MAT02014 - Planejamento de Experimentos II
Introdução aos delineamentos fatoriais (cont.)

Rodrigo Citton P. dos Reis
rodrigocpdosreis@gmail.com

Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística

Departamento de Estatística
Porto Alegre, 2018

Análise estatística do modelo de efeitos fixos

Relembrando

■ **TABLE 5.2**

General Arrangement for a Two-Factor Factorial Design

		Factor *B*			
		1	2	. . .	b
Factor *A*	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$		$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
	⋮				
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

- O **modelo de efeitos**:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}, i = 1, \ldots, a, j = 1, \ldots,$$

Somas de quadrados

- $y_{i..}$ denota o total de todas as observações sob o i-ésimo nível do fator A;
- $y_{.j.}$ denota o total de todas as observações sob o j-ésimo nível do fator B;
- $y_{ij.}$ denota o total de todas as observações na ij-ésima célula;
- $y_{...}$ denota o total global de todas as observações.

Somas de quadrados

- Defina $\bar{y}_{i..}, \bar{y}_{.j.}, \bar{y}_{ij.}$ e $\bar{y}_{...}$ como as correspondentes médias da linha, coluna, célula e global:

$$y_{i..} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk} \quad \bar{y}_{i..} = \frac{y_{i..}}{bn}, i = 1, 2, \ldots, a$$

$$y_{.j.} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk} \quad \bar{y}_{.j.} = \frac{y_{i..}}{an}, j = 1, 2, \ldots, b$$

$$y_{ij.} = \sum_{k=1}^{n} y_{ijk} \quad \bar{y}_{ij.} = \frac{y_{ij.}}{n}, i = 1, 2, \ldots, a, j = 1, 2, \ldots, b$$

$$y_{...} = \sum^{a} \sum^{b} \sum^{n} y_{ijk} \quad \bar{y}_{...} = \frac{y_{...}}{abn}$$

$$\sum_{i=1} \sum_{j=1} \sum_{k=1} \qquad abn$$

Somas de quadrados

- A **soma de quadrados total** deve ser escrita como:

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{...})^2 = bn \sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2 + an \sum_{j=1}^{b} (\bar{y}$$

$$n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}$$

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{...})^2$$

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E.$$

Graus de liberdade

Efeito	**Graus de liberdade**
A	a - 1
B	b - 1
AB	(a - 1)(b - 1)
Erro	ab(n - 1)
Total	abn - 1

Quadrados médios

- Cada soma de quadrados dividida por seus graus de liberdade é um **quadrado médio**.
  - Os valores esperados dos quadrados médios são:

$$E[MS_A] = E\left[\frac{SS_A}{a-1}\right] = \sigma^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E[MS_B] = E\left[\frac{SS_B}{b-1}\right] = \sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$E[MS_{AB}] = E\left[\frac{SS_{AB}}{(a-1)(b-1)}\right] = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau}{(a-1)(b-1}$$

$$E[MS_E] = E\left[\frac{SS_E}{}\right] = \sigma^2$$

$$E[MS_E] = E\left[\frac{\quad}{ab(n-1)}\right] = \sigma^2$$

Análise de variância

- Note que se a hipótese nula de *nenhum efeito de tratamento de linha*, *nenhum efeito de tratamento de coluna* e *nenhuma interação*, então MS_A, MS_B, MS_{AB} e MS_E todos estimam σ^2.
  - No entanto, se existem diferenças entre os efeitos de tratamento de linha, então MS_A é maior que MS_E.
  - De maneira similar, se existem diferenças entre os efeitos de tratamento de coluna, então MS_B é maior que MS_E.
- Assim, para testar a significância dos efeitos principais e do efeito de interação, basta dividir o correspondente quadrado médio pelo quadrado médio do erro.
  - Grandes valores desta razão indicam indícios contra a hipótese nula.

