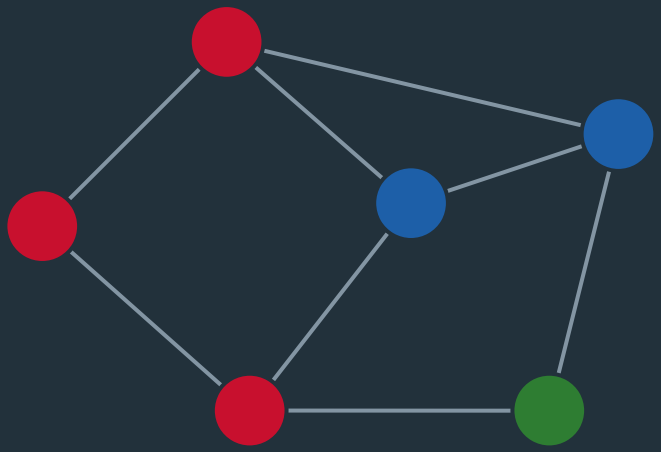
ERGM I: dependence and change statistics



Course B, Week 5

Chapter 21, from the start through reciprocity and the MCMC check

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 Ties that depend on ties**	40 min
0:40	**2 Building the advice model, step by step**	45 min
1:25	Break	10 min
1:35	**3 Lab: the first three models**	45 min
2:20	**4 Lab: gender homophily and the sampler**	40 min

Which local tendencies go with the advice network, once the others are taken into account?

Reciprocity

Advice is an exchange among peers: requests are returned.

Closure

Partners keep one another in line: advice closes into triangles.

Homophily

Opportunity and similarity: advice stays within offices and practices.

Status

Deference: advice flows toward partners and is rarely returned.

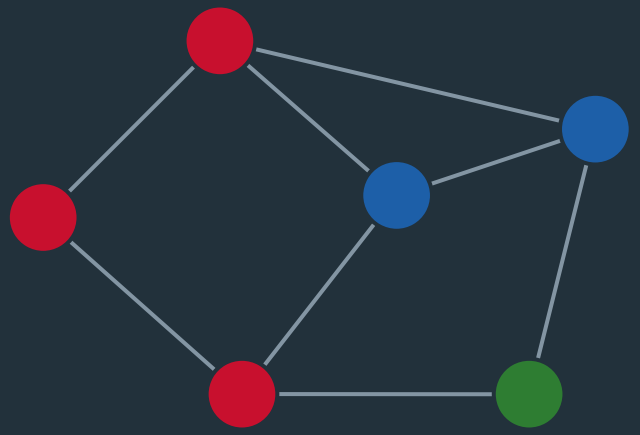
QAP weighed predictors, but has no term for a lawyer who asks a colleague because that colleague asked her. She is acting on other advice ties. An ERGM makes such dependence part of the model, and its coefficients weigh the four tendencies against each other.

Week 4: QAP regression. Week 5: dependence and change statistics. Week 6: closure, comparison and fit.

1

Ties that depend on ties

40 min · Chapter 21: dependence, the ERGM, change statistics, estimation, the terms, the stages of a specification



Four tendencies, four claims about dependence

Tendency	A tie from i to j is more likely when	It depends on
Reciprocity	a tie from j to i exists	one specific other tie
Transitivity	i and j share a partner k	the joint presence of two other ties
Popularity	j already has many ties	the whole configuration of ties at j
Homophily	i and j share an attribute	only the dyad and its attributes

The first three cannot be reduced to i, j and their attributes. Each depends on the state of other ties. That is where QAP stops and an ERGM begins. Homophily reduces to the dyad, and QAP already handles it.

Logit and QAP treat the dependence as a nuisance; an ERGM models it

Logit on all dyads

Treats the dependence as noise. The 4,970 rows count as independent evidence and the standard errors are too small.

QAP regression

Keeps the dependence in the reference distribution and does not model it. Tells whether predictors matter, not how ties depend on ties.

ERGM

Makes the dependence the object of the model: the probability of a whole network is a function of counts of local patterns, such as mutual dyads and closed triangles.

A tie is more likely because another tie exists. That is a reason to put the dependence into the model, not noise to be corrected afterwards.

