



Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas  
ANOVA

Exercício

Aprofundamento

## Comparações múltiplas e ANOVA

Teste paramétrico para vários grupos (desfecho quantitativo)

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

## Sumário

- 1 **Comparações múltiplas**
  - O acaso prega peças
  - Comparações múltiplas
- 2 **Análise de Variância (ANOVA)**
  - ANOVA um fator (One-way ANOVA)
  - O teste F
  - Pós teste
  - Two-way ANOVA
- 3 **Exercício**
  - Exercício
- 4 **Aprofundamento**
  - Aprofundamento



Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

Exercício

Aprofundamento



Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

Exercício

Aprofundamento

Discussão da aula passada

Discussão da leitura obrigatória da aula passada



Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

Coincidências

Comparações  
múltiplas

ANOVA

Exercício

Aprofundamento

A probabilidade de *você* acertar na Mega Sena...

... é grande ou pequena?

A probabilidade de *alguém* acertar na Mega Sena...  
... é grande ou pequena?

### 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?)

Coincidências podem não ser tão raras assim

- 1 Você formulou a hipótese **após** observar o agrupamento de casos
- 2 Você só destacou a escola por causa do agrupamento
- 3 Agrupamentos ocorrem ao acaso
- 4 Definir população:
  - População de escolas (cidade, estado...?)
  - Tempo de observação (mês, ano, década...?)

Considerando o tempo, e o número de escolas da população...

... um agrupamento deste tamanho é realmente improvável?

### 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?

Coincidências podem ocorrer ao testar múltiplas hipóteses

Como comparar dois grupos?

“Comparar” é um termo vago...

... precisamos de um critério bem definido!

### Para comparar quanto às variâncias dos grupos

Podemos usar

- Teste F
- 

### Para comparar quanto às médias dos grupos

Teste t

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

Teste ...

- Vimos que o **teste t** pode ser usado para comparar duas médias
- Assumindo que atendemos às premissas do teste t, consideramos:
  - variabilidade dos grupos<sup>1</sup>
  - tamanho do estudo (n)<sup>2</sup>

### Requisitos não óbvios (além das médias)

desvio padrão + n = erro padrão

<sup>1</sup> Se possível, semelhantes. Caso contrário, correção de Welch.

<sup>2</sup> Componente do DP e do SEM. Usado como GL para o t crítico.

### Exercício

Um cirurgião testa duas drogas para auxiliar a recuperação pós cirúrgica, e mensura a área cicatrizada (y) em uma semana.

São considerados os tratamentos A e B e um Placebo.

Foram selecionados 8 participantes para cada um dos três grupos.

- Dependente:
  - numérica contínua
- Independentes:
  - grupo (categórica nominal – 3 níveis)

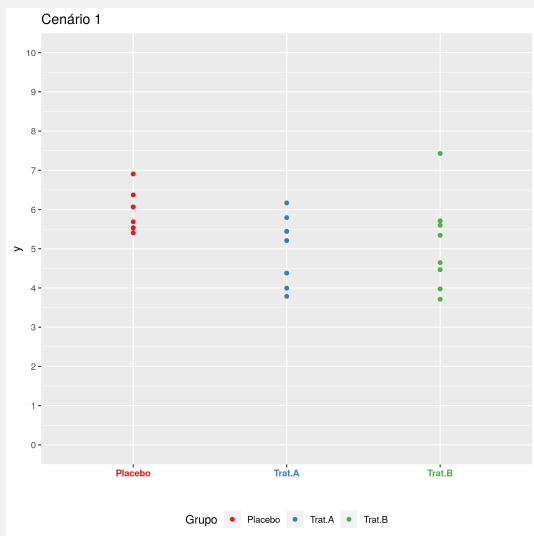
### Esta relação pode ser expressa como

Área cicatrizada ~ Grupo de tratamento

Você consegue decidir **visualmente**...

... se 3 grupos têm médias diferentes?

