Week 16 Lab Exercises

Philip Leifeld

GV903 Advanced Research Methods
University of Essex, Department of Government

Pietryka and DeBats (2017) have found previously lost 19th century individual voting records and social network data for two municipalities through archival research. They combine both to test whether the proximity of voters to party elites in the social network among residents plays a role in election turnout and whether it makes them vote favourably for the respective party.

The replication data for this article are available here: `https://doi.org/10.7910/DVN/UTAP1V/SC5DLE`. The replication file reads in the tab-separated data using the function `read_tsv` in the `readr` package. This creates a "tibble", the `tidyverse` equivalent of a data frame. To avoid problems in the tasks below, convert the tibble into a data frame using the `as.data.frame` function first. Please read the article carefully and complete the following tasks:

1. Replicate the results shown in Table 1 of the article without using the `%>%` operator or the `dplyr` package. It is acceptable if the table is not formatted in exactly the same way as shown in the article because the journal uses a journal-specific table style. Show both the `R` code and the table.

2. Show the equations for Akaike's Information Criterion and for the Bayesian Information Criterion. Insert the values for Models 1, 2, and 3 for Alexandria, and manually calculate the results (without the use of `R`, but based on the information from Table 1 in the article; this should produce six values). Explain what the two criteria do and how they are used, and interpret the values.

3. Interpret the coefficient for *social proximity to elites* as odds-ratios, both in terms of the substantive meaning and the size of the effect, for both models in which it is featured (i. e., Model 3 in both cases). You may use simple mathematical functions to aid the interpretation, but no predicted probabilities or marginal effects.

4. Compute and plot the predicted probabilities of voting conditional on the effect of social proximity to elites (similar to Figure 2a in the paper, but for the full range of the proximity variable) in Model 3 for Alexandria for a typical voter, separately for a home-owner and a non-home-owner, including 95 per cent confidence intervals. Use `ggplot2` for the plot. Show the code and the resulting figure. Pick two points on the curves and interpret them substantively (manually, approximately). Also interpret the overall finding of the two curves.

5. Based on Models 2 and 3 for Newport in Table 1, manually calculate a likelihood ratio test, showing all necessary equations and the result, including critical value, test statistic, and p value. Explain what you did, and interpret the result.

6. Show receiver operating characteristics for Models 1, 2, and 3 for Newport. Compute the area under the curve. How much improvement in terms of model fit do we get each time? Show the `R` code and explain the results and improvement.

References

Pietryka, M. T. and DeBats, D. A. (2017). It's not just what you have, but who you know: Networks, social proximity to elites, and voting in state and local elections. *American Political Science Review*, 111(2):360–378.