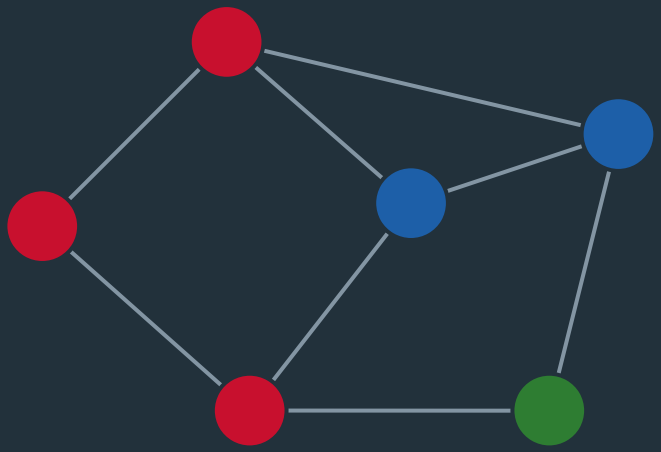
ERGM II: closure, comparison and fit

Course B, Week 6

Chapter 21, from the transitivity step to the end · one lecture and one lab, about 90 minutes each

Instructors are free to use, adapt and share these slides.



Today: a lecture, a break and a lab

0:00	**1 Closure: the last term of the model**	40 min
0:40	**2 Comparing models and judging fit**	40 min
1:20	Break	10 min
1:30	**3 Lab: goodness of fit**	45 min
2:15	**4 Lab: a rival to closure, and a new statistic**	35 min

Week 5 ended with reciprocity in and closure out

Steps 1 and 2

Edges alone is the density (−1.52). Office and practice homophily and partner effects are positive; the model is dyad-independent, so MPLE is exact.

Step 3

Mutual: 1.08, an odds ratio of about 3.0. Ties now depend on ties, so MCMLE. Effective sample sizes 237 to 307: the chain mixes.

Still open

Does advice cluster into closed triads, net of attributes and reciprocity? Today's term: gwesp. Then: which model, and does it fit?

The question for today: a model that fits its own statistics is guaranteed to look right on them. How do we judge a model on what it does not contain?

Three stored fits carry the week

```
. nwwebuse lazega, nwclear
. nwergm lazega_adv, edges nodematch(office practice) ///
>     nodeofactor(status) nodeifactor(status) seed(12345)
. estimates store covariates
. nwergm lazega_adv, edges mutual nodematch(office practice) ///
>     nodeofactor(status) nodeifactor(status) seed(12345)
. estimates store reciprocity
```

Both fits are Week 5's. The third, closure, is built in the first block.

Estimates must be stored under these names: block 2 compares them.

Dyad-independent

covariates: log-likelihood −2004.49, 5 parameters

With reciprocity

reciprocity: log-likelihood −1971.79, 6 parameters

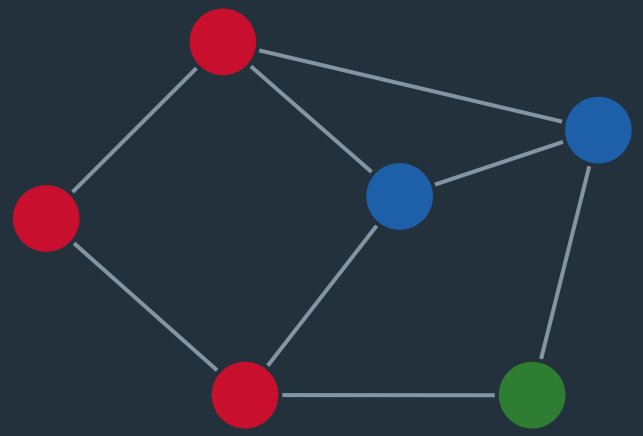
seed() on every fit

so numbers repeat everywhere

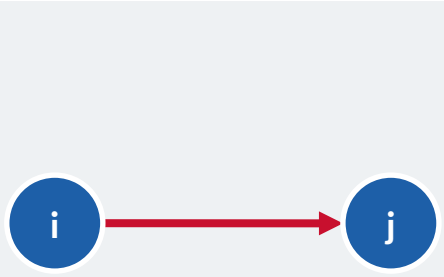
1

Closure: the last term of the model

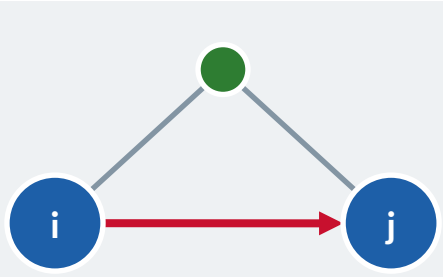
40 min · Chapter 21: Step 4 (transitivity), reading the final model, the diagnostics of the final fit



