(Goodreau et al. 2009)

Goodreau, Steven M, James A Kitts, Carter Butts, Marijtje van Duijn, Krista Gile, Deven Hamilton, Pavel Krivitsky Steven Goodreau, and Martina Morris. 2009. “Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks.” *Demography* 46 (1): 103–25.

Introduction summary:

Goodreau et al. (2009) examine the application of ERGMs to analyze adolescent social networks. This study provides insightful methodologies that I adapt to evaluate the grant proposal network. ERGMs are akin to logistic regression but differ in accounting for the recursive dependency of tie probabilities, necessitating complex statistical computations (Goodreau et al. 2009).

Sociality refers to the intrinsic tendency of an individual to form friendships. It is influenced by various factors such as personality, sociodemographic characteristics, or even external circumstances. Goodreau et al., considered sociality as a social process that contributes to the outcome degree. While degree is a directly measurable attribute (count of connections), sociality is more of an inferred characteristic based on observed patterns of tie formation. For instance, if a person has a high degree (many connections), it could be inferred that they have high sociality, but this is not always a direct correlation, as other factors like organizational structure, external incentives, or opportunities for interaction can also influence the degree. In ERGMs, the edges term captures the overall sociality or propensity to form ties in the network. It is used as a baseline measure of the tendency of network actors to form connections, regardless of specific attributes.

Goodreau et al. (2009) study included the exploration of the propensity of individuals to form connections based on shared attributes, examining both assortative mixing (homophily) and disassortative mixing (heterophily). They differentiated between uniform and differential homophily.

\_\_\_ Reading Notes\_\_

Social networks can make predictions about demographic attributes and behaviors (Goodreau et al. 2009). These are micro-level foundations of population-level relational structure.

ERGMs model the probability that a relationship exists as a linear function of predictors. While similar to logistic regression, tie probabilities are recursively dependent, requiring computationally intensive statistics.

Triad closure, a friend of a friend is my friend, researchers measure the excess number of relational triangles. A single triangle count may reflect many social processes, making it dependent on other triangles.

Sociality is a person's overall propensity to make friends. It's not synonymous with the degree. The degree is the outcome. Sociality is a stochastic process shaped by both sociality and other factors. Highly social categories are disproportionately represented within partnerships. Sociality is based on counts of ties observed. “The s term acts as an intercept, and the coefficient for s represents the conditional log-odds of a tie for the reference category.” “s represents the log-odds of a tie and thus has the number of non-ties in the denominator, the magnitude of s is directly affected by network density.” “we expect the coefficient on s to decline with school size.” [The edges term!] “The coefficients on the k and h statistics should be invariant to school size because they are defined relative to s.”

[In the grant network, the variables that shape a faculty member’s grant proposal “sociality” include their discipline's need to acquire grant proposals, their experience as a researcher, and their department’s workload policy. Perhaps their personality and other personal factors like physical appearance. I do not define these attributes in my thesis, but they might be useful to consider for future projects.]

Selective mixing: a person's propensity to make friends based on the combination of individual attributes. Assortative mixing (homophily) is when there is a greater propensity to form relationships with those similar to one’s trait. Disassortative mixing (heterophily) is selection by an opposite or different trait. [The college and quartile attribute captures this in the grant proposal network.] Homophily is an outcome trait. “We consider two selective mixing dynamics. The first is a homogeneous propensity for assortative mixing across attribute categories (“uniform homophily”). The second is a propensity that is specific to individual categories (“differential homophily”).” [In my study uniform homophily is used for the College term and there are only ## types of colleges. I use differential homophily on the quartile rank attribute.]

Triad closure: a person’s friend’s choice of friends. Three persons form a set. When two persons have ties with one individual, the two tend also to form a tie. [If one person proposes two grants with two different people, the two different people tend also to propose a grant together.] structural balance theory. Triadic closure is when the two people encounter each other through the shared time with the third person. This is different from transitivity, which is based on structural balance theory. In transitivity, a cognitive process causes two people to value each other. [For example, a triad is created due to similar research areas.] GWESP statistic

Sociality, homophily, and transitivity can induce one another. Estimates require controlling for the others using the attribute composition of all relationships.

Dyadic independence models only contain node attribute terms and are very similar to traditional logistic models. The probability of a tie does not depend on the value of other ties, only the actors involved. Maximum pseudolikelihood estimation (MPLE) is equivalent to maximum likelihood estimation.

The endogenous process of tie formation involves dyadic dependence. These are recursive dependencies. “Markov chain Monte Carlo (MCMC) simulation methods generate a sample from a space of possible networks to estimate a vector.”

[I use MPLE estimation for models with only dyadic independent terms and MCMC estimation for dyadic dependence models. I need to include an explanation in my thesis.]

“clustering in networks is often represented as a tendency to close triangles” The clustering coefficient tells of the number of triangles divided by the number of triads with two of more ties.

Model degeneracy occurs when we specify a model that is unlikely to produce the observed network. Specify a better-fitting model.

Geometrically weighted edgewise shared partner distribution (GWESP) captures a decreasing marginal impact in the effect of triangles on tie formation. “Two actors “share” a partner if both have a tie to the same partner, and each shared partner forms a triangle if the original pair are tied.” “The GWESP statistic defines a parametric form of a shared partner count distribution that gives each additional shared partner a declining positive impact on the probability of two persons forming a tie.

“We consider three models based upon this set of statistics: a demographic attribute (DA) model that considers only attribute-based, dyadic-independent processes (s, ki, hi); a triad closure (TC) model that considers only a homogeneous, dyadic-dependent process (s, w); and a full model combining the two (s, ki, hi, w).” “Our hypothesis is that selective mixing and triad closure terms (hi, w) will be positive in all models.” “We also expect that both selective mixing and triad closure are important to friendship formation, so we expect the full model to fit better than either partial model.” “Because the full model fits best for all individual schools, we consider its coefficients to be our best estimates for the true magnitude of sociality, selective mixing, and triad closure.”

Goodness of fit: