Exponential Random Graph Models in **Statnet**

Lorien Jasny[1]

[1]Q-Step Centre, Exeter University
l.jasny@exeter.ac.uk

ERGM Workshop, Heriot-Watt University, 11/05/2022
materials at
https://github.com/ljasny/HeriotWattWorkshop

Exponential Random Graph Models!

Conceptual Introduction

- Many key questions regarding social systems are ***relational***

Conceptual Introduction

- Many key questions regarding social systems are ***relational***
- All ties are not equal

Conceptual Introduction

- Many key questions regarding social systems are ***relational***
- All ties are not equal
  - Chance of an (i, j) edge may depend on the properties of i and j

Conceptual Introduction

- Many key questions regarding social systems are ***relational***
- All ties are not equal
  - Chance of an (i, j) edge may depend on the properties of i and j
  - Can also depend on other (i, j) relationships

Conceptual Introduction

- Many key questions regarding social systems are ***relational***
- All ties are not equal
  - Chance of an (i, j) edge may depend on the properties of i and j
  - Can also depend on other (i, j) relationships
- The statistical approach

Conceptual Introduction

- Many key questions regarding social systems are ***relational***
- All ties are not equal
  - Chance of an (i, j) edge may depend on the properties of i and j
  - Can also depend on other (i, j) relationships
- The statistical approach
  - Assume that what we see reflects processes with many potential outcomes

Conceptual Introduction

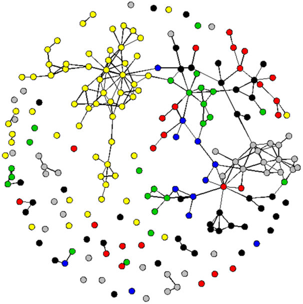
- Many key questions regarding social systems are ***relational***
- All ties are not equal
  - Chance of an (i, j) edge may depend on the properties of i and j
  - Can also depend on other (i, j) relationships
- The statistical approach
  - Assume that what we see reflects processes with many potential outcomes
  - Posit models that reflect our uncertainty about unknowns

Conceptual Introduction

- Many key questions regarding social systems are ***relational***
- All ties are not equal
  - Chance of an (i, j) edge may depend on the properties of i and j
  - Can also depend on other (i, j) relationships
- The statistical approach
  - Assume that what we see reflects processes with many potential outcomes
  - Posit models that reflect our uncertainty about unknowns
  - Reason from observations and prior knowledge to unknown quantities

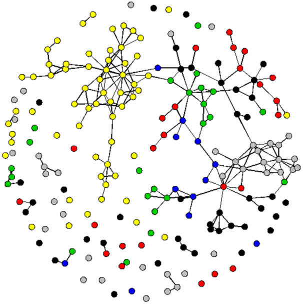
Conceptual Introduction

- Social systems are complex
  - Many parts that affect each other
  - Substantial heterogeneity



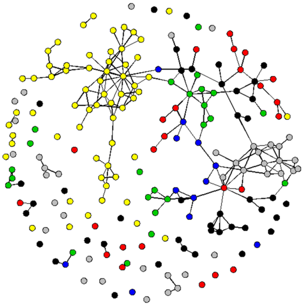
Conceptual Introduction

- Social systems are complex
  - Many parts that affect each other
  - Substantial heterogeneity
- Many mechanisms involved



Conceptual Introduction

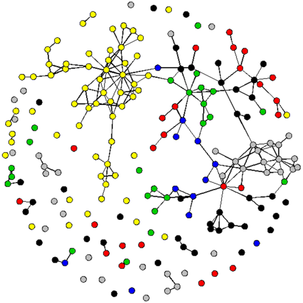
- Social systems are complex
  - Many parts that affect each other
  - Substantial heterogeneity
- Many mechanisms involved
- We're not good at measuring them
  - Usually only see small chunks
  - Error prone observations



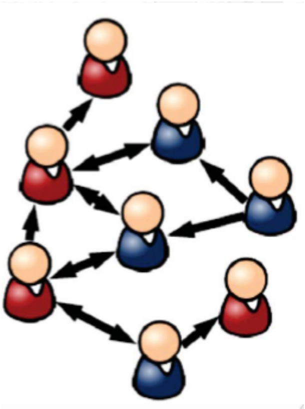
Conceptual Introduction

- Social systems are complex
  - Many parts that affect each other
  - Substantial heterogeneity
- Many mechanisms involved
- We're not good at measuring them
  - Usually only see small chunks
  - Error prone observations

The network we see may result from many mechanisms AND noise AND unobserved factors

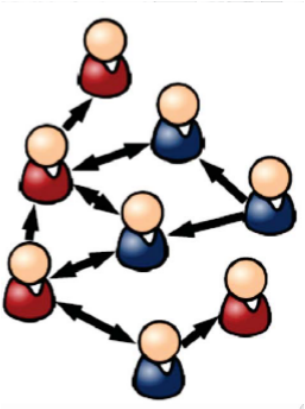


Example: the Reds and Blues



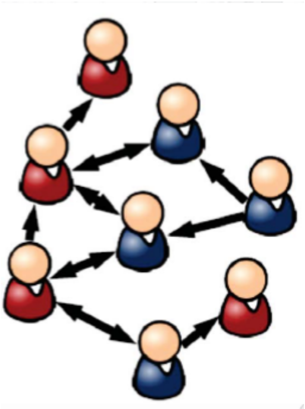
- Consider a hypothetical community with two groups

