

## Categorical Lab #5 Assignment

Please use the “turnout” dataset within the Zelig package, which details a random subset of 2000 individuals (out of 15,837 surveyed) during the 1992 presidential election. There are 5 variables in these data: **race** (whether the participant was “white” or nonwhite [“others”]), **age** (how old the participant was at the time), **educate** (the participant’s years of education), **income** (a categorical measure of the participant’s level of income), and **vote** (whether the participant voted or not [1 = voted, 0 = did not vote]).

To access this dataset, first call the Zelig package [i.e., `library(Zelig)`], then use `data(turnout)` to “attach” the dataset to R. You may then use the `turnout` dataset in R.

1. Run a logistic regression, with **vote** as the dependent variable and **educate** as the independent variable. Please report the beta estimate of the effect of **educate** on **vote**, and provide a meaningful interpretation of this estimate.
2. Provide a 95% profile likelihood ratio confidence interval and a 95% Wald confidence interval of the beta estimate in Question 1. Translate the intervals into odds and provide a meaningful interpretation of each confidence interval. According to each confidence interval, is the estimate significant?
3. Create a probability plot for the logistic regression in Question 1. Provide this plot. Calculate the probability for the following questions: What is the probability of voting at a high school education (12 years)? At a college education (16 years)? At the mean level of education?
4. What is the median effective level of education for this model? In other words, what is the value of **education** at which the probability of voting is 50%? What significance does this point have for the rate of change of the probability?