

# **Análise de variância e correlação**

## **Métodos Quantitativos Aplicados à Ciência Política**

Frederico Bertholini

07.dez.2020

## **1** Formalizando Hipóteses

## **2** Testando diferentes hipóteses

## **3** Medidas de associação

## **4** Modelo Linear Simples

## **5** Modelos Lineares Multivariados

Análise de  
variância e  
correlação

Frederico  
Bertholini

**Formalizando  
Hipóteses**

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

# Formalizando Hipóteses

# Diferença entre médias (amostras não pareadas)

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

$H_0$  : A média de notas de casados e solteiros é igual ou  $H_0 : \mu_c - \mu_s = 0$  ou  $H_0 : \mu_c = \mu_s$

$H_1$  : A média de notas de casados e solteiros é diferente ou  $H_1 : \mu_c - \mu_s \neq 0$  ou  $H_1 : \mu_c \neq \mu_s$

Variável **dependente**: Notas

Variável **independente**: Situação conjugal

O que eu quero testar? Se a situação conjugal *faz diferença* na nota.

É **efeito**? Não! (Pearl, 2020) Inferência vs. Causalidade

# Como testar na prática? Distribuições:

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

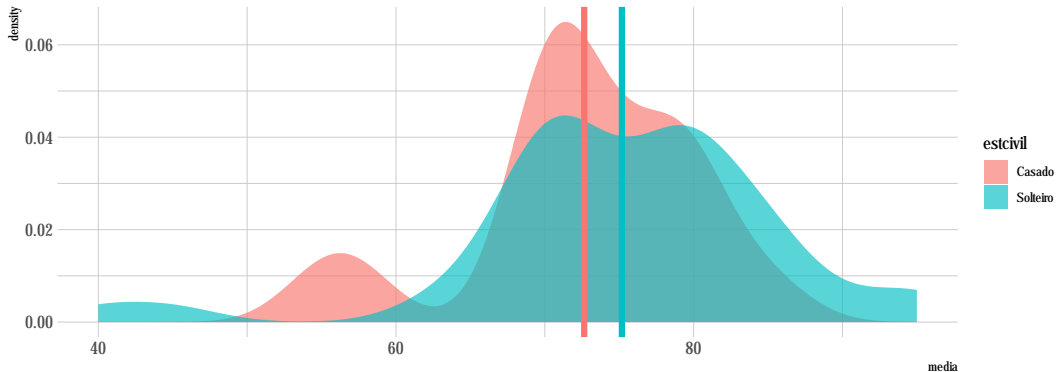
Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
dfe %>% drop_na(estcivil) %>%  
  ggplot(aes(fill=estcivil,x=media,color=estcivil,group=estcivil)) +  
  geom_density(color=NA,alpha=.65) +  
  geom_vline(data=. %>% group_by(estcivil) %>% summarise(media=mean(media,na.rm = T)),  
             size=2,aes(xintercept=media,color=estcivil)) +  
  guides(color="none") + theme_ipsum_mod
```



# Como testar na prática? Vamos construir intervalos

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

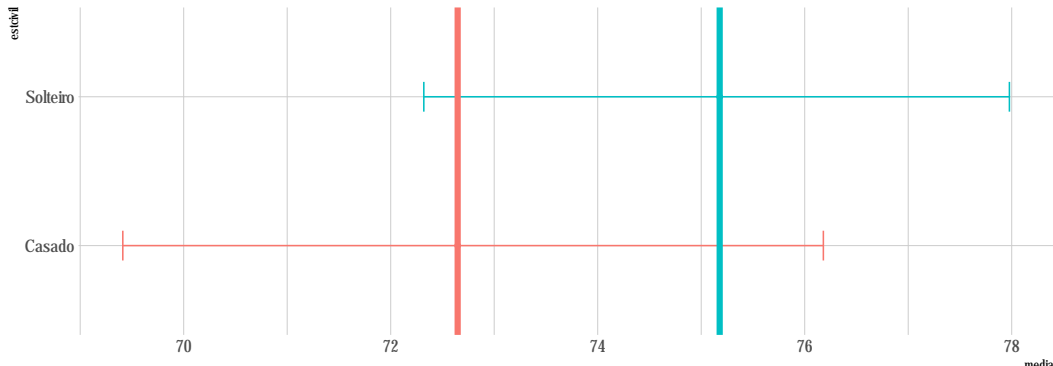
Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
dfe %>% drop_na(estcivil) %>%  
  ggplot(aes(fill=estcivil,x=media,color=estcivil,y=estcivil)) +  
  stat_summary(fun=mean, geom="point") +  
  stat_summary(fun.data=mean_cl_boot, geom="errorbar", width=0.2) +  
  geom_vline(data=., %>% group_by(estcivil) %>% summarise(media=mean(media,na.rm = T)),  
            size=2,aes(xintercept=media,color=estcivil)) +  
  theme_ipsum_mod +theme(legend.position = "none")
```



