



Vall d'Hebron  
Institut de Recerca

VHIR

# Vall d'Hebron Institut de Recerca

**Ricardo Gonzalo**  
[ricardo.gonzalo@vhir.org](mailto:ricardo.gonzalo@vhir.org)

**13/02/2019**

**Module 2**

**Session 3.1: Principles of Experimental Design.  
Analysis of Variance**

# Table of contents

1. Basic ideas of experimental design
2. Experimental Design conditions
3. Experimental Design types
4. Introduction to ANOVA
5. How does ANOVA ‘work’?
6. ANOVA assumptions
7. Beyond ANOVA
  1. Multiple comparisons
  2. Non parametric ANOVA
8. Exercises

# Table of contents

1. Basic ideas of experimental design
2. Experimental Design conditions
3. Experimental Design types
4. Introduction to ANOVA
5. How does ANOVA ‘work’?
6. ANOVA assumptions
7. Beyond ANOVA
  1. Multiple comparisons
  2. Non parametric ANOVA
8. Exercises

## Experimental design should be a mandatory step in every experiment

- Experimental design is a structured, organized method for determining the relationship between the different factors affecting an experimental process, and the output of that process.



Sir Ronald A. Fisher  
Father of modern Mathematical Statistics  
and Developer of Experimental Design  
and ANOVA

*“To consult the statistician after an experiment is finished, is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.”*

## Types of variability that play role in an experiment:



- **Planned variability:** These are the differences in response between treatments applied.



- **Noise variability:** Differences between two consecutive measures. We cannot avoid that.



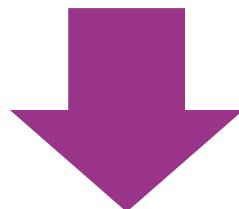
- **Variability not planned:** Produce a systematic variation in the results. A priori the reason is not known.

## A good experimental design should...

- Avoid systematic error: e.g. samples from one group processed with instrument A and samples from the other group processed with instrument B.
- Be precise: try to maintain the random error as low as possible
- Allow estimation of error: enough replicates in each treatment
- Have broad validity: our experimental units should reflect the population about which we wish to draw inference

## A good experimental design should...

- Avoid systematic error: e.g. samples from one group processed with instrument A and samples from the other group processed with instrument B.
- Be precise: try to maintain the random error as low as possible
- Allow estimation of error: enough replicates in each treatment
- Have broad validity: our experimental units should reflect the population about which we wish to draw inference



Randomization

Replication

Local Control

# Table of contents

1. Basic ideas of experimental design
2. Experimental Design conditions
3. Experimental Design types
4. Introduction to ANOVA
5. How does ANOVA ‘work’?
6. ANOVA assumptions
7. Beyond ANOVA
  1. Multiple comparisons
  2. Non parametric ANOVA
8. Exercises

## Randomization

- The method for assigning treatments to units involves *randomization* ("any individual experimental subject has the same chance as any other individual of finding itself in each experimental group")
- It is one of the most important elements of a well-designed experiment
- Made valid most of the statistical analysis usually performed

e.g. 4 treatments to be assigned to 16 units

1. Use sixteen identical slips of paper, 4 marked with A, 4 with B, and so on to D. Put the slips of paper into a basket and mix them thoroughly. For each unit, we draw a slip of paper from the basket and use the treatment marked on the slip
2. Treatment A is assigned to the first four units we have encounter, treatment B to next four units, and so on.

Saying “randomly assign...” is sometimes easier to say than to do, especially in complex designs.

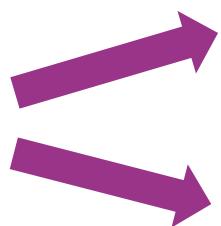
Some tools may help

- Research Randomizer  
<http://www.randomizer.org/>
  
- Interactive Statistical Calculation pages  
<http://statpages.org/> (look por “Experimental design”)

## Replication

- It is the basis of all experimental design
- It is the repetition of the basic experiment with another experimental units

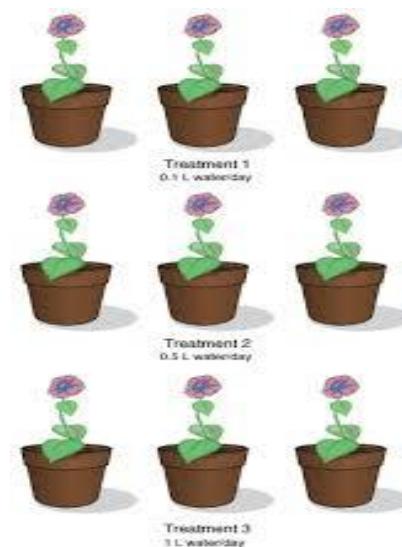
How many replicates I need?



the more replicates we have, the more confident we can be that differences between groups are real and not simply due to chance effects

More replicates increase in time/money cost

- To be in mind:
  1. To Know the variability of the technology used
  2. Previous works with similar technology
  3. Directly correlated with the precision of the experiment



## Local control

- When the experimental units are not homogeneous or the process to analyze them neither are (Kits lot numbers, batch,...)



We are not interested in to find out the differences between the levels of the blocks



## Effect of treatment A and B?

| Sample | Treatment | Sex    | Batch |
|--------|-----------|--------|-------|
| 1      | A         | Male   | 1     |
| 2      | A         | Male   | 1     |
| 3      | A         | Male   | 1     |
| 4      | A         | Male   | 1     |
| 5      | B         | Female | 2     |
| 6      | B         | Female | 2     |
| 7      | B         | Female | 2     |
| 8      | B         | Female | 2     |

| Sample | Treatment | Sex    | Batch |
|--------|-----------|--------|-------|
| 1      | A         | Male   | 1     |
| 2      | A         | Female | 2     |
| 3      | A         | Male   | 2     |
| 4      | A         | Female | 1     |
| 5      | B         | Male   | 1     |
| 6      | B         | Female | 2     |
| 7      | B         | Male   | 2     |
| 8      | B         | Female | 1     |

## Effect of treatment A and B?

| Sample | Treatment | Sex    | Batch |
|--------|-----------|--------|-------|
| 1      | A         | Male   | 1     |
| 2      | A         | Male   | 1     |
| 3      | A         | Male   | 1     |
| 4      | A         | Male   | 1     |
| 5      | B         | Female | 2     |
| 6      | B         | Female | 2     |
| 7      | B         | Female | 2     |
| 8      | B         | Female | 2     |

| Sample | Treatment | Sex    | Batch |
|--------|-----------|--------|-------|
| 1      | A         | Male   | 1     |
| 2      | A         | Female | 2     |
| 3      | A         | Male   | 2     |
| 4      | A         | Female | 1     |
| 5      | B         | Male   | 1     |
| 6      | B         | Female | 2     |
| 7      | B         | Male   | 2     |
| 8      | B         | Female | 1     |



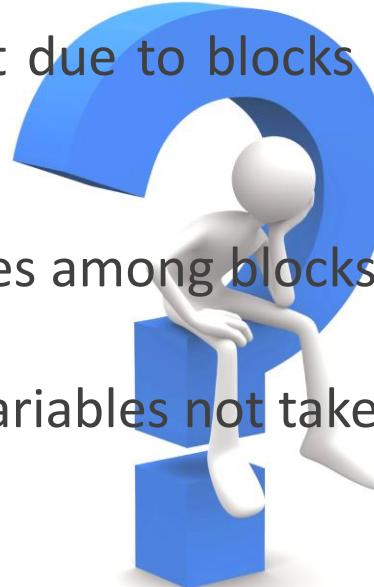
Treatment is confounded  
between sex and batch



Treatment is well balanced

## Local control or randomize?

- Local control assure you that differences are not due to blocks in the sample
- Local control eliminate the noise due to differences among blocks
- Randomization is good for balance effects from variables not taken into account from the beginning.



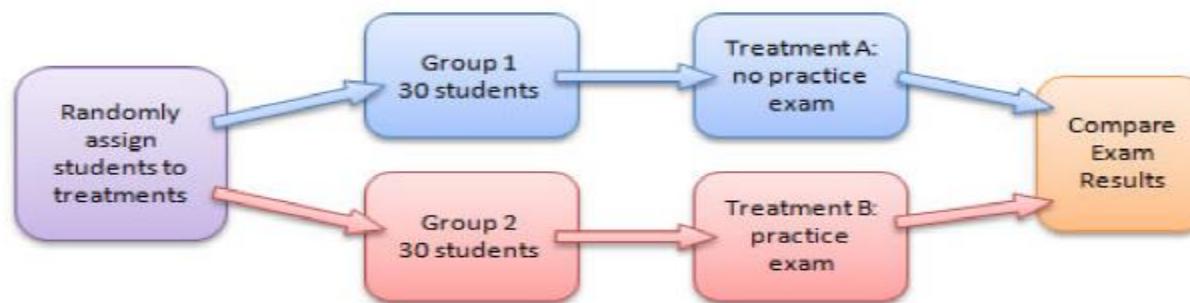
“Block what you can, randomize what you cannot” (George Box, 1978)

# Table of contents

1. Basic ideas of experimental design
2. Experimental Design conditions
3. Experimental Design types
4. Introduction to ANOVA
5. How does ANOVA ‘work’?
6. ANOVA assumptions
7. Beyond ANOVA
  1. Multiple comparisons
  2. Non parametric ANOVA
8. Exercises

## COMPLETE RANDOMIZED DESIGN (CRD)

- Simplest of all designs
  - Uses randomization and replication.
- ✓ Treatments are allocated at random to experimental units



