

# Testes de hipóteses

## Curso de Bioestatística com R

---

Ornella Scardua Ferreira


✉ [ornscar@gmail.com](mailto:ornscar@gmail.com)  [@ornscar](#)  [@ornscar](#)

## Sobre mim



*Amo gráficos como amo cavalos. Gosto de música ruim e de cinema (bom). Sou apaixonada pelo Botafogo e pelo Bayern de Munique. Prefiro Vila Velha a qualquer lugar no mundo. Não tenho sonhos, mas um dia espero ver a Palestina livre.*

# Cronograma

- Aplicação dos testes no :
  - Atividade 1: comparação de médias entre dois grupos (Teste T/Teste de Mann-Whitney);
  - Atividade 2: comparação de proporções entre dois grupos (Teste Qui-quadrado/Teste Exato de Fisher);
  - Atividade 3: comparação de médias entre dois grupos pareados (Teste T pareado/Teste de Wilcoxon);
  - Atividade 4: comparação de médias entre mais de dois grupos (ANOVA/Teste de Kruskal-Wallis).
- Testes com `{gtsummary}`.

# Atividade 1

*Objetivo: relacionar IMC (numérico) com diabetes.*

**Base de dados: diabetes.csv.**

- 1º passo: carregar os dados no .

```
# carregando os dados
dados1 <- readr::read_csv("dados/diabetes.csv", show_col_types = FALSE)
```

- 2º passo: verificar normalidade dos dados, por exemplo, por meio do Teste de Shapiro-Wilk.

```
# teste de shapiro-wilk  
shapiro.test(dados1$BMI)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  dados1$BMI  
## W = 0.88331, p-value < 2.2e-16
```

## Conclusão!

O Teste de Shapiro-Wilk testa a hipótese nula ( $H_0$ ) que afirma que os dados seguem distribuição normal. Logo, dado o valor-p menor do que 0,05, há evidências suficientes para rejeitar  $H_0$  e concluir que a variável IMC não é normalmente distribuída.

- 3º passo: aplicar **Teste T** ou **Teste de Mann-Whitney** para comparar média entre os grupos.

### Teste T (para caso paramétrico)

```
# teste t
t.test(BMI ~ Diabetes_012, data = dados1, alternative = "two.sided", conf.level = 0.95)

##
##      Welch Two Sample t-test
##
## data:  BMI by Diabetes_012
## t = 6.0535, df = 184, p-value = 7.782e-09
## alternative hypothesis: true difference in means between group pre-diabetes ou diabetes
## 95 percent confidence interval:
##  2.907587 5.719199
## sample estimates:
## mean in group pre-diabetes ou diabetes      mean in group sem diabetes
##                31.65625                    27.34286
```

### Atenção!

O Teste T exige a suposição de uma distribuição de probabilidade para a variável testada (geralmente, normal). Por isso, uma vez que o Teste de Shapiro-Wilk apontou para a não normalidade da variável IMC na etapa anterior, a aplicação do Teste T, neste caso, torna-se inverossímil.

## Teste de Mann-Whitney (para caso não paramétrico)

```
# teste de mann-whitney
wilcox.test(BMI ~ Diabetes_012, data = dados1, alternative = "two.sided", conf.level = 0.05)

##
##      Wilcoxon rank sum test with continuity correction
##
## data:  BMI by Diabetes_012
## W = 91526, p-value = 3.379e-13
## alternative hypothesis: true location shift is not equal to 0
```

### Conclusão!

O Teste de Mann-Whitney é utilizado para testar se duas amostras independentes foram retiradas de populações com médias iguais ( $H_0$ ), caso os pressupostos exigidos pelo Teste T forem violados. Portanto, ao nível de significância de 5%, há evidências suficientes para rejeitar  $H_0$  e concluir que o IMC médio entre os grupos com e sem diabetes é diferente.

# Atividade 2

*Objetivo: relacionar nível de colesterol (categórico) com diabetes.*

**Base de dados: diabetes.csv.**

- 1º passo: carregar os dados no .

```
# carregando os dados
dados1 <- readr::read_csv("dados/diabetes.csv", show_col_types = FALSE)
```

- 2º passo: criar tabela de contingência 2x2.

```
# tabela 2x2
t22 <- table(dados1$HighChol, dados1$Diabetes_012); t22
```

```
##
##               pre-diabetes ou diabetes sem diabetes
## High cholesterol                99                321
## No high cholesterol             61                519
```



- 3º passo: aplicar o **Teste Qui-quadrado** ou o **Teste Exato de Fisher**.