Example: the Reds and Blues



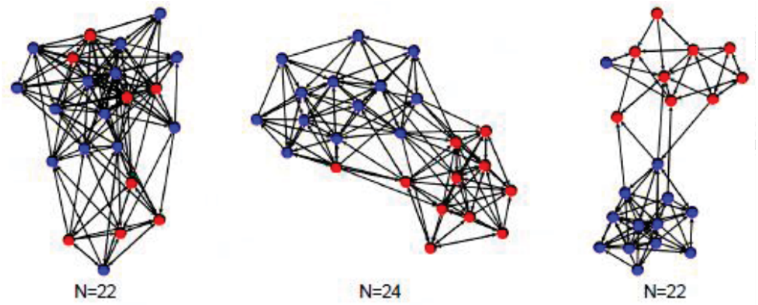
- Consider a hypothetical community with two groups
- We are concerned with cooperation and trust during a period of upheaval

Example: the Reds and Blues

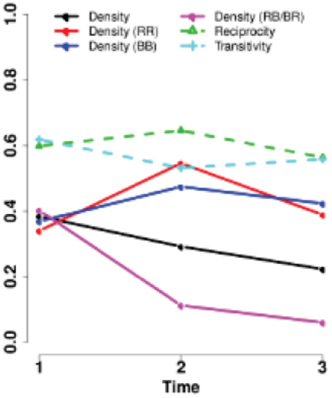


- Consider a hypothetical community with two groups
- We are concerned with cooperation and trust during a period of upheaval
- Our information is limited, but presume that we can observe networks of friendship within representative subgroups...

A Polarization Puzzle

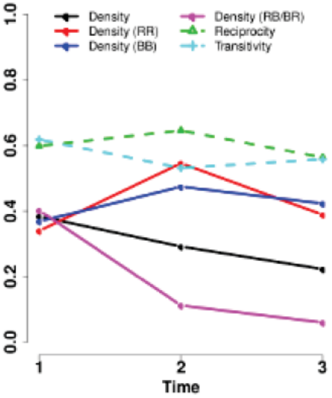


Descriptives



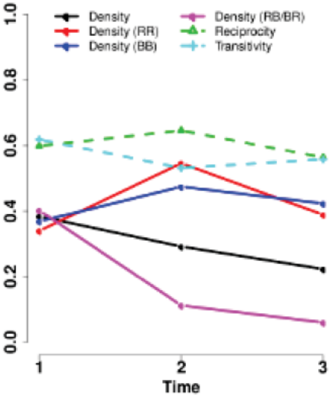
- Without statistical approach, we are limited to description

Descriptives



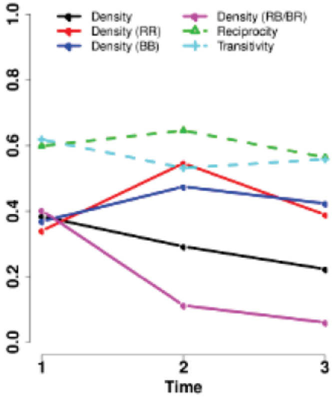
- Without statistical approach, we are limited to description
- Density seems to fall slightly, all this masks an in/out group difference

Descriptives



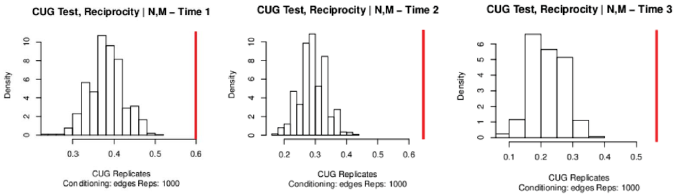
- Without statistical approach, we are limited to description
- Density seems to fall slightly, all this masks an in/out group difference
- Moderately reciprocal, transitive networks with little change

Descriptives

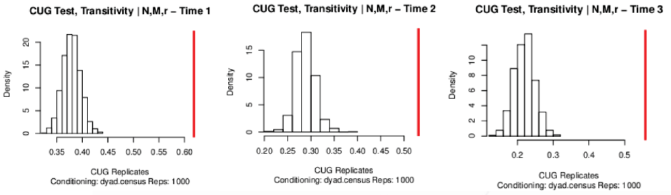


- Without statistical approach, we are limited to description
- Density seems to fall slightly, all this masks an in/out group difference
- Moderately reciprocal, transitive networks with little change
- But, are these changes significant?

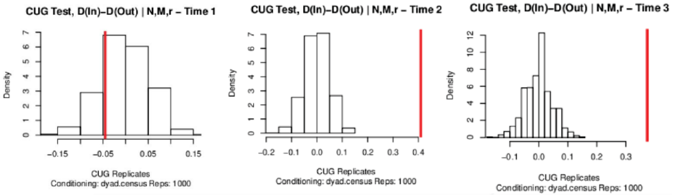
Baseline Models



Baseline Models



Baseline Models



Baseline Models

Baseline models only take us so far

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.
- given “rejection” of the null hypothesis, no clear path forward

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.
- given "rejection" of the null hypothesis, no clear path forward

Solution: Parametric models

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.
- given "rejection" of the null hypothesis, no clear path forward

Solution: Parametric models

- Identify candidate structural mechanisms

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.
- given "rejection" of the null hypothesis, no clear path forward

Solution: Parametric models

- Identify candidate structural mechanisms
- Parameterize using graph statistics

Baseline Models

Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.
- given “rejection” of the null hypothesis, no clear path forward

