Modelos de Regressão e Previsão

Modelo de Regressão Linear Simples: Estimação

Prof. Carlos Trucíos
carlos.trucios@facc.ufrj.br
ctruciosm.github.io

Faculdade de Administração e Ciências Contábeis
Universidade Federal do Rio de Janeiro

Aula 2

MRLS: Estimação

MRLS: Estimação

Na prática, nunca conhecemos os valores de β_0, β_1 então precisamos estima-los utilizando os dados. Estes valores estimados serão denotados por $\hat{\beta}_0$ e $\hat{\beta}_1$, respectivamente.

MRLS: Estimação

Na prática, nunca conhecemos os valores de β_0, β_1 então precisamos estima-los utilizando os dados. Estes valores estimados serão denotados por $\hat{\beta}_0$ e $\hat{\beta}_1$, respectivamente.

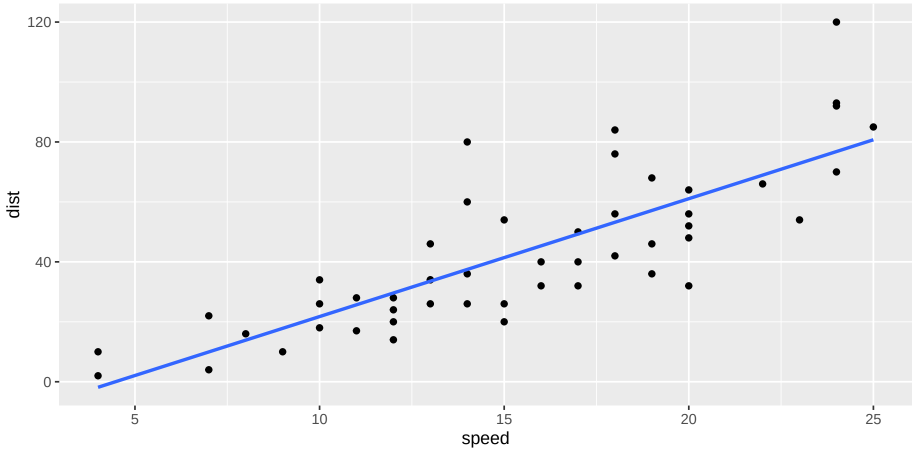
Sejam $(x_1, y_1), \ldots, (x_n, y_n)$ uma amostra aleatória (a.a.) de tamanho n da população, então

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

onde i indica a i-ésima observação.

MRLS: Estimação

Estamos interessados numa reta do tipo:



Métodos de estimação

- Usaremos as suposições descritas na aula anterior:

Métodos de estimação

- Usaremos as suposições descritas na aula anterior:
- $E(u) = 0$

Métodos de estimação

- Usaremos as suposições descritas na aula anterior:
- $E(u) = 0$
- $Cov(xu) = E(xu) = 0$ (consequencia de $E(u|x) = E(u)$)

Métodos de estimação

- Usaremos as suposições descritas na aula anterior:
- $E(u) = 0$
- $Cov(xu) = E(xu) = 0$ (consequencia de $E(u|x) = E(u)$)
- Então,
$$E(u) = E(y - \beta_0 - \beta_1 x) = 0$$
e
$$E(xu) = E(x(y - \beta_0 - \beta_1 x)) = 0$$

Métodos de estimação

- Usaremos as suposições descritas na aula anterior:
- $E(u) = 0$
- $Cov(xu) = E(xu) = 0$ (consequencia de $E(u|x) = E(u)$)
- Então,

$$E(u) = E(y - \beta_0 - \beta_1 x) = 0$$

e

$$E(xu) = E(x(y - \beta_0 - \beta_1 x)) = 0$$

Métodos de estimação

- Usaremos as suposições descritas na aula anterior:
- $E(u) = 0$
- $Cov(xu) = E(xu) = 0$ (consequencia de $E(u|x) = E(u)$)
- Então,

$$E(u) = E(y - \beta_0 - \beta_1 x) = 0$$

e

$$E(xu) = E(x(y - \beta_0 - \beta_1 x)) = 0$$

Método dos momentos

Resolver o sistema resultante de igualar os momentos populacionais com os momentos amostrais

