

Problem Set 2

Jacqueline Bouvier 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.

```
1 # Creating factor variables
2 climateSupport$choice <- as.numeric(climateSupport$choice == "Supported")
3 climateSupport$countries <- relevel(factor(climateSupport$countries ,
4                                           ordered = FALSE) ,
5                                     ref = "20 of 192")
6 climateSupport$sanctions <- relevel(factor(climateSupport$sanctions ,
7                                           ordered = FALSE) ,
8                                       ref = "None")
9
10 add_model <- glm(choice ~ countries + sanctions ,
11                 data = climateSupport ,
12                 family = binomial(link = "logit"))
13 summary(add_model)
```

] We need to create factor variables for each variable output, before we could run the additive model. Then were were able to run the glm()

Table 1: Additive model of ClimateSupport

	<i>Dependent variable:</i>
	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
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

We can see from the table the coefficients of this additive model. Our intercept would be -0.27266 meaning that on average when no countries (0) participate in the agreement and there are no sanctions (none) - the log odds of someone supporting the policy is -0.27266. If we are to look at the explanatory variables individually, all others hold constant, we would see the following results. On average, while all other variables are held constant...

-Countries 80 of 192 - On average, while all other variables are held constant, 80 of 192 countries in participation would see an increase of 0.336 log odds in terms of supporting a policy

-Countries 160 of 192 - On average, while all other variables are held constant, 160 of 192 countries in participation would see an increase of 0.648 log odds in terms of supporting a policy

-Sanction 5percent - On average, while all other variables are held constant, if there is sanctions of missing emissions reduction targets of 5percent, we could see an increase of the log odds of support of the policy by 0.192

-Sanction 15percent - On average, while all other variables are held constant, if there is sanctions of missing emissions reduction targets of 15percent, we could see a decrease of the log odds of support of the policy by 0.133

-Sanction 20percent - On average, while all other variables are held constant, if there is sanctions of missing emissions reduction targets of 20percent, we could see a decrease of the log odds of support of the policy by 0.304

It could be assumed that the greater number of countries participating, the higher the support for the policies would be. Though if there is an increase in terms of sanctions it could be assumed that the support would decrease, according to the regression.

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)

```
1 climateSupport$sanctions <- relevel(factor(climateSupport$sanctions ,
2                                           ordered = FALSE) ,
3                                           ref = "5%")
4 add_model_pt2 <- glm(choice ~ countries + sanctions ,
5                       data = climateSupport ,
6                       family = binomial(link = "logit"))
7 summary(add_model_pt2)
8 # sanctions 15% -0.32510
9 stargazer(add_model_pt2, title = "Additive model of ClimateSupport
10 pt2", type = "latex")
11 climateSupport$sanctions <- relevel(factor(climateSupport$sanctions ,
12                                           ordered = FALSE) ,
13                                           ref = "15%")
14 add_model_pt <- glm(choice ~ countries + sanctions ,
15                     data = climateSupport ,
16                     family = binomial(link = "logit"))
17 summary(add_model_pt)
18 stargazer(add_model_pt, title = "Additive model of ClimateSupport
19 check", type = "latex")
20 #sanctions 5% 0.3251
```

] Running two additive models to look at increase in sanctions from 5percent to 15percent - we can assume that there is an associated decrease of 0.3251 as we increase the percentage. So, on average we could assume a decrease in support in policy by 0.3251 as we increase percentage of sanctions from 5percent to 15percent.

Table 2: Additive model of ClimateSupport pt2

	<i>Dependent variable:</i>
	choice
countries80 of 192	0.336*** (0.054)
countries160 of 192	0.648*** (0.054)
sanctionsNone	−0.192*** (0.062)
sanctions15%	−0.325*** (0.062)
sanctions20%	−0.495*** (0.062)
Constant	−0.081 (0.053)
Observations	8,500
Log Likelihood	−5,784.130
Akaike Inf. Crit.	11,580.260
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Table 3: Additive model of ClimateSupport check

	<i>Dependent variable:</i>
	choice
countries80 of 192	0.336*** (0.054)
countries160 of 192	0.648*** (0.054)
sanctions5%	0.325*** (0.062)
sanctionsNone	0.133** (0.062)
sanctions20%	−0.170*** (0.062)
Constant	−0.406*** (0.054)
Observations	8,500
Log Likelihood	−5,784.130
Akaike Inf. Crit.	11,580.260
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

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

```

1 # Prediction equation from week 4 slides
2 # 1/(1+exp[beta0+beta1*Xi])
3 beta_0 <- -0.27266 # intercept of add_model
4 beta_1 <- 0.33636 # coefficient of countries 80 of 192
5
6 1/(1+(exp(-(beta_0+(beta_1))))))
7 # 0.5159196
8
9 test_probability <- predict(add_model,
10                             newdata = data.frame(countries = "80 of 192",
11                                                     sanctions = "None"),
12                             type = "response") # testing the probability with
13 predict()
14 print(test_probability) # 0.5159191

```

] Using the prediction equation from week 4 slides, and the coefficients from the model the probability was calculated out for an individual that would support a policy if there are 80 of 192 countries and no sanctions. This was then checked using the predict function in R. In a given individual the odds of them supporting a policy when 80 of 192 countries are participating with no sanctions is 51.58 percent that the individual could support.

- (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 # Model with interactions (*) - countries and sanctions are
2   interacting with each other and will get extra coefficients
3 interaction <- glm(choice ~ countries * sanctions,
4                    data = climateSupport,
5                    family = binomial(link = "logit"))
6 summary(interaction)
7 stargazer(interaction, title = "Interaction", type = "latex")
8 # x^2/Likelihood ratio test - we can use the additive model and
9   the interaction model
10 test <- anova(add_model, interaction,
11               test = "LRT")
12 print(test)
13 stargazer(test, title = "Deviance Table", type = "latex")

```

]

Seeing that the p-value (0.391) from the LRT test (table below) is not less than 0.01 we could potentially conclude that adding interaction doesn't change the answers from 2a and 2b significantly. We could see possible small changes in the coefficients and probability of each but would not be considered to be significant. Comparing table 1 and table 5 we can see the slight difference in coefficients that would change the calculation of 2a and 2b

Table 4: Deviance Table

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

Table 5: Interaction

	<i>Dependent variable:</i>
	choice
countries80 of 192	0.323*** (0.108)
countries160 of 192	0.561*** (0.108)
sanctions5%	0.219** (0.107)
sanctionsNone	0.097 (0.108)
sanctions20%	−0.156 (0.110)
countries80 of 192:sanctions5%	0.147 (0.154)
countries160 of 192:sanctions5%	0.182 (0.151)
countries80 of 192:sanctionsNone	0.052 (0.152)
countries160 of 192:sanctionsNone	0.052 (0.153)
countries80 of 192:sanctions20%	−0.145 (0.152)
countries160 of 192:sanctions20%	0.109 (0.154)
Constant	−0.372*** (0.078)
Observations	8,500
Log Likelihood	−5,780.983
Akaike Inf. Crit.	11,585.970