Solution: Parametric models

- Identify candidate structural mechanisms
- Parameterize using graph statistics
- Fit models to data
  - Compare alternatives
  - Interpret parameter estimates
  - Assess adequacy

Baseline Models

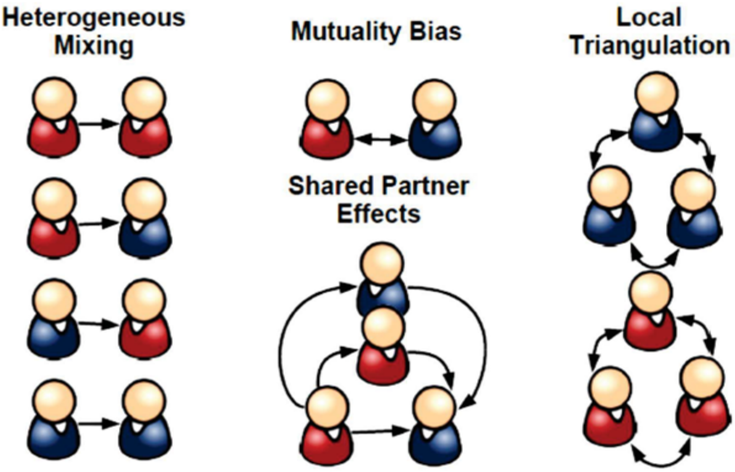
Baseline models only take us so far

- Few statistics lend themselves to conditioning
- Difficult to look at multiple biases at once
- Hard to account for sampling, error, etc.
- given "rejection" of the null hypothesis, no clear path forward

Solution: Parametric models

- Identify candidate structural mechanisms
- Parameterize using graph statistics
- Fit models to data
  - Compare alternatives
  - Interpret parameter estimates
  - Assess adequacy
- Can apply/extend for prediction, etc.

Sample Mechanisms



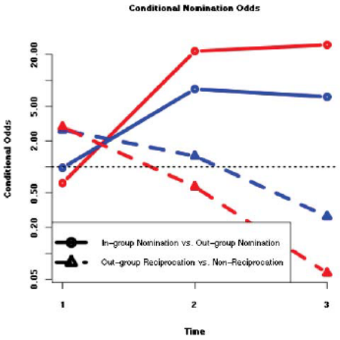
Evaluating Competing Explanations

Edges	Mixing	Mutuals	GWESP	LocalTri	AIC	Rank
1	0	0	0	0	1777.684	15
1	1	0	0	0	1565.073	14
1	0	1	0	0	1516.578	13
1	0	0	1	0	1227.656	2
1	0	0	0	1	1478.532	12
1	1	1	0	0	1428.158	11
1	1	0	1	0	1279.456	6
1	1	0	0	1	1416.441	10
1	0	1	1	0	1234.932	3
1	0	1	0	1	1348.794	9
1	0	0	1	1	1290.241	7
1	1	1	1	0	1216.762	1
1	1	1	0	1	1339.640	8
1	1	0	1	1	1238.285	5
1	0	1	1	1	1236.924	4

Interpreting Mechanisms

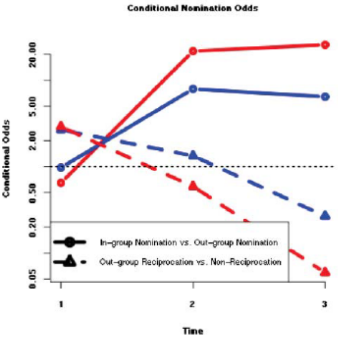
	Time 1 MLE	(SE)	Time 2 MLE	(SE)	Time 3 MLE	(SE)
Red→Red	-1.853	(0.291)	0.557	(0.226)	-1.069	(0.363)
Red→Blue	-1.421	(0.277)	-2.521	(0.428)	-4.317	(0.752)
Blue→Red	-1.501	(0.286)	-1.705	(0.354)	-2.809	(0.417)
Blue→Blue	-1.527	(0.198)	0.364	(0.226)	-0.948	(0.269)
Mutuals	2.484	(0.328)	1.992	(0.335)	1.489	(0.399)
GWESP	-0.030	(0.019)	-0.427	(0.031)	-0.018	(0.104)
GWESP(α)	1.218	(1.248)	0.744	(0.111)	0.598	(6.572)

Interpreting the Mechanisms



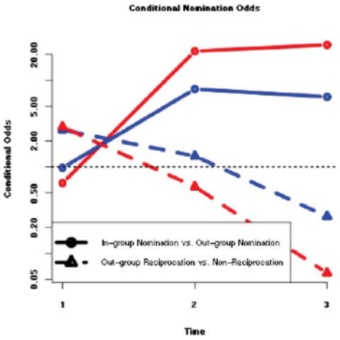
- Sharp decline in out-group nomination propensity with growing numbers of in-group nominations

Interpreting the Mechanisms



- Sharp decline in out-group nomination propensity with growing numbers of in-group nominations
- Decline in mutuality – initially both groups willing to reciprocate, by time 3, neither is!

Interpreting the Mechanisms



- Sharp decline in out-group nomination propensity with growing numbers of in-group nominations
- Decline in mutuality – initially both groups willing to reciprocate, by time 3, neither is!
- our network was actually 3rd, 4th, and 5th grade public school students (Parker and Asher 1993)

Logistic Network Regression

- A classic starting point:

Logistic Network Regression

- A classic starting point:
  - why not treat edges as independent, with log-odds as a linear function of covariates?

