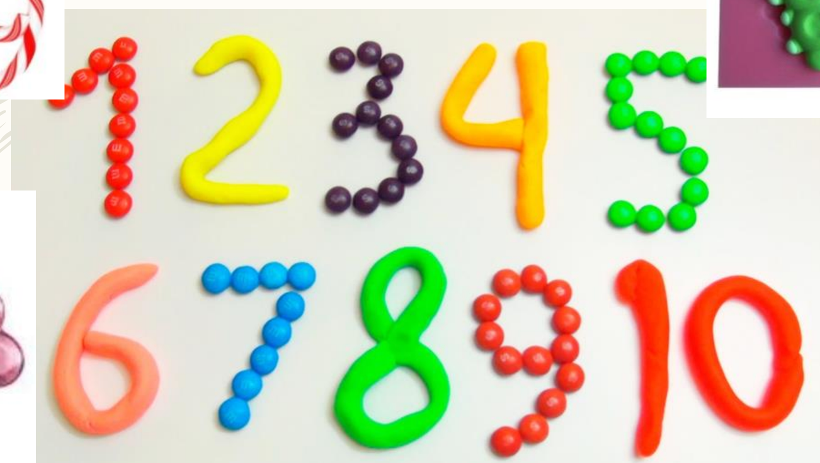
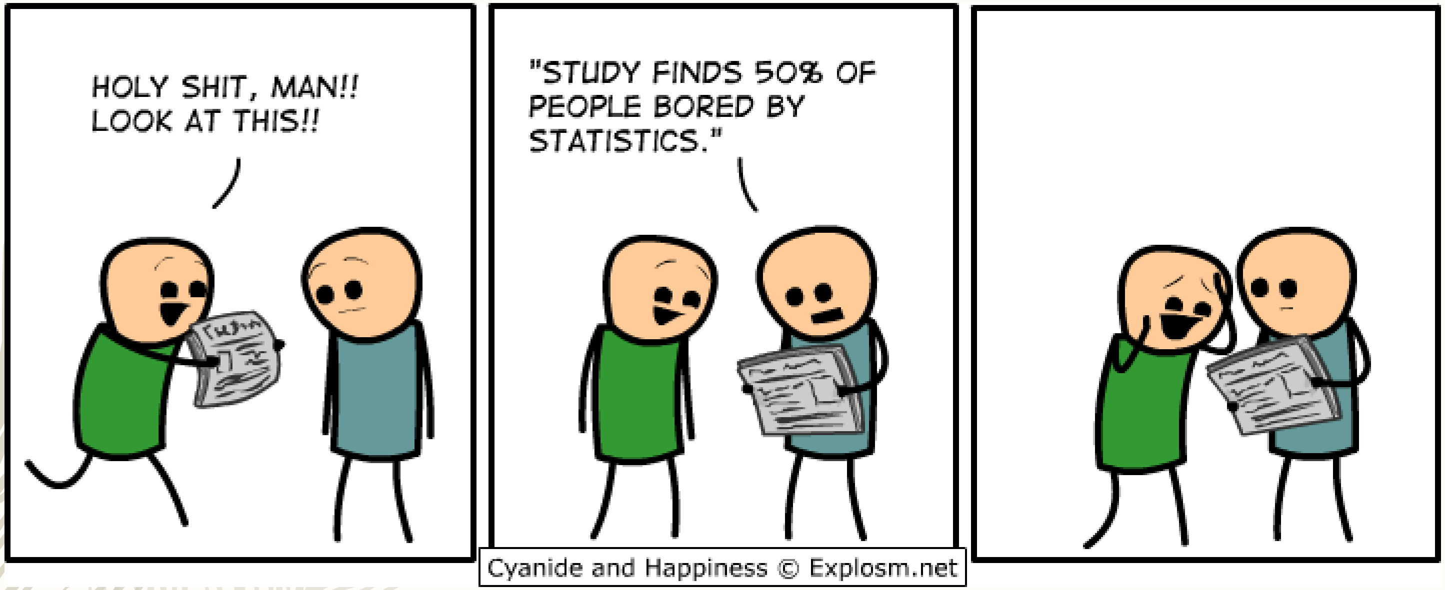
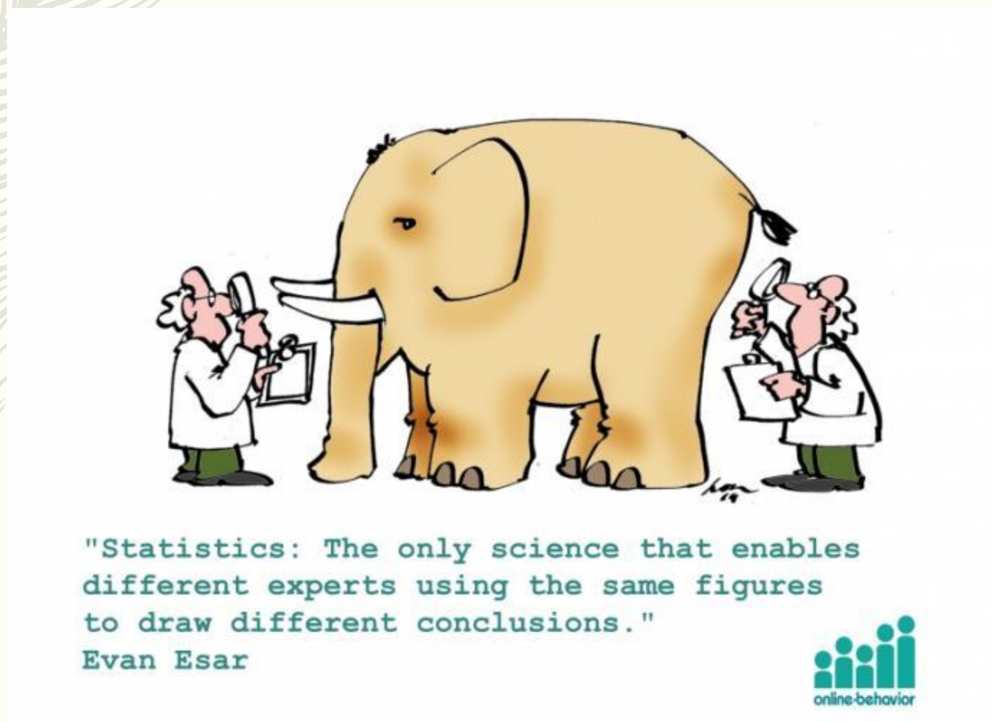
Goodies*



* Goodies related to animals, plants and numbers…


HOLY SHIT, MAN!! LOOK AT THIS!!
"STUDY FINDS 50% OF PEOPLE BORED BY STATISTICS."
Cyanide and Happiness © Explosm.net


"Statistics: The only science that enables different experts using the same figures to draw different conclusions."
Evan Esar
online-behavior

Home
Explore
Notifications
Messages
Bookmarks
Lists
Profile
More
Tweet

Tweet

Stuart Ritchie
@StuartJRitchie

I made this meme for our stats class last week and I thought you might like to see it.

11:54 AM · Oct 27, 2019 · Twitter Web App

370 Retweets **1.9K** Likes

PROGRAMME

09:00-09:10 *Welcome words* (L. Carriço, Dean)
09:10-09:25 *Facts and figures about research @ CIÊNCIAS* (M Santos-Reis, Vice Dean for Research)

SESSION I - Top Notch Science
09:25-09:40 *Out of this world atmospheres* (Pedro Machado)
09:40-09:55 *Active matter* (Nuno Araújo)
09:55-10:10 *The 1755 earthquake and the closing of the Atlantic Ocean* (João C. Duarte)
10:10-10:25 *Transcutaneous electric stimulation of the spinal cord: a modelling study* (Pedro C. Miranda)

10:25-11:00 Coffee-break

11:00-11:15 *Glycofighting bacteria: a new mode of action* (Rodrigo Almeida)
11:15-11:30 *A new mechanism to inhibit amyloid aggregation in Alzheimer's Disease* (Cláudio M. Gomes)
11:30-11:45 *How Mediterranean and Tropical forests react to groundwater change?* (Cristina Antunes)
11:45-12:00 *Vulnerability & Blame: making sense of unauthorized access to smartphones* (Diogo Marques)

12:00-14:30 Bring a sandwich, look at the posters and have a speed date
Speed dating the (great) experts behind great scientists! (12:30-13:30 - C3 Building, Atrium)
Speed dating a statistician (13:00-14:00 - C3 Building, Atrium)
Speed dating mathematicians (13:00-14:30 - C6 Building, Room 6.1.8)
THE LAST M_{e}I.LE_{I} (13:00-14:30 - C6 Building, Room 6.1.8)
(More info about the speed dating sessions available here).

/ Sociedade / Eventos /

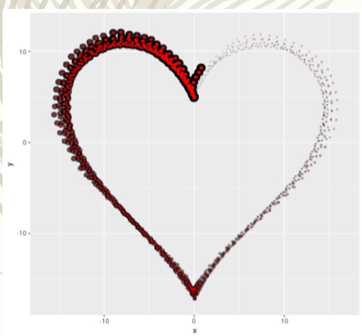
Ciências Research Day - programme now available

Share (0) Tweet (0) LinkedIn (0)

Out 30 2019

Localização
Edifício C3, FCUL, Lisboa

A melhor Ciência faz-se em CIÊNCIAS!



