Meta Network Analysis

Weihua An

EMORY UNIVERSITY

Outline

1. Fit random graph models on each network.
   - Exponential random graph model (ERGM)
   - Stochastic actor oriented model (SAOM)
2. Combine the results across networks through meta analysis.
   - Univariate Meta Regression
     - Fixed effects model
     - Random effects model
   - Multivariate Meta Regression
     - Fixed effects model
     - Random effects model

Exponential Random Graph Models

Network analysis focuses on modeling the connections between units.

$$w = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

- Model terms: sender effects, receiver effects, homophily, and endogenous tie formation processes (e.g., reciprocity, transitivity)
- Model estimation
  - One approach is to use logistic regression to model the cells in the matrix. The maximum pseudo-likelihood estimation (MPLE) can account for Markov dependence in tie formation, but not higher-order dependence.

$$\text{logit}\left[P(w_{ij} = 1 | w_{ij}^{L})\right] = \theta_1 + \theta_2 \times \text{Girl}_i + \theta_3 \times \text{Girl}_j + \theta_4 \times \text{Same Gender}_{ij} + \theta_5 \times \text{Reciprocity} + \theta_6 \times \text{Transitivity}$$

Estimation by MCMLE

ERGM assumes the probability of observing a network w is as follows.

$$P(W = w|\theta, X) = \frac{\exp\{\theta' S(w, X)\}}{K(\theta, \mathbf{W})}, \tag{1}$$

where $K(\theta, \mathbf{W}) = \sum_{W \in \mathbf{W}} \exp\{\theta' S(W, X)\}$ is a normalizing factor. $\mathbf{W}$ represents all possible networks that can be formed by the N nodes. The log likelihood is as follows.

$$\ell(\theta) = \theta' S(w, X) - \log K(\theta, \mathbf{W}) \tag{2}$$

The log likelihood cannot be maximized because the normalizing factor is unknown. Suppose one makes a guess of θ_0. One can get the following log likelihood ratio.

$$\ell(\theta) - \ell(\theta_0) = (\theta - \theta_0)' S(w, X) - \log \frac{K(\theta, \mathbf{W})}{K(\theta_0, \mathbf{W})} \tag{3}$$

$$\frac{K(\theta, \mathbf{W})}{K(\theta_0, \mathbf{W})} = \sum_{W \in \mathbf{W}} \frac{\exp\{\theta_0' S(W, X)\}}{K(\theta_0, \mathbf{W})} \frac{\exp\{\theta' S(W, X))\}}{\exp\{\theta_0' S(W, X))\}} = E_{\theta_0} \exp\{(\theta - \theta_0)' S(W, X)\} \tag{4}$$

If one can sample networks $(W_1, W_2, ..., W_m)$ based on θ_0. Then

$$\ell(\theta) - \ell(\theta_0) \approx (\theta - \theta_0)' S(w, X) - \log \left[\frac{1}{m} \sum_{i=1}^{m} \exp\{(\theta - \theta_0)' S(W_i, X)\} \right] \tag{5}$$

Maximizing this equation w.r.t θ leads to the Monte Carlo Maximum Likelihood Estimation (MCMLE) (Hunter and Handcock 2006). That an ERGM is essentially logit $\left[P(w_{ij} = 1|w_{ij}^c)\right]$ indicates a Gibbs sampling scheme for sampling networks.

Example: ERGM

main effects are independent of each other: likely correlated but here not taken into account

Table 1. ERGM Results for the Friendship and Advice-seeking Networks of the Managers

