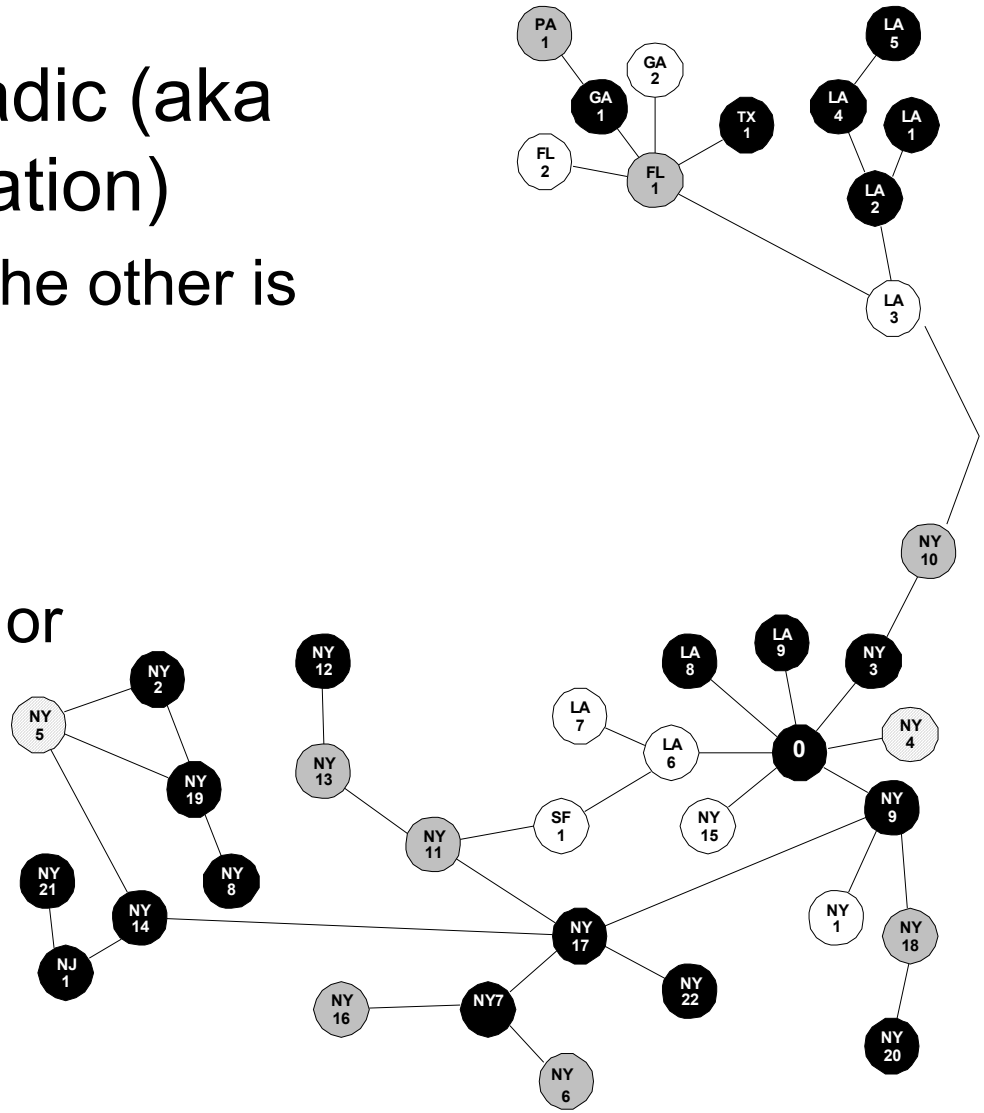
Testing Network Hypotheses

Units of Analysis

- Dyadic
  - The raw data
  - Observations are pairs of actors
  - Proximity leads to interaction which leads to friendship
  - Padgett's multiplexity argument
  - Granovetter's transitivity assumption
- Monadic
  - Aggregated
  - Observations are actors
  - Central people get promoted faster, resist colds better, etc.
- Group
  - Aggregated
  - Observations are groups of actors (e.g., teams, orgs)
  - Centralized teams perform well on simple tasks

Mixed Dyadic-Monadic

- Mixed dyadic/monadic (aka network autocorrelation)
  - One var is dyadic, the other is monadic
  - Explaining ties via actor qualities
  - testing for diffusion or influence



Statistical Issues

- Samples non-random
- Often work with populations
- Observations not independent
- Distributions unknown

Solutions

- Non-independence
  - Model the non-independence explicitly as in HLM
    - Assumes you know all sources of dependence
  - Permutation tests
- Non-random samples/populations
  - Permutation tests

Logic of Permutation Test

- Compute test statistic (e.g., corr or mean diff)
- Construct a distribution of test statistics for null hypothesis
  - E.g., compute thousands of random variables, similar to real vars, and correlate them
- Count proportion of statistics on random variables that are as large as the observed
  - This is the p-value of the test

Where the permutations come in

- When we construct the random variables that we compare against, how to ensure that they are just like your actual variables?
  - Same distribution of values etc.
- Either by sampling from the values you actually have (jack-knife) or by permuting the values of the variables randomly

Permutation vs Classical

Permutation Test

- Tests independence of variables
- Usable with very small datasets

Classical Test

- Tests independence of variables AND random sample AND normality of population distribution
- Results generalize to population

Node-level & Whole Network Level

- Use ordinary permutation test

Dyadic Hypotheses

- QAP correlation & MR-QAP multiple regression
  - All variables are actor-by-actor matrices
  - We use one relation (dyadic variable) to predict another

Friendship

	Jim	Jill	Jen	Joe
Jim	-	1	0	1
Jill	1	-	1	0
Jen	0	1	-	1
Joe	1	0	1	-

Proximity

	Jim	Jill	Jen	Joe
Jim	-	3	9	2
Jill	3	-	1	15
Jen	9	1	-	3
Joe	2	15	3	-

Dyadic/Monadic Hypotheses

- Categorical
  - Join-count statistic (dichotomous vars only)
  - Generalized chi-square (both categorical)
  - Anova (cat monadic, any dyadic)
- Interval
  - Moran's I
  - Geary's C