Logistic Network Regression

- A classic starting point:
  - why not treat edges as independent, with log-odds as a linear function of covariates?
  - Special case of standard logistic regression

Logistic Network Regression

- A classic starting point:
  - why not treat edges as independent, with log-odds as a linear function of covariates?
  - Special case of standard logistic regression
  - Dependent variable is a network adjacency matrix
- Model form:
  $log(\frac{P(Y_{ij}=1)}{P(Y_{ij}=0)} = \theta_1 X_{ij1} + \theta_2 X_{ij2} + ... + \theta_m X_{ijm} = \theta^T X_{ij})$
- Where Y_{ij} is the value of the edge from i to j on the dependent relation, X_{ijk} is the value of the kth predictor for the (i, j) ordered pair, and $\theta_1 ... \theta_m$ are coefficients

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting
    - No way to model conditional dependence among edges (clustering, reciprocity)

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting
  - No way to model conditional dependence among edges (clustering, reciprocity)
  - Cannot handle exotic support constraints

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting
  - No way to model conditional dependence among edges (clustering, reciprocity)
  - Cannot handle exotic support constraints
- A more general framework: discrete exponential families

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting
  - No way to model conditional dependence among edges (clustering, reciprocity)
  - Cannot handle exotic support constraints
- A more general framework: discrete exponential families
  - Very general way of representing discrete distributions

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting
  - No way to model conditional dependence among edges (clustering, reciprocity)
  - Cannot handle exotic support constraints
- A more general framework: discrete exponential families
  - Very general way of representing discrete distributions
  - Turns up frequently in statistics, physics, etc.

Beyond the Logistic Case

- The logistic model can be quite powerful, but still very limiting
  - No way to model conditional dependence among edges (clustering, reciprocity)
  - Cannot handle exotic support constraints
- A more general framework: discrete exponential families
  - Very general way of representing discrete distributions
  - Turns up frequently in statistics, physics, etc.
  - ERGM is more like a language of models than a specific book

Exponential Random Graph Models

$$P(Y = y|t, \theta, Y, X) = \frac{exp(\theta^T t(y,X))}{\sum y' \in Y exp(\theta^T t(y',X))} I_Y(y)$$

Exponential Random Graph Models

$$P(Y = y|t, \theta, Y, X) = \frac{exp(\theta^T t(y,X))}{\sum y' \in Y exp(\theta^T t(y',X))} I_Y(y)$$

$$\uparrow$$

Probability that a random graph drawn from Y is the realized graph y

Exponential Random Graph Models

Given sufficient statistics t, the parameters θ, the countable support Y, and the covariates X

$$P(Y = y|t, \theta, Y, X) = \frac{exp(\theta^T t(y,X))}{\sum y' \in Y exp(\theta^T t(y',X))} I_Y(y)$$

Probability that a random graph drawn from Y is the realized graph y

Exponential Random Graph Models

Given sufficient statistics t, the parameters θ, the countable support Y, and the covariates X

The empirical realization of covariates, statistics, and parameters

$$P(Y = y|t, \theta, Y, X) = \frac{exp(\theta^T t(y,X))}{\sum y' \in Y exp(\theta^T t(y',X))} I_Y(y)$$

Probability that a random graph drawn from Y is the realized graph y

Exponential Random Graph Models

Given sufficient statistics t, the parameters θ, the countable support Y, and the covariates X

The empirical realization of covariates, statistics, and parameters

$$P(Y = y|t, \theta, Y, X) = \frac{exp(\theta^T t(y,X))}{\sum y' \in Y exp(\theta^T t(y',X))} I_Y(y)$$

Probability that a random graph drawn from Y is the realized graph y

Normalizing factor counting over every other graph in the support

Exponential Random Graph Models

Given sufficient statistics t, the parameters θ, the countable support Y, and the covariates X

The empirical realization of covariates, statistics, and parameters

$$P(Y = y|t, \theta, Y, X) = \frac{exp(\theta^T t(y,X))}{\sum y' \in Y exp(\theta^T t(y',X))} I_Y(y)$$

Probability that a random graph drawn from Y is the realized graph y

Normalizing factor counting over every other graph in the support

An indicator that Y is in the support

Conditional Log-Odds of an Edge

$$\frac{P(Y{=}y_{ij}^{+}|t,\theta,Y,X)}{P(Y{=}y_{ij}^{-}|t,\theta,Y,X)}$$

Conditional Log-Odds of an Edge

$$\frac{P(Y{=}y_{ij}^{+}|t,\theta,Y,X)}{P(Y{=}y_{ij}^{-}|t,\theta,Y,X)} = \frac{exp(\theta^T t(y_{ij}^{+},X))}{\sum y'{\in}Y\, exp(\theta^T t(y',X))} * \frac{\sum y'{\in}Y\, exp(\theta^T t(y',X))}{exp(\theta^T t(y_{ij}^{-},X))}$$

Conditional Log-Odds of an Edge

$$\frac{P(Y{=}y_{ij}^{+}|t,\theta,Y,X)}{P(Y{=}y_{ij}^{-}|t,\theta,Y,X)} = \frac{exp(\theta^T t(y_{ij}^{+},X))}{\sum y'\in Y exp(\theta^T t(y',X))} * \frac{\sum y'\in Y exp(\theta^T t(y',X))}{exp(\theta^T t(y_{ij}^{-},X))}$$

$$\frac{exp(\theta^T t(y_{ij}^{+},X))}{exp(\theta^T t(y_{ij}^{-},X))} = exp(\theta^T(t(y_{ij}^{+},X) - t(y_{ij}^{-},X)))$$