@accidental__aRt by @csmarcum

SESSION II - Recognising Excellence (ERC grantees)
14:30-14:40 *Why this, why now, why me?* (Joaquim Gaspar)
14:40-14:50 *Competition under (niche) construction: an ERC project (not so) easy to construct* (Sara Magalhães)
14:50-15:00 *Where's Wally?: Spotting the next ERC grantees at CIÊNCIAS* (Henrique Leitão)

SESSION III – Networking and Science for Society
15:00-15:15 *Intelligent infection management and precision antibiotherapy* (Ricardo Dias)
15:15-15:30 *Estimating the efficacy of mass rescue operations in ocean areas with vehicle routing models and heuristics* (Rui de Deus)
15:30-15:45 *CoastNet - Portuguese Coastal Monitoring Network* (José L. Costa)
15:45-16:00 *SmartHub Energy* (Miguel Brito)

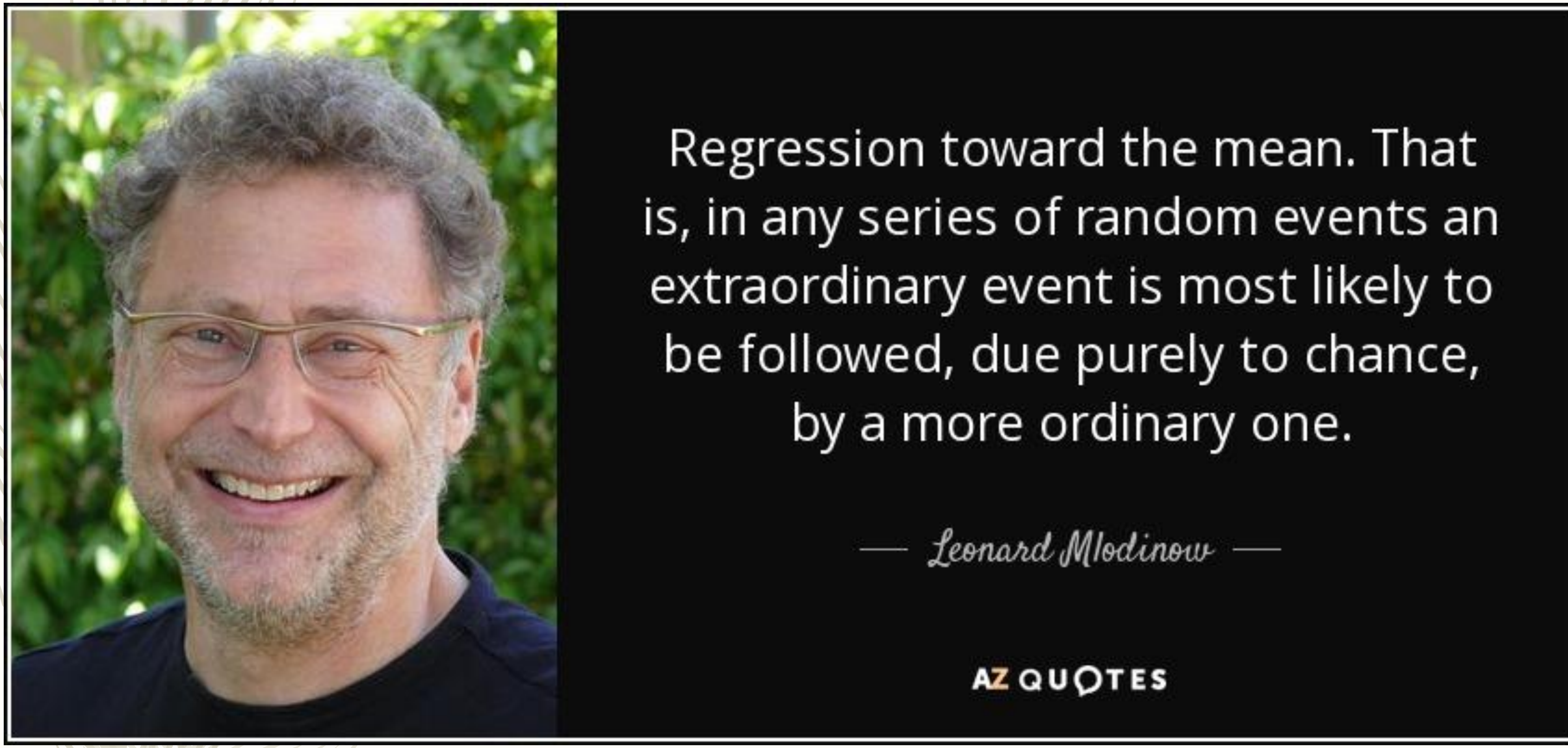
16:00-16:30 Coffee Break

16:30-16:45 *Ciências at the core of European efforts to push the boundary of physics* (António Amorim)
16:45-17:00 *Making the added value of networking tangible* (Raquel Conceição)
17:00-17:15 *The Art of spinning-off* (Fadhil Musa)

SESSION IV - Challenging Ideas for Ciências: Creative Minds Contest
17:15-17:45 *Pitch talks*

17:45-18:00 *Closing remarks and Awards* (Pedro Almeida, Vice Dean for Communication and Image)

Ecologia Numérica - Aula Teórica 13 – 29-10-2018



https://www.azquotes.com/quote/1393944

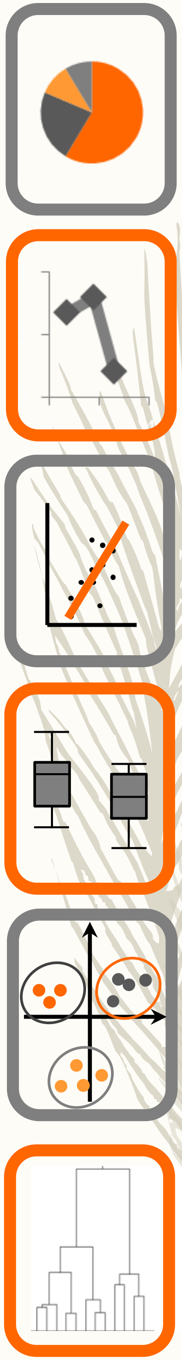
testes de hipóteses a uma ou duas amostras (emparelhadas ou não)

Síntese

1 amostra
- Teste *t* — Paramétrico (P)
- Teste Wilcoxon — Não P (NP)

2 amostras:
- Teste *t* — P
- Teste Mann-Whitney — NP

2 amostras emparelhadas:
- Teste *t* para amostras emparelhadas — P
- Teste Wilcoxon para amostras emparelhadas — NP



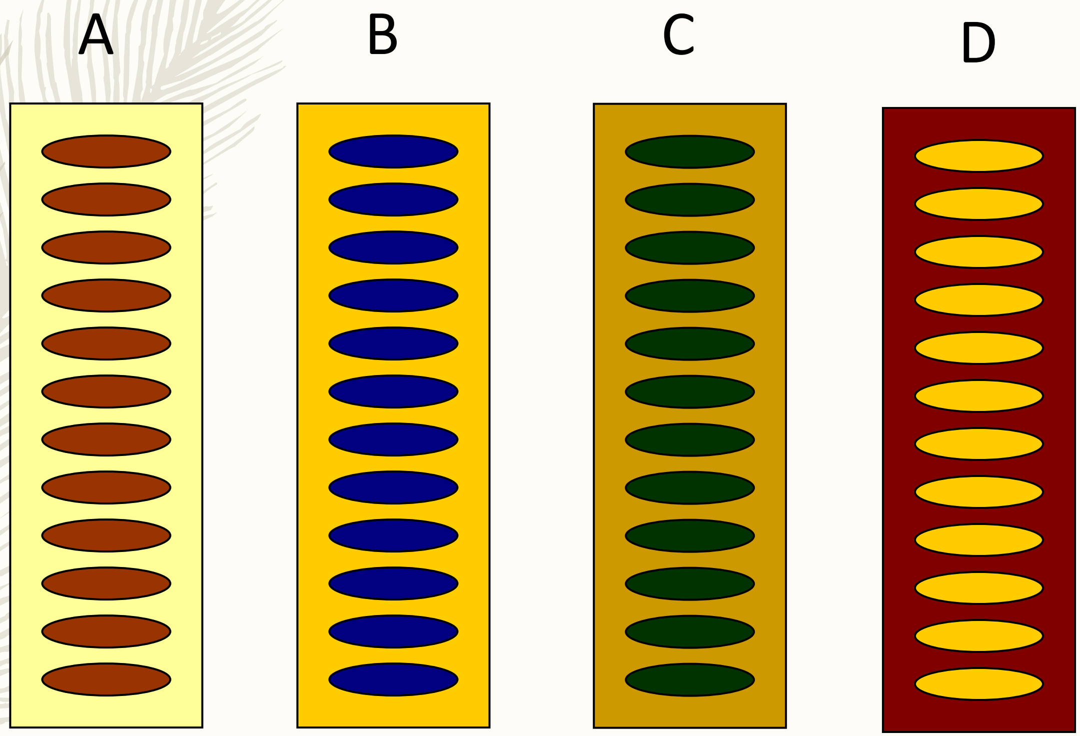
Ecologia Numérica

testes a mais de duas amostras
análise de variância e
equivalente não paramétrico

anova e equivalente não paramétrico

- Quais os testes mais correntes para situações de 3 ou mais amostras?
- Quais as condições para a sua aplicação?
- Como interpretar os seus resultados?