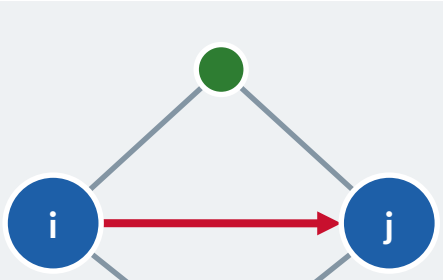
A closed triad is a tendency: the shared partners of a tie



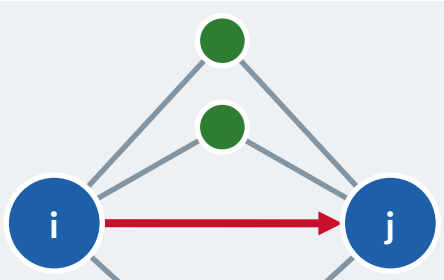
no shared partner



one



two



three

Closure

An arc i to j is likelier when i and j already share contacts: a lawyer asks the lawyers her colleague asks.

The count is a joint property

How many shared partners an arc has depends on other ties. Toggle one tie and the counts of other arcs change: that is dependence.

gwesp: each further shared partner adds less

Shared partners of the arc	Weight in the statistic	Added by the last partner
0	0	
1	1.000	1.000
2	1.393	0.393
3	1.548	0.155
4	1.609	0.061
5	1.633	0.024

gwesp(.5): weight of an arc with k shared partners, exp(.5) × {1 − (1 − exp(−.5))^k}

gwesp(.5) is the closure term of the whole chapter; its coefficient is labeled gwesp_.5.

Why not a triangle count?

A raw count weights every triad alike and tends to give degenerate models (Chapter 22, next week).

The decay is a choice

.5 here. A larger decay keeps rewarding many shared partners; a smaller one stops early. Ex. 6 tries other values.

Step 4: closure added to the advice model

```
. nwergm lazega_adv, edges mutual nodematch(office practice) ///
>     nodeofactor(status) nodeifactor(status) gwesp(.5) seed(12345)
                    | Coefficient  Std. err.      z    P>|z|
              edges |  -4.662469   .1821854   -25.59   0.000
             mutual |   1.031625   .1244439     8.29   0.000
    nodematch_office |   1.016895   .0839039    12.12   0.000
  nodematch_practice |   .7571778   .0784148     9.66   0.000
nodeofactor_status_2 |   .8242996   .0907388     9.08   0.000
nodeifactor_status_2 |   .4019506   .0892556     4.50   0.000
           gwesp_.5 |   .5597618   .1091626     5.13   0.000
```

`MCMLE converged after 5 iteration(s).` Log-likelihood −1954.32, AIC 3922.6, BIC 3968.2.

Positive and clear: net of attributes and reciprocity, an advice arc is likelier when it closes paths through shared contacts.

0.56

gwesp_.5, standard error 0.11, z = 5.13

- **Then store it**

  estimates store closure: the third name of the week.

Closure changes the other coefficients

Term	Step 3 (reciprocity)	Step 4 (closure)
edges	−3.961	−4.662
mutual	1.083	1.032
nodematch_office	**1.199**	**1.017**
nodematch_practice	**0.870**	**0.757**
nodeofactor_status_2	**1.019**	**0.824**
nodeifactor_status_2	**0.586**	**0.402**
gwesp_.5		0.560

Coefficients of differently specified models are not on a common scale: the drop is not a share of homophily that turned out to be closure.

Attribute effects shrink

Ties within an office also tend to be mutual and to close triangles; without those terms the attribute coefficients absorb the overlap.

Mutual hardly moves

1.08 to 1.03: reciprocity and closure are separate tendencies here.

Four tendencies compete for the same ties

4.1 → 2.8

odds ratio of the same office, before and after reciprocity and closure

2.8 → 2.1

odds ratio of the same practice

1.08 → 1.03

the mutual coefficient: hardly moves

What it means

The size of a homophily coefficient depends on whether the local patterns are in the model. Ties inside an office are also mutual and closed more often.

What it does not mean

That a third of the homophily was closure. Coefficients of differently specified models are not on a common scale.

