Exponential Random Graph Models for Social Networks

ERGM Introduction

Carter T. Butts
Department of Sociology and
Institute for Mathematical Behavioral Sciences
University of California, Irvine

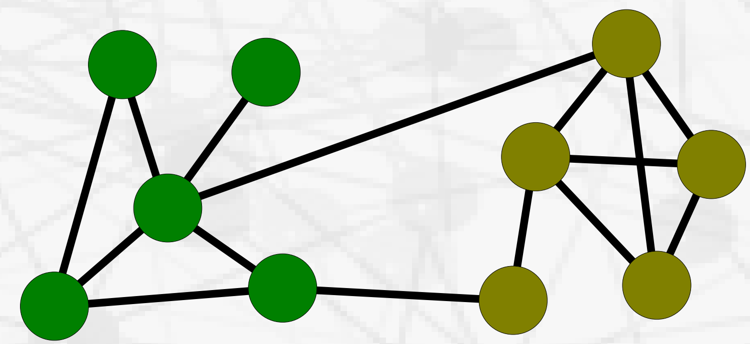
2011 Political Networks Conference
June 15, 2011
Ann Arbor, MI
http://polnet2011.statnet.org/

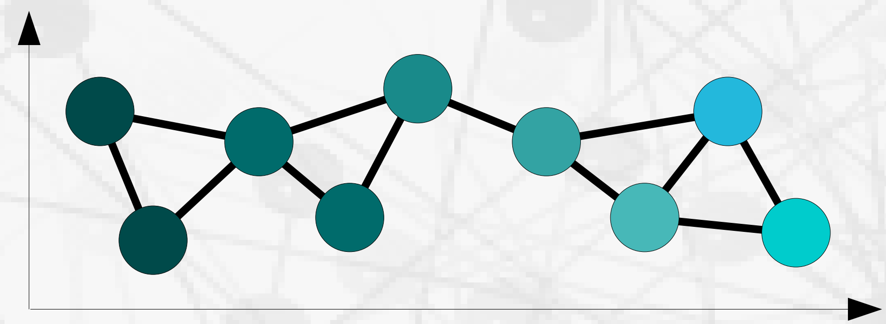
From Description to Modeling

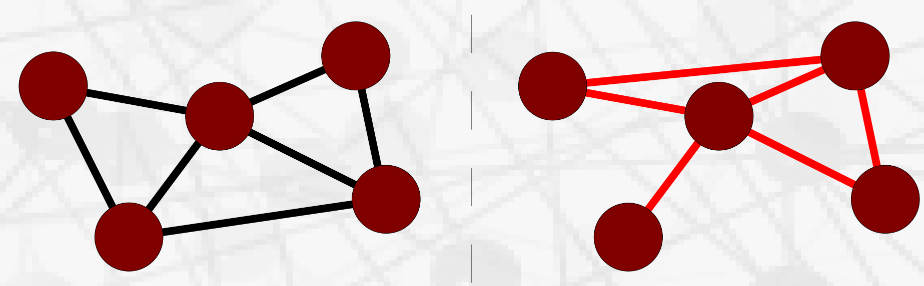
- **Ultimately, want to do more than describe networks**
- **<u>Network modeling</u>: predict the formation and structure of social networks**
- **Many examples**
  - Conditional uniform graphs, Bernoulli graphs
  - Holland and Leinhardt's p_1
  - Degree distribution models, growth models, etc.
- **ERGM: a general representation for such models**
  - Draws on theory of statistical exponential families
  - Not really a "type" of model (in a scientific sense), but a way of representing and working with new and existing models!