Análise de variância

- Se assumimos que o modelo de efeitos é adequado e que os termos de erro ϵ_{ijk} são independentes com distribuição normal com variância constante σ^2, então:
  - $MS_A/MS_E \sim F_{a-1,ab(n-1)}$,
  - $MS_B/MS_E \sim F_{b-1,ab(n-1)}$
  - $MS_{AB}/MS_E \sim F_{b-1,ab(n-1)}$

Análise de variância

■ **TABLE 5.3**

The Analysis of Variance Table for the Two-Factor Factorial, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A treatments	SS_A	$a - 1$	$MS_A = \dfrac{SS_A}{a - 1}$	$F_0 = \dfrac{MS_A}{MS_E}$
B treatments	SS_B	$b - 1$	$MS_B = \dfrac{SS_B}{b - 1}$	$F_0 = \dfrac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a - 1)(b - 1)$	$MS_{AB} = \dfrac{SS_{AB}}{(a - 1)(b - 1)}$	$F_0 = \dfrac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n - 1)$	$MS_E = \dfrac{SS_E}{ab(n - 1)}$	
Total	SS_T	$abn - 1$		

Exemplo da bateria

```
##    MT  T   y
## 1   1 15 130
## 2   2 15 150
## 3   3 15 138
## 4   1 70  34
## 5   2 70 136
## 6   3 70 174
```

Exemplo da bateria

```
exbat.av <- aov(y ~ MT*T, data = D)
summary(exbat.av)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## MT           2  10684    5342   7.911  0.00198 ** 
## T            2  39119   19559  28.968 1.91e-07 ***
## MT:T         4   9614    2403   3.560  0.01861 *  
## Residuals   27  18231     675                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

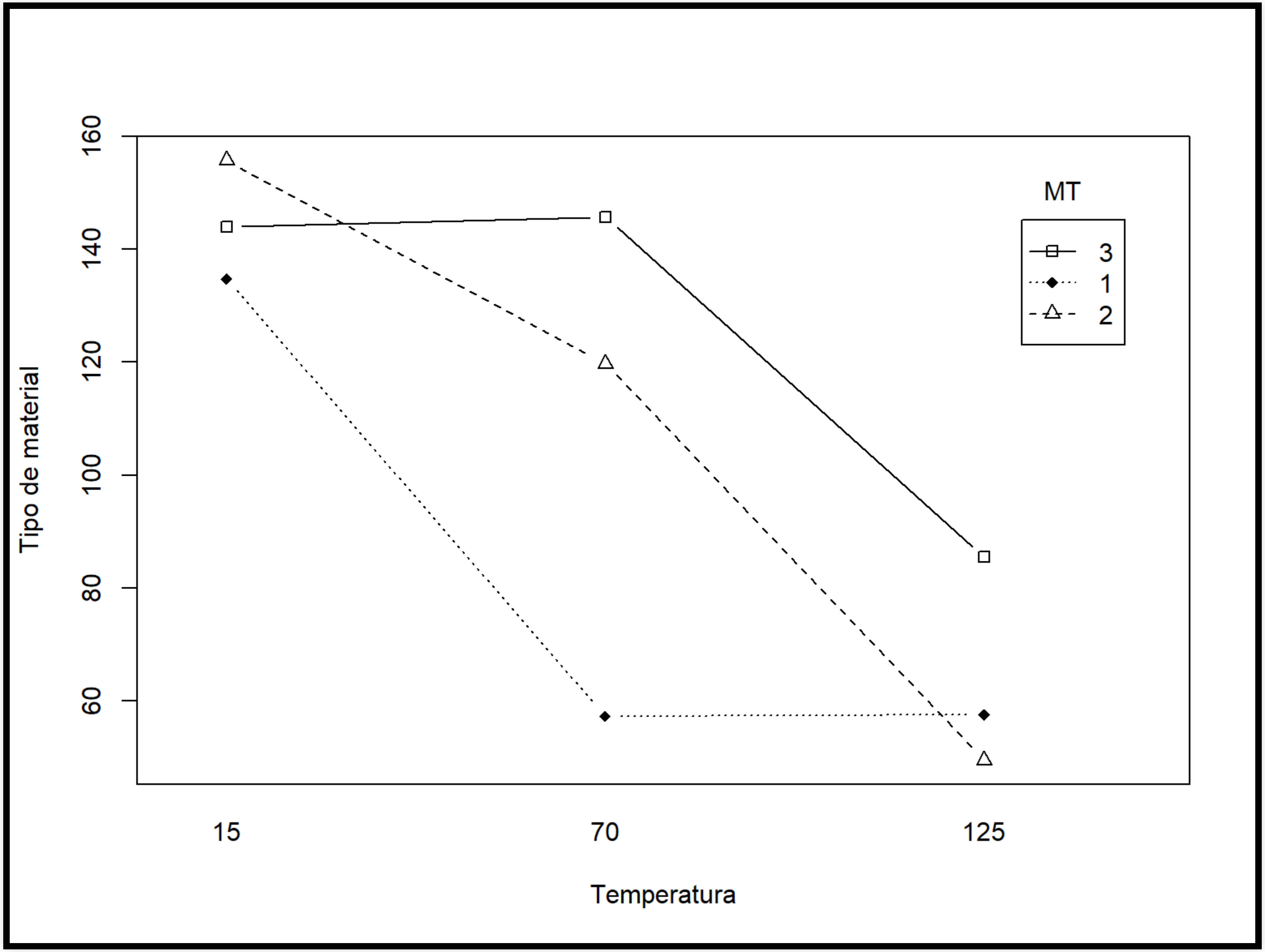
Exemplo da bateria