**Response = mean + treatment effect + error**

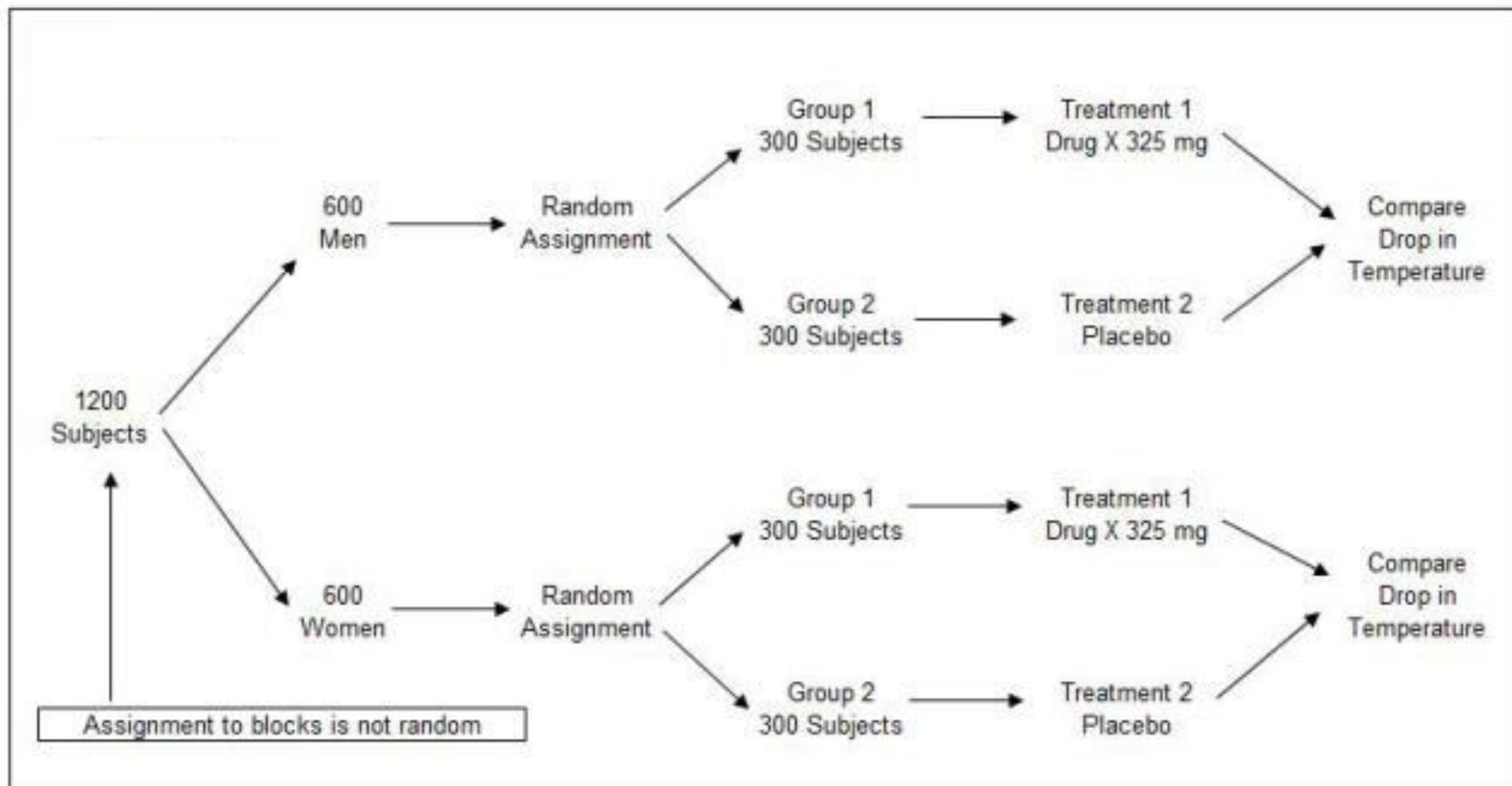
## RANDOMIZED BLOCK DESIGN (RBD)

- Blocking may be used to divide the whole experimental material into homogeneous strata or sub-groups known as blocks.
- Blocking to “remove” the effect of nuisance factors



Then the experimental units are randomly assigned treatments.

**Response = mean + treatment effect + block effect + error**



## FACTORIAL DESIGN

- Effects of several factors of variation are studied simultaneously.
- The treatments are all the combinations of different factors under study.
- The effects of each of the factors and the *interaction* effects, which are the variations in the effect of one factor as a result to different levels of other factors, are studied.

**Response = mean + treat-1 effect + treat-2 effect + treat-1:2-Interaction + error**

The simplest factorial experiment contains two levels for each of two factors.

Study the effects of two drugs at **two** different administration doses



200 mg/kg



400 mg/kg



200 mg/kg



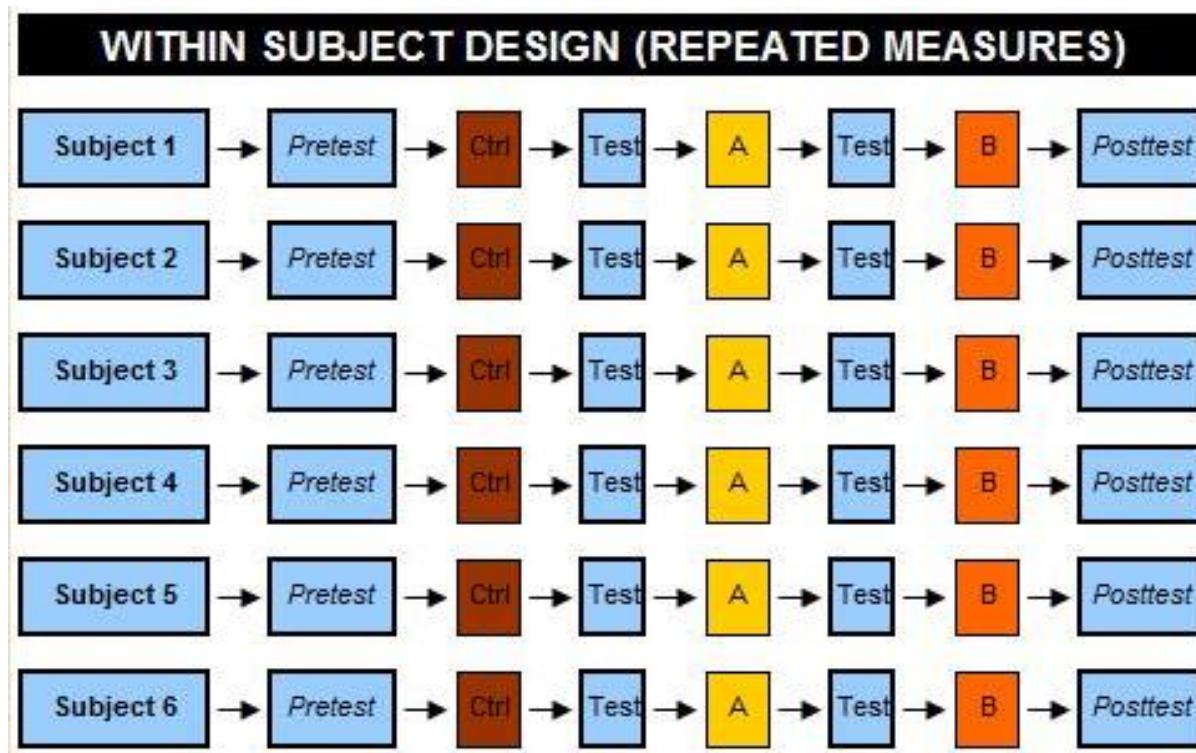
400 mg/kg

## REPEATED MEASURES

When **more than one measure** is taken on each experimental unit one has a within subjects design:

- Measures on the same individual are correlated
- New source of variability to be accounted for:  
**variability within subjects**
- Same possibilities as with other designs
- Problems:
  - some patients can pull out before complete the whole study
  - Individuals become better/worse at a task over time

## Repeated measures Example:



# Table of contents

1. Basic ideas of experimental design
2. Experimental Design conditions
3. Experimental Design types
4. Introduction to ANOVA
5. How does ANOVA ‘work’?
6. ANOVA assumptions
7. Beyond ANOVA
  1. Multiple comparisons
  2. Non parametric ANOVA
8. Exercises

Until now, we were comparing two samples, but what happens if we have more to compare? Could we still use Student's t test?

A pharmaceutical laboratory wants to test which of three drugs are the best:



drug 1



drug 2



drug 3

To know which of the drugs is the best one, one could think to perform the following comparison using a t test:

Comparisons with t test would be:

Chance of Type I error



Comparisons with t test would be:

Chance of Type I error



VS



$$\longrightarrow \alpha = 5\%$$



VS



$$\longrightarrow \alpha = 10\%$$



VS



$$\longrightarrow \alpha = 15\%$$

Comparisons with t test would be:

Chance of Type I error



VS



$$\longrightarrow \alpha = 5\%$$



VS



$$\longrightarrow \alpha \approx 10\%$$



VS



$$\longrightarrow \alpha \approx 15\%$$

Would be easier to accept the null hypothesis when it was wrong (more false positives)

# Table of contents

- 1. Basic ideas of experimental design**
- 2. Experimental Design conditions**
  1. Randomization
  2. Replication
  3. Local control
- 3. Experimental Design types**
  1. Completely Randomized Design (CRD)
  2. Randomized Block Design (RBD)
  3. Factorial Experiments
  4. Repeated Measures (Within Subjects Designs)
- 4. Introduction to ANOVA**
- 5. How does ANOVA ‘work’?**
- 6. ANOVA assumptions**
- 7. Beyond ANOVA**
  1. Multiple comparisons
  2. Non parametric ANOVA
- 8. Exercises**

## ANOVA tests the following hypotheses:

From linear model of ANOVA:  $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$

$H_0:$  The means of all the groups are equal:  
 $H_0: \mu_1 = \mu_2 = \dots = \mu_a$

