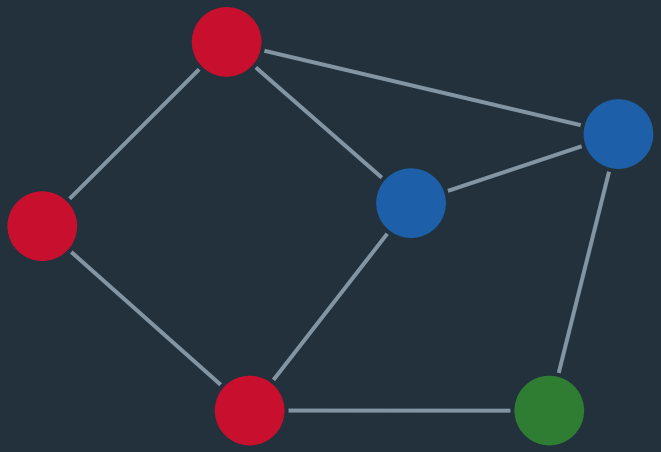
Permutation tests and QAP regression

Course A, Week 11 · Course B, Week 4

Chapters 19 and 20 · a lecture and a lab

One deck for both courses

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



Today: a lecture, a break and a lab

0:00	**1 Permutation tests: relabel the nodes**	35 min
0:35	**2 QAP regression: several predictors, one model**	45 min
1:20	Break	10 min
1:30	**3 Lab: permutation tests**	40 min
2:10	**4 Lab: QAP regression**	45 min

Two steps: association, then association net of the rest

Chapter 19: does advice go with friendship?

Is a relationship between two networks, or between a network and an attribute, stronger than relabeling the nodes would produce?

Chapter 20: and net of office, practice, status?

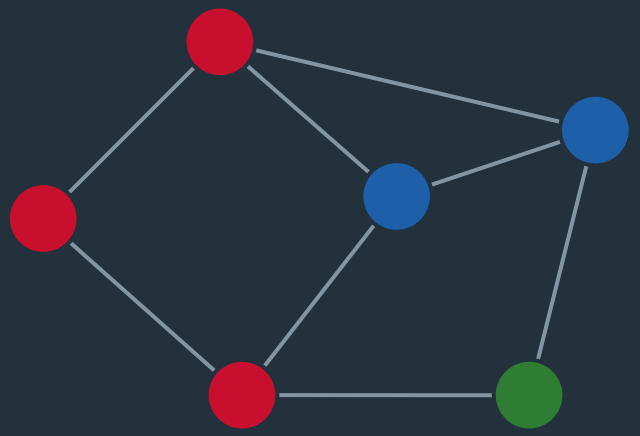
Does one network predict another once several other networks and attributes are held fixed in the same model?

	Course A, Week 11	Course B, Week 4
Follows	Week 10: baselines and the CUG test	Week 3: null models (Chapter 18)
Stress	Significant is not predictive accuracy	Odds ratios, fitted probabilities, what QAP leaves out
Leads to	Week 12: the QAP model in the capstone	Week 5: ERGMs, where ties depend on ties

1

Permutation tests: relabel the nodes

35 min · Chapter 19: why ordinary tests fail, relabeling, nwpermute, nwcorrelate, nwmixing



Two claims: advice goes with friendship, and with office

Advice goes with friendship

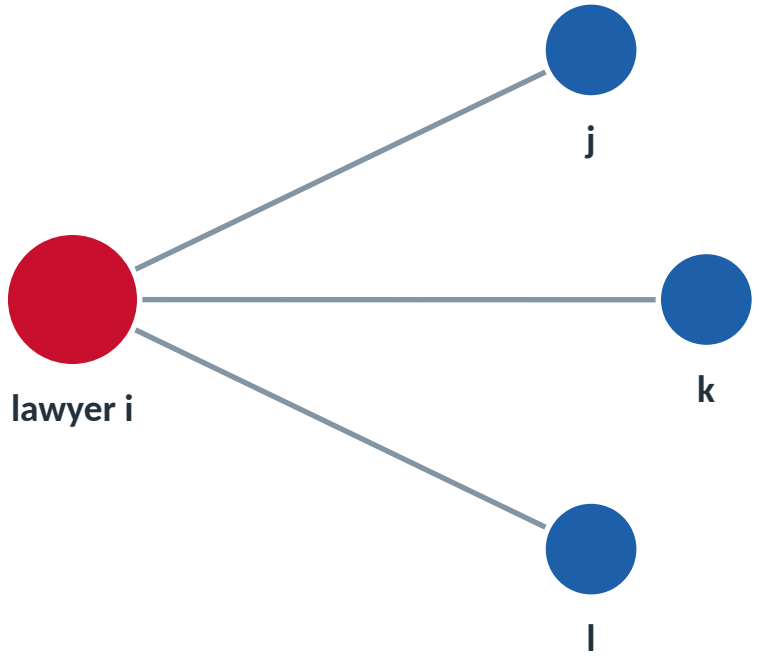
Lazega's cooperation among peers: advice is embedded in friendship. Compares two networks on the same 71 lawyers.

Office predicts who asks whom

The homophily argument. Compares a network with a node attribute. Three quarters of the advice arcs lie inside an office.

Each comparison must say what it preserves and what it randomizes. Preserved: the observed network, with all its internal dependence. Randomized: which lawyer is which.

4,970 rows are not 4,970 pieces of evidence



The dyads (i,j), (i,k) and (i,l) share the lawyer i.

