STATISTICAL CONSIDERATIONS FOR MODELING EPIDEMICS WITH DEPENDENT PROCESSES

Slides: Steven M. Goodreau, Ph.D., Samuel M. Jenness, Ph.D., Martina Morris, Ph.D.

Changing network size and composition

- As social networks change in size (say, for instance, as a village of n = 5,000 nodes grows to n = 10,000 nodes), which of the following do you think is generally preserved?
  - Number of edges? e
  - Mean degree? $2e/n$
  - Density? $e/\binom{n}{2}$

Changing network size and composition

- Applying the coefficients as is from a given *stergm* fit to a network of changing size will lead to preservation of density across time
- For one-mode networks: preserving mean degree instead requires a simple transformation of the edges coefficient in the formation model:

$$\theta_{new} = \theta_{old} + \ln(N_{old}) - \ln(N_{new})$$

- This is mathematically equivalent to partitioning the original edges term into an offset equal to ln(*N*) and a residual, and then updating the offset as *N* changes.
- EpiModel handles this for you
- If you are going to code your own models outside EpiModel, you must handle this

Changing network size and composition

- As network composition changes, balancing will happen automatically – the explicit pairing of individuals requires this.
- Nevertheless – one does not have always have straightforward control over this
  - e.g. with just an edges term in the model, two sexes will automatically "meet in the middle"
- Can change parameterizations to obtain different dynamic behavior
  - worth thinking through the behavior you expect, and what you see for your model
- Some theory to guide you can be found in Morris (1991), Koehly, Goodreau and Morris (2004), Krivitsky, Handcock and Morris (2011)

Relational dissolution through death

- We fit our dynamic network using static data, with a process for dissolving relationships governed by a coefficient derived from relational duration
- All of this was done in a context that contained no information about death – another process that terminates relationships
- If we simply layer death on to our model (even with the size correction on the previous slide) we will see two measures drop down below the expected values we want:
  - relationship durations
  - number of relationships
- Some aspects of this might be desired
  - e.g. if we could interview dead people we might find their past relationships to be shorter than those of the same birth cohort in our sample who are still alive,
- but others are likely not.

Relational dissolution through death

- An approximate correction for this is:
  1. Calculate dissolution coefficients as before (without considering death)
  2. Estimate formation coefficients conditional on these dissolution coefficients.
  3. Calculate new dissolution coefficients that reflect the log-odds of a relationship sustaining conditional on both actors living, which equals:

$$\operatorname{logit}\left[1-\frac{P(E_t)-P(N_t)}{P(\neg N_t)}\right]$$

where:

$P(E_t)$ = the overall prob. of a tie dissolving at time *t* from any cause = 1/D

$P(N_t)$ = the prob. of either incident node dying at time *t*

Relational dissolution through death

- The complicating factor is that $P(N_t)$ may very well change over time as your disease prevalence changes
- But then again, it's probably OK for relationship lengths to shorten, and network density to decline, slightly as death increases
- Again, EpiModel handles this for you
- If you are going to code your own models outside EpiModel, you must handle this
- Bigger point:

DIAGNOSE THE HECK OUT

OF YOUR SIMULATIONS!!!

Review of offsets and corrections

When approximating the fit of a formation STERGM conditional on dissolution STERGM...	...subtract dissolution coefficients from corresponding formation ones (`edapprox=TRUE`)
When network size N changes and you want to preserve mean degree...	...add the *ln* of the old N and subtract the *ln* of the new N to the edges coefficient in the formation model (or equivalently, use an edges offset and update it with *ln* of new N)
To adjust for deaths in simulating from a STERGM model estimated from a cross-sectional network and durations	...use $\operatorname{logit}\left[1-\frac{P(E_t)-P(N_t)}{P(\neg N_t)}\right]$ in calculating your dissolution coefficients