# Estatística para Cursos de Engenharia e Informática

Pedro Alberto Barbetta / Marcelo Menezes Reis / Antonio Cezar Bornia

São Paulo: Atlas, 2004

# Cap. 9 – Comparação entre tratamentos

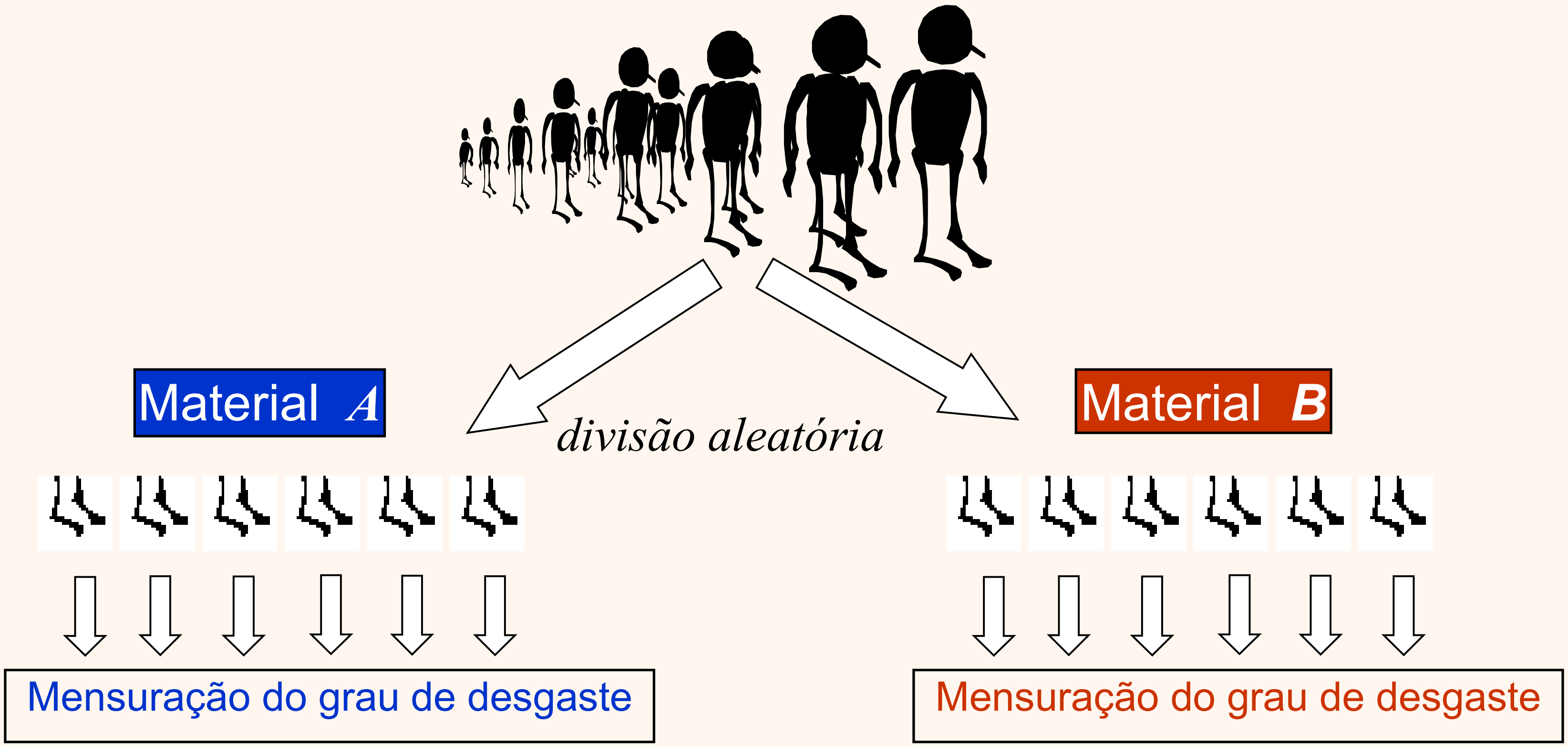
APOIO:
Fundação de Apoio à Pesquisa Científica e Tecnológica do Estado de Santa Catarina (FAPESC)
Departamento de Informática e Estatística – UFSC (INE/CTC/UFSC)

# Amostras independentes

- **Exemplo 9.1** – Considere o problema de comparar dois materiais (*A* e *B*), para sola de tênis, em termos do grau de desgaste após um certo período de uso. Seguem dois projetos de experimentos alternativos:

- **Projeto I** – Um grupo de indivíduos usa tênis com solas feitas com o material *A*; e outro grupo usa tênis com solas feitas com o material *B*.

# Amostras independentes

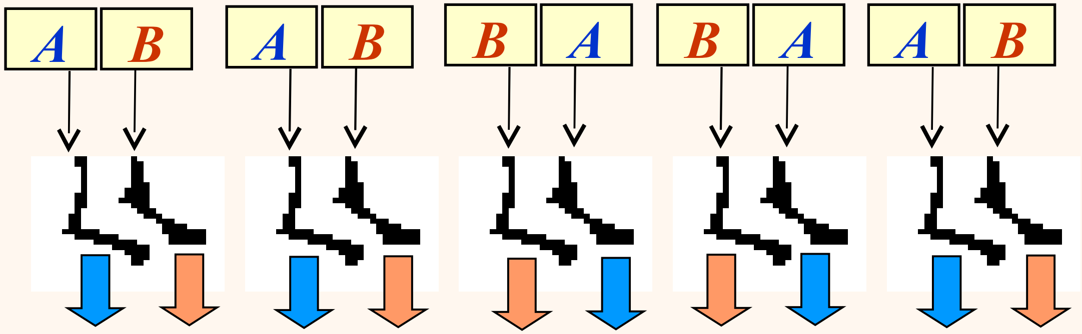


# Amostras pareadas (se $g > 2$, “em blocos”)

- **Exemplo 9.1:**
- **Projeto II** – Fabricam-se, para a realização do experimento, pares de tênis com os dois tipos de sola, isto é, um dos pés com o material *A* e o outro pé com o material *B*. Em cada par, o material usado em cada pé (direito ou esquerdo) é decidido por sorteio