## Cenário 1 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

Coincidências

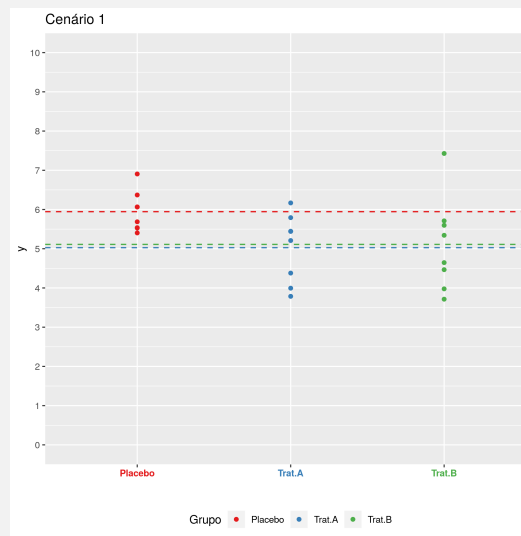
Comparações múltiplas

ANOVA

Exercício

Aprofundamento

Médias: Placebo: 5.945, Tratamento A: 5.027, Tratamento B: 5.110



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

Coincidências

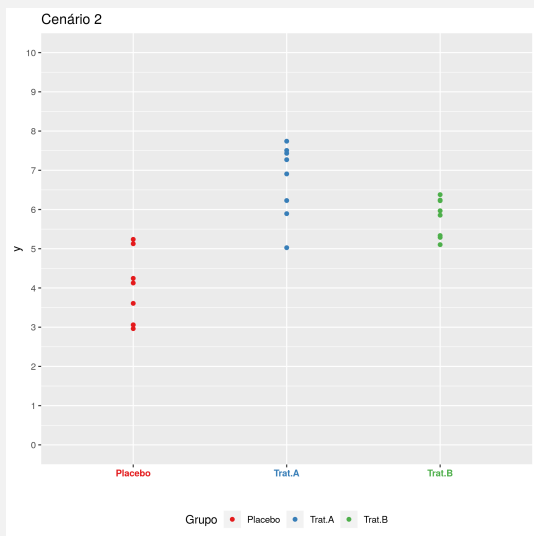
Comparações múltiplas

ANOVA

Exercício

Aprofundamento

## Cenário 2 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

Coincidências

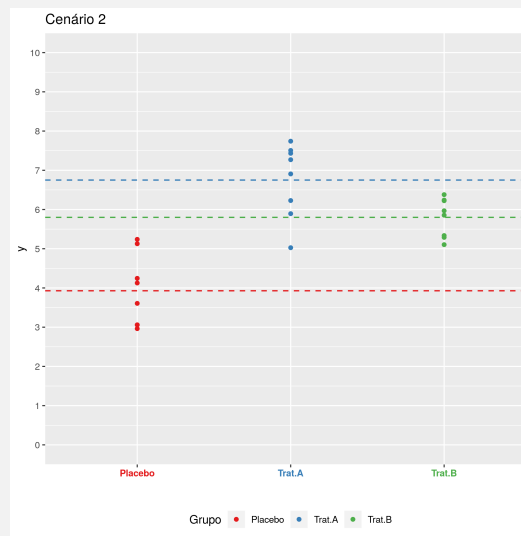
Comparações múltiplas

ANOVA

Exercício

Aprofundamento

Médias: Placebo: 3.928, Tratamento A: 6.751, Tratamento B: 5.799



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

Coincidências

Comparações múltiplas

ANOVA

Exercício

Aprofundamento

## Abordagem mais simples

Uma ideia seria usar o teste t três vezes...

... comparando os grupos, dois a dois.

## Proposta

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

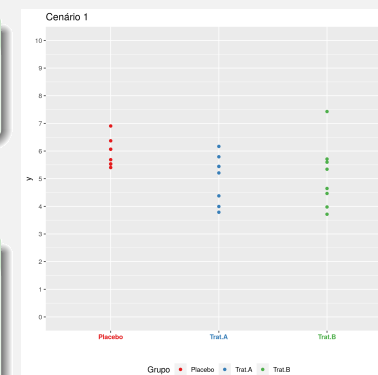
## P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p = 0.025$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.100$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.876$

## Pergunta

Os tratamentos são diferentes do placebo?

E entre si?



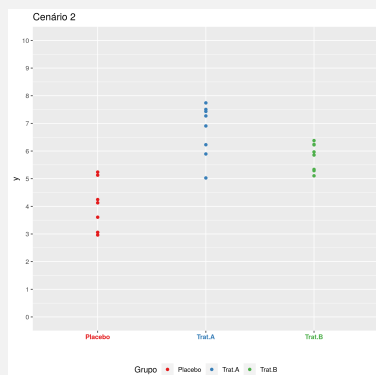
## P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p < 0.0001$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.00037$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.02943$

## Pergunta

Os tratamentos são diferentes do placebo?

E entre si?



Existe um problema oculto aí.

## O problema é...

- A conclusão de que no Cenário 1 os 3 grupos são diferentes está **errada!**
- No Cenário 2, os 2 tratamentos **não são** diferentes entre si!

- O teste t permite a avaliação de **uma** hipótese
- Testamos simultaneamente três<sup>3</sup>
- Isto aumenta a chance de cometermos um erro tipo I (falso positivo)

Nível de significância de cada teste  $\neq$  nível de significância global.

<sup>3</sup> Leia várias vezes o Capítulo 13!