### Teste Qui-quadrado

```
# teste qui-quadrado  
chisq.test(t22)
```

```
##  
##      Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  t22  
## X-squared = 29.923, df = 1, p-value = 4.494e-08
```

### Conclusão!

O Teste Qui-quadrado testa a hipótese ( $H_0$ ) que afirma que não existe associação entre duas certas variáveis (são independentes) contra a hipótese ( $H_1$ ) que afirma que existe associação entre essas duas variáveis (são dependentes). Portanto, ao nível de significância de 5%, há evidências suficientes para rejeitar  $H_0$  e concluir que colesterol está associado ao diagnóstico de diabetes.

## Teste Exato de Fisher (para casos em que há caselas com frequências abaixo de 5)

```
# teste exato de fisher  
fisher.test(t22, alternative = "two.sided", conf.level = 0.95)
```

```
##  
##      Fisher's Exact Test for Count Data  
##  
## data:  t22  
## p-value = 4.326e-08  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
##  1.827891 3.783037  
## sample estimates:  
## odds ratio  
##  2.621399
```

### Atenção!

Quando a frequência esperada de alguma das caselas não é maior do que 5, o Teste Exato de Fisher é utilizado como alternativa. A interpretação é análoga ao Teste Qui-quadrado.

# Atividade 3

*Objetivo: comparar IMC pré e pós-tratamento para anorexia em meninas.*

**Base de dados: anorexia.csv.**

- 1º passo: carregar os dados no .

```
# carregando os dados
dados2 <- readr::read_csv("dados/anorexia.csv", show_col_types = FALSE)
```

- 2º passo: verificar normalidade dos dados, por exemplo, por meio do Teste de Shapiro-Wilk.

```
# teste de shapiro-wilk  
shapiro.test(dados1$BMI)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  dados1$BMI  
## W = 0.88331, p-value < 2.2e-16
```

## Conclusão!

Dado o valor-p menor do que 0,05, há evidências suficientes para rejeitar  $H_0$  e concluir que a variável IMC não é normalmente distribuída.

- 3º passo: aplicar **Teste T pareado** ou **Teste de Wilcoxon**.

### Teste T (para caso paramétrico)

```
# teste t pareado
t.test(dados2$imc[dados2$momento == "pre tratamento"], dados2$imc[dados2$momento == "
      paired = TRUE, alternative = "two.sided", conf.level = 0.95)

##
##      Paired t-test
##
## data:  dados2$imc[dados2$momento == "pre tratamento"] and dados2$imc[dados2$momento == "
## t = -4.1849, df = 16, p-value = 0.0007003
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -4.964412 -1.625984
## sample estimates:
## mean difference
##      -3.295198
```

### Atenção!

Assim como o Teste T, o Teste T pareado também exige a suposição de uma distribuição de probabilidade para a variável testada. Por isso, uma vez que o Teste de Shapiro-Wilk apontou para a não normalidade da variável IMC na etapa anterior, a aplicação do Teste T pareado, neste caso, torna-se inverossímil.

## Teste de Wilcoxon (para caso não paramétrico)

```
# teste de wilcoxon
wilcox.test(dados2$imc[dados2$momento == "pre tratamento"], dados2$imc[dados2$momento
           paired = TRUE, alternative = "two.sided", conf.level = 0.95)

##
##      Wilcoxon signed rank exact test
##
## data:  dados2$imc[dados2$momento == "pre tratamento"] and dados2$imc[dados2$momento
## V = 11, p-value = 0.0008392
## alternative hypothesis: true location shift is not equal to 0
```

### Conclusão!

O Teste de Wilcoxon é uma alternativa ao Teste T para amostras pareadas quando os dados não atendem às exigências deste último. Em palavras, nesta atividade estão sendo testadas as hipóteses:  $H_0$ - há diferença média entre os IMCs pré e pós-tratamento e  $H_1$ - não há diferença média entre os IMCs pré e pós-tratamento. Logo, uma vez que o valor-p encontrado é menor do que 0,05, há fortes evidências em rejeitar  $H_0$  e concluir que o tratamento contra anorexia é efetivo em meninas.