Métodos de estimação: MM

Método dos Momentos (MM)

$$n^{-1}\sum_{i=1}^{n}\hat{u}_i = n^{-1}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = E(u) = 0 \tag{1}$$

$$n^{-1}\sum_{i=1}^{n}x_i\hat{u}_i = n^{-1}\sum_{i=1}^{n}(x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)) = E(xu) = 0 \tag{2}$$

Métodos de estimação: MM

Método dos Momentos (MM)

$$n^{-1}\sum_{i=1}^{n}\hat{u}_i = n^{-1}\sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1x_i) = E(u) = 0 \tag{1}$$

$$n^{-1}\sum_{i=1}^{n}x_i\hat{u}_i = n^{-1}\sum_{i=1}^{n}(x_i(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)) = E(xu) = 0 \tag{2}$$

Então, de (1)

$$\bar{y} = \hat{\beta}_0+\hat{\beta}_1\bar{x} \Longrightarrow \hat{\beta}_0 = \bar{y}-\hat{\beta}_1\bar{x}$$

(1) e (2) são chamadas ***condições de primeira ordem***

Métodos de estimação: MM

Método dos Momentos (MM)

$$n^{-1}\sum_{i=1}^{n}\hat{u}_i = n^{-1}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = E(u) = 0 \qquad (1)$$

$$n^{-1}\sum_{i=1}^{n}x_i\hat{u}_i = n^{-1}\sum_{i=1}^{n}(x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)) = E(xu) = 0 \qquad (2)$$

Então, de (1)

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1\bar{x} \Longrightarrow \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

*(1) e (2) são chamadas **condições de primeira ordem***

Das condições de primeria ordem temos que $\bar{\hat{u}} = 0$ e $cov(x, \hat{u}) = 0$

Métodos de estimação: MM

De (2),

$$\sum_{i=1}^{n} \left[x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)\right] = \sum_{i=1}^{n} \left[x_i(y_i - \underbrace{(\bar{y} - \hat{\beta}_1 \bar{x})}_{\hat{\beta}_0} - \hat{\beta}_1 x_i)\right] = 0 \times n = 0$$

Métodos de estimação: MM

De (2),

$$\sum_{i=1}^{n}\left[x_i(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)\right]=\sum_{i=1}^{n}\left[x_i(y_i-\underbrace{(\bar{y}-\hat{\beta}_1\bar{x})}_{\hat{\beta}_0}-\hat{\beta}_1x_i)\right]=0\times n=0$$

Então,

$$\sum_{i=1}^{n}[x_i(y_i-\bar{y})]-\hat{\beta}_1\sum_{i=1}^{n}[x_i(x_i-\bar{x})]=0$$

Métodos de estimação: MM

De (2),

$$\sum_{i=1}^{n}\left[x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)\right] = \sum_{i=1}^{n}\left[x_i(y_i - \underbrace{(\bar{y} - \hat{\beta}_1 \bar{x})}_{\hat{\beta}_0} - \hat{\beta}_1 x_i)\right] = 0 \times n = 0$$

Então,

$$\sum_{i=1}^{n}[x_i(y_i - \bar{y})] - \hat{\beta}_1 \sum_{i=1}^{n}[x_i(x_i - \bar{x})] = 0$$

Logo,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}[x_i(y_i - \bar{y})]}{\sum_{i=1}^{n}[x_i(x_i - \bar{x})]} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

Métodos de estimação: MM

Então,

Estimadores MM

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} = \frac{cov(x,y)}{var(x)} = \hat{\rho}_{xy}\frac{\hat{\sigma}_y}{\hat{\sigma}_x}$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

Desde que $\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 > 0$.