- Ao testar uma hipótese, assumimos 5% de erro tipo I
- Mas se testarmos 100 hipóteses, o número esperado de falsos positivos é 5!
- Se os testes forem independentes, a probabilidade de **pelo menos 1** erro tipo I é 99.4%.<sup>4</sup>

<sup>4</sup> [https://en.wikipedia.org/wiki/Multiple\\_comparisons\\_problem](https://en.wikipedia.org/wiki/Multiple_comparisons_problem)

## Pensar é obrigatório

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<sup>5</sup>.

Nível de significância de cada teste  $\neq$  nível de significância global.

<sup>5</sup> Leia várias vezes o Capítulo 13!

## 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, consideramos:
  - variabilidade dos grupos<sup>1</sup>
  - tamanho do estudo (n)<sup>2</sup>

### Requisitos não óbvios (além das médias)

desvio padrão + n = erro padrão

<sup>1</sup> Se possível, semelhantes. Caso contrário, correção de Welch.

<sup>2</sup> Componente do DP e do SEM. Usado como GL para o t crítico.

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas  
Coincidências  
Comparações  
múltiplas

ANOVA

Exercício

Aprofundamento

### 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

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas  
ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

- Dependente:
  - numérica contínua
- Independente:
  - grupo (categórica nominal – 3 níveis)

### Esta relação pode ser expressa como

$LH \sim \text{Grupo}$

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

### Versão simplificada (apenas variáveis)

$LH \sim \text{Grupo}$

### Modelo completo

$LH = \text{média global} + \text{efeito do fator grupo} + \varepsilon$

Hipótese:  $\varepsilon$  é um erro aleatório <sup>6</sup> normalmente distribuído e centrado em zero – a incerteza que não pode ser controlada.

<sup>6</sup> residual – não é explicado pela relação entre as variáveis do modelo

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento



## Exemplo 13.5

**Table 30.1.** LH Levels in Three Groups of Women

Group	log(LH) $\pm$ SEM	N
Nonrunners	0.52 $\pm$ 0.027	88
Recreational runners	0.38 $\pm$ 0.034	89
Elite runners	0.40 $\pm$ 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	0.9268	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.

- 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 intragrupos
- A ANOVA tenta *desembarçar* esta decomposição, assumindo a hipótese nula.

- O nome *Análise de Variância* vem do critério usado para comparar as médias
- O teste é baseado na razão 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

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

Se as médias forem iguais, a variância intragrupo deve ser “igual” à variância intergrupo...

... nesse caso a *razão* entre as variâncias deve ser próxima de 1

$$F = \frac{\text{variância intergrupos}}{\text{variância intragrupos}}$$

### Interpretação da estatística *F*

Uma razão muito maior que 1 indica que há mais variância entre os grupos do que o esperado

<sup>7</sup> Mesma ideia do qui-quadrado.

<sup>8</sup> O teste leva em conta dois graus de liberdade: numerador e denominador

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

### 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:  $F = 0.4634/0.0814 = 5.69 \gg 1$   
(mesmo considerando o  $n$  de cada grupo)
- $p = 0.0039$

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

### Resposta

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

Ainda não estamos prontos para redigir o resultado!

- O teste de ANOVA é apenas a primeira parte!<sup>9</sup>

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** é o grupo discrepante
- Para esta outra pergunta, precisamos de outro método

<sup>9</sup> Está com saudade do teste t?

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste  
Two-way ANOVA

Exercício

Aprofundamento

- 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

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste  
Two-way ANOVA

Exercício

Aprofundamento

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

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste  
Two-way ANOVA

Exercício

Aprofundamento

- Faremos os múltiplos testes t, com ajuste de p-valor
- Os dois mais usados são Bonferroni e Tukey
- O ajuste de Bonferroni multiplica o p-valor<sup>10</sup> pelo número de comparações, mas seus IC são muito grandes
- O ajuste de Tukey é mais conservador, mas pode acusar diferenças significativas com mais frequência
- Infelizmente não há consenso sobre critérios de escolha

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas

ANOVA

ANOVA um fator  
(One-way ANOVA)

O teste F

Pós teste  
Two-way ANOVA

Exercício

Aprofundamento

<sup>10</sup> ou, analogamente, divide o nível de significância

**Exemplo**

García-Arenzana et al. (2014) testaram associação de 25 variáveis dietárias e a densidade mamográfica (relevante p/ câncer de mama).

5 das variáveis parecem significativas.

**Bonferroni**

Ao dividir 0.05 pelo número de comparações, obtemos  $\alpha = 0.05/25 = 0.002$ .

