Assignments $\mathcal{N}^{\underline{o}}$ 3

released: 13.11.2013   **due:** 19.11.2013, 12:00h

Task 1: Different edge statistics in the ERGM   5 points

(1) Recall the "boys&girls" planted partition model we used in the last assignment sheet. Determine the parameters θ_1 for the homophily edge statistic and θ_2 for the heterophily edge statistic as a function of the associated density.

(2) Consider now a slightly different model with the edge statistic in the entire network and the homophily edge statistic. How can we derive the parameter θ for the edge statistic and θ_1 for the homophily statistic?
(There is linear dependence with the parameters in (1))

Task 2: Dyad Dependency in ERGMs   5 points

Let $\mathcal{G}$ the set of undirected, loopless graphs with $n = 3$ vertices and consider an exponential random graph model $(\mathcal{G}, P)$ with only one statistic, namely

(a) $t(G)$ (the number of triangles) with associated parameter value $\ln 3$.

(b) $m_a(G)$ (the number of edges connecting actors with the same attribute value) with associated parameter value $\ln 3$. In our case, let a divide the node set $\{1, 2, 3\}$ into *even* and *odd* numbers.

For each case separately, a) triangle statistic and b) homophily statistic, prove whether edge probabilities are dependent or independent.

Task 3: R: Grey's Anatomy and ERGM   **10 points**

For those who haven't heard of the television show before, *Grey's Anatomy* is a widely popular, award-winning medical drama. Meeting conventional medical drama expectations, the show quite regularly features members of its attractive cast "hooking up". In this task we will examine the "hook up network" of the show with the ERGM and try to find network statistics that might play a role in the tie formation process.

Preparatory step:

- Download the adjacency matrix and the node attributes of the grey's anatomy hook up network from the lecture homepage.

(1) In the lecture, we discussed a way to transform the p value of a $G(n, p)$ to the edge parameter θ of an ERGM. Calculate θ and sample a few networks from this ERGM and plot them. What differences to the original network can be observed?
(plot male/female with different colors)

(2) Consider now the set of female actors V_f and male actors V_m. Calculate p_{ff}, p_{mf}, p_{mm} for a planted partition model and derive the parameters $\theta_{ff}, \theta_{mf}, \theta_{mm}$ analogously to (1) for the *nodemix("sex")* statistics. What goes wrong and how could you work around the problem? Sample from the model and plot them again. What has changed in contrast to (1) and what are differences to the original model?

(3) Estimate the parameters for edges, triangles, *nodematch("sex")* (i.e. the tendency of homosexual relationships) and the *degree(1)* (i.e. the tendency to be monogamous) with the *ergm* function. Explain the parameter for the triangle statistic.
(This model should now be vary acurate. If you sample from this model, the networks should look similar to the original network.)

(4) We've so far modeled the network on the number of edge, triangles and the propensity toward homosexuality. What other aspects could be meaningful to analyze? (Textual answer is sufficient)
(Look at the attribute table of the actors.)

Send your R-Script to david.schoch@uni-konstanz.de
A script with an introduction to ergms can be found on the homepage.