Onde

- $\hat{\rho}_{xy}$ é a correlação amostral entre x e y;
- $\hat{\sigma}_x$ é o desvio padrão amostral de x e
- $\hat{\sigma}_y$ é o desvio padrão amostral de y.

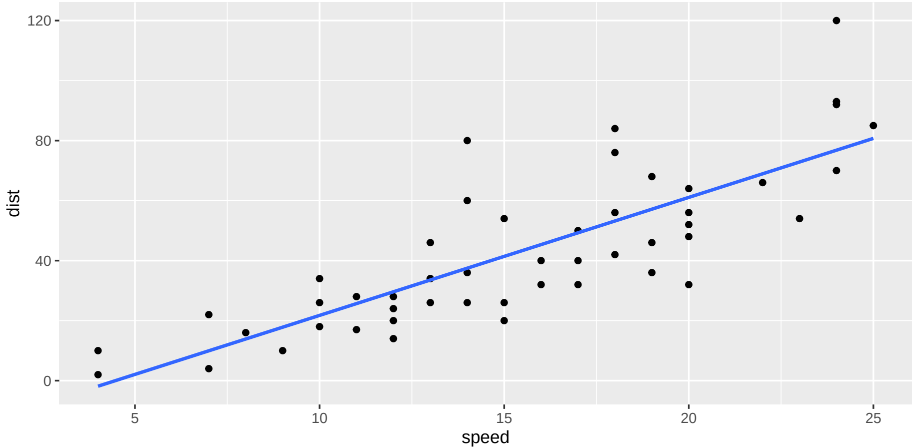
Métodos de estimação: MM - Apêndice

1. $$\sum_{i=1}^{n}[x_i(x_i-\bar{x})] = \sum_{i=1}^{n}(x_i-\bar{x})^2?$$
   $$\sum_{i=1}^{n}[x_i(x_i-\bar{x})] = \sum_{i=1}^{n}[(x_i-\bar{x}+\bar{x})(x_i-\bar{x})] =$$
   $$\sum_{i=1}^{n}(x_i-\bar{x})^2 + \bar{x}\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})}_{0}$$

2. $$\sum_{i=1}^{n}[x_i(y_i-\bar{y})] = \sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]?$$
   $$\sum_{i=1}^{n}[x_i(y_i-\bar{y})] = \sum_{i=1}^{n}[(x_i-\bar{x}+\bar{x})(y_i-\bar{y})] =$$
   $$\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})] + \underbrace{\sum_{i=1}^{n}[\bar{x}(y_i-\bar{y})]}_{\bar{x}\sum_{i=1}^{n}(y_i-\bar{y})} = \sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]$$

Métodos de estimação: MQO

Mínimos Quadrados Ordinaios (MQO)



Métodos de estimação: MQO

Outra forma de obter $\hat{\beta}_0$ e $\hat{\beta}_1$ é atraves de MQO.

- A ideia é minimizar a soma dos quadrados dos resíduos $\hat{u}_i$ onde $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$

Métodos de estimação: MQO

Outra forma de obter $\hat{\beta}_0$ e $\hat{\beta}_1$ é atraves de MQO.

- A ideia é minimizar a soma dos quadrados dos resíduos $\hat{u}_i$ onde $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$
- Seja $SQR = \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} \big(y_i - b_0 - b_1 x_i\big)^2$, então

$$\hat{\beta}_0, \hat{\beta}_1 = \underset{b_0, b_1}{\operatorname{argmin}}\, SQR$$

Métodos de estimação: MQO

Outra forma de obter $\hat{\beta}_0$ e $\hat{\beta}_1$ é atraves de MQO.

