MAE 327
Planejamento e Pesquisa II

Profa. Júlia Maria Pavan Soler

pavan@ime.usp.br

IME/USP

Planejamento e Pesquisa II

- ✓ **Princípios do Planejamento de Experimentos, Estrutura de Dados, Exemplos de Planejamentos Experimentais**

- ✓ **Planejamento I: Delineamentos Completamente Aleatorizados (DCA) com Um ou mais Fatores Cruzados, Balanceados e não Balanceados**

Passos da Análise de Dados

Conhecer o Planejamento Experimental e o Objetivo do Estudo
Entender a Estrutura de Dados
Adotar um modelo Estrutural e Distribucional
Análises de Diagnóstico do Modelo (Análises alternativas?)
Obter a Tabela de ANOVA
Conclusão dos Testes Globais da ANOVA
Comparações Múltiplas (Correções para Múltiplos Testes)
Conclusão e Interpretação dos resultados

Dados de *Clorofila a*

Dados: Medidas de clorofila a dissolvida na água

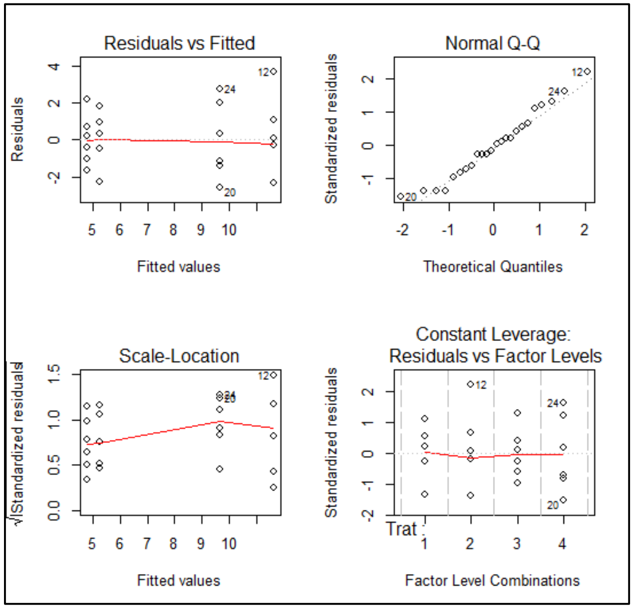
T1	T2	T3	T4
6,2	12,7	7,0	8,3
4,8	11,3	4,4	7,1
3,0	9,3	3,8	11,7
5,6	9,5	5,0	10,0
7,1	11,7	5,5	8,5
4,8	15,3	3,2	12,4

$$y_{ij} = \mu_j + e_{ij} = \mu_1 + \tau_j + e_{ij};$$

$$j = 2,3,4; \;\; i = 1,\dots,6$$

$$e_{ij} \overset{iid}{\sim} N\left(0;\sigma^2\right);$$

$$y_{ij} \sim N\left(\mu_j;\sigma^2\right)$$



$$H_0 : \mu_j = \mu; \quad H_1 : \text{Existe pelo menos uma diferença}$$

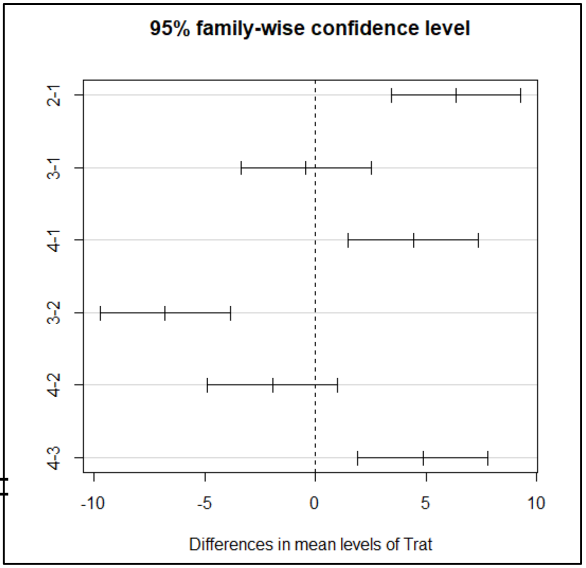
Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trat	3	199.937	66.646	20.16	2.924e-06
Residuals	20	66.117	3.306		

Coefficients

(Intercept)	Trat2	Trat3	Trat4
5.250	6.350	-0.433	4.417

$$\left(\mu_2 = \mu_4\right) > \left(\mu_1 = \mu_3\right) \Leftarrow$$



Dados de *Clorofila a*

Dados: Medidas de clorofila a dissolvida na água

T1	T2	T3	T4
6,2	12,7	7,0	8,3
4,8	11,3	4,4	7,1
3,0	9,3	3,8	11,7
5,6	9,5	5,0	10,0
7,1	11,7	5,5	8,5
4,8	15,3	3,2	12,4

1 Fator em 4 níveis ⇒ Fatorial 2x2

T1	T2	T3	T4	
30%	30%	100%	100%	Luminosidade
SN	N	SN	N	Nutrientes

$$y_{ijk} = \mu_{jk} + e_{ijk} = \mu_{11} + \tau_j + \beta_k + \gamma_{jk} + e_{ijk};$$

$$e_{ij} \overset{iid}{\sim} N\left(0;\sigma^2\right); \quad y_{ij} \sim N\left(\mu_{jk};\sigma^2\right); \quad j = k = 2; \; i = 1,\ldots,6$$

y_{ijk} : resposta da unidade *i* submetida aos níveis *j* de Luminosidade e *k* de Nutrientes

μ_{11} : valor esperado da resposta para luminosidade 30% e Sem Nutrientes (SN)

τ_j : desvio em relação a μ_{11} devido ao efeito (principal) de luminosidade em 100%

β_k : desvio em relação a μ_{11} devido ao efeito (principal) de nutrientes

γ_{jk} : efeito de interação entre os fatores. É o desvio do efeito aditivo dos fatores

e_{ijk} : efeito aleatório, suposto normal, independente e homocedástico.

Dados de *Clorofila a*

Dados: Medidas de clorofila a dissolvida na água

T1	T2	T3	T4
6,2	12,7	7,0	8,3
4,8	11,3	4,4	7,1
3,0	9,3	3,8	11,7
5,6	9,5	5,0	10,0
7,1	11,7	5,5	8,5
4,8	15,3	3,2	12,4

Tabela ANOVA: Modelo de 1 Fator em 4 níveis

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trat	**3**	**199.937**	66.646	20.16	2.924e-06
Residuals	20	66.117	3.306		

Coefficients (codificação: Trat=1,2,3,4)

(Intercept)	Trat2	Trat3	Trat4
5.250	6.350	-0.433	4.417

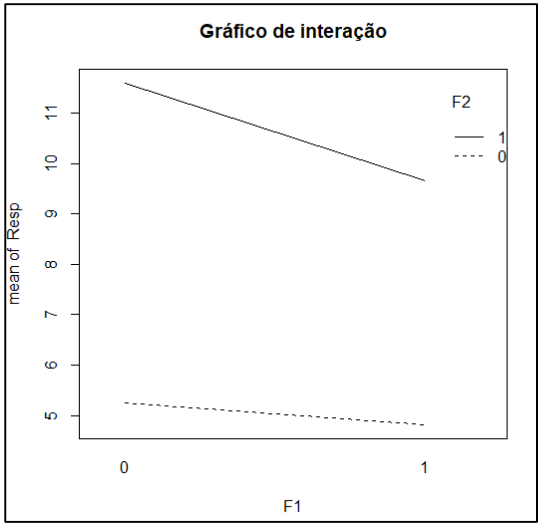


Tabela Anova: Modelo Fatorial 2x2 (Interação0

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
F1(Lumi)	**1**	**8.402**	8.402	2.5415	0.1266
F2(Nutri)	**1**	**188.160**	188.160	56.9176	2.849e-07
F1:F2	**1**	**3.375**	3.375	1.0209	0.3244
Residuals	20	66.117	3.306		

