.Domain.Various	0	0.003174 **
## nodefactor.Continent.Asia	0	0.870675
## nodefactor.Continent.Australia and NZ	0	0.464828
## nodefactor.Continent.Europe	0	0.625303
## nodefactor.Continent.International	0	0.030228 *
## nodefactor.Continent.North America	0	0.620462
## nodefactor.Continent.South America	0	0.966297
## nodefactor.Continent.Unknown	0	0.340454
## nodematch.Gender	0	0.014319 *
## nodematch.Domain	0	< 1e-04 ***
## nodematch.Continent	0	< 1e-04 ***
## ---		

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 814478  on 587522  degrees of freedom
## Residual Deviance: 16143  on 587496  degrees of freedom
##
## AIC: 16195    BIC: 16489    (Smaller is better.)
#check if the order makes a difference here
m_homophily3 <- ergm (net ~ edges + mutual + nodefactor( 'Domain' ) + nodefactor( 'Continent' )+ nodema

## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 2.298
## Iteration 2 of at most 20:
## The log-likelihood improved by 2.147
## Iteration 3 of at most 20:
## The log-likelihood improved by 1.62
## Iteration 4 of at most 20:
## The log-likelihood improved by 1.843
## Iteration 5 of at most 20:
## The log-likelihood improved by 1.984
## Iteration 6 of at most 20:
## The log-likelihood improved by 1.719
## Iteration 7 of at most 20:
## The log-likelihood improved by 1.558
## Iteration 8 of at most 20:
## The log-likelihood improved by 1.934
## Iteration 9 of at most 20:
## The log-likelihood improved by 2.174
## Iteration 10 of at most 20:
## The log-likelihood improved by 1.853
## Iteration 11 of at most 20:
## The log-likelihood improved by 1.923
## Iteration 12 of at most 20:
## The log-likelihood improved by 2.007
## Step length converged once. Increasing MCMC sample size.
## Iteration 13 of at most 20:
## The log-likelihood improved by 2.818
## Iteration 14 of at most 20:
## The log-likelihood improved by 2.108
## Iteration 15 of at most 20:
## The log-likelihood improved by 1.966
## Iteration 16 of at most 20:
## The log-likelihood improved by 1.916
## Step length converged once. Increasing MCMC sample size.
## Iteration 17 of at most 20:
## The log-likelihood improved by 2.069
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC.  To examine model diagnostics and check for degeneracy, use the mcmc.
summary(m_homophily3)

##

```

```

## =====
## Summary of model fit
## =====
##
## Formula:  net ~ edges + mutual + nodefactor("Domain") + nodefactor("Continent") +
##          nodematch("Gender") + nodematch("Domain") + nodematch("Continent") +
##          gwidegree(0.4, fixed = T)
##
## Iterations:  17 out of 20
##
## Monte Carlo MLE Results:
##
##                                     Estimate Std. Error
## edges                               -5.831822   0.649245
## mutual                               3.460960   0.177488
## nodefactor.Domain.Community           1.476674   0.139386
## nodefactor.Domain.Elementary and primary education -0.112800   0.200210
## nodefactor.Domain.Entrepreneurship     0.359577   0.087984
## nodefactor.Domain.Government          -0.207487   0.352570
## nodefactor.Domain.Health               0.402996   0.140097
## nodefactor.Domain.Higher Education     0.114911   0.077946
## nodefactor.Domain.Languages            0.413592   0.093049
## nodefactor.Domain.Library             -0.157184   0.250259
## nodefactor.Domain.Organization        -0.008141   0.097102
## nodefactor.Domain.Other               -0.363711   0.255502
## nodefactor.Domain.Secondary education -0.039272   0.089016
## nodefactor.Domain.Undergraduate       -1.767754   0.884596
## nodefactor.Domain.Unknown             -0.563012   0.098987
## nodefactor.Domain.Various              0.259037   0.109225
## nodefactor.Continent.Asia              0.080776   0.338378
## nodefactor.Continent.Australia and NZ  0.254680   0.316598
## nodefactor.Continent.Europe           -0.084697   0.319363
## nodefactor.Continent.International     0.750646   0.381826
## nodefactor.Continent.North America    -0.104168   0.316009
## nodefactor.Continent.South America     0.063380   0.319969
## nodefactor.Continent.Unknown          0.294004   0.320619
## nodematch.Gender                       0.122963   0.061437
## nodematch.Domain                       0.595399   0.075147
## nodematch.Continent                    0.389666   0.073012
## gwidegree                             -2.481798   0.102558
##
##                                     MCMC % p-value
## edges                                 0 < 1e-04 ***
## mutual                                 1 < 1e-04 ***
## nodefactor.Domain.Community            0 < 1e-04 ***
## nodefactor.Domain.Elementary and primary education 0 0.57316
## nodefactor.Domain.Entrepreneurship     0 < 1e-04 ***
## nodefactor.Domain.Government            0 0.55620
## nodefactor.Domain.Health                0 0.00402 **
## nodefactor.Domain.Higher Education     0 0.14042
## nodefactor.Domain.Languages             0 < 1e-04 ***
## nodefactor.Domain.Library               0 0.52995
## nodefactor.Domain.Organization          0 0.93318
## nodefactor.Domain.Other                 0 0.15459
## nodefactor.Domain.Secondary education  0 0.65908
## nodefactor.Domain.Undergraduate        0 0.04568 *

```



```

## nodefactor.Domain.Unknown          0 < 1e-04 ***
## nodefactor.Domain.Various          0 0.01771 *
## nodefactor.Continent.Asia          0 0.81133
## nodefactor.Continent.Australia and NZ 0 0.42115
## nodefactor.Continent.Europe        0 0.79085
## nodefactor.Continent.International 0 0.04931 *
## nodefactor.Continent.North America 0 0.74167
## nodefactor.Continent.South America 0 0.84298
## nodefactor.Continent.Unknown       0 0.35915
## nodematch.Gender                   0 0.04534 *
## nodematch.Domain                   0 < 1e-04 ***
## nodematch.Continent                0 < 1e-04 ***
## gwidegree                           0 < 1e-04 ***
## ---

```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Null Deviance: 814478 on 587522 degrees of freedom
```

```
## Residual Deviance: 15631 on 587495 degrees of freedom
```

```
##
```

```
## AIC: 15685 BIC: 15989 (Smaller is better.)
```

```
#check if the order makes a difference here
```

```
m_alt <- ergm (net ~ edges + mutual + nodematch("Gender", diff=T) + nodematch("Domain") + nodematch ("C
```

```
## Observed statistic(s) nodematch.Continent.Africa and nodematch.Continent.International are at their :
```

```
## Starting maximum likelihood estimation via MCMLE:
```

```
## Iteration 1 of at most 20:
```

```
## The log-likelihood improved by 4.93
```

```
## Iteration 2 of at most 20:
```

```
## The log-likelihood improved by 2.804
```

```
## Iteration 3 of at most 20:
```

```
## The log-likelihood improved by 1.695
```

```
## Iteration 4 of at most 20:
```

```
## The log-likelihood improved by 2.13
```

```
## Iteration 5 of at most 20:
```

```
## The log-likelihood improved by 0.774
```

```
## Iteration 6 of at most 20:
```

```
## The log-likelihood improved by 2.62
```

```
## Iteration 7 of at most 20:
```

```
## The log-likelihood improved by 1.454
```

```
## Iteration 8 of at most 20:
```

```
## The log-likelihood improved by 2.62
```

```
## Iteration 9 of at most 20:
```

```
## The log-likelihood improved by 2.125
```

```
## Iteration 10 of at most 20:
```

```
## The log-likelihood improved by 2.307
```

```
## Iteration 11 of at most 20:
```

```
## The log-likelihood improved by 2.929
```

```
## Iteration 12 of at most 20:
```

```
## The log-likelihood did not improve.
```

```
## Iteration 13 of at most 20:
```

```
## The log-likelihood improved by 2.677
```

```
## Iteration 14 of at most 20:
```

```
## The log-likelihood improved by 1.999
```

```
## Iteration 15 of at most 20:
```

```

## The log-likelihood improved by 0.5366
## Iteration 16 of at most 20:
## The log-likelihood improved by 2.552
## Iteration 17 of at most 20:
## The log-likelihood improved by 1.901
## Iteration 18 of at most 20:
## The log-likelihood improved by 2.161
## Iteration 19 of at most 20:
## The log-likelihood improved by 2.514
## Iteration 20 of at most 20:
## The log-likelihood improved by 2.055

## MCMLE estimation did not converge after 20 iterations. The estimated coefficients may not be accurate.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.