The network is the outcome

$$P(Y = y) = \exp\{ \theta' \cdot s(y) \} / \kappa(\theta)$$

s(y): statistics

Counts of local patterns in the candidate network y: edges, mutual dyads, closed triads. They carry the dependence between ties.

θ: weights

One parameter per statistic. A weight of zero says that one more such pattern tells nothing about whether a tie is present.

κ(θ): normalizing constant

The sum of exp{θ'·s(y')} over every possible network on the same nodes. It makes the probabilities add to one.

Chapter 1's mechanisms become terms: reciprocity is mutual, closure is gwesp, homophily is nodematch.

Where the exponential form comes from

Declare dependence

Two possible ties depend on each other when they share an actor (Frank and Strauss 1986). The possible ties are the nodes of a dependence graph.

Hammersley–Clifford

Any distribution consistent with a dependence graph factors over its cliques and has exponential-family form.

Cliques

Single ties, stars of every order and triangles: the basic vocabulary of Markov graph models.

Geometric weights

The raw triangle count is unstable. Geometrically weighted statistics such as gwesp (Hunter and Handcock 2006) replace it.

Nothing later depends on this slide. The practical consequence is the vocabulary of terms: edges, stars, triangles and their geometrically weighted versions.

A coefficient is a change in the conditional log-odds of one tie

```
logit P(Y_ij = 1 | y_-ij)  =  θ'·δ(y)_ij
δ(y)_ij  =  s(y with tie ij)  −  s(y without tie ij)
```

It says

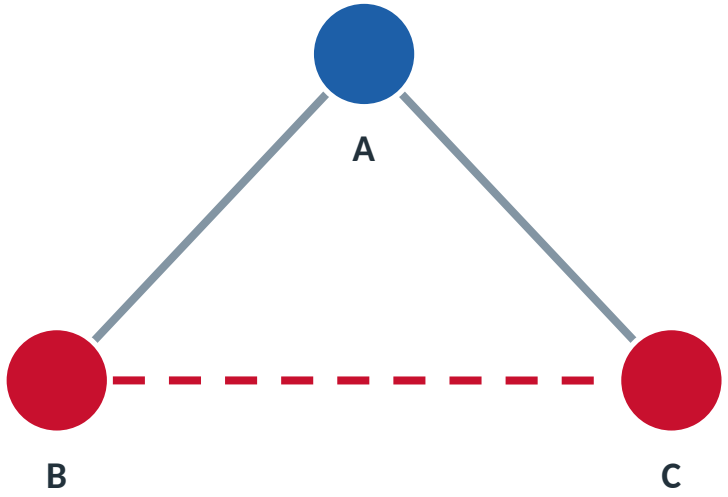
the conditional log-odds of one tie i→j, holding every other tie fixed, change by the coefficient for each unit of its change statistic. The exponential of the coefficient is the odds ratio of that tie.

It does not say

that the coefficient is a probability ratio, or a statement about two random pairs from the network, or a cause. It compares two states of the same tie, given the rest of the network.

A change statistic worked out by hand

Undirected: triangle



Toggle B–C. The triangle count rises from 0 to 1, so the change statistic of the triangle term is 1. A pair with no shared contact has 0.

Directed: mutual



Toggle B→A. An asymmetric dyad becomes mutual, so the change statistic of mutual is 1. The toggle D→B would change nothing in the count of mutual dyads: 0.

Two ways to estimate: MPLE and MCMLE

	MPLE	MCMLE
Used when	edges and dyadic covariates only: no tie depends on another	a structural term such as mutual or gwesp is in the model
How	a logistic regression of every dyad on its change statistics	simulate networks from the model, compare their statistics with the observed ones, move θ
Exact?	yes, for a dyad-independent model	as good as the simulation: check the diagnostics
The header says	Estimation: = MPLE	Estimation: = MCMLE

10^{1496}

directed networks without self-loops on the 71 lawyers: the sum κ(θ) runs over all of them

```
. display 4970*log10(2)
1496.1191
```

The effects catalogue: terms by the claim they make

Kind of claim	Terms in nwergm	Ties depend on ties?
Baseline density	edges	no
Reciprocity (directed)	mutual	yes
Homophily and covariates	nodematch, nodefactor, nodeofactor, nodeifactor, nodemix, nodecov, absdist, edgecov	no
Triadic closure	gwesp(decay); raw counts triangle and ctriple (degeneracy-prone)	yes
Degree and popularity	degree, kstar, concurrent, gwdegree, sender, receiver	mostly yes

125 of the 139 terms of R's ergm, with R's names. This week: edges, nodematch, nodeofactor, nodeifactor and mutual. Week 6: gwesp.