```r
model.tables(exbat.av, type = "means", se = F)
```

```
## Tables of means
## Grand mean
##
## 105.5278
##
##  MT
## MT
##       1      2      3
##  83.17 108.33 125.08
##
##  T
## T
##      15     70    125
## 144.83 107.58  64.17
##
##  MT:T
##    T
## MT  15     70     125
```

Exemplo da bateria

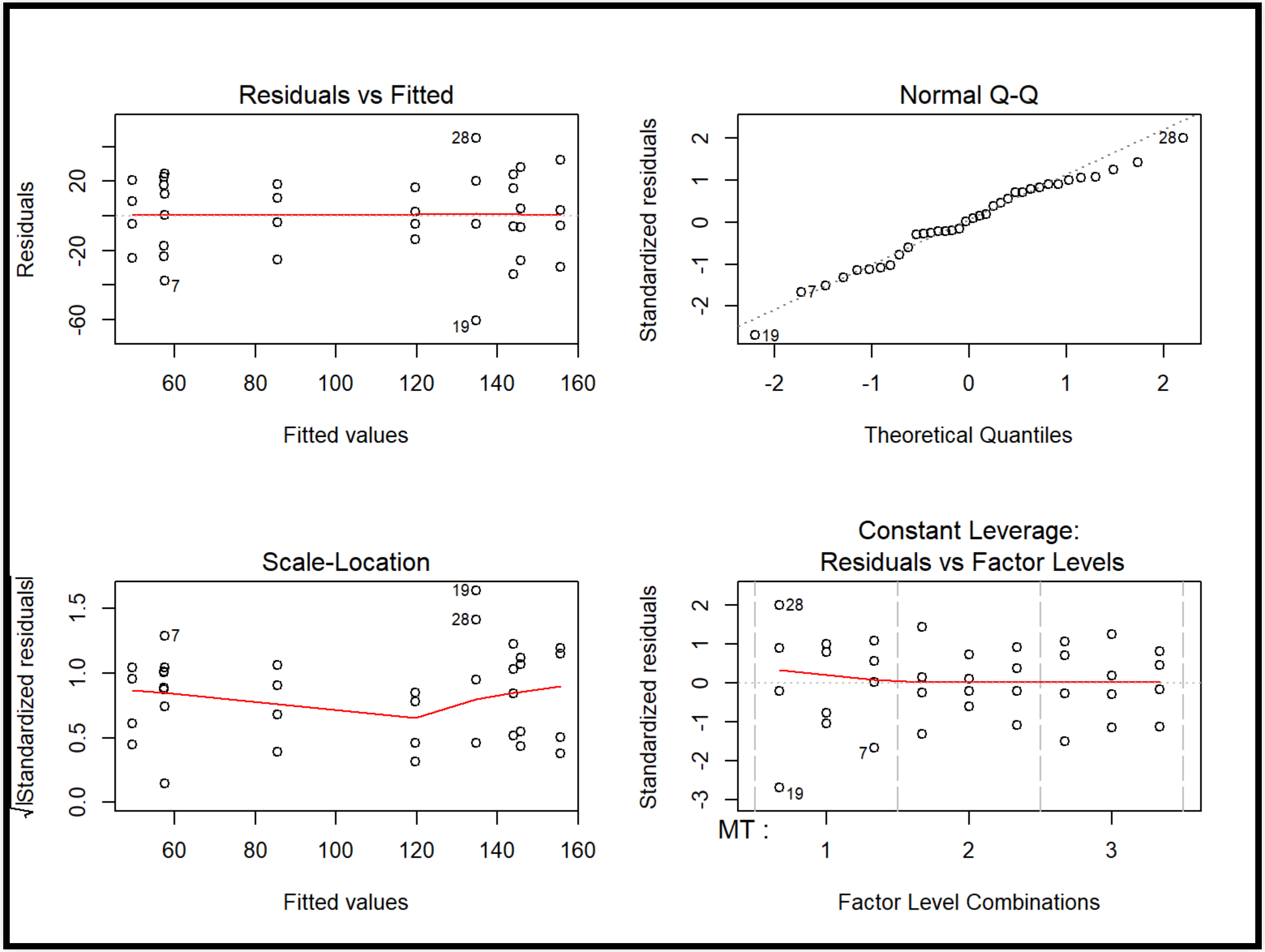
```
with(D, (interaction.plot(T, MT, y, type = "b", pch = c(18,24,22), leg.bt
```


Tipo de material
160
140
120
100
80
60
MT
3
1
2
15
70
125
Temperatura

