SME0816 - Planejamento de Experimentos I: Comparações Múltiplas

Profa. Cibele Russo

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

Revisando

```r
require(readxl)
```

```
## Loading required package: readxl
```

```r
dados <- read_xlsx("plasma_pag_62_ed7.xlsx")
str(dados)
```

```
## tibble [20 x 2] (S3: tbl_df/tbl/data.frame)
##  $ Potencia : num [1:20] 160 160 160 160 160 180 180 180 180 180 ...
##  $ Etch.rate: num [1:20] 575 542 530 539 570 565 593 590 579 610 ...
```

```r
dados$Potencia <- as.factor(dados$Potencia)
str(dados)
```

```
## tibble [20 x 2] (S3: tbl_df/tbl/data.frame)
##  $ Potencia : Factor w/ 4 levels "160","180","200",..: 1 1 1 1 1 2 2 2 2 2 ...
##  $ Etch.rate: num [1:20] 575 542 530 539 570 565 593 590 579 610 ...
```

Os dados

```
dados
```

```
## # A tibble: 20 x 2
##    Potencia Etch.rate
##    <fct>        <dbl>
##  1 160            575
##  2 160            542
##  3 160            530
##  4 160            539
##  5 160            570
##  6 180            565
##  7 180            593
##  8 180            590
##  9 180            579
## 10 180            610
## 11 200            600
## 12 200            651
## 13 200            610
## 14 200            637
## 15 200            629
## 16 220            725
## 17 220            700
## 18 220            715
## 19 220            685
## 20 220            710
```

comparando as funções lm e função aov

```r
names(dados)
```

```
## [1] "Potencia"  "Etch.rate"
```

```r
(mod.lm  <- lm(Etch.rate  ~ Potencia, data=dados))
```

```
##
## Call:
## lm(formula = Etch.rate ~ Potencia, data = dados)
##
## Coefficients:
## (Intercept)  Potencia180  Potencia200  Potencia220
##       551.2         36.2         74.2        155.8
```

```r
(mod.aov <- aov(Etch.rate ~ Potencia, data=dados))
```

```
## Call:
##    aov(formula = Etch.rate ~ Potencia, data = dados)
##
## Terms:
##                  Potencia Residuals
## Sum of Squares   66870.55   5339.20
## Deg. of Freedom         3        16
##
## Residual standard error: 18.26746
## Estimated effects may be unbalanced
```

Saídas lm / aov

```r
names(mod.lm)
```

```
##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "contrasts"     "xlevels"       "call"          "terms"
## [13] "model"
```

```r
names(mod.aov)
```

```
##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "contrasts"     "xlevels"       "call"          "terms"
## [13] "model"
```

```r
class(mod.lm)
```

```
## [1] "lm"
```

```r
class(aov)
```

```
## [1] "function"
```

Quadro da ANOVA

```r
anova(mod.lm)
```

```
## Analysis of Variance Table
##
## Response: Etch.rate
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Potencia   3  66871 22290.2  66.797 2.883e-09 ***
## Residuals 16   5339   333.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(mod.aov)
```

```
## Analysis of Variance Table
##
## Response: Etch.rate
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Potencia   3  66871 22290.2  66.797 2.883e-09 ***
## Residuals 16   5339   333.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modelo superparametrizado (restrição default do R: $\tau_1 = 0$)

$$y_{ij} = \mu + \tau_i + e_{ij},\ e_{ij} \sim N(0, \sigma^2)$$

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

y **ajustado:** $\hat{y}_{ij}$

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i$$

e_{ij} **ajustado:** $\hat{e}_{ij}$

$$\hat{e}_{ij} = y_{ij} - \hat{y}_{ij}$$

Estimativa para θ

$$\widehat{\theta} = \begin{bmatrix} \widehat{\mu} \\ \widehat{\tau}_1 \\ \widehat{\tau}_2 \\ \vdots \\ \widehat{\tau}_a \\ \widehat{\sigma}^2 \end{bmatrix}$$

Estimativa para σ^2

```
sigma(mod.lm)^2
```

```
## [1] 333.7
```

```
sigma(mod.aov)^2
```

```
## [1] 333.7
```

```
y.obs <- dados$Etch.rate

# Número de tratamentos
(a <- length(unique(dados$Potencia)))
```

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

```
# Número de observações
(n <- length(y.obs))
```

```
## [1] 20
```

```
sum((y.obs-fitted(mod.lm))^2)/(n-a)
```

```
## [1] 333.7
```

$\hat{\theta}$ do exemplo

$$\hat{\theta} = \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\tau}_3 \\ \hat{\tau}_4 \\ \hat{\sigma}^2 \end{bmatrix} = \begin{bmatrix} 551,2 \\ 0,00 \\ 36,20 \\ 74,20 \\ 155,80 \\ 333,70 \end{bmatrix}$$

$$\hat{\sigma}^2 = QMRes = 333,70 \quad \text{e} \ \text{g. l. res} = 16 = 20 - 4,$$

- QMRes e g.l. res são quantidades necessárias para o Teste de Tukey.

Interpretação

```r
dummy.coef(mod.lm)
```

```
## Full coefficients are
##
## (Intercept):    551.2
## Potencia:         160   180   200   220
##                   0.0  36.2  74.2 155.8
```

```r
tapply(dados$Etch.rate, dados$Potencia,mean)
```

```
##   160   180   200   220
## 551.2 587.4 625.4 707.0
```

```r
mean(dados$Etch.rate)
```

```
## [1] 617.75
```

Interpretação

- μ é a média do tratamento 1: Potência $= 160$;
- $\tau_1 = 0$ pela restrição do modelo;
- $\tau_i, \; i = 2, 3, 4$ o efeito do tratamento i em relação a média do tratamento 1, isto é, quanto a média do tratamento i difere da média do tratamento 1.
- Observe que com a restrição $\tau_1 = 0$ a estimativa para μ não representa a média geral, e sim, uma constante inerente aos dados, nesse caso, $\mu = \overline{y}_{1.}$, a média do tratamento 1.

As funções abaixo podem auxiliá-lo a descrever o conjunto de dados

```
(d1 <- tapply(dados$Etch.rate, dados$Potencia,summary))
```

```
## $`160`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   530.0   539.0   542.0   551.2   570.0   575.0
##
## $`180`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   565.0   579.0   590.0   587.4   593.0   610.0
##
## $`200`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   600.0   610.0   629.0   625.4   637.0   651.0
##
## $`220`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     685     700     710     707     715     725
```

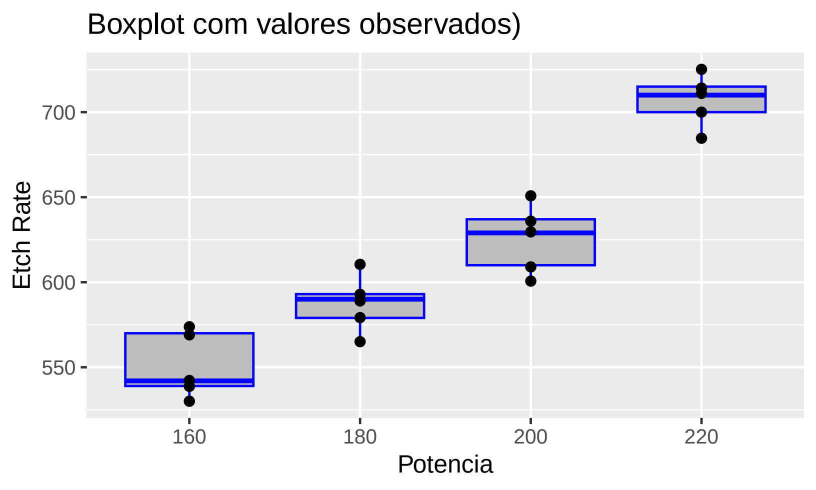
```r
d2 <- tapply(dados$Etch.rate, dados$Potencia,length)
d3 <- tapply(dados$Etch.rate, dados$Potencia,sum)
d4 <- tapply(dados$Etch.rate, dados$Potencia,mean)
d5 <- tapply(dados$Etch.rate, dados$Potencia,var)
d6 <- tapply(dados$Etch.rate, dados$Potencia,sd)
```

```r
data.frame(n=d2,soma=d3,media=d4,
           variancia=d5,desvio.pad=round(d6,1))
```

```
##     n soma media variancia desvio.pad
## 160 5 2756 551.2     400.7       20.0
## 180 5 2937 587.4     280.3       16.7
## 200 5 3127 625.4     421.3       20.5
## 220 5 3535 707.0     232.5       15.2
```

Boxplot com valores observados