A specification is built in stages, each checked before the next

Stage	What is added	What to check
1	edges alone	the network and the software behave; the coefficient is the log-odds of the density
2	covariates: nodematch, nodeofactor, nodeifactor	still dyad-independent: MPLE is exact
3	mutual	MCMLE takes over; read estat mcmcdiag
4	one structural term at a time: gwesp (Week 6)	trace the problem to the term that caused it
5	estat mcmcdiag after every fit	acceptance rate, autocorrelation, effective sample size
6	comparison and goodness of fit (Week 6)	statistics the model does not contain
7	report the whole process	including terms tried and rejected (Chapter 30)

What to take away from block 1

Dependence is the point

Reciprocity, closure and popularity are claims that a tie depends on other ties. The ERGM puts that into the model; QAP and the logit do not.

A network has a probability

$P(Y = y)$ rises with a weighted sum of counts $s(y)$. $\kappa(\theta)$ cannot be computed, so the model is read through the conditional form.

A coefficient is a log-odds change

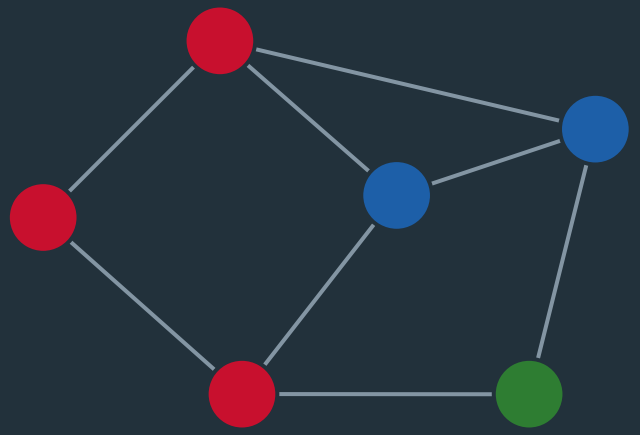
θ is the change in the conditional log-odds of one tie per unit of its change statistic, every other tie held fixed. $\exp(\theta)$ is the odds ratio.

Next: the advice network of the law firm, built up in stages: density, homophily and status, reciprocity.

2

Building the advice model, step by step

45 min · Chapter 21: Steps 1 to 3 on the Lazega advice network, and the check of the MCMC simulation



The advice network as an outcome

892

advice arcs among 71 lawyers

.179

density

192

mutual pairs of 2,485 (508 asymmetric, 1,785 null)

.274

reciprocity: mutual share of the non-null pairs

```
. nwwebuse lazega, nwclear
. nwdyads lazega_adv
    Mutual |      Asym |      Null
-----------+-----------+-----------
       192 |       508 |      1785
    Reciprocity: .274
```

Two questions

Do lawyers ask colleagues in their own office and practice, and do partners seek and give more advice than associates?

Once these are controlled, is advice reciprocated, and does it close into triads?

Step 1: density only

```
. nwergm lazega_adv, edges seed(12345)

Estimation:            =  MPLE

------------------------------------------------
  lazega_adv | Coefficient  Std. err.      z    P>|z|
-------------+----------------------------------
       edges |  -1.519896   .0369634   -41.12   0.000
------------------------------------------------

Log-likelihood: -2338.8785 (df = 1)   AIC:    4679.8   BIC:
4686.3
```

−1.52

= ln{ .1795 / (1 − .1795) }, the log-odds of the density

The null of the later steps. All 71 lawyers alike, advice ties scattered at random.

The table shows no confidence interval so that the names of the terms fit at 80 columns; the option ci adds it.