```

```
summary(m_alt)
```

```

##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + mutual + nodematch("Gender", diff = T) + nodematch("Domain") +
##          nodematch("Continent", diff = T) + simmelianties + cyclicalities
##
## Iterations: 20 out of 20
##
## Monte Carlo MLE Results:
##
##          Estimate Std. Error MCMC % p-value
## edges          -6.55345    0.04718    0 <1e-04
## mutual           3.31390    0.21328    2 <1e-04
## nodematch.Gender.F    0.22669    0.09161    0 0.0133
## nodematch.Gender.M    0.34249    0.08755    0 <1e-04
## nodematch.Gender.Org -0.01825    0.21601    0 0.9327
## nodematch.Gender.Unknown -0.09602    0.20718    0 0.6430
## nodematch.Domain     0.45324    0.07371    0 <1e-04
## nodematch.Continent.Africa -Inf    0.00000    0 <1e-04
## nodematch.Continent.Asia  1.41470    0.71044    1 0.0464
## nodematch.Continent.Australia and NZ 1.31363    0.22579    1 <1e-04
## nodematch.Continent.Europe  0.21499    0.12252    0 0.0793
## nodematch.Continent.International -Inf    0.00000    0 <1e-04
## nodematch.Continent.North America  0.22940    0.10391    0 0.0273
## nodematch.Continent.South America -0.48604    0.41699    0 0.2438
## nodematch.Continent.Unknown -0.01119    0.14834    0 0.9399
## simmelianties       15.35843         NA     NA     NA
## cyclicalities       0.96211    0.07654    1 <1e-04
##
## edges          ***
## mutual          ***
## nodematch.Gender.F    *
## nodematch.Gender.M    ***
## nodematch.Gender.Org
## nodematch.Gender.Unknown

```

```

## nodematch.Domain ***
## nodematch.Continent.Africa ***
## nodematch.Continent.Asia *
## nodematch.Continent.Australia and NZ ***
## nodematch.Continent.Europe .
## nodematch.Continent.International ***
## nodematch.Continent.North America *
## nodematch.Continent.South America
## nodematch.Continent.Unknown
## simmelianties
## cyclicalities ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Null Deviance: 814478 on 587522 degrees of freedom
## Residual Deviance: NaN on 587505 degrees of freedom
##
## AIC: NaN BIC: NaN (Smaller is better.)
##
## Warning: The following terms have infinite coefficient estimates:
## nodematch.Continent.Africa nodematch.Continent.International
#check if the order makes a difference here
#create a binary to control for some of the qual attributes and then model them in here
m_final <- ergm (net ~ edges + mutual + gwidegree(0.4, fixed=T) + nodematch("Gender", diff=T) + nodefac

## Starting maximum likelihood estimation via MCMLE:
## Iteration 1 of at most 20:
## The log-likelihood improved by 3.235
## Iteration 2 of at most 20:
## The log-likelihood improved by 1.911
## Iteration 3 of at most 20:
## The log-likelihood improved by 1.937
## Iteration 4 of at most 20:
## The log-likelihood improved by 2.249
## Iteration 5 of at most 20:
## The log-likelihood improved by 2.246
## Iteration 6 of at most 20:
## The log-likelihood improved by 2.229
## Iteration 7 of at most 20:
## The log-likelihood improved by 1.545
## Iteration 8 of at most 20:
## The log-likelihood improved by 2.337
## Iteration 9 of at most 20:
## The log-likelihood improved by 1.96
## Iteration 10 of at most 20:
## The log-likelihood improved by 1.344
## Iteration 11 of at most 20:
## The log-likelihood improved by 1.548
## Iteration 12 of at most 20:
## The log-likelihood improved by 1.857
## Step length converged once. Increasing MCMC sample size.
## Iteration 13 of at most 20:
## The log-likelihood improved by 0.8568

```

```
## Step length converged twice. Stopping.
## Evaluating log-likelihood at the estimate. Using 20 bridges: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
##
## This model was fit using MCMC. To examine model diagnostics and check for degeneracy, use the mcmc.
```

```
summary(m_final)
```

```
##
## =====
## Summary of model fit
## =====
##
## Formula: net ~ edges + mutual + gwidegree(0.4, fixed = T) + nodematch("Gender",
##         diff = T) + nodefactor("Domain")
##
## Iterations: 13 out of 20
##
## Monte Carlo MLE Results:
```

	Estimate	Std. Error
edges	-5.78473	0.15013
mutual	3.59634	0.16389
gwidegree	-2.53252	0.10763
nodematch.Gender.F	0.15082	0.09103
nodematch.Gender.M	0.19250	0.07663
nodematch.Gender.Org	-0.21463	0.24100
nodematch.Gender.Unknown	1.58599	0.17936
nodefactor.Domain.Community	1.35234	0.14653
nodefactor.Domain.Elementary and primary education	-0.12501	0.22240
nodefactor.Domain.Entrepreneurship	0.38321	0.08743
nodefactor.Domain.Government	-0.33408	0.34008
nodefactor.Domain.Health	0.39091	0.13061
nodefactor.Domain.Higher Education	0.26964	0.07835
nodefactor.Domain.Languages	0.43595	0.08915
nodefactor.Domain.Library	-0.12566	0.29238
nodefactor.Domain.Organization	0.33478	0.09816
nodefactor.Domain.Other	-0.33323	0.25319
nodefactor.Domain.Secondary education	0.00479	0.09291
nodefactor.Domain.Undergraduate	-1.90561	0.94525
nodefactor.Domain.Unknown	-0.41209	0.08974
nodefactor.Domain.Various	0.26585	0.10620

```
##
## MCMC % p-value
## edges 0 < 1e-04 ***
## mutual 1 < 1e-04 ***
## gwidegree 0 < 1e-04 ***
## nodematch.Gender.F 0 0.097541 .
## nodematch.Gender.M 0 0.012006 *
## nodematch.Gender.Org 0 0.373155
## nodematch.Gender.Unknown 0 < 1e-04 ***
## nodefactor.Domain.Community 0 < 1e-04 ***
## nodefactor.Domain.Elementary and primary education 0 0.574071
## nodefactor.Domain.Entrepreneurship 0 < 1e-04 ***
## nodefactor.Domain.Government 0 0.325933
## nodefactor.Domain.Health 0 0.002763 **
## nodefactor.Domain.Higher Education 0 0.000579 ***
## nodefactor.Domain.Languages 0 < 1e-04 ***
```

```

## nodefactor.Domain.Library                0 0.667355
## nodefactor.Domain.Organization          0 0.000648 ***
## nodefactor.Domain.Other                 0 0.188140
## nodefactor.Domain.Secondary education  0 0.958885
## nodefactor.Domain.Undergraduate        0 0.043802 *
## nodefactor.Domain.Unknown              0 < 1e-04 ***
## nodefactor.Domain.Various              0 0.012306 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 814478  on 587522  degrees of freedom
## Residual Deviance: 15751  on 587501  degrees of freedom
##
## AIC: 15793    BIC: 16030    (Smaller is better.)

```

Estimating the Final Model

We then run MCMC diagnostics on the final model, check the degeneracy, and see where the model fails to describe the data. We also may want to take a look at the actual odds of tie formation.

```
mcmc.diagnostics(m_final) # running MCMC diagnostics
```

```

## Sample statistics summary:
##
## Iterations = 16384:4209664
## Thinning interval = 1024
## Number of chains = 1
## Sample size per chain = 4096
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##
##              Mean      SD Naive SE
## edges                29.2712 57.224 0.89413
## mutual                 4.4062  7.875 0.12305
## gwidegree             8.2731 17.041 0.26626
## nodematch.Gender.F    11.3701 16.975 0.26523
## nodematch.Gender.M     6.8049 20.634 0.32241
## nodematch.Gender.Org  -0.2764  4.962 0.07753
## nodematch.Gender.Unknown -1.6990  8.695 0.13586
## nodefactor.Domain.Community 8.4607  9.288 0.14513
## nodefactor.Domain.Elementary and primary education -0.4111  4.858 0.07591
## nodefactor.Domain.Entrepreneurship 16.2874 27.611 0.43142
## nodefactor.Domain.Government 0.4795  3.111 0.04862
## nodefactor.Domain.Health 0.8433  9.736 0.15212
## nodefactor.Domain.Higher Education 9.5496 48.982 0.76535
## nodefactor.Domain.Languages 7.0303 22.281 0.34814
## nodefactor.Domain.Library -1.1006  3.602 0.05629
## nodefactor.Domain.Organization 1.4299 23.043 0.36004
## nodefactor.Domain.Other 0.5498  4.168 0.06513
## nodefactor.Domain.Secondary education 8.0144 18.981 0.29658
## nodefactor.Domain.Undergraduate 0.1594  1.074 0.01678
## nodefactor.Domain.Unknown 2.1687 30.141 0.47095
## nodefactor.Domain.Various 6.7886 14.326 0.22384