Reading the final model as odds ratios

```
. foreach t in mutual nodematch_office nodematch_practice ///
>     nodeofactor_status_2 nodeifactor_status_2 {
  2.     display %-24s "`t'" %8.2f exp(_b[`t'])
  3. }
mutual                    2.81
nodematch_office          2.76
nodematch_practice        2.13
nodeofactor_status_2      2.28
nodeifactor_status_2      1.49
```

No odds ratio for gwesp

Its change statistic depends on how many shared contacts the arc would close.

How to report it

Advice follows office, practice and status, is reciprocated and is clustered, each net of the others, and the model does not say why: shared office and similar tastes look alike in these coefficients.

What the closure coefficient says, and what it does not

It says

Net of homophily, status and reciprocity, an advice arc is more likely when it closes paths through shared contacts, with diminishing returns for each further shared contact (gwesp_.5 = 0.56, standard error 0.11).

It does not say

That lawyers choose a colleague because a colleague of a colleague recommended him, or that closure causes advice. A cross-sectional ERGM describes which configurations are over-represented, not why they arise or in what order the ties formed.

Pooled over arcs

One coefficient for every arc of the network: an average tendency.

Conditional on a model

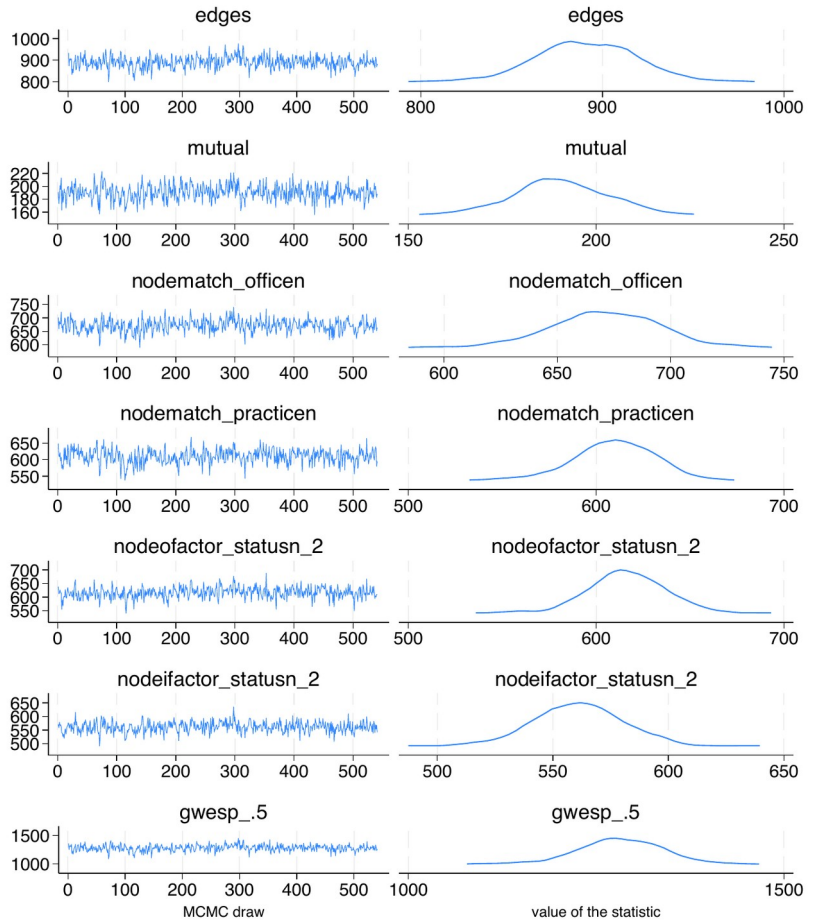
Valid as far as the model is adequate: the second block checks that.

The chain of the final fit: effective sample sizes 292 to 356

```
. estat mcmcdiag, plot
           Statistic |  Autocorr       ESS   Geweke z
               edges |      0.218     348.4      -1.11
              mutual |      0.219     347.1      -1.09
    nodematch_office |      0.259     319.5      -0.66
  nodematch_practice |      0.301     292.2       0.39
nodeofactor_status_2 |      0.236     335.4      -2.25*
nodeifactor_status_2 |      0.231     338.7      -0.76
            gwesp_.5 |      0.207     356.3      -1.09
```

Sample size 541, interval 4,096, acceptance rate 0.4998; columns Mean and SD omitted.

One flagged Geweke z among seven is what chance gives; the trace plots decide. The simulation is centered near the observed network: 891 arcs against 892.