- A ideia é minimizar a soma dos quadrados dos resíduos $\hat{u}_i$ onde $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$
- Seja $SQR = \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} \left(y_i - b_0 - b_1 x_i\right)^2$, então

$$\hat{\beta}_0, \hat{\beta}_1 = \underset{b_0, b_1}{\operatorname{argmin}}\, SQR$$

- Derivando w.r.t b_0 e b_1 temos

$$\frac{\partial SQR}{\partial b_0} = -2 \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i); \quad \frac{\partial SQR}{\partial b_1} = -2 \sum_{i=1}^{n} x_i (y_i - b_0 - b_1 x_i)$$

Métodos de estimação: MQO

Outra forma de obter $\hat{\beta}_0$ e $\hat{\beta}_1$ é atraves de MQO.

- A ideia é minimizar a soma dos quadrados dos resíduos $\hat{u}_i$ onde $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$
- Seja $SQR = \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} \big(y_i - b_0 - b_1 x_i\big)^2$, então

$$\hat{\beta}_0, \hat{\beta}_1 = \underset{b_0, b_1}{\operatorname{argmin}}\, SQR$$

- Derivando w.r.t b_0 e b_1 temos

$$\frac{\partial SQR}{\partial b_0} = -2\sum_{i=1}^{n} \big(y_i - b_0 - b_1 x_i\big); \ \frac{\partial SQR}{\partial b_1} = -2\sum_{i=1}^{n} x_i \big(y_i - b_0 - b_1 x_i\big)$$

- Igualando a zero, chegamos ao mesmo sistema nas Equações (1) e (2).

Métodos de estimação: MQO

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \text{ e } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

Como sabemos que de fato são os valores que minimizam SQR?

Métodos de estimação: MQO

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \text{ e } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

Como sabemos que de fato são os valores que minimizam SQR?

A matriz $\begin{bmatrix} \frac{\partial^2 SQR}{\partial b_0^2} & \frac{\partial^2 SQR}{\partial b_0 \partial b_1} \\ \frac{\partial^2 SQR}{\partial b_1 \partial b_0} & \frac{\partial^2 SQR}{\partial b_1^2} \end{bmatrix}$ deve ser definida positiva (**todos os autovalores devem ser positivos**)

Métodos de estimação: MQO

Uma vez obtidos $\hat{\beta}_0$ e $\hat{\beta}_1$, podemos obter **y chapéu**

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

($\hat{y}$: valores ajustados/previstos da regressão)

Métodos de estimação: MQO

Uma vez obtidos $\hat{\beta}_0$ e $\hat{\beta}_1$, podemos obter **y chapéu**

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

($\hat{y}$: valores ajustados/previstos da regressão)

Veja que

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x \text{ ou equivalentemente } \hat{\beta}_1 = \Delta\hat{y}/\Delta x$$

Métodos de estimação: MQO

Uma vez obtidos $\hat{\beta}_0$ e $\hat{\beta}_1$, podemos obter **y chapéu**

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

($\hat{y}$: valores ajustados/previstos da regressão)

Veja que

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x \text{ ou equivalentemente } \hat{\beta}_1 = \Delta\hat{y}/\Delta x$$

$\hat{\beta}_1$ nos diz quanto varia $\hat{y}$ quando x aumenta em uma unidade

Exemplos no R

Exemplos no R: salarios CEO

Informações sobre o *salario* (anual em milhares de dólares) e o retorno médio sobre o patrimônio líquido dos últimos 3 anos, *roe* (definido como uma porcentagem do patrimônio líquido) de 209 CEOs estão disponíveis no arquivo **ceosal1.txt**.