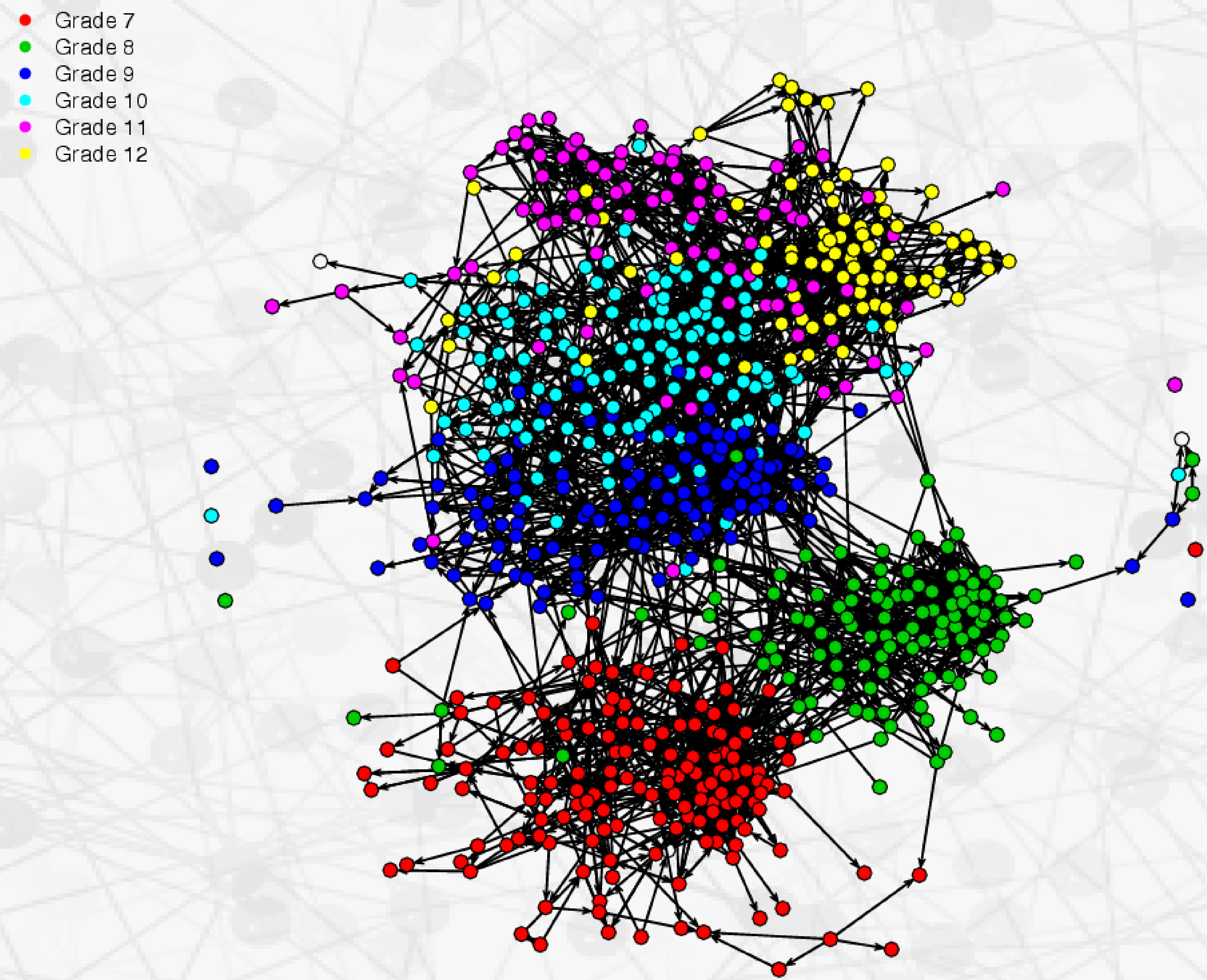
Initial Intuition: Factors in Tie Formation

- **All ties are not equally probable**
  - Chance of an (i,j) edge may depend on properties of i and j
  - Can also depend on other (i,j) relationships
- **Some examples:**
  - Homophily
  - Propinquity
  - Multiplexity

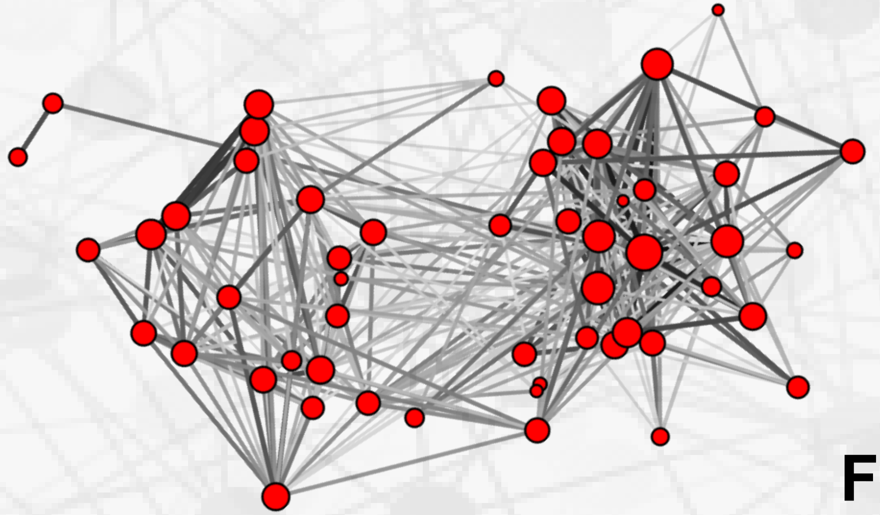








AddHealth Friendship Network, by Grade



Freeman et al. (1988)

