PLSC 504 - Topics in Political Methodology

Exercise Eight
November 15, 2022

Overview

The subject of this exercise is network analysis. We'll only be doing a "real data" part this time; there is no simulation component.

The data for this exercise are drawn from the 2019-2020 session of the Pennsylvania State Senate; the specific topic – one that is a mainstay of network analysis in political science – is bill cosponsorship. Our data are from LegiScan, which tracks bills in the U.S. Congress plus all 50 state legislatures in more-or-less real time.

In the 2019-20 session, senators introduced 1741 bills into that chamber; during that same session, a total of 53 senators served in the chamber.[1] This means there were a total of $1741 \times 53 = 92273$ opportunities for senators to (co-)sponsor legislation. The data available in the "Exercises" folder on the course Github repository are named `PLSC504-2022-ExerciseEight.csv` and contain information on those 92273 opportunities. More specifically, each row is the data contain the following variables:

- `people_id` – a variable indicating the Legiscan identifier for each senator.
- `name` - the senator's name.
- `chamber` – a variable indicating the chamber in which the legislator served (in this instance, uniformly equal to "`Sen`").
- `district` – a character variable indicating the (Senate) district that legislator represented.
- `party` – a character variable, coded "`D`" if that senator was a member of the Democratic party, and "`R`" if that senator was a member of the Republican party.
- `bill_id` – the numeric Legiscan code identifying the bill in question.
- `sponsor` – a binary variable indicating whether (=1) or not (=0) the legislator identified in `people_id` cosponsored the bill identified in `bill_id`.

[1] There are 50 senators in the Pennsylvania Senate; owing to deaths / resignations / etc., three Senate districts – the 33rd, the 41st, and the 48th – had two senators serve during the 2019-20 session.

Exercise

1. Use the data provided to create two *adjacency matrices* for networks of cosponsorship patterns in the PA Senate. The first (call that matrix A) should indicate the presence or absence of cosponsorship among the 53 senators, while the second (matrix B) should indicate the extent of that cosponsorship (by recording the *number* of bills each pair of senators cosponsored with each other).

2. Using matrix A:

   (a) Plot the (undirected) network of cosponsorship among the senators;

   (b) Re-plot the network, distinguishing legislators by their political party (you might do this with labels, colors, or whatever you prefer);

   (c) Calculate the degree, closeness, betweenness, and eigenvector centrality scores of the nodes in the network.

3. Repeat the steps in part (2) using the undirected network defined by matrix B.

4. Finally, fit an ERGM model to the undirected network defined by matrix B, where the principal covariate is whether (=1) or not (=0) the two legislators in question were of the same political party.

As is typically the case, this homework exercise is worth 50 possible points. It is due by 5:00 p.m. EST on Wednesday, November 30, 2022, and should be submitted electronically – via e-mail attachment – to Michael (`tzs5636@psu.edu`) *and* to me (`zorn@psu.edu`).