

STA 365: Applied Bayesian Statistics

Boris Babic
Assistant Professor, University of Toronto

Week 9B: Hierarchical Modeling Lab



- Consider the following common situation: it's 2016, and you are interested in predicting who will win the US Presidential election, Donald Trump, or Hillary Clinton.
- You are going to take a large national poll to make your prediction. Suppose $n = 100,000$.
- You ask each person whether they will vote for Trump or not. Suppose $X = 1$ denotes voting for trump, and suppose that after conducting this survey you learn that $Y = \sum_{i=1}^n X_i = 30,000$
- But, you suspect that not everyone is telling the truth. There are some people who claim they wouldn't vote for trump, even though they actually will when they get to the polls. For example: closet Republicans in New York City.
- Likewise, there are some people who say they will vote for Trump, even though they actually will not when they get to the polls. For example: closet Democrats living in Texas.
- Your (open-ended) challenge is to construct a model that reasonably captures this situation.

- First, you want to think about the likelihood. If there was no possibility of error, then the $Y \sim \text{Bin}(n, p)$. How should we incorporate the possibility of error? (hint: consider another parameter(s) for which you will specify a prior)
- Second, you need to specify your priors.
- Third, you need some data, which in this case is simple: $n = 100,000$ and $Y = 30,000$.
- Fourth, you will fit the model, and obtain posterior estimates.
- Fifth, you will diagnose and then summarize the model.
- Sixth, draw some conclusions now about the possibility that Trump wins, and discuss your modeling and prior assumptions, and how they affect your interpretation of the data.