# Amostras pareadas (se $g > 2$, “em blocos”)

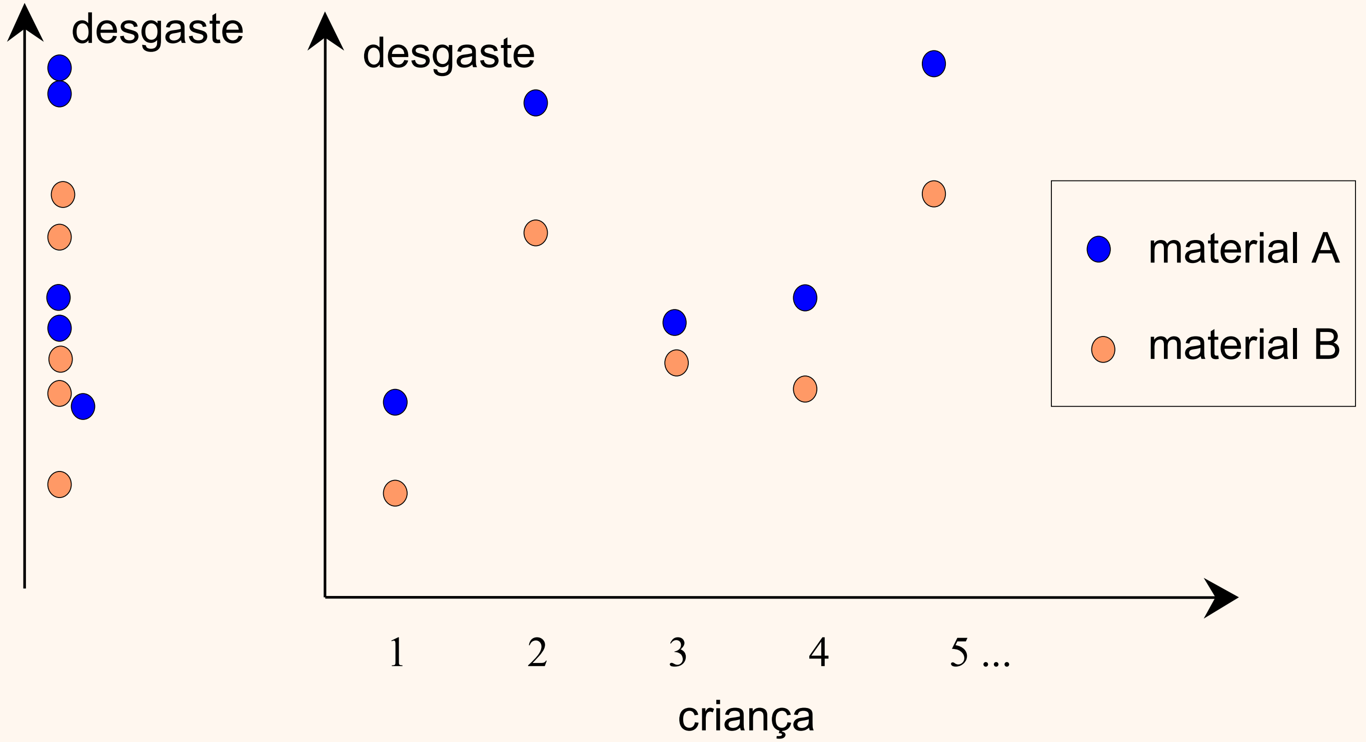
*alocação aleatória de A e B em cada par*




Mensuração do grau de desgaste

# Amostras pareadas

- Importância de considerar os pares na análise:



# Teste t para duas amostras

- $H_0$: $\mu_1 = \mu_2$ e $H_1$: $\mu_1 \neq \mu_2$

onde: $\mu_1$ é o valor esperado da resposta sob o tratamento 1 e

$\mu_2$ é o valor esperado da resposta sob o tratamento 2.

- Na abordagem unilateral, a hipótese alternativa é do tipo $H_1'$: $\mu_1 > \mu_2$ ou $H_1''$: $\mu_1 < \mu_2$.

# Teste t para duas amostras pareadas

- **Exemplo 9.2** Seja o problema de verificar se um novo algoritmo de busca em um banco de dados é mais rápido que o algoritmo atualmente usado. Para se fazer a comparação dos dois algoritmos, planeja-se realizar uma amostra aleatória de 10 buscas experimentais (ensaios). Em cada ensaio, uma dada busca é realizada pelos dois algoritmos e o tempo de resposta de cada algoritmo anotado. Observamos que em cada ensaio os dois algoritmos são usados em condições idênticas, caracterizando 10 pares de observações.

# Teste t para duas amostras pareadas

- $H_0$: em média, os dois algoritmos são *igualmente* rápidos e
- $H_1$: em média, o algoritmo novo é *mais* rápido do que o algoritmo em uso.

Ou:

- $H_0$: $\mu_2 = \mu_1$ e $H_1$: $\mu_1 < \mu_2$

onde: $\mu_1$ é o tempo esperado de resposta do algoritmo novo e

$\mu_2$ é o tempo esperado de resposta do algoritmo antigo.

# Dados:

| Ensaio | Tempo de resposta (s) | | |
|---|---|---|---|
| | Novo $X_1$ | Antigo $X_2$ | Diferença $D = X_2 - X_1$ |
| 1 | 22 | 25 | 3 |
| 2 | 21 | 28 | 7 |
| 3 | 28 | 26 | -2 |
| 4 | 30 | 36 | 6 |
| 5 | 33 | 32 | -1 |
| 6 | 33 | 39 | 6 |
| 7 | 26 | 28 | 2 |
| 8 | 24 | 33 | 9 |
| 9 | 31 | 30 | -1 |
| 10 | 22 | 27 | 5 |

# Teste t para duas amostras pareadas
# Estatística do teste

$$t = \frac{\bar{d} \cdot \sqrt{n}}{s_d}$$

- onde: $n$ é o tamanho da amostra (número de pares);
  $\bar{d}$ é a média das diferenças observadas; e
  $s_d$ é o desvio padrão das diferenças observadas.

- Usa distribuição *t* de *Student* com *gl* = *n* – 1 graus de liberdade (supondo populações com distribuição normal).

# Exemplo 9.2 (continuação)

Valores de $D$: 3, 7, -2, 6, -1, 6, 2, 9, -1, 5

$$n = 10$$

$$\bar{d} = 3{,}4$$

$$s_d = \sqrt{\frac{1}{n-1} \cdot \left( \sum_i d_i^2 - n.\bar{d}^2 \right)} = \sqrt{\frac{246 - (10)(3{,}4)^2}{9}} = 3{,}81$$

$$t = \frac{\bar{d}.\sqrt{n}}{s_d} = \frac{3{,}4.\sqrt{10}}{3{,}81} = 2{,}82$$

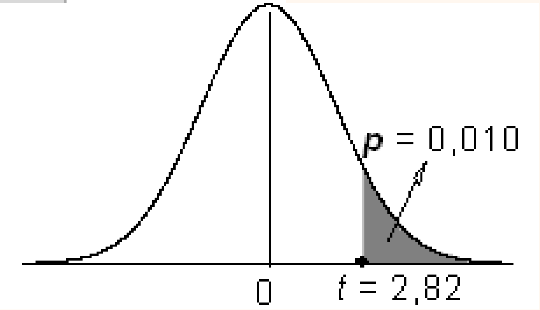
# Exemplo 9.2 (continuação). Teste considerando nível de significância de 5%.

Abordagem do Valor **p:**

dados observados

$t = 2,82$

| | Área na cauda superior | | | | | | |
|---|---|---|---|---|---|---|---|
| *gl* | 0,25 | 0,10 | 0,05 | 0,025 | **0,010** | 0,005 | ... |
| ... | | | | | | | |
| 9 | 0,703 | 1,383 | 1,833 | 2,262 | **2,821** | 3,250 | ... |
| ... | | | | | | | |




Conclusão: rejeita $H_0$.

Ver comentários e abordagem clássica no livro.

# Teste t para duas amostras independentes

**Exemplo 9.3** Desejamos verificar se os catalisadores A e B têm efeitos diferentes no rendimento de uma certa reação química. As hipóteses são:

- $H_0$: em média, os dois catalisadores são *iguais* em termos de rendimento; e
- $H_1$: em média, os dois catalisadores são *diferentes* em termos de rendimento.