Coefficients (codificação:F1(0=30% 1=100%) F2(0=SN 1=N)

(Intercept)	F1	F2	F1:F2
5.250	-0.433	6.350	-1.50

Modelos reduzidos (sem interação → só F2)

Delineamento Aleatorizado em Blocos Completos

	Tratamentos			
Bloco	T_1	T_2	. . .	T_J
B_1	Y_{11}	Y_{12}	. . .	Y_{1J}
B_2	Y_{21}	Y_{22}	. . .	Y_{2J}
	. . .	. . .	Y_{ij}	. . .
B_r	Y_{r1}	Y_{r2}	. . .	Y_{rJ}

"aleatorização restrita" dentro dos blocos

Dentro de cada bloco, há J unidades experimentais, às quais todos os tratamentos são aleatoriamente atribuídos

$n=rJ$ ← **r replicações em cada tratamento**

Exemplo - Blocos

Blocos

- **Homogeneidade das u.e. dentro de blocos**
- **Fator de Controle $\Rightarrow$ não há interesse na variabilidade "Entre" blocos**

- Canteiros
- Pacientes com a mesma gravidade da doença
- Animais de mesma linhagem (mesma mãe)
- 4 pneus de um carro
- Braços ou olhos (direito e esquerdo) de voluntários
- Unidades padronizadas segundo peso, idade, sexo
- Técnico, fornecedor, ...

Exemplo

Dados: Medidas de clorofila *a* em amostras de água filtradas segundo quatro tratamentos

	Tratamento			
Bloco	**T1**	**T2**	**T3**	**T4**
B1	**6,2**	**12,7**	**7,0**	**8,3**
B2	**4,8**	**11,3**	**4,4**	**7,1**
B3	**3,0**	**9,3**	**3,8**	**11,7**
B4	**5,6**	**9,5**	**5,0**	**10,0**
B5	**7,1**	**11,7**	**5,5**	**8,5**
B6	**4,8**	**15,3**	**3,2**	**12,4**

Estrutura de Blocos (hipotético): coluna de água de um rio.

Exemplo

Dados: Crescimento de plantas (cm) segundo a variedade

Canteiro	Var1	Var2	Var3	Var4
1	19,8	21,9	16,4	14,7
2	16,7	19,8	15,4	13,5
3	17,7	21,0	14,8	12,8
4	18,2	21,4	15,6	13,7
5	20,3	22,1	16,4	14,6
6	15,5	20,8	14,6	12,9

Delineamento Aleatorizado em Blocos Completos

Duas Entradas

		Tratamentos		
Bloco	T_1	T_2	...	T_J
B_1	Y_{11}	Y_{12}	...	Y_{1J}
B_2	Y_{21}	Y_{22}	...	Y_{2J}
	...	...	Y_{ij}	...
B_r	Y_{r1}	Y_{r2}	...	Y_{rJ}

2 Fatores

- **Tratamento: fator de interesse**
- **Bloco: fator de controle**

Variável Resposta

Note que em cada combinação dos níveis dos fatores (Tratamento e Bloco) NÃO há replicas. Qual é a implicação disso na modelagem dos dados?

Delineamento Aleatorizado em Blocos Completos

Estrutura dos Dados

- **Fatores**
  - **Tratamento:** ***fator de interesse***
  - **Bloco:** ***fator de controle***
- **Aleatorização Restrita (Dentro dos Blocos)**
- **Variável Resposta**

Vantagem da Blocagem

eliminar o efeito de uma fonte de variação conhecida (bloco)

⬇

redução do "efeito" residual (é esperado ter um ganho em precisão)

Suposições

$N(\mu_1;\sigma^2)$ $N(\mu_2;\sigma^2)$ ... $N(\mu_J;\sigma^2)$ **População**

	Tratamentos				
Bloco	**T1**	**T2**	. . .	**TJ**	**médias**
B1	**Y11**	**Y12**	. . .	**Y1J**	$\bar{Y}_{1.}$
B2	**Y21**	**Y22**	. . .	**Y2J**	$\bar{Y}_{2.}$
	. . .	. . .	**Yij**	. . .	
Br	**Yr1**	**Yr2**	. . .	**YrJ**	$\bar{Y}_{r.}$
médias	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$		$\bar{Y}_{.J}$	$\bar{Y}$

Amostra

- ✓ **Normalidade**
- ✓ **Variância constante**
- ✓ **Independência**

entre todas as u.a., mesmo entre as do mesmo bloco.

Modelo Teórico - Aditividade

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

$$= \underbrace{\mu_j + \beta_i}_{\text{componente fixo}} + e_{ij}$$

$$e_{ij} \overset{iid}{\sim} N\left(0; \sigma^2\right)$$

componente aleatório

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

Efeito do bloco

efeito do tratamento

$$\sum_{j=1}^{J} \tau_j = \sum_{i=1}^{r} \beta_i = 0$$

Suposição: Modelo Aditivo (não há efeito de interação entre os fatores)

Delineamento Completamente Aleatorizado
Delineamento Aleatorizado em Blocos Completos

DCA

T_1	T_2	…	T_J	
-	-		-	
-	-		-	
-	-		-	
-	-		-	
-	-		-	
n_1	n_2	…	n_J	n

Aleatorização irrestrita das *n* unidades experimentais aos *a* tratamentos

- Normalidade, independência e variância constante

DABC

	T_1	T_2	…	T_J
B1	-	-		-
B2	-	-		-
…	-	-		-
Br	-	-		-
	r	r	…	r

Aleatorização restrita das *a* unidades experimentais dentro de cada um dos *r* blocos

- Normalidade, independência e variância constante

- Aditividade

Modelo Teórico / Estimadores

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

identidade útil na obtenção das estatísticas

Resíduo é o efeito de interação entre os fatores

$$\underbrace{y_{ij}}_{\text{observado}} = \underbrace{\bar{y} + \underbrace{(\bar{y}_{.j} - \bar{y})}_{\hat{\tau}_j} + \underbrace{(\bar{y}_{i.} - \bar{y})}_{\hat{\beta}_i}}_{\hat{y}_{ij}\ \text{predito}} + \underbrace{(y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y})}_{\hat{e}_{ij}}$$

Modelo Teórico / Estimadores

$$y_{ij} = \mu_{ij} + e_{ij} = \mu + \tau_j + \beta_i + e_{ij}$$

Sob o modelo aditivo tem-se: $\mu_{ij} = \mu + \tau_j + \beta_i$

Desvio da aditividade: $\mu_{ij} - \left(\mu + \tau_j + \beta_i\right) = e_{ij}$

$$y_{ij} = \bar{y} + (\bar{y}_{.j} - \bar{y}) + \boxed{(\bar{y}_{i.} - \bar{y}) + (y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y})}$$

