Network Modeling of Infectious Disease and Social Diffusion Processes with EpiModel

Sunbelt Conference Workshop
June 23, 2015
Samuel M. Jenness, PhD MPH
University of Washington
Department of Epidemiology
Workshop Materials

http://statnet.github.io/sb/
Acknowledgements

EpiModel Authors
Samuel Jenness
Steven Goodreau
Martina Morris

EpiModel Contributors
Li Wang
Emily Beylerian

EpiModel Users!

Statnet Development Team
Skye Bender-deMoll
Carter Butts
Mark Handcock
David Hunter
Pavel Krivitsky

Funding
R00 HD057533 (NICHD)
R01 HD68395 (NICHD)
T32 HD007543 (NICHD)
R24 HD042828 (NICHD)
EpiModel

- EpiModel is an R software package
- Tools for simulation and analysis of epidemic models
- Supports three model classes
  - Deterministic compartmental models
  - Stochastic individual contact models
  - Stochastic network models
- \texttt{http://epimodel.org/}
EpiModel

• EpiModel is an R software package
• Tools for simulation and analysis of epidemic models
• Supports three model classes
  - Deterministic compartmental models
  - Stochastic individual contact models
  - Stochastic network models
• [http://epimodel.org/](http://epimodel.org/)
Workshop Goals

- Introduce **dynamic modeling** over networks
  - Also called mathematical models or systems models
  - Contrast with purely statistical models

- Provide hands-on experience using **EpiModel** software
  - Estimating statistical models for dynamic networks with temporal ERGMs
  - Simulating infectious disease or social phenomenon on top of dynamic networks

- Explain methods to extend EpiModel for your **research**
Statistical Models

• Start with **data**
• Choose functional framework for **summarizing** data
• Fit **model** to estimate parameters
• Infer population **associations** or casual **effects**
Mathematical Models

- Start with the **parameters**
- Construct the **processes** to get from micro to macro
  - **Micro**: Individual-level biology, behavior, demography
  - **Macro**: Population-level disease incidence and prevalence

Dynamic = Over Time
Statistical vs Mathematical Models

The Epidemic Feedback Loop

Force of Infection

(Rate of contacts) •

(Transmission probability per contact) •

(Probability contacting an infected)
Statistical vs Mathematical Models

Indirect Effects & Herd Immunity

Vaccination coverage (percent)
Statistical vs Mathematical Models
**Statistical vs Mathematical Models**

**Stochastic Network Models**

- Collect *egocentric* network data
- Fit a temporal ERGM with target statistics
- Simulate from that statistical model fit
- Construct the other epidemic or diffusion processes over network

---

Data $\downarrow$

Statistical Model $\Updownarrow$

Mathematical Model
Introductions

Please *briefly* introduce yourself

- Name, department, institution
- Exposure to and experience with:
  - Statnet (sna, network, networkDynamic) for network analysis
  - ERGMs and TERGMs for network modeling
  - Dynamic/mathematical models
  - The R programming language
- Related research project or interest
Workshop Outline

1. Lecture Introduction
2. Lecture From network data to temporal ERGMs
3. Tutorial An SIS epidemic in a closed population
4. Lecture Considerations for open populations
5. Tutorial An SI epidemic in an open population
6. Lab Adding heterogeneity & interventions
7. Lecture Extending EpiModel for novel research
From Survey Data to Network Data

• EpiModel depends on egocentric network data
  - Random sample of population
  - Subjects queried on history of recent (sexual) partnerships
    - Date of first and last contact, whether ongoing, with last three partners
  - Subjects queried on attributes of those partnerships

• Summary statistics from survey data \(\rightarrow\) simulation of complete network consistent with those statistics
  - Fit ERGM with target statistics, simulate from that model fit

• A scalable, flexible, data generating model for dynamic networks
Network Model Parameters

Degree
Network Model Parameters

Assortative Mixing
Network Model Parameters

Dissortative Mixing
Network Model Parameters

Mixing on Multiple Levels
1. Start with Target Population
Egocentric Inference

2. Sample Egos

- 

- 

-
Egocentric Inference

3. Query on Alters
Egocentric Inference

4. Estimate Target Statistics

<table>
<thead>
<tr>
<th>Stat</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td># Edges</td>
<td>4</td>
</tr>
<tr>
<td># Isolate nodes</td>
<td>1</td>
</tr>
<tr>
<td># Concurrent nodes</td>
<td>1</td>
</tr>
<tr>
<td>Age homophily</td>
<td>1 year</td>
</tr>
<tr>
<td>Shape homophily</td>
<td>2</td>
</tr>
<tr>
<td>Color homophily</td>
<td>0</td>
</tr>
</tbody>
</table>
5. Fit an ERGM with Target Statistics

**Graph probability**

\[
\Pr(Y = y) = \frac{\exp \{ \theta' g(y) \}}{\mathcal{K}(\theta)}
\]

- \( g(y) = \) vector of network statistics
- \( \theta = \) vector of coefficients
- \( \mathcal{K} = \) normalizing constant
5. Fit an ERGM with Target Statistics

Graph probability

\[
Pr(Y = y) = \frac{\exp\{\theta' g(y)\}}{\mathcal{K}(\theta)}
\]

\(g(y)\) = vector of network statistics
\(\theta\) = vector of coefficients
\(\mathcal{K}\) = normalizing constant

formula = nw ~ edges +
  isolates +
  concurrent +
  absdiff("age") +
  nodematch("shape") +
  nodematch("color")

targets = c(4, 1, 1, 4*1, 2, 0)*100
Egocentric Inference

6. Simulate from the Model

- MCMC-based simulation similar to that used in estimation
- Simulations generate one network cross-section
- Summary of network stats consistent on average with targets
Egocentric Inference