### Atenção!

Como a variável que descreve o momento do tratamento está em uma única variável, para que o teste pareado performe de maneira correta é necessário que o primeiro indivíduo do primeiro tratamento seja o primeiro indivíduo do segundo tratamento, que o segundo indivíduo do primeiro tratamento seja o segundo indivíduo do segundo tratamento e assim por diante.

# Atividade 4

*Objetivo: relacionar IMC com autoavaliação da saúde geral.*

**Base de dados: diabetes.csv.**

- 1º passo: carregar os dados no .

```
# carregando os dados
dados1 <- readr::read_csv("dados/diabetes.csv", show_col_types = FALSE)
```

- 2º passo: verificar normalidade dos dados por meio do Teste de Shapiro-Wilk.

```
# teste de shapiro-wilk  
shapiro.test(dados1$BMI)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  dados1$BMI  
## W = 0.88331, p-value < 2.2e-16
```

## Conclusão!

Dado o valor-p menor do que 0,05, há evidências suficientes para rejeitar  $H_0$  e concluir que a variável IMC não é normalmente distribuída.



- 3º passo: aplicar **ANOVA** ou **Teste de Kruskal-Wallis**.

### ANOVA (para caso paramétrico)

```
# anova
summary(aov(BMI ~ saude_geral, data = dados1))
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## saude_geral    4   1373    343.3    8.869 4.88e-07 ***
## Residuals   995  38511     38.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Atenção!

A ANOVA exige o pressuposto de normalidade dos dados. Por isso, uma vez que o Teste de Shapiro-Wilk apontou para a não normalidade da variável IMC na etapa anterior, a aplicação da ANOVA, neste caso, não é adequada.

## Teste de Kruskal-Wallis (para caso não paramétrico)

```
# teste de kruskal-wallis
kruskal.test(BMI ~ saude_geral, data = dados1)

##
##      Kruskal-Wallis rank sum test
##
## data:  BMI by saude_geral
## Kruskal-Wallis chi-squared = 31.982, df = 4, p-value = 1.929e-06
```

### Conclusão!

O Teste de Kruskal-Wallis é adequado para testar se mais de dois grupos independentes têm médias iguais ( $H_0$ ). Assim, ao nível de significância de 5%, há evidências suficientes para rejeitar  $H_0$  e concluir que pelo menos uma das categorias de saúde geral tende a ter valor de IMC diferente das demais.

- Aplicando testes com {gtsummary}

```
library(gtsummary)

dados1 |>
  dplyr::mutate(
    Sex = forcats::fct_recode(Sex, "Fem:
    Smoker = forcats::fct_recode(Smoker,
    Diabetes_012 = forcats::fct_recode(I
  ) |>
  tbl_summary(
    by = Diabetes_012,
    include = c(Sex, Smoker, BMI),
    label = list(Sex ~ "Sexo", Smoker ~
    statistic = list(
      all_continuous() ~ "{mean}",
      all_categorical() ~ "{n} ({p}%)"
    ),
    digits = list(
      all_continuous() ~ 2,
      all_categorical() ~ c(0, 1)
    )
  ) |>
  add_p(pvalue_fun = label_style_pvalue(
  modify_header(
    label ~ "**Variável**",
    p.value ~ "**p**"
  ) |>
  modify_spanning_header(c("stat_1", "st
```

Variável	Diagnóstico de diabetes		p <sup>2</sup>
	Pré/Com N = 160 <sup>1</sup>	Sem N = 840 <sup>1</sup>	
<b>Sexo</b>			0.157
Masculino	80 (50.0%)	369 (43.9%)	
Feminino	80 (50.0%)	471 (56.1%)	
<b>Tabagista</b>			0.058
Não	79 (49.4%)	483 (57.5%)	
Sim	81 (50.6%)	357 (42.5%)	
<b>IMC</b>	31.66	27.34	<0.001
<sup>1</sup> n (%); Mean			
<sup>2</sup> Pearson's Chi-squared test; Wilcoxon rank sum test			

### Dica!

Para alterar o teste, faça, por exemplo, `add_p(test = everything() ~ "wilcox.test")`, `add_p(test = list(all_continuous() ~ "wilcox.test", all_categorical() ~ "chisq.test")` ou `add_p(test = my_custom_test)`.

# Meu obrigada!



 [ornscar@gmail.com](mailto:ornscar@gmail.com)

 [@ornscar](#)

 [@ornscar](#)