# Como testar na prática? Teste-t

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
t_test_results <- dfe %>%  
  t_test(formula = media ~ estcivil,  
         order = c("Casado", "Solteiro"))  
t_test_results
```

```
# A tibble: 1 x 6  
  statistic t_df p_value alternative lower_ci upper_ci  
    <dbl> <dbl>   <dbl> <chr>          <dbl>    <dbl>  
1    -1.03  40.4    0.311 two.sided    -7.52     2.46
```

# Outro exercício

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

$$H_0 : \mu_{3joan} = \mu_{3joad}$$

$$H_1 : \mu_{3joan} \neq \mu_{3joad}$$

Variável **dependente**: Notas

Variável **independente**: Turma (apenas 3joan e 3joad)



# Intervalos

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

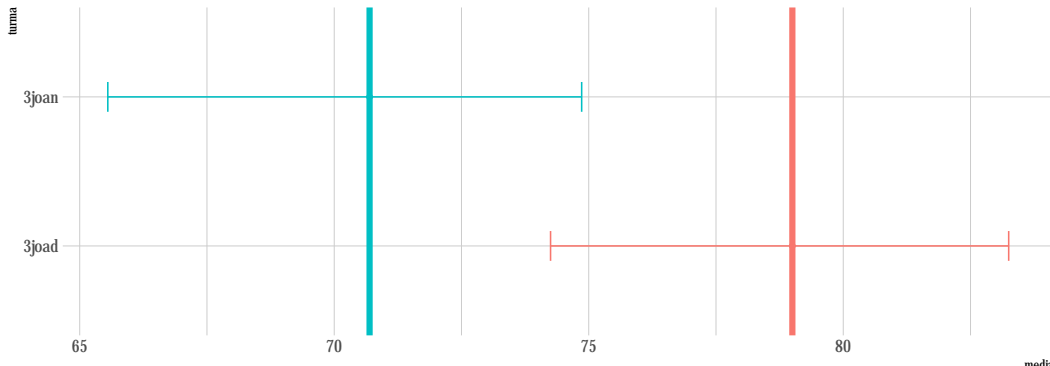
Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
dfe %>% dplyr::filter(turma %in% c("3joan","3joad")) %>%  
  ggplot(aes(fill=turma,x=media,color=turma,y=turma)) +  
  stat_summary(fun=mean, geom="point") +  
  stat_summary(fun.data=mean_cl_boot, geom="errorbar", width=0.2) +  
  geom_vline(data=. %>% group_by(turma) %>% summarise(media=mean(media,na.rm = T)),  
            size=2,aes(xintercept=media,color=turma)) +  
  theme_ipsum_mod +theme(legend.position = "none")
```



# Como testar na prática? Teste-t

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
t_test_results <- dfe %>%  
  dplyr::filter(turma %in% c("3joan", "3joad")) %>%  
  t_test(formula = media ~ turma)  
t_test_results
```

```
# A tibble: 1 x 6
```

	statistic	t_df	p_value	alternative	lower_ci	upper_ci
	<dbl>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>
1	2.37	36.0	0.0232	two.sided	1.20	15.4

# Olhando no infer graficamente

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
# calculate the observed statistic
media_turmas <- dfe %>%
  dplyr::filter(turma %in% c("3joan", "3joad")) %>%
  specify(media ~ turma) %>%
  calculate(stat = "t", order = c("3joan", "3joad"))

# generate the null distribution with the theoretical t
distribuicao_teorica <- dfe %>%
  dplyr::filter(turma %in% c("3joan", "3joad")) %>%
  specify(media ~ turma) %>%
  hypothesize(null = "independence") %>%
  calculate(stat = "t", order = c("3joan", "3joad"))
```