$\hat{\beta}_i + \hat{e}_{ij}$: **se reduz ao resíduo do DCA, somente com o fator Tratamento**

$$\hat{e}_{ij} = y_{ij} - \left[\bar{y}_{.j} + \bar{y}_{.i} - \bar{y}\right]:$$

Estima o desvio do modelo aditivo $\Rightarrow$ é o efeito de interação entre os fatores tratamento e bloco (na ausência de réplicas para cada combinação dos níveis dos dois fatores)

Modelo Teórico / Somas de Quadrados

$$y_{ij} = \mu + \tau_j + \beta_i + e_{ij}; \quad \sum_{j=1}^{J} \tau_j = \sum_{i=1}^{r} \beta_i = 0; \quad e_{ij} \overset{iid}{\sim} N\left(0; \sigma^2\right)$$

Considere a seguinte identidade útil na obtenção das estatísticas da tabela de ANOVA

$$y_{ij} = \bar{y} + (\bar{y}_{.j} - \bar{y}) + (\bar{y}_{i.} - \bar{y}) + (y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y})$$

Elevando ao quadrado ambos os lados da igualdade e aplicando os somatórios é possível mostrar a correspondente partição da SQTotal:

$$\sum_{ij}\left(y_{ij} - \bar{y}\right)^2 = \sum_{j} r\left(\bar{y}_{.j} - \bar{y}\right)^2 + \sum_{i} J\left(\bar{y}_{i.} - \bar{y}\right)^2 + \sum_{ij}\left(y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}\right)^2$$

SQTotal **SQTratamento** **SQBloco** **SQResíduo**

Hipótese de Interesse

$$Y_{ij} \sim N(\mu_j + \beta_i ; \sigma^2); \quad \mu_j = \mu + \tau_j, \sum_{j=1}^{J} \tau_j = 0$$

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \ldots = \mu_J = \mu \quad \Leftrightarrow \quad H : \tau_1 = \tau_2 = \ldots = \tau_J = 0 \\ H_1\text{: existe pelo menos uma diferença entre as médias} \end{cases}$$

Existe evidência amostral de diferenças entre as médias de tratamento ?

As diferenças ENTRE médias de tratamento são maiores que a variação DENTRO dos tratamentos (corrigida/ajustada pelo efeito de bloco) ?

Hipótese de Interesse

$$Y_{ij} \sim N(\mu_j + \beta_i ; \sigma^2)$$

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \ldots = \mu_J = \mu \\ H_1\text{: existe pelo menos uma diferença} \end{cases}$$

Em geral, não há interesse em testar o efeito do fator Bloco (apesar de, teoricamente, ser possível testar!)

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_r = \beta$$

Qual é a estatística deste teste ?

Porém, note que não é possível testar o efeito de interação entre os fatores tratamento e bloco! Este efeito define o resíduo do modelo.

Tabela de ANOVA

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_J = \mu$$

(r-1)+(J-1)(r-1) = rJ-J

g.l. da interação

F.V.	g l	SQ	QM	F	p
TRAT	J-1	$\sum r\,(\bar{y}_{.j} - \bar{y})^2$	SQTrat/(J-1)	$\frac{\text{QMTrat}}{\text{QMRes}}$	
BLOCO	r-1	$\sum J\,(\bar{y}_{i.} - \bar{y})^2$			
RESÍDUO	(J-1)(r-1)	$\sum_{ij} (y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y})^2$	SQRes/(J-1)(r-1)		
TOTAL	rJ-1	$\sum_{ij} (y_{ij} - \bar{y})^2$			

$$F = \frac{QMTrat}{QM\,Res} \overset{H_0}{\sim} F_{(J-1),(J-1)(r-1)}$$

Como tomar decisão sobre H_0 ?

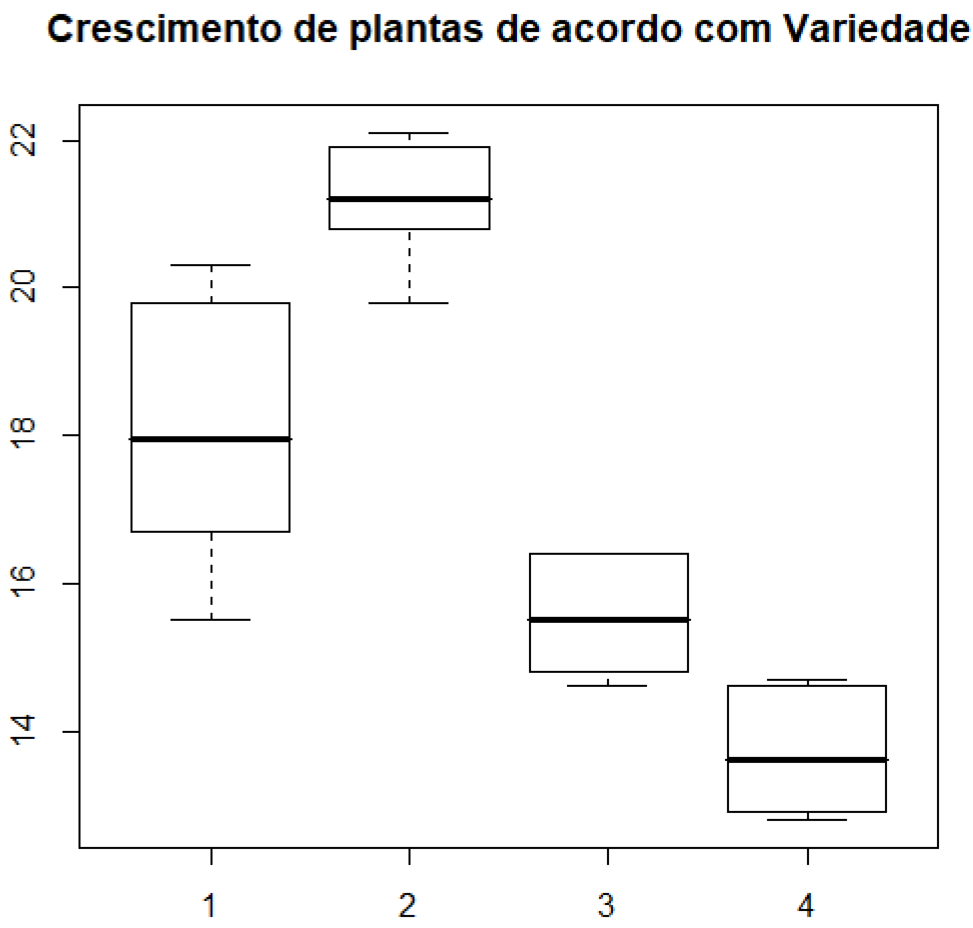
Em geral, não há interesse em testar o efeito do fator Bloco e, além das premissas clássicas, a validade do modelo aditivo deve ser investigada.

Exemplo

<u>Dados</u>: Crescimento de plantas (cm) segundo a variedade

Canteiro	Var1	Var2	Var3	Var4
1	19,8	21,9	16,4	14,7
2	16,7	19,8	15,4	13,5
3	17,7	21,0	14,8	12,8
4	18,2	21,4	15,6	13,7
5	20,3	22,1	16,4	14,6
6	15,5	20,8	14,6	12,9

Modelo ANOVA - DABC



	Var1	Var2	Var3	Var4
Média	18.03	21.17	15.53	13.70
Desvio Padrão	1.82	0.84	0.77	0.81

Hipóteses ?

Suposições ?