Conditional Log-Odds of an Edge

$$\frac{P(Y=y_{ij}^{+}|t,\theta,Y,X)}{P(Y=y_{ij}^{-}|t,\theta,Y,X)} = \frac{exp(\theta^T t(y_{ij}^{+},X))}{\sum y' \in Y exp(\theta^T t(y',X))} * \frac{\sum y' \in Y exp(\theta^T t(y',X))}{exp(\theta^T t(y_{ij}^{-},X))}$$

$$\frac{exp(\theta^T t(y_{ij}^{+},X))}{exp(\theta^T t(y_{ij}^{-},X))} = exp(\theta^T(t(y_{ij}^{+},X) - t(y_{ij}^{-},X)))$$

$$= \frac{P \; |\text{the rest of the graph}}{P \; |\text{the rest of the graph}} = exp(\theta^T * \Delta\text{change score})$$

ERG Fitting using `ergm`

- Dedicated statnet package for fitting, simulating models in ERG form

ERG Fitting using `ergm`

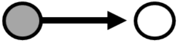
- Dedicated statnet package for fitting, simulating models in ERG form
- Basic call structure: `ergm(y~term1(arg)+term2(arg))`
  - y is a network
  - term1, term2, etc are the “sufficient statistics”, or terms written in the `ergm` package
  - see “ergm-terms”

ERG Fitting using `ergm`

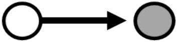
- Dedicated statnet package for fitting, simulating models in ERG form
- Basic call structure: `ergm(y~term1(arg)+term2(arg))`
  - y is a network
  - term1, term2, etc are the "sufficient statistics", or terms written in the `ergm` package
  - see "ergm-terms"
- Output: ergm object
  - Summary, print and other methods can be used to examine it
  - Simulate command can also be used to take draws from the fitted model

Dyadic independent terms

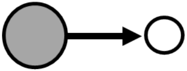
Edge – the baseline probability of a tie



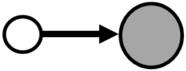
Outdegree (Sender) for an attribute



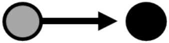
Indegree (Receiver) for an attribute



Outdegree (Sender) for a valued parameter

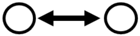


Outdegree (Sender) for a valued parameter

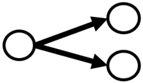


Mixing terms

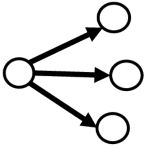
Dyadic dependent terms



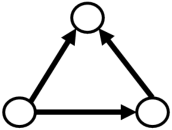
Reciprocity



Out 2-star
(popularity)

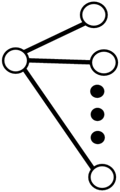


Out 3-star
(more
popularity)

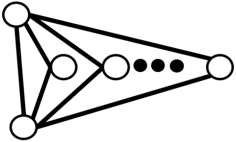


Transitive
Triad

Higher Order Terms



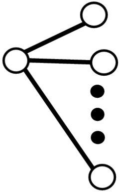
Geometrically Weighted Stars (altkstar or gwdegree)



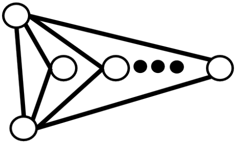
Geometrically Weighted Edgewise Shared Partners (gwesp)

- Diminishing returns makes sense (every three-star has 3 two-stars)

Higher Order Terms



Geometrically Weighted Stars (altkstar or gwdegree)



Geometrically Weighted Edgewise Shared Partners (gwesp)

- Diminishing returns makes sense (every three-star has 3 two-stars)
- Makes fitting the MCMC much easier - we'll see why next...

Interpreting Coefficients

```
Formula:   samplk3 ~ edges + mutual

Iterations:  2 out of 20

Monte Carlo MLE Results:
       Estimate Std. Error MCMC % p-value
edges   -2.1505     0.2181      0  <1e-04 ***
mutual   2.2879     0.4782      0  <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     Null Deviance: 424.2  on 306  degrees of freedom
 Residual Deviance: 267.9  on 304  degrees of freedom

AIC: 271.9    BIC: 279.3    (Smaller is better.)
```

- The log-odds of an unreciprocated edge is -2.15

Interpreting Coefficients

```
Formula:   samplk3 ~ edges + mutual

Iterations:  2 out of 20

Monte Carlo MLE Results:
       Estimate Std. Error MCMC % p-value
edges   -2.1505     0.2181      0  <1e-04 ***
mutual   2.2879     0.4782      0  <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     Null Deviance: 424.2  on 306  degrees of freedom
 Residual Deviance: 267.9  on 304  degrees of freedom

AIC: 271.9    BIC: 279.3    (Smaller is better.)
```

- The log-odds of an unreciprocated edge is -2.15
- The probability of an unreciprocated edge is $\frac{exp(-2.15)}{1+exp(-2.15)} = 0.10$

Interpreting Coefficients

```
Formula:   samplk3 ~ edges + mutual

Iterations:  2 out of 20

Monte Carlo MLE Results:
       Estimate Std. Error MCMC % p-value
edges   -2.1505     0.2181      0  <1e-04 ***
mutual   2.2879     0.4782      0  <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     Null Deviance: 424.2  on 306  degrees of freedom
 Residual Deviance: 267.9  on 304  degrees of freedom

AIC: 271.9    BIC: 279.3    (Smaller is better.)
```

