The MRP model for voting libreay party in 2019 Canada Election

Xiaomeng,Liu

Github Link: https://github.com/liu273/STA304-Final-Project.git

Abstract

Canada Election Survey, which is a study to research the selections and opinions of Canadians during and after the 2019 federal election. In this research, it uses the data files of phone surveys and multilevel regression with poststratification model to predict the proportion of the people voting to liberal party, which is the party that won in the 2019 Canada federal election. (Stephenson,p4)

Keywords

1.Survey data; 2.Canada Election; 3. Politics ; 4. Public Opinion; 5.Phone survey.

Introduction

The 2019 Canadian federal election was held on October 21, 2019, which elects members of the House of Commons to the Canadian Parliament. The Liberal Party won 157 seats to form a minority government in this election. (Wikimedia Foundation, 2)

In order to determin what influence the people's attitude to vote to Liberal Party, I selected 3 predictors(province, sex and age) which use to create a model to describe the whether or not these predictors effect their opinions of voting and predict the proportion of voting of liberal party in the 2019 election..

This paper gathers the data from two datasets, ces2019 phone survey and anadian General Social Survey(GSS). The dataset of ces2019 phone survey is the one file of Canada Election Survey which provides a set of data that gathers Canadians' attitudes and opinions on a variety of social research, and to make the data publicly available to researchers in sociology, economics, communications. (Stephenson,p4). And the GSS program conducts telephone surveys for all non-institutionalized persons 15 years of age and older, in the ten Canada provinces. The study's object are gathering the data on social trends,and to development of scientific knowledge for some social policy issues of current or emerging interest.(General Social Survey,P3)

Methodology:

-Data: The model follows the binomial distribution, since the dependent variable (whether or not people decide to vote to liberal party in the 2019 Canada election) in the research is binary. Since we should select common variables in GSS dataset and ces2019 phone survey dataset to do the model analysis. The following predictors (independent variables) in this model were used Age. It used to describe the age of the voters in the 2019 Canada federal election, which must be higher than 18 years old, and the mean of all people in the survey is 50.89.

Sex. This predictor is binary, since in this model, just consider the situation of the people is male or female.

Province. This predictor I selected after compared the two model's AIC(model_1 only has predictors of sex and age, model_2 has one more predictor, province.). Since model_2's AIC has less value which means it can fit the model better. And from Table1, shows that some provinces,such as Manitoba, Newfoundland ,Labrador and Prince Edward Island have the p-value which is smaller than α, the predictors are significant, then I keep it in the final model.

-Model:

Table 1: table1

variables	estimate	sd_error	p_value
Manitoba	0.973	0.261	1.87e-04
Newfoundland and Labrado	1.367	0.263	2.00e-07
Prince Edward Island	1.377	0.263	2.00e-07
age	0.007	0.002	2.00e-03

Table 2: table2

predict_prop	actural_prop
0.242	0.231

This paper used the model is Multilevel regression and poststratification (MRP) ,which is a technique for estimating subnational when the samples are large nationals. In MRP, the parameters apply to different levels, and in this research, age and sex are level 1, and the province is level 2. (M Downes, p12)The post stratification's formula is

$$\hat{y}^{PS} = \frac{\sum N_j \hat{y}_j}{\sum N_j}$$

, where $\hat{y}_j$ is the estimate in each cell(provinces), and N_j is the population size of the jth cell (provinces) based on demographics.

```
## # A tibble: 1 x 1
##   predict
##     <dbl>
## 1   0.242
```

Results: We estimate that the proportions of voters in favour of voting for liberal party in the 2019 federal election will be 0.242. This is based on the post-stratification analysis of the proportion of voters in favour of liberal party modelled by a MRP model, which accounted for age,sex and province.(Samantha-Jo Caetano, p3) Then compared with the actual probability of the people voting to liberal party in the election is 0.231,we can conclude that the model is fitted.

Discussion: -Summary This paper uses the multilevel logistic regression model and the post stratification technique to predict the possibility of voting for liberal party. At the beginning, we had the hypothesis that the proportions of voting of liberal party is related to voters' age,sex and province. Then,we use the raw datasets of GSS and CES , cleaned and manipulated the census data and survey data. The next steps is constructing the MRP model to determine how the variables of age, sex and province can effect Canadain's intention in voting to liberal party in the election.

-Conclusion Based on the MRP model, it predicts that the proportions of voters in favour of voting for liberal party in the 2019 federal election is 0.242 according to the variables of voters' age, sex and provinces.Since the prediction are really close to the actual proportion of the Canadian who decides to vote to liberal party in this election. This is a useful model that we can use to predict more results of Canadian elections in future.

-Weakness&Next step

Although most situations for Canadian elections seem like adopting this model.There is a biggest weakness in the MRP model is the predictors are too few. This model only considers 3 variables that may affect the people's selection in the 2019 Canada Election, which means it may be a loss of success rate of prediction. In the next step, we should consider more situations and predictors to build the MRP model.

References: <1> Stephenson, Laura B, Western University Harell, Allison, Université du Québec à Montréal Rubenson, Daniel, Ryerson University Loewen, Peter John, University of Toronto. Canadian Election Study, 2019, Phone Survey ,Study Documentation. Scholars Portal (SP) , Ontario Council of University Libraries, 2020.

<2> "2019 Canadian Federal Election." Wikipedia, Wikimedia Foundation, 17 Dec. 2020, en.wikipedia.org/wiki/2019_Canadian

<3> General Social Survey, Cycle 31 : Families Public Use Microdata File Documentation and User's Guide. Rep. Ottawa, 2020. Print.

<4> M Downes. An Introduction to Multilevel Regression and Post-Stratification for Estimating Constituency Opinion. 9 Apr. 2018.

<5> Samantha-Jo Caetano. STA 304 Final Project – Additional Instructions,2020.

<6> "Welcome to My.access – Please Choose How You Will Connect." My.access - University of Toronto Libraries Portal, sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/sdaweb/html/gss.htm.