```
## Loading required package: ggplot2
```



Análise no R

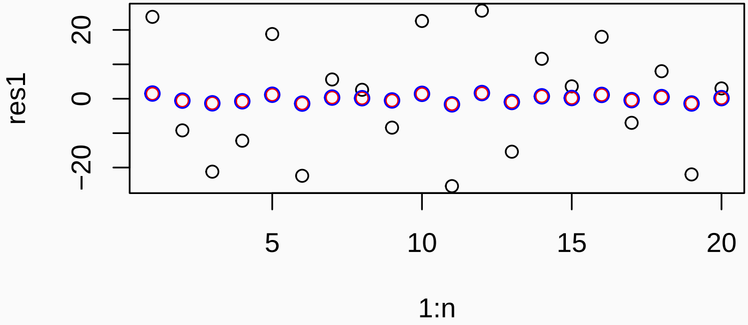
```r
mod1 <- aov(dados$Etch.rate ~ dados$Potencia)
anova(mod1)
## Analysis of Variance Table
##
## Response: dados$Etch.rate
##                Df Sum Sq Mean Sq F value    Pr(>F)
## dados$Potencia  3  66871 22290.2  66.797 2.883e-09 ***
## Residuals      16   5339   333.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Obtenção dos resíduos: ordinário, padronizado, estudentizado

```r
res1  <- mod.lm$res # Resíduo ordinário
res2  <- rstandard(mod.lm) # Resíduos padronizado
res3  <- rstudent(mod.lm)  # resíduo estudentizado
data.frame(res1,res2,res3)
```

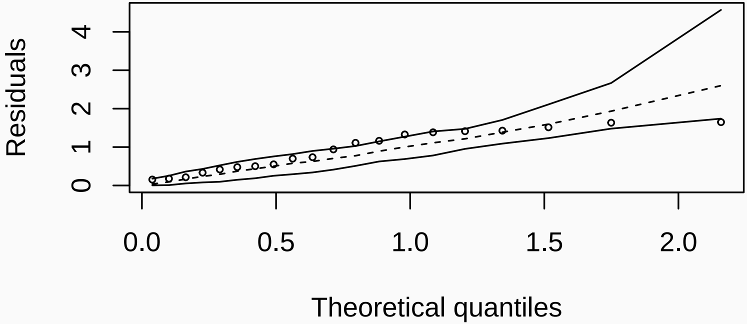
```
##     res1       res2       res3
## 1   23.8  1.4566455  1.5143741
## 2   -9.2 -0.5630730 -0.5506764
## 3  -21.2 -1.2975161 -1.3281303
## 4  -12.2 -0.7466838 -0.7359089
## 5   18.8  1.1506275  1.1632571
## 6  -22.4 -1.3709604 -1.4130125
## 7    5.6  0.3427401  0.3330817
## 8    2.6  0.1591293  0.1541984
## 9   -8.4 -0.5141102 -0.5019482
## 10  22.6  1.3832012  1.4273336
## 11 -25.4 -1.5545712 -1.6336291
## 12  25.6  1.5668119  1.6488129
## 13 -15.4 -0.9425353 -0.9390477
## 14  11.6  0.7099617  0.6985080
## 15   3.6  0.2203329  0.2136608
## 16  18.0  1.1016646  1.1095957
## 17  -7.0 -0.4284251 -0.4172209
## 18   8.0  0.4896287  0.4776731
## 19 -22.0 -1.3464790 -1.3845225
## 20   3.0  0.1836108  0.1779680
```

Graficamente



#

```
## Loading required package: hnp
## Loading required package: MASS
## Gaussian model (lm object)
```

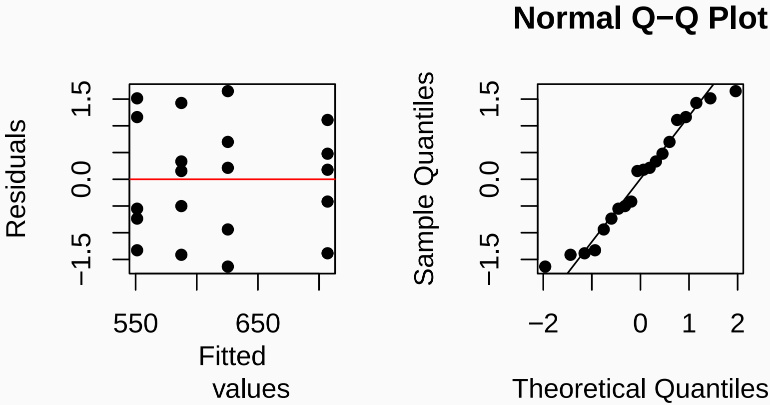


Verificação dos pressupostos básicos

Graficamente

```
##    Potencia taxa   fit     res
## 1       160  575 551.2  1.5144
## 2       160  542 551.2 -0.5507
## 3       160  530 551.2 -1.3281
## 4       160  539 551.2 -0.7359
## 5       160  570 551.2  1.1633
## 6       180  565 587.4 -1.4130
## 7       180  593 587.4  0.3331
## 8       180  590 587.4  0.1542
## 9       180  579 587.4 -0.5019
## 10      180  610 587.4  1.4273
## 11      200  600 625.4 -1.6336
## 12      200  651 625.4  1.6488
## 13      200  610 625.4 -0.9390
## 14      200  637 625.4  0.6985
## 15      200  629 625.4  0.2137
## 16      220  725 707.0  1.1096
## 17      220  700 707.0 -0.4172
## 18      220  715 707.0  0.4777
## 19      220  685 707.0 -1.3845
## 20      220  710 707.0  0.1780
```

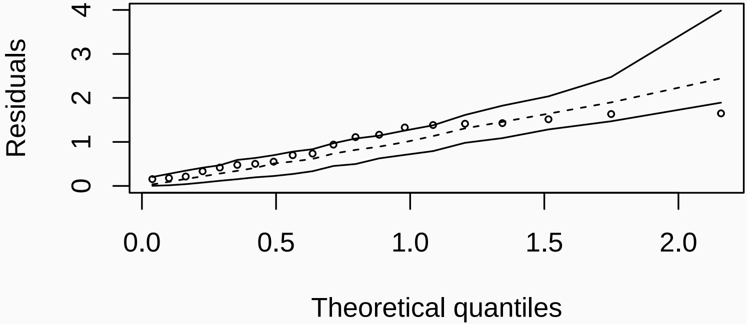
Gráficos



Normal Q−Q Plot
Residuals
1.5
0.0
−1.5
550
650
Fitted values
Sample Quantiles
1.5
0.0
−1.5
−2
0
1
2
Theoretical Quantiles

```r
#```{r,echo=F,eval=T,fig.asp=0.30}
require(hnp)
hnp(mod.lm,resid.type="student",how.many.out=T)
```

```
## Gaussian model (lm object)
```




```
## Total points: 20
## Points out of envelope: 5 ( 25 %)
```

- Idealmente, todos os pontos devem estar dentro do envelope simulado
- Alguns autores sugerem uma tolerância, por exemplo 10% de um conjunto de dados "pequeno" estiverem fora do envelope simulado ou 5% de um conjunto de dados "grande". (Clarice G. B. Demétrio em 03/05/2021).

Verificação dos pressupostos básicos: Homocedasticidade

Teste de Bartlett

```
bartlett.test(Etch.rate ~Potencia, data=dados )
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  Etch.rate by Potencia
## Bartlett's K-squared = 0.43349, df = 3, p-value = 0.933
```

Verificação dos pressupostos básicos: normalidade

Teste de Shapiro-Wilk

```r
shapiro.test(rstudent(mod.lm))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  rstudent(mod.lm)
## W = 0.94514, p-value = 0.2993
```

