Problem Set 2

Applied Stats II

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]

Alt method: Load package 'mgcv' and use gam function to run additive model

- library(mgcv)
- altadditivemodel ¡- gam(choice ., family = 'binomial', data = climateSupport)
- summary(altadditivemodel)
- ?gam

Conventional method: Use glm() function

- additivemodel ¡- glm(choice ., family = 'binomial', data = climateSupport)
- summary(additivemodel)

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.

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)

   R METHOD

   Try for 5 percent sanctions

   - highparticipation ¡- climateSupport[climateSupport(dollar)countries == '160 of 192',]

   - fivepercentdata ¡- highparticipation[highparticipation(dollar)sanctions == '5percent',]

   - fivepercentdata

   Use predict function with 'response' type to estimate probabilities

   - fivepercentprobabilities ¡- predict(additivemodel, newdata = fivepercentdata, type = "response")

   -summary(fivepercentprobabilities)

   ANS = Mean = 0.6382

   0.6382*100

   Probability = [1] 63.82percent

   100-63.82 ANS = [1] 36.18percent

   The odds ratio is... - 63.82/36.18 [1] Odds = 1.763958/1

   For the 160/192 policy, when 5percent sanctions are applied, the odds of support is 1.763958/1

   Now try for 15percent sanctions...

   - fifteenpercentdata ¡- highparticipation[highparticipation(dollar)sanctions == '15percent',]

   - fifteenpercentdata

Use predict function with 'response' type to estimate probabilities

- fifteenpercentprobabilities ¡- predict(additivemodel, newdata = fifteenpercentdata, type = "response")

-summary(fifteenpercentprobabilities)

ANS = mean = 0.5603

0.5603*100

ANS = Probability = [1] 56.03percent

- 100-56.03

- ANS [1] 43.97

The odds ratio is...

- 56.03/43.97

- ANS = [1] Odds = 1.274278/1

For the 160/192 policy, when 15percent sanctions are applied, the odds of support is 1.274278/1

Overall, the majority support of countries increases the odds of support for agreements. Sanctions also have an effect

In answer to the question posed, when 5percent sanctions rise to 15percent for the 160/192 policy, the odds of support declines from 1.76/1 to 1.27/1

There is a decrease in odds of support by 0.48968 when this is so... - 1.763958-1.274278

ANS = 0.48968

(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)

R METHOD

Try for 5percent sanctions

- lowparticipation ¡- climateSupport[climateSupport(dollar)countries == '20 of 192',]

- fivepercentdata2 ¡- lowparticipation[lowparticipation(dollar)sanctions == '5fivepercentdata2

Use predict function with 'response' type to estimate probabilities

- fivepercentprobabilities2 ¡- predict(additivemodel, newdata = fivepercentdata2, type = "response")

-summary(fivepercentprobabilities2)

ANS = mean = 0.4798

- 0.4798*100

ANS = Probability = [1] 47.98 percent

- 100-47.98

ANS = [1] 52.02

The odds ratio is...

- 47.98/52.02

ANS = [1] Odds = 1/0.9223376

For the 160/192 policy, when 5percent sanctions are applied, the odds of support is 1/0.9223376

Now try for 15percent sanctions...

- fifteenpercentdata2 ¡- lowparticipation[lowparticipation(dollar)sanctions == '15

- fifteenpercentdata2

Use predict function with 'response' type to estimate probabilities

- fifteenpercentprobabilities2 ¡- predict(additivemodel, newdata = fifteenpercentdata2, type = "response")

-summary(fifteenpercentprobabilities2)

ANS = mean = 0.3999

- 0.3999*100

ANS = Probability = [1] 39.99percent

- 100-39.99

ANS = [1] 60.1 percent

The odds ratio is... - 39.99/60.1

ANS = [1] Odds = 0.665391/1

For the 160/192 policy, when 15percent sanctions are applied, the odds of support is 0.665391/1

Therefore, when 5

Decrease in odds of support by 0.25

- 0.92-0.67

ANS = 0.48968

Overall, the odds of support for an agreement with only 20 particpating countries is always less likely than the odds that it recieves no support, regardless of sanction percentages.

Therefore, in conjunction with results from question (a), country participation is likley an explanatory variable of support. However, the effect of sanctions (from 5-15 percent) does decrease support further, much-like in question (a)

Sanctions therefore might have an effect also.

