Lab Report 5: Logistic Regression

Understanding Political Beliefs

DUE: November 18 at 5 PM

After the November 2018 elections, Penn State's McCourtney Institute conducted a poll of a random sample of Americans to understand how well Americans understand each other's political beliefs and how that affects the state of political discourse in America today. You can read in the dataset with this R code:

```
data <- read.csv("https://raw.githubusercontent.com/ilaydaonder/POLS209/Lab-Report-5/nov18survey.csv")
```

Your task is to analyze these data using logistic regressions following the steps below. **No memo is required! Just briefly answer the questions.**

Here are the important variables:

- **hardtorelate** is coded on a 0-1 point scale in response to the question about how easy it is to understand people with different political views. Respondents are coded 1 if they find it difficult to see things from the other person's point of view and 0 if they find it easy to see things from the other person's point of view.
- **age** is is the respondent's age in years.
- **democrat** and **republican** are coded 0 if not a member of that political party and 1 if the respondent identifies as a member of that political party. (If an observation is 0 on both, assume they are an Independent)
- **gender** is coded 0 if the respondent is male and 1 if the respondent is female.
- **ideo5** is the respondent's ideology on a 5-point scale. Higher values indicate more conservative respondents. A value of 3 indicates a political moderate, and a value of 5 indicates a strongly conservative respondent.
- **educ** is the respondent's ideology on a 6-point scale. Higher values indicate more educated respondents. A value of 2 means the respondent has completed high school.

Use your work from the last lab and the lecture slides on logistic regression!

1. **Estimate two logistic regressions**, both with **hardtorelate** as the dependent variable. For the first model, include only **age** as the independent variable. For the second model, include **age, democrat, republican, gender, ideo5,** and **educ**. Use stargazer to make a table[1] (just use the `text` type, no need to make it any fancier).

2. **Based on the results from the second model**, what features of Americans are statistically related to their willingness to relate to an opposing point of view?

3. **Interpret the coefficients** for **democrat** and **republican**. Remember that unlike linear regression, we only know the *direction* of the relationships from the coefficients, not the magnitude. Thus, interpret the direction of the relationship between these two variables and the dependent variable, also keeping in mind that the excluded category is Independent.

4. **What is the predicted probability** of finding it hard to relate to people with different political views for a 35-year-old, high-school educated, politically moderate Democratic man? Again, use the second model, and like the last lab, use the `predict()`[2] function. Copy and paste your R code with your answer.

[1] Remember from the previous lab report, first load the stargazer package using `library(stargazer)`, and then create your table using `stargazer(model1, model2, type = "text")`.

[2] Code for this was in the logistic regression slides.