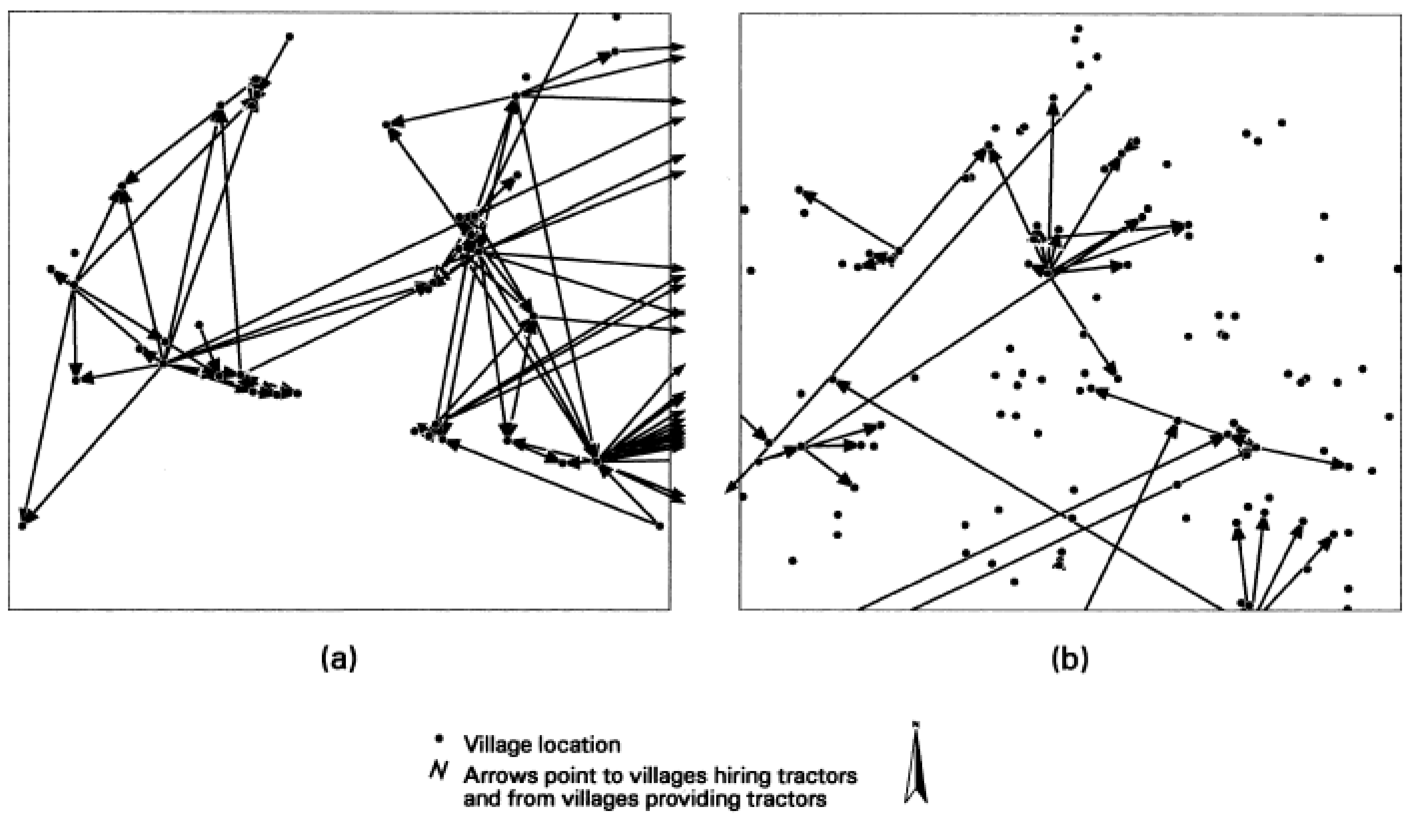
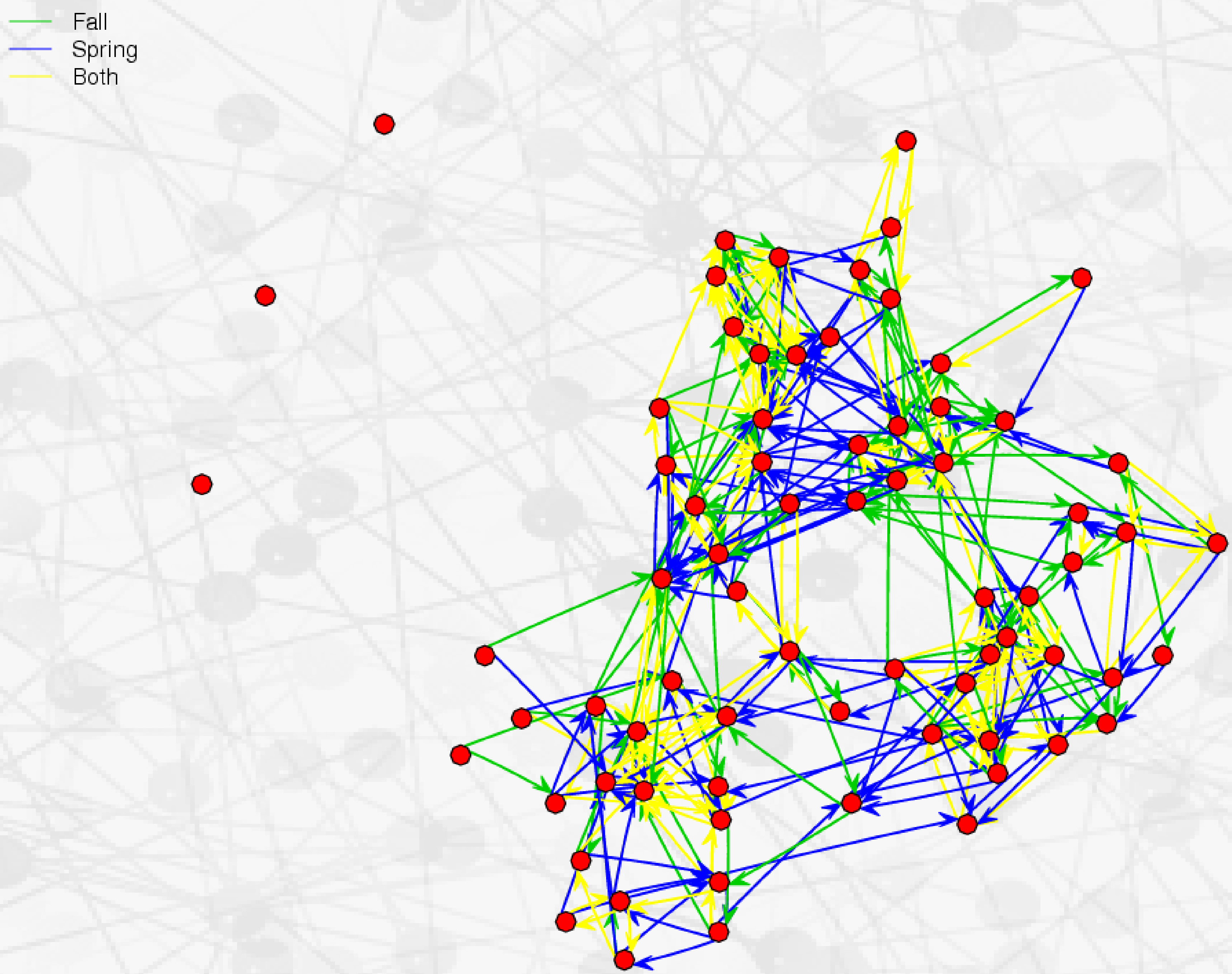


