Delineamento Quadrado Latino

Profa. Cibele Russo

(Referências: Montgomery (2012), Notas de aula de Roseli Leandro; Clarice Demétrio; Marinho Andrade)

Compare os croquis a seguir:

1. Delineamento em blocos completos casualizados (DBCC)
2. Delineamento em quadrado latino

Quais as diferenças?

- Quais princípios básicos da experimentação estão sendo contemplados em cada um deles?
- Lembre-se dos princípios básicos da experimentação:
    1. Repetição
    2. Casualização
    3. Controle local

Delineamento em blocos completos casualizados

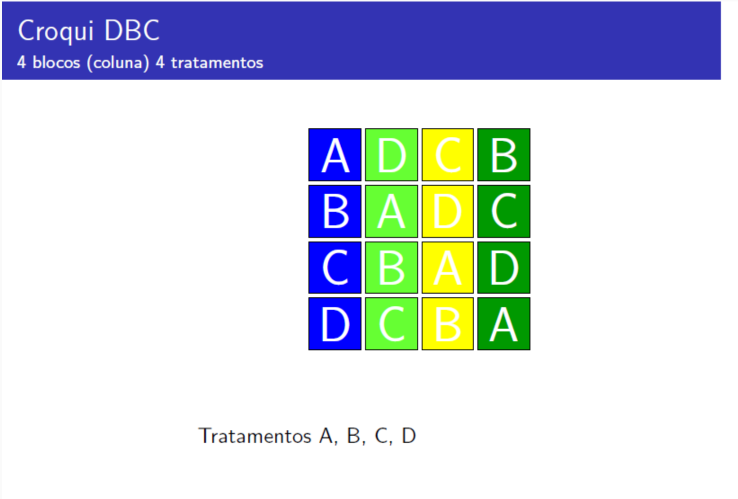


Figure 1: Delineamento blocos completos casualizado

Delineamento quadrado latino

Croqui Quadrado Latino 4×4

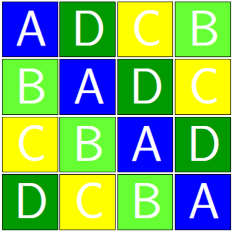
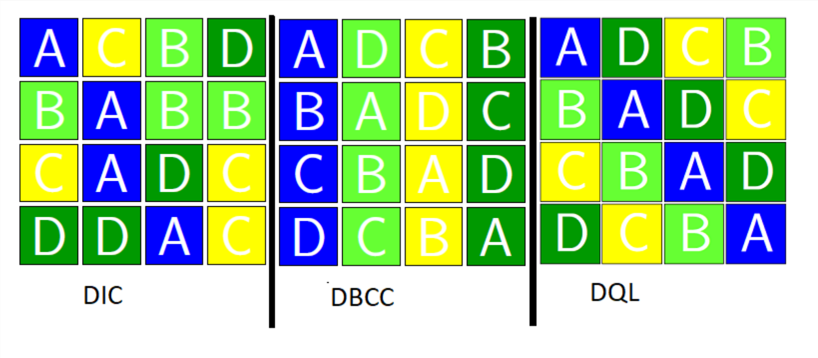


Figure 2: Delineamento Quadrado Latino

Compare os croquis dos diferentes delineamentos



Croquis na indústria

- como você colocaria a disposição os croquis dos delineamentos
    1) DIC
    2) DBCC
    3) DQL

na indústria?

- Exemplifique com situações práticas

Número de fatores de perturbação e delineamentos associados

- O delineamento em blocos completos casualizados (DBCC) reduz o erro residual de um experimento,
- "remove" do erro residual a variabilidade devida ao fator de perturbação conhecido e controlável.
- Existem outros tipos de delineamento que podem controlar
    1. dois fatores de perturbação (Delineamento em quadrado latino)
    2. três fatores de perturbaçao (Delineamento Greco-Latino)
    3. existem outros.

Delineamento Quadrado Latino

- No delineamento quadrado latino todos os três princípios são contemplados
- O controle local é considerando em duas direções: linhas e colunas
- cada tratamento pode ser alocado uma única vez na linha e coluna como mostra o croqui apresentado.

