

Problem Set 2

Applied Stats II

Due: February 19, 2023

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 19, 2023. 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.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.

MY ANSWER FOR QUESTION 1

```
1
2 load(url("https://github.com/ASDS-TCD/StatsII_Spring2023/blob/main/
   datasets/climateSupport.RData?raw=true"))
3
4 # how the data is being stored in the dataset
5
6 str(climateSupport)
7 names(climateSupport)
8 # inspect
9 head(climateSupport)
10 tail(climateSupport)
11
12 # response variable – choice or observe changes (what’s measured)
13 # explanatory variable – countries and sanctions or what changes as a
   result
14 # (what changes)
15
16 # tables shows inputs as factors or only 3 countries with sanctions
17 # 20 counties and with no sanctions (baseline)
18
19 table(climateSupport$countries)
20 table(climateSupport$sanctions)
21
22 # factors to numeric values
23 # converting a vector or a factor to a numeric vector
24 climateSupport$choice <- as.numeric(as.factor(climateSupport$choice))-1
25
26 # 1 = supported and 0 = did not support
27
28 # countries
29 climateSupport$countries <- as.numeric(as.factor(climateSupport$countries
   ))-1
30
31 # sanctions
32 climateSupport$sanctions <- as.numeric(as.factor(climateSupport$sanctions
   ))-1
33
34 # fit an additive model. provide the summary output
35 # depends on countries + sanctions
36 # the data here is binary 1,0 for a logit regression
```

```

37 additive_model <- glm(choice ~ countries + sanctions, data =
    climateSupport,
38                         family = binomial(logit))
39
40 summary(additive_model)
41 # Results
42
43 # Coefficients:
44 # Estimate Std. Error z value Pr(>|z|)
45 # (Intercept) -0.14458      0.04518  -3.200  0.00137 **
46 # countries      0.32436      0.02689  12.062 < 2e-16 ***
47 # sanctions     -0.12353      0.01964  -6.291 3.15e-10 ***
48
49 # Signif. codes:  0      ***      0.001      **      0.01      *      0.05      .
    0.1      1
50 # Null deviance: 11783  on 8499  degrees of freedom
51 # Residual deviance: 11597  on 8497  degrees of freedom
52 # AIC: 11603
53
54 # Number of Fisher Scoring iterations: 4
55
56 # calculate p-value of overall Chi-Square statistic
57 1-pchisq(11783-11597, 8499-8497)
58 # [1] 0
59
60 # since this p-value is less than .05, we reject the null hypothesis
61 # there is a significant relationship between the combination of
62 # countries and sanctions and the final choice selected
63
64 #####
65
66 # trying to discount the global null hypothesis
67
68 summary_glm <- summary(additive_model)
69 anova(additive_model, glm, test = "Chisq")
70
71 # The null hypothesis (H0) of the ANOVA is no difference in means,
72 # and the alternative hypothesis (Ha) is that the means are different
73 # from one another.
74 # here the anova gives p-value of 2.2e-16 or 2 ^ -16
75 # and. can reject the null hypothesis
76
77 exp(summary_glm$coefficients[1,1])/(1 + exp(summary_glm$coefficients[1
    ,1]))
78 b1 = summary_glm$coefficients [1,1]
79 exp(b1 + summary_glm$coefficients[2,1])/(1 + exp(b1 + summary_glm$
    coefficients
80     [2,1]))
81
82
83 levels(climateSupport$sanctions)

```

```

84
85 climateSupport$sanctions <- relevel(factor(climateSupport$sanctions ,
      ordered = F), ref = "5%")
86
87 climateSupport$sanctions
88
89 # fitting the additive model
90
91 c_logit <- glm(formula = choice ~ countries + sanctions, family = "
      binomial", data = climateSupport)
92 summary(c_logit)
93
94 # or c_logit <- glm(formula = choice ~ countries + sanctions, family = "
      binomial", data = climateSupport)
95 # summary(c_logit)
96 #
97
98
99 # Coefficients:
100 #
      Estimate Std. Error  z value  Pr(>|z|)
101 # (Intercept)    0.24743    0.04406    5.616   1.95e-08 ***
102 # countries.L     0.45845    0.03810   12.033  < 2e-16 ***
103 # countries.Q    -0.00995    0.03806   -0.261   0.79374
104 # sanctionsNone  -0.19185    0.06216   -3.086   0.00203 **
105 # sanctions15%   -0.32510    0.06224   -5.224   1.76e-07 ***
106 # sanctions20%   -0.49542    0.06228   -7.955   1.79e-15 ***
107 # ———
108 # Signif. codes:  0      ***      0.001      **      0.01      *      0.05      .
      0.1              1
109
110 # (Dispersion parameter for binomial family taken to be 1)
111
112 # Null deviance: 11783  on 8499  degrees of freedom
113 # Residual deviance: 11568  on 8494  degrees of freedom
114
115 # 11783 = null deviance and 11568 = residual deviance
116 # find the p-value for the chi-square test statistic
117 pchisq(11783-11568, 5, lower.tail = F)
118
119 1-pchisq(11783-11568, 8499-8494)
120
121 # this gives
122 # 1.749304e-44
123
124
125
126 # logit model
127 # period functions as omnibus selector (the kitchen sink additive model)
128 # ~ . will select countries (ord) and sanctions (fct)
129
130 climate_logit <- glm(choice ~ ., family = binomial(link="logit"), data =