# Visualizando

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

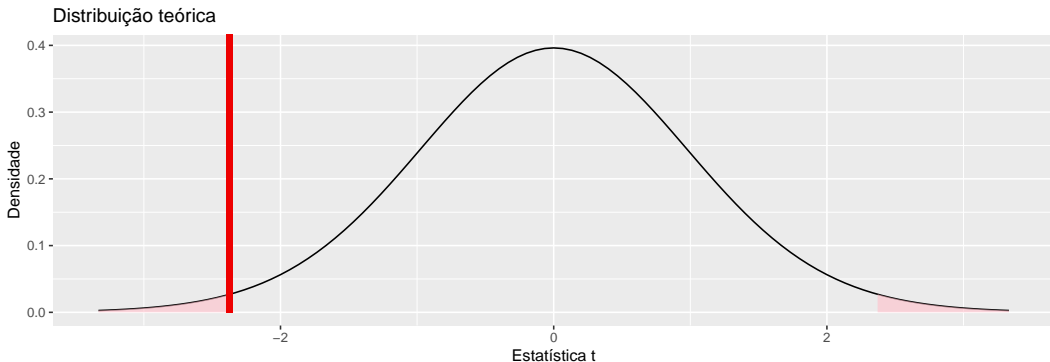
Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
# visualize the randomization-based null distribution and test statistic!  
distribuicao_teorica %>%  
  visualize(method = "theoretical") +  
  shade_p_value(media_turmas, direction = "two-sided") +  
  labs(title = "Distribuição teórica", x = "Estatística t", y = "Densidade")
```



# Usando ggpubr

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

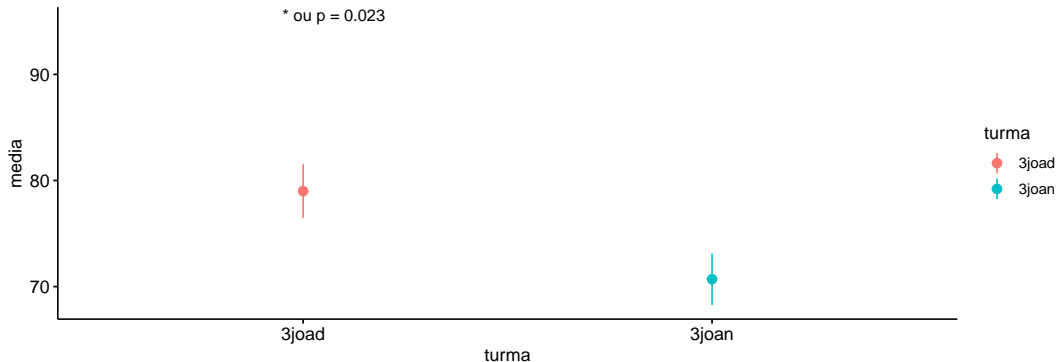
Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

```
library(ggpubr)
dfe %>% dplyr::filter(turma %in% c("3joad", "3joan")) %>%
  ggerrorplot(x = "turma", y = "media", color = "turma",
              position = position_dodge(0.5)) +
  stat_compare_means(aes(label = paste0(..p.signif.., " ou p = ", ..p.format..)),
                    method = "t.test") + theme(legend.position = "right")
```



# Tamanho do efeito

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

**Análise de  
variância e  
correlação**

**Frederico  
Bertholini**

Formalizando  
Hipóteses

**Testando  
diferentes  
hipóteses**

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

# Testando diferentes hipóteses



# Por que ANOVA?

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k, \quad H_A : \mu_i \neq \mu_j \text{ para pelo menos um par } i \text{ e } j$$

# O que é ANOVA?

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

**Testando  
diferentes  
hipóteses**

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

# Na prática

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

$H_0$  : A média de notas das turmas é igual ou  $H_0 : \mu_{3\text{joad}} = \mu_{3\text{joan}} = \mu_{5\text{joan}}$

$H_A$  : A média de notas de pelo menos uma das turmas é diferente ou  $H_A : \mu_{3\text{joad}} \neq \mu_{3\text{joan}} \neq \mu_{5\text{joan}}$

Variável dependente: Notas

Variável independente: Situação conjugal

**Análise de  
variância e  
correlação**

**Frederico  
Bertholini**

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

**Medidas de  
associação**

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

## Medidas de associação

**Análise de  
variância e  
correlação**

**Frederico  
Bertholini**

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

**Modelo  
Linear  
Simples**

Modelos  
Lineares  
Multivaria-  
dos

# Modelo Linear Simples

# Modelo Linear Simples

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

**Modelo  
Linear  
Simples**

Modelos  
Lineares  
Multivaria-  
dos

Nesta primeira aula, veremos um dos modelos mais básicos, mas também um dos mais utilizados, de regressão: o modelo linear simples estimado por mínimos quadrados ordinários (MQO).<sup>1</sup> Entre outros, aprenderemos a visualizar relações lineares bivariadas; a usar a função básica para estimar modelos lineares, `lm`; e a obter resultados detalhados de um modelo. Também usaremos alguns dados, que estão na pasta de materiais desta aula.

---

<sup>1</sup>Para um explicação rápida de como esse método de estimação funciona, ver aqui.

# Visualizando relações entre duas variáveis

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Podemos usar gráficos para visualizar relações entre variáveis. Podemos, por exemplo, criar duas variáveis contínuas, geradas aleatoriamente com a função `rnorm`, e plotar sua relação com um gráfico de pontos.