Step 2: homophily and status

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

Estimation:            =  MPLE

------------------------------------------------------------------
          lazega_adv | Coefficient  Std. err.      z    P>|z|
---------------------+--------------------------------------------
               edges |  -4.095448   .1309373   -31.28   0.000
    nodematch_office |   1.422373   .0887433    16.03   0.000
  nodematch_practice |   1.037945   .0837376    12.40   0.000
nodeofactor_status_2 |   1.125832   .0842818    13.36   0.000
nodeifactor_status_2 |   .7831546   .0819723     9.55   0.000
------------------------------------------------------------------
. estimates store covariates
```

Still dyad-independent. MPLE is exact. Names ending in _2 refer to partners, the second level of status.

Reading step 2: four odds ratios

Term	Coefficient	Odds ratio	An arc is more likely when
`nodematch_office`	1.42	4.1	the two lawyers share an office
`nodematch_practice`	1.04	2.8	the two lawyers share a practice
`nodeofactor_status_2`	1.13	3.1	the sender is a partner
`nodeifactor_status_2`	0.78	2.2	the receiver is a partner

Net of density only. Lawyers who share an office and a practice also share colleagues and return each other's requests. A model without terms for that credits those ties to the attributes.

Without tie dependence, the ERGM is a logistic regression

```
. nwexpand office, name(sameoffice)
. nwexpand practice, name(samepractice)
. nwexpand partner, mode(sender) name(frompartner)
. nwexpand partner, mode(receiver) name(topartner)
. preserve
. nwtoedge lazega_adv samepractice sameoffice frompartner topartner
. drop if _ego == _alter
. logit lazega_adv sameoffice samepractice frompartner topartner, nolog
. restore
```

Term	nwergm	logit
edges / _cons	−4.095448	−4.095448
nodematch_office	1.422373	1.422373
nodematch_practice	1.037945	1.037945
nodeofactor_status_2	1.125832	1.125832
nodeifactor_status_2	.7831546	.7831545

The same numbers. The likelihood factors by dyad, so the logit is the ERGM. With mutual in the model it no longer is.

Step 3: reciprocity makes the model dyad-dependent

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

Estimation:            =  MCMLE
MCMC sample size:      =  913

------------------------------------------------------------
          lazega_adv | Coefficient  Std. err.      z    P>|z|
---------------------+--------------------------------------
               edges |  -3.961329   .1247155   -31.76   0.000
              mutual |   1.083464   .1279072     8.47   0.000
    nodematch_office |   1.198955   .0882039    13.59   0.000
  nodematch_practice |   .8695592   .0769769    11.30   0.000
nodeofactor_status_2 |   1.018712   .0872432    11.68   0.000
nodeifactor_status_2 |    .586494    .084506     6.94   0.000
------------------------------------------------------------
MCMLE converged after 7 iteration(s).
. estimates store reciprocity
```

MCMLE. mutual makes an arc depend on the reverse arc. The estimator now simulates networks from the model.

Other coefficients shrink. Office 1.42 to 1.20, practice 1.04 to .87, partner receiving .78 to .59: part of the homophily was reciprocity.

Reading the mutual coefficient

1.08

mutual coefficient, Step 3

2.95

its odds ratio: about 3.0

It says

compared with an otherwise identical pair, an arc from i to j has about 3.0 times the odds of being present when the arc from j to i is present, given the other terms in the model.

It does not say

that an arc is three times as probable: with a probability near 0.2 the ratio of probabilities is nearer 2 than 3. Nor is it the coefficient of every model: it changes to 1.03 once closure is added in Step 4.

An ERGM as a network in equilibrium

One toggle at a time

The sampler proposes to switch one tie and accepts with a probability computed from the change statistics.

The stationary distribution

Run long enough, the chain visits every network with a stable frequency. That distribution is the ERGM.

The dynamics are implicit

One snapshot cannot say whether lawyers came to return requests, or whether such pairs stayed in the data.

	ERGM (this week)	SAOM (Week 9)
Oriented to	ties: a distribution over whole networks	actors: opportunities to change their own ties
Data	one network	a panel of waves

Checking the MCMC simulation: estat mcmcdiag

```
. estat mcmcdiag

MCMC diagnostics, final sample   Sample size     =       913
                                 Burn-in         =    16,384
                                 Interval        =     2,048
                                 Acceptance rate =    0.5148

          Statistic |      Mean        SD  Autocorr       ESS   Geweke z
--------------------+-----------------------------------------------------
              edges |    893.27     28.06     0.538     307.4      1.10
             mutual |    191.98     12.26     0.522     287.2      1.22
    nodematch_office |    675.01     23.40     0.550     265.6      0.43
  nodematch_practice |    611.74     23.13     0.589     237.1      0.96
nodeofactor_status_2 |    616.84     21.43     0.533     279.3      1.20
nodeifactor_status_2 |    563.08     21.04     0.532     304.6      1.26
```

The sampler

Records every 1,024th proposal after a burn-in of 16,384.