**Conclusão**

**Após o ajuste, apenas 1 significativo.**

<sup>11</sup> <http://www.biostathandbook.com/multiplecomparisons.html>

**p-valores não ajustados**

Dietary variable	P value
Total calories	<0.001
Olive oil	0.008
Whole milk	0.039
White meat	0.041
Proteins	0.042
Nuts	0.06
Cereals and pasta	0.074
White fish	0.205
Butter	0.212
Vegetables	0.216
Skimmed milk	0.222
Red meat	0.251
Fruit	0.269
Eggs	0.275
Blue fish	0.34
Legumes	0.341
Carbohydrates	0.384
Potatoes	0.569
Bread	0.594
Fats	0.696
Sweets	0.762
Dairy products	0.94
Semi-skimmed milk	0.942
Total meat	0.975
Processed meat	0.986

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

**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

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

**Exemplo 13.5**

**Table 30.1.** LH Levels in Three Groups of Women

Group	log(LH) $\pm$ SEM	N
Nonrunners	0.52 $\pm$ 0.027	88
Recreational runners	0.38 $\pm$ 0.034	89
Elite runners	0.40 $\pm$ 0.049	28

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

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

**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:  $F = 0.4634/0.0814 = 5.69 \gg 1$   
(mesmo considerando o  $n$  de cada grupo)
- $p = 0.0039$

Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## 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?

Quais são as variáveis?

- Dependente:
  - numérica contínua
- Independentes:
  - grupo (categórica nominal – 3 níveis)
  - gênero (categórica nominal binária)**

Esta relação pode ser expressa como

Área cicatrizada ~ Grupo de tratamento + **Gênero**

- Vimos como usar o ANOVA com **uma** var. independente categórica

- O teste ANOVA permite qualquer quantidade de variáveis independentes! E de qualquer tipo<sup>12</sup>

- Vejamos o exemplo inicial da aula, com duas var. independentes

## Nova pergunta

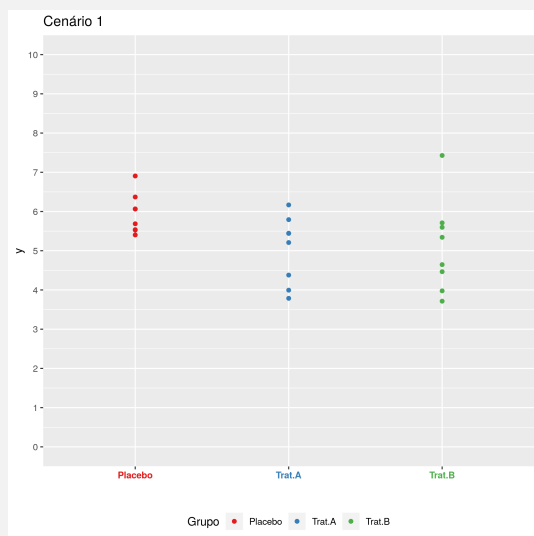
Os tratamentos são diferentes, mesmo controlando pelo Gênero?

<sup>12</sup> Na verdade, ANOVA e Regressão Linear são Múltipla são siameses

Você consegue decidir **visualmente**...

... se 3 grupos têm médias diferentes?

## Cenário 1 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA  
Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

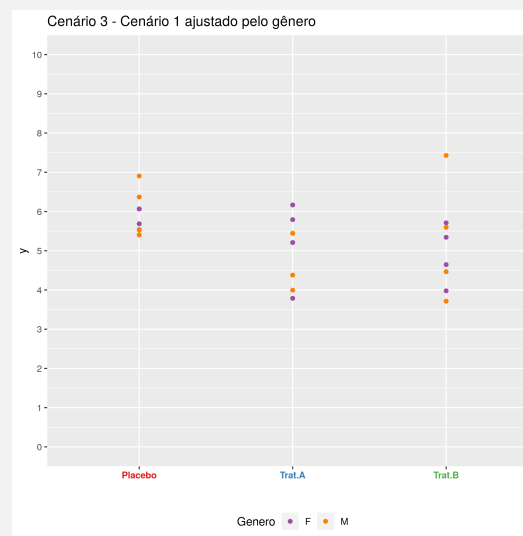
Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## Cenário 3 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA  
Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

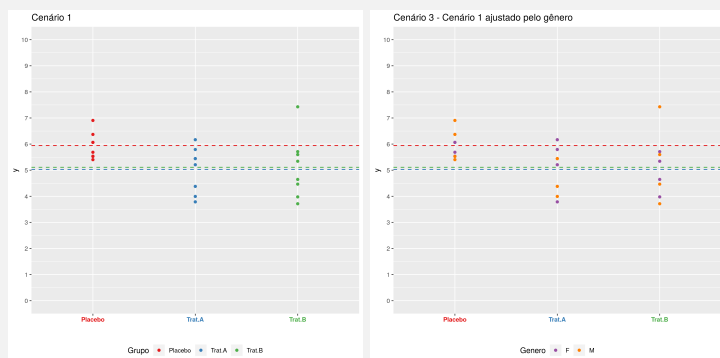
Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## Cenário 3 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA  
Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