Ou, ainda:

$H_0$: $\mu_1 = \mu_2$ e $H_1$: $\mu_1 \neq \mu_2$,

onde

$\mu_1$: rendimento esperado com o catalisador A; e

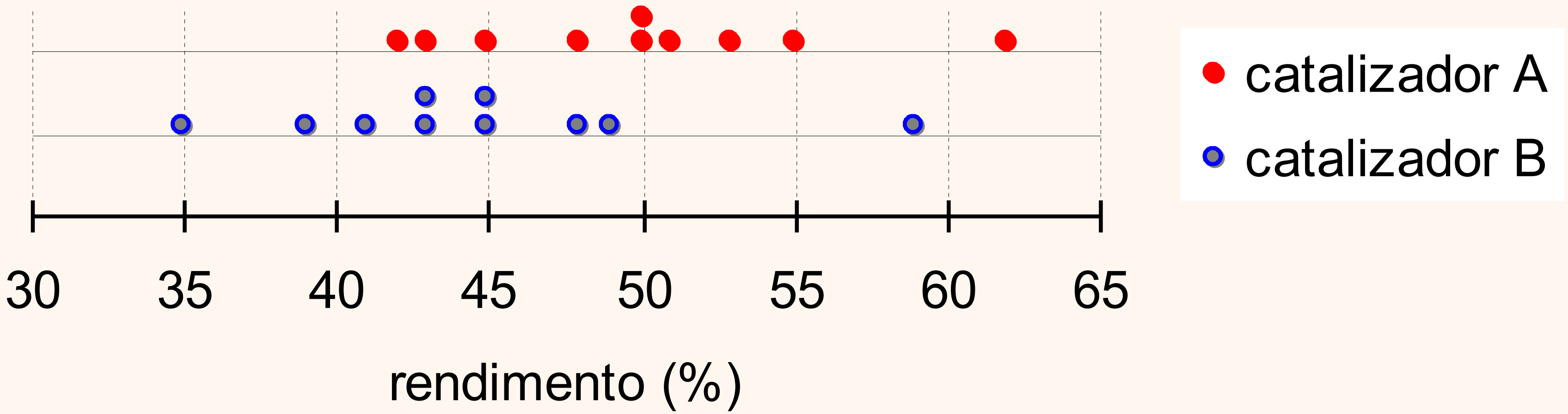
$\mu_2$: rendimento esperado com o catalisador B.

# Exemplo 9.3 – amostras:

**Tabela 9.2** Rendimentos (%) de uma reação química em função do catalisador utilizado.

| catalisador A | catalisador B |
|---|---|
| 45 51 50 62 43<br>42 53 50 48 55 | 45 35 43 59 48<br>45 41 43 49 39 |

# Exemplo 9.3 – amostras:



# Teste t para duas amostras independentes
# Estatística do teste

Se $n_1 = n_2 = n$:

$$s_a^2 = \frac{s_1^2 + s_2^2}{2}$$

$$t = \left(\bar{x}_1 - \bar{x}_2\right) \cdot \sqrt{\frac{n}{2s_a^2}}$$

$n$: tamanho da amostra em cada grupo;

$\bar{x}_1$ média da amostra 1

$\bar{x}_2$ média da amostra 2

$s_1^2$ variância da amostra 1

$s_2^2$ variância da amostra 2

$s_a^2$ variância agregada das duas amostras

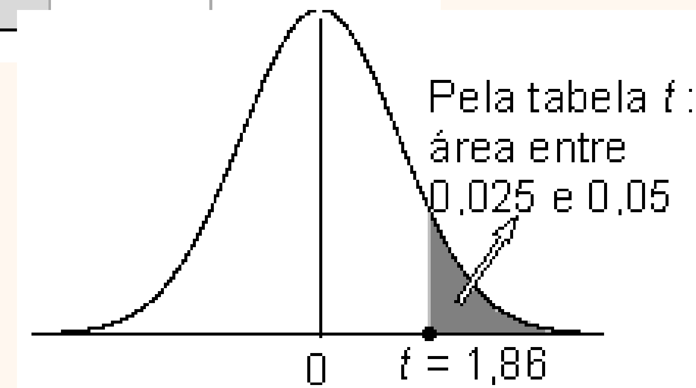
Usa distribuição $t$ de *Student* com $gl = 2n - 2$ graus de liberdade (supondo populações com distribuição normal).

# Exemplo 9.3 (continuação). Abordagem valor p

dados observados

$t = 1,86$

| | Área na cauda superior | | | | | |
|---|---|---|---|---|---|---|
| *gl* | 0,25 | 0,10 | 0,05 | 0,025 | 0,010 | 0,005 ... |
| ... | | | | | | |
| 18 | 0,688 | 1,330 | 1,734 | 2,101 | 2,552 | 2,878 ... |
| ... | | | | | | |




Portanto, 0,05 < Valor **p** < 0,10

➔ Aceita $H_0$.

Ver comentários, abordagem clássica e exemplo com $n_1 \neq n_2$ no livro.

# Comparação entre vários tratamentos. Amostras independentes

- Análise de variância (ANOVA), que supõe:
  - as observações devem ser independentes;
  - as variâncias populacionais devem ser iguais nos *g* grupos; e
  - a distribuição das observações em cada grupo deve ser normal.

# Exemplo 9.4: Comparação de três tipos de rede.

- Considere o problema de comparar 3 tipos de rede de computadores, C1, C2 e C3, em termos do tempo médio de transmissão de pacotes de dados entre duas máquinas.

- **Experimento (projeto completamente aleatorizado com um fator):** 8 replicações com cada tipo de rede, aleatorizando a ordem dos 24 ensaios e mantendo fixos os demais fatores controláveis.

# Exemplo 9.4 : Projeto do experimento.

ensaios de 1 a 8: **C1**
ensaios de 9 a 16: **C2**
ensaios de 17 a 24: **C3**

| Seqüência dos testes | número do ensaio | Uso da rede |
|---|---|---|
| 1 | 16 | C2 |
| 2 | 14 | C2 |
| 3 | 24 | C3 |
| 4 | 6 | C1 |
| ... | ... | ... |
| 24 | 11 | C3 |

# Exemplo 9.4. Dados do experimento:

| Seqüência dos testes | número do ensaio | Rede | Tempo de resposta **(y)** |
| --- | --- | --- | --- |
| 1 | 16 | C2 | **7,8** |
| 2 | 14 | C2 | **8,2** |
| 3 | 24 | C3 | **6,3** |
| 4 | 6 | C1 | **7,2** |
| ... | ... | ... | ... |
| 24 | 11 | C2 | **7,8** |

## Exemplo 9.4: Perguntas a serem respondidas pela análise estatística.

- Existe diferença real (significativa) entre os 3 tipos de rede?

- Qual é a estimativa do tempo de resposta para cada tipo de rede?

# Exemplo 9.4: Dados do experimento

| | Tipo de rede | | |
|---|---|---|---|
| Replicação | **C1** | **C2** | **C3** |
| 1 | 7,2 | 7,8 | 6,3 |
| 2 | 9,3 | 8,2 | 6,0 |
| 3 | 8,7 | 7,1 | 5,3 |
| 4 | 8,9 | 8,6 | 5,1 |
| 5 | 7,6 | 8,7 | 6,2 |
| 6 | 7,2 | 8,2 | 5,2 |
| 7 | 8,8 | 7,1 | 7,2 |
| 8 | 8,0 | 7,8 | 6,8 |
| Soma | 65,7 | 63,5 | 48,1 |
| **Média** | **8,21** | **7,94** | **6,01** |