Queremos estudar a relação

$$\text{salario} = \beta_0 + \beta_1 \text{roe} + u$$

Exemplos no R: salarios CEO

```
CEOSAL1 = read.table("./DadosMRP/ceosal1.txt")[,c(1,4)]
colnames(CEOSAL1) = c("salario", "roe")
head(CEOSAL1,3)
```

```
##   salario  roe
## 1    1095 14.1
## 2    1001 10.9
## 3    1122 23.5
```

Exemplos no R: salarios CEO

```
CEOSAL1 = read.table("./DadosMRP/ceosal1.txt")[,c(1,4)]
colnames(CEOSAL1) = c("salario", "roe")
head(CEOSAL1,3)
```

```
##   salario  roe
## 1    1095 14.1
## 2    1001 10.9
## 3    1122 23.5
```

- Para o CEO 1, temos um salário anual de 1095.000 USD e um retorno médio sobre o patrimônio líquido de 14.1%
- Para o CEO 2, temos um salário anual de 1001.000 USD e um retorno médio sobre o patrimônio líquido de 10.9%

Exemplos no R: salarios CEO

```r
summary(CEOSAL1)
```

```
##     salario             roe       
##  Min.   :  223   Min.   : 0.50  
##  1st Qu.:  736   1st Qu.:12.40  
##  Median : 1039   Median :15.50  
##  Mean   : 1281   Mean   :17.18  
##  3rd Qu.: 1407   3rd Qu.:20.00  
##  Max.   :14822   Max.   :56.30  
```

Exemplos no R: salarios CEO

$$\hat{\beta}_1 = \frac{cov(x,y)}{var(x)} = \hat{\rho}\frac{\hat{\sigma}_y}{\hat{\sigma}_x} \quad \text{e} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

```
# betas (hat)
b1_hat = cov(CEOSAL1)[1,2]/var(CEOSAL1$roe)
b0_hat = mean(CEOSAL1$salario) - b1_hat*mean(CEOSAL1$roe)
c(b0_hat,b1_hat)
```

```
## [1] 963.19133  18.50119
```

```
# Outra forma beta1_hat
cor(CEOSAL1)[1,2]*(sd(CEOSAL1$salario)/sd(CEOSAL1$roe))
```

```
## [1] 18.50119
```

Exemplos no R: salarios CEO

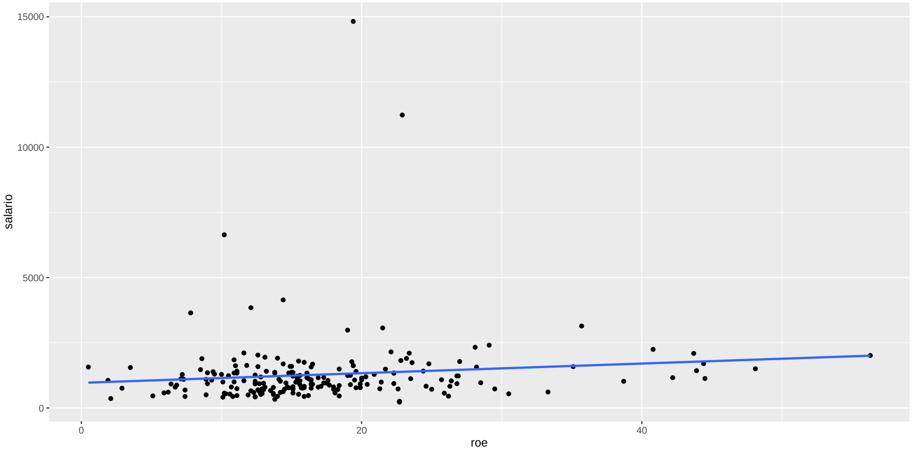
```r
lm(salario~roe, data = CEOSAL1)
```

```
##
## Call:
## lm(formula = salario ~ roe, data = CEOSAL1)
##
## Coefficients:
## (Intercept)          roe
##       963.2         18.5
```

$$\widehat{\text{salario}} = 963.2 + 18.5 \text{ roe}$$

Se o roe aumentar 1%, espera-se que o salario anual aumente em 18.500 USD.

Exemplos no R: salarios CEO



Exemplos no R: resultados eleitorais US

Os gastos de campanha de 173 disputas eleitorais entre dois partidos para o *House of Representatives dos US* estão contidos no arquivo **vote1.txt**. *votosA* representa a porcentagem de votos recebidos pelo candidato A (Há apenas 2 candidatos A e B em cada disputa) e *despesasA* representa a porcentagem das despesas totais em campanha do total que cabe ao candidato A.