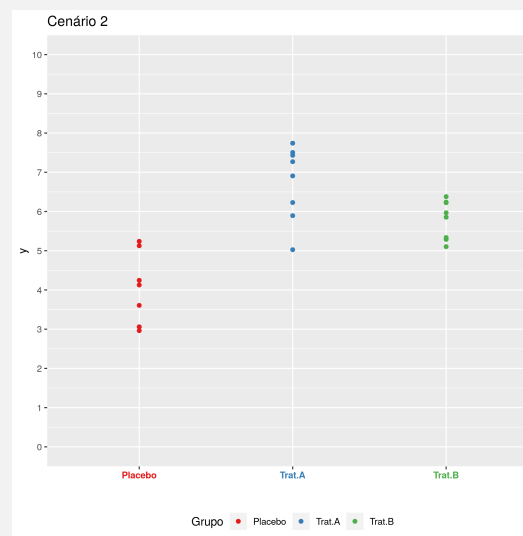
Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## Cenário 2 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA  
Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

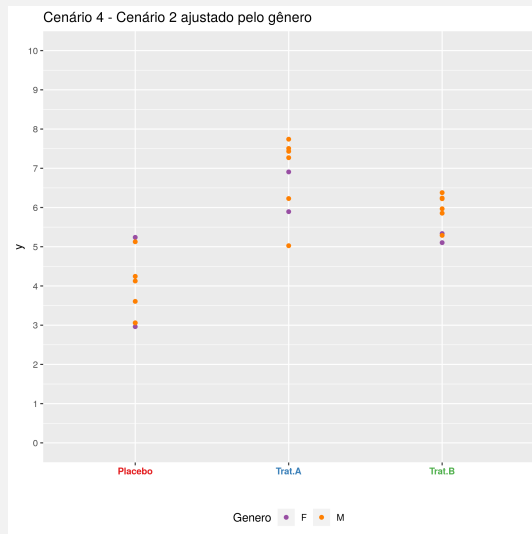
Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## Cenário 4 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

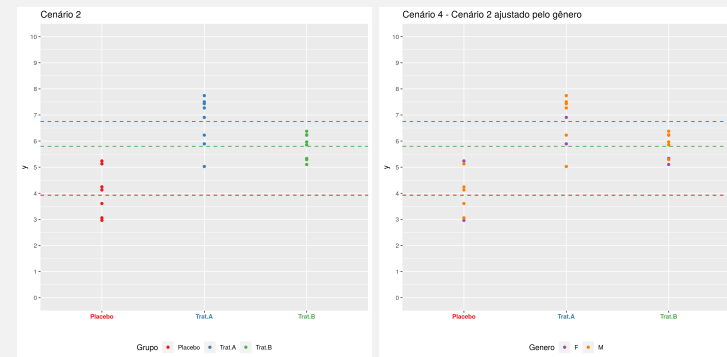
Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## Cenário 4 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

Hora de testar seus conhecimentos



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

ANOVA um fator (One-way ANOVA)

O teste F

Pós teste

Two-way ANOVA

Exercício

Aprofundamento

## Exercício



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

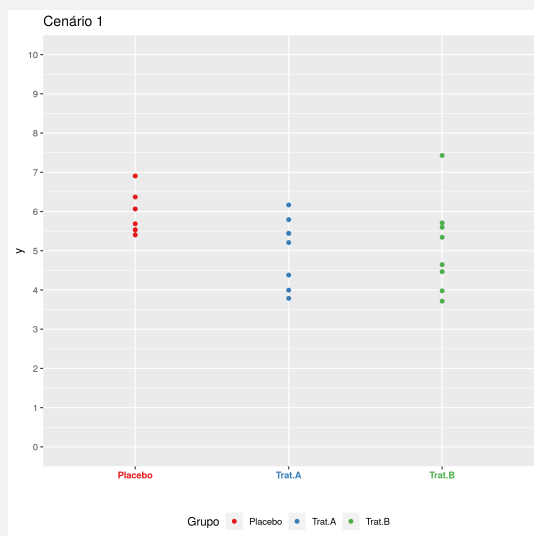
### Exercício

Um cirurgião testa duas drogas para auxiliar a recuperação pós cirúrgica, e mensura a área cicatrizada ( $y$ ) em uma semana.

São considerados os tratamentos A e B e um Placebo.

Foram selecionados 8 participantes para cada um dos três grupos.

## Cenário 1 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

## Cenário 1



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

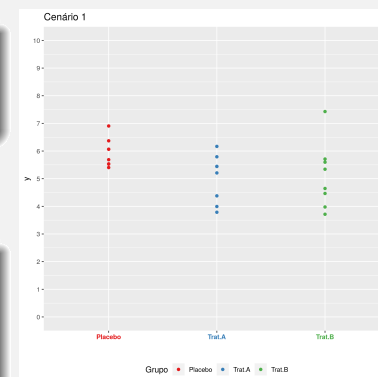
### P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p = 0.025$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.100$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.876$

### Pergunta

Os tratamentos são diferentes do placebo?