# Modelo da ANOVA
# $g$ = 3 grupos

| | tratamento | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | |
| | $y_{11}$ | $y_{21}$ | $y_{31}$ | |
| | $y_{12}$ | $y_{22}$ | $y_{32}$ | |
| | ... | ... | ... | Média global: |
| | $y_{1n}$ | $y_{2n}$ | $y_{3n}$ | |
| Média: | $\bar{y}_{1.}$ | $\bar{y}_{2.}$ | $\bar{y}_{3.}$ | $\bar{y}_{..}$ |

$$y_{ij} = \mu + \tau_i + e_{ij}$$

i = 1, 2, 3
j = 1, 2, ..., n

observação

média global

efeito do tratamento i

erro aleatório

$$\mu_i = \mu + \tau_i$$ = média do fator i

# Hipóteses

$H_0: \tau_1 = \tau_2 = ... = \tau_g = 0$ ou $\mu_1 = \mu_2 = ... = \mu_g$

$H_1: \tau_i \neq 0$ ou $\mu_i \neq \mu_j$

para algum i para algum par (i, j)

As observações

Sob $H_1$:

$$y_{ij} = \mu + \tau_i + e_{ij}$$

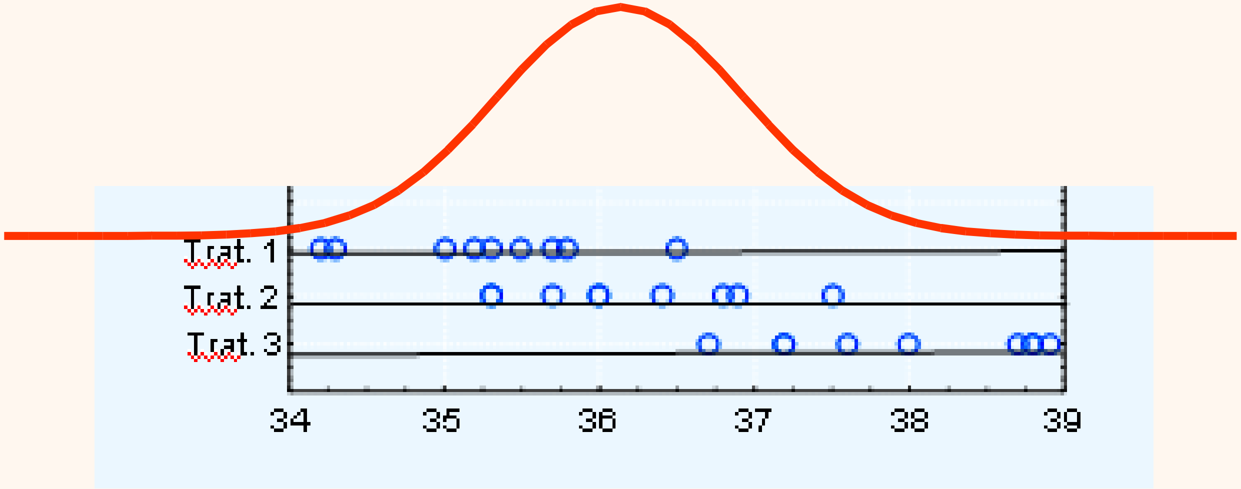
Sob $H_0$:

$$y_{ij} = \mu + e_{ij}$$

# Hipóteses e modelo subjacente

Sob $H_0$: $\tau_1 = \tau_2 = \ldots = \tau_g = 0$

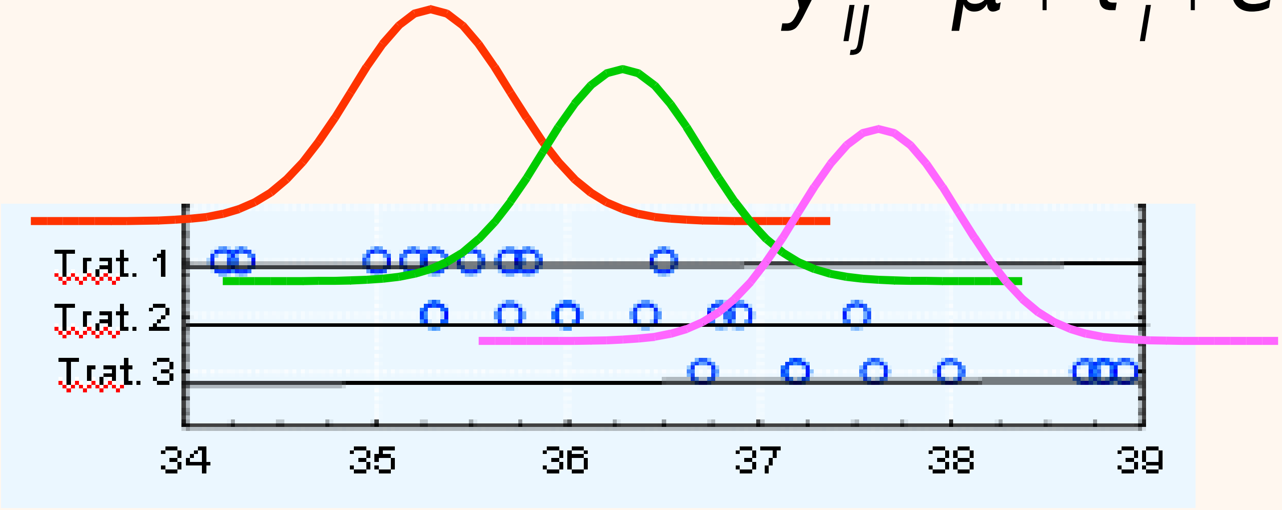
$$y_{ij} = \mu + \tau_i + e_{ij} \longrightarrow y_{ij} = \mu + e_{ij}$$



# Hipóteses e modelo subjacente

Sob $H_1$: $\tau_i \neq 0$ para algum i

$$y_{ij} = \mu + \tau_i + e_{ij}$$



# Análise de variância (ANOVA) com um fator

| Replicação | Tratamento | | | |
|---|---|---|---|---|
| | **1** | **2** | **...** | ***g*** |
| 1 | $y_{11}$ | $y_{21}$ | ... | $y_{g1}$ |
| 2 | $y_{12}$ | $y_{22}$ | ... | $y_{g2}$ |
| ... | ... | ... | ... | ... |
| $n$ | $y_{1n}$ | $y_{2n}$ | ... | $y_{gn}$ |
| **Soma** | $\mathbf{y_{1.}}$ | $\mathbf{y_{2.}}$ | **...** | $\mathbf{y_{g.}}$ | $y_{..} = \sum_i y_{i.}$ |
| **Média** | $\bar{y}_{1.}$ | $\bar{y}_{2.}$ | **...** | $\bar{y}_{g.}$ | $\bar{y}_{..} = \frac{1}{g}\sum_i y_{i.}$ |

Soma de quadrados total:

$$SQ_{Tot} = \sum_{i=1}^{g} \sum_{j=1}^{n} \left( y_{ij} - \bar{y}_{..} \right)^2$$

Graus de liberdade:

$$gl = N - 1$$

onde: $N = ng$

# Análise de variância (ANOVA) com um fator

| Replicação | Tratamento 1 | 2 | ... | $g$ | |
|---|---|---|---|---|---|
| 1 | $y_{11}$ | $y_{21}$ | ... | $y_{g1}$ | |
| 2 | $y_{12}$ | $y_{22}$ | ... | $y_{g2}$ | |
| ... | ... | ... | ... | ... | |
| $n$ | $y_{1n}$ | $y_{2n}$ | ... | $y_{gn}$ | |
| **Soma** | $y_{1.}$ | $y_{2.}$ | ... | $y_{g.}$ | $y_{..}=\sum_i y_{i.}$ |
| **Média** | $\bar{y}_{1.}$ | $\bar{y}_{2.}$ | ... | $\bar{y}_{g.}$ | $\bar{y}_{..}=\frac{1}{g}\sum_i y_{i.}$ |

