Statistical Analysis of Networks
Statistics 218
Professor: Mark S. Handcock

Homework 5

Due Friday, December 14, 2018

1) *Model for the Social Space of Monks:* Sampson (1969) recorded the social interactions among a group of monks while resident as an experimenter on vision, and collected numerous sociometric rankings. During his stay, a political "crisis in the cloister" resulted in the expulsion of four monks (Nos. 2, 3, 17, and 18) and the voluntary departure of several others - most immediately, Nos. 1, 7, 14, 15, and 16. (In the end, only 5, 6, 9, and 11 remained). This data is available in the `ergm` package

```
library(ergm)
data(sampson)
help(sampson)
```

Consider the cumulative ties for "liking" over three periods. For this, a tie from monk A to monk B exists if A nominated B as one of his three best friends at any of the three time points.

`R` code to do the computations in the home work are available on the *Homework* page as `samplike.R`.

a) Fit a two-dimensional latent space model. Plot the resulting positions. Describe the groups you can see in the data. Is it clear how many groups there are? Are there monks that do not fall into groups? Give an interpretation of the distance between the first two monks.

b) White *et. al* (p. 749-754) suggests some groups of monks. Look at White *et. al*'s paper (p. 749-754) to see what they found and what Sampson, the ethnographer, found. Do the groups found by the model correspond to those in White *et. al*?

c) Create a plot of 12 samples from the posterior distribution of positions in social space. Do they vary - that is, are we uncertain (based on this data) about the precise positions? Is the magnitude of the uncertainty in positions enough to cloud the positions of the monks?

2) *Joint Modeling of the Sub-groups and Social Space of Monks:* Consider again the data on the social interactions among a group of monks. We will now fit a model for sub-groups that also allows a latent space aspect.

`R` code to do the computations in the home work are available on the *Homework* page as `samplike.cluster.R`.

Note: You will need the `mclust` and `shapes` packages for this to work. You can install them using `install.packages("mclust")` and `install.packages("shapes")`.

a) Fit a two-dimensional latent cluster model. Plot the resulting positions. Describe the groups you can see in the data. Is it clear how many groups there are? Are there monks that do not fall into groups?
Warning: This may take 10-30 minutes so take a break while it is running!

b) White *et. al* (p. 749-754) suggests some groups of monks. Do the groups found by the joint model correspond to those in White *et. al*?

c) Print out the probabilities of group membership for each monk. Is there much ambiguity in the membership for any monk?

d) Plot the groups for the three group model and the four group model. The plotted group colors are the most likely groups for each monk. Why do only three colors appear for the four group model? Can you give an interpretation of the fourth group?

e) Plot samples from the posterior distribution of positions in social space. Do they vary - that is, are we uncertain (based on this data) about the precise positions? Is the magnitude of the uncertainty in positions enough to cloud the positions of the monks?

3) *"Block" models for the Social Relations of Monks:* An alternative and older approach to modeling is that of grouping the monks into "groups" (or "blocks") based on a clustering the sociomatrix.

a) Using the `sna` package, fit a hierarchical clustering model to find clusters of monks. Plot the results.

Use this to produce a plot of the "blocks" using the `blockmodel` function.

How do the groups compare to those of the social space model? How do they compare to those of White *et. al*?

b) Using the `blockmodels` package on CRAN, fit a stochastic block model to the monks.

How do the groups compare to those of the other models? How do they compare to those of White *et. al*?

c) Briefly summarize the latent space and block model fits. What are the strengths and weaknesses of each?