- A sociable or senior lawyer raises the probability in every dyad that involves her at once.
- Fed into a correlation or a regression, 4,970 ordered pairs count as 4,970 separate pieces of evidence, when 71 actors generate them.
- The effective amount of information is smaller, so the uncertainty is understated, sometimes severely.

Relabeling keeps the structure, breaks the link to identity

Property	Random-graph CUG test	Node-label permutation
Number of nodes	preserved	preserved
Number of ties	preserved under density conditioning	preserved
Degree sequence	generally changes	preserved up to relabeling
Reciprocity	may change (fixed under census)	preserved
Triadic structure	may change	preserved
Alignment with another network or attribute	randomized by generating new networks	randomized by relabeling

Use the CUG test for a feature of one network. Use relabeling when the claim is about how two observed structures line up and each should keep its own shape.

One relabeling, by hand

```
. set seed 12345
. nwpermute lazega_fr, generate(fr_perm)
. nwcorrelate lazega_adv lazega_fr
    .41767433
. nwcorrelate lazega_adv fr_perm
    .00295191
```

Correlations as printed in a one-column table. Chapter 19, node-label permutation.

The relabeled network is as clustered and as reciprocal as the original. Only its alignment with the lawyers' other relations is gone.

.418

advice with friendship

.003

advice with the relabeled friendship

- **Everything else is unchanged**

  fr_perm has the same 575 arcs, density .116, reciprocity .148 and transitivity .186 as lazega_fr; only the link between identity and position is broken.

Correlating two networks: nwcorrelate with permutations()

```
. set seed 12345
. nwcorrelate lazega_adv lazega_fr, permutations(1000)

 Correlation     P-value   [95% permutation interval]
-------------------------------------------------------
   .41767433      0.0020      -.03147169   .03327747
```

.418

observed correlation

±.03

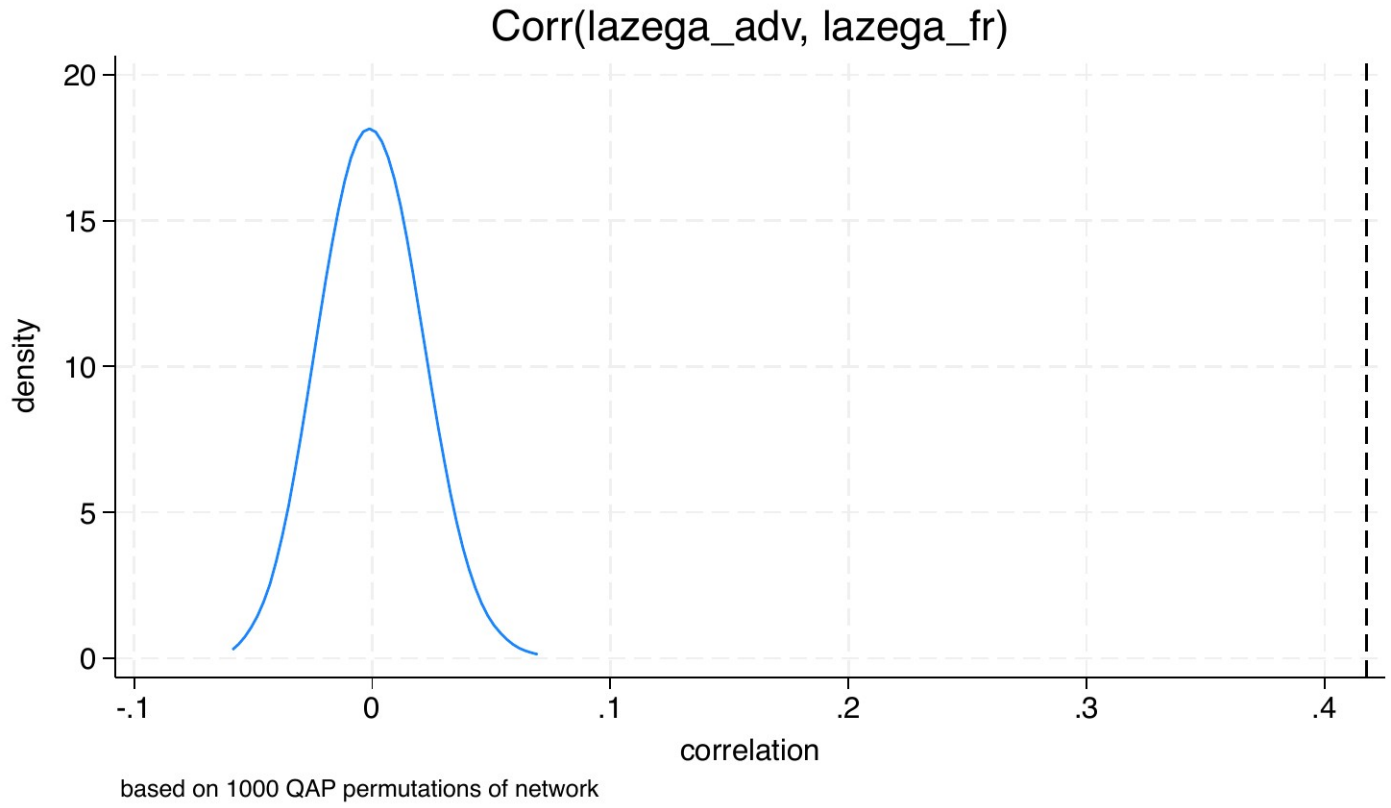
95% of 1,000 relabelings lie between −.031 and .033

p = .002 is the floor for 1,000 relabelings. The correlation is computed over all eligible cells, including the many pairs with neither relation, so .42 does not mean 42% of advice ties are friendship ties.

