## Using Exponential Random Graph Models to build Structural Dependence Tests between Behavior and Network Data

George G. Vega Yon, MS

Kayla de la Haye, PhD

In general, depending on what assumptions we make, we can build tests for identifying structural dependence between a graph *G* and a vector of outcome variables *Y* by either performing a permutation of *Y*, or sampling from a particular graph distribution by rewiring *G*. Overall, the problem with conducting a permutation test relies on the fact that when doing so we are essentially assuming that *Y* is completely random and thus not controlling for any of the observed features of the nodes in the graph. On the other hand, rewiring-based tests, while in principle can be conducted such that the chosen graph distribution preserves some features of the network (its degree sequence, for example), fails to capture more complex properties of a network.

In this presentation, we show a novel statistical method that improves graph sampling based approaches for structural dependence tests. Here, instead of looking at canonical algorithms used to draw samples from networks with a particular set of prescribed properties such as degree sequence or triadic counts, we use a semi-parametric approach based on sampling from an Exponential Random Graph family fitted to the observed graph. Exponential Random Graph Models (ERGMs) have the advantage that they allow us characterize a social network using more complex structures such as homophilic ties, triangle counts, ego-centric features, mutuality, etc., furthermore, our approach goes beyond the aforementioned sampling approaches by, in principle, increasing the sample size of the sampling distribution—for instance, in a small network of size 4 the number of networks with the same prescribed degree sequence as any given network is significantly smaller than the 4096 (2^[n\*(n-1)]) possible networks of size 4 that an Exponential Family characterizes in an ERGM—while at the same time restricting it to the relevant sampling space of those networks that are more likely to arise from such data.

From the statistical point of view, we show that, under the assumption of conditional mean independence, we can do inference regarding the association between *G* and *Y*. Moreover, in general our test allows using almost any definition of test statistic that the researcher can think of, which goes from using a mere correlation (testing for example the association between the number of balanced triads and some behavior *Y*), to more complex statistics such as the resulting set of coefficients obtained from a Generalized Linear Model.

Finally, we show an application of the method using experimental data on small team networks with which, giving its small size, allows us to apply an exact version of the test.

Title: Inference and Generalisability in Modelling Samples of Networks and Multi-Level Network Data

Organizer: Pavel Krivitsky

Co-organizer: Marijtje A. J. van Duijn

Session Description: Sociometric data that we collect are increasingly rich, and we increasingly analyse not single networks but ensembles of networks. Data using the same name generator on disjoint sets of actors in disjoint but similar settings have been collected about classrooms, schools, households, firms, legislative bodies, and other such replicable scenarios. Given such data, we often wish to pool the information from these multiple networks, and to draw conclusions generalisable to a broader population of networks in those settings. Methods to do so range from post-hoc meta-analyses to full hierarchical multi-level models.  
  
These joint analyses raise a number of methodological questions, however. Some of them are questions that are asked in any situation that involves sampling from a population:  
\* What does it mean to draw a representative sample of networks?  
\* Can networks selected using different procedures be analysed together, and how?  
\* What "population" quantities are actually being estimated when metanalyses are performed or multilevel models fit?  
  
Others are specific to social networks:  
\* Can the same model be fit to all of the networks in the ensemble?  
\* How can parameter estimates from networks that vary in size and/or composition be compared?  
  
The goal of this session is to bring attention to these questions and to propose methods and diagnostics for joint estimation of models for multiple networks or for networks with multi-level structure. We welcome contributions on any of the above questions, related questions, or applications in which generalisability or inference to populations of networks play a role.

Title: Events, Situations and Small Networks

Organizer: Christian Stegbauer

Co-organizer: Iris Clemens

Session Description: Situations are a basic component of network analysis. Relationships arises from the coming together of people (face-to-face or also medially). So we can state that networks have their origin in situations. When relationships are measured, e.g. by network generators in classical network analysis, they can be seen as a surrogate of a sequence of common situations. On the downside, situations can also be considered as one of the modes in two-mode network analysis. Situations can be modeled as events; however, what happens at an event is usually of no interest at all in network analysis.  
  
The more network research is interested in these contents, the more likely it is to open up new and extended fields of investigation. Network research can explain how different cultures emerge and are distributed. In situations, ties are negotiated. Although relationships do follow general rules, all relationships differ from each other. In situations and repeated situations (chains of situations) a separate culture is negotiated. Such "small cultures" differ in regard to e.g. their symbols, preferences and behavior.  
  
Nevertheless, situations are always framed, and they follow general context rules in addition to joint negotiations. In situations with few participants, negotiations are easier to conduct because conventions are easier to modify. For this reason, the potential for social innovation is most likely to be found there. Another problem is that of transmission: When cultural changes begin in situations, how is it possible that they become common sense?  
  
The aim of this session is to bring together ideas and empiric work on the mentioned topics in network research: events, situations, the emergence, transmission and diffusion of culture, small networks and the impact of framing for the structure of networks. Submissions in the described field are very welcome.