	A. Friendship Network			B. Advising Network		
	Est	SE	P	Est	SE	P
Main Effects						
Age (Receiver Effect)	-0.02	0.03	0.47	0.01	0.02	0.53
Age (Sender Effect)	0.07	0.03	0.02	0.00	0.02	0.77
Tenure	-0.03	0.02	0.10	0.02	0.02	0.30
Dept.2	-0.52	0.25	0.04	0.08	0.25	0.74
Dept.3	-0.17	0.28	0.55	0.67	0.26	0.01
Dept.4	0.26	0.24	0.29	0.24	0.24	0.31
Homophily						
Age Difference	-0.03	0.02	0.17	-0.02	0.02	0.23
Tenure Difference	-0.02	0.02	0.47	-0.04	0.02	0.03
Same Dept	1.26	0.27	0.00	0.58	0.26	0.03
Endogenous Network Formation						
Mutuality	4.61	0.70	0.00	-0.14	0.30	0.64
Edges	-4.20	0.94	0.00	-0.65	0.85	0.44

between network variation

Meta Network Analysis

What if we have fitted random graph models on multiple networks and we want to combine the results across networks? What if we want to understand why the networks exhibit different patterns and how the network patterns are affected by various ecological factors? Meta network analysis helps to combine and compare results of network models.

- Univariate Meta Regression
  - Fixed effects model
  - Random effects model
- Multivariate Meta Regression
  - Fixed effects model
  - Random effects model

Univariate Meta Regression

Dependent variable: estimated coefficients on one variable (e.g., reciprocity).

- Fixed effects model

Same mean = intersect

$$\hat{\theta}_{ki} = \theta_i + \boldsymbol{x_k'}\boldsymbol{\beta_i} + e_{ki} \tag{6}$$

where $\hat{\theta}_{ki}$ denotes the ith estimated coefficient in the kth network, θ_i a common effect to be estimated, $\boldsymbol{x_k}$ the characteristics of the kth network, and e_{ki} an error term with a zero mean and a variance $\hat{\sigma}^2_{ki}$ (which is known). Assuming independence and normality of the error terms:

$$\hat{\theta}_{ki} \sim \text{Normal}\left(\theta_i + \boldsymbol{x_k'}\boldsymbol{\beta_i}, \hat{\sigma}^2_{ki}\right). \tag{7}$$

- Random effects model

$$\hat{\theta}_{ki} \sim \text{Normal}\left(\theta_i + \boldsymbol{x_k'}\boldsymbol{\beta_i}, \hat{\sigma}^2_{ki}\right) \tag{8}$$

$$\theta_i \sim \text{Normal}\left(\mu_i, v_i^2\right) \tag{9}$$

This is like a random intercept model, where μ_i is the mean effect of the ith variable. v_i^2 (to be estimated) measures the between-network variation of the effects. If it equals to zero, the model falls back to the FE model.

Assumption: Independence across coefficients of different variables.

Multivariate Meta Regression

it will comes as a matrix

Dependent variable: coefficients of all (or selected multiple) variables

- Fixed effects model

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{k}} \sim \text{Normal}_{\boldsymbol{l}}\left(\boldsymbol{\theta} + \boldsymbol{X}'_k \beta, \boldsymbol{\Sigma}_{\boldsymbol{k}}\right), \tag{10}$$

  The coefficients in the kth network are assumed to follow a multivariate normal distribution. $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ represents the vector of coefficients in the kth network. $\boldsymbol{\Sigma}_{\boldsymbol{k}}$ is the variance-covariance matrix of the coefficients in the kth network.

between network variation

- Random effects model

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{k}} \sim \text{Normal}_{\boldsymbol{l}}\left(\boldsymbol{\theta} + \boldsymbol{X}'_k \beta, \boldsymbol{\Sigma}_{\boldsymbol{k}}\right), \text{where } \boldsymbol{\theta} \sim \text{Normal}_{\boldsymbol{l}}(\boldsymbol{\mu}, \boldsymbol{\Omega}) \tag{11}$$

  where $\boldsymbol{\theta}$ represents the mean effect and $\boldsymbol{\Omega}$ the between-network covariation. If $\boldsymbol{\Omega}$ equals to zero, the model falls back to the FE model.

These models can account for correlations among the coefficients of different variables, but are difficult to estimate when the number of networks is small.

