Chapter 3 - Multiple Regression Analysis - Estimation

Thalles Quinaglia Liduares

04/03/2022

Exercise 3.9

Upload packages

```
library(lmreg)
library(wooldridge)
```

Upload database

```
data<-wooldridge::charity
```

Use the data in CHARITY.RAW to answer the following questions

(i) Estimate the equation

$$gift = \beta_0 + \beta_1 mailsyear + \beta_2 giftlast + \beta_3 propresp + u$$

by OLS and report the results in the usual way, including the sample size and R-squared. How does the R-squared compare with that from the simple regression that omits `giftlast` **and** `propresp` **?**

```
lm1<-lm(gift~mailsyear+giftlast+propresp, data)

summary(lm1)
```

```
##
## Call:
## lm(formula = gift ~ mailsyear + giftlast + propresp, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -52.893  -7.050  -3.650   1.397 241.206
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.551518   0.803034  -5.668 1.54e-08 ***
## mailsyear    2.166259   0.331927   6.526 7.53e-11 ***
## giftlast     0.005927   0.001432   4.138 3.58e-05 ***
## propresp    15.358605   0.874539  17.562  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.43 on 4264 degrees of freedom
## Multiple R-squared:  0.08336,    Adjusted R-squared:  0.08271
## F-statistic: 129.3 on 3 and 4264 DF,  p-value: < 2.2e-16
```

```
lm2<-lm(gift~mailsyear, data)

summary(lm2)
```

```
##
## Call:
## lm(formula = gift ~ mailsyear, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -11.287  -7.976  -5.976   2.687 245.999
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.0141     0.7395   2.724  0.00648 **
## mailsyear     2.6495     0.3431   7.723  1.4e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.96 on 4266 degrees of freedom
## Multiple R-squared:  0.01379,    Adjusted R-squared:  0.01356
## F-statistic: 59.65 on 1 and 4266 DF,  p-value: 1.404e-14
```

The estimated equation is given by

$$\widehat{gift} = -4.55 + 2.16mailsyear + 0.005giftlast + 15.35propresp$$

The variability in response variable `gift` is only 8.3% explained by the predictor variables.

In the simple regression estimation this value was 1.3%. Hence, the inclusion of predictor variables increases the predictability of the model.

(ii) Interpret the coefficient on `mailsyear`. Is it bigger or smaller than the corresponding simple regression coefficient?

The coefficient of `mailsyear` gives the marginal increase in the variable `gifts` for each additional mail sent. In the simple regression model this value is 2.64 against 2.14 in the multiple regression.

(iii) Interpret the coefficient on `propresp`. Be careful to notice the units of measurement of `propresp`.

This coefficient measures, for a marginal increase in the `propresp`, the variation in mountant of `gifts`.

(iv) Now add the variable `avggift` to the equation. What happens to the estimated effect of `mailsyear`?

```
lm3<-lm(gift~mailsyear+giftlast+propresp+avggift,data)

summary(lm3)
```

```
## 
## Call:
## lm(formula = gift ~ mailsyear + giftlast + propresp + avggift,
##     data = data)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -197.016   -5.883   -2.065    3.031  221.256
## 
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -7.32776    0.75822  -9.664  < 2e-16 ***
## mailsyear    1.20117    0.31242   3.845 0.000122 ***
## giftlast    -0.26086    0.01076 -24.251  < 2e-16 ***
## propresp    16.20464    0.81753  19.821  < 2e-16 ***
## avggift      0.52695    0.02108  24.996  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 13.47 on 4263 degrees of freedom
## Multiple R-squared:  0.2005, Adjusted R-squared:  0.1998
## F-statistic: 267.3 on 4 and 4263 DF,  p-value: < 2.2e-16
```

In this case, the coefficient of `mailsyear` reduces from 2.16 to 1.20.

(v) In the equation from part (iv), what has happened to the coefficient on giftlast? What do you think is happening?

The sign of coefficient of the variable `giflast` changes from positive to negative. Indicating a possible negative effect of variable `giftlast` in a model with more exogenous variables.