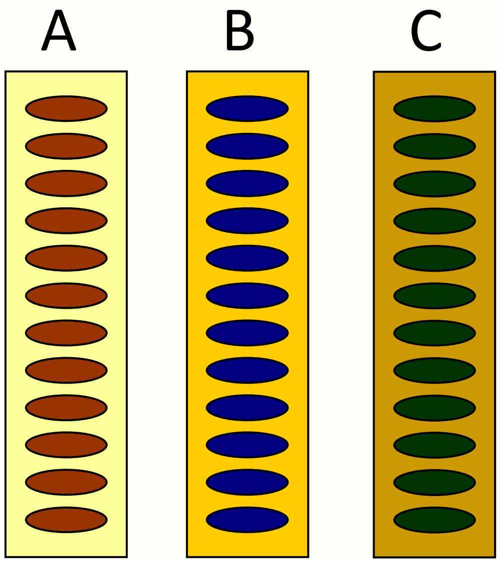
Testes a mais de duas amostras



Testes a mais de duas amostras

Realizar todos os testes de pares de amostras, duas a duas:

A vs B
B vs C
A vs C



Mas isso não é uma boa ideia...Porque não?

Testes a mais de duas amostras

Testes a duas amostras:
A vs B
B vs C
A vs C

Porque não?

Para o conjunto das 3 hipóteses a probabilidade de correctamente não rejeitar todas as H_0 seria
$0.95^3 = 0.86$

ou seja, $\alpha = 0.14$!!!!

Hipótese nula		Hipótese nula: Não Rejeitar	Hipótese nula: Rejeitar
Hipótese nula	Verdadeira	**Não há erro**	α
	Falsa	β	**Não há erro**

A análise de variância (ANOVA)

H_0: $\mu_A=\mu_B=\mu_C$

H_1: Há pelo menos uma média diferente das outras = as médias não são todas iguais

Fontes de variação:

Intra-amostra ou intra-grupos (erro)

Entre-amostras ou entre-grupos

Fontes de variação:

Somas dos quadrados

Total:

$$SQ_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{X})^2$$

Entre-grupos:

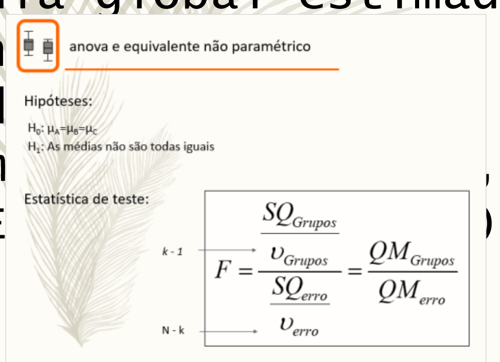
$$SQ_G = \sum_{i=1}^{k} n_i(\overline{X}_i - \overline{X})^2$$

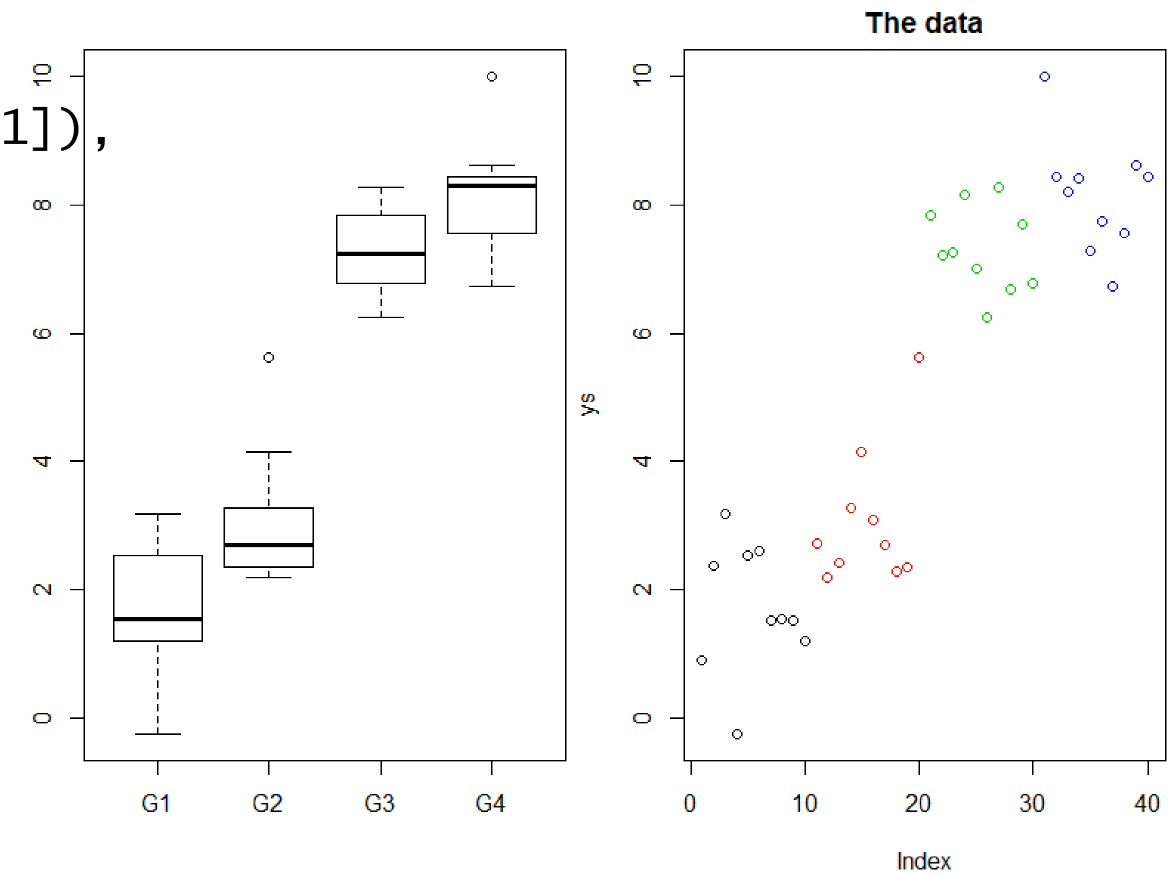
Intra-grupos (erro):

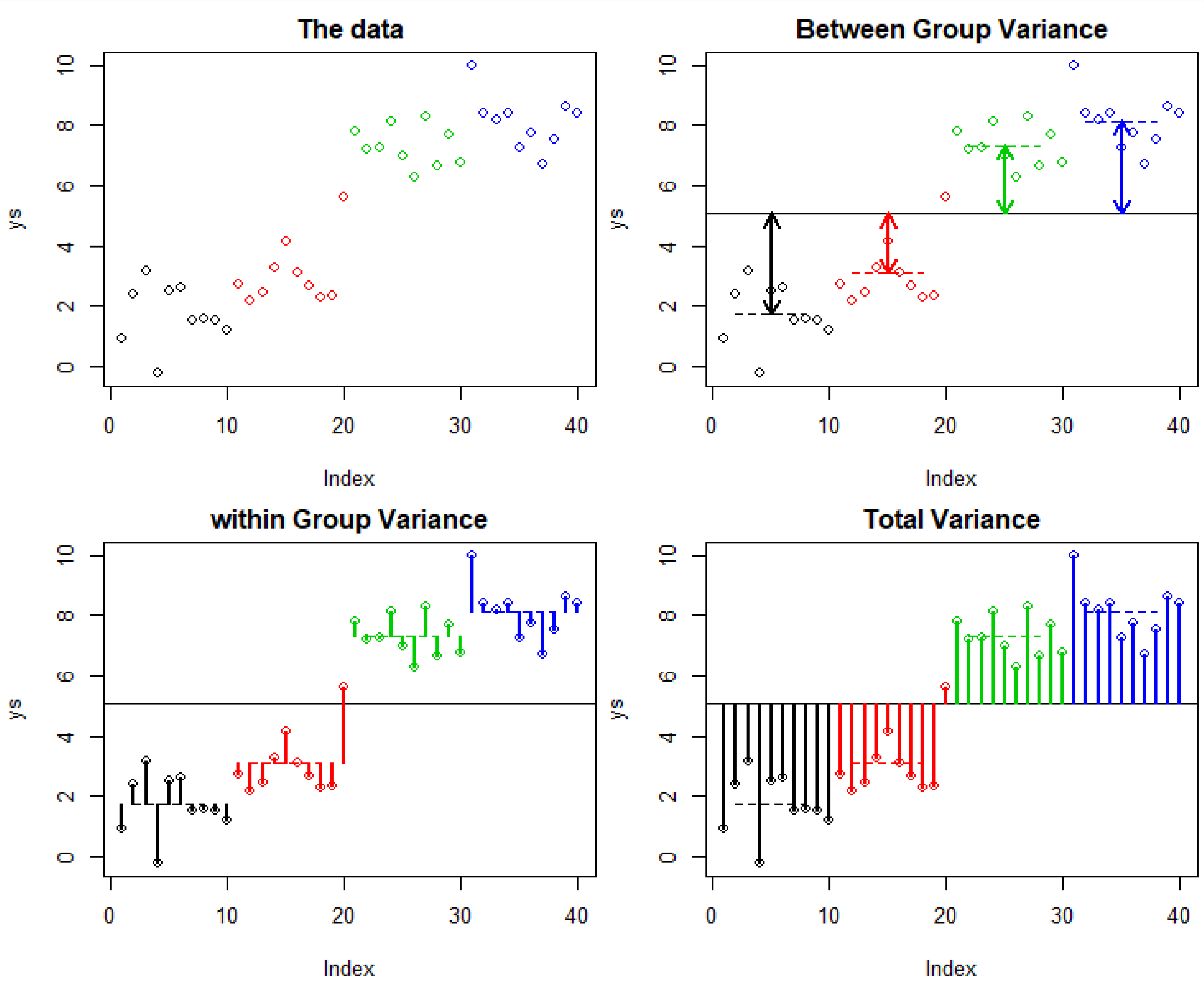
$$SQ_{erro} = \sum_{i=1}^{k}\left[\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2\right]$$

