

# Resolução de Exercício de Regressão Linear no R

Prof. Weligton Gomes

2023-06-19

## Exercícios da Nota de Aula sobre Regressão Linear no R

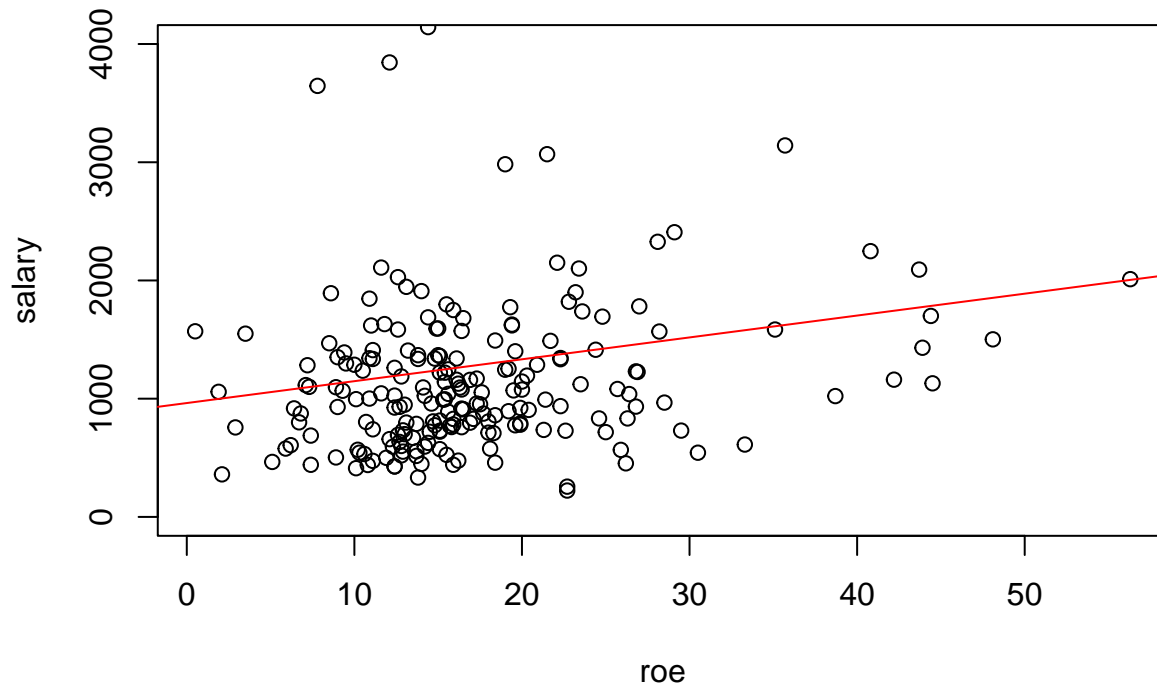
```
library(wooldridge)
```

### Exemplo 2.3:

```
data("ceosal1")
summary(ceosal1)
```

```
##      salary      pcsalary      sales      roe
## Min.   : 223      Min.   : -61.00      Min.   : 175.2      Min.   : 0.50
## 1st Qu.: 736      1st Qu.: -1.00      1st Qu.: 2210.3      1st Qu.:12.40
## Median : 1039      Median :  9.00      Median : 3705.2      Median :15.50
## Mean   : 1281      Mean   : 13.28      Mean   : 6923.8      Mean   :17.18
## 3rd Qu.: 1407      3rd Qu.: 20.00      3rd Qu.: 7177.0      3rd Qu.:20.00
## Max.   :14822      Max.   :212.00      Max.   :97649.9      Max.   :56.30
##      pcroe      ros      indus      finance
## Min.   : -98.9      Min.   : -58.0      Min.   :0.0000      Min.   :0.0000
## 1st Qu.: -21.2      1st Qu.: 21.0      1st Qu.:0.0000      1st Qu.:0.0000
## Median : -3.0      Median : 52.0      Median :0.0000      Median :0.0000
## Mean   : 10.8      Mean   : 61.8      Mean   :0.3206      Mean   :0.2201
## 3rd Qu.: 19.5      3rd Qu.: 81.0      3rd Qu.:1.0000      3rd Qu.:0.0000
## Max.   :977.0      Max.   :418.0      Max.   :1.0000      Max.   :1.0000
##      consprod      utility      lsalary      lsales
## Min.   :0.0000      Min.   :0.0000      Min.   :5.407      Min.   : 5.166
## 1st Qu.:0.0000      1st Qu.:0.0000      1st Qu.:6.601      1st Qu.: 7.701
## Median :0.0000      Median :0.0000      Median :6.946      Median : 8.217
## Mean   :0.2871      Mean   :0.1722      Mean   :6.950      Mean   : 8.292
## 3rd Qu.:1.0000      3rd Qu.:0.0000      3rd Qu.:7.249      3rd Qu.: 8.879
## Max.   :1.0000      Max.   :1.0000      Max.   :9.604      Max.   :11.489
```

```
attach(ceosal1)
CEOregress <- lm(salary ~ roe, data = ceosal1)
plot(roe, salary, ylim=c(0,4000))
abline(CEOregress, col='red')
```



```
summary(CE0regress)
```

```
##
## Call:
## lm(formula = salary ~ roe, data = ceosal1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1160.2  -526.0  -254.0   138.8 13499.9
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    963.19     213.24   4.517 1.05e-05 ***
## roe             18.50       11.12   1.663  0.0978 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1367 on 207 degrees of freedom
## Multiple R-squared:  0.01319,    Adjusted R-squared:  0.008421
## F-statistic: 2.767 on 1 and 207 DF,  p-value: 0.09777
```

