Chapter 2 - The Simple Regression Model

Thalles Quinaglia Liduares

09/02/2022

Exercise 2.1

```
library(wooldridge)
library(stargazer)

data1<-wooldridge::k401k

attach(data1)

View(data1)
```

(i) Find the average participation rate and the average match rate in the sample of plans.

```
ave_prate<-round(mean(prate),2)

ave_prate
```

```
## [1] 87.36
```

```
ave_mrate<-round(mean(mrate),2)

ave_mrate
```

```
## [1] 0.73
```

The average participation and match rate is 87.36% and 0.73%, respectively.

(ii) Now, estimate the simple regression equation

$$\widehat{prate} = \hat{\beta}_0 + \hat{\beta}_1 mrate$$

and report the results along with the sample size and R-squared.

```
library(lmreg)

lm1<-lm(prate~mrate, data1)

summary(lm1)
```

```
##
## Call:
## lm(formula = prate ~ mrate, data = data1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -82.303  -8.184   5.178  12.712  16.807
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  83.0755     0.5633  147.48   <2e-16 ***
## mrate         5.8611     0.5270   11.12   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.09 on 1532 degrees of freedom
## Multiple R-squared:  0.0747, Adjusted R-squared:  0.0741
## F-statistic: 123.7 on 1 and 1532 DF,  p-value: < 2.2e-16
```

The estimated intercept and mrate coefficient are 87.07 and 5.86, respectively.

(iii) Interpret the intercept in your equation. Interpret the coefficient on mrate.

The intercept represents the $prate$ value if $mrate = 0$. In this case, would be equal a 83.07%.

The mrate coefficient, represents the inclination of the curve. For 1% increase in mrate, the prate value would increase aproximatelly 89%.

(iv) Find the predicted prate when $mrate = 3.5$. Is this a reasonable prediction? Explain what is happening here.

$$prate = 83.07 + 5.86 \times 3.5$$

$$prate = 103.58$$

This value is not reasonable because it is larger than 100%.

(v) How much of the variation in prate is explained by mrate? Is this a lot in your opinion?

The R^2 is 0.074. Hence, only 7.4% of the variation in the dependent variable is explained by the independent variable.

It's a very low percent value of prediction.