

Problem Set 2

Vismante Dringelyte

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.

First, I prepare the data by recoding the **countries** and **sanctions** variables to factors.

```
1 df$countries <- factor(df$countries, levels = c("20 of 192", "80 of 192",
  "160 of 192"), ordered = FALSE)
2 df$sanctions <- factor(df$sanctions, levels = c("None", "5%", "15%", "20%
  "), ordered = FALSE)
```

Then, I fit the model

```
1 model <- glm(choice ~ countries + sanctions,
2               data = df,
3               family = binomial(link = "logit"))
```

This model returns these results:

<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	

Table 1:

Table 2: Coefficient Names and P-Values

Coefficient	P-Value
(Intercept)	3.64×10^{-7}
countries80 of 192	4.05×10^{-10}
countries160 of 192	$< 2 \times 10^{-16}$
sanctions5%	0.00203
sanctions15%	0.03183
sanctions20%	1.01×10^{-6}

The global null hypothesis is that, holding all else equal, the number of countries participating and the level of sanctions will not have an effect on an individual's support of a climate policy.

The model equation is:

$$\begin{aligned}
\log\text{-odds}(\text{choice}) = & -0.27266 \\
& + 0.33636 \times \text{countries80 of 192} \\
& + 0.64835 \times \text{countries160 of 192} \\
& + 0.19186 \times \text{sanctions5\%} \\
& - 0.13325 \times \text{sanctions15\%} \\
& - 0.30356 \times \text{sanctions20\%}
\end{aligned} \tag{1}$$

The intercept shows that when there are no sanctions and 20 countries participate, the expected odds for supporting the policy are $\exp(-0.273)=0.761$. On average, people show support for the policy.

Holding all other variables constant, 80 of 192 countries participating increases the log odds of supporting the policy by 0.366.

Holding all other variables constant, 160 of 192 countries participating increases the log odds of supporting the policy by 0.648.

Holding all other variables constant, sanctions of 5% increase the log odds of supporting the policy by 0.192.

Holding all other variables constant, sanctions of 15% decrease the log odds of supporting the policy by 0.133.

Holding all other variables constant, sanctions of 20% decrease the log odds of supporting the policy by 0.273.

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)

$$\begin{aligned}
\text{log-odds(choice)} &= -0.27266 \\
&+ 0.33636 \times \text{countries80 of 192} \\
&+ 0.64835 \times \text{countries160 of 192} \\
&+ 0.19186 \times \text{sanctions5\%} \\
&- 0.13325 \times \text{sanctions15\%} \\
&- 0.30356 \times \text{sanctions20\%} \\
&= -0.27266 \\
&+ 0.33636 \times 0 \\
&+ 0.64835 \times 1 \\
&+ 0.19186 \times 1 \\
&- 0.13325 \times 0 \\
&- 0.30356 \times 0 \\
&= 0.56755
\end{aligned} \tag{2}$$

$$\begin{aligned}
\text{log-odds(choice)} &= -0.27266 \\
&+ 0.33636 \times \text{countries80 of 192} \\
&+ 0.64835 \times \text{countries160 of 192} \\
&+ 0.19186 \times \text{sanctions5\%} \\
&- 0.13325 \times \text{sanctions15\%} \\
&- 0.30356 \times \text{sanctions20\%} \\
&= -0.27266 \\
&+ 0.33636 \times 0 \\
&+ 0.64835 \times 1 \\
&+ 0.19186 \times 0 \\
&- 0.13325 \times 1 \\
&- 0.30356 \times 0 \\
&= 0.24244
\end{aligned} \tag{3}$$

$$0.24244 - 0.56755 = -0.32511 \tag{4}$$

So, when 160 countries are participating, increasing the penalty from 5% to 15% decreases the log-odds of supporting the policy by 0.325.

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

$$\begin{aligned}
\text{log-odds(choice)} &= -0.27266 \\
&+ 0.33636 \times \text{countries}_{80 \text{ of } 192} \\
&+ 0.64835 \times \text{countries}_{160 \text{ of } 192} \\
&+ 0.19186 \times \text{sanctions}_{5\%} \\
&- 0.13325 \times \text{sanctions}_{15\%} \\
&- 0.30356 \times \text{sanctions}_{20\%} \\
&= -0.27266 \\
&+ 0.33636 \times 1 \\
&+ 0.64835 \times 0 \\
&+ 0.19186 \times 0 \\
&- 0.13325 \times 0 \\
&- 0.30356 \times 0 \\
&= 0.0637
\end{aligned} \tag{5}$$

When 80 of 192 countries are participating and there are no sanctions, the log-odds of an individual supporting the policy are 0.0637

- (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.

First I ran a GLM with an interaction.

```

1 model2 <- glm(choice ~ countries * sanctions,
2               data = df,
3               family = binomial(link = "logit"))

```

Then I used ANOVA to conduct the test.

```

1 lrt <- anova(model, model2, test = "LRT")

```

This produced these results:

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

The p-value is quite high, so we fail to reject the null hypothesis that the multiplicative model provides a better fit than the additive one.