Problem Set 2

Applied Stats II

Due: February 18, 2024

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 23:59 on Sunday February 18, 2024. No late assignments will be accepted.

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.RData 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.

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)

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

   (c) 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.

```
1  #######################
2  # Problem 1
3  #######################
4  # load data
5  load(url("https://github.com/ASDS-TCD/StatsII_Spring2024/blob/main/
       datasets/climateSupport.RData?raw=true"))
6  data <- climateSupport
7
8  # Load data
9  load(url("https://github.com/ASDS-TCD/StatsII_Spring2024/blob/main/
       datasets/climateSupport.RData?raw=true"))
10 data <- climateSupport
11
12 # Convert Ord.Factor variables to Factor variables
13 data$countries <- factor(data$countries, levels = c("20 of 192", "80
        of 192", "160 of 192"), ordered = FALSE)
14 data$sanctions <- factor(data$sanctions, levels = c("None", "5%", "
       15%", "20%"), ordered = FALSE)
15
16 # Fit logistic regression model
17 model <- glm(choice ~ countries + sanctions, data = data, family =
       binomial)
18
19 # Display summary output
20 summary(model)
21
22 stargazer(model)
```

	Dependent variable:
	choice
countries80 of 192	0.336***
	(0.054)
countries160 of 192	0.648***
	(0.054)
sanctions5%	0.192***
	(0.062)
sanctions15%	−0.133**
	(0.062)
sanctions20%	−0.304***
	(0.062)
Constant	−0.273***
	(0.054)
Observations	8,500
Log Likelihood	−5,784.130
Akaike Inf. Crit.	11,580.260
Note:	*p<0.1; **p<0.05; ***p<0.01

According to the result:
$logit(P(\text{support policy})) = -0.27266 + 0.33636 \times \text{countries80 of 192} + 0.64835 \times \text{countries160 of 192} + 0.19186 \times \text{sanctions5\%} - 0.13325 \times \text{sanctions15\%} - 0.30356 \times \text{sanctions20\%}$

```r
# Fit null model with only intercept
null_model <- glm(choice ~ 1, data = data, family = binomial)

# Fit full model with predictors
full_model <- glm(choice ~ countries + sanctions, data = data, family =
    binomial)

# Perform likelihood ratio test using ANOVA
lr_test <- anova(null_model, full_model, test = "Chisq")

# Display the results
print(lr_test)

stargazer(lr_test)
```

Table 2:

Statistic	N	Mean	St. Dev.	Min	Max
Resid. Df	2	8,496.500	3.536	8,494	8,499
Resid. Dev	2	11,675.830	152.134	11,568.260	11,783.410
Df	1	5.000		5	5
Deviance	1	215.150		215.150	215.150
Pr(>Chi)	1	0.000		0	0

H0:The coefficient of countries 80 of 192, countries 160 of 192, sanctions5%, sanctions15%, sanctions20% = 0
Due to the p-value equal or less than 0.05, we can reject the H0, that means at least one of the coefficients is not zero. So we can assume the model makes sense.

2(a)To interpret how increasing sanctions from 5% to 15% changes the odds of supporting the policy,we need to calculate the difference in coefficients between 15% and 5% sanctions:
Change in log odds = coefficient for sanctions15% - coefficient for sanctions5% =-0.13325-0.19186=-0.32511

$$OddsRatio = e^{changeinlogodds} = e^{-0.32511}$$

Since the odds ratio is less than 1, it indicates a decrease $e^{-0.32511}$ in the odds of supporting the policy when sanctions increase from 5% to 15%.

2(b)

```
1 # Predict probability for 80 of 192 countries with no sanctions
2 prob <- predict(model, newdata = data.frame(countries = "80 of 192", sanctions
      = "None"), type = "response")
3
4 # Display the estimated probability
5 print(prob)
```

```
        1
  0.5159191
```

according to the $P = \frac{1}{e^{-logit(P)}} = \frac{1}{e^{-logit-0.273+0.336}} = 0.516$, So the probability is 0.5159191.

2(c)

```
1 # Fit the model with interaction term
2 model_interaction <- glm(choice ~ countries * sanctions, data = data, family =
      binomial)
3
4 # Perform likelihood ratio test
5 lr_test_2 <- anova(model, model_interaction, test = "Chisq")
6
7 # Display the results
8 print(lr_test_2)
9
10 stargazer(lr_test_2)
```

Table 3:

Statistic	N	Mean	St. Dev.	Min	Max
Resid. Df	2	8,491.000	4.243	8,488	8,494
Resid. Dev	2	11,565.110	4.450	11,561.970	11,568.260
Df	1	6.000		6	6
Deviance	1	6.293		6.293	6.293
Pr(>Chi)	1	0.391		0.391	0.391

H0: The interaction term has no effect on the model.
Due to the p-value=0.391 not equal or more than 0.05, so we should accept the H0.
There is no evidence to suggest that the effect of interaction term between countries and sanctions does not significantly improve the model fit.