



Prof^a.

Dra. Edneide Ramalho

Monitora:

Lorena Caana

Estatística Paramétrica e Não-Paramétrica com uso do software R



Teste t



- O **Teste t** é usado para comparar duas médias.
- Tipos de **Teste t**:
 - Teste t para amostras **independentes**
 - Teste t para amostras **pareadas**



Teste t para amostras independentes



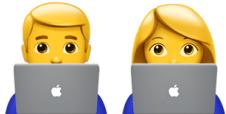
Teste t para amostras independentes

- É usado para comparar as **médias de dois grupos independentes**.
- Por exemplo, podemos comparar o peso médio de indivíduos agrupados por gênero.



Suposições do teste

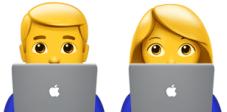
- 1. Independência das observações**
- 2. Sem outliers extremos**
- 3. Normalidade**
- 4. Homogeneidade de variâncias**



Preparando o ambiente

- ❑ Criar uma pasta em nosso computador e um projeto no Rstudio dentro desta pasta.

- ❑ Criar o primeiro script e salvar.



Preparando o ambiente

- ❑ Vamos precisar dos seguintes pacotes:

```
library(tidyverse)
```

```
library(ggpubr)
```

```
library(rstatix)
```



Base de dados

■ Carregando e explorando a base de dados:

```
library(readxl)
```

```
genderweight <- read_excel("genderweight.xlsx")
```

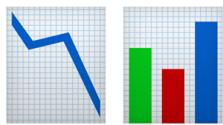
```
glimpse(genderweight)
```

```
genderweight %>%
  sample_n(5)
```



Base de dados

```
table(genderweight$group)
```



Estatísticas de resumo

```
genderweight %>%
  group_by(group) %>%
  get_summary_stats(weight, type = "mean_sd")
```



Visualização

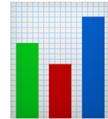
```
bxp_peso <- genderweight %>%
  ggboxplot(x = "group", y = "weight")
bxp_peso
```



Visualização

Com barra de erros:

```
bxp_peso <- genderweight %>%
  ggboxplot(x = "group", y = "weight",
  bxp.errorbar = TRUE)
bxp_peso
```



Visualização

■ Com jitter:

```
bxp_peso <- genderweight %>%
  ggboxplot(x = "group", y = "weight",
add = "jitter")
bxp_peso
```



Checando as suposições do teste

1. Independência das observações

- ❑ Não há dependência na aferição do peso entre os indivíduos, ou seja, a medida de peso de um indivíduo não interfere na medida do outro.
- ❑ As medidas para pesos dos homens são independentes das medidas dos pesos das mulheres.



Checando as suposições do teste

2. Identificando outliers extremos

```
genderweight %>%
  group_by(group) %>%
  identify_outliers(weight)
```



Checando as suposições do teste

3. Checando normalidade

▢ Teste de Shapiro-Wilk

```
genderweight %>%
  group_by(group) %>%
  shapiro_test(weight)
```



Checando as suposições do teste

3. Checando normalidade

- ❑ QQ plot

```
genderweight %>%  
  ggqqplot(x = "weight", facet.by  
= "group")
```

- ❑ Tanto o qq plot quanto os teste de Shapiro-Wilk mostram que os dados são normalmente distribuídos.



Checando as suposições do teste

4. Igualdade de variâncias

- ❑ Podemos usar o teste de **Levene**.
- ❑ As variâncias são iguais se o p-valor > 0.05.

```
genderweight %>%  
  levene_test(weight ~ group)
```



Computando o teste

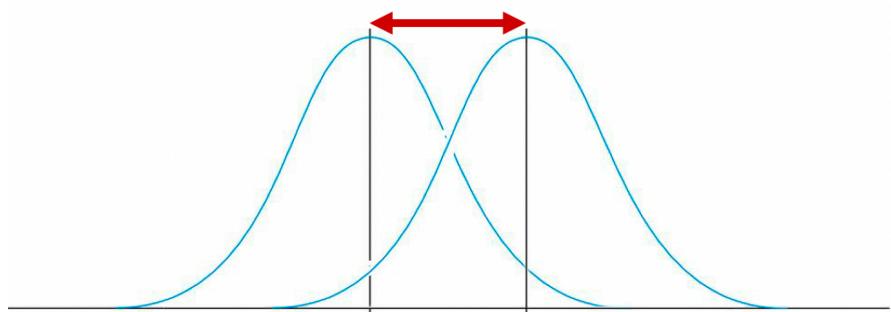
```
teste_peso_genero <- genderweight %>%
  t_test(weight ~ group, var.equal =
FALSE) %>%
  add_significance()
teste_peso_gênero
```