- The log-odds of an reciprocated edge is -2.15+2.29=.14

Interpreting Coefficients

```
Formula:   sampk3 ~ edges + mutual

Iterations:  2 out of 20

Monte Carlo MLE Results:
       Estimate Std. Error MCMC % p-value
edges   -2.1505     0.2181      0  <1e-04 ***
mutual   2.2879     0.4782      0  <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     Null Deviance: 424.2  on 306  degrees of freedom
 Residual Deviance: 267.9  on 304  degrees of freedom

AIC: 271.9    BIC: 279.3    (Smaller is better.)
```

- The log-odds of an reciprocated edge is -2.15+2.29=.14
- The probability of an reciprocated edge is $\frac{exp(.14)}{1+exp(.14)} = 0.53$

Model Fit and Model Assessment

- We've seen how to construct and fit nontrivial ERGs
  - Started with dyadic independent terms
  - Added basic dependence terms
  - Fit the whole thing via MLE
- Now we turn to degeneracy checking and model assessment
  - Looking under the hood to make sure that the engine is still running - and occasionally, getting out to turn the crank
  - Checking the results to make sure that the model makes sense

The role of Simulation in ERG Research

- Simulation is central to ERG modeling
  - Even simple models too complex to get analytical solutions - need to use simulation to study model behaviour, make predictions
  - ERG computations too difficult to perform directly (that support term in the denominator) - simulation used purely for computational purposes
- Implication: we need to know something about ERG simulation to use tools effectively

Mc, MC, and MCMC in one slide

- Markov chain
  - Stochastic process such that $P(X_i|X_{i-1}, X_{i-2}, ...) = P(X_i|X_{i-1})$

Mc, MC, and MCMC in one slide

- Markov chain
  - Stochastic process such that $P(X_i|X_{i-1}, X_{i-2}, ...) = P(X_i|X_{i-1})$
  - meaning that only the previous state (i-1) matters in predicting the current state (i)

Mc, MC, and MCMC in one slide

- Markov chain
  - Stochastic process such that $P(X_i|X_{i-1}, X_{i-2}, ...) = P(X_i|X_{i-1})$
  - meaning that only the previous state (i-1) matters in predicting the current state (i)
- Monte Carlo procedure
  - Any procedure which uses randomization to perform computation, having a fixed execution time and uncertain output

Mc, MC, and MCMC in one slide

- Markov chain
  - Stochastic process such that $P(X_i|X_{i-1}, X_{i-2}, ...) = P(X_i|X_{i-1})$
  - meaning that only the previous state (i-1) matters in predicting the current state (i)
- Monte Carlo procedure
  - Any procedure which uses randomization to perform computation, having a fixed execution time and uncertain output
- Markov chain Monte Carlo (MCMC)
  - Family of procedures using Markov chains to perform computations and/or simulate target distributions

ERG MCMC

- When we need to simulate ERGs, we turn to MCMC
  - Every 'step' in the Markov chain is changing one edge from on (1) to off (0) or vice versa
  - Then, the probability of the next step given the current state of the chain is the change score we saw before
  - General procedure: start with a 'seed' graph (random or data)
    - Early "burn-in" draws contaminated by an initial state - discard
    - need to ensure that sample is large enough to have good properties
    - both aspects sloppily called "convergence" – the chain has "converged" when approximation is adequate
  - mostly automated, but important to use diagnostics to verify behavior

What happens when you run `ergm`

What happens when you run `ergm`

- Little gnomes make an initial guess at θ using the MPLE

What happens when you run `ergm`

- Little gnomes make an initial guess at θ using the MPLE
- More gnomes simulate $y_1, ...y_n$ based on initial guess

What happens when you run `ergm`

- Little gnomes make an initial guess at θ using the MPLE
- More gnomes simulate $y_1, ...y_n$ based on initial guess
- This simulated sample is used to find θ using MLE

What happens when you run `ergm`

- Little gnomes make an initial guess at θ using the MPLE
- More gnomes simulate $y_1, ...y_n$ based on initial guess
- This simulated sample is used to find θ using MLE
- Possibly, the previous two steps are iterated a few times for good measure (since initial estimate may be off)

A Puzzle

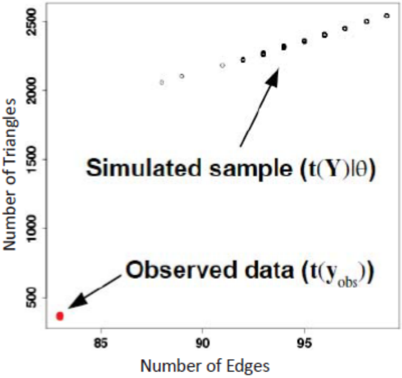
Lots of interest early on in a very (at first glance) simple model:

```
ergm(net~edges+triangle)
```

But some puzzling results when we simulated from the model

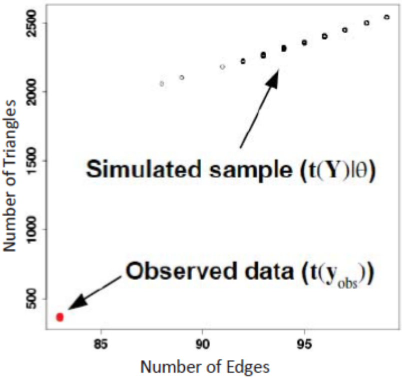
A Puzzle

- The simulated networks look nothing like the observed data
- Even when the correct coefficients are not simulated (was done on an example with 7 nodes) the networks simulated from that model show the same result (Ke Li, 2015)