Soma de quadrados dos tratamentos:

$$SQ_{Trat}=\sum_{i=1}^{g}\sum_{j=1}^{n}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2=n\sum_{i=1}^{g}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2$$

Graus de liberdade:

$$gl = g - 1$$

# Análise de variância (ANOVA) com um fator

| Replicação | Tratamento 1 | 2 | ... | g | |
|---|---|---|---|---|---|
| 1 | $y_{11}$ | $y_{21}$ | ... | $y_{g1}$ | |
| 2 | $y_{12}$ | $y_{22}$ | ... | $y_{g2}$ | |
| ... | ... | ... | ... | ... | |
| $n$ | $y_{1n}$ | $y_{2n}$ | ... | $y_{gn}$ | |
| **Soma** | $\boldsymbol{y_{1.}}$ | $\boldsymbol{y_{2.}}$ | ... | $\boldsymbol{y_{g.}}$ | $y_{..} = \sum_i y_{i.}$ |
| **Média** | $\bar{y}_{1.}$ | $\bar{y}_{2.}$ | ... | $\bar{y}_{g.}$ | $\bar{y}_{..} = \frac{1}{g} \sum_i y_{i.}$ |

Soma de quadrados do erro:

$$SQ_{Erro} = \sum_{i=1}^{g} \sum_{j=1}^{n} \left( y_{ij} - \bar{y}_{i.} \right)^2$$

Graus de liberdade:

$$gl = N - g$$

# Análise de variância (ANOVA) com um fator

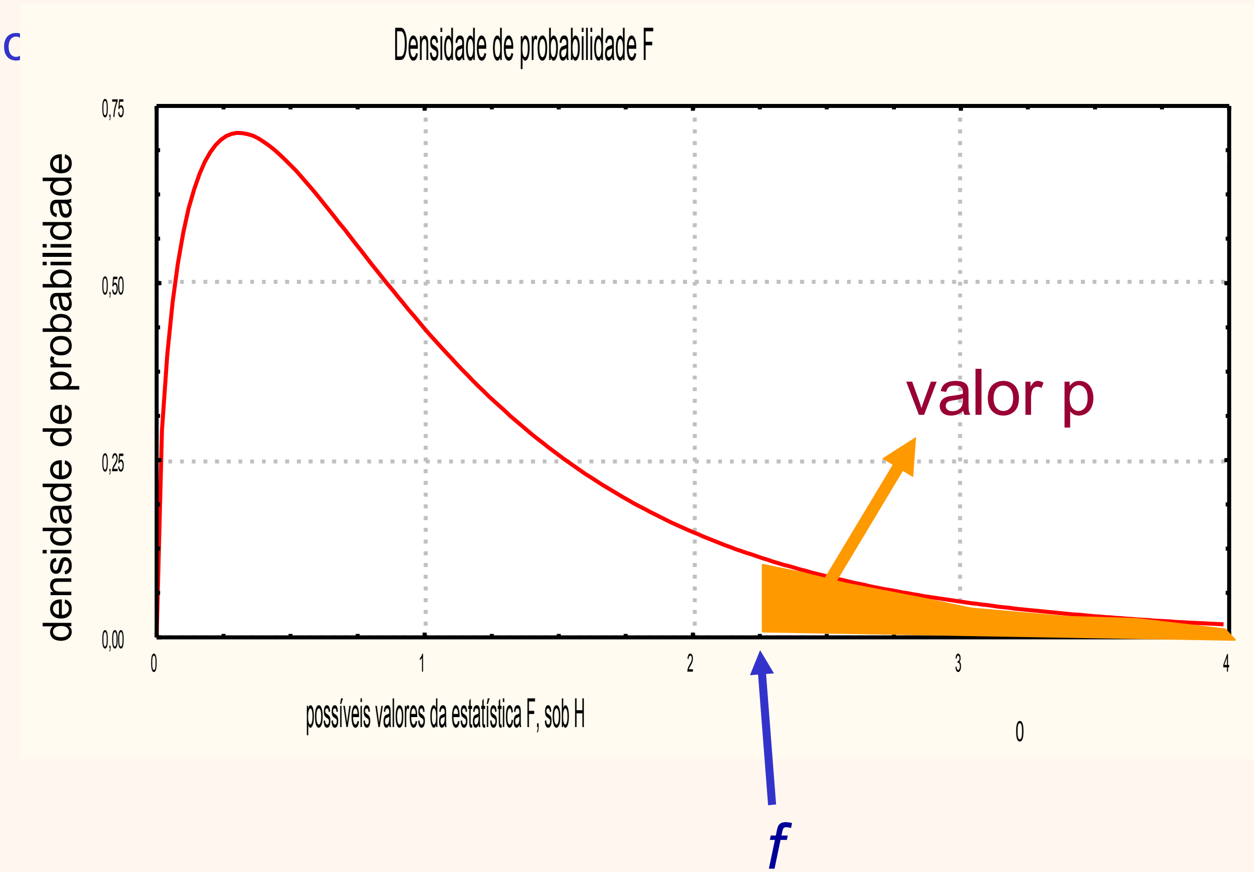
Fórmulas equivalentes às anteriores

| Fonte de variação | Somas de quadrados | gl | Quadrados médios | Razão f |
|---|---|---|---|---|
| Entre tratamentos | $SQ_{Trat} = \sum_{i=1}^{g} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}$ | $g-1$ | $QM_{Trat} = \frac{SQ_{Trat}}{gl_{Trat}}$ | $f = \frac{QM_{Trat}}{QM_{Erro}}$ |
| Dentro trat. (Erro) | $SQ_{Erro} = SQ_{Tot} - SQ_{Trat}$ | $N-g$ | $QM_{Erro} = \frac{SQ_{Erro}}{gl_{Erro}}$ | |
| Total | $SQ_{Tot} = \sum_{i=1}^{g}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{N}$ | $N-1$ | | |

Estatística do teste (possíveis valores da razão *f*): $f = \frac{QM_{Trat}}{QM_{Erro}}$

# Teste F

- Se $H_0$: $\tau_1 = \tau_2 = \ldots = \tau_g = 0$ for verdadeira e considerando as suposições anteriormente enunciadas, a estatística $\boldsymbol{f}$ tem distribuição F com (g - 1) graus de liberdade no numerador e (N - g) graus de liberdade no deno



# Regra de decisão. Abordagem valor **p**

$\alpha$ = nível de significância

(probab. tolerável de se rejeitar $H_o$ quando esta for verdadeira)