```
set.seed(123) # Para replicar o conteúdo da aula

x <- rnorm(100)
y <- rnorm(100)

plot(y ~ x)
```

É possível ver que as variáveis que criamos, `y` e `x`, não parecem ter relação. Na verdade, os pontos no gráfico estão dispersos, o que dificulta detectar algum padrão.

Para investigarmos isso detidamente, usamos um modelo linear estimado por mínimos quadrados ordinários (MQO). A função que estima esse tipo de modelo chama-se `lm`, e ela está contida no R-base. Portanto, não precisamos carregar nenhum pacote para usá-la, basta executar o seguinte código:

```
lm(y ~ x)
```

A utilização da função `lm` é simples. Em primeiro lugar, precisamos passar como argumento da função uma fórmula contendo a especificação do nosso modelo. No exemplo, `y` é nossa variável dependente e

## Visualizando relações entre variáveis em um banco

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Daqui até o fim dessa aula, usaremos um banco de dados como exemplo para praticarmos o básico de construção de modelos bivariados. Esse arquivo, com algumas variáveis selecionadas do Atlas do Desenvolvimento Humano para municípios no ano de 2010

(<http://atlasbrasil.org.br/2013/pt/download/>), chama-se `atlas_condensado.csv`, e está nos materiais complementares dessa aula. Como ele está no formato `.csv`, podemos carregá-lo com a função `read_csv` do pacote `readr`, salvando o banco num objeto `atlas`.

```
library(readr)
atlas <- read_csv("atlas_condensado.csv")
```

A função `read_csv` já indica o nome e a classe de cada variável no objeto `atlas`, onde salvamos os dados. Usando `names` e `head`, é possível ver mais detalhadamente esse `data.frame`.

```
names(atlas)
head(atlas)
```

Em particular, vamos trabalhar com duas dessas variáveis: a `t_agua`, que indica a percentagem de domicílios com água encanada, e a `mort1`, que indica o número de crianças que não sobreviveram ao primeiro ano de idade em cada 1000 crianças nascidas no município. Nosso objetivo será investigar se existe relação entre disponibilidade de água encanada (nossa variável independente) e mortalidade infantil (nossa variável dependente) usando um modelo linear por MQO. Como essas variáveis estão dentro de um banco de dados, existe um modo mais fácil de visualizar a relação entre elas usando o



# Estimativas

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Para examinarmos formalmente a relação entre `t_agua` e `mort1`, vamos recorrer novamente à função `lm`. Como temos um `data.frame` agora, e não mais dois vetores em objetos diferentes, usaremos um argumento da função `lm` para passar as variáveis do banco para ela. O argumento chama-se `data =`, e precisamos apenas passar a ele o nome do `data.frame` que contém nossas variáveis.

```
meu_modelo <- lm(mort1 ~ t_agua, data = atlas)
```

Note que podemos estimar o mesmo modelo sem usar o argumento `data`, mas precisaremos usar em seu lugar indexadores (o que é mais trabalhoso e não recomendado):

```
lm(atlas$mort1 ~ atlas$t_agua)
```

Para exibir de forma condensada os resultados do nosso modelo salvo em `meu_modelo`, basta digitar o nome do objeto no console:

```
meu_modelo
```

Os dois modos produzem o mesmo resultado, ainda que o segundo seja o mais adequado. Mas, afinal, o que significa esse output da função `lm`. Podemos resumir-lo em duas partes:

- **1** Abaixo de `Call:`, a função `lm` apenas exibe a chamada que usamos para rodar nosso modelo, incluindo aqui a fórmula usada;

# Inferência

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simple

Modelos  
Lineares  
Multivaria-  
dos

No mais das vezes, apenas avaliar um modelo pelos seus coeficientes não é suficiente. Usando nosso modelo estimado, podemos facilmente dar um passo adiante e fazer inferências a partir deles, considerando a incerteza na estimação dos nossos parâmetros.

Para tanto, o R dispõe de uma função, que é carregada por padrão a partir do pacote `stats`, chamada `confint`. Com ela, é possível calcular um intervalo de confiança de 95% (ou qualquer outro valor) para os coeficientes de nosso modelo.