- **Size and surprise are different**

  The correlation gives the size of the association in the observed networks; the p-value says how unusual it is among relabelings.

The whole permutation distribution



1,000 relabelings of the friendship network against advice; dashed line at the observed .418. Chapter 19.

Centered on zero

Relabeling destroys the alignment, so the permuted correlations fall between −.03 and .03.

The line is far outside

None of the 1,000 relabelings reached .418: $p \le 2/1001$.

What the correlation says, and what it does not

It says

Lawyers linked by friendship were much more often linked by advice too than a random alignment of the two networks would give. None of the 1,000 relabelings matched .418; the observed arrangement counts as one more draw, with the smaller tail doubled (2/1001).

It does not say

That the probability is zero: 2/1001 is the floor for 1,000 draws. Nor that friendship caused advice or the reverse.

- **A bivariate association**

  It cannot tell a direct link between friendship and advice from a shared cause. Lawyers of one office may both befriend and consult each other because they share a floor. Partialling out a third network or attribute is the subject of block 2.

A network against an attribute: nwmixing

```
. nwmixing lazega_adv, attribute(office) permutations(1000)
   E-I Index: -.511   p-value: .002
. nwmixing lazega_adv, attribute(status) permutations(1000)
   E-I Index: -.152   p-value: .002
(tables of the arcs by office and status omitted)
```

The p-value of the E-I index is this permutation test. With the default 100 relabelings the floor was .02; with 1,000 it is .002.

Attribute	Arcs inside	Relabeling	E-I
Office	76%	52.5%	−.511
Status	58%	49%	−.152

The reference point is what relabeling gives, not zero: an E-I of about −.05 for office, +.014 for status. Office sorting is strong, status sorting modest.

What may be exchanged is part of the design

Exchangeability

Relabeling treats every lawyer as interchangeable with every other. Office, status and practice make them differ, and relabeling also destroys the alignment with those attributes.

A statement about this firm

The 71 lawyers are the whole firm at one time, not a sample. The p-value says how unusual the alignment is among relabelings of this very network.

Too few permutations

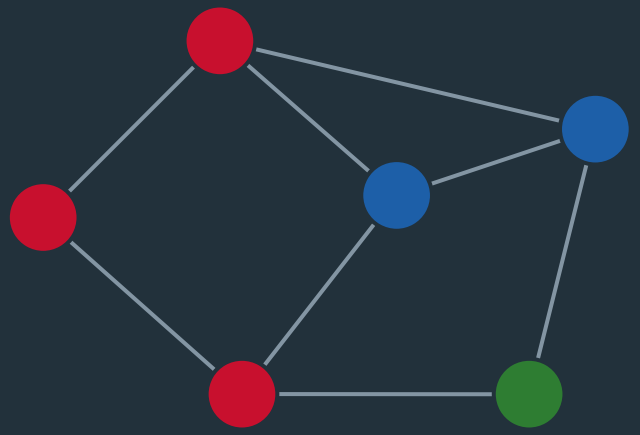
The two-sided p cannot be smaller than 2/(R+1): .002 for 1,000, about .02 for 100. A value at the floor means more are needed if a finer value matters.

The firm so far: advice and friendship correlate at .418 against a permutation range of −.031 to .033; 76% of advice arcs are inside an office against 52.5% under relabeling. Is friendship still there once office, practice and status are in the same model?

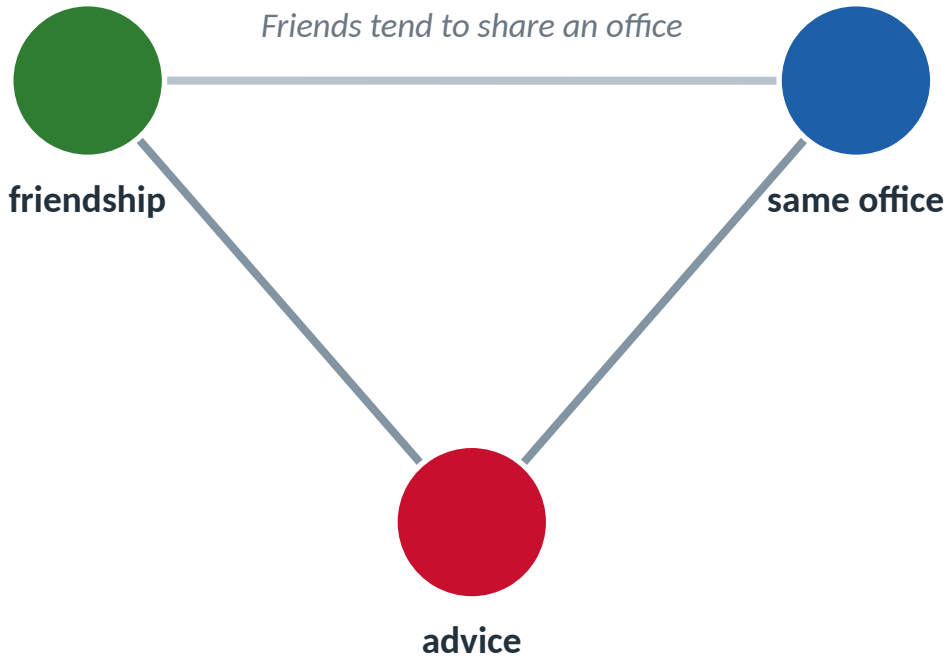
2

QAP regression: several predictors, one model

45 min · Chapter 20: nwqap, odds ratios, qapspp, mode(), type(regress), predict(), Glasgow



Friendship, or the office in disguise?



