

Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento

Felipe Figueiredo

1 Comparações múltiplas

2 Análise de Variância (ANOVA)

- ANOVA um fator (One-way ANOVA)
- O teste F
- Pós teste
- Two-way ANOVA

3 Exercício

4 Encerramento

Como comparar três ou mais grupos?

- “Comparar” é um termo vago - precisamos de um critério bem definido!

Para comparar quanto às variâncias dos grupos

Podemos usar

- Teste de Levene
- Teste de Bartlett

Para comparar quanto às médias dos grupos

Pay attention

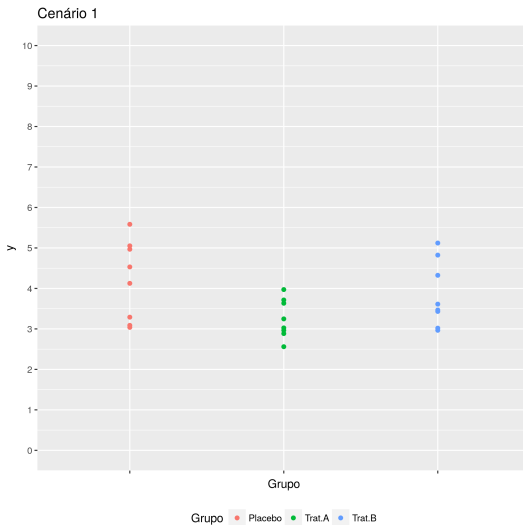
- Vimos que o **teste t** pode ser usado para comparar duas médias
- Assumindo que atendemos às premissas do teste t, precisamos levar em conta:
 - variabilidade dos grupos
 - tamanho do estudo (n)

Requisitos não óbvios (além das médias)

desvio padrão + n = erro padrão

O que é necessário para decidir se 3 (ou mais) grupos possuem médias diferentes?

Esses 3 grupos têm médias diferentes?



Médias: Placebo: 4.210, Tratamento A: 3.250, Tratamento B: 3.845



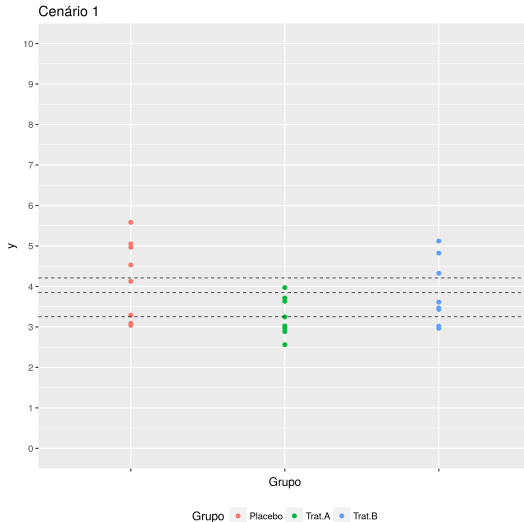
Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento



E estes 3 grupos?

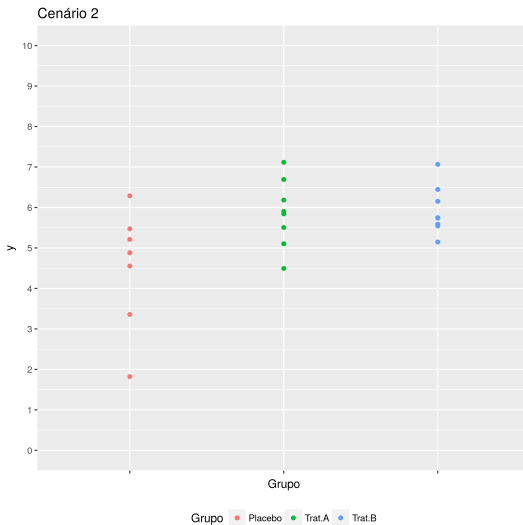
Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento



Médias: Placebo: 4.559, Tratamento A: 5.855, Tratamento B: 5.928



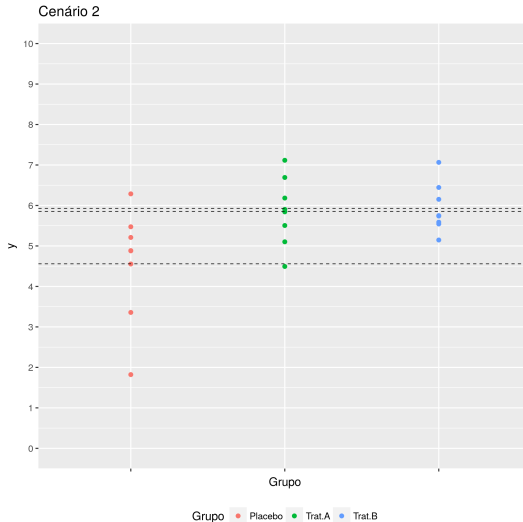
Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento



Comparação entre 3 (ou mais) grupos



Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento

Abordagem mais simples

Uma ideia seria usar o teste t três vezes, comparando os grupos aos pares.

Testar se há diferenças significativas, e seus respectivos tamanhos.

Exemplo

- 1 Placebo x Tratamento A
- 2 Placebo x Tratamento B
- 3 Tratamento A x Tratamento B

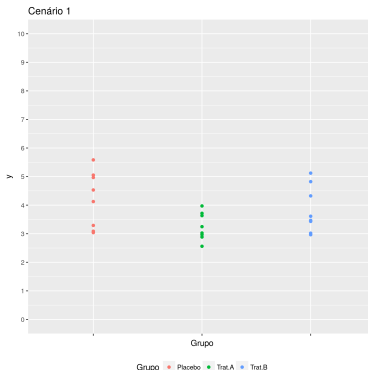
Exemplo 1

P-valores dos 3 testes t

- 1 Placebo x Trat. A $\Rightarrow p = 0.02652$
- 2 Placebo x Trat. B $\Rightarrow p = 0.4331$
- 3 Trat. A x Trat. B $\Rightarrow p = 0.09686$

Pergunta

Qual é a conclusão correta quanto à comparação destes grupos?



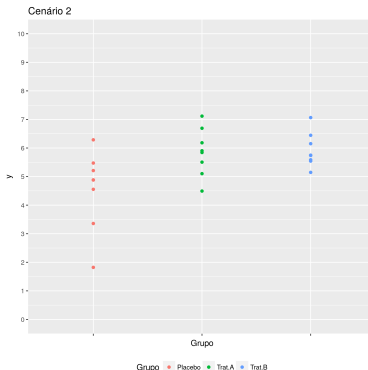
Exemplo 2

P-valores dos 3 testes t

- 1 Placebo x Trat. A $\Rightarrow p = 0.0399$
- 2 Placebo x Trat. B $\Rightarrow p = 0.02235$
- 3 Trat. A x Trat. B $\Rightarrow p = 0.8432$

Pergunta

E no segundo cenário?
Os tratamentos são diferentes
do placebo? E entre si?



Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento

Existe um problema oculto aí.

O problema é...

O problema é...

A conclusão de que no Exemplo 1 os 3 grupos são diferentes está **errada**!

¹Leia várias vezes o Cap 13!

O problema é...

A conclusão de que no Exemplo 1 os 3 grupos são diferentes está **errada!**

- O teste t permite a avaliação de **uma** hipótese
- Testamos simultaneamente várias ¹
- Isto aumenta a chance de cometermos um erro tipo I (falso positivo)
- Múltiplos testes superestimam o p-valor do método

¹Leia várias vezes o Cap 13!

- Os testes estatísticos (e fórmulas) não “sabem” o que foi levado em conta no estudo.
- *Só o pesquisador sabe*
- A metodologia da análise precisa levar em conta todo o planejamento do estudo.

Exemplo 13.2

5 crianças de uma escola tiveram leucemia, ano passado.

- Isto é uma coincidência?
- Esse agrupamento de casos sugere a presença de toxina ou efeito ambiental que causou a doença?

Qual é a probabilidade de se observar 5 casos *nesta* escola, em um ano?

- Considerando a incidência de leucemia, isto parece ser um dado extraordinário
- Esta é a pergunta errada, *após* observar os casos nesta escola.
- Se escola não é especial, é preciso considerar outras escolas
- Além disso, outras doenças (por ex., asma é um fator?).

Exemplo 13.2

5 crianças de uma escola tiveram leucemia, ano passado.

- Isto é uma coincidência?
- Esse agrupamento de casos sugere a presença de toxina ou efeito ambiental que causou a doença?