An example with 4 groups, 10 observations per group

```
set.seed(1234)
#observações por tratamento
n=10
#número de tratamentos
ng=4
N=n*ng
group=rep(paste0("G",1:ng),each=n)
means=c(2.1,3.2,7.7,8.9)
#gerar os dados
ys=c(rnorm(n,mean=means[1]),
rnorm(n,mean=means[2]),
rnorm(n,mean=means[3]),
rnorm(n,mean=means[4]))
#media global estimada
gran
#med
estm                            ,
INDE                            )
```







The data
Between Group Variance
within Group Variance
Total Variance
ys
Index
0
10
20
30
40
0
2
4
6
8
10

Fontes de variação:

Somas dos quadrados

$$SQ_{TOTAL} = SQ_{GRUPOS} + SQ_{erro}$$

Hipóteses:

H_0: $\mu_A = \mu_B = \mu_C$

H_1: As médias não são todas iguais

Estatística de teste:

$$F = \frac{\dfrac{SQ_{Grupos}}{\upsilon_{Grupos}}}{\dfrac{SQ_{erro}}{\upsilon_{erro}}} = \frac{QM_{Grupos}}{QM_{erro}}$$

k - 1 → υ_{Grupos}

N - k → υ_{erro}

Fontes de variação:

Soma dos quadrados

$$SQ_{TOTAL} = SQ_{GRUPOS} + SQ_{erro}$$

$$gl_{TOTAL} = gl_{GRUPOS} + gl_{erro}$$

$$N - 1 = k - 1 + N - k$$

k é o número de grupos, *N* é o número total de observações

anova e equivalente não paramétrico

Estatística de teste:

$$F = \frac{QM_{Grupos}}{QM_{erro}}$$

Valor crítico:

$$F_{\alpha,\upsilon_g,\upsilon_{erro}}$$

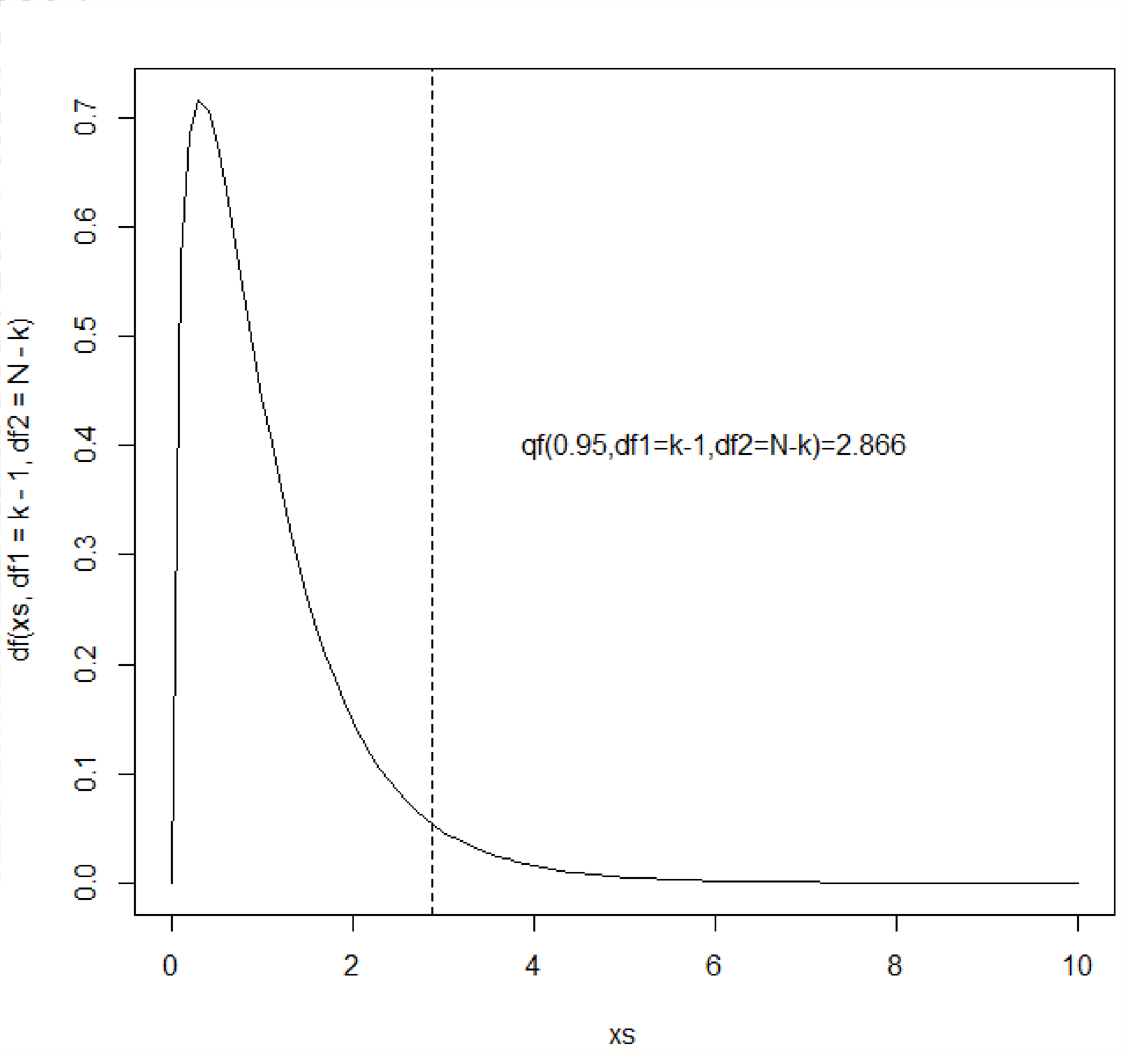
Critério de decisão:

Rejeitar H_0 se:

$$F > F_{\alpha,\upsilon_g,\upsilon_{erro}}$$

Não rejeitar H_0 caso contrário

A distribuição F



K=4

N=40

```
> #ANOVA BY HAND
> #Total sum of squares
> sum((ys-grandmean)^2)
[1] 326.4652
> #between groups sum of squares
> sum(10*(estmeans-grandmean)^2)
[1] 296.0989
> #within groups sum of squares
> sum((rep(estmeans,each=10)-ys)^2)
[1] 30.36632
```

```
> aov(ys~group)
Call:
   aov(formula = ys ~ group)

Terms:
                    group Residuals
Sum of Squares  296.09888  30.36632
Deg. of Freedom         3        36

Residual standard error: 0.9184274
Estimated effects may be unbalanced
> summary(aov(ys~group))
            Df Sum Sq Mean Sq F value Pr(>F)    
group        3 296.10   98.70     117 <2e-16 ***
Residuals   36  30.37    0.84                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> (296.10/3)/(30.37/36)
[1] 116.997
> 1-pf((296.10/3)/(30.37/36),3,36)
[1] 0
> 1-pf(98.7/0.84,3,36)
[1] 0
```

Rejeitamos H0 para qualquer nivel de significância!

