Week 6 Lab Exercises

Philip Leifeld

GV903 Advanced Methods – University of Essex, Department of Government

Use RStudio with either LaTeX or Markdown and `knitr` to write up your answers for the following tasks, and embed your `R` code in the text document.

Background: Two health policy debates in the UK

Fergie et al. (2019), Buckton et al. (2019), and Hilton et al. (2020) conducted a content analysis of two contentious policy debates in the health sector. They first analysed the policy debate on minimum-unit pricing for alcohol (MUP), which was implemented in Scotland in May 2018 (Fergie et al., 2019). Then they analysed the debate on the soft drinks industry levy (SDIL), informally known as the "sugar tax", which was implemented UK-wide in April 2018 (Buckton et al., 2019). Finally, they compared both policy debates (Hilton et al., 2020). Their approach was to hand-code the statements of political actors about various arguments and policy beliefs in national newspapers. There are 237 policy actors, who revealed their positions on a number of arguments related to the two policies.

Ideological positions of actors and cross-issue advocacy

In a new research paper, the same authors now want to analyse the determinants of the ideological positions of the 237 actors. They apply a Bayesian item–response theory (IRT) model, which generates estimates of the ideological ideal point of each actor on a latent underlying ideological dimension along with a highest-posterior density (HPD) interval, which is roughly the Bayesian equivalent of a 95 % confidence interval. For example, if an actor has an estimated ideal point of -0.87, we can derive that the actor is quite far in the pro-regulation camp, while for example a value of 1.04 would be a rather strong anti-regulation position. A value around zero would mean a neutral or ambiguous ideological position with regard to the two policies. You do not need to worry about how the ideal points were computed; focus instead on how these ideal points can be explained. The question is now what determines where an actor stands on this ideological scale, as measured by the IRT model. Your task is to estimate and interpret a regression model for this purpose.

The dataset and tasks

The dataset `healthscaling.csv` contains the names, ideological positions, the lower and upper bounds of the HPD interval, dummy variables for whether an actor is active in the SDIL debate and in the MUP debate, respectively, the number of statements the actor makes in each of the two policy debates, and the actor type. With these data, complete the following tasks:

1. Create a variable for the width of the HPD interval of the estimate and add it to the dataset. It expresses the extent to which an actor expresses a vague ideological position. You can call it "spread".

2. Estimate a linear regression model in `R` to explain the ideological position of actors. Explain the ideological position of actors using the spread of expressed policy beliefs, the type of actor, the overall frequency with which an actor is active (in both policy debates combined), and an indicator of whether an actor is active in both debates. We can suspect that actors who tend towards the extreme left or right end of the ideological scale may have wider HPD intervals because extreme actors may not be inclined to hold or express moderate beliefs and there are few very extreme beliefs. That is, we expect a non-linear, U-shaped effect of spread. Therefore you need to include a second-order polynomial effect for the spread variable to capture this non-linear relationship. Create the necessary new variables and add them to the dataset; estimate a linear model in `R`; report the results using a regression table with the `texreg` package; and show the code for these steps in your answer, along with a brief explanation of what you did.

3. Interpret the effects: Which actor types clearly tend towards the pro-regulation camp? Which actor types tend towards the anti-regulation camp? For which actor types can we not be sure? Where are European Union actors on the ideological scale? What else can we learn from the model about cross-policy involvement, actors' activity, and their vagueness of stated beliefs? Interpret the non-linear relationship by looking at the polynomial coefficients.

4. Evaluate the goodness of fit of the model using the F test and R^2 statistic reported under the linear model. What do these numbers tell us? Also calculate the R^2 statistic step by step in `R` and show the relevant equations with inserted values (where possible and where it does not take up whole pages) in your answer.

5. Use the `predict` and `plot` functions in `R` to interpret the functional form of the relationship between the spread/vagueness of actors' policy beliefs and their ideological position, keeping all other variables constant at typical or average values. Feel free to show either a single average scenario or multiple "typical" scenarios, and describe/interpret the effect graphically.

6. Re-compute the predictions manually, including confidence and prediction intervals.

References

Buckton, C. H., Fergie, G., Leifeld, P., and Hilton, S. (2019). A discourse network analysis of UK newspaper coverage of the "sugar tax" debate before and after the announcement of the Soft Drinks Industry Levy. *BMC Public Health*, 19(1):490. `https://doi.org/10.1186/s12889-019-6799-9`.

Fergie, G., Leifeld, P., Hawkins, B., and Hilton, S. (2019). Mapping discourse coalitions in the Minimum Unit Pricing for alcohol debate: A discourse network analysis of UK newspaper coverage. *Addiction*, 114(4):741–753. `https://doi.org/10.1111/add.14514`.

Hilton, S., Buckton, C. H., Henrichsen, T., Fergie, G., and Leifeld, P. (2020). Policy congruence and advocacy strategies in the discourse networks of Minimum Unit Pricing for alcohol and the Soft Drinks Industry Levy. *Addiction*. Forthcoming: `https://doi.org/10.1111/add.15068`.