Calculando o tamanho do efeito

- O tamanho do efeito mede a “força” da relação entre duas variáveis, ou quantifica o tamanho da diferença entre dois grupos.
- O tamanho do efeito para o teste t é chamado de d de Cohen, e é calculado como:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{DP_{\text{agrupado}}}$$



d de Cohen	Efeito
0.2 – 0.5	Pequeno
0.5 – 0.8	Médio
≥ 0.8	grande



Calculando o tamanho do efeito

```
genderweight %>%  
  cohens_d(weight ~ group, var.equal =  
FALSE)
```

- ▣ O tamanho do efeito é grande, $d = 6.57$.



Calculando o tamanho do efeito

```
genderweight %>%  
  cohens_d(weight ~ group, var.equal =  
FALSE)
```

- ▣ O tamanho do efeito é grande, $d = 6.57$.



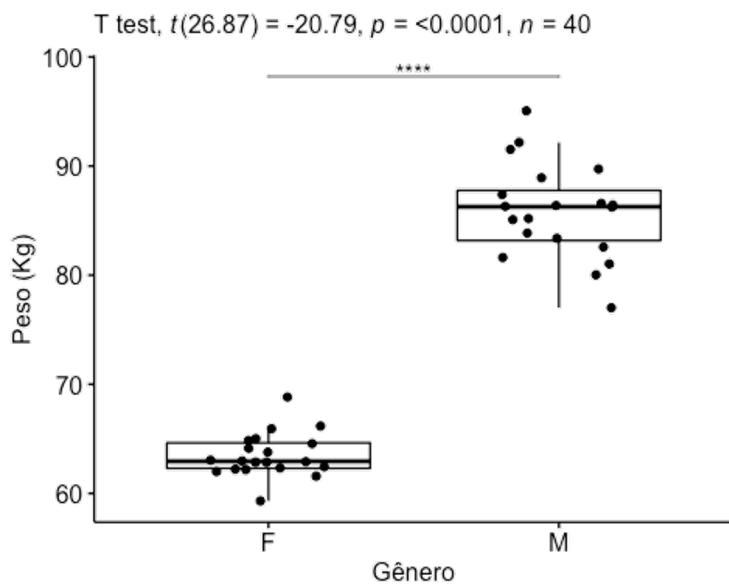
Perguntas/Dúvidas





Reportando o resultado

- A média de peso para as mulheres foi de 63,5 kg ($dp = 2,03$ kg), enquanto que para os homens, a média de peso foi de 85,8 kg ($dp = 4,3$ kg). O test t de Welch para duas amostras foi utilizado e mostrou que a diferença entre os pesos foi estatisticamente significativa, $t(26.9) = -20.8$, $p < 0.0001$, $d = 6.57$; em que $t(26.9)$ denota a estatística de teste que teve 26.9 graus de liberdade.





Reportando o resultado

```
teste_peso_genero <- teste_peso_genero %>%
  add_xy_position(x = "group")

bxp_peso +
  stat_pvalue_manual(teste_peso_genero,
tip.length = 0) +
  labs(subtitle =
get_test_label(teste_peso_genero, detailed
= TRUE),
    x = "Gênero", y = "Peso (Kg)")
```



Teste t para amostras pareadas



Teste t pareado

- ❑ O teste pareado compara duas médias de dois grupos relacionados.
- ❑ Por exemplo, comparar a média de peso de 20 ratos antes e depois de um tratamento.
- ❑ Comparar o desempenho de alunos antes de depois de um treinamento.
- ❑ Comparar o tempo médio de processamento de máquinas antes e depois de um reparo.



Suposições do teste

- **Os dois grupos são pareados:** medimos duas vezes para o mesmo sujeito experimental.
 - **Sem outliers extremos:** na diferença entre os grupos.
 - **Normalidade:** a diferença entre os pares seguem a distribuição normal.