Qual é a probabilidade de se observar 5 casos *nesta* escola, em um ano?

Pergunta correta

Qual é a probabilidade de se observar 5 casos *em alguma* escola, em um ano?

E agora, José?

Como levar em conta as comparações múltiplas sem ser induzido ao erro, pelo teste t?



- Vimos que o **teste t** pode ser usado para comparar duas médias
- Assumindo que atendemos às premissas do teste t, precisamos levar em conta:
 - variabilidade dos grupos
 - tamanho do estudo (n)

Requisitos não óbvios (além das médias)

desvio padrão + n = erro padrão

1 Comparações múltiplas

2 Análise de Variância (ANOVA)

- ANOVA um fator (One-way ANOVA)
 - O teste F
 - Pós teste
 - Two-way ANOVA

3 Exercício

4 Encerramento

Exemplo 13.5

Hetland, et. al (1993) pesquisaram alterações hormonais em mulheres corredoras. Mediram o nível de hormônio luteinizante (LH) em três grupos:

- 1 sedentárias
- 2 corredoras recreacionais
- 3 corredoras de elite

Exemplo 13.5

Table 30.1. LH Levels in Three Groups of Women

| Group | $\log(\text{LH}) \pm \text{SEM}$ | N |
|----------------------|----------------------------------|----|
| Nonrunners | 0.52 ± 0.027 | 88 |
| Recreational runners | 0.38 ± 0.034 | 89 |
| Elite runners | 0.40 ± 0.049 | 28 |

- Com estas informações, podemos construir uma tabela ANOVA
- H_0 : todas as médias são iguais

Exemplo 13.5

Table 30.2. InStat Results for One-Way ANOVA

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square |
|-----------------------------|--------------------|----------------|-------------|
| Treatments (between groups) | 2 | 0.92681 | 0.4634 |
| Residuals (within groups) | 202 | 16.450 | 0.0814 |
| Total | 204 | 17.377 | |

F = 5.690

The P value is 0.0039, considered very significant.

Variation among column means is significantly greater than expected by chance.

- A razão entre as Somas dos Quadrados:
 $0.93/17.38 = 5.3\%$
- 5.3% da variabilidade pode ser explicada pelas diferenças *entre os grupos*
- (lembra do r^2 ?)

- Este método é chamado one-way (ou 1-way) ANOVA, pois tem um fator categórico
- A premissa é que pode-se *modelar* a relação entre um desfecho quantitativo e um preditor categórico + um erro aleatório
- A variável dependente do exemplo é o LH
- A (única) variável independente é o Grupo

- Quando os grupos têm médias diferentes, parte da variabilidade total é devido a esta diferença
- O resto da variabilidade é devido apenas às variâncias intra-grupos
- A ANOVA tenta *desembaraçar* esta decomposição, assumindo a hipótese nula.

- O nome *Análise de Variância* vem do critério usado para comparar as médias
- O teste de hipótese é baseado na comparação entre as variâncias intra- e inter grupos
- Estas variâncias aparecem na tabela como “Média dos Quadrados”
- Lembrete: a variância é a média dos desvios elevados ao quadrado

1 Comparações múltiplas

2 Análise de Variância (ANOVA)

- ANOVA um fator (One-way ANOVA)
- O teste F
- Pós teste
- Two-way ANOVA

3 Exercício

4 Encerramento

- Se as médias forem iguais, a variância intra-grupo deve ser “igual” à variância inter-grupo
- Calculando-se a *razão* entre a variância, esperamos que seja próximo de 1
- $$\text{razão} = F = \frac{\text{Entre grupos}}{\text{Intra grupos}}$$
- Uma razão muito maior que 1 indica que há mais variância entre os grupos do que o esperado
- Obs: o teste leva em conta os graus de liberdade do numerador e denominador

Exemplo 13.5

Table 30.2. InStat Results for One-Way ANOVA

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square |
|-----------------------------|--------------------|----------------|-------------|
| Treatments (between groups) | 2 | 0.9268 | 0.4634 |
| Residuals (within groups) | 202 | 16.450 | 0.0814 |
| Total | 204 | 17.377 | |

$F = 5.690$

The P value is 0.0039, considered very significant.
Variation among column means is significantly greater than expected by chance.