```
confint(meu_modelo, level = 0.95)
```

A função retorna os intervalos mínimo e máximo de cada estimativa, o que pode ser usado para fazer afirmações do tipo de que, com 95% de probabilidade, nossa estimativa do efeito de `t_gua` sobre `mort1` está situada no intervalo indicado por 2.5 % e 97.5 %.

Para alterar a cobertura do intervalo de confiança, basta modificar o valor passado ao argumento `level` da função (0.95 equivale a um intervalo de 95%):

```
confint(meu_modelo, level = 0.90)
```

## Obtendo resultados detalhados

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Ainda que exibir os resultados da forma como fizemos seja útil para uma inspeção rápida, existem outras informações úteis que precisamos acessar para avaliarmos um modelo. Para obtê-las, usamos a função `summary`, que já vimos anteriormente.

```
summary(meu_modelo)
```

Afora a linha iniciada por `Call:`, que já vimos, essa função retorna uma série de novas informações. Vamos focar especialmente em duas:

- **1 Coefficients:** reporta informações sobre as estimativas e inferências de nosso modelo, incluindo aqui a estimativa do efeito de nossas variáveis independentes e intercepto (abaixo de `Estimate`); o erro-padrão (`Std. Error`) de cada estimativa, que indica a incerteza envolvida nelas (e é usada no cálculo de intervalos de confiança); T-valor, `t value`, uma métrica normalmente usada para calcular P-valores em modelos de regressão (i.e., que é igual à estimativa dividido pelo erro-padrão); e, finalmente, P-valor, `Pr(>|t|)`, a famosa estatística usada para testar se o efeito indicado por cada estimativa é ( $P\text{-valor} > 0.05$ ) ou não ( $P\text{-valor} < 0.05$ ) fruto de variação aleatória.<sup>2</sup>;
- **2 Multiple R-squared:** e **Adjusted R-squared:** são métricas simples que indicam a quantia de variação explicada pelo nosso modelo, onde 1 indica que toda a variação é explicada (e, logo, nossas variáveis predizem perfeitamente os valores de  $Y$ );

---

<sup>2</sup>A título de curiosidade, o P-valor é calculado com base na estatística T de cada variável e outra informação que cobriremos a seguir, os graus de liberdade do modelo.

## Acessando informações de um objeto lm

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Assim como outras informações no R, nosso modelo salvo no objeto `meu_modelo` contém diversas informações dentro dele. Por exemplo, podemos extrair um vetor com nossas estimativas usando:

```
meu_modelo$coefficients
```

Essa e outras informações salvas dentro do objeto podem ser vistas com `names`:

```
names(meu_modelo)
```

Também é possível salvar o *output* da função `summary` para fazer a mesma coisa.

```
resumo <- summary(meu_modelo)
```

Feito isso, temos à disposição todas as estatísticas reportadas pela função `summary` dentro de um mesmo objeto, o que pode ser usado, por exemplo, para extrair rapidamente apenas o R-quadrado de nosso modelo, como abaixo.