*salary* é o salário anual de CEO em milhares de dólares e *#roe#* é o retorno médio sobre o patrimônio em percentual. O parâmetro  $\beta_1$  irá medir a variação no salário anual quando o retorno médio sobre o patrimônio aumentar em um ponto percentual.

um *roe* = 0, teremos um salário previsto de 963,19 ou US\$ 963.191, que é o intercepto. Ademais, se  $\Delta roe = 1$ , então  $\Delta salary = 18,5$  ou US\$ 18.501. Podemos, por fim, desenhar a reta de regressão com o código abaixo.

## Exercício 2.10:

```
data("k401k")
# View(k401k)
```

```
summary(k401k)
```

```
##      prate      mrate      totpart      totelg
## Min.   : 3.00   Min.   :0.0100   Min.   : 50.0   Min.   : 51.0
## 1st Qu.: 78.03   1st Qu.:0.3000   1st Qu.: 156.2   1st Qu.: 176.0
## Median : 95.70   Median :0.4600   Median : 276.0   Median : 330.0
## Mean   : 87.36   Mean   :0.7315   Mean   : 1354.2   Mean   : 1628.5
## 3rd Qu.:100.00   3rd Qu.:0.8300   3rd Qu.: 749.5   3rd Qu.: 890.5
## Max.   :100.00   Max.   :4.9100   Max.   :58811.0   Max.   :70429.0
##      age      totemp      sole      ltotemp
## Min.   : 4.00   Min.   : 58   Min.   :0.0000   Min.   : 4.060
## 1st Qu.: 7.00   1st Qu.: 261   1st Qu.:0.0000   1st Qu.: 5.565
## Median : 9.00   Median : 588   Median :0.0000   Median : 6.377
## Mean   :13.18   Mean   : 3568   Mean   :0.4876   Mean   : 6.686
## 3rd Qu.:18.00   3rd Qu.: 1804   3rd Qu.:1.0000   3rd Qu.: 7.498
## Max.   :51.00   Max.   :144387   Max.   :1.0000   Max.   :11.880
```

## Estimação do modelo

```
modelo<-lm(prate~mrate, data = k401k)
summary(modelo)
```

```
##
## Call:
## lm(formula = prate ~ mrate, data = k401k)
##
## 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
```

## Interprete o intercepto de sua equação. Interprete o coeficiente de mrate.

Resposta: 83.075455 é o valor esperado da taxa de participação (prate) quando a taxa de contribuição (mrate) é zero.

Ache o prate predito quando mrate = 3, 5. Essa predição é razoável? Explique o que está acontecendo.

Resposta:

```
modelo$coefficients[1] + modelo$coefficients[2] * 3.5
```

```
## (Intercept)
##      103.5892
```

De acordo com as estimações, se  $\Delta mrate = 3.5$ , haverá uma taxa de participação ou no número de contas ativas no plano da ordem de, aproximadamente, 103,6%, o que não faz sentido, pois a participação total dos funcionários da empresa encerraria um total de 100,0%.

**Quanto da variação em *prate* é explicado por *mrte*? Em sua opinião, isso é bastante?**

Resposta: O  $R^2$  desta estimação foi de 0.0747 implicando que 7,47% das variações em *prate* são capazes de explicar as variações de *mrte*.