Quadro da ANOVA

mod = mod_anula1

```r
anova(mod.lm)
```

```
## Analysis of Variance Table
##
## Response: Etch.rate
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Potencia   3  66871 22290.2  66.797 2.883e-09 ***
## Residuals 16   5339   333.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclusão:

- Rejeita-se H_0 no nível de significância $\alpha = 5\%$
- Em média as taxas de gravação diferem entre si.
- Existe diferença entre pelo menos um contraste de dois tratamentos.
- Ou ainda, pelo menos duas taxas de gravação diferem entre si.

- Dando prosseguimento à análise.

Testes de comparações múltiplas

- Ao comparar mais de duas médias, um teste ANOVA informa se as médias são significativamente diferentes uma das outras, mas não informa quais médias diferem umas das outras.
- Quando a hipótese nula,H_0, é rejeitada. Deve-se dar continuidade à análise estatística dos dados utilizando testes de comparações múltiplas para localizar onde ocorreu a diferença. Os **procedimentos de comparação múltipla (MCPs)**, fornecem informações mais detalhadas sobre as diferenças entre os médias.

Métodos de comparação múltipla (MCP, ou MCM)

- Uma variedade de métodos de comparação múltipla estão disponíveis
  - Teste t;
  - Teste de Tukey;
  - Teste de Scheffé;
  - Teste de Duncan;
  - Teste de Dunnet;
  - Correção de Bonferroni;

entre outros.

O que é um contraste de médias?

Um contraste é uma combinação linear dos parâmetros da forma:

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i$$

em que as constantes $c_i, i = 1, \dots, a$ satisfazem

$$\sum_{i=1}^{a} c_i = 0.$$

Exemplo de contrastes

1. $$\begin{aligned} H_0: \quad \mu_3 \quad &= \quad \mu_4 \quad \text{ou equivalentemente,} \\ H_0: \quad \mu_3 - \mu_4 \quad &= \quad 0 \end{aligned}$$

$$c_1 = c_2 = 0, c_3 = 1, c_4 = -1$$

2. $$\begin{aligned} H_0: \quad \mu_1 + \mu_2 \quad &= \quad \mu_3 + \mu_4 \quad \text{ou equivalentemente;} \\ H_0: \quad \mu_1 + \mu_2 - \mu_3 - \mu_4 \quad &= \quad 0 \end{aligned}$$

$$c_1 = c_2 = 1, c_3 = c_4 = -1$$

Teste de Tukey

- é um teste de comparação de contrastes que envolvem duas médias.
- é um teste de comparação de médias duas a duas
- Se o número de tratamentos é igual a "a" tem-se

$$m = \frac{a(a-1)}{2}$$

comparações de médias duas a duas.

O teste de Tukey foi desenvolvido por John Wilder Tukey e apresentado em 1949 no artigo titulado “Comparing Individual Means in the Analysis of Variance” (Biometrics. 5 (2): 99-114. JSTOR 3001913).

- https://en.wikipedia.org/wiki/John_Tukey
- John Wilder Tukey (1915 - 2000) foi um matemático e estatístico americano famoso pelo algoritmo Fast Fourier Transform (FFT) e pelo gráfico boxplot.

Teste de Tukey (TSD - Tukey Significant Difference)

- O Teste proposto por Tukey (1953) é também conhecido como:
  - teste de Tukey da diferença honestamente significativa (honestly significant difference)(HSD) e,
  - teste de Tukey da diferença totalmente significativa (wholly significant difference)(WSD).
  - As hipóteses a serem testadas são:

$$\begin{cases} H_0: & \mu_i & = & \mu_j \\ H_1: & \mu_i & \neq & \mu_j,\ i > j \end{cases}$$

Dados balanceados

- Para tamanhos de amostras iguais (dados balanceados), o teste de Tukey declara duas médias significativamente diferentes se o valor absoluto de suas diferenças amostrais ultrapassar a diferença mínima significativa (DMS) ou Diferença significativa de Tukey (DST) (TSD-Tukey Significant Difference)

Ou seja,

- O procedimento de Tukey utiliza a estatística da amplitude studentizada

$$Q = \frac{|\overline{Y}_i - \overline{Y}_j|}{\sqrt{\frac{QMRes}{k}}}$$

e declara duas médias significativamente diferentes se

$$|\overline{y_{i.}} - \overline{y_{j.}}| > TSD = DMS$$

em que

$$DMS = TSD = q_{\alpha}(a, na - a)\sqrt{\frac{QMRes}{n}},$$

e $q(a, na - a)$ é a amplitude studentizada e a é o número de tratamentos e n é o número de unidades experimentais, ou ainda, $na - a$ é o número de graus de liberdade do resíduo.

- Em outras palavras, rejeita-se a igualdade da média de dois níveis de tratamento se

$$|\overline{y_{i.}} - \overline{y_{j.}}| > DMS$$

.

Intervalo de confiança (dados balanceados)

- Também pode-se rejeitar H_0 quando o intervalo de confiança de $100(1-\alpha)$ para a diferença entre o par de médias correspondente não incluir zero, indicando que o valor p para o teste de igualdade de médias é menor do que $0,05$.

$$\overline{y_{i.}} - \overline{y_{j.}} - q_\alpha(a, na-a)\sqrt{\frac{QMRes}{n}} \leq \mu_i - \mu_j \leq$$

$$\leq \overline{y_{i.}} - \overline{y_{j.}} + q_\alpha(a, na-a)\sqrt{\frac{QMRes}{n}}, \quad i \neq j.$$

Dados desbalanceados

Quando o tamanho das amostras são diferentes (dados não balanceados), o teste de Tukey é modificado e é chamado de Teste de Tukey-Kramer. Esse teste não é exato, mas é minimamente conservativo no sentido em que a FWER real é muitas vezes menor que α.

O teste de Tukey-Kramer declara duas médias significativamente diferentes se o valor absoluto de suas diferenças amostrais ultrapassar a diferença mínima significativa (DMS)

$$DMS = \frac{q_\alpha(a, \sum_{l=1}^{a} n_l - a)}{\sqrt{2}} \sqrt{QMRes \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

Intervalo de confiança (dados desbalanceados)

$$\overline{y_{i.}} - \overline{y_{j.}} - \frac{q_\alpha(a, \sum_{l=1}^{a} n_l - a)}{\sqrt{2}} \sqrt{QMRes\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \leq \mu_i - \mu_j \leq$$

$$\leq \overline{y_{i.}} - \overline{y_{j.}} + \frac{q_\alpha(a, \sum_{l=1}^{a} n_l - a)}{\sqrt{2}} \sqrt{QMRes\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

O teste de Tukey-Kramer também fornece intervalos uniformemente mais curtos que qualquer um do outros de uma etapa para a família de todas as comparações duas a duas.

- O teste de Tukey tem-se mostrado analiticamente ótimo, no sentido que, entre todos os procedimentos que resultam em intervalos de confiança com mesmo tamanho para todas diferenças duas a duas com coeficiente de confiança da família de pelo menos $1 - \alpha$, o teste de Tukey resulta em intervalos menores.

- Isso quer dizer que, se a família consiste em todas comparações duas a duas e o teste de Tukey pode ser usado, ele resultará em intervalos menores que qualquer outro método de comparação múltipla de uma etapa.

Exemplo: Plasma

- $a = 4$, $trat = Potencia = 160, 180, 200, 220$
- $m = \frac{a(a-1)}{2} = \frac{4(4-1)}{2} = 6$ comparações, a saber:

	μ_{160}	μ_{180}	μ_{200}
μ_{180}	$\lvert\mu_{180} - \mu_{160}\rvert$		
μ_{200}	$\lvert\mu_{200} - \mu_{160}\rvert$	$\lvert\mu_{200} - \mu_{180}\rvert$	
μ_{220}	$\lvert\mu_{220} - \mu_{160}\rvert$	$\lvert\mu_{220} - \mu_{180}\rvert$	$\lvert\mu_{220} - \mu_{200}\rvert$

Teste Tukey passo a passo (dados balanceados)

```r
(mean.trat <- tapply(dados$Etch.rate, dados$Potencia,mean))
```

```
##   160   180   200   220
## 551.2 587.4 625.4 707.0
```

```r
mat.comp <-matrix(c(
         mean.trat[2]-mean.trat[1],
         mean.trat[3]-mean.trat[1],
         mean.trat[4]-mean.trat[1],
         0,
         mean.trat[3]-mean.trat[2],
         mean.trat[4]-mean.trat[2],
         0,
         0,
          mean.trat[4]-mean.trat[3]),3,3)
row.names(mat.comp) <- c(180,200,220)
colnames(mat.comp)  <- c(160,180,200)
```

Teste Tukey passo a passo (dados balanceados)

```
mat.comp
```

```
##       160   180  200
## 180  36.2   0.0  0.0
## 200  74.2  38.0  0.0
## 220 155.8 119.6 81.6
```

O teste de Tukey passo a passo (DMS)

```
a <- 4 # Número de tratamentos
k <- 5 # Número de repetições por tratamento (dados balanceados)
n <- length(dados$Etch.rate) # Número de unidades experimentais

alpha <- 0.05 # Nível de significância do teste

qtukey(1-alpha,a,n-a)   # Amplitude studentizada
```

```
## [1] 4.046093
```

```
(QMRes <- sigma(mod.lm)^2)
```

```
## [1] 333.7
```

```
(DMS <- qtukey(1-alpha,4,n-4)*sqrt(QMRes/k))
```

```
## [1] 33.05438
```

O teste de Tukey passo a passo

```r
mat.comp
```

```
##       160   180  200
## 180  36.2   0.0  0.0
## 200  74.2  38.0  0.0
## 220 155.8 119.6 81.6
```