7. Add Time!
Egocentric Inference

7. Add Time!

- Incidence = prevalence / duration
- STERGMs fit two ERGMs
  - One for formation and one for persistence of edges
- Mean edge duration as fixed (offset) coefficient
- Default EpiModel method bypasses full STERGM
  - Uses cross-sectional ERGM with manual coefficient adjustment
  - The “Edges Dissolution Approximation”
1. Lecture  Introduction
2. Lecture  From network data to temporal ERGMs
3. Tutorial  An SIS epidemic in a closed population
4. Lecture  Considerations for open populations
5. Tutorial  An SI epidemic in an open population
6. Lab  Adding heterogeneity & interventions
7. Lecture  Extending EpiModel for novel research
Workshop Outline

1. Lecture Introduction
2. Lecture From network data to temporal ERGMs
3. Tutorial An SIS epidemic in a closed population
4. Lecture Considerations for open populations
5. Tutorial An SI epidemic in an open population
6. Lab Adding heterogeneity & interventions
7. Lecture Extending EpiModel for novel research
Independent vs Dependent Simulations

• Independent simulations
  - Network structure does not depend on epidemiology, demography, or other exogenous processes
    - The epidemiology still depends on network structure!
  - Closed populations and fixed nodal attributes

• Dependent simulations
  - Network structure does depend on exogenous processes
  - Open populations: births, deaths, and migration
  - Time-varying attributes: disease status and aging
Implication 1

Network Resimulation

Independent Models

Network Simulation → Epidemic Simulation

$t_1$ to $t_n$

Dependent Models

Net $\rightarrow$ Epi $\rightarrow$ Net $\rightarrow$ Epi $\rightarrow$ Net $\rightarrow$ Epi $\ldots$ $\rightarrow$ Net $\rightarrow$ Epi

$t_1$ to $t_n$
Implication 2

Formation Model Coefficient Adjustment

• What happens to mean degree when population size changes?
  - Growing population = growing mean degree
    - Person moving from 10k town to 10k city increases degree by 10-fold

• EpiModel includes a edges coefficient adjustment as a function of population size

\[ \theta_{t_2} = \theta_{t_1} + \log(N_{t_1}) - \log(N_{t_2}) \]

  - Growing population = shrinking density, preserved mean degree
Implication 3

Dissolution Model Coefficient Adjustment

- STERGMs with demography
  - STERGMs assume a fixed node set where edge dissolution is endogenous
  - Death is an exogenous method of edge dissolution
  - Edge duration usually estimated on living populations
  - Without adjustment, mean degree would fall below empirically observed

- Dissolution coefficient adjustment for deaths/exits
  - Increase the dissolution coefficients (really, edge persistence coefficients)
  - Analytically solved for dyadic independent dissolution models
Workshop Outline

1. **Lecture**  Introduction
2. **Lecture**  From network data to temporal ERGMs
3. **Tutorial**  An SIS epidemic in a closed population
4. **Lecture**  Considerations for open populations
5. **Tutorial**  An SI epidemic in an open population
6. **Lab**  Adding heterogeneity & interventions
7. **Lecture**  Extending EpiModel for novel research
Workshop Outline

1. **Lecture**  Introduction
2. **Lecture**  From network data to temporal ERGMs
3. **Tutorial**  An SIS epidemic in a closed population
4. **Lecture**  Considerations for open populations
5. **Tutorial**  An SI epidemic in an open population
6. **Lab**  Adding heterogeneity & interventions
7. **Lecture**  Extending EpiModel for novel research
Lab  Adding Heterogeneity & Interventions

• Working in small groups (2 to 3 people) for 30 - 45 minutes
  - Build off the examples in Tutorial 1 or Tutorial 2
  - Use a different network model parameterization
  - Model an SIR disease
  - Explore different epidemic parameters or initial conditions:
    - Add an intervention to the model with the inter.eff and inter.start parameters
  - Try a bipartite network (this will be tough)

• Brief report back on findings
Workshop Outline

1. Lecture  Introduction
2. Lecture  From network data to temporal ERGMs
3. Tutorial An SIS epidemic in a closed population
4. Lecture  Considerations for open populations
5. Tutorial An SI epidemic in an open population
6. Lab     Adding heterogeneity & interventions
7. Lecture  Extending EpiModel for novel research
Extending EpiModel

• Built-in models shown here are for teaching purposes
  - Epidemiology of real diseases much more complex
  - Novel research questions require programming modules controlling mechanics of interest

• EpiModel has a “plug-n-play” API to write modules
  - Modules may replace existing modules: transmission, mortality, summary statistics
  - New modules may supplement existing modules: aging, disease progression, complex intervention
Extending EpiModel Resources

- Advanced Extension Models tutorials
- Network Modeling for Epidemics summer courses
  - Seattle and Belgium in 2015
- Email listserv

http://epimodel.org/
Scaling Up with EpiModelHPC

- Extension package for simulating network models on high-performance computing systems
- Designed for Linux-based OpenPBS systems like Moab or Torque

http://github.com/statnet/EpiModelHPC
Coming Soon… EpiModelHIV

• Extension modules for EpiModel specifically for modeling HIV infection
  - Modules geared towards heterosexual transmission in Sub-Saharan Africa

• Modules include:
  - Natural disease progression impact on CD4 and HIV viral load trajectories
  - Inter-host transmission risk dependent on disease stage and VL
  - Anti-retroviral therapy treatment

• Modules expanded for modeling HIV in MSM next
Infectious Diseases & Social Networks Session
10:40 – Noon, Thursday @ Viscount Room


• Samuel Jenness. Effectiveness of Male Circumcision for HIV-1 Prevention Depends on Contact Network Structure.