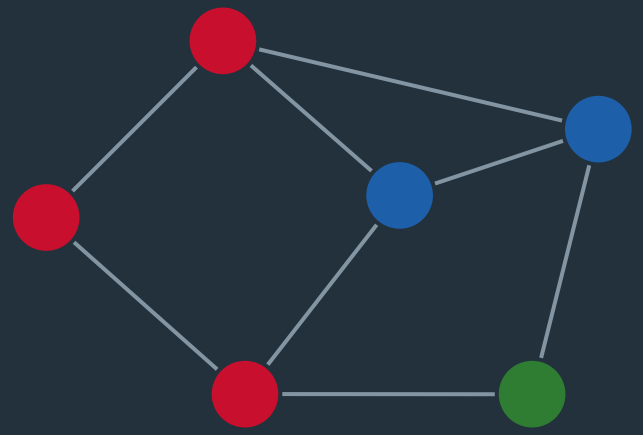


estat mcmcdiag, plot (Figure 21.1)

2

Comparing models and judging fit

40 min · Chapter 21: AIC and BIC, lrtest, predict, estat gof, what can go wrong



Two checks, two questions

Question	Command	Look at
Did the simulation behind the estimates mix?	`estat mcmcdiag`	effective sample size, autocorrelation, Geweke
Which of two models is better?	`estimates stats, lrtest`	AIC, BIC, likelihood ratio
Does the model reproduce the observed network?	`estat gof`	degree, shared partners, distances

Convergence is not a verdict

The message records only that the algorithm stopped.

Fit is judged on what the model does not contain

Agreement on its own statistics is guaranteed.

Each term earns its place: AIC and BIC

```
. estimates stats covariates reciprocity closure
      Model |          N   ll(null)  ll(model)      df        AIC        BIC
------------+---------------------------------------------------------------
  covariates |      4,970  -3444.941  -2004.493       5   4018.985   4051.541
 reciprocity |      4,970  -3444.941  -1971.793       6   3955.585   3994.652
     closure |      4,970  -3444.941  -1954.322       7   3922.643   3968.222
```

Both criteria agree

AIC and BIC prefer each larger model: reciprocity, then closure.

Exact or estimated

Exact for dyad-independent fits; bridge sampling for MCMLE, with a Monte Carlo standard error e(ll_se).

N is 4,970

The BIC counts the 71 × 70 ordered pairs as the sample size.

The likelihood-ratio test agrees: chi-square 34.94

```
. lrtest reciprocity closure
Likelihood-ratio test
Assumption: reciprocity nested within closure

 LR chi2(1) =  34.94
Prob > chi2 = 0.0000
```

34.94

χ^2 on one degree of freedom: the gwesp term

- Only the models that are nested: the smaller one's terms are a subset
- The χ^2 reference is approximate for a single network and an estimated log-likelihood; 34.94 leaves no doubt
- After method(mple) with tie dependence the likelihood is a pseudolikelihood: do not compare it

Better is not adequate

The test says the closure model beats the reciprocity model. It does not say that the closure model reproduces the network: that is the next question.

predict: the conditional probability of each tie

```
. predict, frame(pr) pr
. frame pr: summarize p
    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
           p |      4,970     .1722427    .1679888   .0093548   .9024792
```

It says

How plausible each arc is, given every other arc of the observed network.

It does not say

How well the model would predict arcs that were never observed: each probability conditions on the rest of the network.

0.172

mean conditional probability of a tie over 4,970 pairs; range 0.009 to 0.90

Conditional

The quantity the coefficients are read through.

Goodness of fit: simulate, then compare

```
. estat gof, gof(model distance odegree idegree ///
>     espartners) seed(12345) plot
```

Simulate

100 networks from the fitted coefficients, with the MCMC settings of the fit.

Compare

Observed value, minimum, mean and maximum of the simulated values, on each statistic.

Monte Carlo p-value

Twice the smaller of the share of simulated values at least as large and the share at most as large as the observed one. Small: the observed value is unusual among the simulations.

Panels: the model's statistics, minimum geodesic distance, out-degree, in-degree, edgewise shared partners.

The model's own statistics fit, as they must

Statistic	Observed	Mean	Min	Max	MC p
edges	892	893.20	819	993	0.96
mutual	192	192.55	161	233	0.98
nodematch_office	674	679.34	608	752	0.80
nodematch_practice	610	611.17	561	711	0.98
nodeofactor_status_2	616	615.94	570	667	0.92
nodeifactor_status_2	562	561.23	516	642	0.94
gwesp_.5	1284.86	1288.29	1134.59	1496.89	0.94

Guaranteed

The estimation matched these statistics by construction. All p-values are 0.80 or higher.