Fig. 3. Tractor hiring network in two regions of Nang Rong.

Faust et al. (1999)



Boy's School Friendship Network (Coleman, 1964)

Logistic Network Regression

- **A classic starting point: why not treat edges as independent, w/log-odds as a linear function of covariates?**
  - Special case of standard logistic regression
  - Dependent variable is a network adjacency matrix
- **Model form:**

$$\log\left(\frac{\Pr(\mathbf{Y_{ij}}=1)}{\Pr(\mathbf{Y_{ij}}=0)}\right)=\theta_1 \mathbf{X_{ij1}}+\theta_2 \mathbf{X_{ij2}}+\ldots+\theta_m \mathbf{X_{ijm}}=\theta^T \mathbf{X_{ij}}$$

  - where $\mathbf{Y}_{ij}$ is the value of the edge from i to j on the dependent relation, $\mathbf{X}_{ijk}$ is the value of the kth predictor for the (i,j) ordered pair, and $\theta_1,\ldots,\theta_m$ are coefficients
    - $\log(p/(1-p))=\text{logit}(p)$, maps $(0,1)$ to $(-\infty,\infty)$

Moving Beyond the Logistic Case

- **The logistic model can be quite powerful, but still very limiting**
  - No way to model conditional dependence among edges
    - E.g., true triad closure bias, reciprocity
  - Cannot handle exotic support constraints
    - What if your network *must* be transitive (e.g., sports contests, entailments), an interval graph (e.g., life history graphs), etc?
- **A more general framework: discrete exponential families**
  - Very general way of representing discrete distributions
  - Turns up frequently in statistics, physics, etc.