Carro, motorista: Comparação de aditivos

Carro, motorista: Comparação de aditivos

Gado, raça: comparação de variedade de capim

Gado, raça: comparação de variedade de capim

Delineamento quadrado latino

- O delineamento quadrado latino é usado para eliminar duas fontes incômodas de variabilidade; isto é,
- sistematicamente permite o bloqueio em duas direções.
- Assim, as linhas e colunas representam duas restrições à randomização (casualização).
- Em qualquer experimento, a variabilidade decorrente de um fator incômodo (nuisance factor) pode afetar os resultados.

Delineamento quadrado latino

- Em geral, um quadrado latino para p tratamentos ou um Quadrado latino $p \times p$, é um quadrado que contém p linhas e p colunas.
- Cada uma das p^2 caselas resultantes contém uma das letras p que corresponde aos tratamentos e cada letra ocorre uma vez e apenas uma vez, em cada linha e coluna.

Alguns exemplos de quadrados latinos

4 × 4

A	B	D	C
B	C	A	D
C	D	B	A
D	A	C	B

5 × 5

A	D	B	E	C
D	A	C	B	E
C	B	E	D	A
B	E	A	C	D
E	C	D	A	B

6 × 6

A	D	C	E	B	F
B	A	E	C	F	D
C	E	D	F	A	B
D	C	F	B	E	A
F	B	A	D	C	E
E	F	B	A	D	C

Quadrado latino padrão

- Um quadrado latino no qual a primeira linha e a primeira coluna está escrita em ordem alfabética é chamado **quadrado latino padrão**.
- Um quadrado latino padrão pode ser obtido escrevendo-se a primeira linha e a primeira coluna em ordem alfabética e então, escrevendo-se cada linha sucessiva deslocando-se as letras em uma posição. Veja na Tabela 4.13.

Delineamento quadrado latino

■ **TABLE 4.13**

Standard Latin Squares and Number of Latin Squares of Various Sizes[a]

Size	3 × 3	4 × 4	5 × 5	6 × 6	7 × 7	$p \times p$
Examples of standard squares	*A B C* *B C A* *C A B*	*A B C D* *B C D A* *C D A B* *D A B C*	*A B C D E* *B A E C D* *C D A E B* *D E B A C* *E C D B A*	*A B C D E F* *B C F A D E* *C F B E A D* *D E A B F C* *E A D F C B* *F D E C B A*	*A B C D E F G* *B C D E F G A* *C D E F G A B* *D E F G A B C* *E F G A B C D* *F G A B C D E* *G A B C D E F*	*ABC . . . P* *BCD . . . A* *CDE . . . B* ⋮ *PAB . . .* $(P-1)$
Number of standard squares	1	4	56	9408	16,942,080	—
Total number of Latin squares	12	576	161,280	818,851,200	61,479,419,904,000	$p!(p-1)! \times$ (number of standard squares)

[a]Some of the information in this table is found in Fisher and Yates (1953). Little is known about the properties of Latin squares larger than 7 × 7.

O modelo estatístico

The **statistical model** for a Latin square is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \end{cases}$$

The analysis of variance consists of partitioning the total sum of squares of the $N = p^2$ observations into components for rows, columns, treatments, and error,

$$SS_T = \boxed{SS_{\text{Rows}}} + \boxed{SS_{\text{Columns}}} + \boxed{SS_{\text{Treatments}}} + \boxed{SS_E}$$

with respective degrees of freedom

$$p^2 - 1 = \boxed{p - 1} + \boxed{p - 1} + \boxed{p - 1} + \boxed{(p - 2)(p - 1)}$$

- SS representa soma de quadrados

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk}$$

em que

- y_{ijk} é a observação da $i-$ésima linha e $k-$ésima coluna para $j-$ tratamento;
- μ é a média geral (considerando-se restrição soma zero);
- τ_j é o efeito do tratamento j;
- α_i é o efeito da linha i;
- β_k é o efeito da coluna k;
- ϵ_{ijk} é o erro aleatório são erros aleatórios independentes e identicamente distribuídos, $\epsilon_{ijk} \overset{i.i.d.}{\sim} N(0, \sigma^2)$