Model Comparison

- Overall model fitness measure, like Akaike's Information Criterion (AIC) or Bayesian Information Criterion (BIC)
  no difference between estimates cross networks
- Cochran Q test (Gasparrini et al., 2012). A large Q (i.e., a small P-value) indicates significant heterogeneity and random effects models are preferred.
- The I^2 statistic shows the proportion of variation in the coefficients across networks that is attributable to heterogeneity rather than sampling error. A larger I^2 indicates more heterogeneity.

Example

univariate random effect

Table 2. Meta Analysis of the ERGM Results

	A. Friendship			B. Advising			C. Combined			
	Est	SE	P	Est	SE	P	Est	SE	P	Q
Main Effects										
Age (Receiver Effect)	-0.02	0.03	0.47	0.01	0.02	0.53	0.22	0.24	0.36	0.00
Age (Sender Effect)	0.07	0.03	0.02	0.00	0.02	0.77	0.03	0.02	0.17	0.07
Tenure	-0.03	0.02	0.10	0.02	0.02	0.30	0.03	0.06	0.62	0.00
Dept.2	-0.52	0.25	0.04	0.08	0.25	0.74	-0.19	0.27	0.49	0.04
Dept.3	-0.17	0.28	0.55	0.67	0.26	0.01	-0.06	0.26	0.81	0.32
Dept.4	0.26	0.24	0.29	0.24	0.24	0.31	0.27	0.17	0.11	0.94
Homophily										
Age Difference	-0.03	0.02	0.17	-0.02	0.02	0.23	0.07	0.10	0.47	0.00
Tenure Difference	-0.02	0.02	0.47	-0.04	0.02	0.03	0.23	0.24	0.36	0.00
Same Dept	1.26	0.27	0.00	0.58	0.26	0.03	0.73	0.62	0.24	0.05
Endogenous Network Formation										
Mutuality	4.61	0.70	0.00	-0.14	0.30	0.64	2.25	2.30	0.33	0.00
Edges	-4.20	0.94	0.00	-0.65	0.85	0.44	-2.04	2.10	0.33	0.00

Note: The Q column shows the p-value of the Cochran's Q test for effect homogeneity.

Example: SAOM

The Stochastic Actor-Oriented Model (Steglich, Snijders, and Pearson 2010) models the co-evolution of networks and behaviors.

Table 1. SAOM Results for the Network Dynamics in a Selected Class

	Function	Dynamics	OR	95% CI			CT
1	Evaluation	outdegree (density)	0.08	-3.54	0.25		0.08
2	Evaluation	reciprocity	4.85	4.24	14.36	***	0.00
3	Evaluation	transitive triplets	1.48	1.37	4.38	***	-0.01
4	Evaluation	indegree - popularity (sqrt)	0.93	0.42	2.75		-0.08
5	Evaluation	outdegree - popularity (sqrt)	0.66	0.08	1.95		0.00
6	Evaluation	boy alter	1.13	0.81	3.35		0.00
7	Evaluation	same boy	3.58	3.01	10.59	***	-0.02
8	Evaluation	height alter	1.01	1.00	3.00		-0.05
9	Evaluation	height similarity	1.54	0.50	4.55		0.04
10	Evaluation	ranking alter	1.60	1.26	4.73	**	0.03
11	Evaluation	same ranking	1.40	1.08	4.15	*	-0.02
12	Evaluation	personality alter	1.09	0.76	3.22		0.01
13	Evaluation	same personality	0.83	0.55	2.45		0.01
14	Evaluation	family alter	0.92	0.38	2.71		0.00
15	Evaluation	same family	1.63	1.19	4.83	*	-0.01
16	Maintenance	smoking alter	1.47	-1.14	4.34		-0.04
17	Creation	smoking alter	0.19	-7.61	0.56		0.02
18	Maintenance	smoking ego	14.96	11.60	44.27		-0.01
19	Creation	smoking ego	0.02	-4.90	0.06		0.03
20	Maintenance	same smoking	1.15	-1.32	3.42		-0.04
21	Creation	same smoking	0.21	-7.23	0.63		0.06
22	Evaluation	treat alter	0.70	0.24	2.07		0.04
23	Evaluation	same treat	0.95	0.63	2.80		0.01

Note: The shown estimates are odds ratios and their 95% confidence intervals. The last column shows the convergence t-ratios. An absolute value smaller than 0.1 generally indicates a good convergence (Ripley et al., 2014). Significance pattern: ***, P < 0.001; **, P < 0.01; *, P < 0.05.

