

# Regressão linear simples: um exemplo

*Fernando Mayer*

*18 de setembro, 2015*

## Contents

Definição do modelo	1
Resumo dos dados	1
Ajuste do modelo	1
Análise dos resíduos	2
Predição	3

## Definição do modelo

Abaixo tem-se o ajuste do modelo de regressão linear simples

$$Y = \beta_0 + \beta_1 x + \epsilon$$

para dados de distância para parada completa de um veículo em função da velocidade no instante de acionamento dos freios. Esse conjunto de dados está objeto `cars`.

## Resumo dos dados

Um resumo dos dados pode ser visto na tabela abaixo:

speed	dist
Min. : 4.0	Min. : 2.00
1st Qu.:12.0	1st Qu.: 26.00
Median :15.0	Median : 36.00
Mean :15.4	Mean : 42.98
3rd Qu.:19.0	3rd Qu.: 56.00
Max. :25.0	Max. :120.00

## Ajuste do modelo

Os parâmetros do modelo linear podem ser obtidos pela função `lm()`, como pode ser visto abaixo

```
## Ajuste do modelo
mod <- lm(dist ~ speed, data = cars)
```

Um resumo do ajuste do modelo pode ser visualizado abaixo

```
## Resumo do modelo
summary(mod)
```

Call:

```
lm(formula = dist ~ speed, data = cars)
```

Residuals:

Min	1Q	Median	3Q	Max
-29.069	-9.525	-2.272	9.215	43.201

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-17.5791	6.7584	-2.601	0.0123 *
speed	3.9324	0.4155	9.464	1.49e-12 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.38 on 48 degrees of freedom

Multiple R-squared: 0.6511, Adjusted R-squared: 0.6438

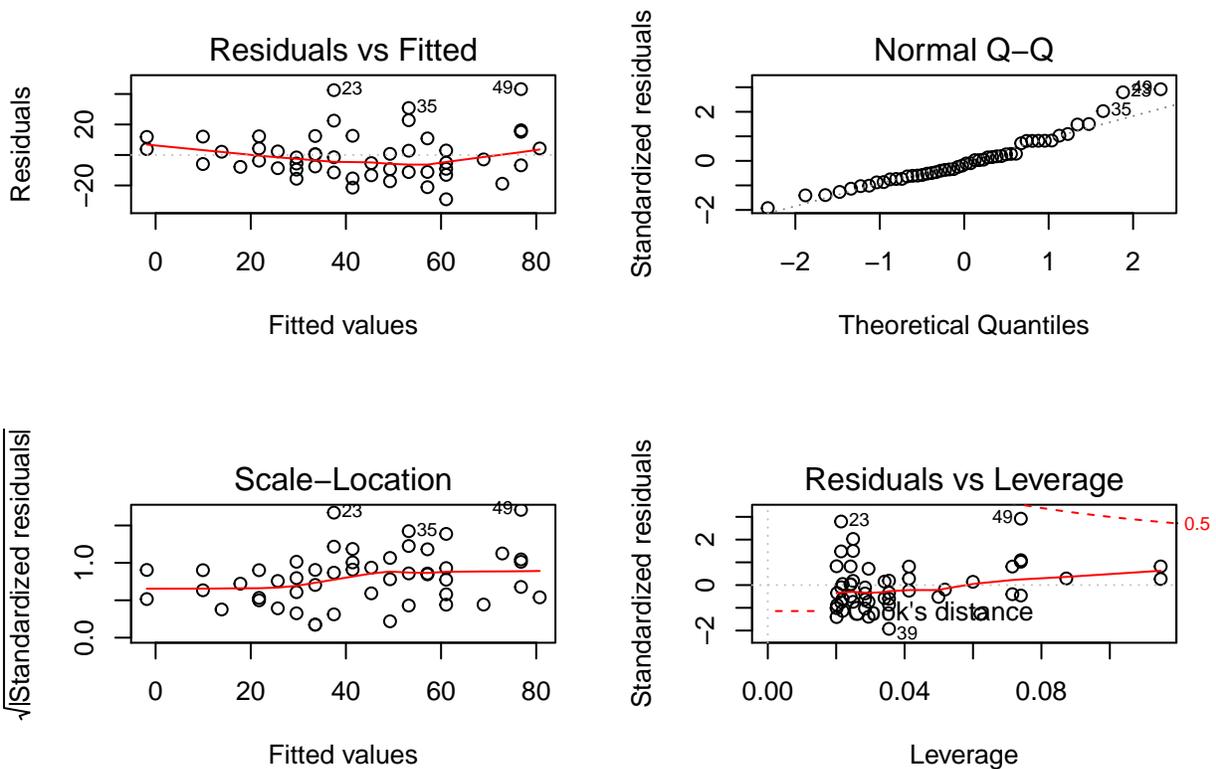
F-statistic: 89.57 on 1 and 48 DF, p-value: 1.49e-12

## Análise dos resíduos

Antes de interpretar o modelo e fazer inferência, é importante que sejam verificados os pressupostos considerados através da análise dos resíduos.

Por padrão se fizermos um `plot()` no modelo, temos os seguintes gráficos para avliar os resíduos:

```
plot(mod)
```



## Predição

Assumindo que o ajuste do modelo está adequado, pode-se obter o gráfico dos valores preditos sobre o diagrama de dispersão dos valores observados.

```
## Intervalo de valores para predição
pred <- data.frame(speed = seq(4, 25, by = 0.5))
## Valores preditos
ypred <- predict(mod, newdata = pred, interval = "confidence")
## Inclui valores preditos no data frame
pred <- cbind(pred, ypred)
```

Agora podemos visualizar o modelo ajustado aos dados

```
## Gráfico de dispersão
plot(dist ~ speed, data = cars,
     xlab = "Velocidade ao acionar os freios",
     ylab = "Distância para parada completa")
## Intervalo de predição
with(pred,
     matlines(x = speed, y = cbind(fit, lwr, upr),
             lty = c(1, 2, 2), col = 1))
```