Matricialmente

$$Y = X\theta + \epsilon$$

- $X_{n \times p}$ matriz do delineamento
- $\theta_{(1+p+p+p)\times 1}$ vetor de parâmetros

$$\theta' = [\mu, \alpha_1, \alpha_2, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_p, \tau_1, \tau_2, \ldots, \tau_p]$$

- $$\hat{\sigma^2} = \frac{SQRes}{glRes} = QMRes$$

ANOVA Quadrado latino

TABLE 4.10
Analysis of Variance for the Latin Square Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}} = \frac{1}{p}\sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Treatments}}}{p-1}$	$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$
Rows	$SS_{\text{Rows}} = \frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Rows}}}{p-1}$	
Columns	$SS_{\text{Columns}} = \frac{1}{p}\sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Columns}}}{p-1}$	
Error	SS_E (by subtraction)	$(p-2)(p-1)$	$\frac{SS_E}{(p-2)(p-1)}$	
Total	$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$	p^2-1		

Resíduo

$$\epsilon_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..k} + 2\bar{y}_{...}$$

- $\bar{y}_{i..} = \frac{\sum_{j=1}^{p}\sum_{k=1}^{p} y_{ijk}}{p}$
- $\bar{y}_{.j.} = \frac{\sum_{i=1}^{p}\sum_{k=1}^{p} y_{ijk}}{p}$
- $\bar{y}_{..k} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p} y_{ijk}}{p}$
- $\bar{y}_{...} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p} y_{ijk}}{p}$

Exemplo ilustrativo

Exercício 4.22 página 180 Montgomery Ed. 8

O efeito de cinco ingredientes diferentes (A, B, C, D, E) está sendo estudado no tempo de reação de um processo químico. Cada lote de material novo é grande o suficiente para permitir cinco corridas (sequências, repetições). Além disso, cada execução requer aproximadamente uma hora é meia para ser executada. Assim somente cinco execuções podem ser feitas em um dia pois a CLT (Consolidação das Leis do Trabalho) define que a jornada diária de trabalho deve ser de no máximo oito horas). O experimentador decide executar o experimento como um quadrado latino para que os efeitos do dia e do lote possam ser sistematicamente controlados. Os dados obtidos encontram-se na Tabela 01

Analise os dados desta experiência (use $\alpha = 0,05$) e tire conclusões.

Tabela 01

	Dia				
Lote	1	2	3	4	5
1	A = 8	B = 7	D = 1	C = 7	E = 3
2	C = 11	E = 2	A = 7	D = 3	B = 8
3	B = 4	A = 9	C = 10	E = 1	D = 5
4	D = 6	C = 8	E = 6	B = 6	A = 10
5	E = 4	D = 2	B = 3	A = 8	C = 8

Orientação para estudo

- Faça análise descritiva
- Verifique os pressupostos básicos
- Interprete o quadro da ANOVA
- Realize um teste de comparações múltiplas

Entrada de dados

```r
require(readxl)
```

```
## Carregando pacotes exigidos: readxl
```

```r
dad <- read_xlsx("pag_180_4_22.xlsx")
str(dad)
```

```
## tibble [25 x 4] (S3: tbl_df/tbl/data.frame)
##  $ Lote       : num [1:25] 1 2 3 4 5 1 2 3 4 5 ...
##  $ Dia        : num [1:25] 1 1 1 1 1 2 2 2 2 2 ...
##  $ Ingrediente: chr [1:25] "A" "C" "B" "D" ...
##  $ Tempo      : num [1:25] 8 11 4 6 4 7 2 9 8 2 ...
```

Formatação do arquivo de dados

```
## # A tibble: 25 x 4
##     Lote   Dia Ingrediente Tempo
##    <dbl> <dbl> <chr>       <dbl>
##  1     1     1 A               8
##  2     2     1 C              11
##  3     3     1 B               4
##  4     4     1 D               6
##  5     5     1 E               4
##  6     1     2 B               7
##  7     2     2 E               2
##  8     3     2 A               9
##  9     4     2 C               8
## 10     5     2 D               2
## # i 15 more rows
```