Model Degeneracy

- Almost all the graphs are the same (usually complete/empty)
- The probability of a given statistic pushes the MCMC to always/never add edges



More Broadly

- Simulation can fail in several (essentially four) ways
  - Insufficient burn-in - starting point still affects results
  - Insufficient post-burn samples - sample hasn't converged
  - degeneracy
  - Sample does not cover observed graph - you couldn't generate your given graph from any combination of sufficient statistics

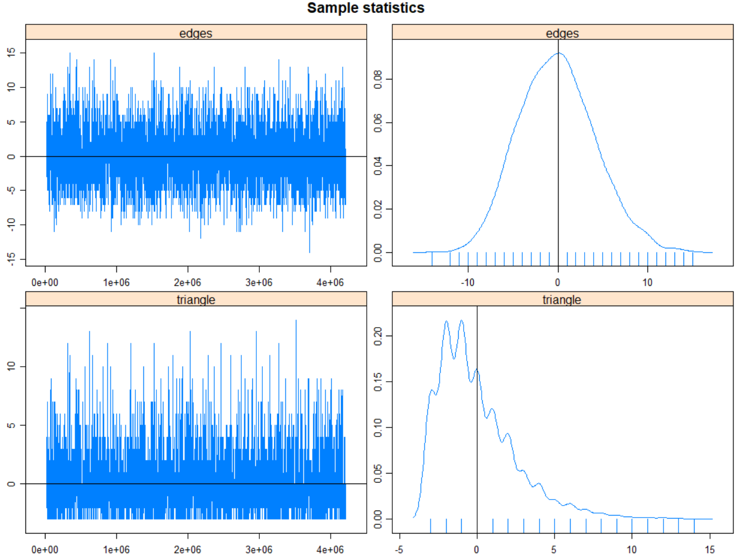
Assessing Simulation Quality

- No foolproof method, but several heuristics
- in `ergm`, primary tool is `mcmc.diagnostics`
- calculates various diagnostics on MCMC output
- Can also directly plot statistics (from the MCMC) vs observed values

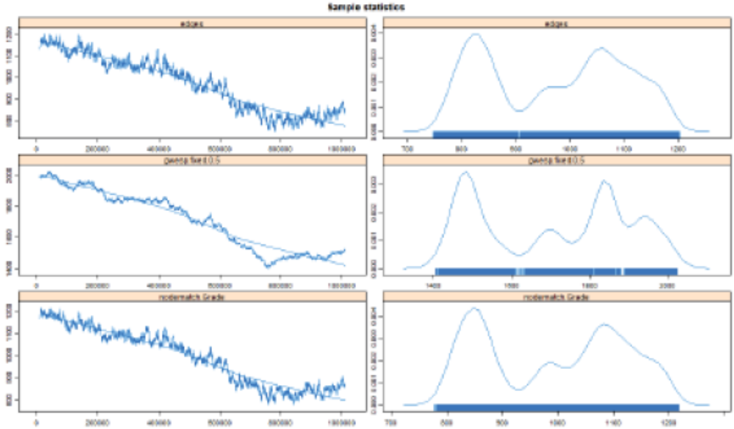
What if things go wrong?

- Different MCMC controls are set using the sequence `control=control.ergm(`*terms*`)`
- For burn-in issues, increase MCMC.burnin parameter
- For post-burn convergence, increase MCMC.samplesize
- If none of these work, may need to change the model

Diagnostics



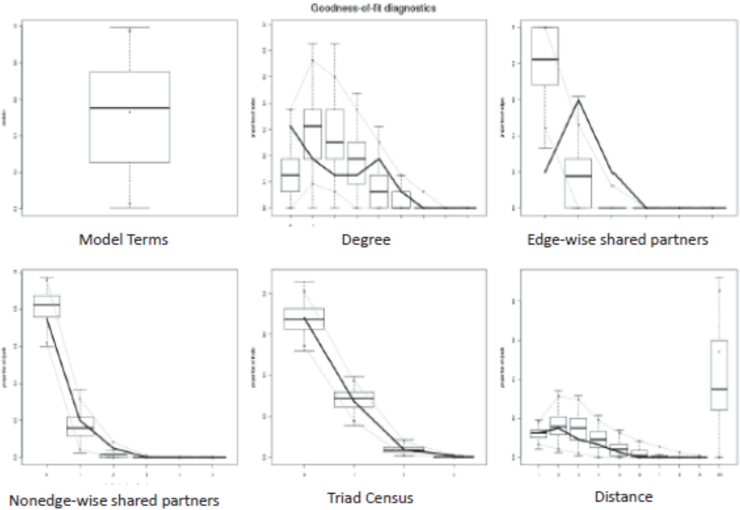
Diagnostics



Assessing Adequacy

- How does one assess model adequacy? Simulation!
  - Simulate draws from fitted model
  - Compare observed graph to simulated graphs on measures of interest
  - Verify that observed properties are well-covered by simulated ones (e.g. not in 5% tails)
- What properties should be considered?
  - This is application-specific - no single uniform answer
  - Start with "in-model" statistics - ERG must get means right, but should still verify non-pathological distributions (remember the triangles)
  - "out-of-model" statistics can be common low-level properties (e.g. degree, triad census) or theoretically motivated quantiles

Example - a model only with edges



What if model is inadequate?