No effects among the different treatments:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_a = 0$$

$H_1:$  Not all the means are equal

## ANOVA tests the following hypotheses:

$H_1$ : Not all the means are equal



If we reject  $H_0$  we only could say that there is a difference among the groups, but we couldn't say between which of them



Can follow up with “multiple comparisons”

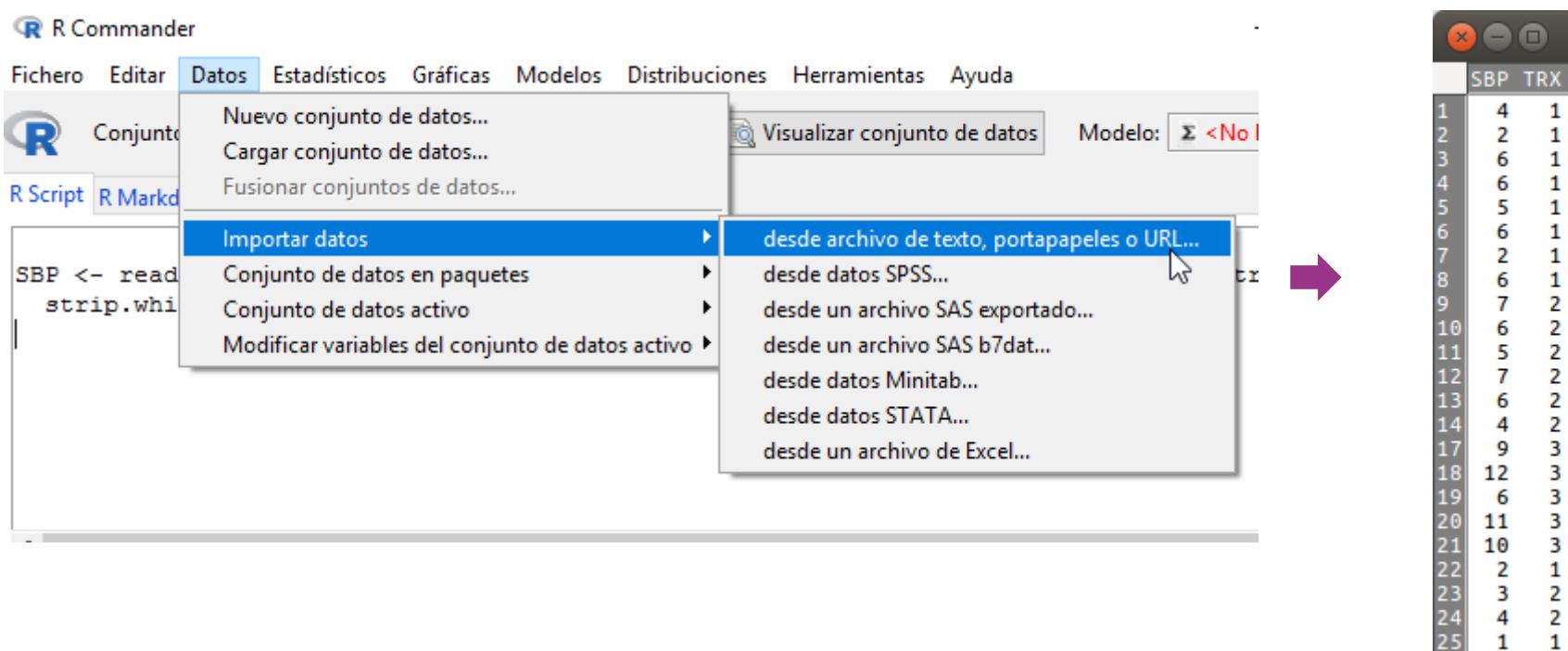
## Example. Complete Randomized Design

Which of the three drugs is the best to reduce the SBP? Three drugs are randomly assigned to 24 patients (same characteristics), and SBP is monitored after a month.

| Sample | Drug1 | Drug2 | Drug3 | One factor |
|--------|-------|-------|-------|------------|
| 1      | 4     | 7     | 9     |            |
| 2      | 2     | 6     | 12    |            |
| 3      | 6     | 5     | 6     |            |
| 4      | 6     | 7     | 11    |            |
| 5      | 5     | 6     | 10    |            |
| 6      | 6     | 4     | 11    |            |
| 7      | 2     | 7     | 9     |            |
| 8      | 6     | 5     | 10    |            |

Three levels

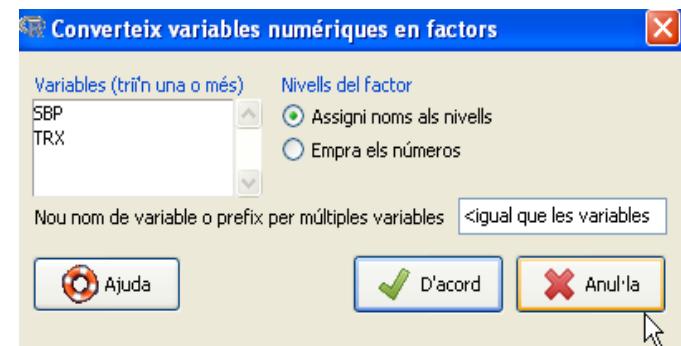
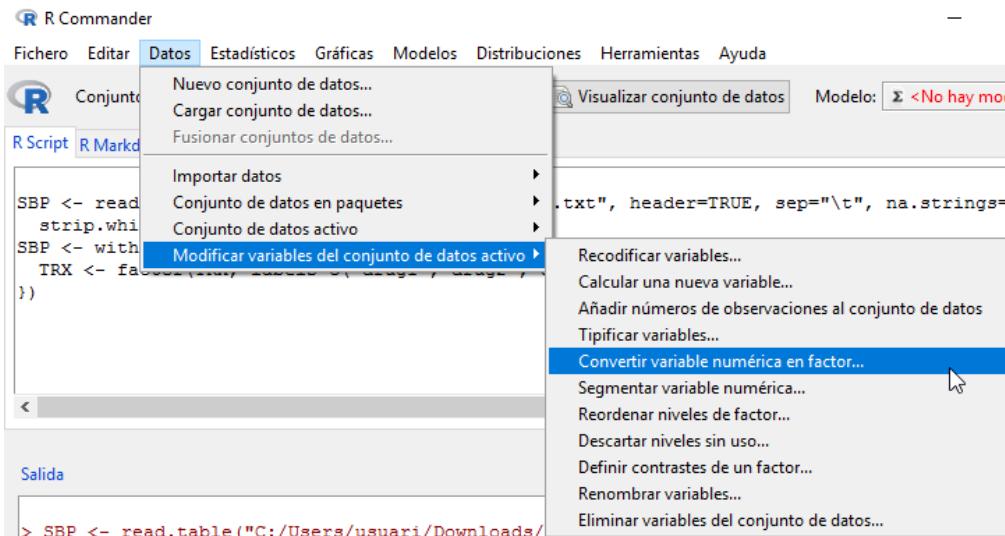
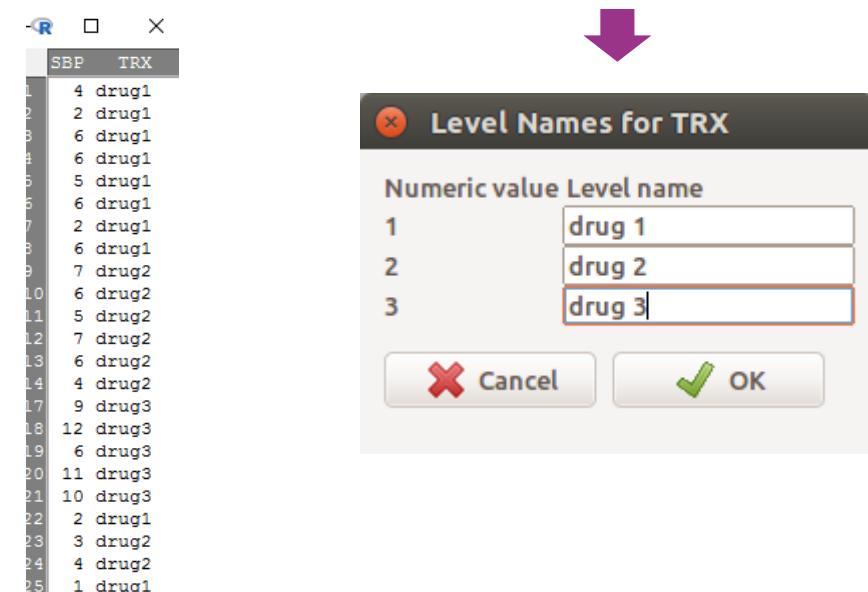
Always is good to “look” graphically the data. Let’s try with R-commander:



The screenshot shows the R Commander interface. The 'Datos' menu is selected, and the 'Importar datos' option is highlighted. A submenu is open, listing various data import options: 'desde archivo de texto, portapapeles o URL...', 'desde datos SPSS...', 'desde un archivo SAS exportado...', 'desde un archivo SAS b7dat...', 'desde datos Minitab...', 'desde datos STATA...', and 'desde un archivo de Excel...'. A purple arrow points from the 'Importar datos' submenu towards a data preview window on the right. The preview window displays a table with columns 'SBP' and 'TRX', containing 25 rows of data.

|    | SBP | TRX |
|----|-----|-----|
| 1  | 4   | 1   |
| 2  | 2   | 1   |
| 3  | 6   | 1   |
| 4  | 6   | 1   |
| 5  | 5   | 1   |
| 6  | 6   | 1   |
| 7  | 2   | 1   |
| 8  | 6   | 1   |
| 9  | 7   | 2   |
| 10 | 6   | 2   |
| 11 | 5   | 2   |
| 12 | 7   | 2   |
| 13 | 6   | 2   |
| 14 | 4   | 2   |
| 17 | 9   | 3   |
| 18 | 12  | 3   |
| 19 | 6   | 3   |
| 20 | 11  | 3   |
| 21 | 10  | 3   |
| 22 | 2   | 1   |
| 23 | 3   | 2   |
| 24 | 4   | 2   |
| 25 | 1   | 1   |