```r
DMS
```

```
## [1] 33.05438
```

Conclusão

```r
abs(mat.comp) > DMS
```

```
##       160   180   200
## 180 TRUE FALSE FALSE
## 200 TRUE  TRUE FALSE
## 220 TRUE  TRUE  TRUE
```

- Conclui-se que existem diferenças entre os pares de médias de potência:

$$
\begin{array}{ccc}
(\mu_1, \mu_2) & (\mu_1, \mu_3) & (\mu_1, \mu_4) \\
(\mu_2, \mu_3) & (\mu_2, \mu_4) & (\mu_3, \mu_4)
\end{array}
$$

ao nível de significância $\alpha = 5\%$.

Resumindo:

```r
data.frame(medias=sort(mean.trat,decreasing=T),
           Letras=c("a","b","c","d"))
```

```
##     medias Letras
## 220  707.0      a
## 200  625.4      b
## 180  587.4      c
## 160  551.2      d
```

Letras diferentes implicam em tratamentos diferentes.

No R

- No R existem vários pacotes para realizar comparações múltiplas cada um com suas particularidades
    - agricolae
    - DescTools
    - multcomp
    - multicompView
    - ExpDes
    - ExpDes.pt
    - entre outros

Teste de comparação de médias no R

Teste de Tukey (Direto)

- sem necessidade de pacotes específicos

```
k <- 5 # Número de repetições
mod.Tukey <- aov(Etch.rate ~ Potencia-1, data=dados )
model.tables(mod.Tukey, type="means")
```

```
## Tables of means
##
##  Potencia
## Potencia
##   160   180   200   220
## 551.2 587.4 625.4 707.0
```

```
mod.Tukey$df
```

```
## [1] 16
```

```
(q <- qtukey(0.95, a, mod.Tukey$df))
```

```
## [1] 4.046093
```

```
(w <- qtukey(0.95, a, mod.Tukey$df)*sigma(mod.Tukey)/sqrt(k))
```

```
## [1] 33.05438
```

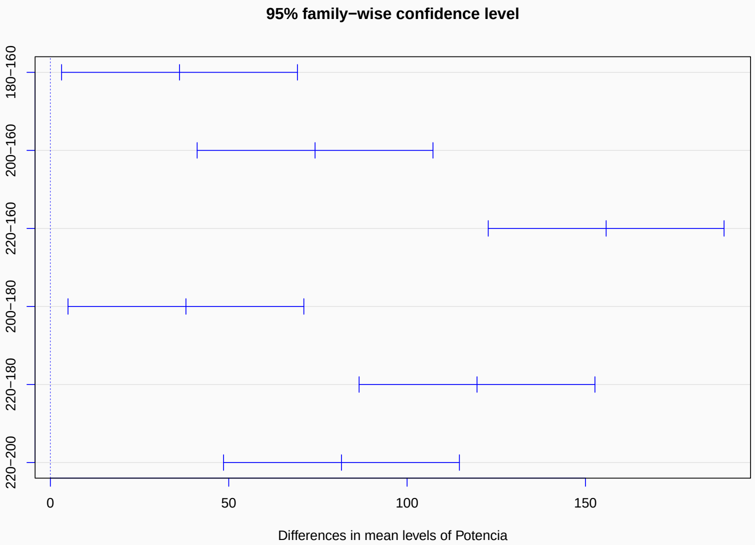
Teste de Tukey

```r
TukeyHSD(mod.Tukey, "Potencia", ordered = TRUE)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = Etch.rate ~ Potencia - 1, data = dados)
##
## $Potencia
##          diff        lwr       upr     p adj
## 180-160  36.2   3.145624  69.25438 0.0294279
## 200-160  74.2  41.145624 107.25438 0.0000455
## 220-160 155.8 122.745624 188.85438 0.0000000
## 200-180  38.0   4.945624  71.05438 0.0215995
## 220-180 119.6  86.545624 152.65438 0.0000001
## 220-200  81.6  48.545624 114.65438 0.0000146
```

Graficamente,

```r
plot(TukeyHSD(mod.Tukey, "Potencia"),col="blue")
```



Teste de Tukey é muito simples no R

- Observe que é necessário utilizar a função **aov**

```
modelo <- aov(Etch.rate ~ Potencia ,data=dados )
tukey<-TukeyHSD(modelo)
#TukeyHSD() package=base R objeto ajuste aov
tukey
```

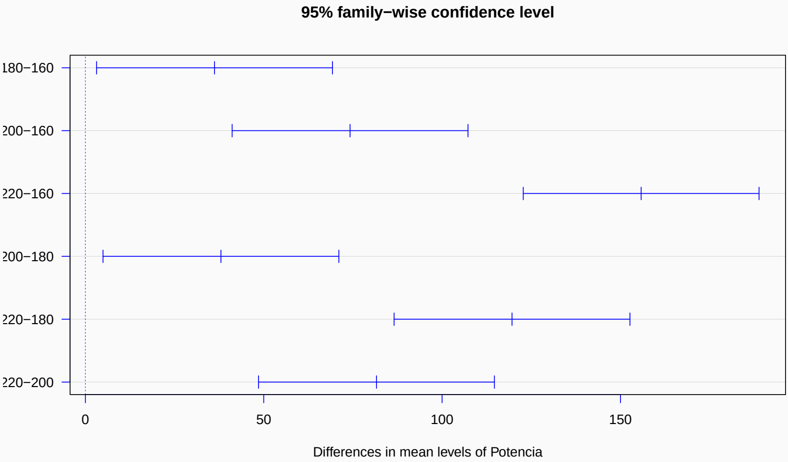
```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Etch.rate ~ Potencia, data = dados)
##
## $Potencia
##          diff        lwr       upr     p adj
## 180-160  36.2   3.145624  69.25438 0.0294279
## 200-160  74.2  41.145624 107.25438 0.0000455
## 220-160 155.8 122.745624 188.85438 0.0000000
## 200-180  38.0   4.945624  71.05438 0.0215995
## 220-180 119.6  86.545624 152.65438 0.0000001
## 220-200  81.6  48.545624 114.65438 0.0000146
```

```
mat.comp
```

```
##       160   180  200
## 180  36.2   0.0  0.0
## 200  74.2  38.0  0.0
## 220 155.8 119.6 81.6
```

Gráfico sem necessidade de pacotes especiais

```
plot(tukey,las=1,col="blue")
```



Tukey com letras e gráficos

package=agricolae

- Observe as saídas

```
require(agricolae)
```

```
## Loading required package: agricolae
```

```
out <- HSD.test(modelo, "Potencia",console=F)
names(out)
```

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

Estatísticas

```
print(out$statistics)
```

```
##   MSerror Df   Mean       CV      MSD
##     333.7 16 617.75 2.957095 33.05438
```

em que

- MSerror = QMRes
- Df = glres
- Mean = média geral (média da variável resposta)
- CV = coeficiente de variação
- MSD = DMS = diferença mínima significativa

Coeficiente de variação

- O coeficiente de variação do modelo (CV) é definido por

$$CV = \frac{\sqrt{QMRes}}{\overline{y}} 100\%$$

No exemplo do plasma:

$$CV = \frac{s}{\overline{y}} = \frac{18,26746}{617,75} \approx 2,96\%$$

```
(CV=sigma(modelo)/mean(dados$Etch.rate))
```

```
## [1] 0.02957095
```

Interpretação do coeficiente de variação

- O coeficiente de variação mede o quanto da variabilidade dos dados não foi explicada pelo modelo
- a variabilidade residual dos dados como uma porcentagem da média da variável resposta.

R-squared (R^2)

$$R^2 = \frac{SQTrat}{SQT}$$

é a proporção da variabilidade dos dados "explicada" pelo tratamento.

Assim no exemplo do plasma etching o fator (tratamento=potência) explica em torno de 93% da variabilidade da variável resposta etch.rate.

- No exemplo do plasma:

$$R^2 = \frac{SQTrat}{SQRes} = \frac{66871}{66871 + 5339} \approx 0,93$$

Quadro ANOVA para verificação dos valores

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

```
##              Df  Sum Sq Mean Sq F value   Pr(>F)    
## (Intercept)   1 7632301 7632301 22871.7  < 2e-16 ***
## Potencia      3   66871   22290    66.8 2.88e-09 ***
## Residuals    16    5339     334                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

Parâmetros

```
print(out$parameters)
```

```
##    test   name.t ntr StudentizedRange alpha
##   Tukey Potencia   4         4.046093  0.05
```

- StudentizedRange = amplitude studentizada tabelada

```
a <- 4
(n <- length(dados$Etch.rate))
```

```
## [1] 20
```

```
(q.tab <- qtukey(0.95,a,n-a))
```

```
## [1] 4.046093
```

```
glRes <- 16  # n-a
```

Lembre-se

- $a = 4$ fatores=tratamento=potência
- $n = 20$ número de unidades amostrais
- $n - a = 16$ graus de liberdade do resíduo (glRes)