```
## NULL
```

Exemplo da bateria

```
par(mfrow = c(2,2))
plot(exbat.av)
```


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
MT :
Factor Level Combinations

Comparações múltiplas

- Quando a análise de variância indica que as médias de linha ou coluna diferem, usualmente é de interesse fazer comparações entre médias individuais de linha ou coluna para descobrir diferenças específicas.
  - Métodos de comparações múltiplas.

Comparações múltiplas

Exemplo da bateria

- Suponha que estejamos interessados em detectar diferenças entre as médias dos três tipos de materiais.
  - Como a interação é significativa, faremos esta comparação em apenas um nível de temperatura (nível 2: 70ºF).
- O **teste de Tukey** pode ser utilizado para comparar os níveis do fator “Tipo de material” mantendo fixa a temperatura.
- Assumiremos que a melhor estimativa para a variância do erro é o MS_E.
- Ao nível de significância de 5%

$$T_{0,05} = q_{0,05}(3, 27)\sqrt{\frac{MS_E}{n}},$$

em que $q_{0,05}(3, 27)$ é o percentil 95% da distribuição q (*distribuição da estatística da amplitude studentizada*) com três comparações e 27 graus de liberdade (associados ao MS_E).

Comparações múltiplas

```
qtukey(0.95, 3, 27) * sqrt(sum(exbat.av$residuals^2)/(3*3*3)/4)
```

```
## [1] 45.557
```

```
model.tables(exbat.av, "means")$tables$`MT:T`[,2][3] - model.tables(exbat
```

```
##    2    1
## 26.0 88.5
```

```
model.tables(exbat.av, "means")$tables$`MT:T`[,2][2] - model.tables(exbat
```

```
##    2
## 62.5
```

Para casa (para aula)

- Realize o procedimento de comparações múltiplas para comparar as médias dos tipos de materiais nos outros níveis de temperatura.
- Realize o procedimento de comparações múltiplas para comparar as médias das temperaturas fixando os níveis de tipo de material.
- Pense de que outra forma você poderia avaliar a estatística de teste F_0.
    - Implement sua ideia.


YAY STATISTICS