Duas formas de fazer a mesma coisa (e uma *preview* sobre a relação entre modelos de regressão e análise de variância)

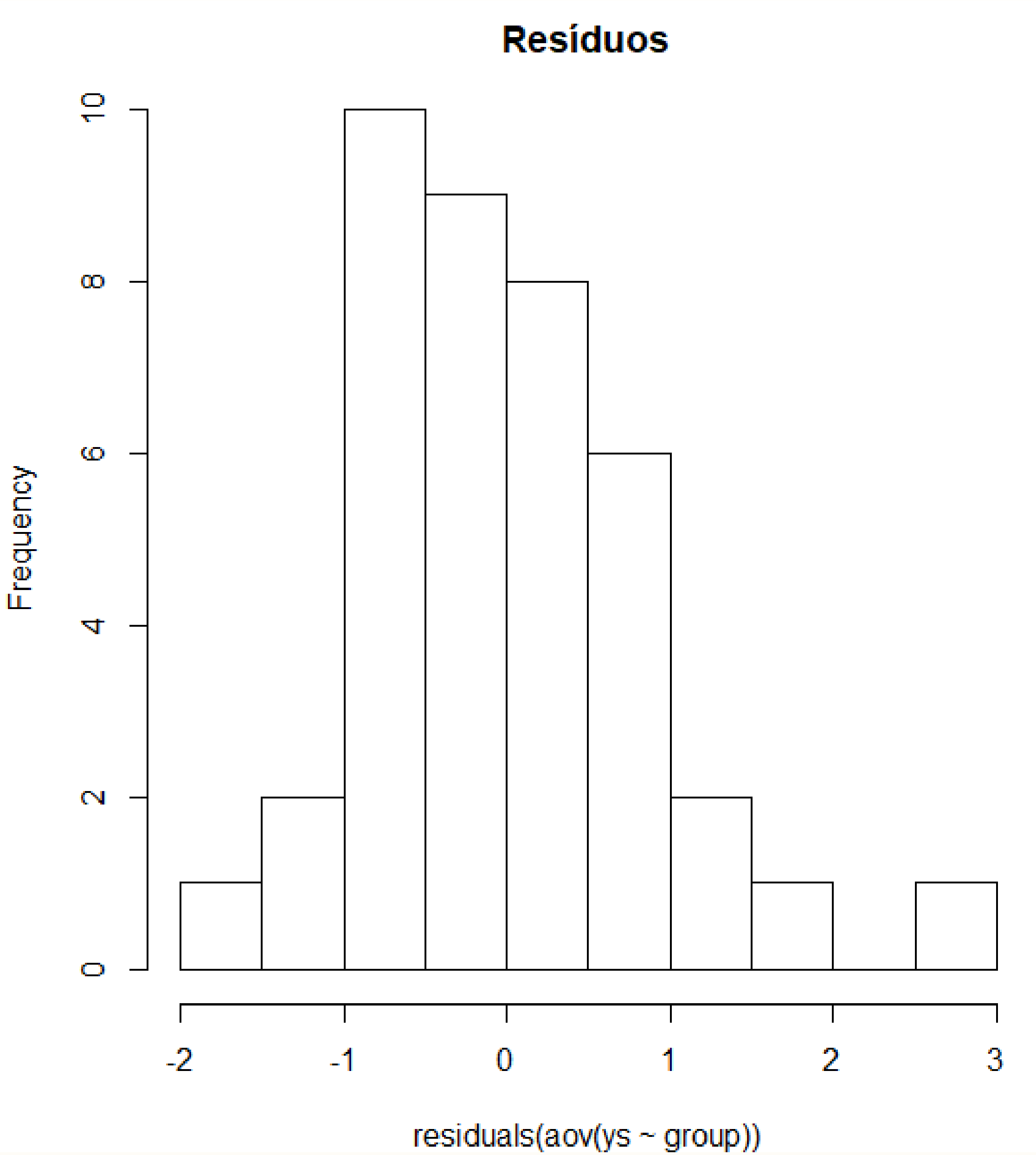
```
> anova(lm(ys~group))
Analysis of Variance Table

Response: ys
          Df  Sum Sq Mean Sq F value    Pr(>F)    
group      3 296.099  98.700  117.01 < 2.2e-16 ***
Residuals 36  30.366   0.844                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> aov(ys~group)
Call:
   aov(formula = ys ~ group)

Terms:
                    group Residuals
Sum of Squares  296.09888  30.36632
Deg. of Freedom         3        36

Residual standard error: 0.9184274
Estimated effects may be unbalanced
```

Pressupostos da ANOVA

- Os resíduos provêm duma população Gaussiana
- As variâncias dos diferentes tratamentos são homogéneas

No entanto, à semelhança do referido para o teste *t*, a análise de variância é bastante robusta, i.e. o seu desempenho não é profundamente afectado por desvios moderados dos pressupostos

```
hist(residuals(aov(ys~group)),main="Resíduos")
```



Pressupostos da ANOVA

As amostras provêm de populações normais com variância igual

Como verificar os pressupostos?

O que fazer caso não sejam cumpridos?

Teste sobre diferença no valor médio entre mais de duas amostras

Avaliar se os pressupostos são cumpridos

Não

Transformação dos dados

Não

Sim

Testes não paramétricos

Teste de Kruskal-Wallis

Testes paramétricos

ANOVA

Teste de Kruskal-Wallis: abordagem não-paramétrica

Hipóteses:

H_0: as amostras provêm da mesma população

H_1: as amostras não provêm da mesma população

Estatística de teste:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i \bar{r}_{i.}^{2} - 3(N+1)$$

Teste de Kruskal-Wallis

Quando há empates deve ser utilizado um factor de correcção:

$$C = 1 - \frac{\sum t}{N^3 - N}$$

m=número de empates
ti = número de observações empatadas no grupo i

onde

$$\sum t = \sum_{i=1}^{m} (t_i^3 - t_i)$$

A estatística de teste deve ser corrigida do seguinte modo:

$$H_C = \frac{H}{C}$$

Teste de Kruskal-Wallis

Estatística de teste:

$$H_C = \frac{H}{C}$$

Valor crítico:

$$\chi^2{}_{\alpha,gl}$$

sendo gl=k-1 (k=número de grupos)

Critério de decisão:

Rejeitar H_0 se:

$$\chi^2 > \chi^2_{\alpha,gl}$$

Não rejeitar H_0 caso contrário

```
> 1-pchisq(32.0839,3)
[1] 5.024662e-07
```

```
> kruskal.test(x=ys,g=as.factor(group))

        Kruskal-Wallis rank sum test

data:  ys and as.factor(group)
Kruskal-Wallis chi-squared = 32.084, df = 3, p-value = 5.025e-07
```

```
> ranks=rank(ys)
> meanranks=tapply(X=ranks,INDEX=group,FUN=mean)
> H=12/(N*(N+1))*(sum(meanranks^2*10))-3*(N+1)
> H
[1] 32.0839
```

Decisão: rejeitar H0 para os níveis usuais de significância

Procedimentos de teste a posteriori

Quando a H_0 é rejeitada numa hipótese envolvendo 3 ou mais amostras, não é sempre óbvio qual ou quais das amostras diferem das outras.

Há, por isso, a necessidade de efectuar testes *a posteriori* de comparações múltiplas

Testes a posteriori

ANOVA

- Teste de Tukey (tipo Tukey)
- Teste de Newman-Keuls
- Teste de Scheffé

Teste de Kruskal-Wallis

- Teste de Dunn

Teste de Tukey

Estatística de teste:

$$q = \frac{\bar{X}_B - \bar{X}_A}{\sqrt{\frac{s^2}{n}}}$$

Valor crítico:

$$q_{\alpha,\upsilon,k}$$

Sendo gl=N-k (N=número total de observações; k=número de grupos)

Critério de decisão:

Rejeitar H_0 se:

$$q > q_{\alpha,\upsilon,k}$$

Não rejeitar H_0 caso contrário

Teste de Dunn

Estatística de teste:

$$Q = \frac{\overline{R}_B - \overline{R}_A}{SE}$$

onde

$$SE = \sqrt{\left(\frac{N(N+1)}{12} - \frac{\sum t}{12(N-1)}\right)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

Valor crítico:

$$Q_{\alpha,k}$$

sendo k=número de grupos

Critério de decisão:

Rejeitar H_0 se:

$$Q > Q_{\alpha,k}$$

Não rejeitar H_0 caso contrário

R IMPLEMENTATIONS
OF
A POSTERIORI TESTS
FOR
MULTIPLE COMPARISONS

Teste de Tukey

```
> myanova=aov(ys~group)
> TukeyHSD(myanova)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = ys ~ group)

$group
           diff        lwr      upr     p adj
G2-G1 1.3649867  0.2587888 2.471185 0.0105651
G3-G1 5.5952106  4.4890127 6.701408 0.0000000
G4-G1 6.4169643  5.3107665 7.523162 0.0000000
G3-G2 4.2302239  3.1240260 5.336422 0.0000000
G4-G2 5.0519776  3.9457798 6.158176 0.0000000
G4-G3 0.8217538 -0.2844441 1.927952 0.2066495
```

Apenas os grupos 3 e 4 parecem não ser diferentes entre si, o que faz sentido (ver gráfico)