```

```

##                                     Time-series SE
## edges                             5.33698
## mutual                             1.38424
## gwidegree                          1.30053
## nodematch.Gender.F                 1.65369
## nodematch.Gender.M                 1.74904
## nodematch.Gender.Org               0.27097
## nodematch.Gender.Unknown           0.66180
## nodefactor.Domain.Community        1.12030
## nodefactor.Domain.Elementary and primary education 0.23666
## nodefactor.Domain.Entrepreneurship 3.14648
## nodefactor.Domain.Government       0.19604
## nodefactor.Domain.Health           0.96282
## nodefactor.Domain.Higher Education 4.44503
## nodefactor.Domain.Languages        2.43915
## nodefactor.Domain.Library          0.16639
## nodefactor.Domain.Organization     1.91187
## nodefactor.Domain.Other            0.20215
## nodefactor.Domain.Secondary education 1.15608
## nodefactor.Domain.Undergraduate    0.03714
## nodefactor.Domain.Unknown         2.00254
## nodefactor.Domain.Various         1.70842
##
## 2. Quantiles for each variable:
##
##                                     2.5%    25%    50%
## edges                             -77.00 -11.000 30.00
## mutual                             -8.00  -1.000  4.00
## gwidegree                          -23.86  -3.929  8.08
## nodematch.Gender.F                 -21.00   0.000 11.00
## nodematch.Gender.M                 -33.00  -7.000  6.00
## nodematch.Gender.Org               -9.00   -4.000 -1.00
## nodematch.Gender.Unknown           -17.00  -8.000 -2.00
## nodefactor.Domain.Community        -9.00   2.000  8.00
## nodefactor.Domain.Elementary and primary education -9.00  -4.000 -1.00
## nodefactor.Domain.Entrepreneurship -39.00  -2.000 18.00
## nodefactor.Domain.Government       -4.00   -2.000  0.00
## nodefactor.Domain.Health           -18.00  -6.000  1.00
## nodefactor.Domain.Higher Education -85.00 -25.000  9.00
## nodefactor.Domain.Languages        -36.00  -9.000  7.00
## nodefactor.Domain.Library          -7.00   -4.000 -1.00
## nodefactor.Domain.Organization     -43.00 -14.000  0.00
## nodefactor.Domain.Other            -7.00   -2.000  0.00
## nodefactor.Domain.Secondary education -28.00  -5.000  8.00
## nodefactor.Domain.Undergraduate    -1.00   -1.000  0.00
## nodefactor.Domain.Unknown         -55.00 -18.000  2.00
## nodefactor.Domain.Various         -19.00  -3.000  6.00
##                                     75%    97.5%
## edges                             68.25 141.00
## mutual                             9.00  23.00
## gwidegree                          19.69 41.04
## nodematch.Gender.F                 22.00 46.00
## nodematch.Gender.M                 20.00 47.00
## nodematch.Gender.Org               3.00 10.00

```

```

## nodematch.Gender.Unknown          4.00  17.00
## nodefactor.Domain.Community        14.00  28.00
## nodefactor.Domain.Elementary and primary education  3.00  10.00
## nodefactor.Domain.Entrepreneurship 35.00  70.00
## nodefactor.Domain.Government        2.00   7.00
## nodefactor.Domain.Health            7.00  20.00
## nodefactor.Domain.Higher Education 43.00 108.00
## nodefactor.Domain.Languages         22.25  51.00
## nodefactor.Domain.Library           1.00   7.00
## nodefactor.Domain.Organization      17.00  47.00
## nodefactor.Domain.Other             3.00   9.00
## nodefactor.Domain.Secondary education 21.00  44.00
## nodefactor.Domain.Undergraduate     1.00   3.00
## nodefactor.Domain.Unknown           21.00  67.00
## nodefactor.Domain.Various           16.00  37.00
##
##
## Sample statistics cross-correlations:
##
##                                     edges
## edges                             1.00000000
## mutual                             0.52528837
## gwidegree                          0.81424078
## nodematch.Gender.F                 0.56606948
## nodematch.Gender.M                 0.55597949
## nodematch.Gender.Org               0.13091305
## nodematch.Gender.Unknown           0.07818562
## nodefactor.Domain.Community        0.33486451
## nodefactor.Domain.Elementary and primary education 0.17843832
## nodefactor.Domain.Entrepreneurship 0.55342666
## nodefactor.Domain.Government       0.13304904
## nodefactor.Domain.Health           0.34425574
## nodefactor.Domain.Higher Education 0.74227766
## nodefactor.Domain.Languages        0.46390231
## nodefactor.Domain.Library          0.05049426
## nodefactor.Domain.Organization     0.45630171
## nodefactor.Domain.Other            0.09770270
## nodefactor.Domain.Secondary education 0.45842960
## nodefactor.Domain.Undergraduate    -0.02602773
## nodefactor.Domain.Unknown          0.42083511
## nodefactor.Domain.Various          0.40309779
##
##                                     mutual
## edges                             0.525288369
## mutual                             1.000000000
## gwidegree                          0.389083556
## nodematch.Gender.F                 0.433557988
## nodematch.Gender.M                 0.330126450
## nodematch.Gender.Org               0.024759964
## nodematch.Gender.Unknown           -0.098659380
## nodefactor.Domain.Community        0.359898406
## nodefactor.Domain.Elementary and primary education 0.080661983
## nodefactor.Domain.Entrepreneurship 0.360685703
## nodefactor.Domain.Government       0.094722598
## nodefactor.Domain.Health           0.195447754
## nodefactor.Domain.Higher Education 0.363876722

```

```

## nodefactor.Domain.Languages                0.318156905
## nodefactor.Domain.Library                  0.007208079
## nodefactor.Domain.Organization             0.221777771
## nodefactor.Domain.Other                    0.006227561
## nodefactor.Domain.Secondary education     0.238553453
## nodefactor.Domain.Undergraduate          -0.067257155
## nodefactor.Domain.Unknown                 0.030190069
## nodefactor.Domain.Various                 0.339268316
##                                           gwidegree
## edges                                     0.81424078
## mutual                                    0.38908356
## gwidegree                                 1.00000000
## nodematch.Gender.F                       0.45374346
## nodematch.Gender.M                       0.39511442
## nodematch.Gender.Org                     0.09889622
## nodematch.Gender.Unknown                 0.18734000
## nodefactor.Domain.Community               0.19632883
## nodefactor.Domain.Elementary and primary education 0.14759771
## nodefactor.Domain.Entrepreneurship       0.40918004
## nodefactor.Domain.Government              0.12765676
## nodefactor.Domain.Health                  0.25539353
## nodefactor.Domain.Higher Education        0.57143227
## nodefactor.Domain.Languages               0.33046817
## nodefactor.Domain.Library                 0.05987266
## nodefactor.Domain.Organization            0.35516072
## nodefactor.Domain.Other                   0.09849102
## nodefactor.Domain.Secondary education     0.38871027
## nodefactor.Domain.Undergraduate          -0.01097118
## nodefactor.Domain.Unknown                 0.48358400
## nodefactor.Domain.Various                 0.30345890
##                                           nodematch.Gender.F
## edges                                     0.566069480
## mutual                                    0.433557988
## gwidegree                                 0.453743460
## nodematch.Gender.F                       1.000000000
## nodematch.Gender.M                       0.114999563
## nodematch.Gender.Org                     -0.022418020
## nodematch.Gender.Unknown                 -0.077705304
## nodefactor.Domain.Community               0.337384504
## nodefactor.Domain.Elementary and primary education 0.121542414
## nodefactor.Domain.Entrepreneurship       0.378093885
## nodefactor.Domain.Government              0.033892640
## nodefactor.Domain.Health                  0.190306882
## nodefactor.Domain.Higher Education        0.441491471
## nodefactor.Domain.Languages               0.423933704
## nodefactor.Domain.Library                 0.087004436
## nodefactor.Domain.Organization            0.080378196
## nodefactor.Domain.Other                   0.087215986
## nodefactor.Domain.Secondary education     0.280528898
## nodefactor.Domain.Undergraduate          0.008523978
## nodefactor.Domain.Unknown                 0.093971293
## nodefactor.Domain.Various                 0.225523526
##                                           nodematch.Gender.M
## edges                                     0.555979492

```



```

## mutual 0.330126450
## gwidegree 0.395114416
## nodematch.Gender.F 0.114999563
## nodematch.Gender.M 1.000000000
## nodematch.Gender.Org -0.042138050
## nodematch.Gender.Unknown -0.165273707
## nodefactor.Domain.Community 0.184727203
## nodefactor.Domain.Elementary and primary education 0.148171975
## nodefactor.Domain.Entrepreneurship 0.403621516
## nodefactor.Domain.Government 0.040563286
## nodefactor.Domain.Health 0.196054391
## nodefactor.Domain.Higher Education 0.584247690
## nodefactor.Domain.Languages 0.164265486
## nodefactor.Domain.Library 0.039546446
## nodefactor.Domain.Organization 0.064222512
## nodefactor.Domain.Other 0.003893365
## nodefactor.Domain.Secondary education 0.280426088
## nodefactor.Domain.Undergraduate -0.050522730
## nodefactor.Domain.Unknown 0.007680186
## nodefactor.Domain.Various 0.325366545
## nodematch.Gender.Org 0.1309130527
## edges 0.0247599640
## mutual 0.0988962204
## gwidegree -0.0224180201
## nodematch.Gender.F -0.0421380502
## nodematch.Gender.M 1.0000000000
## nodematch.Gender.Org -0.0005843348
## nodematch.Gender.Unknown 0.0290346517
## nodefactor.Domain.Community 0.0598566971
## nodefactor.Domain.Elementary and primary education 0.0097399516
## nodefactor.Domain.Entrepreneurship 0.0182030194
## nodefactor.Domain.Government 0.0262445498
## nodefactor.Domain.Health -0.0073436601
## nodefactor.Domain.Higher Education 0.0481424516
## nodefactor.Domain.Languages -0.0598367233
## nodefactor.Domain.Library 0.5165095565
## nodefactor.Domain.Organization 0.0600543017
## nodefactor.Domain.Other 0.0177876471
## nodefactor.Domain.Secondary education 0.0117073985
## nodefactor.Domain.Undergraduate 0.0542679771
## nodefactor.Domain.Unknown -0.0484105963
## nodefactor.Domain.Various 0.0781856223
## edges -0.0986593799
## mutual 0.1873400007
## gwidegree -0.0777053037
## nodematch.Gender.F -0.1652737071
## nodematch.Gender.M -0.0005843348
## nodematch.Gender.Org 1.0000000000
## nodematch.Gender.Unknown -0.0517745082
## nodefactor.Domain.Community -0.0328180496
## nodefactor.Domain.Elementary and primary education -0.0704651622
## nodefactor.Domain.Entrepreneurship 0.0827961967
## nodefactor.Domain.Government

