Homework 4: Permutation Testing

CSC 495, Prof. Robin Burke
Spring 2015
Due: 5/21/2015[1]

Objective:

To gain experience with CUG and QAP tests.

What to do:

1. Download the archive file `hwk4.zip` and uncompress. This archive contains several graph data files: `dolphin.graphml`, `krack-advice.graphml`, and `krack-friend.graphml`. It also contains two utilities which are my reimplementations of the CUG and QAP routines from the SNA library: `mycugtest.R` and `myqaptest.R`. There is also a starter file `hwk4.R` that loads the utilities.
2. Run RStudio.

Part I: Dolphin Network

3. Load the dolphin network from `dolphin.graphml`. This is an undirected network. The Sex attribute is 0 for female, 1 for male, and 2 for a few dolphins where this variable was not known.
4. Compute the (global) transitivity of this graph. It should come out around 0.31.
5. Our first task is to test if this transitivity is significantly different from what would be expected in a random network of the same density. In other words, can we reject the null hypothesis that the level of transitivity here could be due to chance edge formation?
    a. Use the `mycugtest` routine to test the transitivity of the graph against density-conditioned random neworks. You will need to pass "edges" as the cmode parameter and you'll need to supply type= "global" so that transitivity will be calculated correctly each time.
    b. Use `print.cug.text` and `plot.cug.test` to examine the results of the test.
    c. Can we reject the null hypothesis (p=0.05)? Why or why not?
6. Next, test the degree centralization of the network. Recall that centralization is the extent of concentration of degree in nodes in a network.
    a. Define a function that takes a graph, computes degree centralization and returns just the centralization value. (Remember that `centralization.degree` returns a named list.)
    b. Using this function, compute a CUG test for degree centralization.
    c. Use `print.cug.text` and `plot.cug.test` to examine the results of the test.
    d. What is the null hypothesis in this test? Can it be rejected? Why or why not?
7. Next, test the assortativity of the network by sex. (Are male dolphins more likely to be found in the company of other male dolphins?)
    a. You need to use `assortativity.nominal`, but note: the `types` parameter for `assortativity.nominal` does not allow zero as a value. You will need to add 1 to the `Sex` attribute before passing it to `mycugtest`.
    b. Use `print.cug.text` and `plot.cug.test` to examine the results of the test.
    c. What is the null hypothesis in this test? Can it be rejected? Why or why not?
8. Using your degree centralization function from step 6, perform a QAP test on the dolphin network.
    a. Use the `summary.qaptest` and `plot.qaptest` to examine the results.
    b. Explain why QAP is not a valid test for degree centralization. How do the results demonstrate this?
9. Next, test the assortativity of the dolphin network by sex using the QAP method.
    a. Use the `summary.qaptest` and `plot.qaptest` to examine the results.
    b. What is the null hypothesis in this test? Can it be rejected? Why or why not?

[1] Due dates are before class: 5:30 pm CDT

Part II

10. In this part, we will look at the Krackhardt data from Lab 1. I have already created the networks and stored them in graphml files.
11. Load the advice and friend graphs.
12. Make the friend graph undirected by collapsing the edges (weak version).
13. Use the leading eigenvector method to find communities in the undirected network.
14. Plot the friend network with the edges colored according to their community membership.
15. Using the CUG test, examine the assortativity of the advice network by the friend communities. (Are the managers more likely to ask advice from those in their friend group?) As before, state the null hypothesis and what your test finds relative to that hypothesis.
16. Using the CUG test, examine the assortativity of the advice network by the tenure attribute.
    a. Note that this is scalar assortativity, so you will use the `assortativity` function, not `assortativity.nominal`. The `assortativity` function requires two parameters: `types1` and `types2`, because it is possible to use different values on each end of a directed edge. For this question, set `types1` to be the tenure attribute and `types2=NULL`.
    b. What is the null hypothesis in this test? Can it be rejected? Why or why not?
17. Repeat the analysis in step 15, but using the QAP test instead.
18. Repeat the analysis in step 16, but using the QAP test instead.
19. Create an R markdown file hwk4.Rmd and copy your R code into it with appropriate annotations. Include your answers to the hypothesis testing questions in steps 5-9 and 15-18.
20. Run `hwk4.Rmd` and save the resulting HTML as `hwk4.html`.
21. Combine `hwk4.R`, `hwk4.Rmd`, and `hwk4.html` into a zip archive. No need to include the utility files or the data files.

What to turn in

Upload your zip archive to the Homework 4 dropbox on D2L before class on 5/21.

Rubric

Step 5, 1 pt (0.5 mechanics, 0.5 explanation)
Step 6, 1 pt
Step 7, 1 pt
Step 8, 1 pt
Step 9, 1 pt
Step 14, 1 pt
Step 15, 1 pt
Step 16, 1 pt
Step 17, 1 pt
Step 18, 1 pt