Base sample 256 draws, extended until the effective sample size is 64.

Here the interval grew to 2,048.

Reading the diagnostics

Column	What it measures	In this fit
Mean, SD	the simulated draws of each statistic; compare with the observed counts	893 arcs against 892 observed; 192 mutual pairs, as observed
Autocorr	correlation between two successive recorded draws	0.52 to 0.59
ESS	the number of independent draws that carry the same information; a few hundred is comfortable, a few dozen is not	237 to 307
Geweke z	first tenth of the draws against the last half; |z| above 1.96 is flagged	largest 1.26, none flagged

A slow chain shows in Autocorr and ESS and in the trace plots, not in the tests. The convergence message records only that the algorithm stopped.

Convergence is not adequacy of the simulation

Converged

The message records only that the algorithm stopped. Read estat mcmcdiag after every fit.

ESS small

Raise mcmleeffectivesize(), lengthen mcmcinterval(), or raise mcmcsamplesize(). Read the diagnostics again.

Statistics stop varying

Or the ESS stays small however long the chain: the cause is the model, not the chain (degeneracy, Week 7).

The matrix r(ess) holds the effective sample sizes for a loop over settings. The lab asks you to compare two chains of different lengths.

What to take away from block 2

Build in stages

Density, then covariates (MPLE, exact), then reciprocity (MCMLE). Each step changes the others: office fell from 1.42 to 1.20 when mutual entered.

Read odds ratios

Net of the other terms, an arc has about 3.0 times the odds when the reverse arc exists. The ratio of probabilities is smaller.

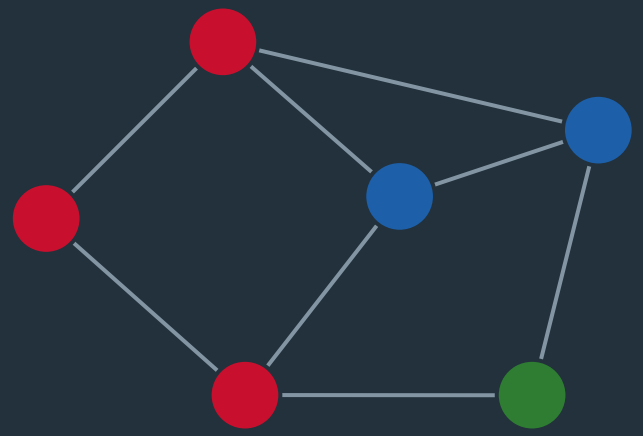
Check the sampler

Convergence says the algorithm stopped. Read Autocorr and ESS after every MCMLE fit.

Week 6 continues with closure (Step 4: transitivity), model comparison and goodness of fit.

Break

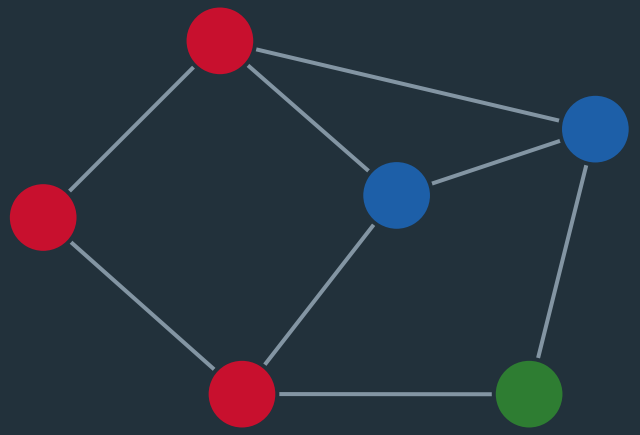
10 minutes



3

Lab: the first three models

45 min · Run along: Steps 1 to 3 and estat mcmcdiag; Ch. 21, Ex. 1 and 2



Run along 1: the three steps

```
. nwwebuse lazega, nwclear
. nwergm lazega_adv, edges seed(12345)
. 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
. estat mcmcdiag
```

Checkpoint	Value
Step 1: edges	−1.519896
Step 2: nodematch_office	1.422373
Step 2: log-likelihood	−2004.4927
Step 3: mutual	1.083464
Step 3: MCMLE iterations	7
Step 3: log-likelihood	−1971.7926
mcmcdiag: interval	2,048
mcmcdiag: smallest ESS	237.1

Predict first. Before step 3: will the office coefficient rise, fall, or stay when mutual enters?