```

```

## nodefactor.Domain.Health -0.0451418739
## nodefactor.Domain.Higher Education -0.1342556647
## nodefactor.Domain.Languages -0.0750438714
## nodefactor.Domain.Library -0.0219687180
## nodefactor.Domain.Organization -0.0223221657
## nodefactor.Domain.Other -0.0219836771
## nodefactor.Domain.Secondary education -0.0311519533
## nodefactor.Domain.Undergraduate 0.0621192965
## nodefactor.Domain.Unknown 0.7150731094
## nodefactor.Domain.Various -0.0264838729
## nodefactor.Domain.Community
## edges 0.334864507
## mutual 0.359898406
## gwidegree 0.196328825
## nodematch.Gender.F 0.337384504
## nodematch.Gender.M 0.184727203
## nodematch.Gender.Org 0.029034652
## nodematch.Gender.Unknown -0.051774508
## nodefactor.Domain.Community 1.000000000
## nodefactor.Domain.Elementary and primary education 0.017024675
## nodefactor.Domain.Entrepreneurship 0.073976620
## nodefactor.Domain.Government -0.027055899
## nodefactor.Domain.Health -0.044633276
## nodefactor.Domain.Higher Education 0.202833026
## nodefactor.Domain.Languages 0.149569182
## nodefactor.Domain.Library 0.024937096
## nodefactor.Domain.Organization 0.115523557
## nodefactor.Domain.Other 0.041928885
## nodefactor.Domain.Secondary education 0.262798032
## nodefactor.Domain.Undergraduate -0.020682992
## nodefactor.Domain.Unknown -0.005009924
## nodefactor.Domain.Various 0.327492395
## nodefactor.Domain.Elementary and primary education
## edges 0.17843832
## mutual 0.08066198
## gwidegree 0.14759770
## nodematch.Gender.F 0.12154241
## nodematch.Gender.M 0.14817197
## nodematch.Gender.Org 0.05985669
## nodematch.Gender.Unknown -0.03281805
## nodefactor.Domain.Community 0.01702467
## nodefactor.Domain.Elementary and primary education 1.00000000
## nodefactor.Domain.Entrepreneurship 0.07699861
## nodefactor.Domain.Government -0.02942782
## nodefactor.Domain.Health 0.08465982
## nodefactor.Domain.Higher Education 0.11426477
## nodefactor.Domain.Languages 0.05199152
## nodefactor.Domain.Library -0.00632626
## nodefactor.Domain.Organization 0.07982453
## nodefactor.Domain.Other 0.00723377
## nodefactor.Domain.Secondary education 0.07415218
## nodefactor.Domain.Undergraduate 0.00376641
## nodefactor.Domain.Unknown 0.01899700
## nodefactor.Domain.Various 0.08398319

```

```

##                                     nodefactor.Domain.Entrepreneurship
## edges                               0.553426663
## mutual                               0.360685703
## gwidegree                            0.409180044
## nodematch.Gender.F                   0.378093885
## nodematch.Gender.M                   0.403621516
## nodematch.Gender.Org                 0.009739952
## nodematch.Gender.Unknown             -0.070465162
## nodefactor.Domain.Community           0.073976620
## nodefactor.Domain.Elementary and primary education 0.076998618
## nodefactor.Domain.Entrepreneurship   1.000000000
## nodefactor.Domain.Government          0.041060326
## nodefactor.Domain.Health              0.255293715
## nodefactor.Domain.Higher Education    0.241407413
## nodefactor.Domain.Languages           0.179540464
## nodefactor.Domain.Library             0.005880984
## nodefactor.Domain.Organization        0.154346094
## nodefactor.Domain.Other               0.083963137
## nodefactor.Domain.Secondary education 0.122491932
## nodefactor.Domain.Undergraduate       -0.052547881
## nodefactor.Domain.Unknown             0.126965634
## nodefactor.Domain.Various             0.197965718
##                                     nodefactor.Domain.Government
## edges                               0.133049044
## mutual                               0.094722598
## gwidegree                            0.127656755
## nodematch.Gender.F                   0.033892640
## nodematch.Gender.M                   0.040563286
## nodematch.Gender.Org                 0.018203019
## nodematch.Gender.Unknown             0.082796197
## nodefactor.Domain.Community           -0.027055899
## nodefactor.Domain.Elementary and primary education -0.029427821
## nodefactor.Domain.Entrepreneurship   0.041060326
## nodefactor.Domain.Government          1.000000000
## nodefactor.Domain.Health              0.079447779
## nodefactor.Domain.Higher Education    0.080771491
## nodefactor.Domain.Languages           0.084648987
## nodefactor.Domain.Library             -0.005739795
## nodefactor.Domain.Organization        0.080808741
## nodefactor.Domain.Other               -0.007095082
## nodefactor.Domain.Secondary education -0.040664017
## nodefactor.Domain.Undergraduate       0.066205296
## nodefactor.Domain.Unknown             0.083209884
## nodefactor.Domain.Various             0.005770427
##                                     nodefactor.Domain.Health
## edges                               0.34425574
## mutual                               0.19544775
## gwidegree                            0.25539353
## nodematch.Gender.F                   0.19030688
## nodematch.Gender.M                   0.19605439
## nodematch.Gender.Org                 0.02624455
## nodematch.Gender.Unknown             -0.04514187
## nodefactor.Domain.Community           -0.04463328
## nodefactor.Domain.Elementary and primary education 0.08465983

```

```

## nodefactor.Domain.Entrepreneurship          0.25529371
## nodefactor.Domain.Government                0.07944778
## nodefactor.Domain.Health                    1.00000000
## nodefactor.Domain.Higher Education          0.23571572
## nodefactor.Domain.Languages                 0.13262507
## nodefactor.Domain.Library                  0.08527069
## nodefactor.Domain.Organization              0.09488388
## nodefactor.Domain.Other                     0.00299658
## nodefactor.Domain.Secondary education       0.06898518
## nodefactor.Domain.Undergraduate            -0.03402223
## nodefactor.Domain.Unknown                   0.08078846
## nodefactor.Domain.Various                   0.03251369
##
## nodefactor.Domain.Higher Education
## edges                                       0.742277656
## mutual                                     0.363876722
## gwidegree                                  0.571432270
## nodematch.Gender.F                         0.441491471
## nodematch.Gender.M                         0.584247690
## nodematch.Gender.Org                       -0.007343660
## nodematch.Gender.Unknown                   -0.134255665
## nodefactor.Domain.Community                 0.202833026
## nodefactor.Domain.Elementary and primary education 0.114264779
## nodefactor.Domain.Entrepreneurship         0.241407413
## nodefactor.Domain.Government               0.080771491
## nodefactor.Domain.Health                   0.235715716
## nodefactor.Domain.Higher Education         1.000000000
## nodefactor.Domain.Languages                 0.223604600
## nodefactor.Domain.Library                  0.082915379
## nodefactor.Domain.Organization              0.151164270
## nodefactor.Domain.Other                     0.004840767
## nodefactor.Domain.Secondary education       0.241121231
## nodefactor.Domain.Undergraduate            -0.063872007
## nodefactor.Domain.Unknown                   0.134252904
## nodefactor.Domain.Various                   0.239727892
##
## nodefactor.Domain.Languages
## edges                                       0.46390231
## mutual                                     0.31815691
## gwidegree                                  0.33046817
## nodematch.Gender.F                         0.42393370
## nodematch.Gender.M                         0.16426549
## nodematch.Gender.Org                       0.04814245
## nodematch.Gender.Unknown                   -0.07504387
## nodefactor.Domain.Community                 0.14956918
## nodefactor.Domain.Elementary and primary education 0.05199152
## nodefactor.Domain.Entrepreneurship         0.17954046
## nodefactor.Domain.Government               0.08464899
## nodefactor.Domain.Health                   0.13262507
## nodefactor.Domain.Higher Education         0.22360460
## nodefactor.Domain.Languages                 1.00000000
## nodefactor.Domain.Library                  -0.06072790
## nodefactor.Domain.Organization              0.20462047
## nodefactor.Domain.Other                     -0.02478975
## nodefactor.Domain.Secondary education       0.17197755
## nodefactor.Domain.Undergraduate            0.00702373

