Some Notes on Modelling Examples

Theodore Dounias

September 17, 2018

Comprehensive Smoking Bans, Barr et al.

This paper studies how smoking bans implemented across 9 US States and 387 Counties effected hospital admission rates for Acute Myocardial Infraction among Medicare enrollees. This paper follows a series of studies that use a wide variety of methods, and tend to show a decrease in admissions. The dataset studied in this specific instance included over 9 million individuals aged 65 and over–as Medicare is a program for the elderly–and their records of admission, sex, specific age, race, and county/state of residence. The authors also included open source data on smoking bans gathered by some advocacy groups. The largest modelling issues that they noted was that many models failed to compensate for a secular trend towards reduction in AMI admittances, or failed to include some factor to compensate for seasonality in the data. A series of random effects poisson regression models was then fit to the data; the state level model that they most articulate the parametrization of includes, apart from typical variables on admission rate, age, and gender, state level effects and county level mixed effects for the intercept, for the time variable, and for when the ban was implemented, and a state-level indicator for each month to adjust for seasonality. After evaluating their models, they reach the conclusion that when relaxing assumptions about linearity of the secular trend behind AMI reduction, the reduction was reduced to zero in all but the highest age group.

The two modelling issues solved here are probably present in CO data as well. There may exist a non-linear secular trend in turnout that I will have to compensate for. This trend may transcend election type (primary, general, coordinated), or county, or voting method, and I will have to find some way for the model to reflect this or risk reaching conclusions for VBM based on other underlying effects. Maybe this can be compensated for by doing multi-level modelling for election type instead of just including two dummy variables. I will also have to be very careful on the modelling of the time variable. Maybe I should add national effects to the model somehow? Particularly for general elections, this might help to reflect national shocks on turnout. In terms of modelling time, my study seems very different to this particular one; here the division of years into months seems useful, as maybe in summer months AMI is more likelly due to weather. However, I am not sure the same circularity exists in elections. I will keep the idea of including a state level effect and a county specific effect, which when dealing with one state is essentially just including mixed effects for counties.

Barr, Christopher D., David M. Diez, Yun Wang, Francesca Dominici, and Jonathan M. Samet. "Comprehensive Smoking Bans and Acute Myocardial Infarction Among Medicare Enrollees in 387 US Counties: 1999–2008." American Journal of Epidemiology 176, no. 7 (October 1, 2012): 642–48. https://doi.org/10.1093/aje/kws267.

Modeling Missing Respons in W Jones Thesis–Chapter 5

This is interesting for my thesis because I have missing data of various "flavors". The first I can think of is missing data on voting method. This obviously falls under the non ignorable category, since it is the variable whose effect I am directly modelling. If it was independent of the response it means my thesis hypothesis is wrong. Whether an individual voted by mail or not is also dependant on other variables; age, race, or gender should play a role, particularly in earlier elections. Also, well, if VBM was offered in the first place.

My other missing data problem comes from missing voters who dropped off the rolls, and I do not think that this algorithm can take care of it, since I quite literally have not data for them.

The thesis contains code for the missing data EM algorythm in Appendix A. I can use that!

Jones, Will. Senior Thesis on Statistics of Routes. Contribute to Wjones127/Thesis Development by Creating an Account on GitHub. HTML, 2017. https://github.com/wjones127/thesis.