E entre si?



Por que este resultado está errado?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

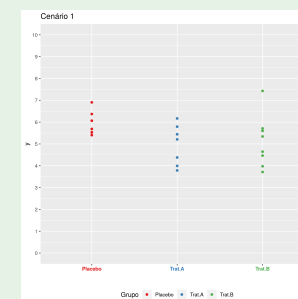
### Resposta

Testamos simultaneamente **3** hipóteses...

... você foi levado ao engano: Placebo é diferente do trat. A

### Cenário 1 – ANOVA One-way

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Grupo	2	4.124	2.0620	2.545	0.102
Residuals	21	17.018	0.8104		



Os 3 tratamentos não diferem além da expectativa ( $p = 0.102$ )



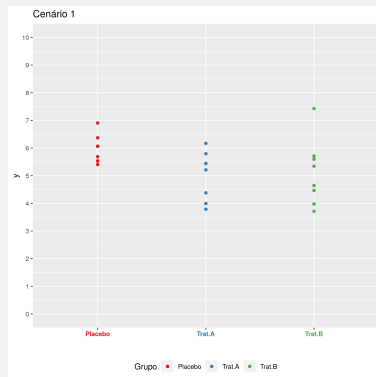
## P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p = 0.025$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.100$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.876$

## Pergunta

Os tratamentos são diferentes do placebo?

E entre si?



## Cenário 1 – Bonferroni

Pairwise comparisons using t tests with non-pooled SD

data: y and Grupo

```

      Placebo Trat.A
Trat.A 0.076    -
Trat.B 0.299    1.000

```

P value adjustment method: bonferroni

Os p-valores de Bonferroni são 3x maiores...

... o placebo não é diferente do tratamento A ( $p = 0.076$ )

## Cenário 1 – Tukey

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

Fit: aov(formula = y ~ Grupo, data = cenario1.long)

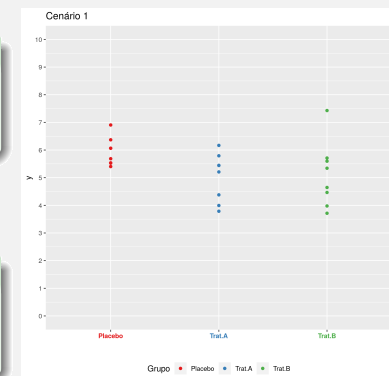
```

$Grupo              diff      lwr      upr    p adj
Trat.A-Placebo -0.91797498 -2.052498 0.2165479 0.1274511
Trat.B-Placebo -0.83482042 -1.969343 0.2997025 0.1767378
Trat.B-Trat.A   0.08315455 -1.051368 1.2176774 0.9813768

```

Os p-valores de Tukey são mais conservadores...

... o placebo não é diferente do tratamento A ( $p = 0.12745$ )



## Cenário 1 – Solução

## P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p = 0.025$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.100$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.876$

## P-valores ajustados (Bonferroni)

- 1 Placebo x Trat. A  $\Rightarrow p = 0.076$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.299$
- 3 Trat. A x Trat. B  $\Rightarrow p = 1.000$

## Cenário 2 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

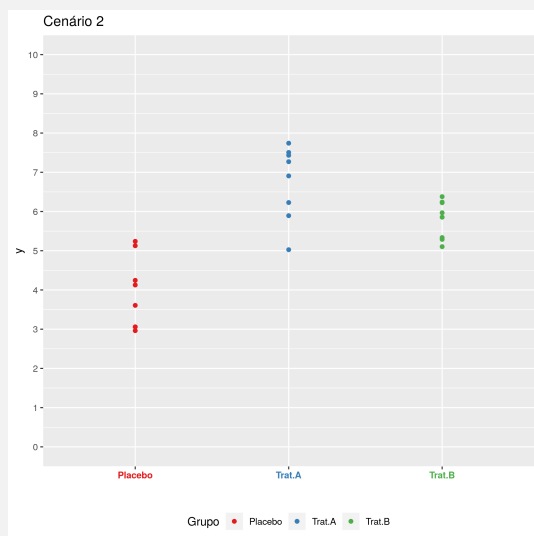
Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento



## Cenário 2



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

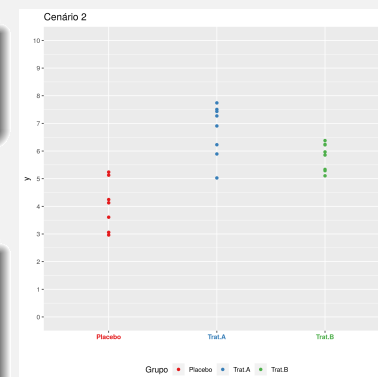
### P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p < 0.0001$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.00037$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.02943$

### Pergunta

Os tratamentos são diferentes do placebo?