```

```

## nodefactor.Domain.Unknown          0.03552940
## nodefactor.Domain.Various          0.04833916
##                                     nodefactor.Domain.Library
## edges                               5.049426e-02
## mutual                              7.208079e-03
## gwidegree                           5.987266e-02
## nodematch.Gender.F                 8.700444e-02
## nodematch.Gender.M                 3.954645e-02
## nodematch.Gender.Org               -5.983672e-02
## nodematch.Gender.Unknown           -2.196872e-02
## nodefactor.Domain.Community         2.493710e-02
## nodefactor.Domain.Elementary and primary education -6.326266e-03
## nodefactor.Domain.Entrepreneurship  5.880984e-03
## nodefactor.Domain.Government        -5.739795e-03
## nodefactor.Domain.Health            8.527069e-02
## nodefactor.Domain.Higher Education  8.291538e-02
## nodefactor.Domain.Languages         -6.072790e-02
## nodefactor.Domain.Library           1.000000e+00
## nodefactor.Domain.Organization      -4.548910e-02
## nodefactor.Domain.Other             2.786601e-02
## nodefactor.Domain.Secondary education -1.988876e-02
## nodefactor.Domain.Undergraduate     -1.087686e-02
## nodefactor.Domain.Unknown           -2.142104e-02
## nodefactor.Domain.Various           8.466649e-05
##                                     nodefactor.Domain.Organization
## edges                               0.4563017149
## mutual                              0.2217777706
## gwidegree                           0.3551607188
## nodematch.Gender.F                 0.0803781956
## nodematch.Gender.M                 0.0642225125
## nodematch.Gender.Org               0.5165095565
## nodematch.Gender.Unknown           -0.0223221657
## nodefactor.Domain.Community         0.1155235566
## nodefactor.Domain.Elementary and primary education 0.0798245366
## nodefactor.Domain.Entrepreneurship  0.1543460945
## nodefactor.Domain.Government        0.0808087414
## nodefactor.Domain.Health            0.0948838808
## nodefactor.Domain.Higher Education  0.1511642696
## nodefactor.Domain.Languages         0.2046204710
## nodefactor.Domain.Library           -0.0454890957
## nodefactor.Domain.Organization      1.0000000000
## nodefactor.Domain.Other             0.0601986359
## nodefactor.Domain.Secondary education 0.1836018787
## nodefactor.Domain.Undergraduate     0.0002294795
## nodefactor.Domain.Unknown           0.0899169532
## nodefactor.Domain.Various           0.1006256053
##                                     nodefactor.Domain.Other
## edges                               0.097702698
## mutual                              0.006227561
## gwidegree                           0.098491017
## nodematch.Gender.F                 0.087215986
## nodematch.Gender.M                 0.003893365
## nodematch.Gender.Org               0.060054302
## nodematch.Gender.Unknown           -0.021983677

```

```

## nodefactor.Domain.Community 0.041928885
## nodefactor.Domain.Elementary and primary education 0.007233770
## nodefactor.Domain.Entrepreneurship 0.083963137
## nodefactor.Domain.Government -0.007095082
## nodefactor.Domain.Health 0.002996580
## nodefactor.Domain.Higher Education 0.004840767
## nodefactor.Domain.Languages -0.024789746
## nodefactor.Domain.Library 0.027866008
## nodefactor.Domain.Organization 0.060198636
## nodefactor.Domain.Other 1.000000000
## nodefactor.Domain.Secondary education 0.051081804
## nodefactor.Domain.Undergraduate -0.027331277
## nodefactor.Domain.Unknown 0.044392080
## nodefactor.Domain.Various 0.054288218
## nodefactor.Domain.Secondary education
## edges 0.45842960
## mutual 0.23855345
## gwidegree 0.38871027
## nodematch.Gender.F 0.28052890
## nodematch.Gender.M 0.28042609
## nodematch.Gender.Org 0.01778765
## nodematch.Gender.Unknown -0.03115195
## nodefactor.Domain.Community 0.26279803
## nodefactor.Domain.Elementary and primary education 0.07415218
## nodefactor.Domain.Entrepreneurship 0.12249193
## nodefactor.Domain.Government -0.04066402
## nodefactor.Domain.Health 0.06898518
## nodefactor.Domain.Higher Education 0.24112123
## nodefactor.Domain.Languages 0.17197755
## nodefactor.Domain.Library -0.01988876
## nodefactor.Domain.Organization 0.18360188
## nodefactor.Domain.Other 0.05108180
## nodefactor.Domain.Secondary education 1.00000000
## nodefactor.Domain.Undergraduate 0.02953683
## nodefactor.Domain.Unknown 0.07878681
## nodefactor.Domain.Various 0.15544870
## nodefactor.Domain.Undergraduate
## edges -0.0260277264
## mutual -0.0672571545
## gwidegree -0.0109711756
## nodematch.Gender.F 0.0085239778
## nodematch.Gender.M -0.0505227304
## nodematch.Gender.Org 0.0117073985
## nodematch.Gender.Unknown 0.0621192965
## nodefactor.Domain.Community -0.0206829922
## nodefactor.Domain.Elementary and primary education 0.0037664150
## nodefactor.Domain.Entrepreneurship -0.0525478810
## nodefactor.Domain.Government 0.0662052963
## nodefactor.Domain.Health -0.0340222261
## nodefactor.Domain.Higher Education -0.0638720072
## nodefactor.Domain.Languages 0.0070237297
## nodefactor.Domain.Library -0.0108768630
## nodefactor.Domain.Organization 0.0002294795
## nodefactor.Domain.Other -0.0273312772

```

```

## nodefactor.Domain.Secondary education          0.0295368326
## nodefactor.Domain.Undergraduate              1.0000000000
## nodefactor.Domain.Unknown                    0.0237631440
## nodefactor.Domain.Various                    -0.0019196493
##
## nodefactor.Domain.Unknown                    nodefactor.Domain.Unknown
## edges                                         0.420835105
## mutual                                       0.030190069
## gwidegree                                    0.483583997
## nodematch.Gender.F                          0.093971293
## nodematch.Gender.M                          0.007680186
## nodematch.Gender.Org                        0.054267977
## nodematch.Gender.Unknown                    0.715073109
## nodefactor.Domain.Community                  -0.005009924
## nodefactor.Domain.Elementary and primary education 0.018997004
## nodefactor.Domain.Entrepreneurship          0.126965634
## nodefactor.Domain.Government                0.083209884
## nodefactor.Domain.Health                    0.080788459
## nodefactor.Domain.Higher Education          0.134252904
## nodefactor.Domain.Languages                 0.035529396
## nodefactor.Domain.Library                   -0.021421042
## nodefactor.Domain.Organization              0.089916953
## nodefactor.Domain.Other                     0.044392080
## nodefactor.Domain.Secondary education       0.078786812
## nodefactor.Domain.Undergraduate            0.023763144
## nodefactor.Domain.Unknown                  1.000000000
## nodefactor.Domain.Various                   0.080650231
##
## nodefactor.Domain.Various                    nodefactor.Domain.Various
## edges                                         4.030978e-01
## mutual                                       3.392683e-01
## gwidegree                                    3.034589e-01
## nodematch.Gender.F                          2.255235e-01
## nodematch.Gender.M                          3.253665e-01
## nodematch.Gender.Org                        -4.841060e-02
## nodematch.Gender.Unknown                    -2.648387e-02
## nodefactor.Domain.Community                  3.274924e-01
## nodefactor.Domain.Elementary and primary education 8.398319e-02
## nodefactor.Domain.Entrepreneurship          1.979657e-01
## nodefactor.Domain.Government                5.770427e-03
## nodefactor.Domain.Health                    3.251369e-02
## nodefactor.Domain.Higher Education          2.397279e-01
## nodefactor.Domain.Languages                 4.833916e-02
## nodefactor.Domain.Library                   8.466649e-05
## nodefactor.Domain.Organization              1.006256e-01
## nodefactor.Domain.Other                     5.428822e-02
## nodefactor.Domain.Secondary education       1.554487e-01
## nodefactor.Domain.Undergraduate            -1.919649e-03
## nodefactor.Domain.Unknown                  8.065023e-02
## nodefactor.Domain.Various                   1.000000e+00
##
## Sample statistics auto-correlation:
## Chain 1
##
## edges      mutual gwidegree nodematch.Gender.F
## Lag 0      1.000000 1.000000 1.000000      1.000000
## Lag 1024  0.9231679 0.9843166 0.8666520      0.9157959

```

```

## Lag 2048 0.8573417 0.9691554 0.7717738 0.8439373
## Lag 3072 0.7963559 0.9540414 0.6908535 0.7811771
## Lag 4096 0.7431480 0.9397253 0.6284173 0.7291950
## Lag 5120 0.7013290 0.9251926 0.5778723 0.6858333
##      nodematch.Gender.M nodematch.Gender.Org nodematch.Gender.Unknown
## Lag 0      1.0000000      1.0000000      1.0000000
## Lag 1024   0.9049342      0.8296553      0.8893481
## Lag 2048   0.8279193      0.7012635      0.7961678
## Lag 3072   0.7624960      0.5990681      0.7226590
## Lag 4096   0.7055890      0.5158060      0.6595153
## Lag 5120   0.6589412      0.4517075      0.6061715
##      nodefactor.Domain.Community
## Lag 0      1.0000000
## Lag 1024   0.9623361
## Lag 2048   0.9282825
## Lag 3072   0.8981062
## Lag 4096   0.8681216
## Lag 5120   0.8394162
##      nodefactor.Domain.Elementary and primary education
## Lag 0      1.0000000
## Lag 1024   0.8007127
## Lag 2048   0.6541440
## Lag 3072   0.5376457
## Lag 4096   0.4441869
## Lag 5120   0.3754919
##      nodefactor.Domain.Employment nodefactor.Domain.Government
## Lag 0      1.0000000      1.0000000
## Lag 1024   0.9258810      0.8196827
## Lag 2048   0.8613121      0.6891870
## Lag 3072   0.8028558      0.5903766
## Lag 4096   0.7510065      0.5125924
## Lag 5120   0.7048265      0.4596795
##      nodefactor.Domain.Health nodefactor.Domain.Higher Education
## Lag 0      1.0000000      1.0000000
## Lag 1024   0.9212351      0.9226188
## Lag 2048   0.8565136      0.8577650
## Lag 3072   0.8016818      0.7986681
## Lag 4096   0.7563360      0.7483667
## Lag 5120   0.7153415      0.7047794
##      nodefactor.Domain.Languages nodefactor.Domain.Library
## Lag 0      1.0000000      1.0000000
## Lag 1024   0.9325507      0.7945902
## Lag 2048   0.8753822      0.6308267
## Lag 3072   0.8282411      0.5054412
## Lag 4096   0.7832273      0.4061327
## Lag 5120   0.7470654      0.3350728
##      nodefactor.Domain.Organization nodefactor.Domain.Other
## Lag 0      1.0000000      1.0000000
## Lag 1024   0.9004615      0.7903174
## Lag 2048   0.8153354      0.6245880
## Lag 3072   0.7448271      0.4981513
## Lag 4096   0.6893699      0.4152040
## Lag 5120   0.6327280      0.3475712
##      nodefactor.Domain.Secondary education