```

```

climateSupport)
131 summary(climate_logit)
132
133 reg_exp <- exp(coef(climate_logit))
134 stargazer(reg_exp, type = "text")
135
136
137 # not supported or supported
138 # the dependent variable is binary(0/1, True/False, Yes/No) in nature
139 t_glm <- glm(choice ~ 1, data = climateSupport, family=binomial(link = "
logit"))
140 summary(t_glm)
141
142 exp(summary_glm$coefficients [1,1])/(1 + exp(summary_glm$coefficients [1
,1]))

```

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.

MY ANSWER FOR QUESTION 2

```

1
2 # something to do with how does the baseline countries = 20 and sanctions = 5%
3
4 # see lecture 4 notes
5
6 ##### Part (a) #####
7
8 # previous results from Question 1
9 # # Coefficients:
10 # Estimate Std. Error z value Pr(>|z|)
11 # (Intercept) -0.14458      0.04518  -3.200  0.00137 **
12 # countries    0.32436      0.02689  12.062 < 2e-16 ***
13 # sanctions   -0.12353      0.01964  -6.291 3.15e-10 ***
14
15 log_odds1 <- exp((-0.14458) + (0.32436)*2 + (-0.12353)*1)/(1 + exp
(( -0.14458) + (0.32436)*2 + (-0.12353)*1))
16 log_odds1
17 # result
18 # 0.5940255
19

```

```

20 log_odds2 <- exp((-0.144558) + (0.32436)*2 + (-0.12353)*2)/(1 + exp
    ((-0.144558) + (0.32436)*2 + (-0.12353)*2))
21 log_odds2
22 # result
23 # 0.5639238
24
25 # interperatation of additive model week 4 lecture
26
27 odds_difference <- log_odds1 - log_odds2
28 odds_difference
29 # result
30 # 0.5639238
31
32
33 ##### Part (b) #####
34
35 # referencing week 4 lecture 4 notes
36 # using the exp(coefficients) function to create the log odds ratio
37 # 5% sanctions
38 odds_1 <- exp((-0.144558) + (0.32436)*0 + (-0.12353)*1)/(1 + exp ((-0.144558)
    + (0.32436)*0 + (-0.12353)*1))
39
40 odds_1
41 # result
42 # 0.4333765
43
44 exp(odds_1)
45 # result
46 # 1.542457
47
48 # 15%
49 odds_2 <- exp((-0.144558) + (0.32436)*0 + (-0.12353)*2)/(1 + exp ((-0.144558)
    + (0.32436)*0 + (-0.12353)*2))
50
51 odds_2
52 # result
53 # 0.4033279
54
55 exp(odds_2)
56 # result
57 # 1.496798
58
59 odds_diff <- odds_1 - odds_2
60 odds_diff
61 # result
62 # 0.03004869
63
64 # interperatation
65 # an increase 0.03004869 of log odds of a participant supporting the climate
    policy when
66 # the sanctions are increased from 5% to 15%

```

```

67
68
69 # or trying another way (just experimenting)
70 # the coefficients are exponentiated: exp(coef(fit))
71 fit <- glm(choice ~ countries + sanctions, data = climateSupport, family="
    binomial")
72 exp(cbind(Odds_and_OR=coef(fit), confint(fit)))
73 # result
74 # 0.456359
75
76
77 ##### Part (c) #####
78
79 # with 80 countries and no sanctions
80
81 odds_3 <- exp((-0.144558) + (0.32436)*1 + (-0.12353)*0)/(1 + exp((-0.144558)
    + (0.32436)*1 + (-0.12353)*0))
82 odds_3
83
84 # result
85 # 0.5448298
86
87 # interaction model
88
89 interaction_model <- glm(choice ~ countries*sanctions, data = climateSupport,
    family = binomial(logit))
90 summary(interaction_model)
91
92 # results
93
94 # Coefficients:
95 # Estimate Std. Error z value Pr(>|z|)
96 # (Intercept) -0.148144 0.057311 -2.585 0.00974 **
97 # countries 0.328007 0.045036 7.283 3.26e-13 ***
98 # sanctions -0.121111 0.030987 -3.908 9.29e-05 ***
99 # countries:sanctions -0.002455 0.024288 -0.101 0.91950

```