Chapter 3 - The Multiple Regression Analysis - Estimation

Thalles Quinaglia Liduares

27/02/2022

Exercise 6

Upload packages

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

Upload database

```
data<-wooldridge::wage2
attach(data)
```

Use the data set in WAGE2.RAW for this problem. As usual, be sure all of the following regressions contain an intercept.

(i) Run a simple regression of IQ on educ to obtain the slope coefficient, say, $\tilde{\delta}$

```
lm1<-lm(data$IQ~data$educ)

summary(lm1)
```

```
##
## Call:
## lm(formula = data$IQ ~ data$educ)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -50.228  -7.262   0.907   8.772  37.373
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  53.6872     2.6229   20.47   <2e-16 ***
## data$educ     3.5338     0.1922   18.39   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.9 on 933 degrees of freedom
## Multiple R-squared:  0.2659, Adjusted R-squared:  0.2652
## F-statistic:   338 on 1 and 933 DF,  p-value: < 2.2e-16
```

The coefficient $\tilde{\delta}$ is equal to 3.53.

(ii) Run the simple regression of log(wage) on educ, and obtain the slope coefficient, $\tilde{\beta}_1$.

```
lm2<-lm(data$lwage~data$educ)

summary(lm2)
```

```
##
## Call:
## lm(formula = data$lwage ~ data$educ)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.94620 -0.24832  0.03507  0.27440  1.28106
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.973063   0.081374   73.40   <2e-16 ***
## data$educ   0.059839   0.005963   10.04   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4003 on 933 degrees of freedom
## Multiple R-squared:  0.09742,    Adjusted R-squared:  0.09645
## F-statistic: 100.7 on 1 and 933 DF,  p-value: < 2.2e-16
```

The $\tilde{\beta}_1$ is equal to 0.05.

(iii) Run the multiple regression of log(wage) on educ and IQ, and obtain the slope coefficients, $\tilde{\beta}_1$ and $\tilde{\beta}_2$, respectively.

```
lm3<-lm(data$lwage~data$educ+data$IQ)

summary(lm3)
```

```
##
## Call:
## lm(formula = data$lwage ~ data$educ + data$IQ)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.01601 -0.24367  0.03359  0.27960  1.23783
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.6582877  0.0962408  58.793  < 2e-16 ***
## data$educ   0.0391199  0.0068382   5.721 1.43e-08 ***
## data$IQ     0.0058631  0.0009979   5.875 5.87e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3933 on 932 degrees of freedom
## Multiple R-squared:  0.1297, Adjusted R-squared:  0.1278
## F-statistic: 69.42 on 2 and 932 DF,  p-value: < 2.2e-16
```

The $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are equal to 0.03 and 0.005, respectively.

(iv) Verify that $\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}_1$.

Given the estimated coefficientes, we have that

$$0.03 + 0.005(3.53) = 0.05$$

Hence, the equality is satisfied.