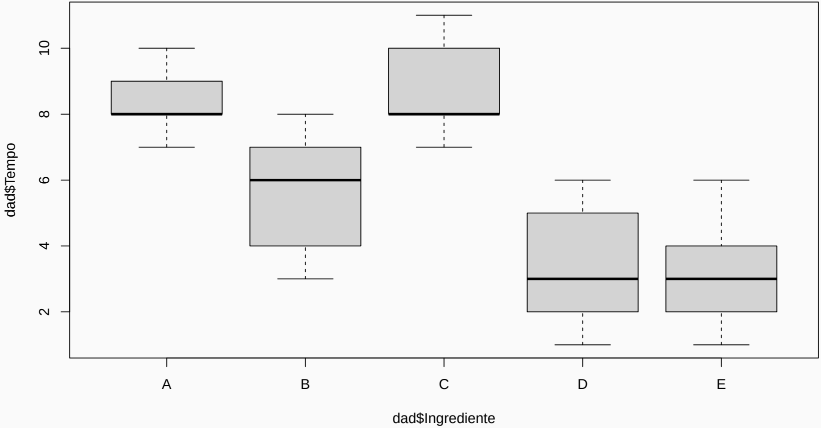
Conversão

```
dad$Lote <- as.factor(dad$Lote)
dad$Dia  <- as.factor(dad$Dia)
dad$Ingrediente <- as.factor(dad$Ingrediente)
str(dad)
```

```
## tibble [25 x 4] (S3: tbl_df/tbl/data.frame)
##  $ Lote       : Factor w/ 5 levels "1","2","3","4",..: 1
##  $ Dia        : Factor w/ 5 levels "1","2","3","4",..: 1
##  $ Ingrediente: Factor w/ 5 levels "A","B","C","D",..: 1
##  $ Tempo      : num [1:25] 8 11 4 6 4 7 2 9 8 2 ...
```

Análise exploratória

```
boxplot(dad$Tempo ~ dad$Ingrediente)
```



Estatística descritiva

```r
(media       <- tapply(dad$Tempo,dad$Ingrediente,mean))
```

```
##   A   B   C   D   E
## 8.4 5.6 8.8 3.4 3.2
```

```r
(sd.dad      <- tapply(dad$Tempo,dad$Ingrediente,sd))
```

```
##        A        B        C        D        E
## 1.140175 2.073644 1.643168 2.073644 1.923538
```

```r
(soma.linha  <- tapply(dad$Tempo,dad$Lote,sum))
```

```
##  1  2  3  4  5
## 26 31 29 36 25
```

```r
(soma.coluna <- tapply(dad$Tempo,dad$Dia,sum))
```

```
##  1  2  3  4  5
## 33 28 27 25 34
```

```r
(soma.Trat   <- tapply(dad$Tempo,dad$Ingrediente,sum))
```

```
##  A  B  C  D  E
## 42 28 44 17 16
```

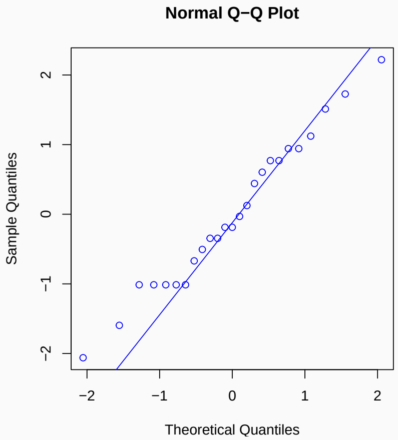
Extraindo resíduos

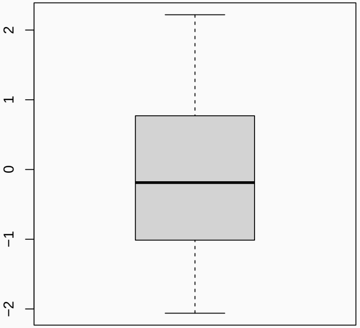
```r
mod1 <- aov(dad$Tempo ~ Lote + Dia + Ingrediente, data=dad)
res.s <- rstudent(mod1)
```