Not evidence of a correct model. The test starts with the other panels.

Statistics the model does not contain: distances do not fit

Distance	Observed	Mean	Min to max	MC p
1	892	893.2	819 to 993	0.96
2	**3,074**	**3,283**	**3,104 to 3,461**	**0.00**
3	**991**	**730**	**527 to 963**	**0.00**
4	13	17.0	0 to 70	1.00
Inf	0	46.8	0 to 209	1.00

Advice paths in the firm are somewhat longer than the model makes them.

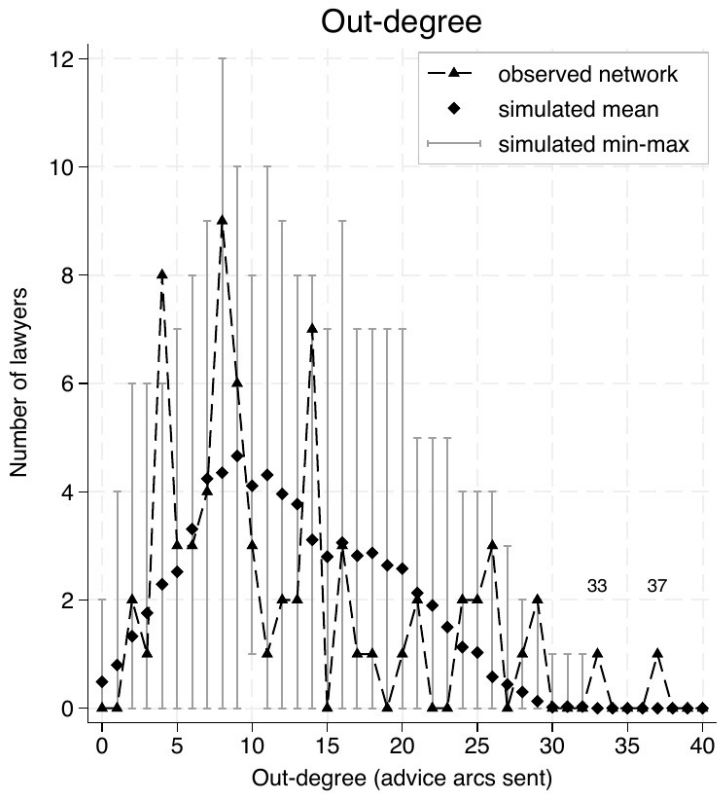
Paths are longer

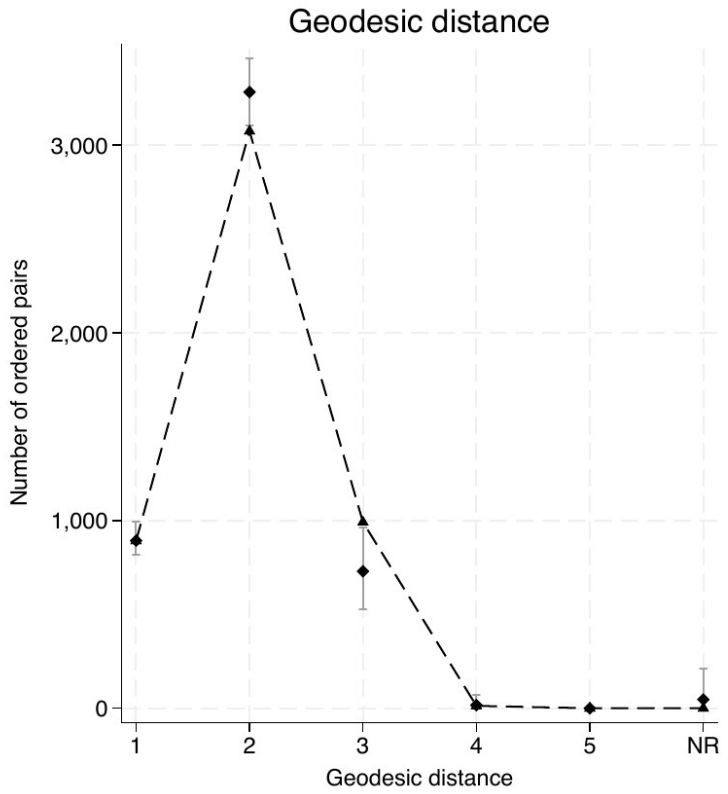
Fewer pairs at distance 2 and more at distance 3 than in any simulated network.