```
resumo$r.squared
```

Isso será bastante útil em algumas situações que veremos adiante. Por enquanto, tente explorar as informações contidas dentro dos objetos `resumo` e `meu_modelo` e tente identificar o que indicam.

# Exercícios II

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

- 1) A base atlas, que carregamos anteriormente, contém outras duas variáveis que podem ter relação com a taxa de mortalidade infantil até um 1 ano por 1000 crianças nascidas. São elas: `t_lixo`, que indica a % de domicílios no município com coleta de lixo; e `t_luz`, que indica a % de domicílios com energia elétrica. Qual sua expectativa do efeito dessas variáveis sobre `mort1`?
- 2) Faça dois gráficos de *scatterplot*, um para investigar a relação entre `mort1` e `t_luz` e, o outro, para investigar a relação entre `mort1` e `t_lixo`;
- 3) Estime e salve dois modelos lineares simples para examinar a relação entre `mort1` e `t_luz` e entre `mort1` e `t_lixo`. Lembre-se: `mort1` é nossa variável dependente;
- 4) Usando `summary`, interprete os resultados desses dois modelos. Os resultados condizem com a sua expectativa prévia?
- 5) Com os mesmos modelos estimados anteriormente, faça inferência a partir do efeito de `t_luz` e `t_lixo` sobre `mort1`. Quais são os intervalos máximos e mínimos, considerando 95% de probabilidade, dos efeitos dessas duas variáveis?

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

**Modelos  
Lineares  
Multivaria-  
dos**

# Modelos Lineares Multivariados

# Modelos Lineares Multivariados

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Embora examinar a relação entre duas variáveis, como vimos na Aula @ref(aula1), seja útil em algumas ocasiões, modelos lineares também servem para examinar o efeito de múltiplas variáveis independentes simultaneamente. Nessa aula, vamos ver como estimar modelos como estes. Após isso, veremos como rodar e interpretar de forma simples modelos lineares multivariados e, também, modelos com transformações de variáveis independentes – o que é útil em alguns casos, como quando examinamos relações quadráticas ou quando desejamos interpretar coeficientes de forma percentual.

# Modelo linear com dois preditores

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

No exemplo da Aula @ref(aula1), examinamos o efeito da variável X sobre Y em todos os municípios brasileiros. Como vimos, essa relação é positiva e, além disso, é possível inferir isso a partir dos dados, dado os intervalos de confiança da estimativa do efeito predito de X sobre Y.

Na mesma base de dados, também temos outra variável que possivelmente tem efeito sobre Y. Essa variável chama-se X, e indica quanto. . .

Vamos carregar nosso banco de dados usando o pacote readr.

```
library(readr)
x <- read_delim("afe.txt", delim = ";")
```

Para termos uma ideia sobre a distribuição da variável X, podemos usar o indexador \$ para acessá-la conjuntamente com a função summary para gerar algumas estatísticas descritivas.

```
summary(x$variavel)
```

Sempre é uma boa ideia examinar graficamente a relação entre variáveis e, nesse caso, podemos fazer isso usando o pacote ggplot2 para plotar a relação entre X e Y.

```
library(ggplot2)

ggplot(banco, aes(x = x, y = y)) +
```



# Modelos com múltiplos preditores

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Usando a mesma sintaxe que vimos acima, podemos facilmente incluir mais variáveis a um modelo linear. Vamos incluir uma variável adicional ao nosso modelo, chamada X, que indica...

```
meu_modelo2 <- lm(y ~ x + x2, data = banco)
```

Aqui também, interpretar os resultados é algo simples.

# Obtendo resultados simplificados

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

Algo que dificulta interpretar modelos multivariados de forma rápida é a quantidade de informações reportada pela função `summary`. Felizmente, existe uma alternativa a ela em um pacote chamado `arm`: a função `display`. Em vez de reportar uma série de estatísticas e informações nem sempre úteis, essa função simplifica os resultados de uma regressão, reportando apenas estimativas, erros-padrão e informações como número de observações e parâmetros estimados. Como veremos adiante, `arm` também oferece algumas funções que simplificam enormemente nossa tarefa de estimar e analisar modelos de regressão.

Como esse pacote não vem por padrão no R, precisamos instalá-lo usando `install.packages`.

```
install.packages(arm)
```

Feito isso, podemos carregar o pacote e usá-lo para reportar os resultados do nosso modelo linear multivariado salvo no objeto `modelo_multi`.

```
library(arm)
```

```
display(modelo_multi)
```

# Interpretando resultados com gráficos

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

arm ainda oferece um jeito mais fácil de visualizar resultados de modelos multivariados. Usando a função `coefplot`, ele plota as estimativas de um modelo, já incluindo intervalos de confiança de 95% e 90%. Para usá-la, basta passar o objeto com o modelo salvo para a função.

```
coefplot(modelo_multi)
```

Além de simples, o resultado gráfico

# Centrando preditores

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

# Tranformando preditores

**Análise de  
variância e  
correlação**

**Frederico  
Bertholini**

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

**Modelos  
Lineares  
Multivaria-  
dos**

# Relações quadráticas e outras

Análise de  
variância e  
correlação

Frederico  
Bertholini

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

Modelos  
Lineares  
Multivaria-  
dos

# Exercícios

**Análise de  
variância e  
correlação**

**Frederico  
Bertholini**

Formalizando  
Hipóteses

Testando  
diferentes  
hipóteses

Medidas de  
associação

Modelo  
Linear  
Simples

**Modelos  
Lineares  
Multivaria-  
dos**