```



```

## Lag 0                1.0000000
## Lag 1024             0.8637854
## Lag 2048             0.7593528
## Lag 3072             0.6696223
## Lag 4096             0.5968343
## Lag 5120             0.5301457
##      nodefactor.Domain.Undergraduate nodefactor.Domain.Unknown
## Lag 0                1.0000000                1.0000000
## Lag 1024             0.6756511                0.8668193
## Lag 2048             0.4418909                0.7559296
## Lag 3072             0.2879208                0.6662231
## Lag 4096             0.1884002                0.5921462
## Lag 5120             0.1239938                0.5326079
##      nodefactor.Domain.Various
## Lag 0                1.0000000
## Lag 1024             0.9209554
## Lag 2048             0.8558088
## Lag 3072             0.8012508
## Lag 4096             0.7529562
## Lag 5120             0.7069675
##
## Sample statistics burn-in diagnostic (Geweke):
## Chain 1
##
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
##
##      edges
##      0.13635
##      mutual
##      0.30489
##      gwidegree
##      -0.51694
##      nodematch.Gender.F
##      2.00935
##      nodematch.Gender.M
##      1.13783
##      nodematch.Gender.Org
##      -0.70290
##      nodematch.Gender.Unknown
##      -1.13595
##      nodefactor.Domain.Community
##      1.63377
## nodefactor.Domain.Elementary and primary education
##      -0.01824
##      nodefactor.Domain.Entrepreneurship
##      1.07273
##      nodefactor.Domain.Government
##      -1.97393
##      nodefactor.Domain.Health
##      -1.67060
##      nodefactor.Domain.Higher Education
##      -0.79660
##      nodefactor.Domain.Languages

```

```

##          3.04768
##          nodefactor.Domain.Library
##          0.11805
##          nodefactor.Domain.Organization
##          -1.08793
##          nodefactor.Domain.Other
##          -0.04613
##          nodefactor.Domain.Secondary education
##          0.73158
##          nodefactor.Domain.Undergraduate
##          0.12539
##          nodefactor.Domain.Unknown
##          -0.32981
##          nodefactor.Domain.Various
##          -0.24767
##
## Individual P-values (lower = worse):
##          edges
##          0.89154524
##          mutual
##          0.76045372
##          gwidegree
##          0.60519760
##          nodematch.Gender.F
##          0.04450024
##          nodematch.Gender.M
##          0.25518998
##          nodematch.Gender.Org
##          0.48211600
##          nodematch.Gender.Unknown
##          0.25597897
##          nodefactor.Domain.Community
##          0.10230771
## nodefactor.Domain.Elementary and primary education
##          0.98544951
##          nodefactor.Domain.Entrepreneurship
##          0.28339295
##          nodefactor.Domain.Government
##          0.04839025
##          nodefactor.Domain.Health
##          0.09480025
##          nodefactor.Domain.Higher Education
##          0.42568544
##          nodefactor.Domain.Languages
##          0.00230618
##          nodefactor.Domain.Library
##          0.90602866
##          nodefactor.Domain.Organization
##          0.27662589
##          nodefactor.Domain.Other
##          0.96320756
##          nodefactor.Domain.Secondary education
##          0.46442623
##          nodefactor.Domain.Undergraduate

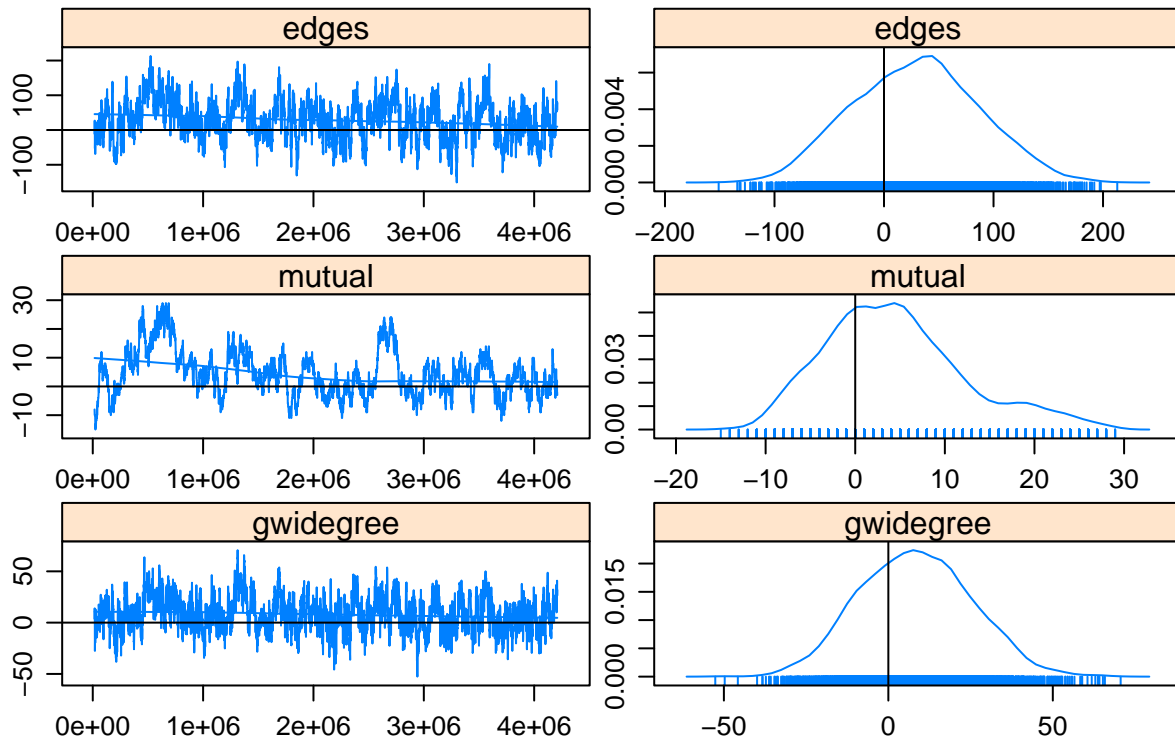
```

```

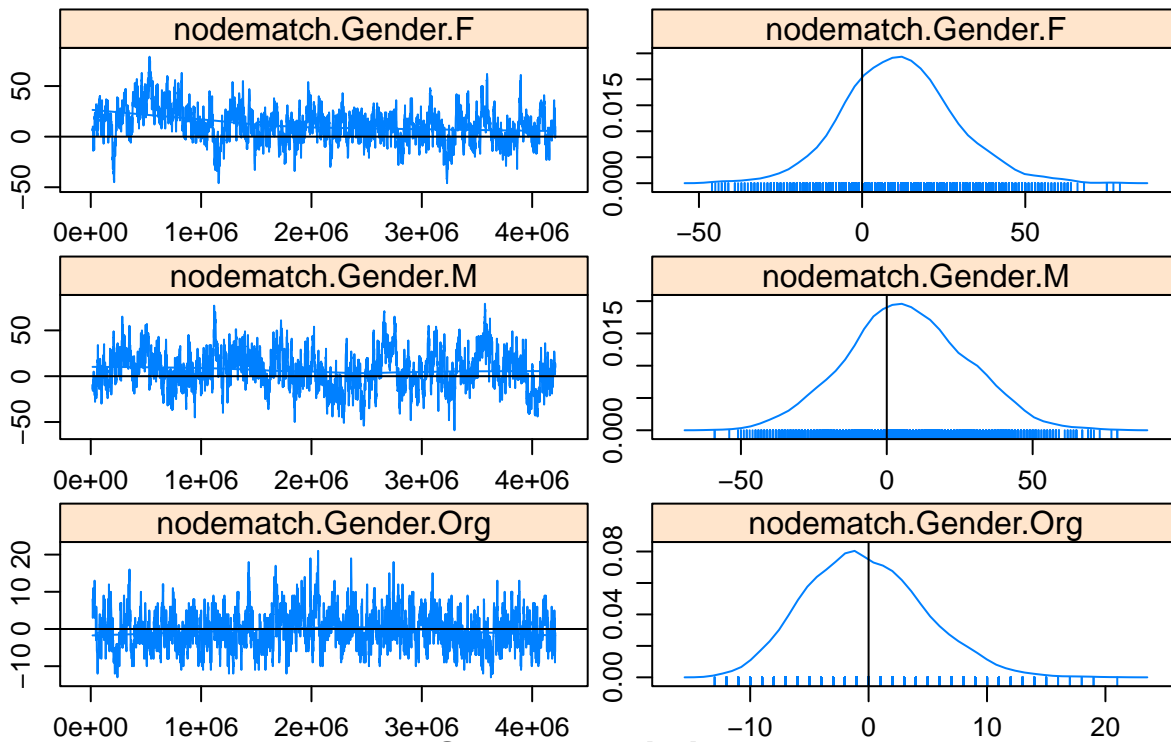
##                               0.90021516
##                               nodefactor.Domain.Unknown
##                               0.74154492
##                               nodefactor.Domain.Various
##                               0.80439224
## Joint P-value (lower = worse): 0.0003274969 .
## Loading required namespace: latticeExtra