Teste de Newman-Keuls

```
> library(agricolae)
> out <- SNK.test(myanova,trt="group")
> out
$statistics
    MSerror Df     Mean       CV
  0.8435089 36 5.061133 18.14668

$parameters
  test name.t ntr alpha
   SNK  group   4  0.05

$snk
      Table CriticalRange
2 2.868158     0.8330056
3 3.456758     1.0039540
4 3.808798     1.1061979

$means
         ys       std  r        Min       Max      Q25      Q50      Q75
G1 1.716843 0.9957875 10 -0.2456977  3.184441 1.288787 1.544458 2.491201
G2 3.081829 1.0673376 10  2.2016136  5.615835 2.378058 2.705899 3.220773
G3 7.312053 0.6660013 10  6.2517951  8.274756 6.824608 7.234383 7.796782
G4 8.133807 0.8942458 10  6.7199604 10.002298 7.602350 8.294651 8.431679

$comparison
NULL

$groups
         ys groups
G4 8.133807      a
G3 7.312053      a
G2 3.081829      b
G1 1.716843      c
```

Teste de Scheffé

```
> library(DescTools)
Warning message:
package 'DescTools' was built under R version 3.4.4
> ScheffeTest(myanova)

  Posthoc multiple comparisons of means : Scheffe Test
    95% family-wise confidence level

$group
           diff     lwr.ci   upr.ci    pval
G2-G1 1.3649867  0.1605647 2.569409  0.0207 *
G3-G1 5.5952106  4.3907886 6.799633 2.8e-14 ***
G4-G1 6.4169643  5.2125424 7.621386 4.2e-16 ***
G3-G2 4.2302239  3.0258019 5.434646 7.9e-11 ***
G4-G2 5.0519776  3.8475557 6.256400 5.7e-13 ***
G4-G3 0.8217538 -0.3826682 2.026176  0.2784

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Teste de Dunn

```
> dunn.test(x=ys,g=group)
  Kruskal-Wallis rank sum test

data: ys and group
Kruskal-Wallis chi-squared = 32.0839, df = 3, p-value = 0


                           Comparison of ys by group
                                (No adjustment)
Col Mean-|
Row Mean |          G1          G2          G3
---------+------------------------------------
      G2 |   -1.224147
         |      0.1104
         |
      G3 |   -3.863714   -2.639567
         |     0.0001*     0.0042*
         |
      G4 |   -5.011352   -3.787205   -1.147638
         |     0.0000*     0.0001*      0.1256