A different MCMLE number usually means a missing seed(12345). The > in the listing is Stata's continuation marker, not part of the command.

Run along 2: how the steps change each other

```
. estimates table covariates reciprocity, ///
>         b(%7.3f) se(%7.3f)
```

Term	covariates	reciprocity
edges	−4.095 (.131)	−3.961 (.125)
nodematch_office	1.422 (.089)	1.199 (.088)
nodematch_practice	1.038 (.084)	.870 (.077)
nodeofactor_status_2	1.126 (.084)	1.019 (.087)
nodeifactor_status_2	.783 (.082)	.586 (.085)
mutual		1.083 (.128)

Every other coefficient falls when mutual enters; the standard errors hardly change. The earlier steps credited reciprocity to the attributes.

Hands-on (20 min)

Ch. 21, Ex. 1

In the final advice model (Step 4, next week), the coefficient of nodematch_office is 1.017, an odds ratio of 2.76. Using the conditional log-odds form of this chapter, state precisely what this says about an advice arc from one lawyer to another of the same office. Then name two things it does not say, for example about two lawyers of the same office chosen at random from the firm, or about why lawyers of one office ask each other.

Hint Write the conditional log-odds of the arc i→j for a pair in the same office and for an otherwise identical pair in different offices, with every other arc held fixed. The two differ by the coefficient times the change statistic of nodematch_office, which is 1 for a pair in the same office.

Ch. 21, Ex. 2

How many directed networks without self-loops can be formed on the 71 lawyers? Give the answer as a power of ten. Explain why the normalizing constant κ(θ), a sum over all of these networks, cannot be computed, and why the coefficients are therefore read through the conditional form, which does not contain κ(θ).

Hint Each of the 71×70 = 4,970 ordered pairs is either tied or not, so there are 2^4970 networks; display 4970*log10(2) gives the power of ten. Then write the conditional probability of one tie as a ratio of two network probabilities and see what happens to κ(θ).

Ch. 21, Ex. 1 (left) and Ex. 2 (right): the answers

It says
for an arc i→j, the conditional odds are 2.76 times as large when i and j share an office, given every other arc and the other terms.

It says
2^4970 networks, about 10^1496 (display 4970*log10(2) prints 1496.1191). κ(θ) is a sum of that many terms, so it cannot be computed.

It does not say
anything about two same-office lawyers drawn at random, or a ratio of probabilities, or why they ask each other (floor, cases, friendship, selection).

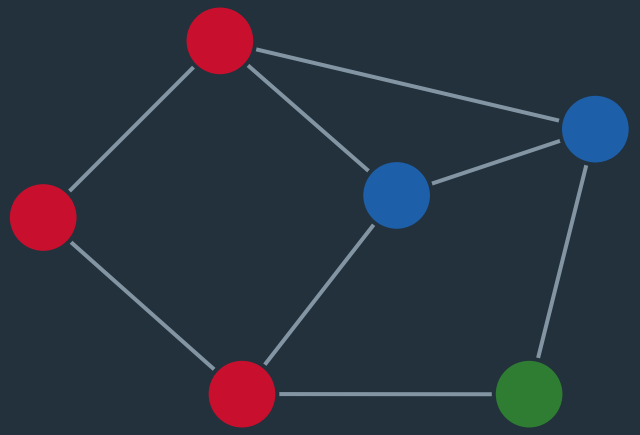
It does not say
that the coefficients are out of reach: in the conditional form κ(θ) cancels, which is why they are read through it.

Why κ cancels: the conditional probability of a tie is the ratio of the probabilities of two networks that differ only in that tie, and both have κ(θ) in the denominator.

