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]

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.

QUESTION 1 - MY ANSWER

```
1
2   library(stargazer)
3
4 load(url("https://github.com/ASDS-TCD/StatsII_Spring2022/blob/main/
      datasets/climateSupport.RData?raw=true"))
5
6 str(climateSupport)
7 names(climateSupport)
8
9
10 # Question 1
11 #
12 # fit an additive model. provide the summary output,
13
14 levels(climateSupport$sanctions)
15
16 climateSupport$sanctions <- relevel(factor(climateSupport$sanctions,
      ordered = F), ref = "5%")
17
18 climateSupport$sanctions
19
20 # fitting the additive model
21
22 c_logit <- glm(choice ~ countries + sanctions, family = "binomial", data
      = climateSupport)
23 summary(c_logit)
24
25
26 # Coefficients:
27 #                  Estimate Std. Error   z value   Pr(>|z|)
28 # (Intercept)       0.24743    0.04406     5.616    1.95e-08 ***
29 #   countries.L     0.45845    0.03810    12.033    < 2e-16 ***
30 #   countries.Q    -0.00995    0.03806    -0.261    0.79374
31 # sanctionsNone    -0.19185    0.06216    -3.086    0.00203 **
32 #   sanctions15%   -0.32510    0.06224    -5.224    1.76e-07 ***
33 #   sanctions20%   -0.49542    0.06228    -7.955    1.79e-15 ***
34 #   ---
35 #   Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .
      0.1            1
```

```
36
37 # (Dispersion parameter for binomial family taken to be 1)
38
39 # Null deviance: 11783  on 8499  degrees of freedom
40 # Residual deviance: 11568  on 8494  degrees of freedom
41
42 # 11783 = null deviance and 11568 = residual deviance
43 # find the p-value for the chi-square test statistic
44 pchisq(11783-11568, 5, lower.tail = F)
45
46 # this gives
47 # 1.749304e-44
48
49
50
51 # logit model
52 # period functions as omnibus selector (the kitchen sink additive model)
53 # ~ . will select countries (ord) and sanctions (fct)
54
55 climate_logit <- glm(choice ~ ., family = binomial(link="logit"), data =
       climateSupport)
56 summary(climate_logit)
57
58 reg_exp <- exp(coef(climate_logit))
59 stargazer(reg_exp, type = "text")
60
61
62 # not supported or supported
63 #  the dependent variable is binary(0/1, True/False, Yes/No) in nature
64 t_glm <- glm(choice ~ 1, data = climateSupport, family=binomial(link = "
       logit"))
65 summary(t_glm)
66
67
```

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)
      ANSWER

   (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)

   (c) What is the estimated probability 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.