Homework 10 – Intro. to Computational Statistics

1. Using the `anes_2008tr.csv` dataset in Course Resources, model `vote_rep` (whether the respondent voted Republican in the last election) as a function of age, race, income, and ideology.

a. What's the probability of voting Republican for a white person of average age, income, and ideology?

b. What's the change in probability of voting Republican for a person of average age, income, and ideology who switches from black to white?

c. Using the e^{β} formula from the lesson, what's the effect on the odds ratio of shifting from black to white?

d. What has a greater effect on voting Republican: an age increase of 50 years, or an incease of one income bracket?

e. Now run the regression with all the other variables in anes_2008tr (except for `voted`). How do your coefficients change? What do you think explains any coefficient that became or lost significance?

2. Construct a simulated y variable with 100 observations where each observation (year) is a function of the previous observation: specifically, y_t is 80% y_{t-1} + 20% random noise with mean 0 and sd 1 (and $y_1 = 1$). Estimate an ARIMA model using `auto.arima()` from the `forecast` package and interpret the results, in particular the ARIMA(p,d,q) numbers and the coefficients reported, if any. What do you think is going on here?

3. Find some existing data that either has temporal or bindary dependent variable data and run a ARIMA or logit model on it and interpret the results in detail. You can use data from the previous assignment, and you can construct a binary variable out of some existing continuous variable if you like.