Usual: $\alpha = 0{,}05 = 5\%$

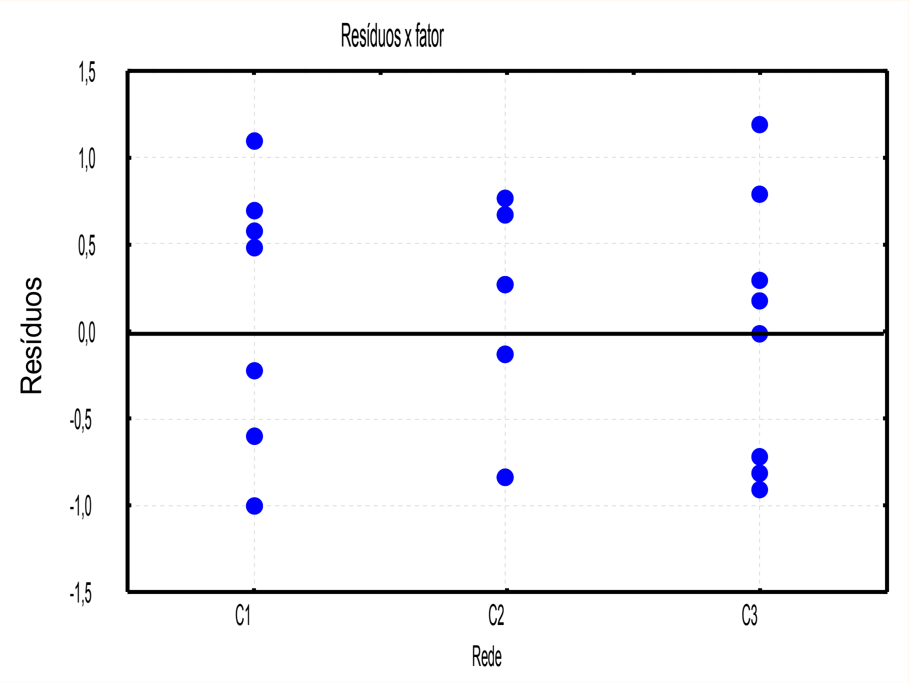
- **p** $\leq \alpha$ ⇒ rejeita $H_0$ (prova-se estatisticamente $H_1$)
- **p** $> \alpha$ ⇒ aceita $H_0$ (os dados não mostram evidência para afirmar $H_1$)

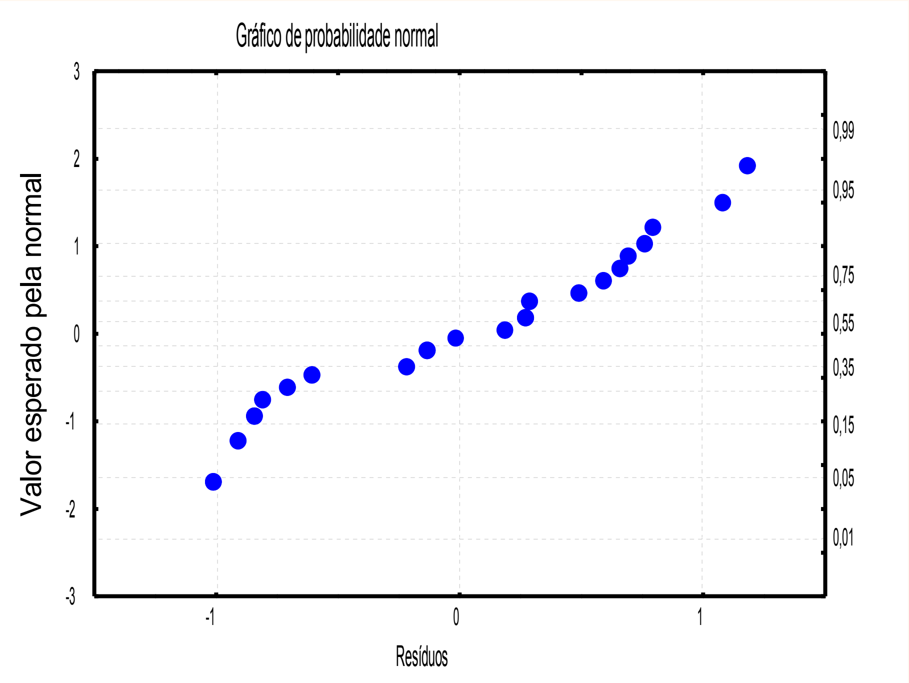
# Teste F

- Ver no livro como usar a Tabela F e como fazer o teste pela abordagem clássica.

# Análise dos resíduos

Avaliação das suposições da ANOVA através de gráficos dos resíduos:

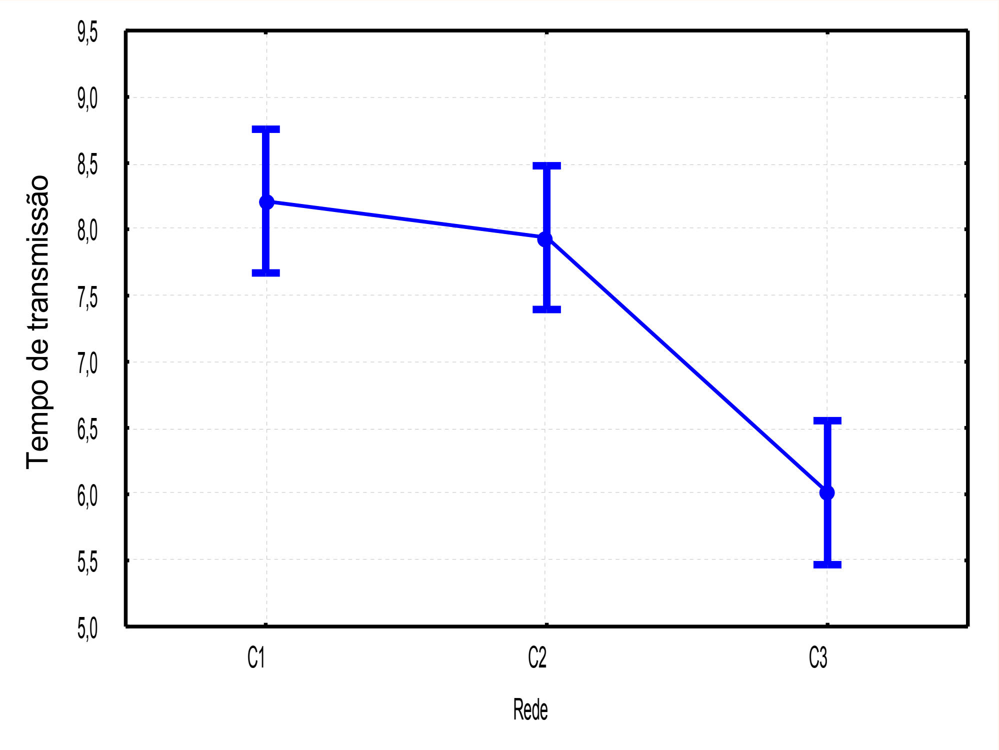


# Estimação das médias

Intervalo de confiança para o valor esperado da resposta sob o *i*-ésimo tratamento (nível de conf. γ):

$$IC(\mu_i, \gamma) = \bar{y}_{i.} \pm t_\gamma \sqrt{\frac{QM_{erro}}{n}}$$



# Teste F para amostras em blocos

Notação para os dados:

| Bloco | Tratamento 1 | 2 | ... | $g$ | Soma |
|---|---|---|---|---|---|
| 1 | $y_{11}$ | $y_{21}$ | ... | $y_{g1}$ | $y_{.1}$ |
| 2 | $y_{12}$ | $y_{22}$ | ... | $y_{g2}$ | $y_{.2}$ |
| ... | ... | ... | ... | ... | |
| $h$ | $y_{1h}$ | $y_{2h}$ | ... | $y_{gh}$ | $y_{.h}$ |
| Soma | $y_{1.}$ | $y_{2.}$ | ... | $y_{g.}$ | $y_{..}=\sum_i y_{i.}=\sum_j y_{.j}$ |

# Modelo para os dados

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

- $\mu$ é a média global da resposta;
- $\tau_i$ é o efeito do $i$-ésimo tratamento;
- $\beta_j$ é o efeito do $j$-ésimo bloco; e
- $\varepsilon_{ij}$ é o efeito aleatório ($i = 1, 2, ..., g$; $j = 1, 2, ..., h$).