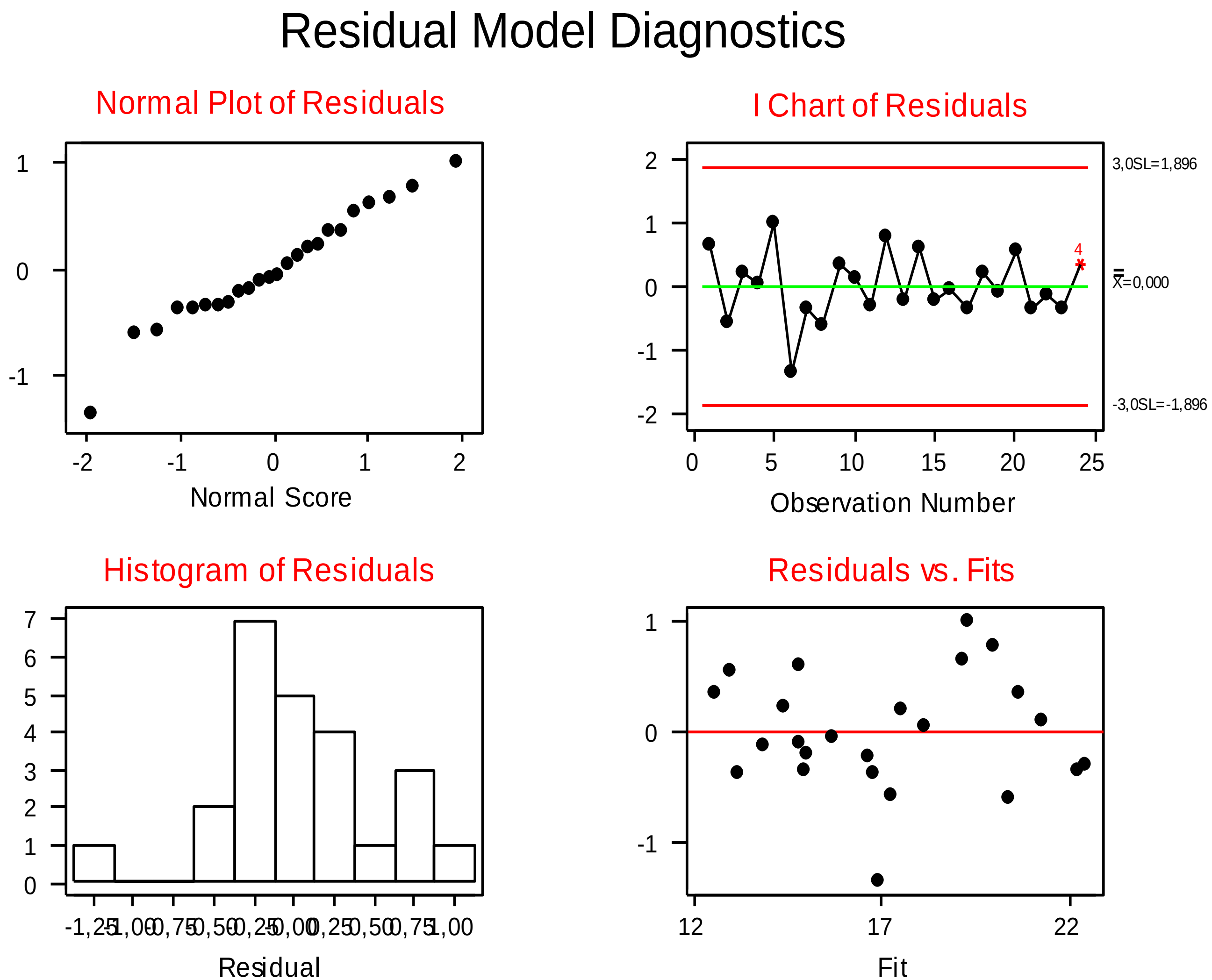
Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
factor(Var)	3	188.538	62.846	144.4369	2.742e-11	***
factor(Bloco)	5	19.793	3.959	9.0981	0.0003857	***
Residuals	15	6.527	0.435			
Total	23	214.858				

Concl. ?

Como é calculado?

```
                                   Individual 95% CI
Var           Mean   --------+---------+---------+---------+---
1            18,03                      (-*-)
2            21,17                                 (--*-)
3            15,53            (-*-)
4            13,70     (-*-)
                    --------+---------+---------+---------+---
                          15,00     17,50     20,00     22,50
```



As suposições do modelo estão satisfeitas ?

Modelos ANOVA – DCA e DABC

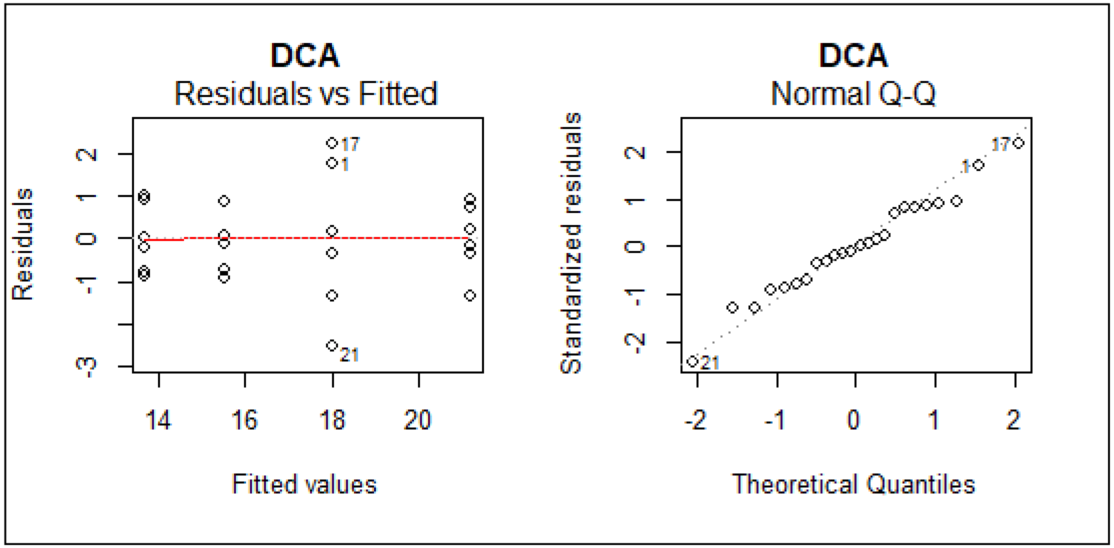
```
Delineamento Completamente Aleatorizado - DCA

              Df Sum Sq Mean Sq F value    Pr(>F)
factor(Var)    3 188.54  62.846  47.755 2.654e-09 ***
Residuals     20  26.32   1.316
```