Estatísticas descritivas

```
print(out$means)
```

```
##     Etch.rate      std r       se Min Max Q25 Q50 Q75
## 160     551.2 20.01749 5 8.169455 530 575 539 542 570
## 180     587.4 16.74216 5 8.169455 565 610 579 590 593
## 200     625.4 20.52559 5 8.169455 600 651 610 629 637
## 220     707.0 15.24795 5 8.169455 685 725 700 710 715
```

Tukey com letras

Tratamentos com mesma letra não diferem entre si

```
print(out$groups)
```

```
##     Etch.rate groups
## 220     707.0      a
## 200     625.4      b
## 180     587.4      c
## 160     551.2      d
```

Veja a diferença entre group=T e group=F

```
out <- HSD.test(modelo, "Potencia",group=F,console=F)
print(out$comparison)
```

```
##           difference pvalue signif.        LCL         UCL
## 160 - 180      -36.2 0.0294       *  -69.25438   -3.145624
## 160 - 200      -74.2 0.0000     *** -107.25438  -41.145624
## 160 - 220     -155.8 0.0000     *** -188.85438 -122.745624
## 180 - 200      -38.0 0.0216       *  -71.05438   -4.945624
## 180 - 220     -119.6 0.0000     *** -152.65438  -86.545624
## 200 - 220      -81.6 0.0000     *** -114.65438  -48.545624
```

```
out1 <- HSD.test(modelo, "Potencia",group=F,console=F)
print(out1$comparison)
```

```
##           difference pvalue signif.        LCL         UCL
## 160 - 180      -36.2 0.0294       *  -69.25438   -3.145624
## 160 - 200      -74.2 0.0000     *** -107.25438  -41.145624
## 160 - 220     -155.8 0.0000     *** -188.85438 -122.745624
## 180 - 200      -38.0 0.0216       *  -71.05438   -4.945624
## 180 - 220     -119.6 0.0000     *** -152.65438  -86.545624
## 200 - 220      -81.6 0.0000     *** -114.65438  -48.545624
```

Comparação saída das funções

1. out1$comparison
2. TukeyHSD

Saída out1$comparison

```r
out1$comparison
```

```
##           difference pvalue signif.        LCL         UCL
## 160 - 180      -36.2 0.0294       *  -69.25438   -3.145624
## 160 - 200      -74.2 0.0000     *** -107.25438  -41.145624
## 160 - 220     -155.8 0.0000     *** -188.85438 -122.745624
## 180 - 200      -38.0 0.0216       *  -71.05438   -4.945624
## 180 - 220     -119.6 0.0000     *** -152.65438  -86.545624
## 200 - 220      -81.6 0.0000     *** -114.65438  -48.545624
```

Saída TukeyHSD

```r
modelo <- aov(Etch.rate ~ Potencia ,data=dados )
TukeyHSD(modelo)
```

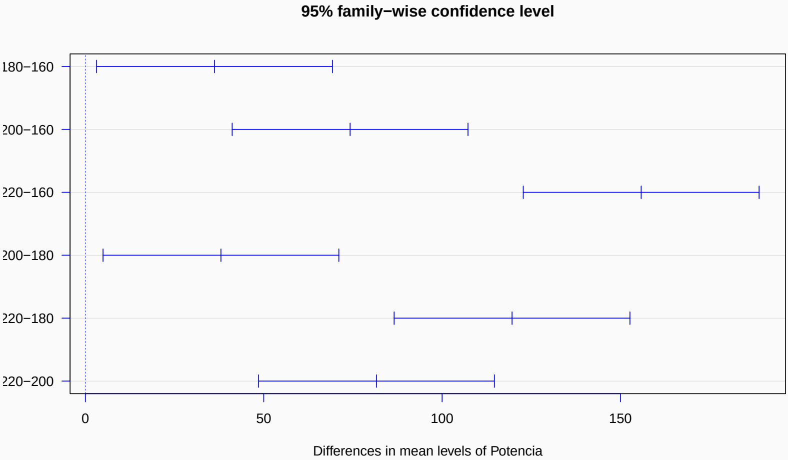
```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Etch.rate ~ Potencia, data = dados)
##
## $Potencia
##          diff        lwr       upr     p adj
## 180-160  36.2   3.145624  69.25438 0.0294279
## 200-160  74.2  41.145624 107.25438 0.0000455
## 220-160 155.8 122.745624 188.85438 0.0000000
## 200-180  38.0   4.945624  71.05438 0.0215995
## 220-180 119.6  86.545624 152.65438 0.0000001
## 220-200  81.6  48.545624 114.65438 0.0000146
```

Diferenças?

- Você observa diferenças entre utilizar
  - TukeyHSD
  - agricolae: HSD.test
  - Quais?
  - Por quê isso ocorre?

Gráfico com package=base

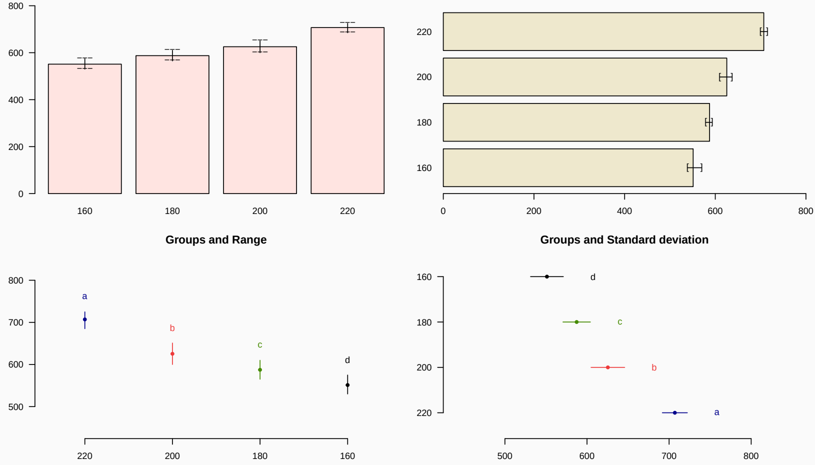
```
plot(TukeyHSD(modelo),las=1,col="blue")
```



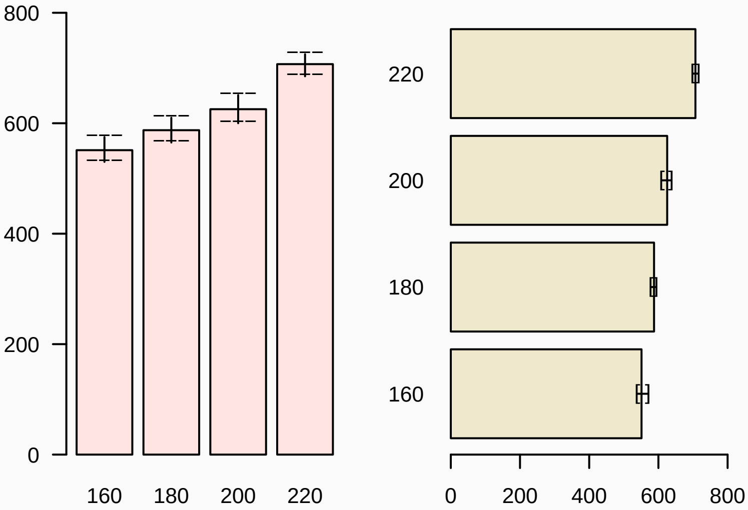
Gráficos com package=Agricolae

```
out <- HSD.test(modelo, "Potencia",group=T,console=F)
oldpar<-par(mfrow=c(2,2),mar=c(3,3,2,1),cex=0.7)
 c1<-colors()[480]; c2=colors()[65]
 bar.err(out$means, variation="range",ylim=c(0,800),col=c1,las=1)
 bar.err(out$means, variation="IQR",horiz=TRUE, xlim=c(0,800),col=c2,las=1)
 plot(out, variation="range",las=1)
 plot(out, horiz=TRUE, variation="SD",las=1)
 par(oldpar)
```