# Teste F para amostras em blocos: quadro da ANOVA

| *Fonte de variação* | *Somas de quadrados* | *gl* | *Quadrados médios* | *Razão f* |
|---|---|---|---|---|
| *Entre tratamentos* | $SQ_{Trat} = \sum_{i=1}^{g} \frac{y_{i.}^2}{h} - \frac{y_{..}^2}{N}$ | $g-1$ | $QM_{Trat} = \frac{SQ_{Trat}}{gl_{Trat}}$ | $f = \frac{QM_{Trat}}{QM_{Erro}}$ |
| *Entre blocos* | $SQ_{Blocos} = \sum_{j=1}^{h} \frac{y_{.j}^2}{g} - \frac{y_{..}^2}{N}$ | $h-1$ | $QM_{Bloco} = \frac{SQ_{Bloco}}{gl_{Bloco}}$ | |
| *Dentro (Erro)* | $SQ_{Erro} = SQ_{Tot} - SQ_{Trat} - SQ_{Bloco}$ | $(g-1)(h-1)$ | $QM_{Erro} = \frac{SQ_{Erro}}{gl_{Erro}}$ | |
| *Total* | $SQ_{Tot} = \sum_{i=1}^{g}\sum_{j=1}^{h} y_{ij}^2 - \frac{y_{..}^2}{N}$ | $N-1$ | | |

# Exemplo 9.5

- Seja o problema de comparar 3 algoritmos de busca em um banco de dados. Realiza-se um experimento com 6 buscas experimentais, sendo que em cada uma é sorteado um número aleatório que indica o registro do banco de dados a ser localizado. Em cada um dos 6 processos de busca, são usados separadamente os três algoritmos em estudo, mas sob as mesmas condições, em termos dos fatores controláveis. São anotados os tempos de resposta ao usuário.
- Hipóteses:
  - $H_0$: em média, os três algoritmos *são igualmente* rápidos; e
  - $H_1$: em média, os três algoritmos *não são igualmente* rápidos

# Exemplo 9.5

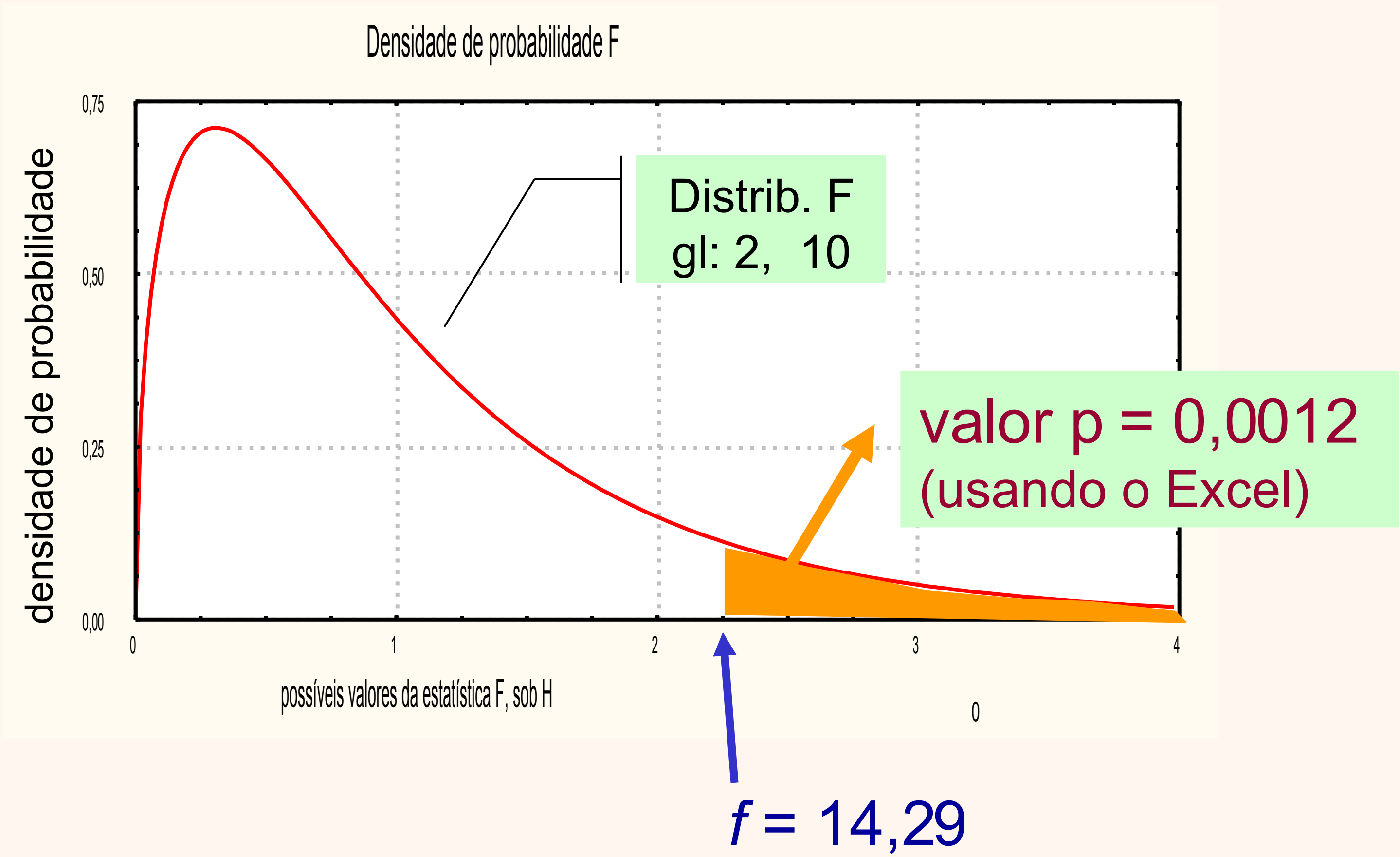
| Ensaio (bloco) | Algoritmo de busca | | |
|---|---|---|---|
| | A1 | A2 | A3 |
| 1 | 8,3 | 8,1 | 9,2 |
| 2 | 9,4 | 8,9 | 9,8 |
| 3 | 9,1 | 9,3 | 9,9 |
| 4 | 9,9 | 9,6 | 10,3 |
| 5 | 8,2 | 8,1 | 8,9 |
| 6 | 10,9 | 11,2 | 13,1 |
| Soma | 55,8 | 55,2 | 61,2 |
| Média | 9,3 | 9,2 | 10,2 |

# Exemplo 9.5

| *Fonte da variação* | *SQ* | *gl* | *QM* | *f* |
|---|---|---|---|---|
| Algoritmos | 3,64 | 2 | 1,82 | 14,29 |
| Blocos | 21,95 | 5 | 4,39 | |
| Erro | 1,27 | 10 | 0,13 | |
| Total | 26,86 | 17 | | |

Qual é a conclusão?

# Exemplo 9.5




- Conclusão?

# ANOVA em projetos fatoriais com 2 fatores

- Ver comentários sobre esse projeto e as hipóteses no livro.