Unreachable pairs

47 per simulated network on average; none in the observed network.

Two lawyers lie beyond every simulated network





33 and 37

out-degrees of two lawyers: no simulated network reaches them

Activity is more dispersed than the model allows. 8 lawyers name exactly 4 colleagues against 2.3 on average.

Observed (dashed line) against 100 simulated networks (Figure 21.2 of Chapter 21).

Shared partners: the clearest failure

34 | 8.7

arcs with 10 shared partners: observed | simulated mean

17 | 1.9

arcs with 12 shared partners

95 | 155

arcs with 2 shared partners: fewer than simulated

Denser clusters than the model

A core of tightly clustered colleagues has many arcs with many shared partners; gwesp(.5) does not produce them. Another decay might (Ex. 6).

Borderline p-values move

Rows with p between 0.02 and 0.06 can change with another seed. The large gaps do not (Ex. 4).

Useful and incomplete: say both

It says

The coefficients describe the over-represented patterns: reciprocity, closure, homophily and status, each net of the others.

It does not say

That the model is correct. It reproduces every statistic it contains (p of 0.80 or higher) and still misses the spread of the out-degrees and the number of pairs at distance 2.

In the write-up

Say both. Add what the gof output points to, refit, run estat gof again. Where the gaps cannot be closed, report them.

Fit on statistics the model does not contain is the main evidence that it is more than a re-description of the data.

What can go wrong

Reporting a coefficient because MCMLE converged

The message records only that the algorithm stopped. A fit can report convergence while the effective sample sizes are a few dozen among thousands of draws. Read estat mcmcdiag after every MCMLE fit.

Treating agreement on the fitted statistics as proof

Agreement is guaranteed, since the estimation matched those statistics by construction. The final advice model fits all of its own statistics and still misses the distances and degrees.

Both are caught by a diagnostic that takes a minute to run.

What the gof output points to

The gof output shows	A candidate repair	In the lab
Shared partners: far too few arcs with two, far too many with ten or more	another gwesp decay: gwesp(.25), gwesp(1)	Ex. 6
Two lawyers with out-degrees 33 and 37, beyond every simulated network	a geometrically weighted out-degree term: gwodegree(.5)	Ex. 6
Closed triads that need not be closure: lawyers who work on the same cases	shared work as a dyadic covariate: edgecov(lazega_cow)	Ex. 9
47 unreachable pairs per simulated network; none observed	a statistic for the reachability of a whole network	Ex. 10

Every added term changes the other coefficients: store, compare, check the fit again.

What to take away from the lecture

2.81

odds ratio of mutual in the final model: advice is returned

2.76

same office (4.1 before reciprocity and closure)

0.56

gwesp_.5: closure, with diminishing returns

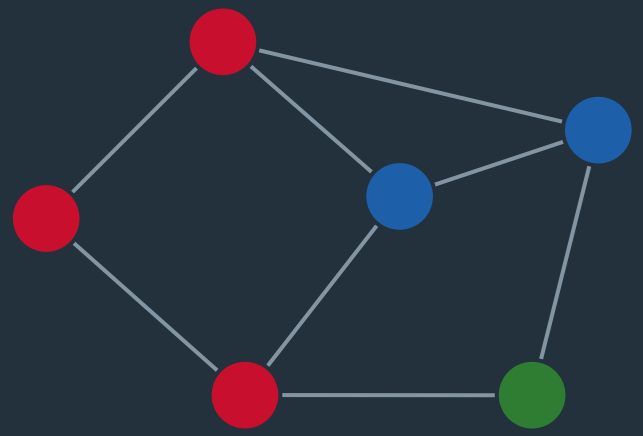
- Build in stages; one dependence term at a time, and check each stage
- Compare with estimates stats and lrtest; compare fit with estat gof
- Judge fit on statistics the model does not contain
- The model is useful and incomplete: say both

The ERGM describes

which local configurations are over-represented net of the others, not why they arise.