An (2015) fitted a SAOM on the two-wave friendship network in each of the 76 classes and then used a univariate meta-regression to examine how the network patterns are affected by an external intervention and school indicators.

$$\hat{\theta}_{ki} = \theta_i + \sum_{t!=2} \delta_t \times T_t + \sum_{s!=1} \eta_s \times S_s + e_{ki}$$

Table 2. Results of the Univariate Meta-Regressions for Network Dynamics Related to Smokers

	(1)			(2)			(3)			(4)			(5)			(6)		
	Maintenance Alter			Creation Alter			Maintenance Ego			Creation Ego			Maintenance Same			Creation Same		
	OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI	
Fixed Effects Model																		
Control	0.56	0.28	1.11	1.21	0.65	2.24	0.01	0.00	0.27 **	9.92	0.50	2.E+2	1.67	0.85	3.29	0.75	0.41	1.38
Random Intervention																		
Central Intervention	0.49	0.26	0.91 *	1.18	0.67	2.08	11.48	0.89	1.E+2	0.10	0.01	1.21	1.04	0.56	1.93	1.20	0.71	2.05
Group Intervention	0.53	0.28	0.99 *	1.43	0.85	2.39	2.83	0.31	25.68	0.40	0.05	3.27	1.41	0.76	2.60	0.90	0.56	1.46
Random Effects Model																		
Control	0.55	0.25	1.21	1.21	0.65	2.24	1.08	0.00	3.E+4	9.92	0.50	2.E+2	1.67	0.85	3.29	0.75	0.41	1.38
Random Intervention																		
Central Intervention	0.52	0.25	1.06	1.18	0.67	2.08	2.E+2	0.03	2.E+6	0.10	0.01	1.21	1.04	0.56	1.93	1.20	0.71	2.05
Group Intervention	0.53	0.26	1.09	1.43	0.85	2.39	58.22	0.01	2.E+5	0.40	0.05	3.27	1.41	0.76	2.60	0.90	0.56	1.46
Model Fitness																		
AIC (Fixed Effects)	**206**			**180**			486			458			**189**			**161**		
AIC (Random Effects)	214			191			**440**			**445**			199			173		
Cochran Q-test	0.70			0.83			0.00			0.00			1.00			1.00		
I-square Statistic	0.01			0.01			0.63			0.59			0.01			0.01		
Samples	55			55			55			55			55			55		

Table 4. Results of the Bivariate Meta-Regression

	(1)				(2)		
	Endowment Alter				Creation Alter		
	OR	95% CI			OR	95% CI	
Fixed Effects Models							
Control	0.66	0.38	1.15		1.18	0.73	1.89
Random Intervention							
Central Intervention	0.57	0.33	0.99	*	1.30	0.82	2.04
Group Intervention	0.62	0.36	1.06		1.35	0.86	2.11
Random Effects Models							
Control	0.65	0.34	1.23		1.20	0.72	2.00
Random Intervention							
Central Intervention	0.60	0.32	1.14		1.29	0.79	2.10
Group Intervention	0.61	0.33	1.13		1.37	0.84	2.23
Between-class covariance							
Standard deviations	0.33				0.16		
Between-class correlation	-0.50						
Model Comparison							
AIC (Fixed Effects)		339					
AIC (Random Effects)		357					
Cochran Q-test		0.52					
I-square Statistic		0.01					
Samples	55						