E entre si?



Por que este resultado está errado?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

### Resposta

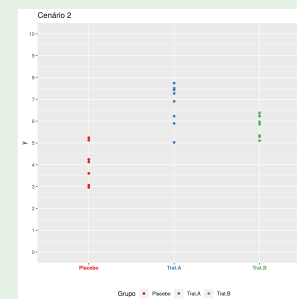
Testamos simultaneamente **3** hipóteses...

... você foi levado ao engano: trat. A é diferente do trat. B

### Cenário 2 – ANOVA One-way

Grupo	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Grupo	2	32.99	16.496	25.04	2.75e-06 ***
Residuals	21	13.83	0.659		

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



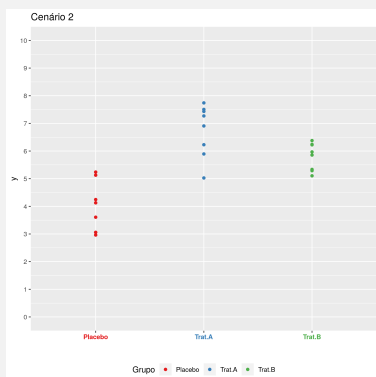
## P-valores dos 3 testes t

- 1 Placebo x Trat. A  $\Rightarrow p < 0.0001$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.00037$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.02943$

## Pergunta

Os tratamentos são diferentes do placebo?

E entre si?



## Cenário 2 – Bonferroni

Pairwise comparisons using t tests with non-pooled SD

data: y and Grupo

```
Placebo Trat.A
Trat.A 8.8e-05 -
Trat.B 0.0011 0.0883
```

P value adjustment method: bonferroni

Os p-valores de Bonferroni são 3x maiores...

... os tratamentos A e B não são diferentes entre si ( $p = 0.0883$ )

## Cenário 2 – Tukey

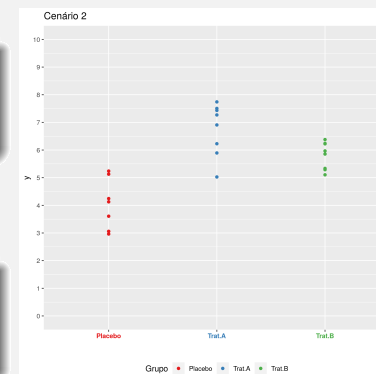
Tukey multiple comparisons of means  
95% family-wise confidence level

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

\$Grupo		diff	lwr	upr	p adj
Trat.A-Placebo	2.8224313	1.7995273	3.8453353	0.0000021	
Trat.B-Placebo	1.8711918	0.8482877	2.8940958	0.0004262	
Trat.B-Trat.A	-0.9512395	-1.9741435	0.0716645	0.0713859	

Os p-valores de Tukey são mais conservadores...

... os tratamentos A e B não são diferentes ( $p = 0.0713859$ )



## Cenário 2 – Solução

## P-valores dos 3 testes t

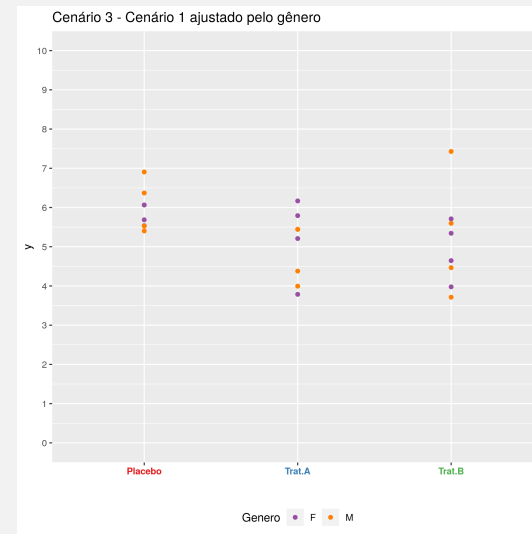
- 1 Placebo x Trat. A  $\Rightarrow p < 0.0001$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.00037$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.02943$

## P-valores ajustados (Bonferroni)

- 1 Placebo x Trat. A  $\Rightarrow p < 0.0001$
- 2 Placebo x Trat. B  $\Rightarrow p = 0.0011$
- 3 Trat. A x Trat. B  $\Rightarrow p = 0.0883$

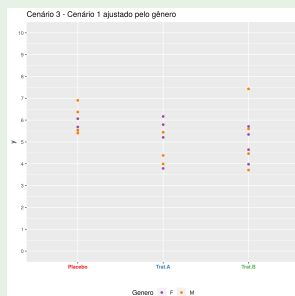
Agora interprete cada um dos dois fatores

Cenário 3 – esses 3 grupos têm médias diferentes?



