Problem Set 2: Applied Stats II

Luna Goldstein

Due: February 28, 2022

Instructions

- Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in R, please include the code you used to get your answers. Please also include the `.R` file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.

- Your homework should be submitted electronically on GitHub in `.pdf` form.

- This problem set is due before class on Monday February 28, 2022. No late assignments will be accepted.

- Total available points for this homework is 80.

We're interested in what types of international environmental agreements or policies people support (Bechtel and Scheve 2013). So, we asked 8,500 individuals whether they support a given policy, and for each participant, we vary the (1) number of countries that participate in the international agreement and (2) sanctions for not following the agreement.

Load in the data labeled `climateSupport.csv` on GitHub, which contains an observational study of 8,500 observations.

- Response variable:

  - `choice`: 1 if the individual agreed with the policy; 0 if the individual did not support the policy

- Explanatory variables:

  - `countries`: Number of participating countries [20 of 192; 80 of 192; 160 of 192]
  - `sanctions`: Sanctions for missing emission reduction targets [None, 5%, 15%, and 20% of the monthly household costs given 2% GDP growth]

Please answer the following questions:

1. Remember, we are interested in predicting the likelihood of an individual supporting a policy based on the number of countries participating and the possible sanctions for non-compliance.

   Fit an additive model. Provide the summary output, the global null hypothesis, and p-value. Please describe the results and provide a conclusion.

   R code:

   ```
   1  reg <- glm(choice ~ .,
   2      data = climateSupport,
   3      family = "binomial")
   4  summary(reg)
   ```

   ### The summary output

Intercept	-0.14458
	(0.00137**)
Countries	0.32436
	(0.0000000000000002***)
sanctions	-0.12353
	(0.000000000315 ***)

   Interpreting the results:

   i. Since all estimated coefficients are significant we are able to reject the null hypothesis. expected odds that judge sides with more feminist decision

   ii. The intercept represent the expected odds that an individual supports a given policy. On average an individual supports a given policy = exp(-0.14458) = 0.8653857.

   iii. Increasing the number of countries increases the log odds of supporting a given policy by 0.32436 and is a statistically reliable relationship.

   iv. Increasing the number of sanctions decreases the log odds of supporting a given policy by 0.12353 which is a statistically reliable relationship.

2. If any of the explanatory variables are significant in this model, then:

   (a) For the policy in which nearly all countries participate [160 of 192], how does increasing sanctions from 5% to 15% change the odds that an individual will support the policy? (Interpretation of a coefficient)

   For 5% : exp((-0.144558) + (0.32436*countries) + (-0.12353*sanctions))
   For 15% : exp((-0.144558) + (0.32436*countries + (-0.12353*sanctions))

We perceive a 0.17 increase in odds when increasing the sanctions from 5% to 15% the individuals are more likely to support the climate policy.

(b) For the policy in which very few countries participate [20 of 192], how does increasing sanctions from 5% to 15% change the odds that an individual will support the policy? (Interpretation of a coefficient)

For 5% : exp((-0.144558) + (0.32436*countries) + (-0.12353*sanctions))
For 15% : exp((-0.144558) + (0.32436*countries + (-0.12353*sanctions))

We perceive a 0.0888 increase in odds when increasing the sanctions from 5% to 15% the individuals are more likely to support the climate policy.

(c) What is the estimated probability that an individual will support a policy if there are 80 of 192 countries participating with no sanctions?

exp((-0.144558) + (0.32436*countries) + (-0.12353*sanctions)/(1+(exp(-0.144558) + (0.32436*countries) + (-0.12353*sanction))

The estimated probability that an individual will support a policy if there are 80 of 192 countries participating with no sanctions is 55%.

(d) Would the answers to 2a and 2b potentially change if we included the interaction term in this model? Why?

- Perform a test to see if including an interaction is appropriate.

```
interaction <- glm(choice ~ . + countries:sanctions,
            data = climateSupport,
            family = "binomial")
summary(interaction)
```

The answers would not change the interaction is not significant. Including an interaction in this scenario is not necessary.