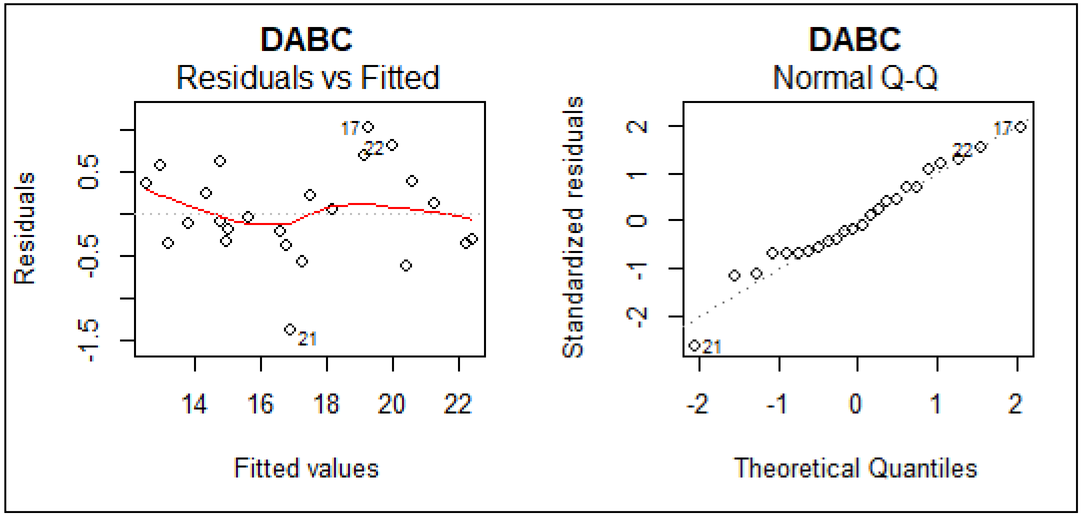
```
Delineamento Aleatorizado em Blocos Completos - DABC

              Df  Sum Sq Mean Sq  F value    Pr(>F)
factor(Var)    3 188.538  62.846 144.4369 2.742e-11 ***
factor(Bloco)  5  19.793   3.959   9.0981 0.0003857 ***
Residuals     15   6.527   0.435
```

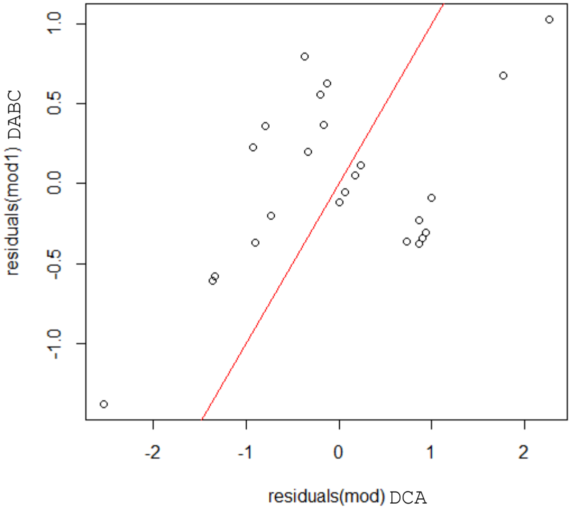
Medida de eficiência: 1.316/0.435 = 3.03

Sucesso da blocagem em reduzir cerca de 3 vezes a variabilidade do erro!

Modelos ANOVA – DCA e DABC







Maior variância dos resíduos sob o DCA (QMRes=1,316) relativamente ao DABC (QMRes=0,435)

DABC - Tukey Simultaneous Tests (efeito do fator Variedade)

Response Variable resp

All Pairwise Comparisons among Levels of var

Hipóteses ?

var = 1 subtracted from:

Level var	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
2	3,133	0,3808	8,23	0,0000
3	-2,500	0,3808	-6,56	0,0001
4	-4,333	0,3808	-11,38	0,0000

(21,17-18,03)/(√2*0.435/6)

var = 2 subtracted from:

Level var	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-5,633	0,3808	-14,79	0,0000
4	-7,467	0,3808	-19,61	0,0000

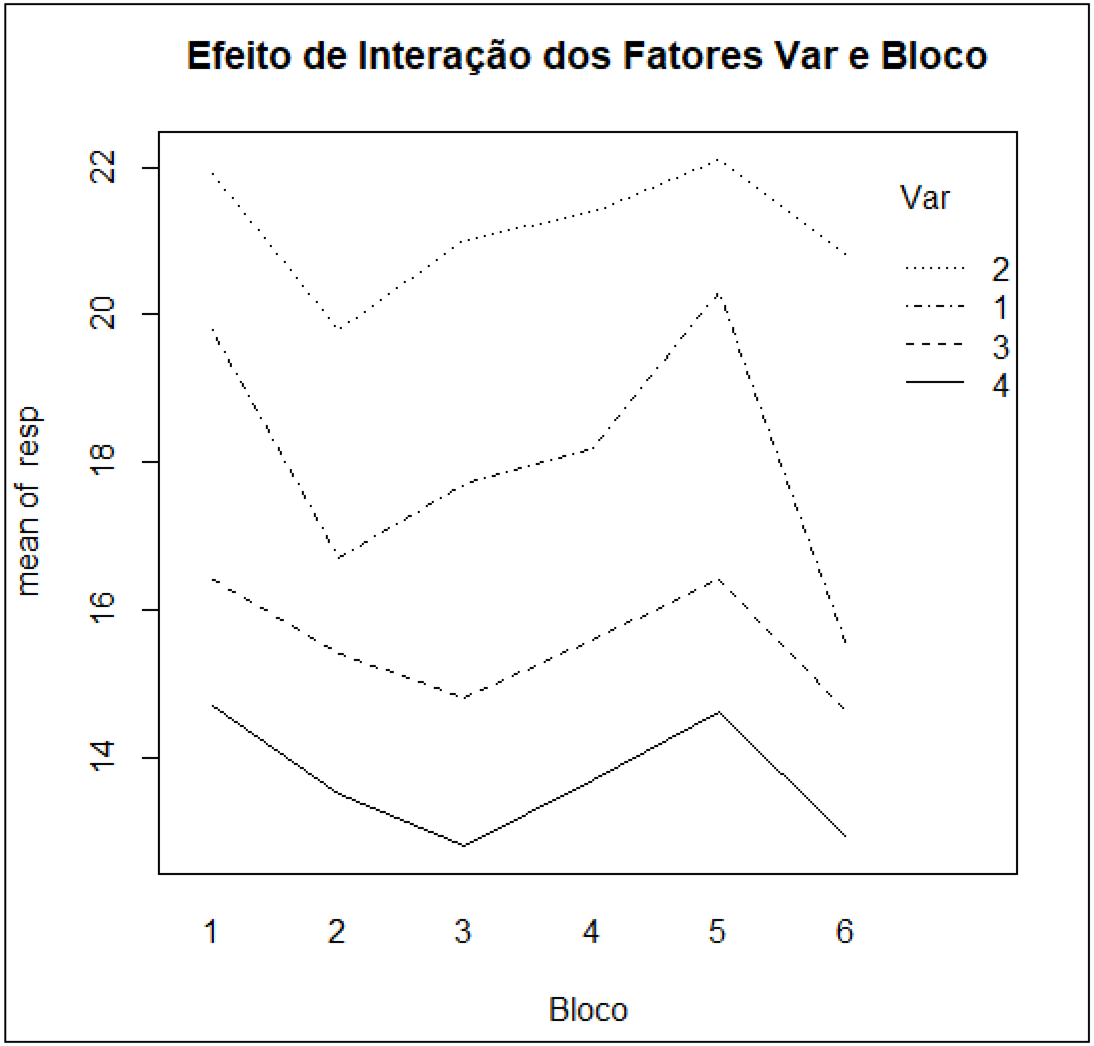
Conclusão?

var = 3 subtracted from:

Level var	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	-1,833	0,3808	-4,814	0,0012

Calcule o erro padrão das diferenças entre médias (sob homocedasticidade).