Exponential Families for Random Graphs (w/Covariates)

- **For random graph G w/countable support $\mathcal{G}$ and covariate set X, pmf can be written in ERG form:**

$$\Pr\left(G=g \middle| \mathbf{t},\theta,\mathcal{G},X\right)=\frac{\exp\left(\theta^T \mathbf{t}(g,X)\right)}{\sum_{g'\in\mathcal{G}} \exp\left(\theta^T \mathbf{t}(g',X)\right)} I_{\mathcal{G}}(g)$$

  - $\theta^{\mathrm{T}}\mathbf{t}$: linear predictor
    - $\mathbf{t}:\mathcal{G}\rightarrow\mathbb{R}^m$: vector of sufficient statistics
    - $\theta\in\mathbb{R}^m$: vector of parameters
    - $\sum_{g'\in\mathcal{G}}\exp(\theta^{\mathrm{T}}\mathbf{t}(g',X))$: normalizing factor

- **Intuition: ERG placed more/less weight on structures with certain features, as determined by $\mathbf{t}$, θ**
  - Framework is complete for pmfs on $\mathcal{G}$, few constraints on $\mathbf{t}$

Exponential Families for Random Graphs (w/Covariates)

- **For random graph G w/countable support $\mathcal{G}$ and covariate set X, pmf can be written in ERG form:**

$$\Pr\left(G=g \middle| \mathbf{t},\theta,\mathcal{G},X\right)=\frac{\exp\left(\theta^T\mathbf{t}(g,X)\right)}{\sum_{g'\in\mathcal{G}}\exp\left(\theta^T\mathbf{t}(g',X)\right)}I_{\mathcal{G}}(g)$$

  - $\theta^{\mathrm{T}}\mathbf{t}$: linear predictor
    - $\mathbf{t}:\mathcal{G}\rightarrow\mathbb{R}^m$: vector of sufficient statistics
    - $\theta\in\mathbb{R}^m$: vector of parameters
    - $\sum_{g'\in\mathcal{G}}\exp(\theta^{\mathrm{T}}\mathbf{t}(g',X))$: normalizing factor

Source of great difficulty!

- **Intuition: ERG placed more/less weight on structures with certain features, as determined by $\mathbf{t}$, θ**
  - Framework is complete for pmfs on $\mathcal{G}$, few constraints on $\mathbf{t}$

Equivalent Expression: Modeling the Adjacency Matrix

- **For adjacency matrix $\mathbf{Y}$ w/countable support $\mathcal{Y}$ and covariate set X, <u>any</u> pmf can be written in ERG form:**

$$\Pr\left(\mathbf{Y}=\mathbf{y}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)=\frac{\exp\left(\theta^T\mathbf{t}(\mathbf{y},X)\right)}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^T\mathbf{t}(\mathbf{y}',X)\right)}I_{\mathcal{Y}}(\mathbf{y})$$

  - $\theta^{\mathrm{T}}\mathbf{t}$: linear predictor
    - $\mathbf{t}:\mathcal{Y}\rightarrow\mathbb{R}^m$: vector of sufficient statistics
    - $\theta\in\mathbb{R}^m$: vector of parameters
    - $\sum_{\mathbf{y}'\in\mathcal{Y}}\exp(\theta^{\mathrm{T}}\mathbf{t}(\mathbf{y}',X))$: normalizing factor
- **Additional notation:**
  - $\mathbf{y}^{+}_{\mathbf{ij}},\mathbf{y}^{-}_{\mathbf{ij}}$: y w/$ij$th cell set to 1 or 0 (respectively)
  - $\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}$: all elements of y other than the ijth

ERGs and Conditional Odds of an Edge

- **Can easily specify the conditional odds of an edge:**

$$\frac{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{+}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{-}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)} = \frac{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)\right)}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}\frac{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)}$$

ERGs and Conditional Odds of an Edge

- **Can easily specify the conditional odds of an edge:**

$$\frac{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{+}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{-}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)} = \frac{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)\right)}{\cancel{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}}\frac{\cancel{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}}{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)}$$

$$= \frac{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)\right)}{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)} = \exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)-\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)\right)$$

ERGs and Conditional Odds of an Edge

- **Can easily specify the conditional odds of an edge:**

$$\frac{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{+}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{-}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)} = \frac{\exp\left(\theta^{T}\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{+},X\right)\right)}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}\left(\mathbf{y}',X\right)\right)}\frac{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}\left(\mathbf{y}',X\right)\right)}{\exp\left(\theta^{T}\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{-},X\right)\right)}$$

$$= \frac{\exp\left(\theta^{T}\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{+},X\right)\right)}{\exp\left(\theta^{T}\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{-},X\right)\right)} = \exp\left(\theta^{T}\left(\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{+},X\right)-\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{-},X\right)\right)\right)$$