## Recode Trx (numeric) variable to a factor:

| Numeric value | Level name |
|---------------|------------|
| 1             | drug 1     |
| 2             | drug 2     |
| 3             | drug 3     |

–Set up the null hypothesis:  $H_0 : \mu_A = \mu_B = \mu_C$

–Load and look the data:

Numerical summaries:

|        | mean | sd       | IQR  | 0% | 25% | 50%  | 75%   | 100% | data:n |
|--------|------|----------|------|----|-----|------|-------|------|--------|
| drug 1 | 4.00 | 2.054805 | 4.00 | 1  | 2   | 4.5  | 6.00  | 6    | 10     |
| drug 2 | 5.25 | 1.488048 | 2.25 | 3  | 4   | 5.5  | 6.25  | 7    | 8      |
| drug 3 | 9.60 | 2.302173 | 2.00 | 6  | 9   | 10.0 | 11.00 | 12   | 5      |

Let's see graphically:

R Commander

Fitxer Edita Dades Estadístics Gràfics Models Distribucions Eines Ajuda

Taula de dades: taula

R Script R Markdown

```
Boxplot (SBP~TRX, data=taula)
with(taula.de.dades,
     labels=rownames(tau...
densityPlot (SBP~TRX, ...
with(taula.de.dades,
     labels=rownames(tau...
stripchart (SBP ~ TRX,
stripchart (SBP ~ TRX,
stripchart (SBP ~ TRX,
```

Output

> Boxplot (SBP~TRX, da

Paleta de colors... Vis

Gràfic seqüencial... "y")

Histograma... l.meth

Density estimate... ", ad

Gràfic de tiges i fulles, l.meth

Caixa de dispersió... " , ad

Gràfic quantil-quantil... l.meth

Diagrama de dispersió... " , ad

Matriu de diagrames de dispersió... " , ad

Gràfic de línies... " , ad

Gràfic condicional XY " , ad

Gràfic de les mitjanes... " , ad

Gràfic de franja " , ad

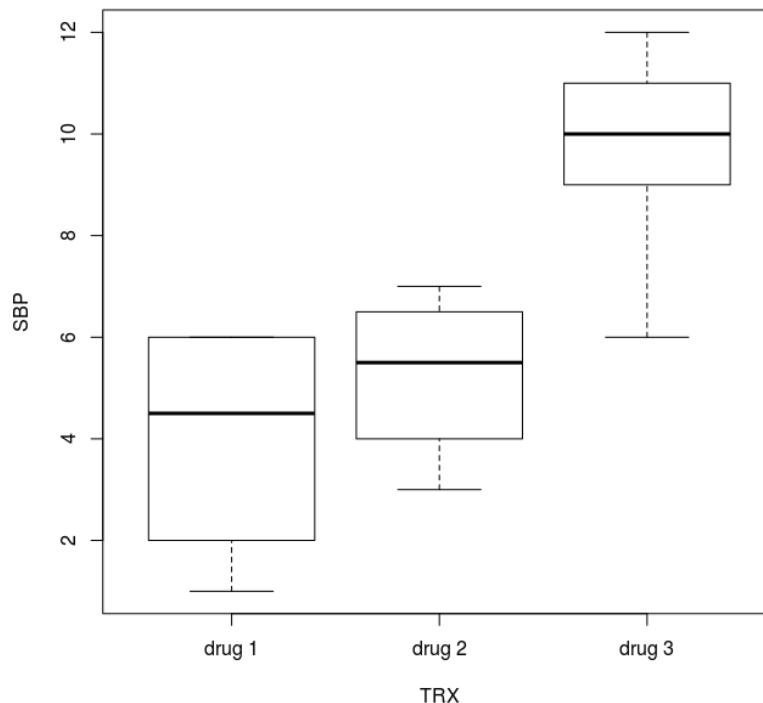
Gràfic de barres... " , ad

Gràfic de sectors... " , ad

Gràfic 3D " , ad

Salva el gràfic a un fitxer... " , ad

R Graphics: Device 2 (ACTIVE)



Screenshot of the R Commander interface showing the selection of a One-way ANOVA analysis.

The R Commander window shows the menu path:

- File
- Edit
- Data
- Statistics**
- Graphs
- Models
- Distributions
- Tools
- Help

The **Means** submenu is open, and the **One-way ANOVA...** option is selected.

The R Script pane contains the following code:

```
})  
showData(SBP)  
maxheight=  
library(abin)  
library(e107)  
numSummariv(SBP[, "SBP"], groups=SBP$TRX, statistics=c("m
```

A large purple arrow points from the R Commander window down to the **One-Way Analysis of Variance** dialog box.

The **One-Way Analysis of Variance** dialog box has the following settings:

- Model name: AnovaModel.2
- Groups (pick one): TRX
- Response Variable (pick one): SBP
- Options:
  - Pairwise comparisons of means
  - Welch F-test not assuming equal variances
- Buttons: Help, Reset, Apply, Cancel, OK

```
> AnovaModel.2 <- aov(SBP ~ TRX, data=sbp)
```

```
> summary(AnovaModel.2)
   Df Sum Sq Mean Sq F value    Pr(>F)
TRX      2 106.5  53.26  14.26 0.000142 ***
Residuals 20  74.7   3.73
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> with(sbp, numSummary(SBP, groups=TRX, statistics=c("mean", "sd")))
   mean      sd data:n
drug 1 4.00 2.054805     10
drug 2 5.25 1.488048      8
drug 3 9.60 2.302173      5
```

## If we use ANOVA to compare means....why is it called “analysis of variance”?

$H_0$ : The means of all the groups are equal

$H_1$ : Not all the means are equal



ANOVA analyzes if there are differences between two variances to deduce if there are differences between the means

ANOVA partitions the total variability in the sample data into two component parts.

- Due to treatments: variation **BETWEEN** groups
- Due to the observations: variation **WITHIN** groups

## Types of analysed variability by ANOVA

| Sample | Drug1 | Drug2 | Drug3 |
|--------|-------|-------|-------|
| 1      | 4     | 7     | 9     |
| 2      | 2     | 6     | 12    |
| 3      | 6     | 5     | 6     |
| 4      | 6     | 7     | 11    |
| 5      | 5     | 6     | 10    |
| 6      | 6     | 4     | 11    |
| 7      | 2     | 7     | 9     |
| 8      | 6     | 5     | 10    |
| Mean   | 4.6   | 5.9   | 9.8   |

$$\text{Total variability} = V_{\text{Between groups}} + V_{\text{Within groups}}$$

## Types of analysed variability by ANOVA

| Sample | Drug1 | Drug2 | Drug3 |
|--------|-------|-------|-------|
| 1      | 4     | 7     | 9     |
| 2      | 2     | 6     | 12    |
| 3      | 6     | 5     | 6     |
| 4      | 6     | 7     | 11    |
| 5      | 5     | 6     | 10    |
| 6      | 6     | 4     | 11    |
| 7      | 2     | 7     | 9     |
| 8      | 6     | 5     | 10    |
| Mean   | 4.6   | 5.9   | 9.8   |

Variations among the conditions represents variation due to “treatment effects”

$$\text{Total variability} = V_{\text{Between groups}} + V_{\text{Within groups}}$$

## Types of analysed variability by ANOVA

| Sample | Drug1 | Drug2 | Drug3 |
|--------|-------|-------|-------|
| 1      | 4     | 7     | 9     |
| 2      | 2     | 6     | 12    |
| 3      | 6     | 5     | 6     |
| 4      | 6     | 7     | 11    |
| 5      | 5     | 6     | 10    |
| 6      | 6     | 4     | 11    |
| 7      | 2     | 7     | 9     |
| 8      | 6     | 5     | 10    |
| Mean   | 4.6   | 5.9   | 9.8   |

Variation among participants within each condition represents “individual differences”

$$\text{Total variability} = V_{\text{Between groups}} + V_{\text{Within groups}}$$

## Types of variability analysed by ANOVA

- Variation BETWEEN groups ( $SS_{\text{Effect}}$ ): for each data value looks at the difference between its group mean and the overall mean



- Variation WITHIN groups ( $SS_{\text{Error}}$ ): for each data value looks at the difference between that value and the mean of its group



**Differences** will be found between groups when:

Variation between  
groups



Variation within  
groups

- Strictly speaking to compare the two sum of squares (SS) they must be standardized.



dividing each sum of squares by their *degrees of freedom*

- The Table of ANOVA

| Source     | DF      | SS                | MS                        | F                      |
|------------|---------|-------------------|---------------------------|------------------------|
| Treatments | $g - 1$ | $SS_{\text{Trt}}$ | $SS_{\text{Trt}}/(g - 1)$ | $MS_{\text{Trt}}/MS_E$ |
| Error      | $N - g$ | $SSE$             | $SSE/(N - g)$             |                        |

The ANOVA F-statistic is a ratio of the Between group variation divided by the Within group variation.

$$F = \text{Between} / \text{Within}$$

- ✓ A large **F** is evidence against  $H_0$ , since it indicates that there are more difference between groups than within groups.

| Source     | DF      | SS                | MS                        | F                      |
|------------|---------|-------------------|---------------------------|------------------------|
| Treatments | $g - 1$ | $SS_{\text{Trt}}$ | $SS_{\text{Trt}}/(g - 1)$ | $MS_{\text{Trt}}/MS_E$ |
| Error      | $N - g$ | $SSE$             | $SSE/(N - g)$             |                        |

```
> summary(AnovaModel1.2)
```

|           | Df | Sum Sq | Mean Sq | F value | Pr(>F)       |
|-----------|----|--------|---------|---------|--------------|
| TRX       | 2  | 106.5  | 53.26   | 14.26   | 0.000142 *** |
| Residuals | 20 | 74.7   | 3.73    |         |              |
| ---       |    |        |         |         |              |