Modelo ANOVA - DABC



No Delineamento Aleatorizado em Blocos Completos (DABC) o resíduo é o Efeito de Interação entre os fatores Tratamento e Bloco.

Pelo gráfico de perfis individuais há indicação de efeito de interação?

Note que o paralelismos dos perfis é uma indicação de ausência de interação!

ANOVA em um DABC

Modelo Aditivo: $$y_{ij} = \mu + \tau_j + \beta_i + e_{ij}; \quad \sum_{j=1}^{J} \tau_j = \sum_{i=1}^{r} \beta_i = 0$$

Como não há réplicas nas combinações dos níveis de tratamento e bloco, o Resíduo corresponde à Interação entre tratamento e bloco (veja como está definido o número de graus de liberdade e a soma de quadrados do resíduo). Porém, a não aditividade do modelo pode ser estudada particionando os (n-1)(k-1) contrastes que definem a interação.

Teste de Tukey de Não-Aditividade em um DABC usando 1 grau de liberdade:

Construir um novo fator (adicionando uma coluna em X) definido como a interação (produto) entre os efeitos principais preditos de tratamento e de bloco

- Construir uma coluna com os coeficientes do efeito de Tratamento (τ_j)
- Construir uma outra coluna com os coeficientes do fator Bloco (β_i)
- Multiplicar essas duas colunas, definindo assim a interação com 1 g.l.
- Ajustar o modelo de ANOVA incluindo este novo fator (($\tau_j * \beta_i$)

DABC- Teste de Tukey do Modelo Aditivo

```
> mod1$coefficients
    (Intercept)     factor(Var)2    factor(Var)3    factor(Var)4
     19.125000         3.133333       -2.500000       -4.333333
factor(Bloco)2 factor(Bloco)3 factor(Bloco)4 factor(Bloco)5 factor(Bloco)6
     -1.850000      -1.625000      -0.975000       0.150000 -2.250000
```

	Bloco	Var	resp	EfBloco	EfVar	Var*Bloco
1	1	1	19.8	0	0	0
2	1	2	21.9	0	3.13	0
3	1	3	16.4	0	-2.50	0
4	1	4	14.7	0	-4.33	0
5	2	1	16.7	-1.85	0	0
6	2	2	19.8	-1.85	3.13	-5.79
7	2	3	15.4	-1.85	-2.50	4.625
8	2	4	13.5	-1.85	-4.33	8.01
9	3	1	17.7	-1.62	0	0
10	3	2	21.0	-1.62	3.13	-5.07
11	3	3	14.8	-1.62	-2.50	4.05
12	3	4	12.8	-1.62	-4.33	7.01
13	4	1	18.2	-0.975	0	0
14	4	2	21.4	-0.975	3.13	-3.05
15	4	3	15.6	-0.975	-2.50	2.44
16	4	4	13.7	-0.975	-4.33	4.22
17	5	1	20.3	0.15	0	0
18	5	2	22.1	0.15	3.13	0.47
19	5	3	16.4	0.15	-2.50	-0.375
20	5	4	14.6	0.15	-4.33	-0.650
21	6	1	15.5	-2.25	0	0
22	6	2	20.8	-2.25	3.13	-7.04
23	6	3	14.6	-2.25	-2.50	5.63
24	6	4	12.9	-2.25	-4.33	9.74

$$X_{\hat{\tau}_j \hat{\beta}_i} = \hat{\tau}_j \hat{\beta}_i$$

	Bloco	Var	resp	EfBloco	EfVar	Var*Bloco
1	1	1	19.8	0	0	0
2	1	2	21.9	0	3.13	0
3	1	3	16.4	0	-2.50	0
4	1	4	14.7	0	-4.33	0
5	2	1	16.7	-1.85	0	0
6	2	2	19.8	-1.85	3.13	-5.79
7	2	3	15.4	-1.85	-2.50	4.625
8	2	4	13.5	-1.85	-4.33	8.01
9	3	1	17.7	-1.62	0	0
10	3	2	21.0	-1.62	3.13	-5.07
11	3	3	14.8	-1.62	-2.50	4.05
12	3	4	12.8	-1.62	-4.33	7.01
13	4	1	18.2	-0.975	0	0
14	4	2	21.4	-0.975	3.13	-3.05
15	4	3	15.6	-0.975	-2.50	2.44
16	4	4	13.7	-0.975	-4.33	4.22
17	5	1	20.3	0.15	0	0
18	5	2	22.1	0.15	3.13	0.47
19	5	3	16.4	0.15	-2.50	-0.375
20	5	4	14.6	0.15	-4.33	-0.650
21	6	1	15.5	-2.25	0	0
22	6	2	20.8	-2.25	3.13	-7.04
23	6	3	14.6	-2.25	-2.50	5.63
24	6	4	12.9	-2.25	-4.33	9.74

$\hat{\beta}_i$ $\quad$ $\hat{\tau}_j$ $\quad$ $X_{\hat{\tau}_j\hat{\beta}_i} = \hat{\tau}_j\hat{\beta}_i$

Tabela de ANOVA para o DABC

$$y_{ij} = \mu + \tau_j + \beta_i + e_{ij}; \quad \sum_{j=1}^{J}\tau_j = \sum_{i=1}^{r}\beta_i = 0$$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(Var)	3	188.538	62.846	144.4369	2.742e-11
factor(Bloco)	5	19.793	3.959	9.0981	0.0003857
Residuals	15	6.527	0.435		

Tabela de ANOVA para o teste de Aditividade

$$y_{ij} = \mu + \tau_j + \beta_i + \gamma X_{\hat{\tau}_j\hat{\beta}_i} + e_{ij}; \quad \sum_{j=1}^{J}\tau_j = \sum_{i=1}^{r}\beta_i = 0$$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(Var)	3	188.538	62.846	137.6560	1.262e-10
factor(Bloco)	5	19.793	3.959	8.6709	0.0006414
Var*Bloco	1	0.135	0.135	0.2958	0.5950899
Residuals	14	6.392	0.457		

$$H_0 : \gamma = 0$$

Concl: Não há evidência para a rejeição do modelo aditivo

Soma de Quadrados da Não-aditividade:

$$SQ(TrBl) = \sum_{i=1}^{r}\sum_{j=1}^{J}\hat{\gamma}^2\left(\bar{Y}_{i.} - \bar{Y}\right)^2\left(\bar{Y}_{.j} - \bar{Y}\right)^2 = \left[\sum_{i=1}^{r}\sum_{j=1}^{J}Y_{ij}\left(\bar{Y}_{i.} - \bar{Y}\right)\left(\bar{Y}_{.j} - \bar{Y}\right)\right]^2 / rJ(SQTr)(SQBl)$$

$$\hat{\gamma} = \left(\sum_{i=1}^{r}\sum_{j=1}^{J}\hat{\tau}_j\hat{\beta}_i Y_{ij}\right) / \left(\sum_{i=1}^{r}\sum_{j=1}^{J}\hat{\tau}_j^2\hat{\beta}_i^2\right)$$

Exemplo

<u>Dados</u>: Medidas de *clorofila a*

	Tratamento				
Bloco	**T1**	**T2**	**T3**	**T4**	**Média**
B1	**6,2**	**12,7**	**7,0**	**8,3**	**8,6**
B2	**4,8**	**11,3**	**4,4**	**7,1**	**6,9**
B3	**3,0**	**9,3**	**3,8**	**11,7**	**6,9**
B4	**5,6**	**9,5**	**5,0**	**10,0**	**7,5**
B5	**7,1**	**11,7**	**5,5**	**8,5**	**8,2**
B6	**4,8**	**15,3**	**3,2**	**12,4**	**8,9**
Média	**5,3**	**11,6**	**4,8**	**9,7**	**7,8**

Estrutura de Bloco

(hipotético)

Tabela de ANOVA - DABC

Compare com a tabela de ANOVA do correspondente DCA.

$$H_0\text{: } \mu_1 = \mu_2 = \ldots = \mu_J = \mu$$

F.V.	g.l.	SQ	QM	F	p
TRAT	3	201.448	67.149	19.79	0.000
BLOCO	5	14.343	2.869	0.85	0.538
RESÍDUO	15	50.346	3.392		
TOTAL	23	266.678			

Com a inclusão do fator bloco, houve ganho em precisão na estimação do efeito do tratamento?