Break

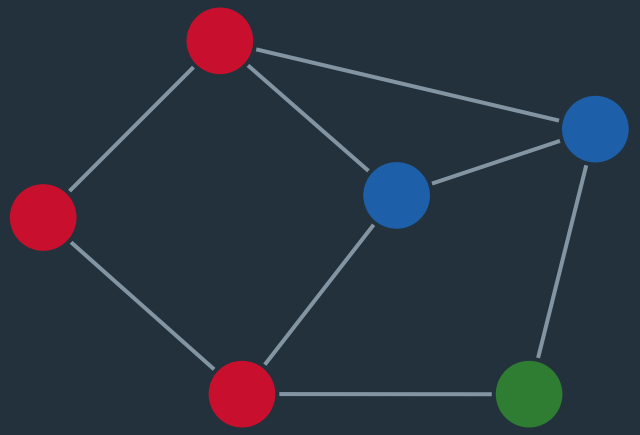
10 minutes



3

Lab: goodness of fit

45 min · Run along: estat gof with more draws, other closure decays, a degree term; Ch. 21, Ex. 4 and 6



Run along: more draws, other closure terms

```
. nwergm lazega_adv, edges mutual nodematch(office practice) ///
>     nodeofactor(status) nodeifactor(status) gwesp(.5) seed(12345)
```

```
. estat gof, gof(distance odegree espartners) nsim(300) seed(12345)
```

Ex. 4: the final model, with 300 simulated networks instead of 100.

```
. nwergm lazega_adv, edges mutual nodematch(office practice) ///
>     nodeofactor(status) nodeifactor(status) gwesp(.25) seed(12345)
. estat gof, gof(odegree espartners) seed(12345)
```

Ex. 6: gwesp(.25); then gwesp(1); then gwesp(.5) gwodegree(.5). Run gof after each fit.

Check the header

estat gof tests the last fit.

Compare by row

The ESP panel changes length from run to run.

Keep the seed

Seed 12345 in every call.

Hands-on (40 min)

Ch. 21, Ex. 4

Rerun estat gof after the final advice model with 300 simulated networks instead of the default 100: estat gof, gof(distance odegree espartners) nsim(300) seed(12345). For the rows that were clearly off with 100 draws (distance 2 and 3, out-degrees 33 and 37, 10 and 12 shared partners), report the simulated mean and the MC p-value. Does the misfit hold up with more draws?

Hint Compare each row with the listing of the goodness-of-fit section. A misfit holds up when the observed value stays outside or at the edge of the simulated range and its p-value stays small; the minimum and maximum can widen with more draws.

Ch. 21, Ex. 6

Refit the final advice model three times: with gwesp(.25) and with gwesp(1) instead of gwesp(.5), and with gwesp(.5) and gwodegree(.5) added (a geometrically weighted out-degree term for the lawyers who ask very many colleagues). Use seed(12345) throughout and run estat gof, gof(odegree espartners) seed(12345) after each fit. Report how the gwesp and mutual coefficients change, and whether the fit of the shared-partner rows and of out-degrees 33 and 37 improves with any of the three. Report a fit that does not converge, if there is one.

Hint The decay sets how quickly additional shared partners stop adding weight: with a larger decay an arc with many shared partners is rewarded more. Count, for each fit, the shared-partner rows with an MC p-value of 0.05 or less.

Ex. 4: the misfit holds up with 300 draws

Row	Observed	Mean, 100 draws	Mean, 300 draws	MC p, 100	MC p, 300
distance 2	3,074	3,283	3,293	0.00	0.01
distance 3	991	730	720	0.00	0.01
out-degree 33	1	0.00	0.00	0.00	0.01
out-degree 37	1	0.00	0.00	0.00	0.00
10 shared partners	34	8.71	8.62	0.00	0.00
12 shared partners	17	1.91	1.79	0.00	0.00

It holds up

Observed values stay at the edge of the simulated range; p-values stay at 0.01 or below.

The range widens

Distance 2: minimum 3,046 now below the observed 3,074. Borderline rows (p near .05) move.

Ex. 6: a larger decay closes the gap; a degree term does not

Model	gwesp	mutual	Log-lik.	AIC	ESP rows, p ≤ .05	Degree 33 | 37, p
gwesp(.5), final	0.560	1.032	−1954.3	3922.6	12 of 18	0.00 | 0.00
gwesp(.25)	0.493	1.068	−1963.4	3940.8	12 of 18	0.02 | 0.00
gwesp(1)	**0.419**	**0.950**	**−1934.4**	**3882.9**	**5 of 22**	**0.14 | 0.02**
gwesp(.5) + gwodegree(.5)	0.592	1.029	−1955.0	3926.1	13 of 18	0.00 | 0.00

gwesp(1) fits best

Lower AIC by 39.7; most shared-partner rows now fit; the chain mixes (ESS 285 to 366).

gwodegree adds nothing

Coefficient 1.93, standard error 2.03; AIC 3926.1, higher than 3922.6.