- Option 1: add terms
    - Which features are poorly captured? Is there a term which would add in such effects?

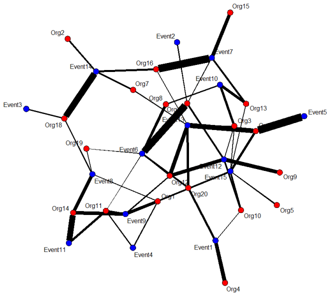
What if model is inadequate?

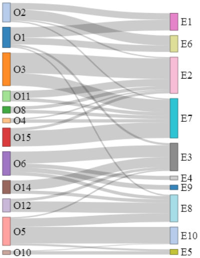
- Option 1: add terms
  - Which features are poorly captured? Is there a term which would add in such effects?
- Option 2: switch terms

What if model is inadequate?

- Option 1: add terms
  - Which features are poorly captured? Is there a term which would add in such effects?
- Option 2: switch terms
- Option 3: do nothing
  - Is the type of inadequacy a problem for your specific question? Can it be tolerated in this case? How good is the overall fit?

Bipartite Data

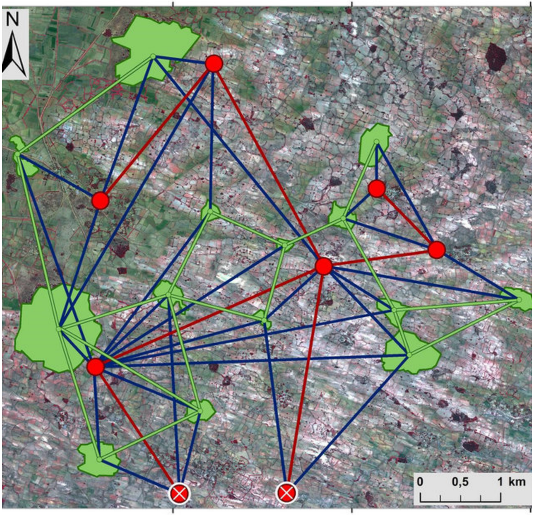




Bipartite Data

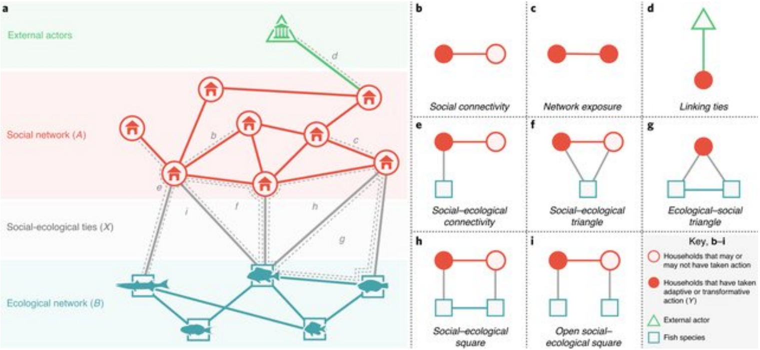
- Many terms already written
- look for B1 or B2 in the term description (1 is first mode, 2 is 2nd mode)
- Simulated networks will not have within-mode ties

Multi-level Data



Bodin, Örjan, and Maria Tengö. "Disentangling intangible social–ecological systems." Global Environmental Change 22.2 (2012): 430-439.

Multi-level Data



Barnes, Michele L., et al. “Social determinants of adaptive and transformative responses to climate change.” Nature Climate Change 10.9 (2020): 823-828.

Multi-level Data

- Tons of confusion over the term ‘multi-level’

Multi-level Data

- Tons of confusion over the term ‘multi-level’
- New functionality in Statnet to write these terms

Multi-level Data

- Tons of confusion over the term ‘multi-level’
- New functionality in Statnet to write these terms
- Essentially, treats social and ecological parts as an attribute and runs a normal ERGM

Multi-level Data

- Tons of confusion over the term ‘multi-level’
- New functionality in Statnet to write these terms
- Essentially, treats social and ecological parts as an attribute and runs a normal ERGM
- We’ll see an example using **F** and **Sum**

Multi-level Data

- Tons of confusion over the term ‘multi-level’
- New functionality in Statnet to write these terms
- Essentially, treats social and ecological parts as an attribute and runs a normal ERGM
- We’ll see an example using **F** and **Sum**
- ...but it’s complicated

One last trick

- Say a term is theoretically very important

One last trick

- Say a term is theoretically very important
- But the term hasn't been written

One last trick

- Say a term is theoretically very important
- But the term hasn't been written
  - One solution - write your own term (ergm-terms package)

One last trick

- Say a term is theoretically very important
- But the term hasn't been written
    - One solution - write your own term (ergm-terms package)
- There is a term, but you can't get it to fit

One last trick

- Say a term is theoretically very important
- But the term hasn't been written
  - One solution - write your own term (ergm-terms package)
- There is a term, but you can't get it to fit
  - Simulate a model with lower order parameters (and not your term of interest)

One last trick

- Say a term is theoretically very important
- But the term hasn't been written
  - One solution - write your own term (ergm-terms package)
- There is a term, but you can't get it to fit
  - Simulate a model with lower order parameters (and not your term of interest)
  - Use the goodness-of-fit method to see how extreme your parameter of interest is in your empirical data compared to a sample/simulation from this model

Code Time!

- The rest! Whew!