4

Lab: gender homophily and the sampler

40 min · Ch. 21, Ex. 3 and Ex. 7: a new term in the larger model, and two chains of different length



Hands-on (25 min)

Ch. 21, Ex. 3

Extend the final advice model with homophily on gender (a string variable like office): nwergm lazega_adv, edges mutual nodematch(office practice gender) nodeofactor(status) nodeifactor(status) gwesp(.5) seed(12345). Report the coefficient of nodematch_gender, its odds ratio and its p-value, and say whether the other coefficients change when it is added. Then run estat mcmcdiag and report the smallest effective sample size and any Geweke test that is flagged.

Hint Store the final model with estimates store closure as in Step 4 and the new fit under another name, for example gender; estimates table closure gender puts the two side by side. A flagged Geweke test among several statistics is read as in the Step 4 diagnostics: look at the trace plots (estat mcmcdiag, plot).

Ch. 21, Ex. 7

Refit the step 3 model (with mutual, without gwesp) with mcmcinterval(50) mcmcburnin(3000) mcmleeffectivesize(0) seed(12345), and again with the interval raised to 1024 and the other options unchanged. After each fit, run estat mcmcdiag and report the autocorrelations and the smallest effective sample size. Which of the two fits would you report, and why could the coefficients alone not have told you?

Hint The effective sample sizes are in the matrix r(ess); mata: min(st_matrix("r(ess)")) prints the smallest. Compare the convergence messages of the two fits as well.

Chapter 21, Ex. 3: gender homophily is small

Term	final model	+ gender
edges	−4.662 (.182)	−4.730 (.189)
mutual	1.032 (.124)	1.019 (.128)
nodematch_office	1.017 (.084)	1.017 (.087)
nodematch_practice	.757 (.078)	.772 (.073)
nodematch_gender		.161 (.077)
nodeofactor_status_2	.824 (.091)	.796 (.089)
nodeifactor_status_2	.402 (.089)	.364 (.087)
gwesp_.5	.560 (.109)	.558 (.108)

1.17

odds ratio of nodematch_gender (p = .036)

277

smallest effective sample size

Small, and the others hardly move. Four Geweke tests are flagged (edges, both status terms, gwesp, |z| from 2.56 to 2.84): look at the trace plots and say so in the report.

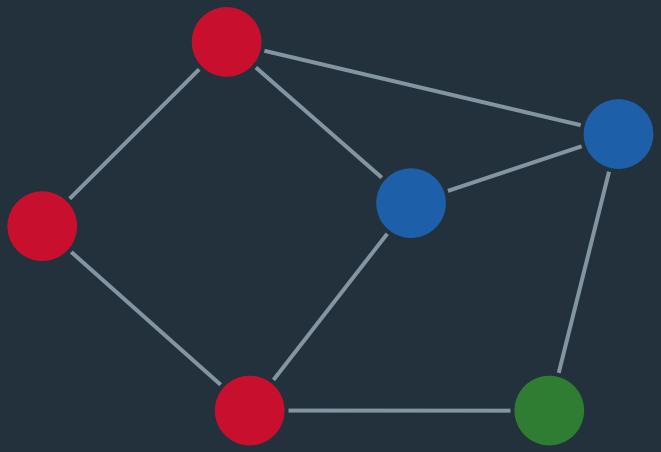
Chapter 21, Ex. 7: same coefficients, different warnings

Fit	Converged	Autocorr	Smallest ESS	edges	mutual
Interval 50, 3,000 draws	after 1 iteration	.98	21.5	−3.948	1.117
Interval 1,024, 3,000 draws	after 1 iteration	.71 to .75	340.6	−3.962	1.076
Default (lecture)	after 7 iterations	.52 to .59	237.1	−3.961	1.083

Report the longer chain. Both fits say converged, both pass the Geweke tests, and the coefficients differ only in the second digit, so the coefficients could not have told you. The effective sample size of about 22 draws among 3,000 does.

Lengthening the interval is the remedy: the longer chain costs more run time but carries about fifteen times the information.

To read after the session



Read Chapter 21 from Step 4 (transitivity) to the end for Week 6

Left over: Ch. 21 Ex. 5 (three tools for one question) and Ex. 8 (predict)

Reread the first part of the chapter and the interpretation boxes

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

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