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

ANOVA's Table for two factor design:

Two  $F$  tests:

1. “Age” effective?

$$F = \frac{\text{variation from block}}{\text{variation due to error}} = \frac{MSB}{MSE}$$

2. “Dose” effective?

$$F = \frac{\text{variation from treatment}}{\text{variation due to error}} = \frac{MSTr}{MSE}$$

## EXAMPLE

ANOVA's Table for two factor design:

| Source of Variation | SS          | df        | MS    | F    | p-Value       |
|---------------------|-------------|-----------|-------|------|---------------|
| Age Group           | 793         | 5         | 158.6 | 4.21 | <b>0.0136</b> |
| Dose Level          | 217         | 3         | 72.3  | 1.92 | <b>0.169</b>  |
| Error               | 564         | 15        | 37.6  |      |               |
| <b>TOTAL</b>        | <b>1574</b> | <b>23</b> |       |      |               |

We fail to reject the claim that there is no significant effect due to dose level on gene expression measurements

# Table of contents

- 1. Basic ideas of experimental design**
- 2. Experimental Design conditions**
  1. Randomization
  2. Replication
  3. Local control
- 3. Experimental Design types**
  1. Completely Randomized Design (CRD)
  2. Randomized Block Design (RBD)
  3. Factorial Experiments
  4. Repeated Measures (Within Subjects Designs)
- 4. Introduction to ANOVA**
- 5. How does ANOVA ‘work’?**
- 6. ANOVA assumptions**
- 7. Beyond ANOVA**
  1. Multiple comparisons
  2. Non parametric ANOVA
- 8. Exercises**

**ANOVA's assumptions have to be checked in the data that they are true before run the model:**

1. Each group is approximately normal (normally distributed).
2. The variances of different samples are homogeneous (homocedasticity)
3. The errors are independent from observation to observation.



**How to check them?**

## Normality of the data.

- Generally deviations from the normality do not seriously affect the validity of the assumptions. Not seriously outliers.
- F-test is very robust against non-normal data.
- The best way to check the normality of the data is a QQ plot (Shapiro-Wilk test is also valid)

## Homocedasticity of the variances.

- F-test is very robust against heterogeneity of variances, especially with fixed factors and equal sample sizes (*balanced designs*)
- There are some statistical test, like Breusch-Pagan and Levene's test, to check that

## Homocedasticity of the variances (in Rcommander)



$H_0$ : error variances are equal

```
> leveneTest(SBP ~ TRX, data=sbp, center="median")
```

```
Levene's Test for Homogeneity of Variance (center = "median")
  Df F value Pr(>F)
group  2 0.6501 0.5327
      20
```

```
> bartlett.test(SBP ~ TRX, data=sbp)
```

**Bartlett test of homogeneity of variances**

```
data: SBP by TRX
Bartlett's K-squared = 1.075, df = 2, p-value = 0.5842
```

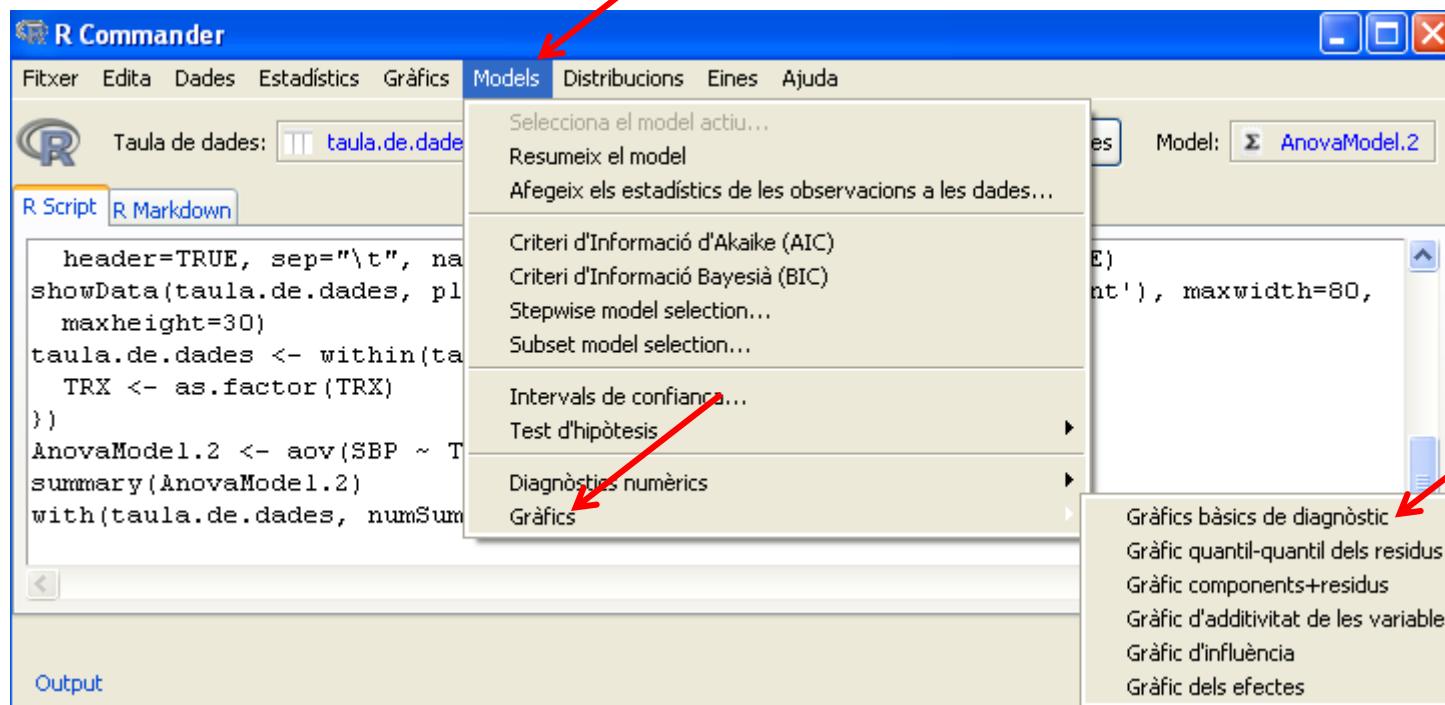
## Independence of the residues.

- Independent observations
  - ✓ No correlation between error terms
  - ✓ No correlation between independent variables and error
- If the residues are independent they won't have to follow any clear pattern when we observe them in a plot.



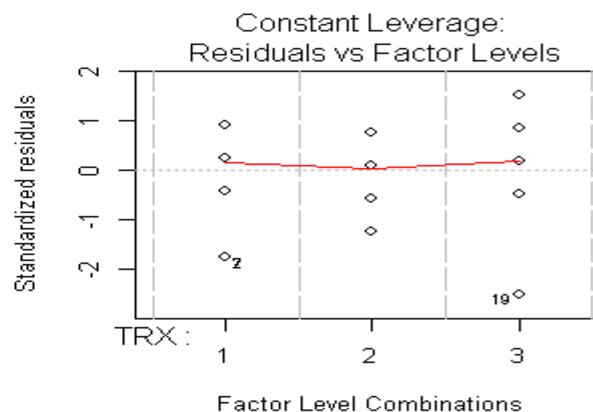
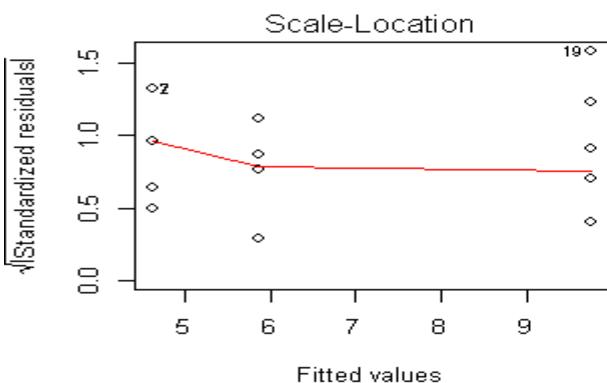
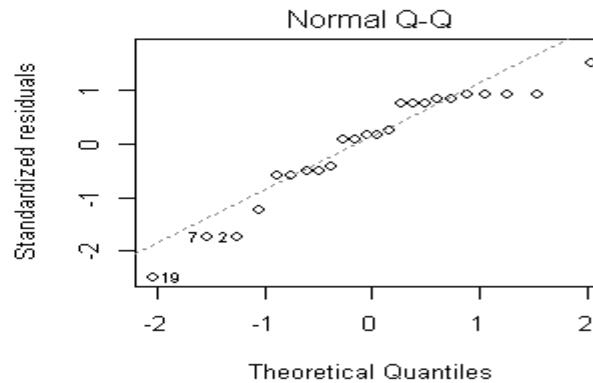
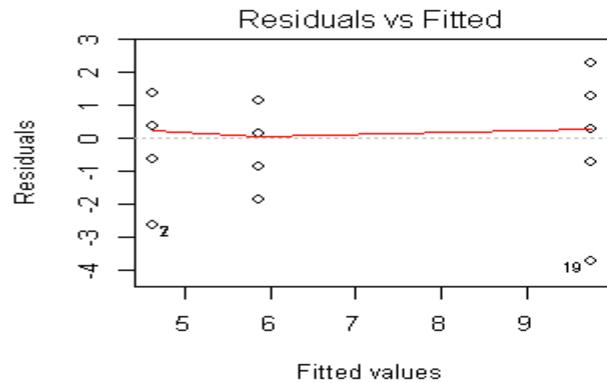
We observe a plot of the residues vs estimates values