  - Log-odds depend on the changescore, $\Delta_{ij}=\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}},X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}},X)$
  - Useful implication: each unit change in $\mathbf{t}_{\mathbf{k}}$ for (i,j) edge present (versus absent) increases the conditional log-odds of (i,j) by θ_k
    - ***Important***: this is only conditionally true! The marginal log-odds of an (i,j) edge can depend on a complex way on other aspects of the graph

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) = \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right)$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1\,\middle|\,\mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}},\mathbf{t},\theta,\mathcal{Y},X\right) = \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}}\,\middle|\,\mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}},\mathbf{t},\theta,\mathcal{Y},X\right)$$

$$= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}}\,\middle|\,\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}}\,\middle|\,\mathbf{t},\theta,\mathcal{Y},X\right)}\left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}}\,\middle|\,\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}}\,\middle|\,\mathbf{t},\theta,\mathcal{Y},X\right)}\right]^{-1}$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$\Pr\left(\mathbf{Y}_{ij}=1 \,\middle|\, \mathbf{Y}^c_{ij}=\mathbf{y}^c_{ij}, \mathbf{t}, \theta, \mathcal{Y}, X\right) = \Pr\left(\mathbf{Y}=\mathbf{y}^+_{ij} \,\middle|\, \mathbf{Y}^c_{ij}=\mathbf{y}^c_{ij}, \mathbf{t}, \theta, \mathcal{Y}, X\right)$$

$$= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^+_{ij} \,\middle|\, \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^-_{ij} \,\middle|\, \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^+_{ij} \,\middle|\, \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^-_{ij} \,\middle|\, \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1}$$

$$= \frac{\exp\left(\theta^T\left(\mathbf{t}(\mathbf{y}^+_{ij}, X)-\mathbf{t}(\mathbf{y}^-_{ij}, X)\right)\right)}{1+\exp\left(\theta^T\left(\mathbf{t}(\mathbf{y}^+_{ij}, X)-\mathbf{t}(\mathbf{y}^-_{ij}, X)\right)\right)}$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)} \\
&= \left[1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)\right)\right)\right]^{-1}
\end{aligned}
$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}},\mathbf{t},\theta,\mathcal{Y},X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}},\mathbf{t},\theta,\mathcal{Y},X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t},\theta,\mathcal{Y},X\right)}\left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t},\theta,\mathcal{Y},X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}},X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}},X)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}},X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}},X)\right)\right)} \\
&= \left[1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}},X)-\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}},X)\right)\right)\right]^{-1} \\
&= \operatorname{logit}^{-1}\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}},X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}},X)\right)\right)
\end{aligned}
$$

Conditional Edge Probability

- From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)} \\
&= \left[1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)\right)\right)\right]^{-1} \\
&= \operatorname{logit}^{-1}\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)
\end{aligned}
$$

Conditional Edge Probability, Cont.

- **So, the conditional probability of an (i,j) edge is simply the inverse logit of $\theta^T\Delta_{ij}$**
  - Obvious idea: to find θ, why not set this up as a logistic network regression problem (regressing $\mathbf{y}$ on Δ)?
    - This is an "autologistic regression," and the resulting estimator is known as a *pseudolikelihood* estimator (Besag, 1975)
  - Problem: the probability here is only conditional – can use for any one ij, but joint likelihood of $\mathbf{y}$ is not generally the product of $\Pr(\mathbf{Y}_{ij}=\mathbf{y}_{ij}|\mathbf{Y}^c_{ij}=\mathbf{y}^c_{ij})$
    - Another view: $\mathbf{y}$ appears on both sides – can't regress w/out accounting for the "feedback" (i.e., dependence) among edges
    - Does work iff edges are independent – i.e., the logistic case
- **Still, useful aid in interpretation**
  - Can consider probability of $\mathbf{y}_{ij}$ under various scenarios to understand *local* model behavior

Fitting ERGs to Data

- **After positing a model, generally want to estimate parameters from data**
  - Benefit of the framework: standardized inferential framework
- **Most common current method: maximum likelihood estimation**
  - Find θ^* that maximizes $\Pr(\mathbf{Y}=\mathbf{y}_{obs}|\mathbf{t},\theta^*,\mathcal{Y})$
    - Exists for regular model so long as observed data is "non-extreme" (in a specific sense to be discussed later); always unique
    - Also has first order interpretation: $\mathbf{E}_{\theta*}\mathbf{t}(\mathbf{Y})=\mathbf{t}(\mathbf{y}_{obs})$
  - Approximate standard errors based on Hessian of maximized likelihood
    - Not guaranteed, but simulation studies suggest to be reasonable
- **Older method: maximum pseudo-likelihood estimation**
  - Based on autologistic approximation; *<u>don't use</u>* unless you have to – can be very bad (and you can't tell)