## Exercício 2

```
data("ceosal2")
```

**Ache o salário médio (*salary*) e a permanência média (*ceoten*) da amostra.**

Resposta:

```
summary(ceosal2)
```

```
##      salary      age      college      grad
## Min.   : 100.0   Min.   :33.00   Min.   :0.0000   Min.   :0.0000
## 1st Qu.: 471.0   1st Qu.:52.00   1st Qu.:1.0000   1st Qu.:0.0000
## Median : 707.0   Median :57.00   Median :1.0000   Median :1.0000
## Mean   : 865.9   Mean   :56.43   Mean   :0.9718   Mean   :0.5311
## 3rd Qu.:1119.0   3rd Qu.:62.00   3rd Qu.:1.0000   3rd Qu.:1.0000
## Max.   :5299.0   Max.   :86.00   Max.   :1.0000   Max.   :1.0000
##      comten      ceoten      sales      profits
## Min.   : 2.0     Min.   : 0.000   Min.   : 29     Min.   : -463.0
## 1st Qu.:12.0     1st Qu.: 3.000   1st Qu.: 561    1st Qu.: 34.0
## Median :23.0     Median : 6.000   Median : 1400   Median : 63.0
## Mean   :22.5     Mean   : 7.955   Mean   : 3529   Mean   : 207.8
## 3rd Qu.:33.0     3rd Qu.:11.000   3rd Qu.: 3500   3rd Qu.: 208.0
## Max.   :58.0     Max.   :37.000   Max.   :51300   Max.   :2700.0
##      mktval      lsalary      lsales      lmktval
## Min.   : 387     Min.   :4.605   Min.   : 3.367   Min.   : 5.958
## 1st Qu.: 644     1st Qu.:6.155   1st Qu.: 6.330   1st Qu.: 6.468
## Median : 1200    Median :6.561   Median : 7.244   Median : 7.090
## Mean   : 3600    Mean   :6.583   Mean   : 7.231   Mean   : 7.399
## 3rd Qu.: 3500    3rd Qu.:7.020   3rd Qu.: 8.161   3rd Qu.: 8.161
## Max.   :45400    Max.   :8.575   Max.   :10.845   Max.   :10.723
##      comtensq      ceotensq      profmarg
## Min.   : 4.0     Min.   : 0.0    Min.   : -203.077
## 1st Qu.:144.0     1st Qu.: 9.0    1st Qu.: 4.231
## Median : 529.0    Median : 36.0    Median : 6.834
## Mean   : 656.7    Mean   :114.1    Mean   : 6.420
## 3rd Qu.:1089.0    3rd Qu.:121.0    3rd Qu.:10.947
## Max.   :3364.0    Max.   :1369.0    Max.   :47.458
```

Quantos CEOs estão em seu primeiro ano na posição de CEO (isto é,  $ceoten = 0$ )? Qual é a permanência mais longa como CEO?

```
table(ceosal2 == 0)
```

```
##  
## FALSE TRUE  
## 2557 98
```

```
max(ceosal2$ceoten)
```

```
## [1] 37
```

## Estime o modelo de regressão simples

```
modelo2 <- lm(salary ~ ceoten, data = ceosal2)  
summary(modelo2)
```

```
##  
## Call:  
## lm(formula = salary ~ ceoten, data = ceosal2)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -977.8 -345.9 -169.4  263.8 4373.9   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)  772.426     65.676  11.761  <2e-16 ***  
## ceoten       11.746       6.148   1.911  0.0577 .    
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 583.2 on 175 degrees of freedom  
## Multiple R-squared:  0.02043,    Adjusted R-squared:  0.01484   
## F-statistic: 3.651 on 1 and 175 DF,  p-value: 0.05769
```

Qual é o aumento do salário predito (aproximado) no salário, dado um ano a mais como CEO?

```
coefficients(modelo2)[2]
```

```
## ceoten  
## 11.74613
```

## Estime o modelo de regressão simples (log-lin)

```
modelo3 <- lm(lsalary ~ ceoten, data = ceosal2)
```

Qual é o aumento da percentagem predita (aproximado) no salário, dado um ano a mais como CEO?

```
summary(modelo3)
```

```
##
## Call:
## lm(formula = lsalary ~ ceoten, data = ceosal2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.15314 -0.38319 -0.02251  0.44439  1.94337
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.505498   0.067991  95.682  <2e-16 ***
## ceoten       0.009724   0.006364   1.528   0.128
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6038 on 175 degrees of freedom
## Multiple R-squared:  0.01316,    Adjusted R-squared:  0.007523
## F-statistic: 2.334 on 1 and 175 DF,  p-value: 0.1284
```