Multiple Regression, Lesson 1

Henri Makika

Julho 01, 2019

Import data with information for CO2, GDP and % manufacturing

```
library(readr)
library(stargazer)
Data_CO2 <- read_csv("Data_CO2.csv")
head(Data_CO2)
```

```
## # A tibble: 6 x 4
##   country   co2    gdp   ind
##   <chr>   <dbl>  <dbl> <dbl>
## 1 ALB      0.99  1200.  19.0
## 2 DZA      2.88  1794.  58.6
## 3 AGO      0.69   656.  72.1
## 4 ATG      4.44 10334.  16.1
## 5 ARG      3.82  7696.  27.6
## 6 ARM      1.13   621.  39.0
```

Linear Regression by OLS

```
countries = Data_CO2
ols1 <- lm(co2 ~ gdp + ind, data = countries)
#stargazer(ols1, type = "text", digits = 8, column.labels = c("Linear model"),
          # keep.stat = c('n','rsq','adj.rsq','f'), out = "mrd.txt")
```

Interpretação dos resultados

Céteris paribus, se o PIB varie de uma unidade, o CO2 varie de 0.003. As variáveis consideradas são estatisticamente significantes, a hipótese nula : $H_0 : \beta_1 = \beta_2 = 0$ e a alternativa : $H_1 : \beta_1 \neq \beta_2 \neq 0$, no nosso caso a hipótese nula é rejeitada. Isso significa que as variáveis consideradas explicam perfeitamente o modelo estimado.

Lesson 2

Model log-lin

```
ols2 = lm(log(co2) ~ gdp + ind, data = countries)
```

Model lin-log

```
ols3 = lm(co2 ~ log(gdp) + ind, data = countries)
```

Model log-log

```
ols4 = lm(log(co2) ~ log(gdp) + ind, data = countries)

stargazer(ols1, ols2, ols3, ols4, type = "text", digits = 8,
          column.labels = c("Linear", "Log-lin", "Lin-log", "Log-log"),
          keep.stat = c('n','rsq','adj.rsq','f'), out = "mrd.txt")
```

```
##
## =========================================================================================
##                                                Dependent variable:
##                              ------------------------------------------------------------
##                                   co2          log(co2)          co2          log(co2)
##                                  Linear        Log-lin        Lin-log        Log-log
##                                   (1)            (2)            (3)            (4)
## -----------------------------------------------------------------------------------------
## gdp                          0.00032095***  0.00009556***
##                               (0.00002616)   (0.00000930)
##
## log(gdp)                                                   2.05877000***   0.82913180***
##                                                            (0.16245000)    (0.03856182)
##
## ind                          0.10506920***  0.04203434***  0.06058868***   0.02469790***
##                               (0.01986168)   (0.00705851)   (0.01978834)    (0.00469729)
##
## Constant                      -1.12589100*  -1.38169600*** -13.42203000*** -6.56817800***
##                               (0.67025630)   (0.23819790)   (1.31058700)    (0.31110270)
##
## -----------------------------------------------------------------------------------------
## Observations                      169            169            169            169
## R2                             0.51415990     0.45473730     0.52915380     0.76423830
## Adjusted R2                    0.50830640     0.44816790     0.52348090     0.76139780
## F Statistic (df = 2; 166) 87.83811000*** 69.22021000*** 93.27836000***  269.05030000***
## =========================================================================================
## Note:                                                        *p<0.1; **p<0.05; ***p<0.01
```

On choisit le modèle log-log pour que le coefficient de détermination R^2 est supérieur de tous les autres modèles adjustés.

Model with binary variable

```
countries$brics = 0

countries$brics[countries$country == "BRA" | countries$country == "RUS" |
                  countries$country == "IND" | countries$country == "CHN" |
                  countries$country == "ZAF"] <- 1
```

Log-log model with binary

```
ols5 = lm(log(co2) ~ log(gdp) + ind + brics, data = countries)
ols6 = lm(co2 ~ gdp + ind + brics, data = countries)
```

```
stargazer(ols5, ols6, type = "text", digits = 8, column.labels = c("Log-log Binary",
                                                                   "Linear"),
          keep.stat = c('n','rsq','adj.rsq','f'), out = "mrd.txt")
```

```
##
## ===========================================================
##                                 Dependent variable:
##                             ------------------------------
##                                log(co2)          co2
##                             Log-log Binary      Linear
##                                   (1)             (2)
## -----------------------------------------------------------
## log(gdp)                     0.83310940***
##                              (0.03792617)
##
## gdp                                          0.00032348***
##                                               (0.00002619)
##
## ind                          0.02395126***   0.10370920***
##                              (0.00462493)     (0.01985613)
##
## brics                        0.93491090***     1.95440900
##                              (0.35628580)     (1.55204200)
##
## Constant                    -6.60398400***    -1.15966700*
##                              (0.30603430)     (0.66961440)
##
## -----------------------------------------------------------
## Observations                      169              169
## R2                             0.77368270       0.51878460
## Adjusted R2                    0.76956790       0.51003520
## F Statistic (df = 3; 165) 188.02170000*** 59.29393000***
## ===========================================================
## Note:                           *p<0.1; **p<0.05; ***p<0.01
```

$\Delta ln(co2) = \frac{\Delta co2}{co2} = 93\%$

Para ter a variação exata

```
exp(ols5$coefficients[4]) -1
```

```
##    brics
## 1.546987
```

$\Delta ln(co2) = \frac{\Delta co2}{co2} = 154\%$ Os Brics emitem mais co2 154% em comparação ao outros países.