ERG Fitting Using `ergm`

- **Dedicated statnet package for fitting, simulating models in ERG form**
- **Basic call structure: `ergm(y~term1(arg)+term2(arg))`**
  - `y`: network object
  - `term1, term2`, etc.: terms to use (see `?"ergm-terms"`)
    - `arg`: where relevant, arguments to the term functions
- **Output: ergm object**
  - Summary, print, and other methods can be used to examine it
  - `simulate` command can also be used to take draws from the fitted model

ERG Parameterization

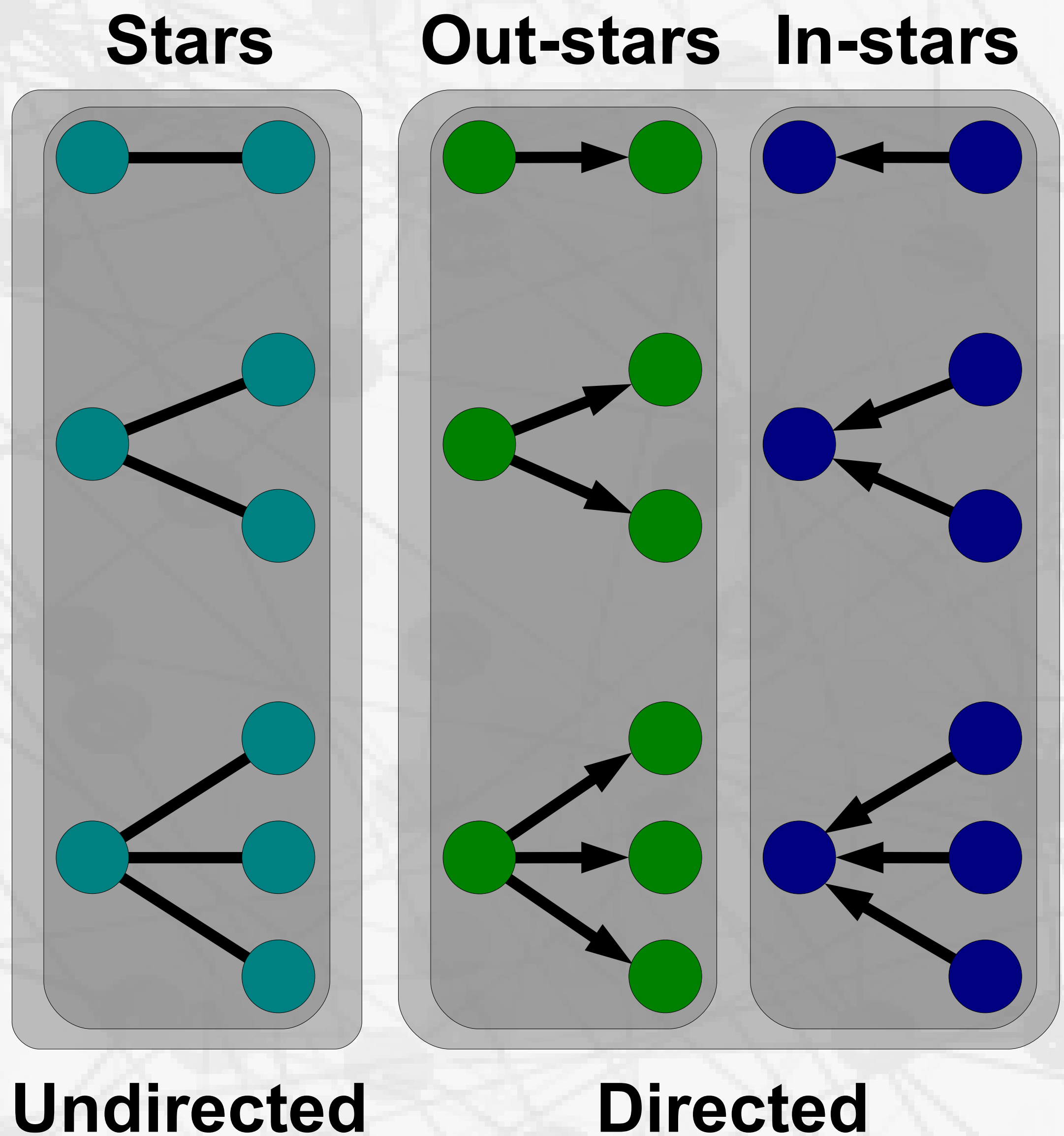
- **ERG form is just a way of writing models – to use it, we must choose a set of terms ($\mathbf{t}$)**
- **Some basics (*dyad independence* terms):**
  - Edge term: $\sum_i \sum_j \mathbf{y}_{ij}$
    - Captures overall tendency of ties to form/not (density effect)
  - Row-sum term: $\sum_i \mathbf{y}_{ij}$
    - Captures net tendency to send ties (sender/expansiveness effect)
  - Col-sum term: $\sum_j \mathbf{y}_{ij}$
    - Captures net tendency to receive ties (receiver/popularity effect)
  - Mutuality term: $\sum_i \sum_{j>i} \mathbf{y}_{ij} \mathbf{y}_{ji}$
    - Captures tendency of ties to reciprocate one another (reciprocity effect)
  - Linear covariates: $\sum_i \sum_j \mathbf{y}_{ij} X_{ij}$
    - Captures tendency of $\mathbf{y}_{ij}$ edges to covary with X_{ij} (covariate effect)