Verificação dos pressupostos básicos

- Análise gráfica
- Testes

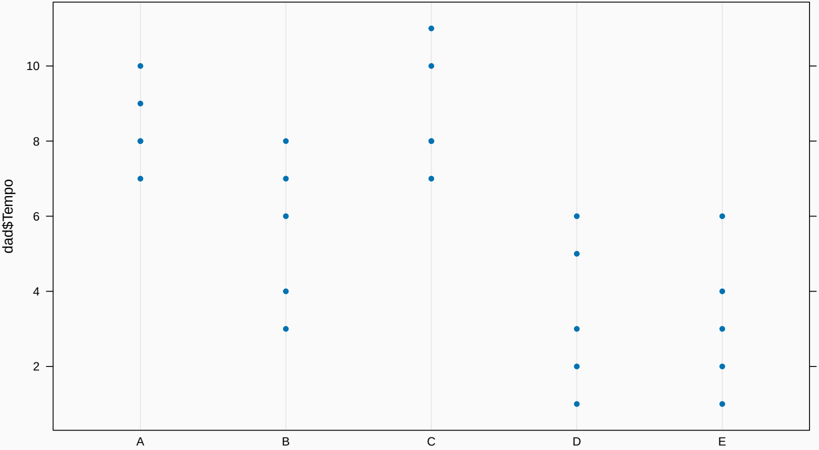
Normalidade



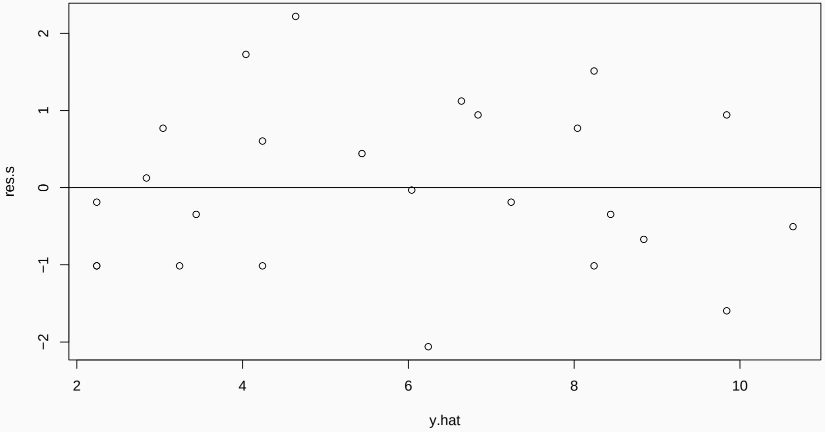


Homogeneidade de variâncias

```
## Carregando pacotes exigidos: lattice
```



Homogeneidade de variâncias



Testes

Teste Shapiro-Wilk para normalidade

```
##
##  Shapiro-Wilk normality test
##
## data:  res.s
## W = 0.97785, p-value = 0.8395
```

Teste Bartlett para homogeneidade de variâncias

```
##
##  Bartlett test of homogeneity of variances
##
## data:  dad$Tempo by dad$Ingrediente
## Bartlett's K-squared = 1.5544, df = 4, p-value = 0.817
```

Teste para independência (?)

```
## Carregando pacotes exigidos: carData

##  lag Autocorrelation D-W Statistic p-value
##    1      -0.1718977      2.319829   0.958
##  Alternative hypothesis: rho != 0
```

Os pressupostos básicos foram verificados.

Quadro ANOVA - Cálculos

```r
(n <- length(dad$Tempo))
```

```
## [1] 25
```

```r
(correcao <- n*mean(dad$Tempo)^2)
```

```
## [1] 864.36
```

```r
(SQTotal  <- sum(t(dad$Tempo)%*%dad$Tempo)-correcao)
```

```
## [1] 206.64
```

```r
(SQLinha  <- sum(soma.linha^2)/5-correcao)
```

```
## [1] 15.44
```

```r
(SQTrat   <- sum(soma.Trat^2)/5-correcao)
```

```
## [1] 141.44
```

```r
(SQColuna <- sum(soma.coluna^2)/5-correcao)
```

```
## [1] 12.24
```

```r
(SQRes    <- SQTotal - SQLinha - SQColuna-SQTrat)
```

```
## [1] 37.52
```