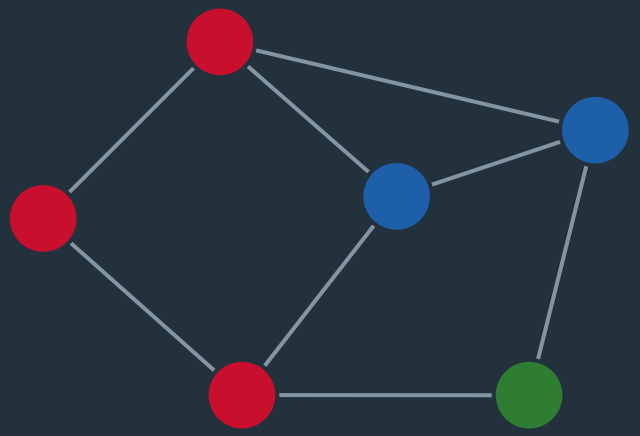
Decay is a choice

State it in the write-up. Coefficients of different decays are not on one scale.

4

Lab: a rival to closure, and a new statistic

35 min · Run along: edgecov(), estimates table, nwreach, nwergm simulate; Ch. 21, Ex. 9 and 10



Hands-on (35 min)

Ch. 21, Ex. 9

Shared work is a rival to closure: lawyers who work on the same cases share contacts, and their advice ties close triangles without anyone turning to the contact of a contact. Add the co-work network as a dyadic covariate to the final advice model, edgecov(lazega_cow), with seed(12345). Report the coefficients of edgecov_lazega_cow, gwesp_.5, mutual and nodematch_office next to those of the final model, and say what the change in gwesp means for reading it as closure.

Hint edgecov() takes the name of a network on the same nodes and adds one term whose change statistic is the value of that network for the pair. Store both fits and compare them with estimates table, as in the gender exercise.

Ch. 21, Ex. 10

The distance panel of estat gof counts the pairs that cannot reach each other (row Inf): none in the observed advice network, 46.8 on average and up to 209 in the simulated networks. Propose a statistic that summarizes the reachability of a whole network, compute it for the observed advice network, and explain what values among networks simulated from the model would show that the model reproduces the firm's reachability.

Hint Start from the Inf row and ask what single number would say, for a whole network, how many of its pairs can reach each other. nwreach (Chapter 7) and nwcomponents (Chapter 9) give the raw material; Chapter 22 shows how to keep the simulated networks. In a loop over them, nwreach stops at the second pass because the network it creates exists already; add the option nwreplace.

Ex. 9: shared work absorbs part of the closure, not all of it

Term	Final model	With co-work
edgecov_lazega_cow		**2.083 (0.092)**
gwesp_.5	0.560 (0.109)	**0.407 (0.104)**
mutual	1.032 (0.124)	0.767 (0.136)
nodematch_office	1.017 (0.084)	0.701 (0.097)

Coefficients with standard errors in parentheses. Log-likelihood −1954.3 and −1656.6; AIC 3922.6 and 3329.1.

8.0

odds ratio of edgecov_lazega_cow: exp(2.083)

A rival focus, not the only one.

Part of the closure is shared work

gwesp falls by about a quarter, 0.56 to 0.41.

Closure is still there

0.41 with a standard error of 0.10: positive, net of co-work.

Ex. 10: the share of pairs that can reach each other

```
. nwreach lazega_adv
. nwsummarize _reach
    Density:  1
. nwergm simulate, coef(b) nsim(100) output(network) seed(1)
. forvalues k = 1/100 {
>     quietly nwreach sim_`k', nwreplace
>     quietly nwsummarize _reach
>     frame post rr (r(density))
> }
```

b holds e(b) of the final model; frame rr was created with: frame create rr double reach.

1 | 0.987

reachability: observed | mean of 100 simulated networks

Observed

All 4,970 pairs reach each other: density of _reach is 1.

Simulated

Min 0.944, max 1; 35 of 100 draws reach every pair.

The verdict

On this statistic the observed value is not unusual; but a typical draw leaves about 63 pairs unreachable.

To read after the session

Next week (Week 7): Chapter 22, Advanced exponential random graph models

Ex. 1, 2, 3, 7 of Chapter 21 if not yet done (Week 5); Ex. 5 and 8 are left over

Write down: the model you would report, its gaps, the terms you tried

Do-files and teaching material: nwcommands.org/book

Instructors are free to use, adapt and share these slides.