Compare o número de graus de liberdade dos delineamentos DCA e DABC definidos para o mesmo conjunto de dados!

Modelos ANOVA

DCA: Delineamento Completamente Aleatorizado

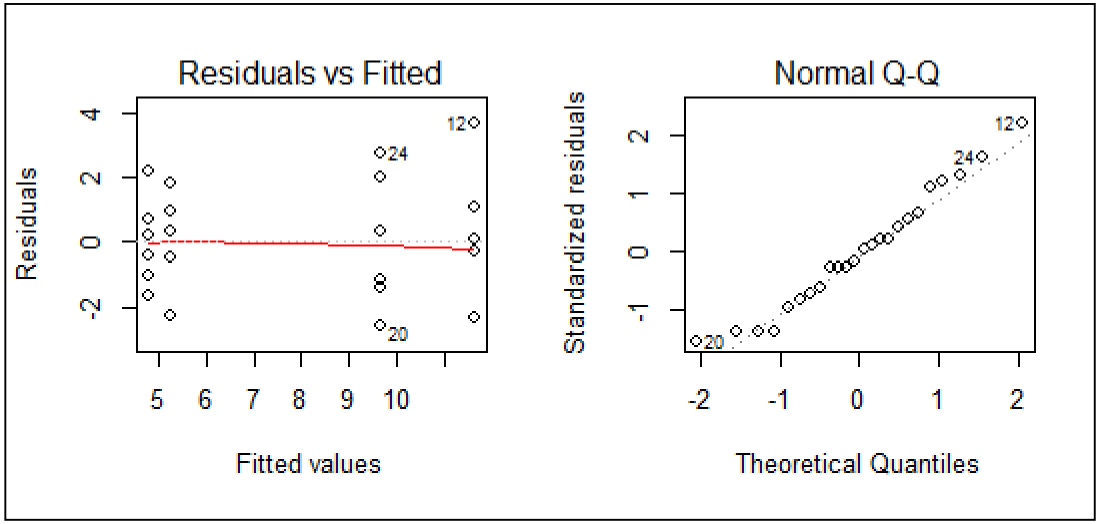
```
Analysis of Variance Table
          Df  Sum Sq Mean Sq F value    Pr(>F)
Trat       3 199.937  66.646   20.16 2.924e-06 ***
Residuals 20  66.117   3.306
```

DABC: Delineamento Aleatorizado em Blocos Completos

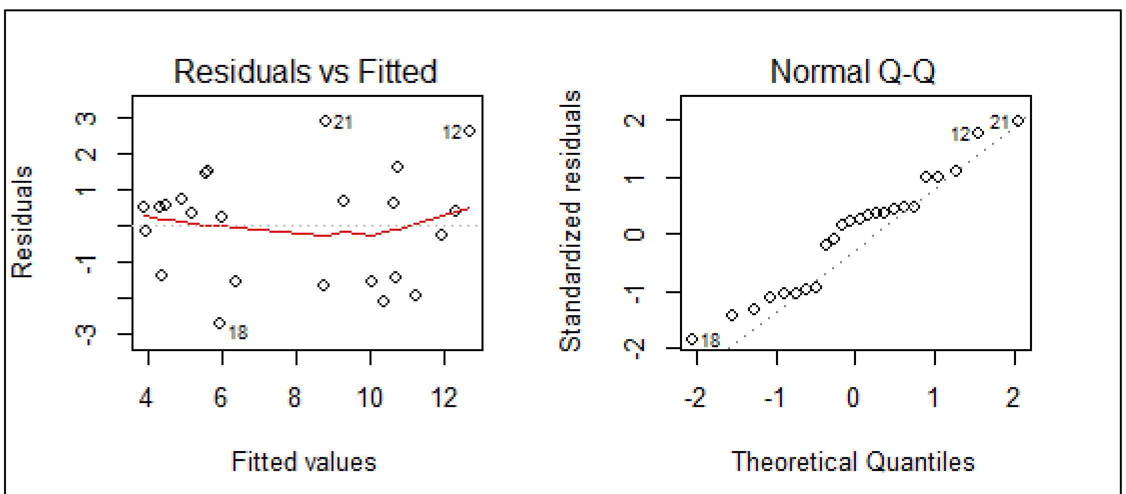
```
Analysis of Variance Table
          Df  Sum Sq Mean Sq F value   Pr(>F)
Trat       3 199.937  66.646 19.3593 2.04e-05 ***
Bloco      5  14.478   2.896  0.8411   0.5412
Residuals 15  51.638   3.443
```

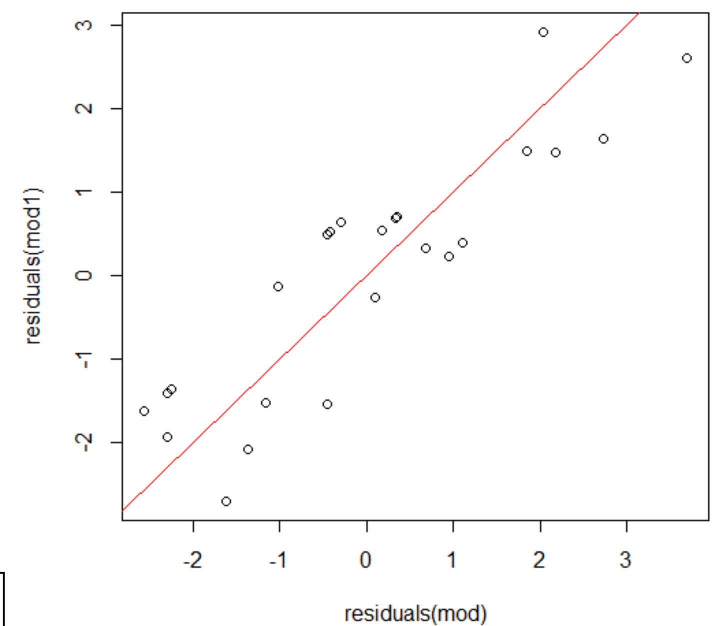
Modelos ANOVA

ANOVA-DCA



ANOVA-DABC





Não há evidência amostral de ganho em precisão devido à inclusão do fator Bloco: sob o DCA (QMRes=3,306) e sob o DABC (QMRes=3,443)