- Lawyers who are friends also tend to share an office, and lawyers who share an office tend to ask each other for advice.
- A bivariate correlation cannot separate the two: friendship may be the office in disguise.
- So: several predictors in one model, each coefficient read holding the others fixed.

A dyadic regression, and why its standard errors fail

$$\text{logit } \Pr(Y_{ij} = 1) = b0 + b1\, X_{ij} + b2\, Z_{ij}$$

Rows share lawyers

The 70 rows in which a given lawyer is the sender move together: a generally sociable lawyer raises the probability in all of them at once. The logit counts 70 independent pieces of evidence.

Predictors correlate

Friends tend to work in the same office, so a coefficient can borrow strength from, or be confused with, a neighbor (Dekker, Krackhardt and Snijders 2007).

Each dyad is one observation; Y is 1 if there is a tie, X and Z are dyadic predictors. The point estimates of an ordinary logit are fine as summaries; the standard errors are not.

QAP keeps the estimates and replaces the inference

1 Fit the logit

Advice on friendship and three same-value indicators. The point estimates stay.

2 Relabel the outcome

Permute the rows and columns of the dependent network together, keep the predictors, refit the whole model on each permutation.

3 Compare

Each coefficient with its own permutation distribution. The p-value counts both tails around its mean; the observed value is one more draw.

The smallest p-value is 1/(R+1) for R permutations. The dependence between dyads that share a lawyer stays in the reference distribution, because whole networks are relabeled (Krackhardt 1987, 1988).

Predicting advice with nwqap

```
. set seed 12345
. nwqap lazega_adv lazega_fr office practice status, ///
>         permutations(500) plot

  Regression            =  logit
  Permutations          =  500
  Number of vertices    =  71
  Number of arcs        =  892

            lazega_adv |      Coef.    P-value
-----------------------+------------------------
             lazega_fr |      2.198     .002
           same_office |      1.062     .002
         same_practice |       .752     .002
           same_status |       .407     .002
                 _cons |     -3.228
```

Attributes become dyads

office, practice and status turn into same_office, same_practice and same_status by the default mode(same).

4,970 rows

The 71-node directed network is reshaped into dyad-level observations.

Odds ratios from the printed coefficients

Predictor	Coefficient	Odds ratio	p
Friendship tie (lazega_fr)	2.198	9.0	.002
Same office	1.062	2.9	.002
Same practice	.752	2.1	.002
Same status	.407	1.5	.002

exp(2.198) ≈ 9.0, exp(1.062) ≈ 2.9, exp(.752) ≈ 2.1, exp(.407) ≈ 1.5.
Arithmetic from the printed coefficients: in Stata, display exp(2.198).

9.0

times the odds of an advice tie: friendship, holding office, practice and status fixed

- **Crude odds ratio: 11.9**

  From the cross-tabulation (358 × 3,861)/(217 × 534). Odds ratios from tables and from models with other predictors are not on a common scale.

What the coefficients say, and what they do not

It says

A friendship tie from i to j multiplies the odds of an advice tie from i to j by about 9.0, holding office, practice and status fixed (p = .002). The three shared attributes multiply the odds by about 2.9, 2.1 and 1.5.