- Razão entre as variâncias:

$$F = 0.4634 / 0.0814 = 5.69 \gg 1 \text{ (mesmo considerando o } n \text{ de cada grupo)}$$

- $p = 0.0039$
- Pergunta: Como você redigiria este resultado?

Resposta

Sabemos apenas que pelo menos um dos grupos é diferente dos outros. Mas qual(is)?

Ainda não estamos prontos para redigir o resultado!

1 Comparações múltiplas

2 Análise de Variância (ANOVA)

- ANOVA um fator (One-way ANOVA)
- O teste F
- **Pós teste**
- Two-way ANOVA

3 Exercício

4 Encerramento

- O teste de ANOVA é apenas a primeira parte!²
- O p-valor do teste F indica o quão raro é encontrar uma discrepância tão grande (ou maior) entre as médias dos grupos, ao acaso
- Mas isso não nos ajuda a saber qual grupo é diferente dos outros.
- Para esta outra pergunta, precisamos de outro método

²Está com saudade do teste t?

- Como vimos, não podemos simplesmente fazer vários testes t
- Mas podemos *ajustar* os p-valores destes testes, para compensar a *inflação* destes resultados
- Isso pode ser feito de várias maneiras

- **Correção de Bonferroni**
- Correção para tendências
- **Teste “honesto” das diferenças, de Tukey (HSD)**
- Método de Scheffe
- Teste de Dunnet

- Os dois mais usados são Bonferroni e Tukey
- O teste de Bonferroni ajusta o p-valor dividindo pelo número de comparações, mas seus ICs são muito grandes
- O teste de Tukey é mais conservador, mas pode acusar diferenças significativas com mais frequência
- Infelizmente não há consenso sobre critérios de escolha

Exemplo 13.5

Table 30.3. InStat Results for Tukey's Post Test

| Comparison | Mean Difference | q | P Value |
|----------------------------|-----------------|--------|-------------|
| Nonrunners vs Recreational | 0.1400 | 2.741 | ** P < 0.01 |
| Nonrunners vs Elite | 0.1200 | 2.741 | ns P > 0.05 |
| Recreational vs Elite | -0.02000 | 0.4574 | ns P > 0.05 |

| Difference | Mean Difference | Lower 95% CI | Upper 95% CI |
|---------------------------|-----------------|--------------|--------------|
| Nonrunners — Recreational | 0.1400 | 0.03823 | 0.2418 |
| Nonrunners — Elite | 0.1200 | -0.02688 | 0.2669 |
| Recreational — Elite | -0.02000 | -0.1667 | 0.1267 |

- Pergunta: Como você redigiria este resultado?

Comparações
múltiplas

ANOVA

ANOVA um fator
(One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Encerramento

1 Comparações múltiplas

2 Análise de Variância (ANOVA)

- ANOVA um fator (One-way ANOVA)
- O teste F
- Pós teste
- Two-way ANOVA

3 Exercício

4 Encerramento

- Nas seções anteriores vimos como executar o ANOVA com uma var. independente categórica
- O teste ANOVA permite qualquer quantidade de variáveis independentes!
- Vejamos o exemplo inicial da aula, com duas: incluindo o Gênero
- Agora a pergunta é dupla: as médias são diferentes, quando estratificamos por uma segunda var. categórica?

Esses 3 grupos têm médias diferentes, controlando por Gênero?

Felipe
Figueiredo

Comparações
múltiplas

ANOVA

ANOVA um fator
(One-way ANOVA)

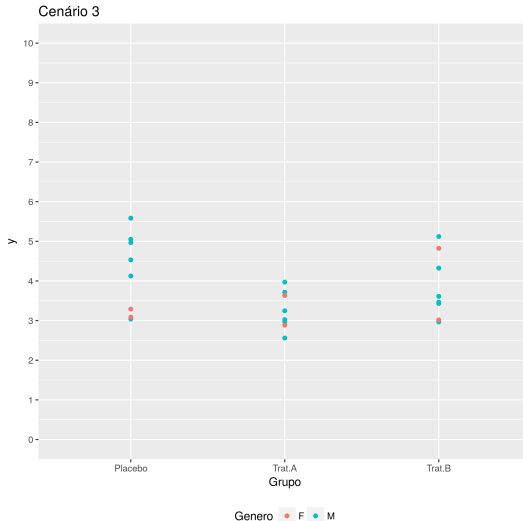
O teste F

Pós teste

Two-way ANOVA

Exercício

Encerramento



Esses 3 grupos têm médias diferentes, controlando por Gênero?