Dados: Medidas de clorofila a

	trat	bloco	resp
1	1	1	6.2
2	2	1	12.7
3	3	1	7.0
4	4	1	8.3
5	1	2	4.8
6	2	2	11.3
7	3	2	4.4
8	4	2	7.1
9	1	3	3.0
10	2	3	9.3
11	3	3	3.8
12	4	3	11.7
13	1	4	5.6
14	2	4	9.3
15	3	4	5.0
16	4	4	10.0
17	1	5	7.1
18	2	5	11.7
19	3	5	5.5
20	4	5	8.5
21	1	6	4.8
22	2	6	15.3
23	3	6	3.2
24	4	6	12.4

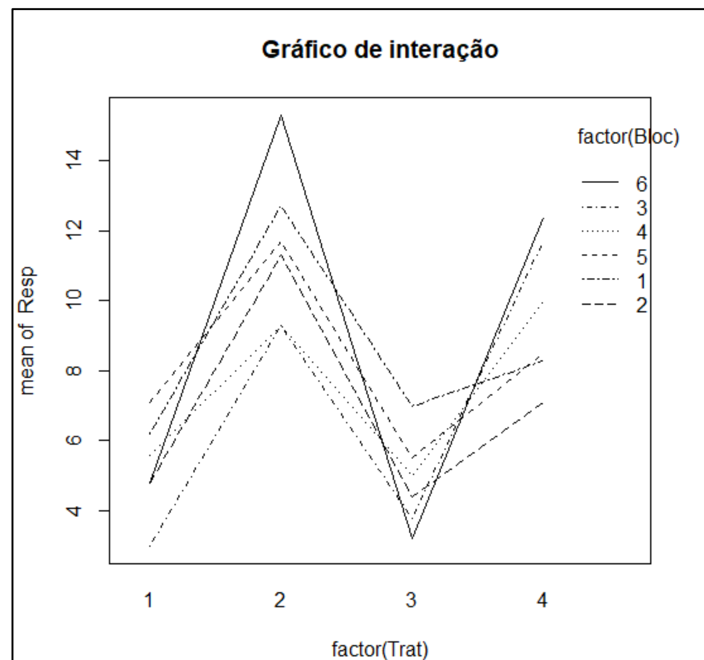
Estimativas dos parâmetros do modelo (no "R")

Intercept	trat2	trat3	trat4
5.9666667	6.3500000	-0.4333333	4.4166667

bloco2	bloco3	bloco4	bloco5	bloco6
-1.6500000	-1.6000000	-1.0750000	-0.3500000	0.3750000

Interprete as estimativas dos parâmetros do modelo estrutural adotado na análise dos dados (sob o DABC)!

Há indicação de efeito de interação entre os fatores Tratamento e Bloco?



Dados: Medidas de clorofila a – Teste de Tukey do Modelo Aditivo

	trat	bloco	resp	tpred	bpred	trbl
1	1	1	6.2	0.0000000	0.000	0.0000000
2	2	1	12.7	6.3500000	0.000	0.0000000
3	3	1	7.0	-0.4333333	0.000	0.0000000
4	4	1	8.3	4.4166667	0.000	0.0000000
5	1	2	4.8	0.0000000	-1.650	0.0000000
6	2	2	11.3	6.3500000	-1.650	-10.4775000
7	3	2	4.4	-0.4333333	-1.650	0.7150000
8	4	2	7.1	4.4166667	-1.650	-7.2875000
9	1	3	3.0	0.0000000	-1.600	0.0000000
10	2	3	9.3	6.3500000	-1.600	-10.1600000
11	3	3	3.8	-0.4333333	-1.600	0.6933333
12	4	3	11.7	4.4166667	-1.600	-7.0666667
13	1	4	5.6	0.0000000	-1.075	0.0000000
14	2	4	9.3	6.3500000	-1.075	-6.8262500
15	3	4	5.0	-0.4333333	-1.075	0.4658333
16	4	4	10.0	4.4166667	-1.075	-4.7479167
17	1	5	7.1	0.0000000	-0.350	0.0000000
18	2	5	11.7	6.3500000	-0.350	-2.2225000
19	3	5	5.5	-0.4333333	-0.350	0.1516667
20	4	5	8.5	4.4166667	-0.350	-1.5458333
21	1	6	4.8	0.0000000	0.375	0.0000000
22	2	6	15.3	6.3500000	0.375	2.3812500
23	3	6	3.2	-0.4333333	0.375	-0.1625000
24	4	6	12.4	4.4166667	0.375	1.6562500

$$X_{\hat{\tau}_j \hat{\beta}_i} = \hat{\tau}_j \hat{\beta}_i$$

$$y_{ij} = \mu + \tau_j + \beta_i + e_{ij}; \quad \sum_{j=1}^{J} \tau_j = \sum_{i=1}^{r} \beta_i = 0$$

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
trat	3	199.937	66.646	19.3593	2.04e-05
bloco	5	14.478	2.896	0.8411	0.5412
Residuals	**15**	**51.638**	**3.443**		

$$y_{ij} = \mu + \tau_j + \beta_i + \gamma X_{\hat{\tau}_j \hat{\beta}_i} + e_{ij}; \quad \sum_{j=1}^{J} \tau_j = \sum_{i=1}^{r} \beta_i = 0$$

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
trat	3	199.937	66.646	19.5807	2.802e-05
bloco	5	14.478	2.896	0.8508	0.5367
trbl	**1**	**3.987**	**3.987**	**1.1715**	**0.2974**
Residuals	**14**	**47.651**	**3.404**		

$$H_0 : \gamma = 0$$

Concl: Não há evidência para a rejeição do modelo aditivo

Testes de Aleatorização

- A **aleatorização** (atribuição aleatória dos “Fatores” às unidades experimentais) permite a construção de uma **Distribuição de Referência** para estatísticas de interesse (que comparam os fatores)

- A Distribuição de Referência pode ser usada para calcular a “significância estatística” do resultado efetivamente observado na amostra e a única premissa usada é a Aleatorização

- É possível mostrar que os valores-p calculados sob a Distribuição de Referência sob aleatorização são próximos daqueles calculados sob a hipótese de Normalidade

No DABC, a aleatorização pode ser usada para justificar a construção do Teste do Efeito dos Tratamentos, mas NÃO pode ser usada para justificar o Teste do Efeito de Blocos, já que esse fator foi fixado e não houve atribuição aleatória de seus níveis às unidades experimentais.

Testes de Aleatorização

Os dados a seguir se referem à diferença na contagem de um tipo de besouro que infesta folhagens, antes e depois de um tratamento por 21 dias. Três tratamentos de controle dessa infestação estão sendo comparados: tratamento cm água (T1), esporos (T2) e um óleo natural (T3). 5 plantas foram selecionadas para o experimento e, de cada uma, 3 folhagens foram escolhidas sendo que, dentro de cada folhagem, 2 quadrantes foram definidos. Em cada planta as folhagens foram aleatoriamente destinadas aos tratamentos.

Trat	Planta 1	Planta 2	Planta 3	Planta 4	Planta 5
T1: Água	-9	18	10	?	-6
	-6	5	9	?	13
T2: Esporos	-4	29	4	-2	11
	7	10	-1	6	-1
T3: Óleo	4	29	14	14	7
	11	36	16	18	15

- Como está definida a unidade experimental?
- Complete a tabela nas caselas faltantes (?). Obtenha a tabela de ANOVA sob DABC, bem como sob um DCA, sob premissas clássicas, e compare a eficiência da blocagem. Considere na análise a média da resposta nos quadrantes de cada folhagem.
- Obtenha o valor-p do Teste de Aleatorização Global de efeito de Tratamentos. Considere o caso de um DABC.