Preparando o ambiente

■ Carregando a base de dados

```
library(readxl)
mice2 <- read_excel("mice2.xlsx")
glimpse(mice2)
```



Preparando o ambiente

- Transformando os dados para o formato longo

```
mice2.long <- mice2 %>%
  gather(key = "group", value =
  "weight", before, after)
```

```
head(mice2.long)
```



Estatísticas de resumo

```
mice2.long %>%
  group_by(group) %>%
  get_summary_stats(weight, type =
  "mean_sd")
```



Visualização

```
bxp_pareado <- ggpaired(mice2.long,
  x = "group",
  y = "weight",
  order = c("before", "after"),
  ylab = "Peso (kg)",
  xlab = "Grupos")
bxp_pareado
```



Suposições do teste

- ❑ Vamos acrescentar uma coluna chamada **diferencias** na base de dados original, porque é nesta variável que vamos checar as suposições:

```
mice2 <- mice2 %>%
  dplyr::mutate(diferenca = before
- after)

head(mice2)
```



Identificando outliers

```
mice2 %>%  
  identify_outliers(diferenca)
```



Checando normalidade

```
mice2 %>%  
  shapiro_test(diferenca)
```



Computando o teste

```
teste_pareado <- mice2.long %>%
  t_test(weight ~ group, paired =
TRUE) %>%
  add_significance()
View(teste_pareado)
```



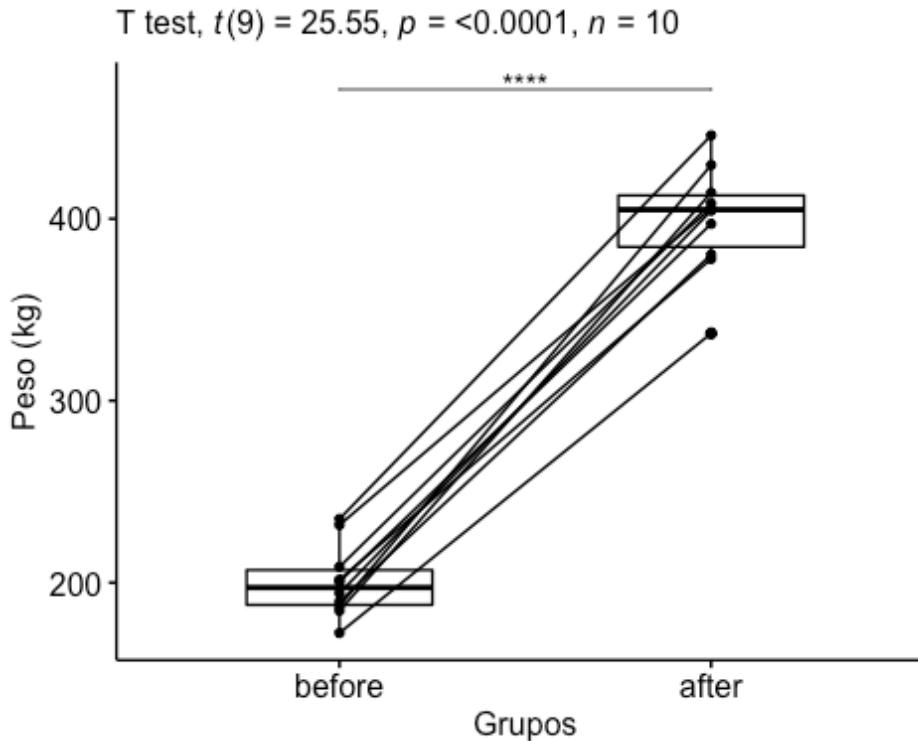
Tamanho do efeito

```
tmn_efeito_pareado <- mice2.long %>%
  cohens_d(weight ~ group, paired =
TRUE )
```



Reportando o resultado

- O peso médio dos ratos teve um crescimento estatisticamente significativo após o tratamento, $t(9) = 25.5$, $p < 0.0001$, $d = 8.07$.





Reportando o resultado

```
teste_pareado <- teste_pareado %>%
  add_xy_position(x = "group")

bxp_pareado +
  stat_pvalue_manual(teste_pareado,
tip.length = 0) +
  labs(subtitle =
get_test_label(teste_pareado, detailed =
TRUE))
```



Resumindo

As diferentes maneiras de realizar um teste t são:

- **Teste t para duas amostras independentes**

```
genderweight %>%
```

```
  t_test(weight ~ group, var.equal = FALSE)
```

- **Teste t para amostras pareadas**

```
mice2.long %>%
```

```
  t_test(weight ~ group, paired = TRUE)
```

- Quando as suposições para a realização do teste t são violadas, testes não paramétricos são recomendados, como o.



Perguntas/Dúvidas