Exemplos no R: resultados eleitorais US

```
VOTE1 = read.table("./DadosMRP/vote1.txt")[,c(4,10)]
colnames(VOTE1) = c("votosA", "despesasA")
head(VOTE1,3)
```

```
##   votosA despesasA
## 1     68     97.41
## 2     62     60.88
## 3     73     97.01
```

Exemplos no R: resultados eleitorais US

```r
VOTE1 = read.table("./DadosMRP/vote1.txt")[,c(4,10)]
colnames(VOTE1) = c("votosA", "despesasA")
head(VOTE1,3)
```

```
##   votosA despesasA
## 1     68     97.41
## 2     62     60.88
## 3     73     97.01
```

- Obs1: Candidato A obteve 68% dos votos e gastou o 97.41% do assignado para a campanha.
- Obs2: Candidato A obteve 62% dos votos e gastou o 60.88% do assignado para a campanha.

Exemplos no R: resultados eleitorais US

```
summary(VOTE1)
```

```
##      votosA          despesasA
##  Min.   :16.0   Min.   : 0.09
##  1st Qu.:36.0   1st Qu.:18.87
##  Median :50.0   Median :50.85
##  Mean   :50.5   Mean   :51.08
##  3rd Qu.:65.0   3rd Qu.:84.26
##  Max.   :84.0   Max.   :99.50
```

Exemplos no R: resultados eleitorais US

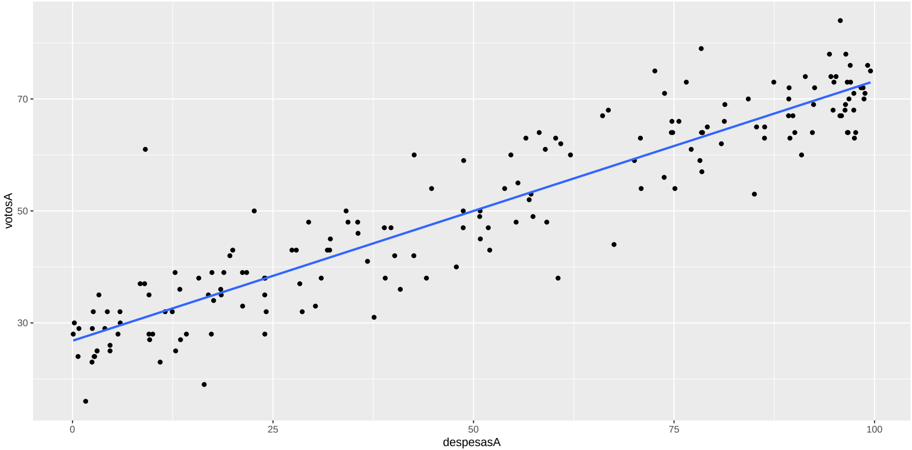
```r
lm(votosA~despesasA, data = VOTE1)
```

```
##
## Call:
## lm(formula = votosA ~ despesasA, data = VOTE1)
##
## Coefficients:
## (Intercept)    despesasA
##     26.8125       0.4638
```

$$\widehat{\text{votosA}} = 26.8125 + 0.4638 \text{ despesasA}$$

Se aumentarmos o gasto de despesas de campanha em 1%, espera-se que o candidato A receba 0.46% a mais de votos.