### Cenário 3 – ANOVA Two-way

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Grupo	2	4.124	2.0620	2.426	0.114
Genero	1	0.020	0.0198	0.023	0.880
Residuals	20	16.998	0.8499		



### Cenário 3 – Tukey

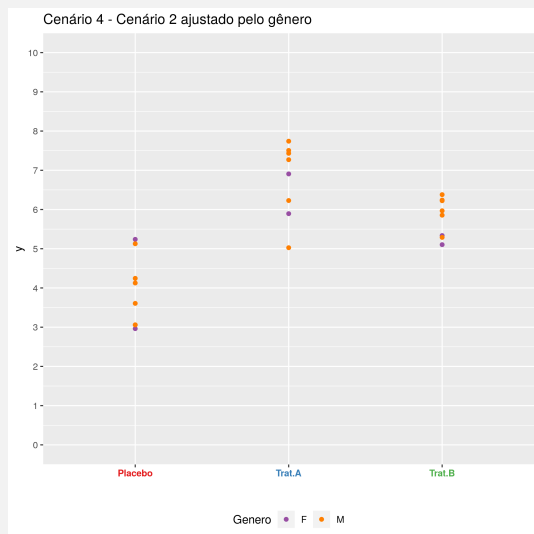
Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: aov(formula = y ~ Grupo + Genero, data = cenariol.long)

```
$Grupo
      diff      lwr      upr    p adj
Trat.A-Placebo -0.91797498 -2.084178 0.2482277 0.1402196
Trat.B-Placebo -0.83482042 -2.001023 0.3313822 0.1915255
Trat.B-Trat.A   0.08315455 -1.083048 1.2493572 0.9822352
```

```
$Genero
      diff      lwr      upr    p adj
M-F 0.05741033 -0.7276764 0.8424971 0.8802907
```

## Cenário 4 – esses 3 grupos têm médias diferentes?



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

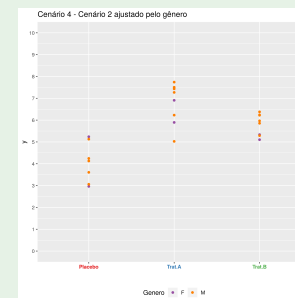
Exercício

Aprofundamento

## Cenário 4 – ANOVA Two-way

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Grupo	2	32.99	16.496	24.760	3.88e-06 ***
Genero	1	0.51	0.509	0.764	0.393
Residuals	20	13.33	0.666		

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



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

## Cenário 4 – Tukey

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	2.8224313	1.789885	3.85497800	0.000030
Trat.B-Placebo	1.8711918	0.838645	2.90373849	0.0005050
Trat.B-Trat.A	-0.9512395	-1.983786	0.08130722	0.0743628

\$Genero	diff	lwr	upr	p adj
M-F	0.3362835	-0.4663601	1.138927	0.3925159



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

## Resumo

- Vimos o modelo **ANOVA com fatores fixos** para comparar médias
- Há também...<sup>13</sup>
  - ANOVA com interações entre os fatores**
  - ANOVA com Medidas Repetidas** quando você mensura do participante em vários momentos diferentes<sup>14</sup> (ex: baseline, pré-op imediato, pós-op imediato, e após 1 ano)
  - ANOVA com fatores aleatórios** permite decompor as variâncias – contribuição de cada fator para a variância total
  - ANOVA com fatores mistos** – fatores fixos E aleatórios
- ANOVA é a base para (livros de) Design of Experiments (DoE)
- Considere sempre usar **desenhos balanceados**<sup>15</sup>!

<sup>13</sup> Todos fora do escopo deste curso

<sup>14</sup> RM ANOVA é mencionado no livro-texto

<sup>15</sup> Grupos com **mesmo tamanho**



Comparações múltiplas e ANOVA

Felipe Figueiredo

Comparações múltiplas

ANOVA

Exercício

Exercício

Aprofundamento

- Dependente (VD): **numérica**

- discreta
- contínua

- Independentes (VI):

- categórica – 2+ níveis
- numérica discreta
- numérica contínua

Esta relação pode ser expressa como

$$VD \sim VI_1 + VI_2 + \dots$$

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas  
ANOVA

Exercício  
Exercício

Aprofundamento

Comparações  
múltiplas e  
ANOVA

Felipe  
Figueiredo

Comparações  
múltiplas  
ANOVA

Exercício

Aprofundamento  
Aprofundamento

### Leitura obrigatória

- Capítulo 13
- Capítulo 30 (atenção às premissas!)

### Leitura recomendada

Kim, Bang, 2016, *Three common misuses of P values*, Dent. Hypotheses. (editorial)