alpha = 0.05
Reject Ho if p <= alpha/2
```

Less power, so actually can't tell if Groups 1 and 2 are different!

anova e equivalente não paramétrico

Síntese

- 1 amostra
  - Teste *t* — P
  - Teste Wilcoxon — NP
- 2 amostras:
  - Teste *t* — P
  - Teste Mann-Whitney — NP
- 2 amostras emparelhadas:
  - Teste *t* *para amostras emparelhadas* — P
  - Teste Wilcoxon para amostras emparelhadas — NP
- 3 ou + amostras:
  - ANOVA — P
  - Teste Kruskal-Wallis — NP

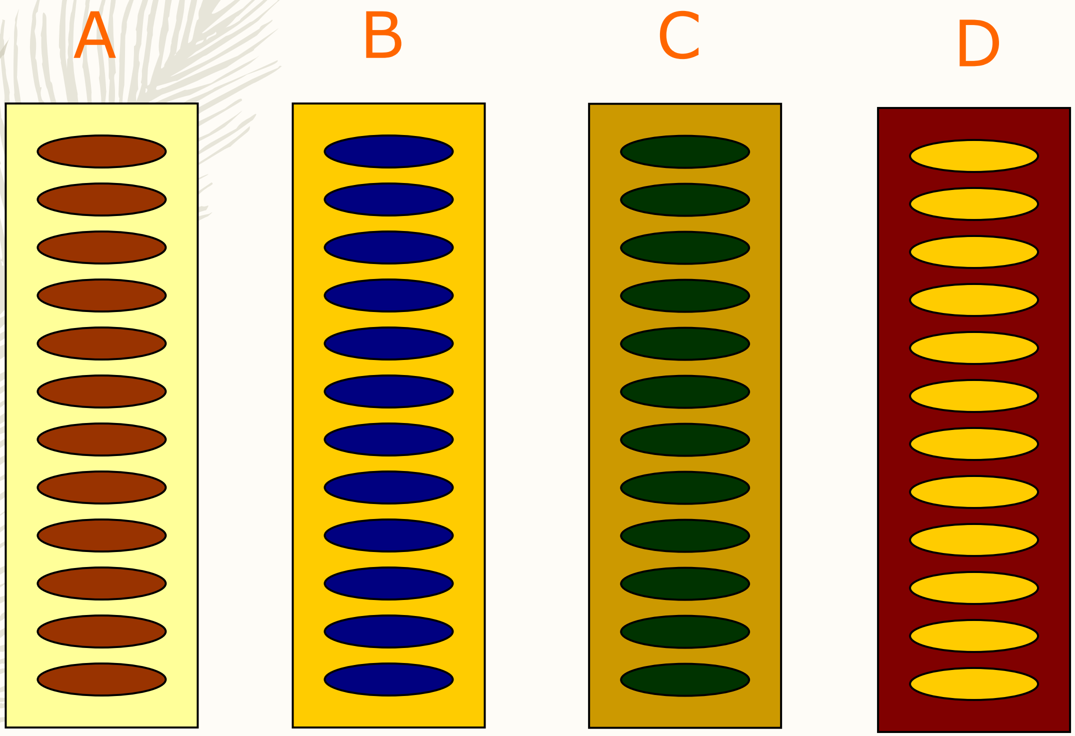
Ecologia Numérica

anova multifactorial

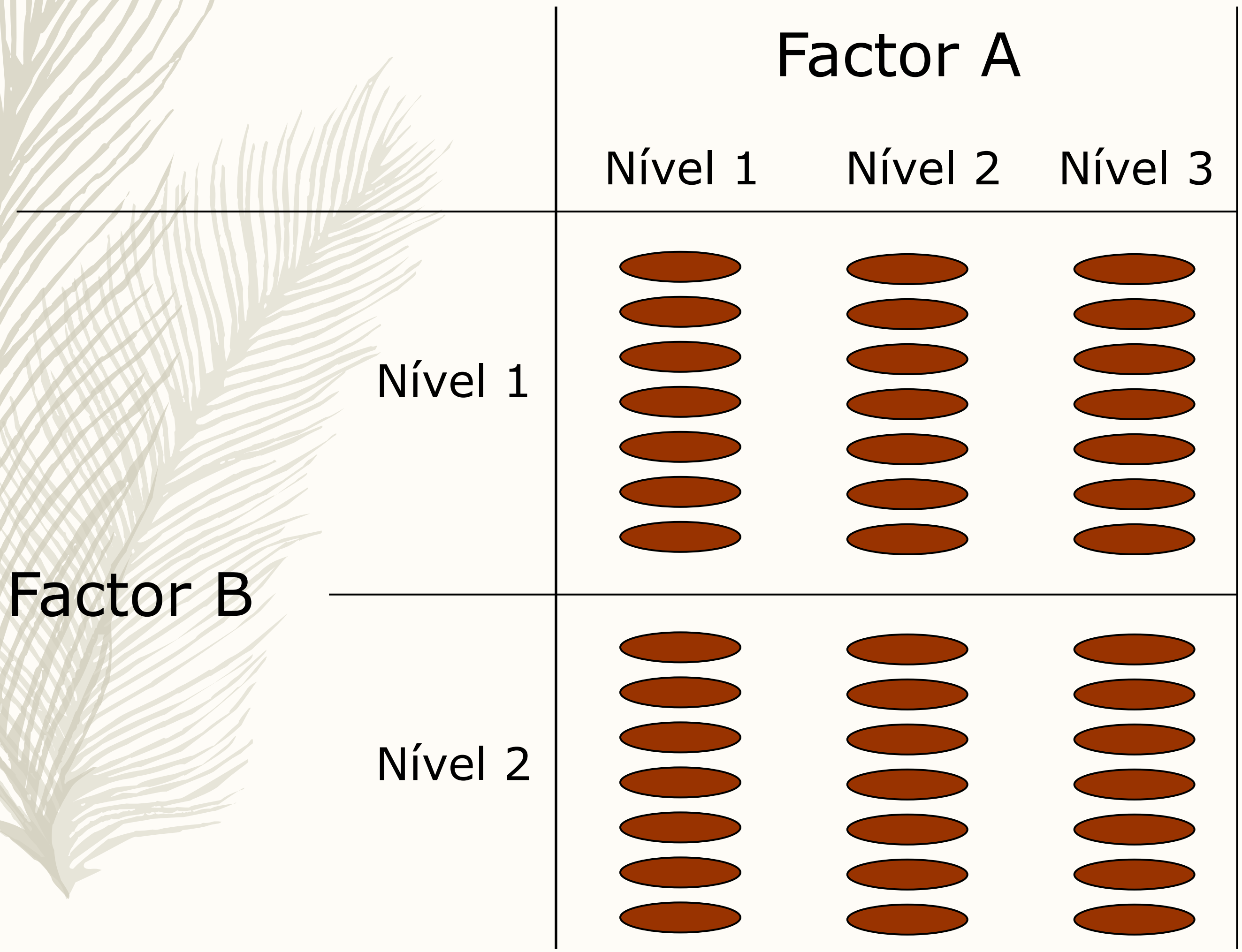
anova multifactorial

- Quais os testes mais correntes para situações com múltiplas amostras e factores?
- Quais as condições para a sua aplicação?
- Como interpretar os seus resultados?

Testes a mais de duas amostras



ANOVA multifactorial



ANOVA multifactorial (H0's c/ 2 factores)

H0: μA=μB=μC - factor A
H1: As médias não são todas iguais – factor A

H0: μA=μB=μC - factor B
H1: As médias não são todas iguais – factor B

H0: não há interacção entre factor A e B
H1: há interacção entre factor A e B

Fontes de variação ANOVA simples:

Somas dos quadrados

$$SQ_{TOTAL} = SQ_{GRUPOS} + Sq_{erro}$$

(às vezes a notação é SS de "sum of squares", às vezes é SQ de "squares")

$$SS_{TOTAL} = SS_{GRUPOS} + SS_{erro}$$

ANOVA multifactorial

Fontes de variação:

Somas dos quadrados

Entre grupos

Intra grupos, ou o que fica por explicar, ou seja, o erro

$$SQ_{TOTAL} = SQ_{GRUPOS} + SQ_{erro}$$

$$SQ_{factor\ A} + SQ_{factor\ B} + SQ_{interacção}$$

Fontes de variação ANOVA simples:
Somas dos quadrados

Total:

$$SQ_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2$$

Entre-grupos:

$$SQ_G = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$

Intra-grupos (erro):

$$SQ_{erro} = \sum_{i=1}^{k}\left[\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2\right]$$

anova multifactorial

A variação total entre grupos pode ser decomposta em mais 3 fontes de variação

Fontes de variação ANOVA simples:

Somas dos quadrados

Total:

$$SQ_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i}(X_{ijk} - \bar{X})^2$$

$gl = N - 1$

Entre-grupos:

$$SQ_G = \sum_{i=1}^{a}\sum_{j=1}^{b} n(\bar{X}_{ij} - \bar{X})^2$$

$gl = ab - 1$

Intra-grupos (erro):

$$SQ_{erro} = \sum_{i=1}^{a}\sum_{j=1}^{b}\left[\sum_{k=1}^{n}(X_{ijk} - \bar{X}_{ij})^2\right]$$

$gl = ab(n-1)$

Com 2 factores, A SQ_G pode ser dividida em:

factor A:

$$SQ_A = bn\sum_{i=1}^{a}(\bar{X}_{i.} - \bar{X})^2$$

$$gl = a - 1$$

factor B:

$$SQ_B = an\sum_{j=1}^{a}(\bar{X}_{.j} - \bar{X})^2$$

$$gl = a - 1$$

interacção

$$SQ_I = SQ_G - SQ_A - SQ_B$$

$$gl = (a-1)(b-1)$$

Hipóteses:

H_0: $\mu_A=\mu_B=\mu_C$

H_1: As médias não são todas iguais

Estatística de teste:

$$F = \frac{\dfrac{SQ_{factor}}{\upsilon_{factor}}}{\dfrac{SQ_{erro}}{\upsilon_{erro}}} = \frac{QM_{factor}}{QM_{erro}}$$

anova multifactorial

Hipóteses:

H0: não há interacção entre os factores

H1: há interacção entre os factores

Estatística de teste:

$$F = \frac{\dfrac{SQ_{AxB}}{\upsilon_{AxB}}}{\dfrac{SQ_{erro}}{\upsilon_{erro}}} = \frac{QM_{AxB}}{QM_{erro}}$$

Estatística de teste:

$$F = \frac{QM_{factor}}{QM_{erro}}$$

$$F = \frac{QM_{AxB}}{QM_{erro}}$$

Valor crítico:

$$F_{\alpha,\upsilon_{num},\upsilon_{erro}}$$

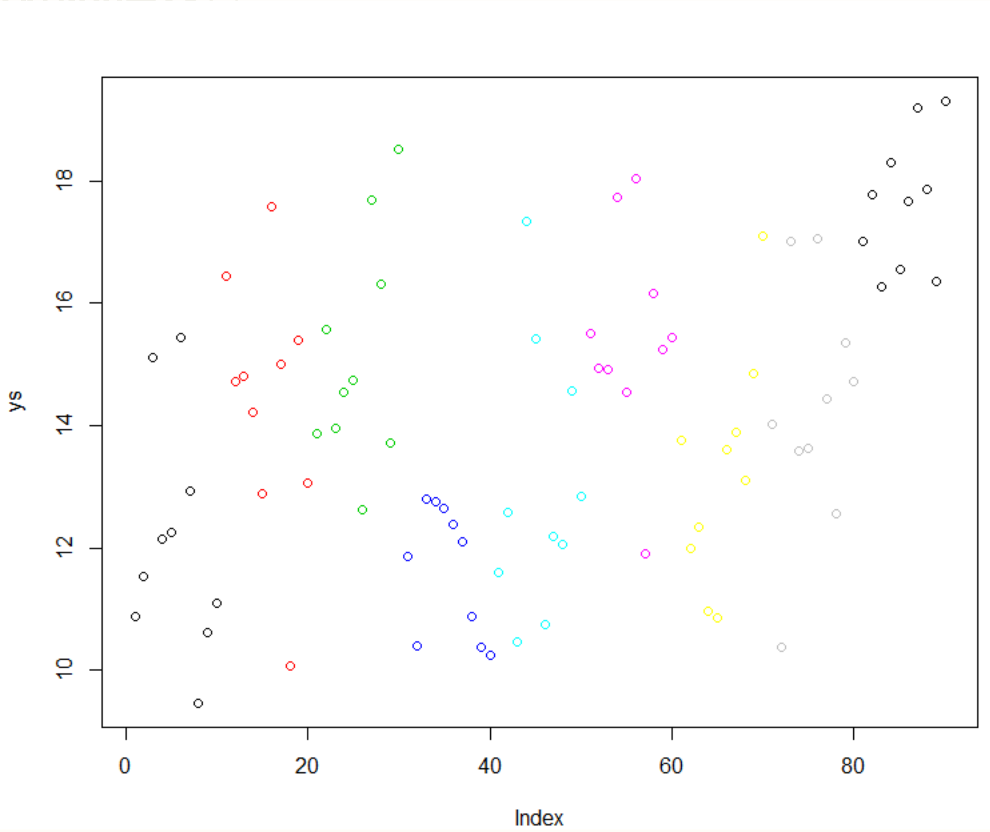
Critério de decisão:

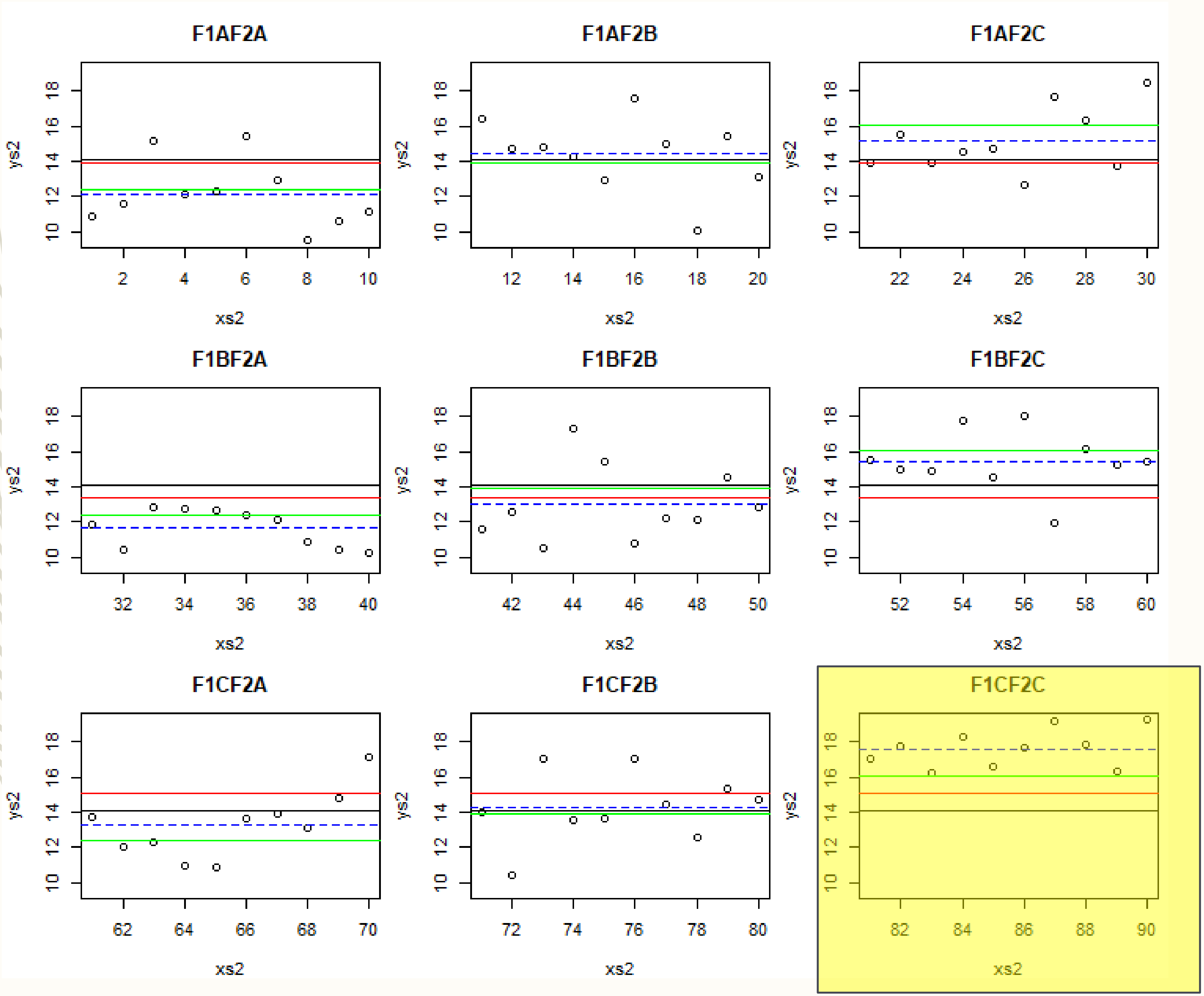
Rejeitar H0 se: $F > F_{\alpha,\upsilon_{num},\upsilon_{erro}}$

Não rejeitar H0 caso contrário

Um exemplo com 2 factores, 3 níveis por factor, 10 observações por grupo (N=90)

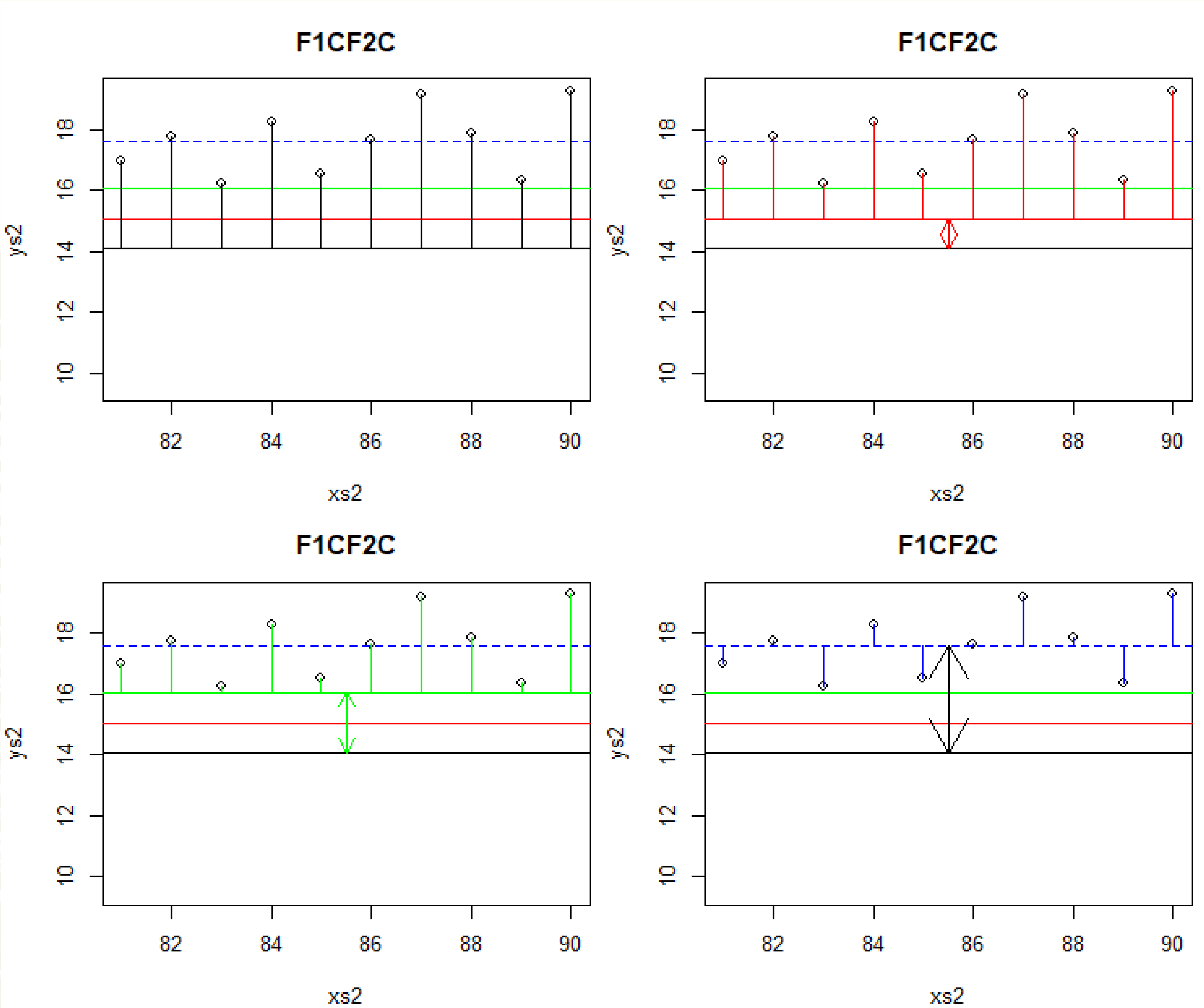
```
set.seed(123)
F1=rep(paste0("F1",c("A","B","C")),each=30)
F2=rep(paste0("F2",c("A","B","C","A","B","C","A","B","C")),each=10)
table(F1,F2)
#simulate the data
ys=3+ifelse(F1=="F1A",3,ifelse(F1=="F1B",2,4))+ifelse(F2=="F2A",6,if
else(F2=="F2B",8,10))+rnorm(90,mean=0,sd=2)
```




F1AF2A
F1AF2B
F1AF2C
F1BF2A
F1BF2B
F1BF2C
F1CF2A
F1CF2B
F1CF2C
ys2
xs2

Total error

Level 1 effect



Level 2 effect

Interaction effect

