POLS201 Spring 2019

Homework, Logit Example, and R Tricks

April 17

Agenda

- Review of the homework
- Live examples of logit regression
- . . . using a website with an excellent demo
- The Nieheisel paper (focus on the table formats)
- If time permits: tables and equation editors in Word
- Note: I will not be on campus during tomorrow's scheduled office hours :(

Homework

- Answers on handout. Let's review them.
- If question 8 seemed unfamiliar to you, don't be alarmed.
- The score will give me some indication if you worked the problems, but the important takeaway: use this as a study guide for the exam.

A useful logit example from the UCLA Digital Education site

- Posted on RStudio Cloud (not an assignment)
- https://stats.idre.ucla.edu/r/dae/logit-regression/

Why do a logit regression?

- Because your dv is binary.
- Can it work if you have a few outcomes but more than two?
  - Not literally, but you can run multiple regressions on different pairs of outcomes. Your data must be formatted as ones and zeroes, however.

The basic format difference:

- myOLS <- lm(admit ~ gre + gpa + rank, data = mydata)
- mylogit <- glm(admit ~ gre + gpa + rank, data = mydata, family = "binomial")

If you like Datacamp, I created a new Custom Track:

- Regression in R. Don't try to do it all.
- But the Intro to Statistics with R: Correlation and Regression might help.

The idea with logit regression

- We express our prediction of y as a probability that y=1 given the values of X (x1, x2, etc.).
- We convert the dependent variable to something called a "log-likelihood".
- . . . which means, our formula changes to something like this:
- $ln(\frac{p}{1-p}) = \alpha_i + X\beta + \epsilon_i$ where $p = prob(y = 1)$
- The right hand side is the same, but the left hand side is no longer just y_i

Do you need to remember this formula? No. But keep in mind

- The coefficients β do not predict a linear relationship. That means they predict a value that doesn't change at a constant rate.
- The very best way to show the predicted effect is graphically. The UCLA example has an example if you want to try.
- For our purposes, the important things to see are these:

Do you need to remember this formula? No. But keep in mind

1. Does the coefficient have statistical significance? It's harder to achieve; hope for the best, see what you see, and report it.
2. Does the sign of the coefficient (+ or -) move in the direction you expect? Example: Republicans are more likely to complain we spend too little on the military (a real finding from a student's work here). It would be weird if the sign suggested the opposite.
3. Is the coefficient approaching twice the standard error?
4. Do you need to make an IV a dummy using as.factor()?

Merging Data Example

- You have one dataset with state names and another with state codes.
- You want to merge them. How?

1. Upload a lookup table that translates two-character codes to full state names.
2. Get your field names consistent.
3. Use one of the join commands.

A Word About Tables

- Use the Nieheisel paper as an illustrative example. You are concerned about reporting two things:
- The estimate of the coefficients and the standard error.
- Customarily, this means:
- Each line shows an explanatory variable, and each column gives a coefficient estimate.
- Below the estimate show the SE in parentheses, on a separate line.

A Word About Tables

- Create tables in your word processor (and you can create as many as you want) that show, in columns,
- all the variations of the models you want to report.
- If you change the DV, report the changes in a separate column. List all your variables but leave the excluded variables for a given model blank in that column

Two features of a word processor that will help

- Tables
- Equation editor

One final thing to remember

- If you don't achieve a result that shows significance, just report what you find.
- Use some of your paper's real estate to explain what kind of data would have produced a result more consistent with your theory.
- In fact: do that anyway (which is mentioned in the paper guide).