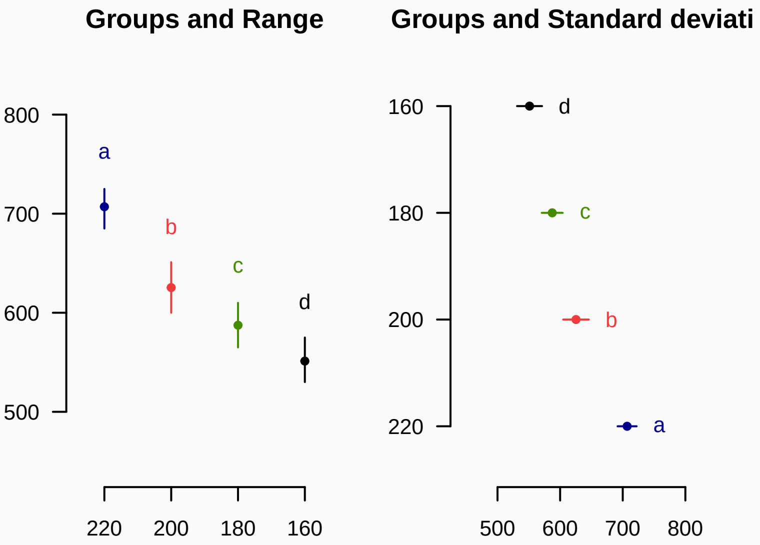
Gráficos com package=Agricolae



Gráficos com package=Agricolae



Gráficos com package=Agricolae



Para ajudar a interpretar os gráficos

```r
tapply(dados$Etch.rate, dados$Potencia,mean)
```

```
##   160   180   200   220
## 551.2 587.4 625.4 707.0
```

```r
tapply(dados$Etch.rate, dados$Potencia,sd)
```

```
##      160      180      200      220
## 20.01749 16.74216 20.52559 15.24795
```

Para ajudar a interpretar os gráficos

```
tapply(dados$Etch.rate, dados$Potencia,range)
```

```
## $`160`
## [1] 530 575
##
## $`180`
## [1] 565 610
##
## $`200`
## [1] 600 651
##
## $`220`
## [1] 685 725
```

Obtenção dos intervalos de confiança

Diferenças de média

```
coeff1 <- c(-1,1,0,0)
coeff2 <- c(-1,0,1,0)
coeff3 <- c(-1,0,0,1)
coeff4 <- c(0,-1,1,0)
coeff5 <- c(0,-1,0,1)
coeff6 <- c(0,0,-1,1)


dif21 <- sum(coeff1*mean.trat)
dif31 <- sum(coeff2*mean.trat)
dif41 <- sum(coeff3*mean.trat)
dif32 <- sum(coeff4*mean.trat)
dif42 <- sum(coeff5*mean.trat)
dif43 <- sum(coeff6*mean.trat)

(dif.media <- c(dif21,dif31,dif41,dif32,dif42,dif43))
```

```
## [1]  36.2  74.2 155.8  38.0 119.6  81.6
```

```r
int.conf <- cbind(dif.media - DMS, dif.media+DMS)

resp <- data.frame(tukey.teste=
              round(tukey$Potencia[,2:3],2),
            int.conf=round(int.conf,2))
colnames(resp) <- c("tukey.inf","tukey.sup","inf","sup")
resp
```

```
##         tukey.inf tukey.sup    inf    sup
## 180-160      3.15     69.25   3.15  69.25
## 200-160     41.15    107.25  41.15 107.25
## 220-160    122.75    188.85 122.75 188.85
## 200-180      4.95     71.05   4.95  71.05
## 220-180     86.55    152.65  86.55 152.65
## 220-200     48.55    114.65  48.55 114.65
```

Obtenção do valor p ajustado

```
valor.p.ajust <- ptukey(dif.media/sqrt(QMRes/k),a,glRes,
                         nranges=1,lower.tail=F)

valor.p.ajust.r <- tukey$Potencia[,4]

data.frame(valor.p.ajust   = round(valor.p.ajust,4) ,
           valor.p.ajust.r = round(valor.p.ajust.r,4) )
```

```
##         valor.p.ajust valor.p.ajust.r
## 180-160        0.0294          0.0294
## 200-160        0.0000          0.0000
## 220-160        0.0000          0.0000
## 200-180        0.0216          0.0216
## 220-180        0.0000          0.0000
## 220-200        0.0000          0.0000
```

3.27. Four catalysts that may affect the concentration of one component in a three-component liquid mixture are being investigated. The following concentrations are obtained from a completely randomized experiment:

Catalyst			
1	2	3	4
58.2	56.3	50.1	52.9
57.2	54.5	54.2	49.9
58.4	57.0	55.4	50.0
55.8	55.3		51.7
54.9			

(**a**) Do the four catalysts have the same effect on the concentration?

(**b**) Analyze the residuals from this experiment.

(**c**) Construct a 99 percent confidence interval estimate of the mean response for catalyst 1.

Exercício

- Considerando-se do exercício 3.27, pag. pdf. 135, do livro do Montgomery referente a catalisadores.
    1. Os catalisadores têm o mesmo efeito sobre a concentração?
    2. Se houver diferença de efeito aplique o Teste de Tukey para fazer comparação entre os pares de médias das concentrações.
    3. Quais suas conclusões?
- OBS: Para as comparações múltiplas interprete o valor-p, os intervalos de confiança, resuma os resultados graficamente e utilizando letras.

Teste de Scheffé

As hipóteses nula e alternativa podem ser expressas como:

$H_0 : \Gamma_u = \sum_{i=1}^{a} c_i \mu_i = 0, u = 1, \ldots, m$

$H_1 : \Gamma_u \neq 0, u = 1, \ldots, m$

Para calcular os contrastes, podemos usar uma combinação linear das médias amostrais:

$$C = c_1 \bar{Y}_1 + c_2 \bar{Y}_2 + \ldots + c_a \bar{Y}_a$$

Onde $\bar{Y}_i$ é a média da amostra i e $c_1, c_2, \ldots, c_a$ são coeficientes que somam zero.

Cálculo da estatística de teste de Scheffé:
A estatística de teste F de Scheffé é calculada como:

$$F = \frac{C^2}{MSE}$$

Onde C^2 é a soma dos quadrados dos contrastes e MSE é a média dos quadrados dos erros.

Teste de Scheffé

Decisão

O valor crítico é selecionado a partir da distribuição F com $df_1 = k - 1$ e $df_2 = df_{\text{error}}$, onde k é o número de grupos e df_{error} são os graus de liberdade do erro.

Se F for maior que o valor crítico, rejeitamos H_0, indicando diferenças significativas entre as médias dos grupos.

Teste de Scheffé

Comparações múltiplas
Se a hipótese nula for rejeitada, o teste de Scheffé pode ser usado para realizar comparações múltiplas entre as médias dos grupos, controlando o erro tipo I.

Interpretação dos resultados:
As diferenças significativas entre os grupos podem ser identificadas e interpretadas com base nos intervalos de confiança para as diferenças entre médias ajustados pelo teste de Scheffé.

- https://eaulas.usp.br/portal/video.action?idItem=17687
- https://eaulas.usp.br/portal/video?idItem=30424
- https://www.uel.br/projetos/experimental/pages/arquivos/Dic_patol.html
- https://www.statology.org/scheffe-test-in-r/

Teste de Dunnet

- Comparações múltiplas com um controle
- Vídeo recomendado: https://eaulas.usp.br/portal/video?idItem=18543 Prof. César Gonçalves de Lima (FZEA/USP).
- Dunnett C. W. (1955) A multiple comparison procedure for comparing several treatments with a control, Journal of the American Statistical Association, 50:1096-1121.

```r
library(DescTools)

## Hollander & Wolfe (1973), 116.
## Mucociliary efficiency from the rate of removal of dust in no
##  subjects, subjects with obstructive airway disease, and subj
##  with asbestosis.
x <- c(2.9, 3.0, 2.5, 2.6, 3.2) # normal subjects
y <- c(3.8, 2.7, 4.0, 2.4)      # with obstructive airway diseas
z <- c(2.8, 3.4, 3.7, 2.2, 2.0) # with asbestosis
```

Controle x

```
DunnettTest(list(x, y, z))
```

```
##
##   Dunnett's test for comparing several treatments with a cont
##     95% family-wise confidence level
##
## $`1`
##        diff     lwr.ci    upr.ci   pval
## 2-1  0.385 -0.6901497 1.4601497 0.5832
## 3-1 -0.020 -1.0336608 0.9936608 0.9982
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

Controle y

```
DunnettTest(list(y,x, z))
```

```
##
##   Dunnett's test for comparing several treatments with a cont
##     95% family-wise confidence level
##
## $`1`
##        diff    lwr.ci    upr.ci   pval
## 2-1 -0.385 -1.453557 0.6835572 0.5717
## 3-1 -0.405 -1.473557 0.6635572 0.5421
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

Controle z

```
DunnettTest(list(z,x,y))
```

```
##
##   Dunnett's test for comparing several treatments with a cont
##     95% family-wise confidence level
##
## $`1`
##      diff     lwr.ci   upr.ci   pval
## 2-1 0.020 -0.9936608 1.033661 0.9982
## 3-1 0.405 -0.6701497 1.480150 0.5535
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

De forma equivalente,

```r
x <- c(x, y, z)

g <- factor(rep(1:3, c(5, 4, 5)),
            labels = c("Normal subjects",
                       "Subjects with obstructive airway disease
                       "Subjects with asbestosis"))
```

continua