```
> #ANOVA A DOIS FACTORES
> set.seed(123)
> F1=rep(paste0("F1",c("A","B","C")),each=30)
> F2=rep(paste0("F2",c("A","B","C","A","B","C","A","B","C")),each=10)
> table(F1,F2)
     F2
F1     F2A F2B F2C
  F1A  10  10  10
  F1B  10  10  10
  F1C  10  10  10
> #simulate the data
> ys=3+ifelse(F1=="F1A",3,ifelse(F1=="F1B",2,4))+ifelse(F2=="F2A",6,ifelse(F2=="F2B",8,10))+rnorm(90,mean=0,sd=2)
> my2WayANOVA=aov(ys~F1*F2)
> summary(my2WayANOVA)
            Df Sum Sq Mean Sq F value   Pr(>F)    
F1           2  44.72   22.36   6.982  0.00159 ** 
F2           2 210.39  105.19  32.851 3.57e-11 ***
F1:F2        4  17.78    4.45   1.388  0.24532    
Residuals   81 259.37    3.20                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Pressupostos da ANOVA:

- Os resíduos provêm duma população normal
- As variâncias são homogéneas

No entanto, à semelhança do referido para o teste t, a análise de variância é bastante robusta, i.e. o seu desempenho não é profundamente afectado por desvios moderados dos pressupostos

Em caso de rejeição da H0 com mais de dois níveis devem efectuar-se...

Testes a posteriori

ANOVA

Teste de Tukey (tipo Tukey)
Teste de Newman-Keuls
Teste de Scheffé

Teste de Tukey

Estatística de teste:

$$q = \frac{\bar{X}_B - \bar{X}_A}{\sqrt{\frac{s^2}{n}}}$$

Valor crítico:

$$q_{\alpha,\upsilon,k}$$

Sendo gl=N-k (N=número total de observações; k=número de grupos)

Critério de decisão:

Rejeitar H0 se: $q > q_{\alpha,\upsilon,k}$

Não rejeitar H0 caso contrário

Se os pressupostos não forem cumpridos, qual a abordagem não-paramétrica?

Teste de Kruskal-Wallis de dois factores

Segundo alguns autores o desempenho deste teste, também conhecido por teste de Scheirer Ray Hare é fraco e não constitui uma boa alternativa à ANOVA de dois factores.

```
> library(rcompanion)
> dados=data.frame(ys=ys,F1=F1,F2=F2)
> scheirerRayHare(ys~F1*F2,data=dados)

DV:  ys
Observations:  90
D:  1
MS total:  682.5

          Df  Sum Sq      H p.value
F1         2  4475.5  6.557 0.03768
F2         2 23885.9 34.998 0.00000
F1:F2      4  2038.1  2.986 0.56014
Residuals 81 30343.1
```