Continuação

```
(SQRes    <- SQTotal - SQLinha - SQColuna-SQTrat)
```

```
## [1] 37.52
```

```
p <- 5
(gl.Trat <- p-1)
```

```
## [1] 4
```

```
(gl.Res  <- (p-2)*(p-1))
```

```
## [1] 12
```

```
(QMTrat <- SQTrat/gl.Trat)
```

```
## [1] 35.36
```

```
(QMRes <- SQRes/gl.Res)
```

```
## [1] 3.126667
```

F calculado e valor p

```
(Fcalc <- QMTrat/QMRes)

## [1] 11.30917

(valor.p <- pf(Fcalc,gl.Trat,gl.Res,lower.tail=F))

## [1] 0.0004876512
```

Conclusão:

Rejeita-se H_0 ao nível de significância $\alpha = 0,05$ ou seja, existem diferenças entre médias de tratamentos.

Resumindo:

```r
anova(mod1)
```

```
## Analysis of Variance Table
##
## Response: dad$Tempo
##             Df Sum Sq Mean Sq F value    Pr(>F)
## Lote         4  15.44   3.860  1.2345 0.3476182
## Dia          4  12.24   3.060  0.9787 0.4550143
## Ingrediente  4 141.44  35.360 11.3092 0.0004877 ***
## Residuals   12  37.52   3.127
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(mod1,intercept = T)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## (Intercept)  1  864.4   864.4 276.448 1.19e-09 ***
## Lote         4   15.4     3.9   1.235 0.347618
## Dia          4   12.2     3.1   0.979 0.455014
## Ingrediente  4  141.4    35.4  11.309 0.000488 ***
## Residuals   12   37.5     3.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Teste de Tukey - Cálculos

```r
layout(1)
graf <- TukeyHSD(mod1,cterms="Ingrediente",order=T)
model.tables(mod1, cterms="Ingrediente",type="means")
```

```
## Tables of means
## Grand mean
##
## 5.88
##
##  Ingrediente
## Ingrediente
##   A   B   C   D   E
## 8.4 5.6 8.8 3.4 3.2
```

```r
(gl.Res <- df.residual(mod1))
```

```
## [1] 12
```

```r
(QMRes  <- sigma(mod1)^2)
```

```
## [1] 3.126667
```

```r
alpha  <- 0.05
a <- 5 # Número de tratamentos
```

continua

```r
names(graf)
```

```
## [1] "Lote"        "Dia"         "Ingrediente"
```

```r
class(graf$Ingrediente)
```

```
## [1] "matrix" "array"
```

```r
graf$Ingrediente[,1]
```

```
## D-E B-E A-E C-E B-D A-D C-D A-B C-B C-A
## 0.2 2.4 5.2 5.6 2.2 5.0 5.4 2.8 3.2 0.4
```

```r
q   <- qtukey(0.95, a, gl.Res)
(dms <- q*sqrt(QMRes/5))
```

```
## [1] 3.564608
```

```r
graf$Ingrediente[,1] > dms
```

```
##   D-E   B-E   A-E   C-E   B-D   A-D   C-D   A-B   C-B   C-A
## FALSE FALSE  TRUE  TRUE FALSE  TRUE  TRUE FALSE FALSE FALSE
```

Teste Tukey

```r
require(agricolae)
```

```
## Carregando pacotes exigidos: agricolae
```

```r
Teste <- HSD.test(mod1,"Ingrediente",group=T)
names(Teste)
```

```
## [1] "statistics" "parameters" "means"      "comparison" "groups"
```

```r
Teste$statistics
```

```
##    MSerror Df Mean       CV      MSD
##   3.126667 12 5.88 30.07208 3.564608
```