```r
DunnettTest(x, g)
```

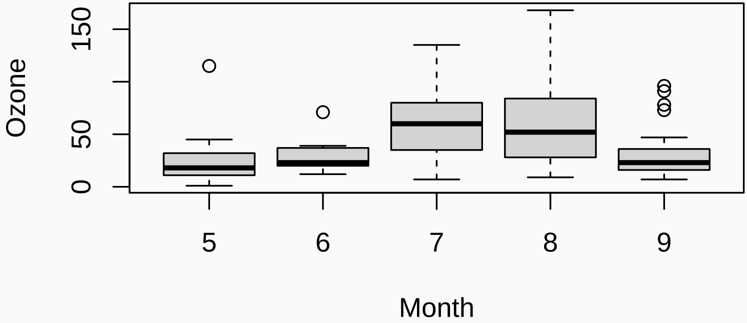
```
##
##   Dunnett's test for comparing several treatments with a control :
##     95% family-wise confidence level
##
## $`Normal subjects`
##                                                              diff     lwr.ci
## Subjects with obstructive airway disease-Normal subjects  0.385 -0.6901497
## Subjects with asbestosis-Normal subjects                 -0.020 -1.0336608
##                                                              upr.ci   pval
## Subjects with obstructive airway disease-Normal subjects 1.4601497 0.5832
## Subjects with asbestosis-Normal subjects                 0.9936608 0.9982
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Equivalentemente,

```r
## Formula interface
layout(1)
str(airquality)
```

```
## 'data.frame':    153 obs. of  6 variables:
##  $ Ozone  : int  41 36 12 18 NA 28 23 19 8 NA ...
##  $ Solar.R: int  190 118 149 313 NA NA 299 99 19 194 ...
##  $ Wind   : num  7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1
##  $ Temp   : int  67 72 74 62 56 66 65 59 61 69 ...
##  $ Month  : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ Day    : int  1 2 3 4 5 6 7 8 9 10 ...
```

```r
boxplot(Ozone ~ Month, data = airquality)
```



```r
DunnettTest(Ozone ~ Month, data = airquality)
```

```
##
##   Dunnett's test for comparing several treatments with a cont
##     95% family-wise confidence level
##
## $`5`
##          diff    lwr.ci   upr.ci    pval
## 6-5  5.829060 -22.43792 34.09604 0.96467
## 7-5 35.500000  15.22870 55.77130 0.00011 ***
## 8-5 36.346154  16.07486 56.61745 8.7e-05 ***
## 9-5  7.832891 -11.90719 27.57297 0.73509
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

```r
DunnettTest(Ozone ~ Month, data = airquality, control="8", conf.
```

```
##
##   Dunnett's test for comparing several treatments with a cont
##     90% family-wise confidence level
##
## $`8`
##            diff    lwr.ci     upr.ci    pval
## 5-8 -36.3461538 -54.26925 -18.423062 6.4e-05 ***
## 6-8 -30.5170940 -55.50965  -5.524537  0.0300 *
## 7-8  -0.8461538 -18.76925  17.076938  0.9999
## 9-8 -28.5132626 -45.96667 -11.059853  0.0019 **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
```

Teste de Dunnet

- Dunnett, C. W. (1955). A multiple comparisons procedure for comparing several treatments with a control. Journal of the American Statistical Association. 50 (272): 1096–1121. doi:10.1080/01621459.1955.10501294
- Dunnett, C. W. (1964). New tables for multiple comparisons with a control. Biometrics. 20 (3): 482–491. JSTOR 2528490. doi:10.2307/2528490

Teste de Duncan

- https://soniavieira.blogspot.com/2016/09/teste-de-amplitudes-multiplas-de-duncan.html

Teste de Duncan no R

- Entrada de dados

```
y <- c(25,17,27,21,15,
       10,-2,12, 4,16,
       18, 8, 4,14, 6,
       23,29,25,35,33,
       11,23, 5,17, 9,
        8,-6, 6, 0, 2)

trat <- as.factor(rep(1:6,each=5))
```

Dados

```r
(dados  <- data.frame(y=y,trat=trat))
```

Normalidade

```
mod <- aov(y ~ trat, data=dados )
shapiro.test(rstudent(mod))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  rstudent(mod)
## W = 0.96257, p-value = 0.3598
```

Homogeneidade de variâncias

```r
bartlett.test(y ~ trat, data=dados )
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  y by trat
## Bartlett's K-squared = 0.85057, df = 5, p-value = 0.9737
```

Interpretação Quadro ANOVA

```
anova(mod)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## trat       5 2354.2  470.83  13.079 3.318e-06 ***
## Residuals 24  864.0   36.00
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

- Como o valor de F é significante ao nível de 5%, existe pelo menos uma média diferente das demais.

Pairwise t.test

```r
attach(airquality)
Month <- factor(Month, labels = month.abb[5:9])
pairwise.t.test(Ozone, Month)
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Ozone and Month
##
##     May     Jun     Jul     Aug
## Jun 1.00000 -       -       -
## Jul 0.00026 0.05113 -       -
## Aug 0.00019 0.04987 1.00000 -
## Sep 1.00000 1.00000 0.00488 0.00388
##
## P value adjustment method: holm
```

Pairwise t.test

```r
pairwise.t.test(Ozone, Month, p.adjust.method = "bonf")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Ozone and Month
##
##     May     Jun     Jul     Aug
## Jun 1.00000 -       -       -
## Jul 0.00029 0.10225 -       -
## Aug 0.00019 0.08312 1.00000 -
## Sep 1.00000 1.00000 0.00697 0.00485
##
## P value adjustment method: bonferroni
```

```r
pairwise.t.test(Ozone, Month, pool.sd = FALSE)
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  Ozone and Month
##
##     May     Jun     Jul     Aug
## Jun 1.00000 -       -       -
## Jul 0.00026 0.01527 -       -
## Aug 0.00195 0.02135 1.00000 -
## Sep 0.86321 1.00000 0.00589 0.01721
##
## P value adjustment method: holm
```

```r
detach()
```

Teste de comparações múltiplas no R

Biblioteca ExpDes

Obtendo os resultados com o pacote ExpDes

```
require(ExpDes)
require(doBy)
```

```r
with(dados , crd(treat=trat, resp=y, mcomp="tukey")) # Tukey
with(dados , crd(treat=trat, resp=y, mcomp="duncan")) # Duncan
with(dados , crd(treat=trat, resp=y, mcomp="snk")) # Student-Newman-Keuls
with(dados , crd(treat=trat, resp=y, mcomp="lsd")) # t-Student (LSD)
with(dados , crd(treat=trat, resp=y, mcomp="lsdb")) # t-Student / Bonferroni
with(dados , crd(treat=trat, resp=y, mcomp="sk")) # Scott-Knott - sob balanceam
```

Biblioteca DescTools

https://www.statology.org/scheffe-test-in-r/

Biblioteca Agricolae

- https://www.uel.br/projetos/experimental/pages/arquivos/Dic_patol.html

Um teste único

	Decisão	
Situação	Rejeito H_0	Não rejeito H_0
H_0 é V	Erro Tipo I	Ok
H_0 é F	Ok	Erro Tipo II

- α é a probabilidade de cometer o erro tipo I

$$Prob[\text{Rejeitar } H_0 | H_0 \text{é verdadeira}]$$

Tabela (Múltiplos testes)

	Decisão		
Situação	Rejeito H_0	Não rejeito H_0	Total
H_0 é V	V	U	m_0
H_0 é F	S	T	m -m_0
	R	m-R	m

- Rejeito H_0 se o teste for declarado significativo,
- Não rejeito H_0 se o teste for declarado não significativo.
- $\alpha = Prob[\text{Rejeitar } H_0 | H_0 \text{é verdadeira}]$
- $1 - \alpha = Prob[\text{Não Rejeitar } H_0 | H_0 \text{é verdadeira}]$

- A probablidade global (Familywise Error), $FWER$, é a probabilidade de cometer pelo menos um erro do Tipo I na família de testes, ou seja,

$$FWER = P[V \geq 1] = 1 - P[V = 0] \leq 1 - (1 - \alpha)^m$$

Visto que, $P[A|B] + P[A^C|B] = 1$ e os testes não são independentes.