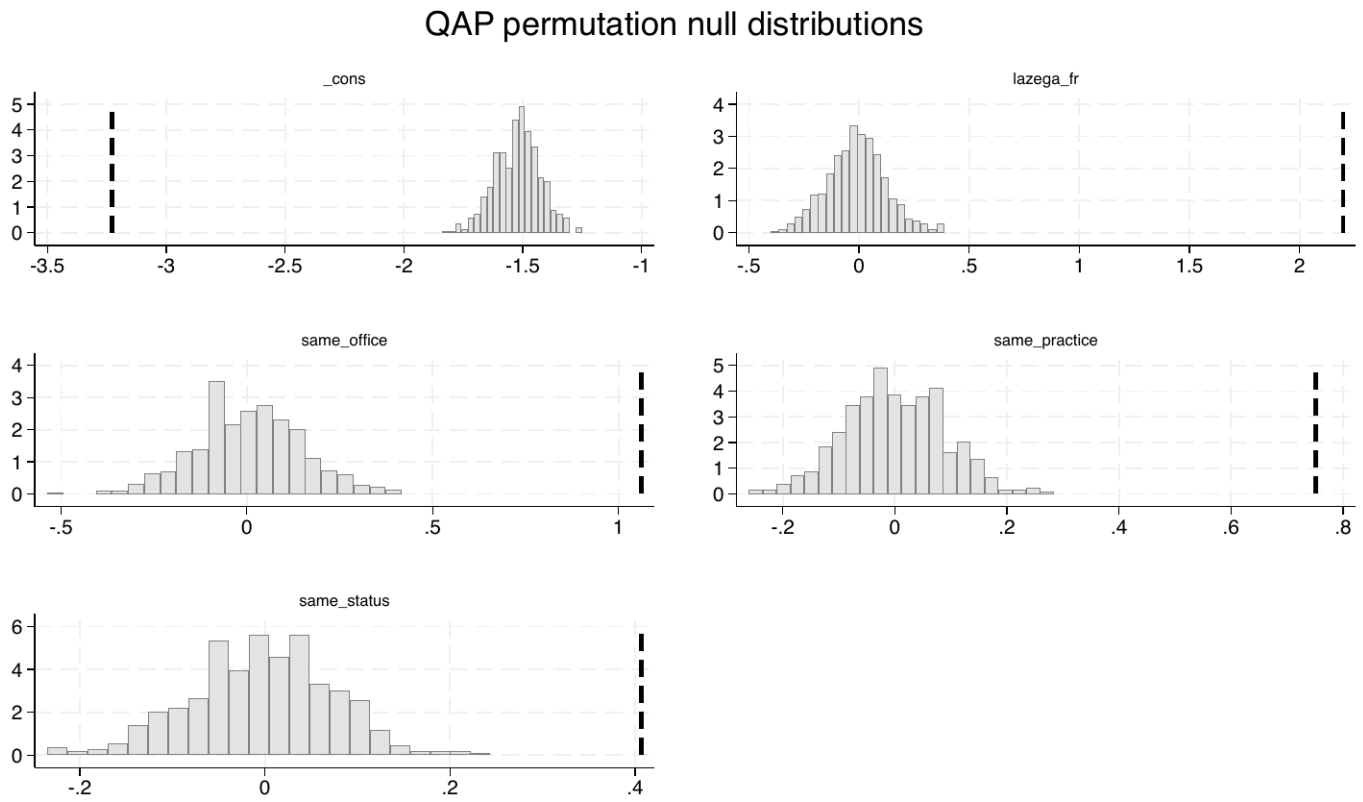
It does not say

How the effects add up for a specific pair: these are odds ratios, not probability ratios. Nor that friendship causes advice.

- **A bivariate test would borrow**

  Separate bivariate tests would answer “is this predictor related to advice?” without holding the others fixed; because friends share offices, each association would be partly borrowed from another predictor.

The permutation distributions behind the p-values



One panel per coefficient, 500 permutations, dashed lines at the observed values. Chapter 20.

Outside every histogram

The dashed line sits outside each panel: what p = .002 looks like.

Status is the closest call

The permuted coefficients reach about .25; the observed .41 lies only a little beyond.

Correlated predictors: double semi-partialling

```
. set seed 12345
. nwqap lazega_adv lazega_fr office practice status, ///
>         permutations(500) qapspp

  Estimation            =  QAP, double semi-partialling

             lazega_fr |      2.198     .002
           same_office |      1.062     .002
         same_practice |       .752     .002
           same_status |       .407     .002
```

Coefficient rows only; the rest of the table is as before. Chapter 20.

Same coefficients, same conclusion. Only the test differs: for these data the result does not depend on the permutation scheme.

Plain QAP

Permutes the whole outcome network: some association, not strictly net of the others.

qapspp

Permutes each predictor's residuals from the others (Dekker et al. 2007): much less sensitive to correlated predictors.

From a nodal attribute to a dyadic predictor: mode()

mode()	Dyadic variable	Use for
`same (default)`	1 if v_i = v_j	A categorical attribute: pooled homophily (office, practice)
`dist`	v_i − v_j	A signed difference
`absdist`	|v_i − v_j|	A continuous or ranked attribute where only the size of the gap matters
`sender / receiver`	v_i / v_j	Directed main effects: does the lawyer who asks, or is asked, have a characteristic?

The choice is part of the research question. A preference for partners and a preference for similar seniority are different ideas; the same indicator measures neither. With several attributes, mode() takes one entry per attribute, in order.

Receivers, seniority gaps and a valued outcome

```
. nwqap lazega_adv lazega_fr partner seniority, ///
>         mode(receiver absdist) permutations(500)

              lazega_fr |      2.472     .002
          receiver_partner |       .647     .002
        absdist_seniority |      -.017     .002
                    _cons |     -1.955
```

```
. nwgen strength = lazega_adv + lazega_cow + lazega_fr
. nwqap strength office practice status, ///
>         type(regress) permutations(500)

            same_office |       .477     .002
          same_practice |       .391     .002
            same_status |       .178     .002
                  _cons |      -.018
```

Friendship: 2.47 here, 2.20 before. A coefficient controls only for what is in the model; the seniority gap also stands in for office and practice, which are left out.

mode(receiver absdist)

A partner has about 1.9 times the odds of an associate of being asked (exp(.647)). Per rank step exp(−.017) ≈ .98; over a gap of 50 ranks ≈ .43.

type(regress)

Same office: .48 more of the three relations on average; practice .39; status .18. A linear summary of a bounded count.

Significant is not the same as accurate: predict()

```
. nwqap lazega_adv lazega_fr office practice status, ///
>         permutations(500) predict(adv_fitted)
. nwtoedge adv_fitted lazega_adv
. gsort -adv_fitted _ego _alter
. tabulate lazega_adv if adv_fitted == adv_fitted[1]

 lazega_adv |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         54       33.54       33.54
          1 |        107       66.46      100.00
      Total |        161      100.00
```

.767

highest fitted probability, shared by 161 dyads

66%

actual advice ties in that group: 107 of 161

All four coefficients are at the floor, yet the model overstates the top group by about ten points. It is an additive logit: the same effects are added for every combination. This shows miscalibration in the fitted data; it says nothing about new data.

An earlier wave as a predictor: the Glasgow girls

```
. nwwebuse glasgow, nwclear
. set seed 12345
. nwqap glasgow2 glasgow1 smoke1 sport1, permutations(200)

             glasgow2 |      Coef.    P-value
----------------------+-------------------------
             glasgow1 |      3.653     .005
           same_smoke1 |       .514      .03
           same_sport1 |       .217     .378
                 _cons |     -4.125
```

Prior friendship is by far the strongest predictor. Only 200 permutations: the p-values are coarse, and one wave of attributes cannot separate selection from influence.

Predictor	Odds ratio
Wave-1 friendship	38.6
Same smoking	1.67
Same sport	1.24

With qapspp: smoking p = .06

Not printed in the book (same seed, 200 permutations): the coefficients are the same, but the semi-partialled p-value of same_smoke1 is .06, not .03.