```

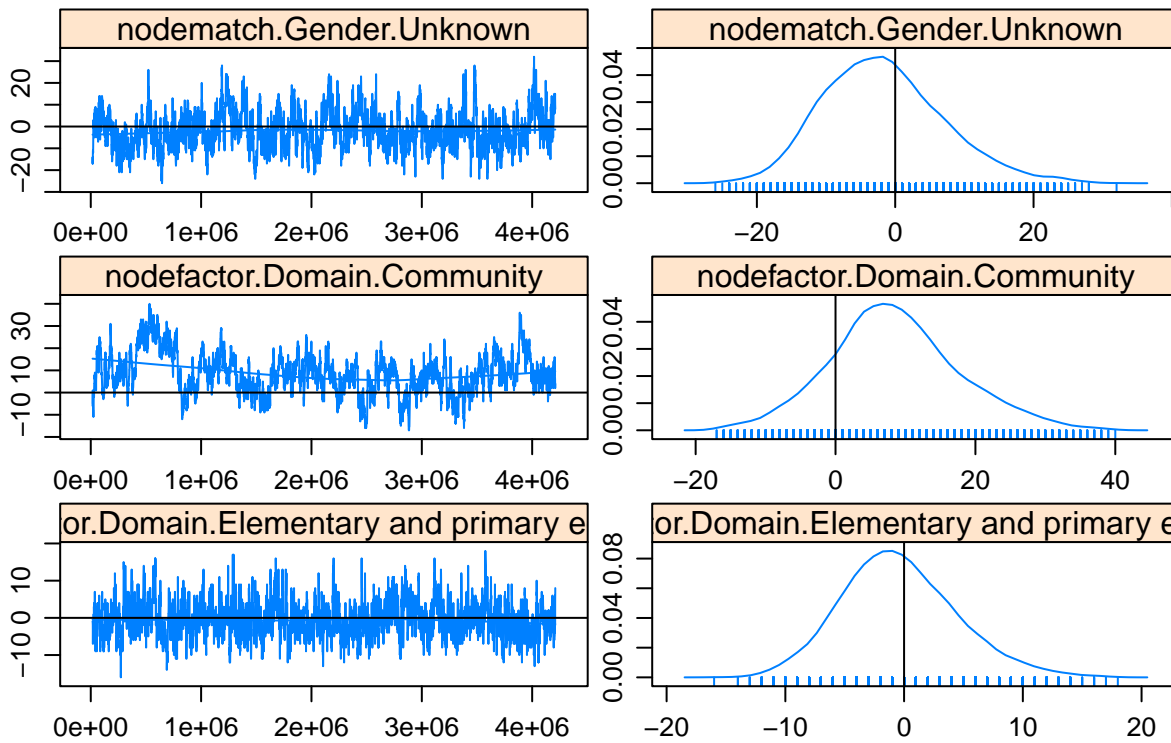
Sample statistics



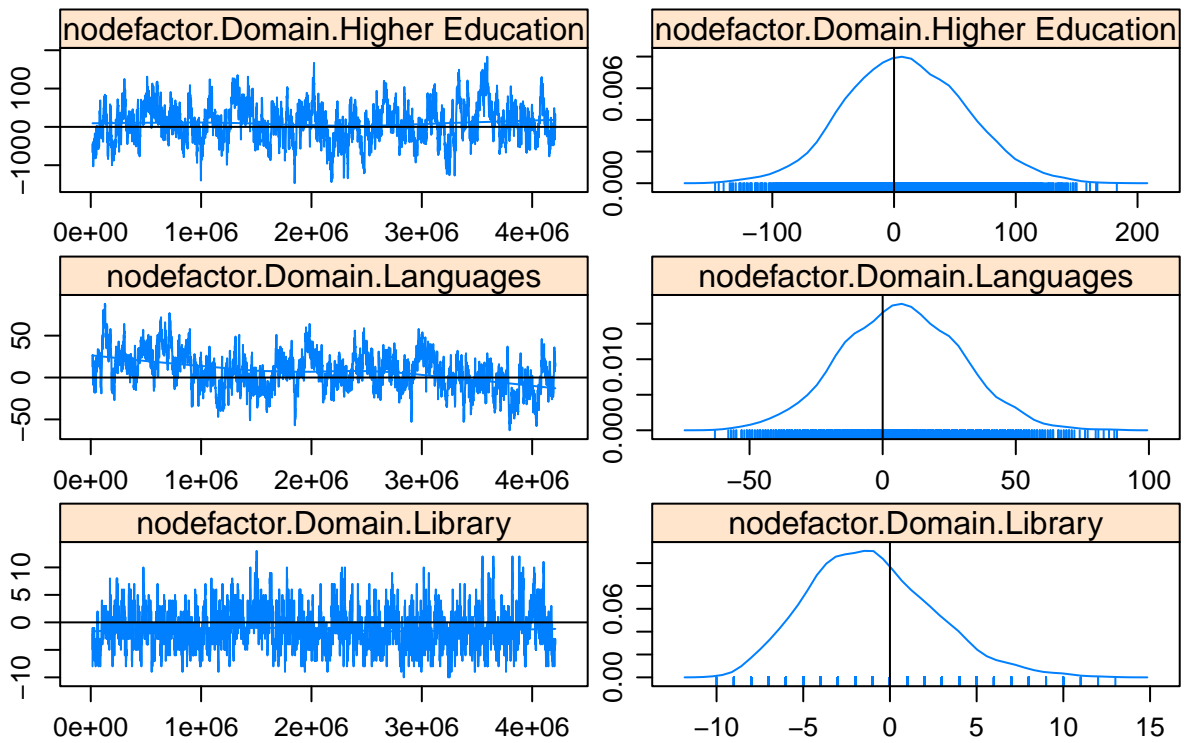
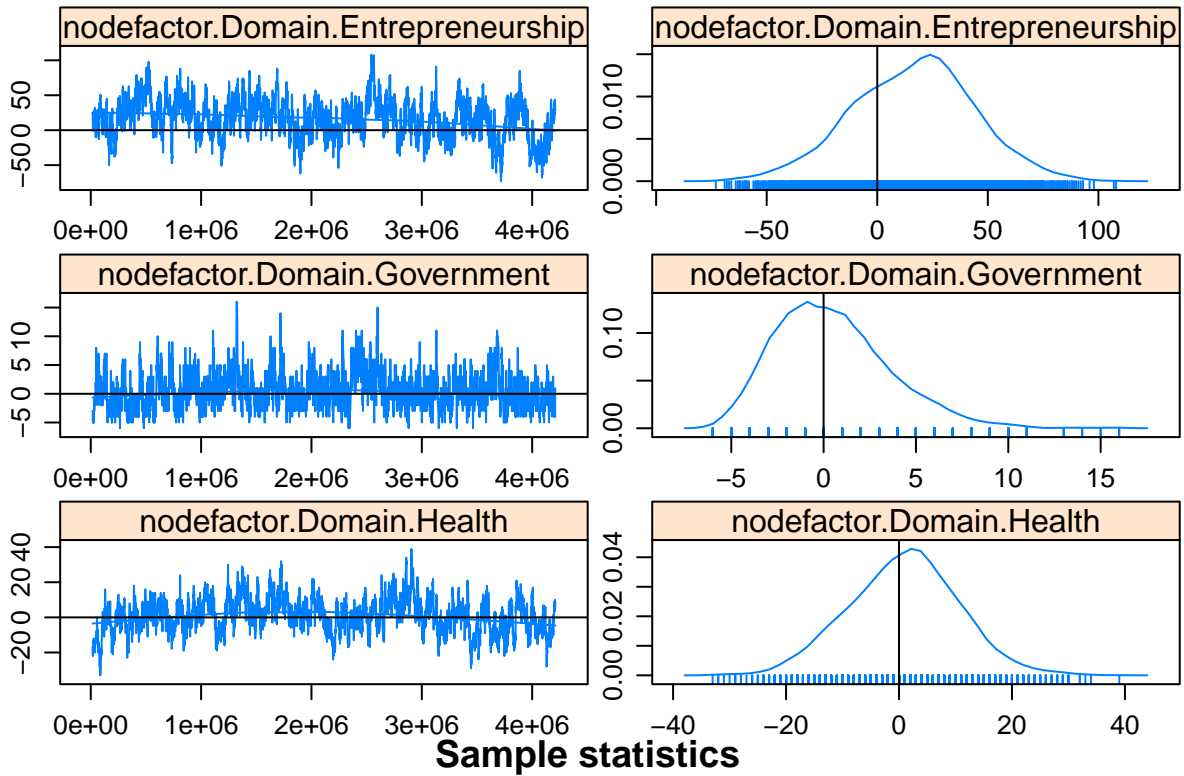
Sample statistics