Múltiplos testes

Classificação de testes de hipóteses múltiplas

- Testando várias hipóteses nulas: Suponha que temos um número m de hipóteses nulas, denotado por: $H_1, H_2, \ldots, H_m$.
- Usando um teste estatístico, rejeitamos a hipótese nula se o teste for declarado significativo.
- Não rejeitamos a hipótese nula se o teste não for significativo.
- A soma de cada tipo de resultado sobre todas as H_i produz as seguintes variáveis aleatórias:

- m é o número total de hipóteses testadas (conhecido)
- m_0 é o número de hipóteses nulas verdadeiras, um parâmetro desconhecido
- $m - m_0$ é o número de hipóteses alternativas verdadeiras, um parâmetro desconhecido
- V é o número de falsos positivos (Erro Tipo I) (também chamado de "falsas descobertas")
- S é o número de verdadeiros positivos (também chamados de "descobertas verdadeiras")
- T é o número de falsos negativos (Erro Tipo II)
- U é o número de verdadeiros negativos
- $R = V + S$ é o número de hipóteses nulas rejeitadas (também chamadas de "descobertas", verdadeiras ou falsas)

Em m testes de hipóteses dos quais m_0 são hipóteses nulas verdadeiras, R é uma variável aleatória observável e S, T, U e V são variáveis aleatórias não observáveis.

- https://en.wikipedia.org/wiki/Multiple_comparisons_problem

Familywise error rate (Taxa de erro familiar)

Portanto, precisamos de uma abordagem diferente, que comece por definir exatamente o que queremos controlar. Este não é um problema novo e, durante décadas, os estatísticos usaram a Familywise error rate (FWER), que garante que a probabilidade de cometer uma única rejeição falsa é limitada por α.

- Alguns testes que controlam FWER
  - Teste de Tukey
  - Teste de Scheffé
  - Teste de Duncan
  - Teste de Dunnett
  - Correção de Bonferroni

$$\begin{cases} H_0: & \mu_i - \mu_j & = 0 \\ H_1: & \mu_i - \mu_j & \neq & 0, \text{ para i} < \text{j} \end{cases}$$

Teste Tukey (p.adj???)

- controla a taxa de erro individual?
- ou a taxa de erro experimental (ou de família)?

O teste Tukey controla a taxa de erro experimental ou familiar

- O Teste de Tukey controla o erro da família, ou family-wise error (FWE) considerando-se o valor p ajustado (p adj).
- O valor *p adj* deve ser comparado com o valor α, em geral, $\alpha = 0,05$.

Alguns comentários

- com testes de hipóteses: quanto mais testes simultâneos você executar, maior será a probabilidade de obter um resultado "significativo". Digamos que você esteja executando testes de 50 simultaneamente com um nível alpha de $0,05$. A probabilidade de observar pelo menos um evento significativo devido apenas ao acaso é:

$$P(\text{significant event}) = 1 - P(\text{no significant event}) \leq 1 - (1 - 0.05)^{50} = 0.92.$$

É quase certo (92%) que há pelo menos um resultado significativo.

- A correção de múltiplos testes refere-se a tornar os testes estatísticos mais rigorosos, a fim de neutralizar o problema dos testes múltiplos. O ajuste mais conhecido é a correção de Bonferroni, mas outros métodos foram desenvolvidos. Esses métodos são normalmente projetados para controlar a taxa de erro familiar ou a taxa de falsas descobertas.
- Brillinger, D. "The Collected Works of John W. Tukey". 1984.

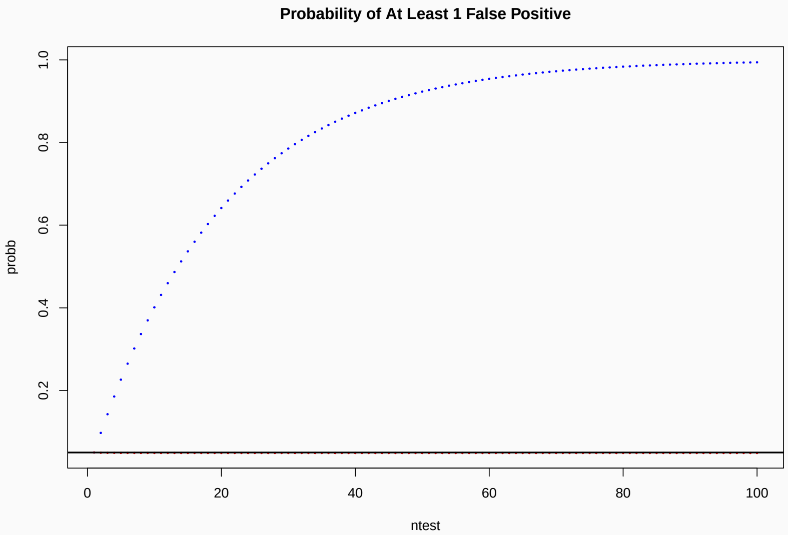
Como calcular a correção de Bonferroni

- O cálculo para este teste post-hoc é na verdade muito simples, é apenas o nível alfa (α) dividido pelo número de testes que você está executando.
- Exemplo de pergunta: Um pesquisador está testando 25 hipóteses diferentes ao mesmo tempo, usando um valor crítico de $0,05$. Qual é a correção de Bonferroni?

Responder: A correção de Bonferroni é $\frac{\alpha}{n} = \frac{0{,}05}{25} = 0,002$

Para este conjunto de testes de 25, você rejeitaria o nulo apenas se seu valor p fosse menor que 0,002.

Graficamente



- https://stat.ethz.ch/R-manual/R-devel/library/stats/html/p.adjust.html

```
?p.adjust.methods
```

- https://multithreaded.stitchfix.com/blog/2015/10/15/multiple-hypothesis-testing/

Artigos

- https://home.uchicago.edu/~amshaikh/webfiles/palgrave.pdf

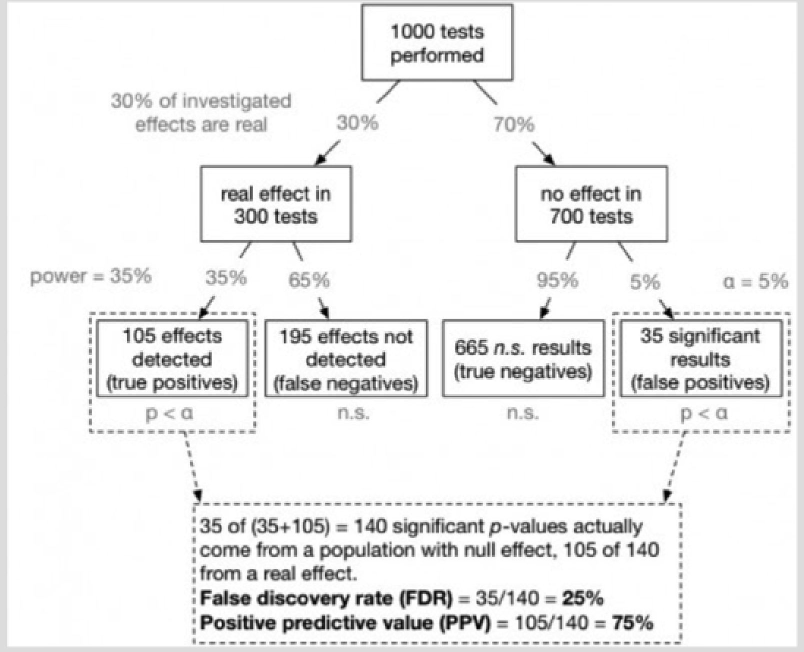
Correção de Bonferroni

- https://pt.wikipedia.org/wiki/Corre%C3%A7%C3%A3o_de_Bonferroni

Qual teste utilizar?

- https://www.statology.org/tukey-vs-bonferroni-vs-scheffe/

Testes que controlam o False discovery rate



- https://www.wikiwand.com/pt/Testes_de_hip%C3%B3teses
- https://en0.wikipedia.org/wiki/Multiple_comparisons_problem#Classification_of_multiple_hypothesis_tests

Correção de Bonferroni

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6099145/ Leia!
- https://www.wikiwand.com/pt/Corre%C3%A7%C3%A3o_de_Bonferroni
- https://www.graphpad.com/support/faq/adjusted-p-values-as-part-of-multiple-comparisons/

- Shaffer, J. P. (1995). Multiple Hypothesis Testing. Annual Review of Psychology. 46: 561–584. doi:10.1146/annurev.ps.46.020195.003021
- Strassburger, K.; Bretz, Frank (2008). Compatible simultaneous lower confidence bounds for the Holm procedure and other Bonferroni-based closed tests. Statistics in Medicine. 27 (24): 4914–4927. PMID 18618415. doi:10.1002/sim.3338
- Sidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. Journal of the American Statistical Association. 62 (318): 626–633. doi:10.1080/01621459.1967.10482935

```r
require(multcomp)
?glht
```

Teste de Dunnet

- https://pt.stackoverflow.com/questions/309395/como-fazer-o-teste-de-dunnett-no-r

Teste de Duncan

- https://search.r-project.org/CRAN/refmans/agricolae/html/duncan.test.html