```r
Teste$means
```

```
##   dad$Tempo      std r        se Min Max Q25 Q50 Q75
## A       8.4 1.140175 5 0.7907802   7  10   8   8   9
## B       5.6 2.073644 5 0.7907802   3   8   4   6   7
## C       8.8 1.643168 5 0.7907802   7  11   8   8  10
## D       3.4 2.073644 5 0.7907802   1   6   2   3   5
## E       3.2 1.923538 5 0.7907802   1   6   2   3   4
```

Teste Tukey (tente groups=F. Verifique qual é o default)

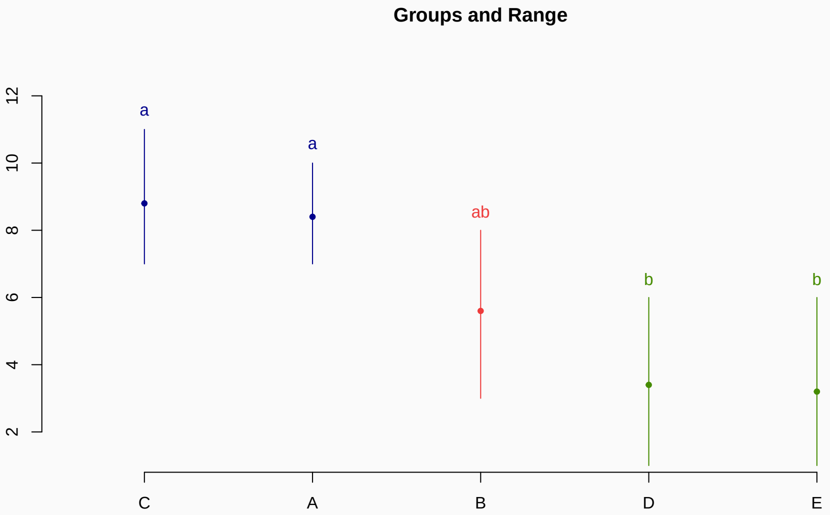
```r
Teste$comparison # Não apresenta porque group=T
```

```
## NULL
```

```r
Teste$groups      # Apresenta porque group=T
```

```
##   dad$Tempo groups
## C       8.8      a
## A       8.4      a
## B       5.6     ab
## D       3.4      b
## E       3.2      b
```

Teste Tukey



Exercício 4.23 Montgomery 8a. Ed. pág. 180

- Um engenheiro industrial está investigando o efeito de quatro métodos de montagem (A, B, C, D) no tempo de montagem para um componente de televisão em cores.
- Quatro operadores são selecionados para o estudo.
- Além disso, o engenheiro sabe que cada método produz tanta fadiga que o tempo necessário para a última montagem pode ser maior que o tempo necessário para a primeiro, independentemente do método. Ou seja, uma tendência se desenvolve no tempo de montagem necessário.
- Para explicar essa fonte de variabilidade, o engenheiro usa o delineamento quadrado latino mostrado na Tabela 01.
- Analise os dados desta experiência (0,05) e tire conclusões apropriadas.

Perguntas?

1) Qual o controle 1 (linhas)?
2) Qual o controle 2 (colunas)?
3) Qual o tratamento?
4) Qual a variável resposta?
5) Qual a unidade da variável resposta?

Respostas

1) Ordem de montagem
2) Operador
3) Método de montagem
4) Tempo de montagem
5) Nada é mencionado! Considerando-se que turno diário é de 8 horas. O tempo deve ser em minutos.

Banco de dados no excel

Ordem	Operador	método	tempo
1	1	C	10
1	2	D	14
1	3	A	7
1	4	B	8
2	1	B	7
2	2	C	18
2	3	D	11
2	4	A	8
3	1	A	5
3	2	B	10
3	3	C	11
3	4	D	9
4	1	D	10
4	2	A	10
4	3	B	12
4	4	C	14

Exercício 01

- fazer a análise do exemplo proposto
    a) Realizando todos os cálculos, interpretando os resultados
    b) Fazer a análise no R

Exercício 02

- Analise os dados da página 159 Montgomery

■ **TABLE 4.9**

Latin Square Design for the Rocket Propellant Problem

	Operators				
Batches of Raw Material	**1**	**2**	**3**	**4**	**5**
1	$A = 24$	$B = 20$	$C = 19$	$D = 24$	$E = 24$
2	$B = 17$	$C = 24$	$D = 30$	$E = 27$	$A = 36$
3	$C = 18$	$D = 38$	$E = 26$	$A = 27$	$B = 21$
4	$D = 26$	$E = 31$	$A = 26$	$B = 23$	$C = 22$
5	$E = 22$	$A = 30$	$B = 20$	$C = 29$	$D = 31$