Chapter 4 - Inference

Thalles Quinaglia Liduares

10/03/2022

Exercise 4.1

Upload packages

```
library(lmreg)
library(dplyr)
library(wooldridge)
library(car)
```

Upload database

```
data<-wooldridge::vote1

attach(data)
```

The following model can be used to study whether campaign expenditures affect election outcomes:

$$voteA = \beta_0 + \beta_1 log(expendA) + \beta_2 log(expendB) + \beta_3 prtystrA + u$$

where `voteA` is the percentage of the vote received by `Candidate A`, `expendA` and `expendB` are campaign expenditures by Candidates A and B, and `prtystrA` is a measure of party strength for Candidate A (the percentage of the most recent presidential vote that went to A's party).

(i) What is the interpretation of β_1?

This case is known as level-log, and is the percentual variation in voteA, for a unit increase in `log(expendA)`.

(ii) In terms of the parameters, state the null hypothesis that a 1% increase in A's expenditures is offset by a 1% increase in B's expenditures.

The null hypothesis is given by $H_0 : \beta_1 + \beta_2 = 0$ and alternative hypothesys by $H_1 : \beta_1 + \beta_2 \neq 0$.

(iii) Estimate the given model using the data in VOTE1.RAW and report the results in usual form. Do A's expenditures affect the outcome? What about B's expenditures? Can you use these results to test the hypothesis in part (ii)?

```
lm1<-lm(voteA~log(expendA)+log(expendB)+prtystrA, data)

summary(lm1)
```

```
##
## Call:
## lm(formula = voteA ~ log(expendA) + log(expendB) + prtystrA,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.3968  -5.4174  -0.8679   4.9551  26.0660
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  45.07893    3.92631   11.48   <2e-16 ***
## log(expendA)  6.08332    0.38215   15.92   <2e-16 ***
## log(expendB) -6.61542    0.37882  -17.46   <2e-16 ***
## prtystrA      0.15196    0.06202    2.45   0.0153 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.712 on 169 degrees of freedom
## Multiple R-squared:  0.7926, Adjusted R-squared:  0.7889
## F-statistic: 215.2 on 3 and 169 DF,  p-value: < 2.2e-16
```

The estimated model, is given by following equation

$$\widehat{voteA} = 45.07 + 6.08log(expendA) - 6.61log(expendB) + 0.15prtystrA$$

The R^2 indicates that 79.2% of variability in `voteA` is explaneid by exogenous variables.

(iv) Estimate a model that directly gives the t statistic for testing the hypothesis in part (ii). What do you conclude? (Use a two-sided alternative.)

```
data$expendA<--data$expendB

lm2<-lm(voteA~expendA+expendB+prtystrA, data)

summary(lm2)
```

```
##
## Call:
## lm(formula = voteA ~ expendA + expendB + prtystrA, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -31.305 -11.419   3.823  10.416  28.455
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 28.677086   5.603793   5.117 8.30e-07 ***
## expendA      0.022917   0.003531   6.491 8.99e-10 ***
## expendB            NA         NA      NA       NA
## prtystrA     0.579163   0.108309   5.347 2.85e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.18 on 170 degrees of freedom
## Multiple R-squared:  0.2945, Adjusted R-squared:  0.2862
## F-statistic: 35.47 on 2 and 170 DF,  p-value: 1.332e-13
```

The coefficient associated to `expendA` is statistically different from zero, as it's p-value<0.05.