Felipe
Figueiredo

Comparações
múltiplas

ANOVA

ANOVA um fator
(One-way ANOVA)

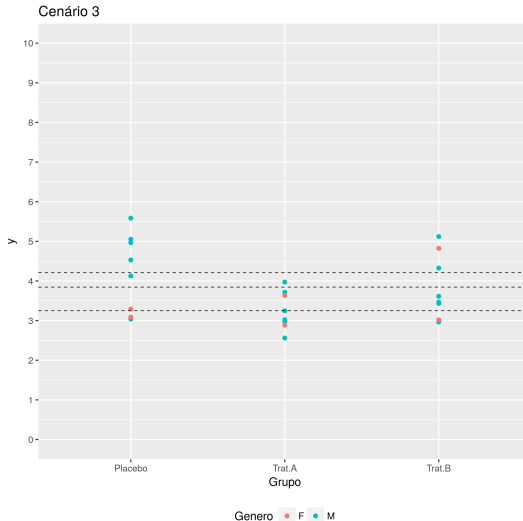
O teste F

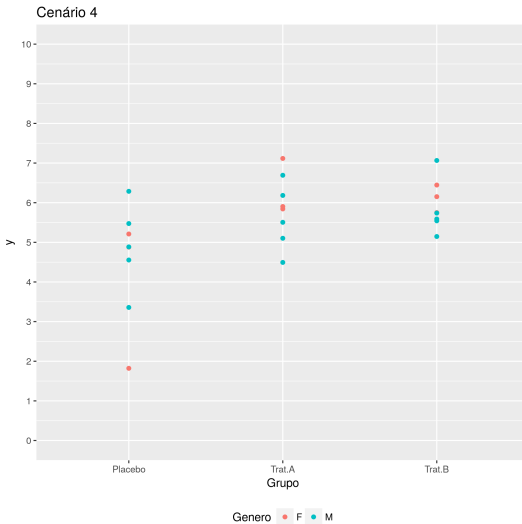
Pós teste

Two-way ANOVA

Exercício

Encerramento





E estes 3 grupos?

Felipe
Figueiredo

Comparações
múltiplas

ANOVA

ANOVA um fator
(One-way ANOVA)