What QAP does not model

Ties that depend on ties

There is no term for “this tie is more likely because two other ties in the same network exist”: a returned request, a triangle closed.

Left in the error

Reciprocity and closure may load onto predictors that correlate with them. Friendship is the likeliest candidate: friends also return each other's requests.

The next step

Exponential random graph models put reciprocity, closure and degree into the model as terms to estimate (Chapter 21).

If the question is only whether friendship and the formal organization go with advice, QAP with double semi-partialling is enough, and the write-up should say that it takes no position on how ties depend on one another.

What to take away from the lecture

Relabel, do not regenerate

Permuting node labels preserves the observed structure and destroys the link between identity and position. The p-value concerns the data in hand.

Same estimates

QAP keeps the logit's point estimates and judges them against permutations. Report the routine, mode() and the number of permutations.

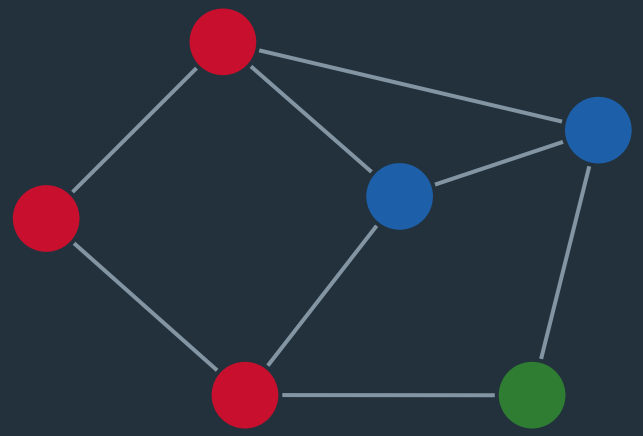
Significant is not accurate

A model can have every coefficient at the floor and still miscalibrate a group of pairs. predict() shows it.

The firm: net of office, practice and status, a friendship tie goes with about 9.0 times the odds of an advice tie; the shared attributes with 2.9, 2.1 and 1.5. Advice overlaps with the formal organization without being reducible to it.

Break

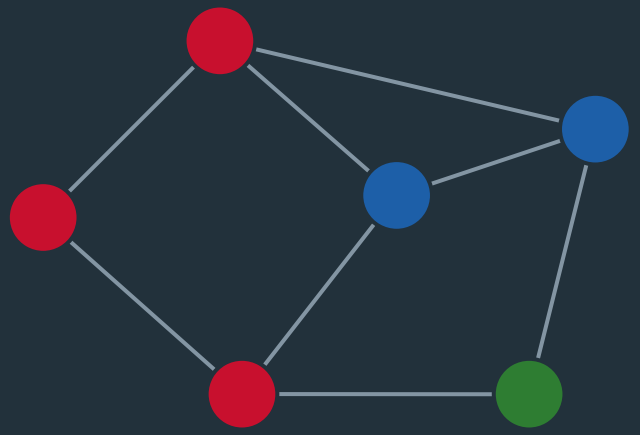
10 minutes



3

Lab: permutation tests

40 min · Run along: nwpermute, nwcorrelate, nwmixing; Ch. 19, Ex. 2, 4 and 5



Run along 1: relabel, correlate, test

```
. nwwebuse lazega, nwclear
. set seed 12345
. nwpermute lazega_fr, generate(fr_perm)
. nwcorrelate lazega_adv lazega_fr
. nwcorrelate lazega_adv fr_perm
. set seed 12345
. nwcorrelate lazega_adv lazega_fr, permutations(1000)
. nwmixing lazega_adv, attribute(office) permutations(1000)
. nwmixing lazega_adv, attribute(status) permutations(1000)
```

Predict first. What will the correlation of advice with the relabeled friendship network be? Then press return.

A mismatch in the interval usually means a missing set seed; a missing network means nwwebuse was not run.

Checkpoint	Value
advice with friendship	.41767433
advice with fr_perm	.00295191
p-value, 1,000 relabelings	0.0020
permutation interval, lower	-.03147169
permutation interval, upper	.03327747
E-I index, office (p)	-.511 (.002)
E-I index, status (p)	-.152 (.002)

Hands-on (20 min)

Ch. 19, Ex. 2

With permutations(1000), what is the smallest two-sided p-value that nwcorrelate can report when the observed correlation is more extreme than every relabeling? Compare it with the p-value printed for the advice–friendship correlation of 0.418, and say what that printed value does and does not tell you about how unusual 0.418 is.

Hint Count the observed arrangement as one more draw, 1,001 in all, and remember that nwcorrelate doubles the smaller tail.

Ch. 19, Ex. 4

Test the correlation of the advice network with the co-work network lazega_cow: set seed 12345, then nwcorrelate lazega_adv lazega_cow, permutations(1000). Report the correlation, the 95% permutation interval and the p-value as an inequality. Larger or smaller than advice and friendship (0.418)? Which would you expect to overlap more with advice, and why?

Hint When no relabeling reaches the observed value, write $p \le 2/1001$ and name the number of permutations. For the last question, think about what the lawyers ask one another for advice about.

Ch. 19, Ex. 5

Is advice sorted by gender? Compute from the group sizes (53 men and 18 women; tabulate gender) the E-I index that relabeling alone would give. Then set seed 12345 and run nwmixing lazega_adv, attribute(gender) permutations(1000). Report the observed E-I index and its p-value: homophily, heterophily or neither?

Hint The share of ordered pairs within a group is $(53{\cdot}52 + 18{\cdot}17)/(71{\cdot}70)$; the E-I index expected under relabeling is one minus twice that share. Compare with this value, not with zero.