## •How to check in R-Commander



- How to check in R-Commander.

aov(SBP ~ TRX)



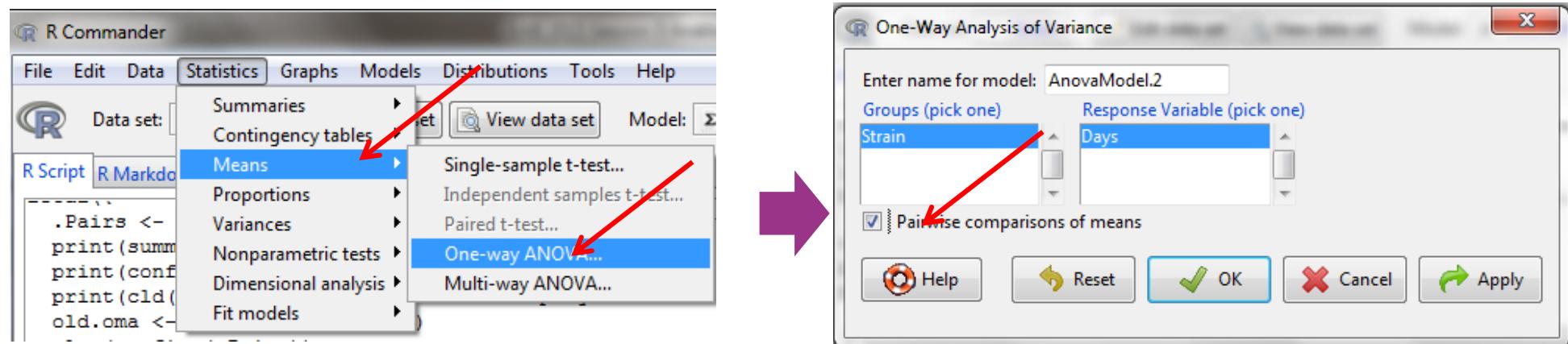
# Table of contents

- 1. Basic ideas of experimental design**
- 2. Experimental Design conditions**
  1. Randomization
  2. Replication
  3. Local control
- 3. Experimental Design types**
  1. Completely Randomized Design (CRD)
  2. Randomized Block Design (RBD)
  3. Factorial Experiments
  4. Repeated Measures (Within Subjects Designs)
- 4. Introduction to ANOVA**
- 5. How does ANOVA ‘work’?**
- 6. ANOVA assumptions**
- 7. Beyond ANOVA**
  1. Multiple comparisons
  2. Non parametric ANOVA
- 8. Exercises**

Until now we only can say if there differences among the groups compared, but we don't know between which groups.

- Usually we are interested in the comparison of the samples or treatments two by two (remember we couldn't use a t-test!!!)
- We need to adjust our p-value threshold because we are doing multiple tests with the same data (type I error probability increases). There are some tests that take into account this.

## Tukey-Kramer in R-Commander



A priori comparisons result is displayed again, plus...

```
> summary(AnovaModel.2)
      Df Sum Sq Mean Sq F value    Pr(>F)
Strain      4 148.2   37.05   29.88 3.39e-09 ***
Residuals  25  31.0    1.24
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> with(Strain, numSummary(Days, groups=Strain, statistics=c("mean", "sd")))
      mean      sd data:n
StrainA 10.000000 1.4142136     6
StrainB  6.166667 1.1690452     6
StrainC  8.666667 1.0327956     6
StrainD 11.500000 1.0488088     6
StrainE  5.666667 0.8164966     6
```

# Tukey-Kramer in R-Commander

## Simultaneous Tests for General Linear Hypotheses

### Multiple Comparisons of Means: Tukey Contrasts

```
Fit: aov(formula = Days ~ Strain, data = Strain)
```

Linear Hypotheses:

|                   |      | Estimate | Std. Error | t value | Pr(> t )     |
|-------------------|------|----------|------------|---------|--------------|
| StrainB - StrainA | == 0 | -3.8333  | 0.6429     | -5.962  | < 1e-04 ***  |
| StrainC - StrainA | == 0 | -1.3333  | 0.6429     | -2.074  | 0.262331     |
| StrainD - StrainA | == 0 | 1.5000   | 0.6429     | 2.333   | 0.167875     |
| StrainE - StrainA | == 0 | -4.3333  | 0.6429     | -6.740  | < 1e-04 ***  |
| StrainC - StrainB | == 0 | 2.5000   | 0.6429     | 3.889   | 0.005444 **  |
| StrainD - StrainB | == 0 | 5.3333   | 0.6429     | 8.296   | < 1e-04 ***  |
| StrainE - StrainB | == 0 | -0.5000  | 0.6429     | -0.778  | 0.934683     |
| StrainD - StrainC | == 0 | 2.8333   | 0.6429     | 4.407   | 0.001467 **  |
| StrainE - StrainC | == 0 | -3.0000  | 0.6429     | -4.666  | 0.000793 *** |
| StrainE - StrainD | == 0 | -5.8333  | 0.6429     | -9.073  | < 1e-04 ***  |
| ---               |      |          |            |         |              |

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

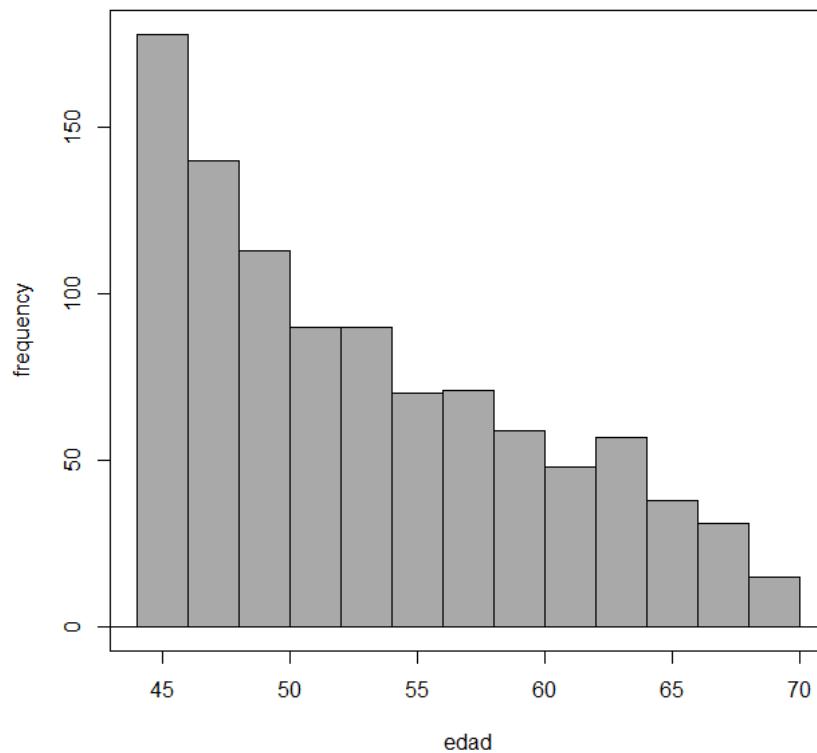
(Adjusted p values reported -- single-step method)

## Kruskal-Wallis Test (Non parametric ANOVA)

- Kruskal-Wallis test is used to find differences among  $k$  experimental samples or treatments
- This test don't assume the normality of the variables.
- It is similar to ANOVA, but it uses the ranges
- Used when number of observations are little and unbalanced designs.

## Kruskal-Wallis Test in Rcommander

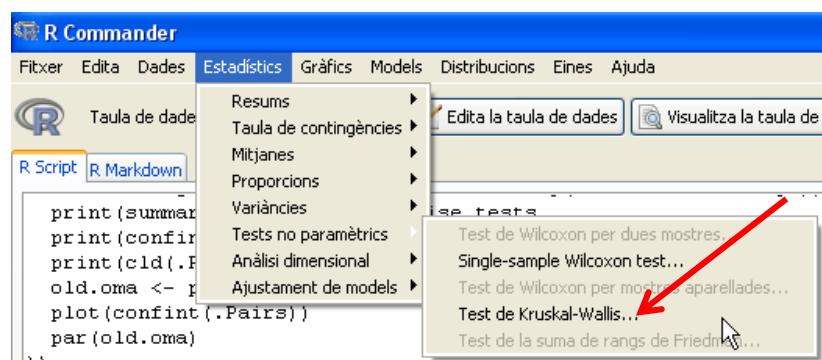
“Osteo” data set. Check if “age” is the same in the women of the study (classified by degree of illness)



## 7. Beyond ANOVA

### Kruskal-Wallis Test in Rcommander

Perform the test.

