

Does Urbanisation Predict Election Outcomes?

A Bayesian's Perspective

Eldaleona Odole Leonor Cunha Anarghya Murthy

Faculty of Statistics

January 18, 2025

Table of Contents

- 1 Introduction
- 2 Dataset Description
- 3 Model Setup
- 4 Priors
- 5 Results - Pooled Model
- 6 Results - Unpooled Model
- 7 Results - Hierarchical Model
- 8 Results - Varying Intercept Model
- 9 Centred vs Non-centred Parameterisation/ Sampling Investigation with STAN
- 10 Raw References

Introduction

- **Research Question:** How does urbanization of a particular district affect result of an election in the US?
- **Variable of Interest:** Winning party in the House of Representatives 2022 General Election (binary)

Dataset

We wanted to consider different factors in the analysis, with our primary focus being the urbanization of each House district. These factors included:

- 1 Demographic Data (US Census Bureau)
- 2 Urbanization (FiveThirtyEight)
- 3 Regional Information (US Census Bureau)
- 4 Election Results (FiveThirtyEight)

We combined different sources in order to create our data set containing 435 instances of 16 unique covariates.

Model Assumptions

There are many people trying to predict US election outcomes, from the wealth of data available about voters. However we wanted to look at the voters in relation to their geography. In order to do this we assumed

- District voting outcomes can be modeled via logistic regression
- Districts are exchangeable within each state and each state is exchangeable within its region
- ???

Model

Let the response variable 'Winning Party' be y , the predictor of interest 'Urban index' be x , and the other covariates be a 15-dimensional vector z . Let i , j , and k be the indices for the district, region, and state respectively.

$$y_{i,j} \sim \text{Ber.}(\text{logit}^{-1}(\theta_j))$$

$$\theta_j := \beta_{0,j} + x_{i,j} * \beta_{1,j} + z_{i,j}^T * \gamma_{1,j}$$

Factor name

Urbanization Index

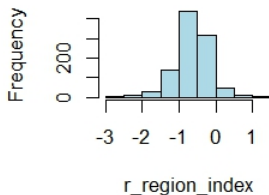
Urbanization Index is a measure created by fivethirtyeight calculated based on populatin density of a given district.



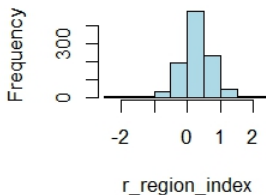


Varying Intercept Model

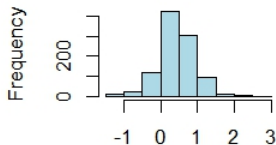
Region 1 Histogram



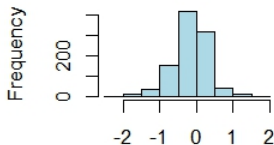
Region 2 Histogram



Region 3 Histogram



Region 4 Histogram



Varying Intercept Model - II



STAN Code

Raw references

- stargazer
- tidybayes
- brms, stan