Chapter 19: what the exercises give

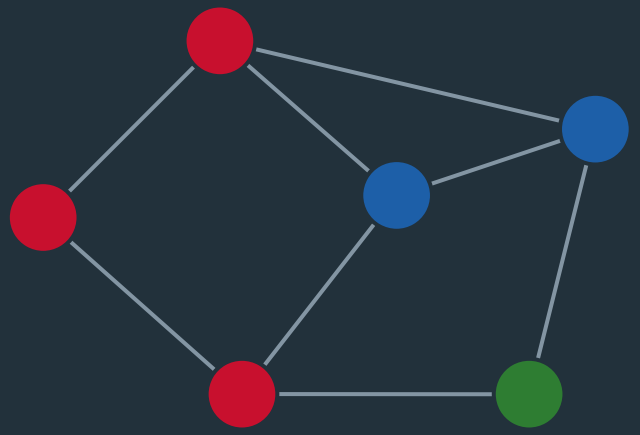
Exercise	Result	Reading
Ch. 19, Ex. 2	smallest two-sided p = 2/1001 = .001998; printed .0020	A value at the floor says only that no relabeling reached .418, not how far in the tail it lies.
Ch. 19, Ex. 4	r = .484, interval −.035 to .034, p ≤ 2/1001	Larger than advice and friendship (.418): lawyers ask the colleagues they work with.
Ch. 19, Ex. 5	expected E-I −.232; observed −.395, p = .004	More ties within a gender than relabeling gives: homophily. Not at the floor.

Compare with what relabeling gives, not with zero. For gender the reference point of the E-I index is −.232 (61.6% of ordered pairs lie within a gender), so the excess is smaller than the raw −.395 suggests.

4

Lab: QAP regression

45 min · Run along: nwqap, qapspp, mode(), type(regress), predict(); Ch. 20, Ex. 1, 2, 4, 5 and 9



Run along 1: the advice model, plain and semi-partialled

```
. nwwebuse lazega, nwclear
. set seed 12345
. nwqap lazega_adv lazega_fr office practice status, ///
>         permutations(500) plot
. set seed 12345
. nwqap lazega_adv lazega_fr office practice status, ///
>         permutations(500) qapspp
```

Predict first. Do the coefficients change when qapspp is added?

With permutations(100) the p-values print as .01 (1/101 = .0099); the coefficients are the same.

Checkpoint	Value
lazega_fr	2.198
same_office	1.062
same_practice	.752
same_status	.407
_cons	-3.228
every p-value (500 permutations)	.002
header of the second run	QAP, double semi-partialling

Run along 2: mode(), a valued outcome and predictions

```
. set seed 12345
. nwqap lazega_adv lazega_fr partner seniority, ///
>         mode(receiver absdist) permutations(500)
. nwgen strength = lazega_adv + lazega_cow + lazega_fr
. nwqap strength office practice status, ///
>         type(regress) permutations(500)
. nwqap lazega_adv lazega_fr office practice status, ///
>         permutations(500) predict(adv_fitted)
. nwtoedge adv_fitted lazega_adv
. gsort -adv_fitted _ego _alter
. tabulate lazega_adv if adv_fitted == adv_fitted[1]
```

Predict first. How many different fitted probabilities can a model with four binary predictors give at most?

Checkpoint	Value
receiver_partner	.647
absdist_seniority	-.017
strength: same_office	.477
strength: same_practice	.391
strength: same_status	.178
highest fitted value	.7667885
dyads with that value	161 (54 / 107)

Hands-on (10 min)

Ch. 20, Ex. 1

From the printed coefficients of the four-predictor advice model compute the fitted probability of an advice tie for two kinds of dyad: (a) a friendship tie, the same office and the same practice, but a different status; (b) the same dyad without the friendship tie. By how many percentage points does the friendship tie raise the probability for this kind of dyad, and why is that difference not the same for every kind of dyad, although the coefficient is?

Hint Add to the constant the coefficients of the predictors that equal 1: for (a), −3.228 + 2.198 + 1.062 + .752. Then turn the log-odds x into a probability with 1/(1 + e^−x), in Stata display invlogit(x).

Ch. 20, Ex. 2

A QAP logit gives a coefficient of 0.5 for a predictor. Compute the odds ratio. If the p-value is based on 200 permutations and the observed coefficient is farther from the mean of the permutation distribution than every permuted one, what is the smallest p-value that nwqap can report? Find it in the Glasgow table.

Hint The odds ratio is e^0.5. For the p-value, use the floor of nwqap, 1/(R + 1) with R permutations; unlike nwcorrelate, nwqap does not double it.

Chapter 20, Ex. 1 and 2: the answers

Dyad	Log-odds	Probability
(a) friends, same office, same practice, other status	.784	.687
(b) the same without friendship	−1.414	.196
no friendship, nothing shared	−3.228	.038
friendship, nothing shared	−1.030	.263

(a) minus (b): 49 points. Friendship alone: 22.5 points. The coefficient adds the same log-odds everywhere; the probability moves most where the dyad starts near the middle of the logistic curve.

Ex. 2: odds ratio

exp(.5) ≈ 1.65. A coefficient of .5 multiplies the odds of a tie by about 1.65.

Ex. 2: the floor

1/(200 + 1) = .005: the p-value of glasgow1. nwqap does not double the tail, unlike nwcorrelate.

Hands-on (15 min)

Ch. 20, Ex. 5

Fit the advice model without lazega_fr: set seed 12345, then nwqap lazega_adv office practice status, permutations(500). Report the three coefficients next to those of the four-predictor model, and say in which direction they moved and why that direction is expected. Why can the size of the change not be read as the share of the friendship–advice association that the attributes carry?