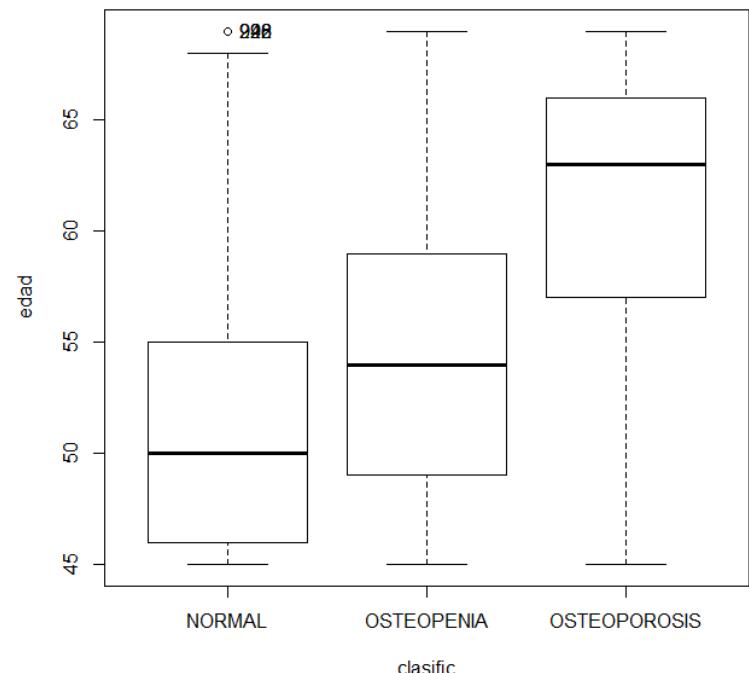


R Commander interface showing the 'Estadístics' menu open. The 'Test de Kruskal-Wallis...' option is highlighted with a red arrow.

```
> kruskal.test(edad ~ clasific, data=osteo)

  Kruskal-Wallis rank sum test

data: edad by clasific
Kruskal-Wallis chi-squared = 112.0766, df = 2, p-value < 2.2e-16
```



# Table of contents

- 1. Basic ideas of experimental design**
- 2. Experimental Design conditions**
  1. Randomization
  2. Replication
  3. Local control
- 3. Experimental Design types**
  1. Completely Randomized Design (CRD)
  2. Randomized Block Design (RBD)
  3. Factorial Experiments
  4. Repeated Measures (Within Subjects Designs)
- 4. Introduction to ANOVA**
- 5. How does ANOVA ‘work’?**
- 6. ANOVA assumptions**
- 7. Beyond ANOVA**
  1. Multiple comparisons
  2. Non parametric ANOVA
- 8. Exercises**

# Exercise

With the dataset osteo, find out if there are some relation between the *age of the women (grupedad)* and *the broadband ultrasound attenuation (bua)*. After that try to analyze among which groups. The desirable steps to follow would be:

1. Exploratory data analysis (numerically and graphically)
2. Run the model.
3. Check for assumptions of the model (graphically)
4. Run post hoc comparison if necessary

## 1. Exploratory data analysis (numerically and graphically)

R Commander

Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda

Conjunto de datos

R Script R Markdown

```
osteos <- read.table("osteos.csv", header=TRUE, sep=";")  
library(mvtnorm, survival, MASS, TH.data, multcomp)
```

Resúmenes ►

- Conjunto de datos activo
- Resúmenes numéricos... **Resúmenes numéricos...**
- Distribución de frecuencias...
- Número de observaciones ausentes
- Tabla de estadísticas...
- Matriz de correlaciones...
- Test de correlación
- Test de normalidad...
- Transformación para normalizar...

osteos <- read.table("osteos.csv", header=TRUE, sep=";")  
library(mvtnorm, survival, MASS, TH.data, multcomp)

Resúmenes numéricos

Datos Estadísticos

Variables (seleccione una o más)

- area
- bua
- edad
- edad\_men
- imc
- menarqui

Resumir por grupos...

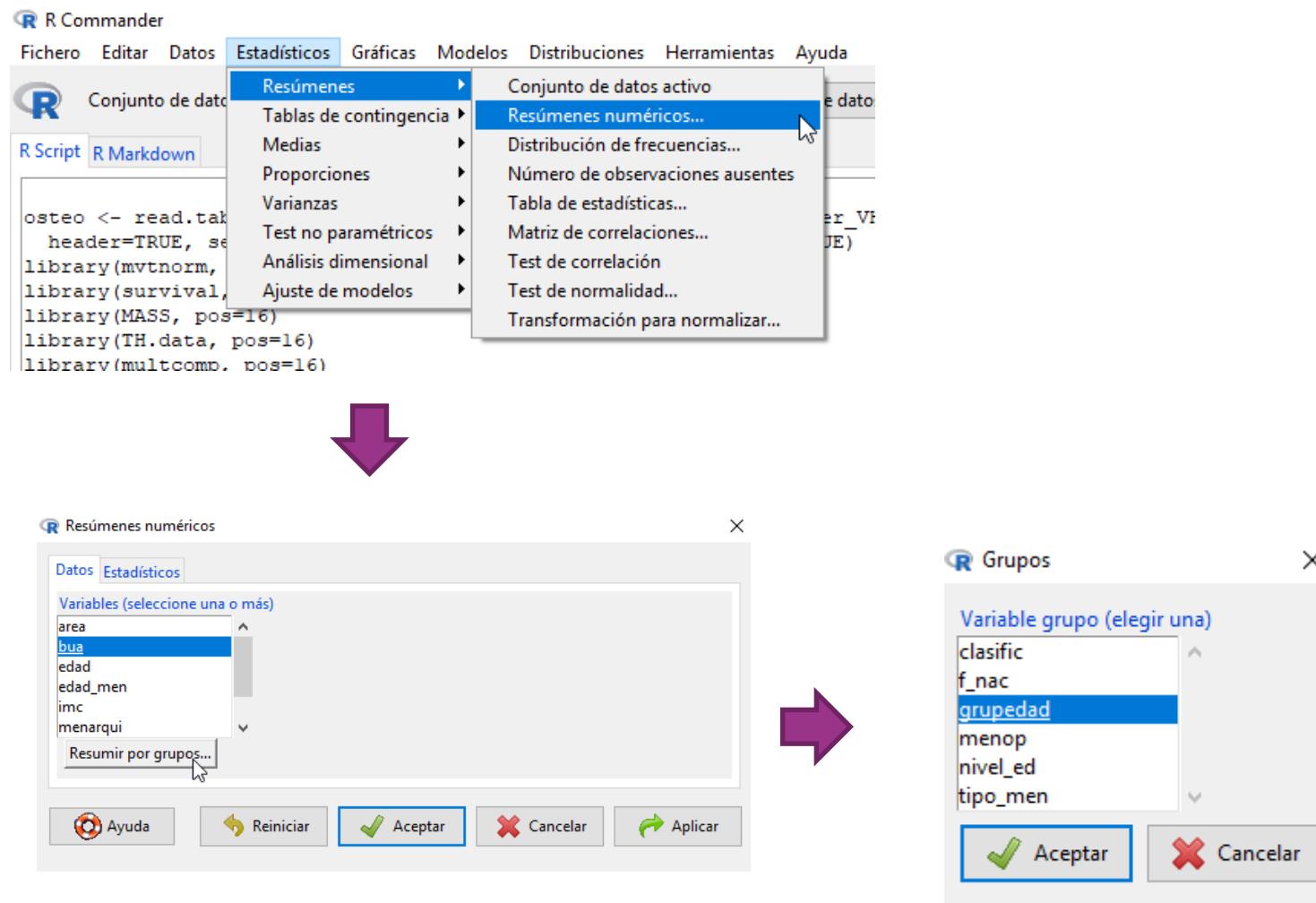
Aceptar Cancelar Aplicar

Grupos

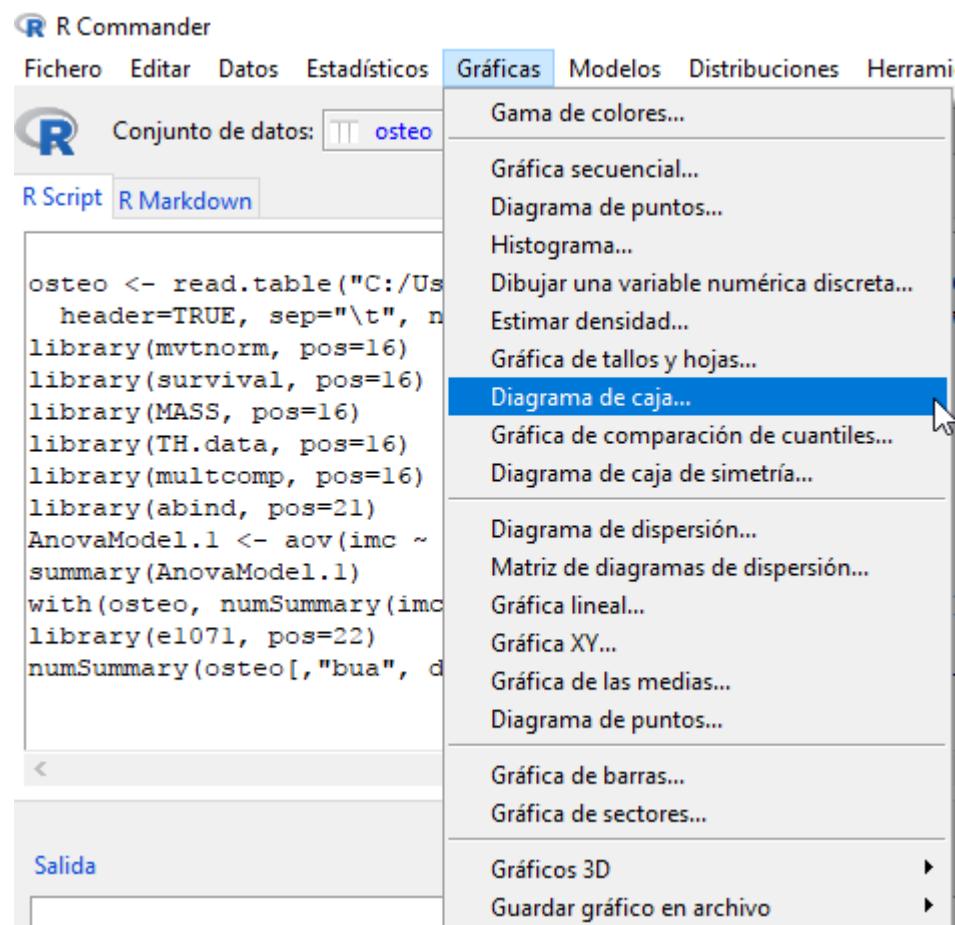
Variable grupo (elegir una)