Exemplos no R: resultados eleitorais US



Qualidade de ajuste

SQT, SQE, SQR

Temos que $y_i - \hat{y}_i = \hat{u}_i \Longrightarrow y_i = \hat{y}_i + \hat{u}_i$, então

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\underbrace{y_i - \hat{y}_i}_{\hat{u}_i} + \hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{u}_i + \hat{y}_i - \bar{y})^2$$

$$\underbrace{\sum_{i=1}^{n}(y_i - \bar{y})^2}_{SQT} = \underbrace{\sum_{i=1}^{n}\hat{u}_i^2}_{SQR} + \underbrace{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}_{SQE} + 2\underbrace{\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i - \bar{y})}_{0}$$

$$SQT = SQR + SQE$$

Qualidade de ajuste

- Como sabemos se a variável x ajuda a explicar bem a variável y?

Qualidade de ajuste

- Como sabemos se a variável x ajuda a explicar bem a variável y?
- Termos um número que nos ajude a entender quão bem a reta de regressão se ajusta aos dados é bastante útil.

Qualidade de ajuste

- Como sabemos se a variável x ajuda a explicar bem a variável y?
- Termos um número que nos ajude a entender quão bem a reta de regressão se ajusta aos dados é bastante útil.
- O R^2 (R-quadrado) nos ajudará a responder esta pergunta.

Qualidade de ajuste

- Como sabemos se a variável x ajuda a explicar bem a variável y?
- Termos um número que nos ajude a entender quão bem a reta de regressão se ajusta aos dados é bastante útil.
- O R^2 (R-quadrado) nos ajudará a responder esta pergunta.

Qualidade de ajuste

- Como sabemos se a variável x ajuda a explicar bem a variável y?
- Termos um número que nos ajude a entender quão bem a reta de regressão se ajusta aos dados é bastante útil.
- O R^2 (R-quadrado) nos ajudará a responder esta pergunta.

$$SQT = SQR + SQE$$

$$1 = \frac{SQR}{SQT} + \underbrace{\frac{SQE}{SQT}}_{R^2}$$

Qualidade de ajuste

R-quadrado ou coeficiente de determinação

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}$$

Qualidade de ajuste

R-quadrado ou coeficiente de determinação

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}$$

- $0 \leq R^2 \leq 1$

Qualidade de ajuste

R-quadrado ou coeficiente de determinação

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}$$

- $0 \leq R^2 \leq 1$
- razão entre a variação explicada e a variação total.

Qualidade de ajuste

R-quadrado ou coeficiente de determinação

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}$$

- $0 \leq R^2 \leq 1$
- razão entre a variação explicada e a variação total.
- $100 \times R^2$ é a porcentagem de variação de y explicada pelo modelo.

Qualidade de ajuste

R-quadrado ou coeficiente de determinação

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}$$

- $0 \leq R^2 \leq 1$
- razão entre a variação explicada e a variação total.
- $100 \times R^2$ é a porcentagem de variação de y explicada pelo modelo.
- R^2 próximo de **0** indica um ajuste **ruim**, R^2 próximo de **1** indica um **bom** ajuste.

Qualidade de ajuste

R-quadrado ou coeficiente de determinação

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}$$

- $0 \leq R^2 \leq 1$
- razão entre a variação explicada e a variação total.
- $100 \times R^2$ é a porcentagem de variação de y explicada pelo modelo.
- R^2 próximo de **0** indica um ajuste **ruim**, R^2 próximo de **1** indica um **bom** ajuste.
- R^2 é o quadrado do coeficiente de correlação entre y e $\hat{y}$.

Qualidade de ajuste

Como é o ajuste em cada um dos modelos?

```
modelo1 = lm(salario~roe, data = CEOSAL1)
cor(CEOSAL1$salario,fitted.values(modelo1))^2
```

```
## [1] 0.01318862
```

```
summary(modelo1)$r.squared
```

```
## [1] 0.01318862
```

```
modelo2 = lm(votosA~despesasA, data = VOTE1)
summary(modelo2)$r.squared
```

```
## [1] 0.8561459
```

Leituras recomendadas

Leituras recomendadas

- Wooldridge, Jeffrey M. *Introdução à Econometria: Uma abordagem moderna*. (2016). Cengage Learning. – **Cap 2.2-2.3**