# ANOVA em projetos fatoriais com 2 fatores

Notação para os dados:

| Fator B | Fator A: 1 | 2 | ... | $g$ | Soma |
|---|---|---|---|---|---|
| 1 | $y_{111}, \ldots, y_{11n}$ | $y_{211}, \ldots, y_{21n}$ | ... | $y_{g11}, \ldots, y_{g1n}$ | $y_{.1.}$ |
| 2 | $y_{121}, \ldots, y_{12n}$ | $y_{221}, \ldots, y_{22n}$ | ... | $y_{g21}, \ldots, y_{g2n}$ | $y_{.2.}$ |
| ... | ... | ... | ... | ... | |
| $h$ | $y_{1h1}, \ldots, y_{1hn}$ | $y_{2h1}, \ldots, y_{2hn}$ | ... | $y_{gh1}, \ldots, y_{ghn}$ | $y_{.h.}$ |
| Soma | $y_{1..}$ | $y_{2..}$ | ... | $y_{g..}$ | $y_{...}=\sum_i y_{i..}=\sum_j y_{.j.}$ |

# ANOVA em projetos fatoriais com 2 fatores

| Fonte de variação | Somas de quadrados | gl | Quadrados médios | Razão f |
|---|---|---|---|---|
| Fator A | $SQ_A = \sum_{i=1}^{g} \frac{y_{i..}^2}{hn} - \frac{y_{...}^2}{N}$ | $g-1$ | $QM_A = \frac{SQ_A}{gl_A}$ | $f = \frac{QM_A}{QM_{Erro}}$ |
| Fator B | $SQ_B = \sum_{j=1}^{h} \frac{y_{.j.}^2}{gn} - \frac{y_{...}^2}{N}$ | $h-1$ | $QM_B = \frac{SQ_B}{gl_B}$ | $f = \frac{QM_B}{QM_{Erro}}$ |
| Interação A*B | $SQ_{AB} = SQ_{Subtot} - SQ_A - SQ_B$ | $(g-1)(h-1)$ | $QM_{AB} = \frac{SQ_{AB}}{gl_{AB}}$ | $f = \frac{QM_{AB}}{QM_{Erro}}$ |
| Erro | $SQ_{Erro} = SQ_{Tot} - SQ_{Subtot}$ | $hg(n-1)$ | $QM_{Erro} = \frac{SQ_{Erro}}{gl_{Erro}}$ | |
| Total | $SQ_{Tot} = \sum_{i=1}^{g}\sum_{j=1}^{h}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{N}$ | $N-1$ | | |

# Exemplo 9.6

- Considere o problema de comparar 3 topologias de rede de computadores (C1, C2 e C3) e 2 protocolos (L1 e L2), em termos do tempo de resposta ao usuário. Realizou-se um experimento com 4 replicações em cada combinação de topologia e protocolo. Deseja-se verificar se há diferenças entre as topologias, entre os protocolos e eventual interação entre topologia e protocolo. Então, quer-se testar as seguintes hipóteses nulas:
  - $H_0^{(A)}$: os tempos esperados de resposta *são iguais* para as três topologias;
  - $H_0^{(B)}$: os tempos esperados de resposta *são iguais* para os dois protocolos;
  - $H_0^{(AB)}$: a mudança de protocolo *não altera* as diferenças médias do tempo de resposta nas três topologias (ausência de interação).

# Exemplo 9.6

Dados:

| | Topologia | | | | |
|---|---|---|---|---|---|
| Protocolo | C1 | C2 | C3 | Soma | Média |
| L1 | 6,2 | 5,9 | 5,9 | | |
| | 7,6 | 8,4 | 6,2 | | |
| | 7,2 | 7,1 | 5,2 | | |
| | 8,8 | 7,1 | 7,2 | $y_{1.} = 82,8$ | 6,90 |
| L2 | 9,0 | 7,1 | 6,2 | | |
| | 8,9 | 8,6 | 6,1 | | |
| | 9,4 | 9,1 | 8,9 | | |
| | 8,0 | 7,8 | 6,8 | $y_{2.} = 95,9$ | 7,99 |
| Soma | $y_{1..} = 65,1$ | $y_{2..} = 61,1$ | $y_{3..} = 52,5$ | $y_{...} = 178,7$ | |
| Média | 8,21 | 7,94 | 6,01 | | 7,45 |

# Exemplo 9.6

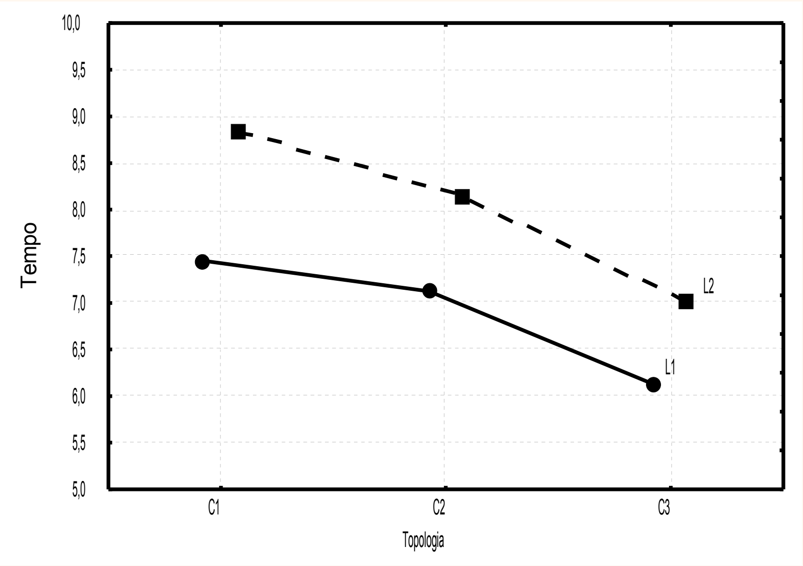
ANOVA:

| Fonte de variação | *SQ* | *gl* | *QM* | *f* |
|---|---|---|---|---|
| Topologia | 10,36 | 2 | 5,18 | 5,44 |
| Protocolo | 7,15 | 1 | 7,15 | 7,51 |
| Interação | 0,26 | 2 | 0,13 | 0,14 |
| Erro | 17,14 | 18 | 0,95 | |
| Total | 34,92 | 23 | | |

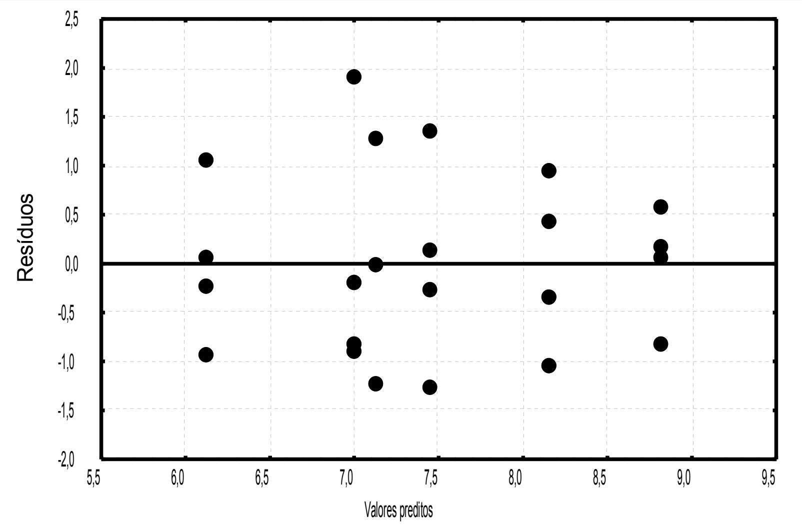
Quais as conclusões?

# Exemplo 9.6

**(a)** *Perfil das médias*




**(b)** *Análise dos resíduos*




Quais as conclusões?

# ANOVA para projetos fatoriais $2^k$ e $2^{k-p}$

- Ver no livro as técnicas e exemplos.