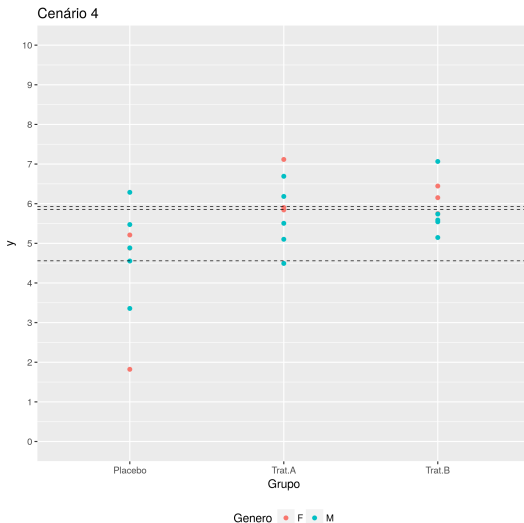
O teste F

Pós teste

Two-way ANOVA

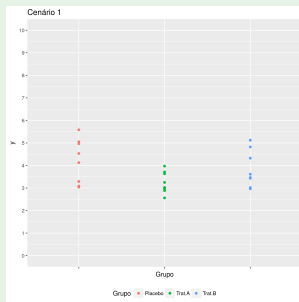
Exercício

Encerramento



Cenário 1 - ANOVA one-way

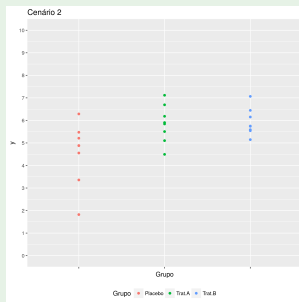
```
          Df Sum Sq Mean Sq F value Pr(>F)
Grupo      2   3.753   1.8763    3.025 0.0701 .
Residuals 21 13.026   0.6203
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Cenário 2 - ANOVA one-way

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|-----------|----|--------|---------|---------|----------|
| Grupo | 2 | 9.499 | 4.749 | 4.775 | 0.0195 * |
| Residuals | 21 | 20.889 | 0.995 | | |

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



Cenário 2 - Tukey

Tukey multiple comparisons of means
95% family-wise confidence level

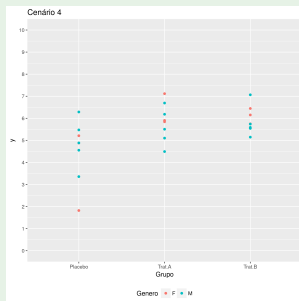
Fit: aov(formula = y ~ Grupo, data = cenario2.long)

| \$Grupo | | diff | lwr | upr | p adj |
|----------------|------------|------------|----------|-----------|-------|
| Trat.A-Placebo | 1.29615978 | 0.0392117 | 2.553108 | 0.0424949 | |
| Trat.B-Placebo | 1.36988994 | 0.1129419 | 2.626838 | 0.0311078 | |
| Trat.B-Trat.A | 0.07373016 | -1.1832179 | 1.330678 | 0.9880276 | |

Cenário 4 - ANOVA two-way (sem interações)

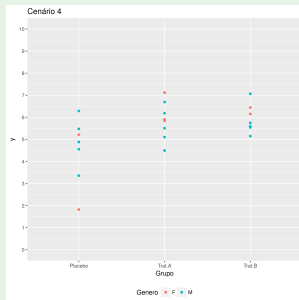
| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|-----------|----|--------|---------|---------|----------|
| Grupo | 2 | 9.499 | 4.749 | 4.548 | 0.0236 * |
| Genero | 1 | 0.002 | 0.002 | 0.002 | 0.9690 |
| Residuals | 20 | 20.887 | 1.044 | | |

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



Cenário 4 - ANOVA two-way (com interações)

```
      Df Sum Sq Mean Sq F value    Pr(>F)
Grupo    2 20.001   10.000    9.935 0.00124 **
Genero    1  3.665    3.665    3.641 0.07246 .
Grupo:Genero  2  3.157    1.579    1.568 0.23557
Residuals 18 18.119    1.007
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Cenário 4 - ANOVA two-way (com interações)

Tukey multiple comparisons of means
95% family-wise confidence level

```
Fit: aov(formula = y ~ Grupo * Genero, data = cenario2.long)
```

```
$Grupo
              diff          lwr          upr          p adj
Trat.A-Placebo 1.5514853  0.2712011 2.831770 0.0164455
Trat.B-Placebo 2.1703237  0.8900395 3.450608 0.0011265
Trat.B-Trat.A  0.6188384 -0.6614458 1.899123 0.4494538
```

```
$Genero
              diff          lwr          upr          p adj
M-F 0.8071626 -0.08158125 1.695906 0.0724633
```

```
$ 'Grupo:Genero'
              diff          lwr          upr          p adj
Trat.A:F-Placebo:F 2.1160544 -0.4873461 4.719455 0.1523427
Trat.B:F-Placebo:F 1.5879521 -1.0154484 4.191353 0.4122593
Placebo:M-Placebo:F 0.7976679 -1.5308843 3.126220 0.8795976
Trat.A:M-Placebo:F 2.0104118 -0.3181404 4.338964 0.1143524
Trat.B:M-Placebo:F 3.3174146  0.9888624 5.645967 0.0030192
Trat.B:F-Trat.A:F -0.5281023 -3.1315028 2.075298 0.9857698
Placebo:M-Trat.A:F -1.3183865 -3.6469387 1.010166 0.4902167
Trat.A:M-Trat.A:F -0.1056426 -2.4341948 2.222910 0.9999896
Trat.B:M-Trat.A:F  1.2013602 -1.1271920 3.529912 0.5849474
Placebo:M-Trat.B:F -0.7902842 -3.1188364 1.538268 0.8835616
Trat.A:M-Trat.B:F  0.4224597 -1.9060925 2.751012 0.9913898
Trat.B:M-Trat.B:F  1.7294625 -0.5990897 4.058015 0.2216761
```

Felipe
Figueiredo

Comparações
múltiplas

ANOVA

Exercício

Encerramento

Leitura obrigatória

- Capítulo 13
- Capítulo 30

Exercícios

Capítulo 13, problema: 1