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

Create estdata dataframe to make predictions

- nosanctions ¡- climateSupport[climateSupport(dollar)sanctions == 'None',]
- estdata ¡- nosanctions[nosanctions(dollar)countries == '80 of 192',]
- estdata

Use predict function with 'response' type to estimate probabilities

- estprobabilities ¡- predict(additivemodel, newdata = estdata, type = "response")
- estprobabilities
- summary(estprobabilities)

ANS = mean = 0.5159

Estimated probability = 0.5159 or 51.59/100percent

Approx. a 52 percent chance that an individual will support a policy if there are 80 of 192 countries participating with no sanctions

(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.

Create interaction model by multiplying x1 (countries) by x2 (sanctions)

- interactionmodel ¡- glm(choice countries*sanctions, family = 'binomial', data = climateSupport)

-interactionmodel(dollar)coefficients

R METHOD

FOR PART (a) USING INTERACTION MODEL

Use highparticpation, fivepercentdata and fifteenpercentdata objects from question (a)

Use predict function with 'response' type to estimate probabilities

- qainteractprobabilities ¡- predict(interactionmodel, newdata = fivepercentdata, type = "response")

- summary(qainteractprobabilities)

ANS = mean = 0.6433

- 0.6433*100

ANS = Probability = [1] 64.33percent

100-64.33

ANS = [1] 35.67percent

The odds ratio is...

- 64.33/35.67

[1] Odds = 1.803476/1 using interaction model

This is marginally greater than 1.763958/1 odds observed in part (a)

Now try for 15percent sanctions on question (a)...

Use predict function with 'response' type to estimate probabilities

- qainteractprobabilities2 ¡- predict(interactionmodel, newdata = fifteenpercentdata, type = "response")

-summary(qainteractprobabilities2)

ANS = mean = 0.5472

- 0.5472*100

ANS = Probability = [1] 54.72percent

- 100-54.72

ANS = [1] 45.28

The odds ratio is...

54.72/45.28

[1] Odds = 1.208481/1 using interaction model

This is slightly lower than 1.274278/1 odds observed in part (a)

Therefore, when 5percent sanctions rise to 15percent for the 160/192 policy using an interactive model, the odds of support still declines from 1.803476/1 to 1.208481/1

- 1.803476-1.208481

Decrease in odds of support by 0.594995.

This is in comparison with a lesser effect seen in part a at 0.48968

Overall, results mixed as intercation increases support marginally with low sanctions but lessens support slightly for high sanctions. Results also quite close to original model without intercation. Regardless, there is an overall effect on the odds when including the interaction as support was decreased by a greater margin than original model when sanctions applied.

FOR PART (B) USING INTERACTIVE MODEL

Use lowparticpation, fivepercentdata and fifteenpercentdata objects from question (b)

Use predict function with 'response' type to estimate probabilities

- qbinteractprobabilities ¡- predict(interactionmodel, newdata = fivepercentdata2, type = "response")

-summary(qbinteractprobabilities)

ANS = mean = 0.4618

- 0.4798*100

ANS = Probability = [1] 46.18

- 100-46.18

ANS = [1] 53.82

The odds ratio is...

- 46.18/53.82

ANS = [1] Odds = 1/0.8580453

For the 160/192 policy, when 5percent sanctions are applied, the odds of support is 1/0.8580453

This is a decrease in odds from the results of question (b) 1/0.92

Now try for 15percent sanctions...

Use predict function with 'response' type to estimate probabilities

- qbinteractprobabilities2 ¡- predict(interactionmodel, newdata = fifteenpercentdata2, type = "response")

-summary(qbinteractprobabilities2)

ANS = mean = 0.4082

- 0.4082*100

ANS = Probability = [1] 40.82percent

- 100-40.82

ANS = [1] 59.18percent

The odds ratio is...

- 40.82/59.18

ANS = [1] Odds = 0.6897601/1

For the 160/192 policy, when 15percent sanctions are applied, the odds of support is 0.6897601/1

This is an increase in support from the odds of question (b) 1/0.67

Therefore, when 5percent sanctions rise to 15percent for the 20/192 policy using an interactive model, the odds of support still declines from 1/0.8580453 to 0.6897601/1

- 0.8580453-0.6897601

Decrease in odds of support by 0.1682852

These results using an interactive model differ from odds in part (b) of 1/0.92 and 1/0.67

However, the results here are still very similar. Main difference is that effect of sanction rise on support is lower at 0.1682852

Overall, the effect of including an intercation term in the model effects the answers to part a and b, but also slightly.

This indicates that an intercation term may not be neccesary as its effect was not so large.