Some Basic Families

- **Several familiar and/or famous model families can be built from the above terms...**
  - N,p: edge term only
  - $U|man$: edge and reciprocity terms
  - `netlogit`: any terms other than mutuality (can rewrite as covariates using dummies)
  - p_1: edge, row-sum, col-sum, and mutuality (or any subset thereof)
    - Can also add covariates; this was done early on for block density (selective mixing) effects
- **Generally flexible, and can be fit using a multinomial logit approach**
  - However, no dependence among dyads allowed...

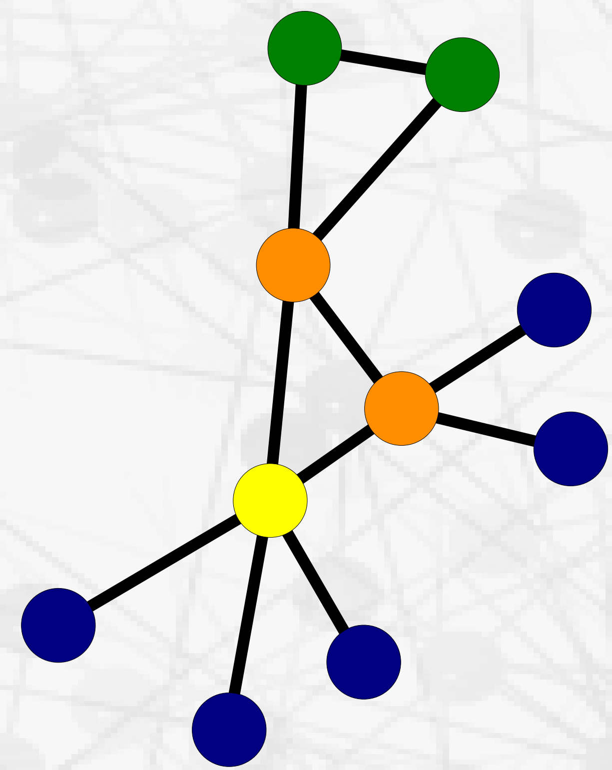
Beyond Independence: the Star Terms

- **Simple subgraph census terms**
  - k-stars: number of subgraphs isomorphic to $K_{1,k}$
  - k-in/out/mixed-stars: number of subgraphs isomorphic to orientations of $K_{1,k}$
- **Interpretations**
  - Tendency of edges to "stick together" on endpoints ("edge clustering")
  - Fixes moments of the degree distribution
    - 1-stars fix mean degree, 2-stars fix variance



Another Way to See Stars: Degree Terms

- **Natural reparameterization of the star terms**
  - *i*th degree term: number of vertices of degree *i*
    - Likewise for indegree, outdegree terms
  - Can be derived from the full set of star terms (and vice versa)



- **Interpretation**
  - Non-parametric model for the degree distribution
  - Note: do not confuse with sender/receiver terms!
    - Latter refer to specific vertices, do not create dependence among edges

$d_0=0,\ d_1=5,\ d_2=2,\ d_3=0,\ d_4=2,\ d_5=1,\ d_6=0,\ d_7=0,\ d_8=0,\ d_9=0$

Triad Census Terms

- **Most basic terms for endogenous clustering**
  - Each term counts subgraphs isomorphic to triads of a given type (i.e., elements of the triad census)
  - In practice, triangles, cycles, and transitives most often used
- **Interpretation**
  - Tendencies towards transitive closure, cycles, etc.
  - Transitivity can be an indicator of latent hierarchy
  - Cyclicity can be an indicator of extended reciprocity