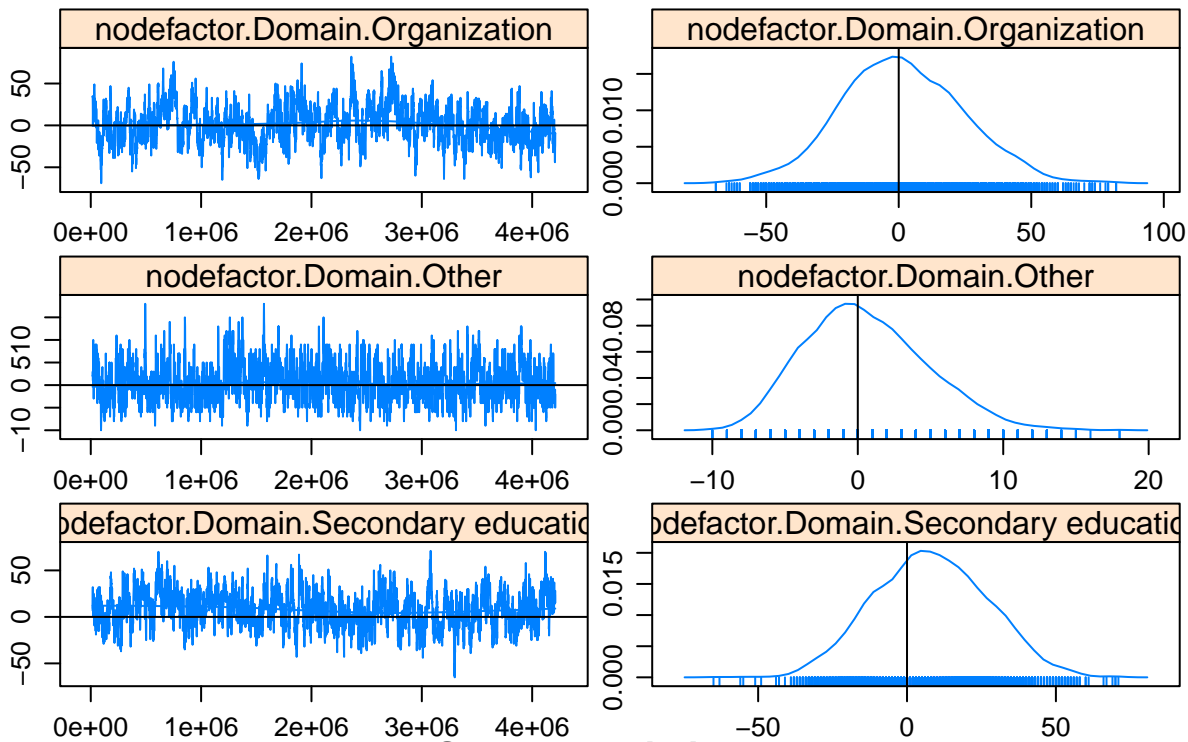
Sample statistics



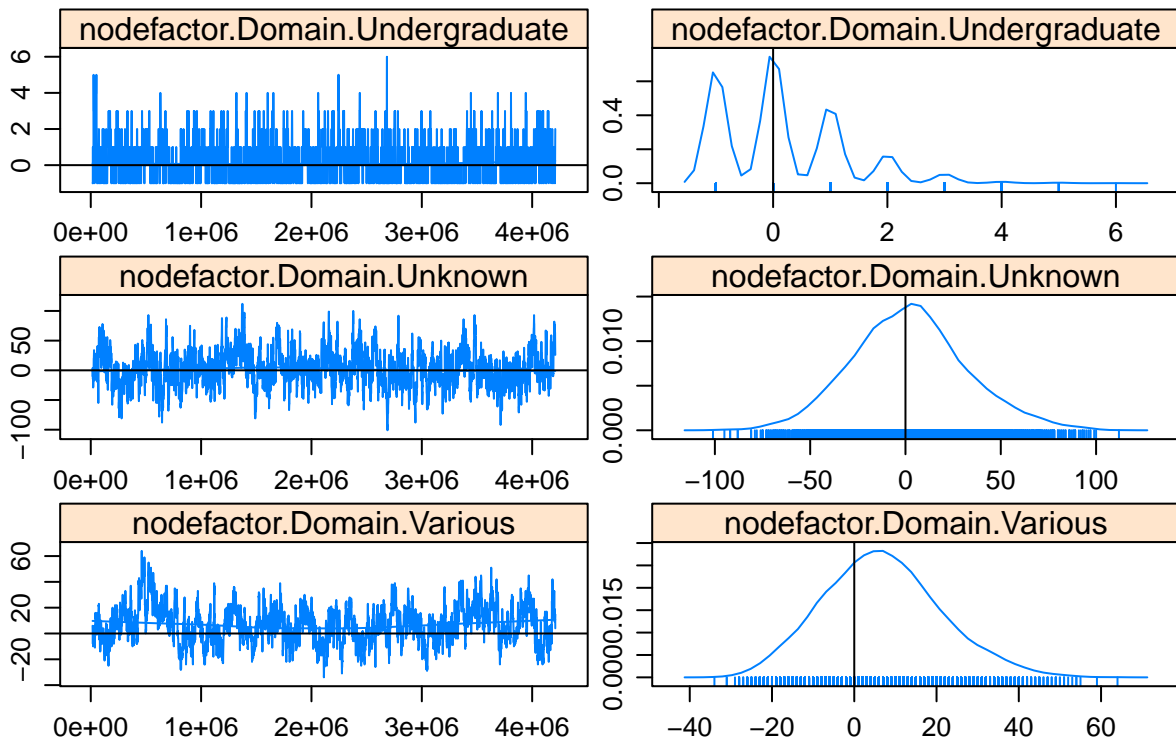
Sample statistics



Sample statistics



Sample statistics



##

MCMC diagnostics shown here are from the last round of simulation, prior to computation of final par

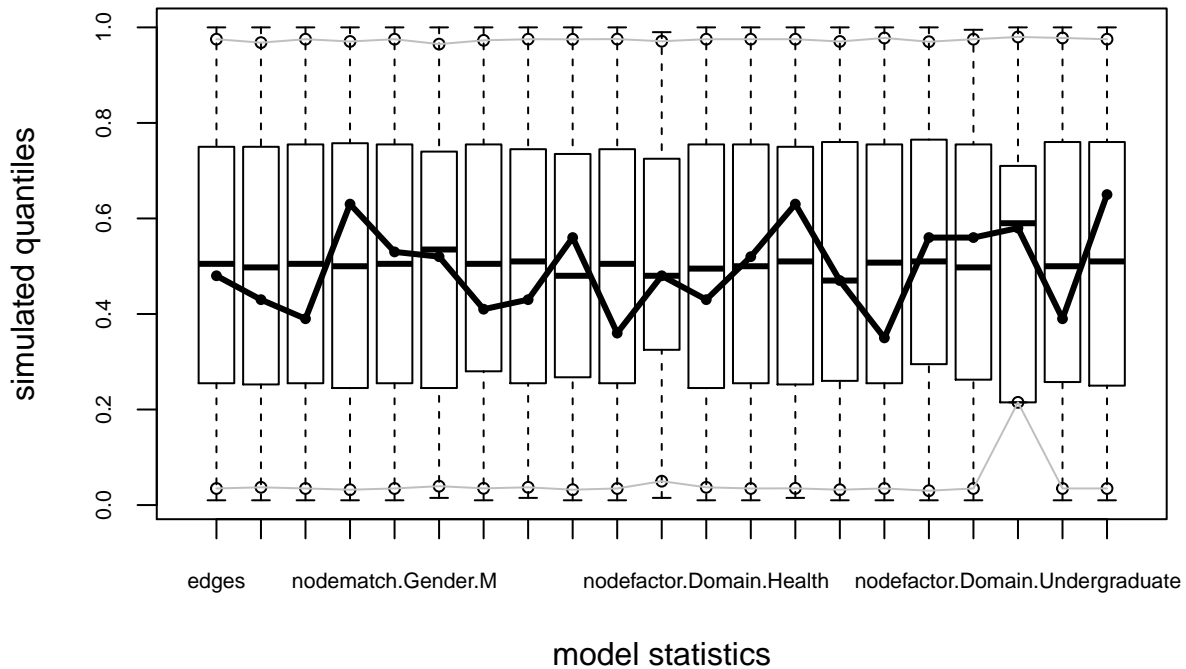
```

mgof <- gof(m_final, GOF=~model) #modelling the network based on our parameters, and comparing observed
plot(mgof) #plotting two degree distributions to compare

mgof.degree <- gof(m_final, GOF=~idegree)
plot(mgof)

```

Goodness-of-fit diagnostics



```

lapply(m_final[1],exp) #checking the odds of the data

```

```

## $coef
##          edges
##          0.00307413
##          mutual
##          36.46457185
##          gwidegree
##          0.07945882
##          nodematch.Gender.F
##          1.16278900
##          nodematch.Gender.M
##          1.21227150
##          nodematch.Gender.Org
##          0.80684170
##          nodematch.Gender.Unknown
##          4.88412830
##          nodefactor.Domain.Community
##          3.86645858
## nodefactor.Domain.Elementary and primary education
##          0.88249189
##          nodefactor.Domain.Entrepreneurship

```

```
## 1.46699164
## nodefactor.Domain.Government
## 0.71600035
## nodefactor.Domain.Health
## 1.47832135
## nodefactor.Domain.Higher Education
## 1.30949686
## nodefactor.Domain.Languages
## 1.54642937
## nodefactor.Domain.Library
## 0.88191498
## nodefactor.Domain.Organization
## 1.39762852
## nodefactor.Domain.Other
## 0.71660583
## nodefactor.Domain.Secondary education
## 1.00480125
## nodefactor.Domain.Undergraduate
## 0.14873159
## nodefactor.Domain.Unknown
## 0.66226765
## nodefactor.Domain.Various
## 1.30453406
```