Problem Set 4 - Solution

Dino Wildi

Due: November 26, 2021

Question 1: Economics

*We would like to study whether individuals with higher levels of income have more prestigious jobs. Moreover, we would like to study whether professionals have more prestigious jobs than blue and white collar workers. Create a new variable **professional** by recoding the variable **type** so that professionals are coded as 1, and blue and white collar workers are coded as 0 (Hint: **ifelse**.) Run a linear model with **prestige** as an outcome and **income**, **professional**, and the interaction of the two as predictors (Note: this is a continuous × dummy interaction.)*

```
1 Prestige$professional <- ifelse(Prestige$type == "prof", 1, 0)
2
3 reg1 <- lm(prestige ~ income + professional + income*
     professional, data = Prestige)
4 summary(reg1)
```

Write the prediction equation based on the result.

$$\hat{y} = 21.142 + 37.781D + 0.003x - 0.002xD + e$$

*Interpret the coefficient for **income**. Interpret the coefficient for **professional**.*

The coefficient for `income` is 0.003. This means that for a non-professional worker, an additional dollar of income will lead to a 0.003 point increase in prestige.
The coefficient for `professional` is 37.781. This means that assuming income equals 0 (i.e. at the intercept), a professional worker will have a prestige score that is 37.781 points higher than a non-professional worker.

*What is the effect of a $1,000 increase in income on prestige score for professional occupations? In other words, we are interested in the marginal effect of income when the variable **professional** takes the value of 1. Calculate the change in $\hat{y}$ associated with a $1,000 increase in income based on your answer for (c).*

The marginal effect of an increase of x on $\hat{y}$ will be $x\beta_1 + xD\beta_2$. In this case, D takes the value of 1, so the marginal effect will be $\hat{y} = 0.003 * 1000 - 0.002 * 1000 = 0.001 * 1000$. If we do not round the results obtained by the regression in (b), the resulting marginal increase will be 0.845 prestige points for an increase of $1000 for a professional occupation.

*What is the effect of changing one's occupations from non-professional to professional when her income is $6,000? We are interested in the marginal effect of professional jobs when the variable **income** takes the value of $6,000$. Calculate the change in $\hat{y}$ based on your answer for (c).*

The marginal effect of changing the value of D from 1 to 0 will be $37.781 - 0.002x$. If x takes a value of 6000, this means that the marginal change in $\hat{y}$ equals $37.781 - 0.002 * 6000 = 23.827$, when not using rounded coefficients.

Question 2: Political Science

Researchers are interested in learning the effect of all of those yard signs on voting preferences.[1] Working with a campaign in Fairfax County, Virginia, 131 precincts were randomly divided into a treatment and control group. In 30 precincts, signs were posted around the precinct that read, "For Sale: Terry McAuliffe. Don't Sellout Virgina on November 5."

Below is the result of a regression with two variables and a constant. The dependent variable is the proportion of the vote that went to McAuliff's opponent Ken Cuccinelli. The first variable indicates whether a precinct was randomly assigned to have the sign against McAuliffe posted. The second variable indicates a precinct that was adjacent to a precinct in the treatment group (since people in those precincts might be exposed to the signs).

Impact of lawn signs on vote share

Precinct assigned lawn signs (n=30)	0.042
	(0.016)
Precinct adjacent to lawn signs (n=76)	0.042
	(0.013)
Constant	0.302
	(0.011)

Notes: R^2=0.094, N=131

Use the results from a linear regression to determine whether having these yard signs in a precinct affects vote share (e.g., conduct a hypothesis test with $\alpha = .05$).

The appropriate hypothesis test to determine if the vote in a precinct is affected by having signs in that precinct is to perform a t-test on the first coefficient ("Precinct assigned lawn signs"). The null hypothesis will be $H_0 : \beta_1 = 0$, whereas $H_A : \beta_1 \neq 0$. The test statistic equals $t = \frac{\hat{\beta_1}}{se_{\hat{\beta_1}}}$, i.e. $\frac{0.042}{0.016} = 2.625$. Using the pt() function on R to determine a p-value, we get

---

[1]Donald P. Green, Jonathan S. Krasno, Alexander Coppock, Benjamin D. Farrer, Brandon Lenoir, Joshua N. Zingher. 2016. "The effects of lawn signs on vote outcomes: Results from four randomized field experiments." Electoral Studies 41: 143-150.

a result of $p = 0.995$, hence we can reject the null hypothesis and conclude there is an effect of vote signs being present on vote share.

```
1 pt(2.625, df = 131−4)
```

Use the results to determine whether being next to precincts with these yard signs affects vote share (e.g., conduct a hypothesis test with $\alpha = .05$).

The test for this question is again a t-test analogous to what is described above. This time, the null hypothesis will be that the coefficient for "Precinct adjacent to lawn signs" will equal 0; the alternative hypothesis is that it will not. The test statistic equals $t = \frac{0.042}{0.013} = 3.231$.

```
1 pt(3.231, df = 131−4)
```

The code above returns a p-value of 0.999, which is sufficient to reject the null hypothesis and conclude that being in a precinct adjacent to lawn signs has a significant effect on vote share.

Interpret the coefficient for the constant term substantively. Evaluate the model fit for this regression. What does this tell us about the importance of yard signs versus other factors that are not modeled?

The constant or intercept is 0.302. This means that on average, in a precinct that was neither assigned a lawn sign nor adjacent to a precinct assigned lawns signs, 30.2% of voters voted for Ken Cuccinelli. This value is notably larger than the effect for the lawn signs (more than 7 times as large), indicating that while the lawn sign effect is significant, it is small. This is further supported by the R^2 value being rather low (0.094), indicating that the lawn signs explained just short of 10% in variation.