Hint Ask how friends are spread over offices, practices and status groups, and what an attribute coefficient can pick up when friendship is no longer in the model. For the last question, reread the paragraph after the interpretation box on comparing odds ratios across models.

Ch. 20, Ex. 4

Fit the Glasgow model twice more, each time after set seed 12345 and with permutations(200): once with sport1 dropped (nwqap glasgow2 glasgow1 smoke1) and once with smoke1 alone. Report the coefficient of same_smoke1 and its p-value in all three models. Does holding wave-1 friendship fixed change the smoking coefficient much? What would a large change have told you that a small one does not?

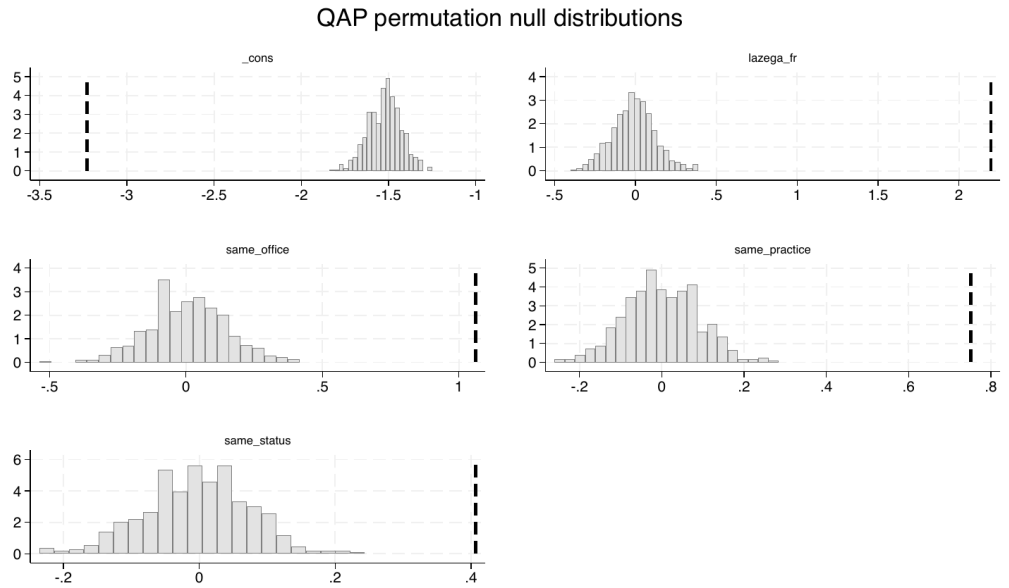
Hint Setting the same seed before each call makes the differences come from the model, not from the draws. Compare the model with smoke1 alone with the one that also contains glasgow1.

Hands-on (5 min)

Ch. 20, Ex. 9

A colleague fits an ordinary logit to all 4,970 dyads of the advice network and reports its standard errors as they come. Write a short paragraph that explains what is wrong with those standard errors, and say which output of nwqap, and of its option plot, you would show a skeptical reader to make the point.

Hint The section on the dyadic regression says why the rows are not independent observations; the section on visualizing the permutation distribution shows the distributions that replace the standard errors.



What to show the skeptic: the plot option of nwqap. Chapter 20.

Chapter 20, Ex. 5, 4 and 9: the answers

Ex. 5: coefficient	with lazega_fr	without
same_office	1.062	1.261
same_practice	.752	.957
same_status	.407	.437

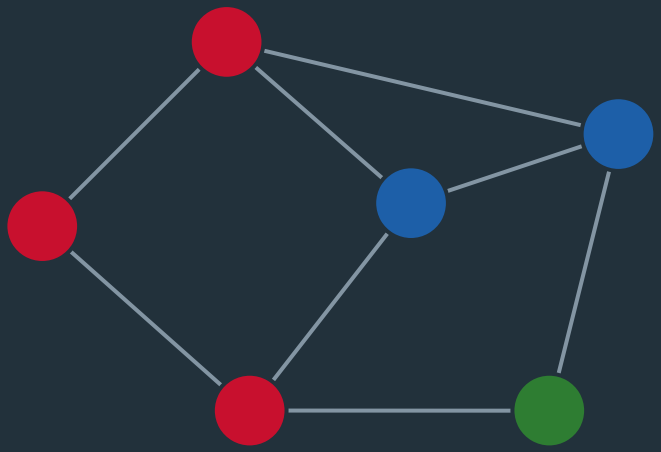
All p = .002; every coefficient moves up. Friends share offices, practices and status, so without friendship the attributes pick up part of its effect. The change is not a share: odds ratios across models are not on a common scale.

Ex. 4: same_smoke1 in	coef.	p
glasgow1 + smoke1 + sport1	.514	.03
glasgow1 + smoke1	.496	.035
smoke1 alone	.544	.025

Holding wave-1 friendship fixed changes little. A large change would have meant that shared smoking mostly marked earlier friendships; a small one, that it does not. Neither separates selection from influence.

Ex. 9 The 70 rows in which a lawyer is the sender share that lawyer, so an ordinary logit counts dependent rows as independent evidence and its standard errors are too small. Show the QAP p-values (refits on permuted networks) and the permutation histograms of the plot option.

To read after the session



Course A: read Chapters 29 and 30 for Week 12

Course B: read Chapter 21, from the start through the reciprocity step and the MCMC check, for Week 5

Left over: Ch. 19 Ex. 1, 3, 6; Ch. 20 Ex. 3, 6, 7, 8

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

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