- clasific
- f\_nac
- grupedad
- menop
- nivel\_ed
- tipo\_men

Aceptar Cancelar



## 1. Exploratory data analysis (numerically and graphically)



The screenshot shows the R Commander interface. The 'Gráficas' tab is selected in the top menu bar. A context menu is open over a script named 'osteob'. The menu items are:

- Gama de colores...
- Gráfica secuencial...
- Diagrama de puntos...
- Histograma...
- Dibujar una variable numérica discreta...
- Estimar densidad...
- Gráfica de tallos y hojas...
- Diagrama de caja...** (highlighted with a blue selection bar)
- Gráfica de comparación de cuantiles...
- Diagrama de caja de simetría...
- Diagrama de dispersión...
- Matriz de diagramas de dispersión...
- Gráfica lineal...
- Gráfica XY...
- Gráfica de las medias...
- Diagrama de puntos...
- Gráfica de barras...
- Gráfica de sectores...
- Gráficos 3D
- Guardar gráfico en archivo

The script pane contains the following R code:

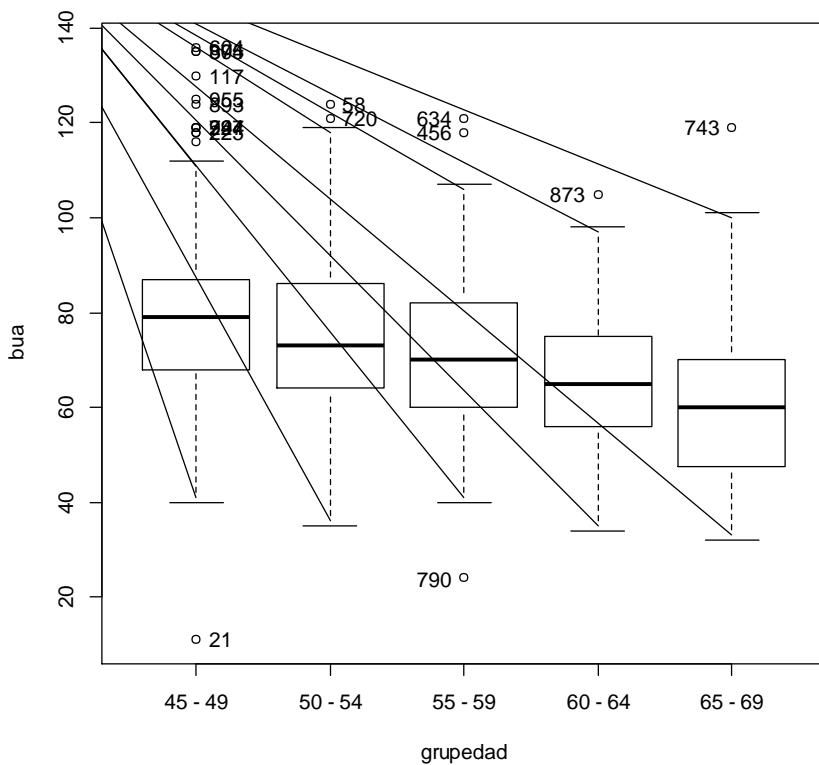
```
osteob <- read.table("C:/Us  
header=TRUE, sep="\t", n  
library(mvtnorm, pos=16)  
library(survival, pos=16)  
library(MASS, pos=16)  
library(TH.data, pos=16)  
library(multcomp, pos=16)  
library(abind, pos=21)  
AnovaModel.1 <- aov(imc ~  
summary(AnovaModel.1)  
with(osteob, numSummary(imc  
library(e1071, pos=22)  
numSummary(osteob[, "bua", d
```

The output pane is labeled 'Salida'.

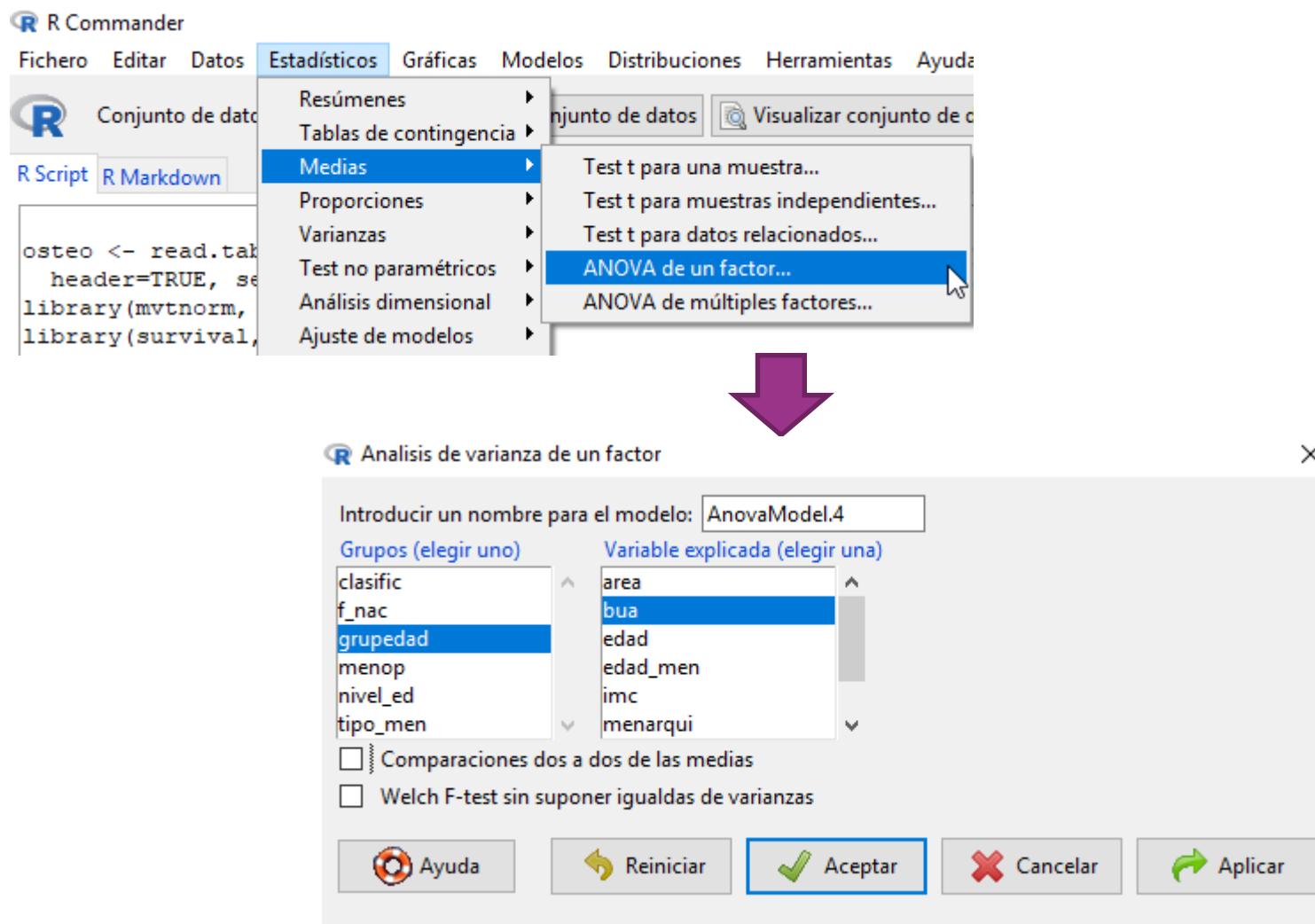
## 1. Exploratory data analysis (numerically and graphically)

```
> numSummary(osteо[, "bua", drop=FALSE], groups=osteо$grupedad, statistics=c("mean", "sd", "IQR", "quantiles"), quantiles=c(0,.25,.5,.75,1))
```

|         | mean     | sd       | IQR   | 0% | 25%   | 50% | 75% | 100% | bua:n |
|---------|----------|----------|-------|----|-------|-----|-----|------|-------|
| 45 - 49 | 78.75926 | 16.24163 | 19.00 | 11 | 68.00 | 79  | 87  | 136  | 378   |
| 50 - 54 | 75.05150 | 15.78921 | 22.00 | 35 | 64.00 | 73  | 86  | 124  | 233   |
| 55 - 59 | 71.43182 | 15.86248 | 22.00 | 24 | 60.00 | 70  | 82  | 121  | 176   |
| 60 - 64 | 64.89147 | 13.91225 | 19.00 | 34 | 56.00 | 65  | 75  | 105  | 129   |
| 65 - 69 | 60.66667 | 15.92704 | 22.25 | 32 | 47.75 | 60  | 70  | 119  | 84    |



## 2. Run the model



The screenshot shows the R Commander interface. The 'Estadísticos' menu is open, and the 'ANOVA de un factor...' option is selected. A large purple arrow points down to the 'Analisis de varianza de un factor' dialog box.

**R Commander**

Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda

Conjunto de datos R Script R Markdown

```
osteо <- read.table("...")  
header=TRUE, sep=","  
library(mvtnorm,  
library(survival,
```

Resúmenes Conjunto de datos Visualizar conjunto de datos

Tablas de contingencia

Medias Test t para una muestra...

Proporciones Test t para muestras independientes...

Varianzas Test t para datos relacionados...

Test no paramétricos ANOVA de un factor...

Análisis dimensional ANOVA de múltiples factores...

Ajuste de modelos

**Analisis de varianza de un factor**

Introducir un nombre para el modelo: AnovaModel.4

Grupos (elegir uno) Variable explicada (elegir una)

clasific  
f\_nac  
**grupedad**  
menop  
nivel\_ed  
tipo\_men

area  
**bua**  
edad  
edad\_men  
imc  
menarqui

Comparaciones dos a dos de las medias

Welch F-test sin suponer igualdades de varianzas

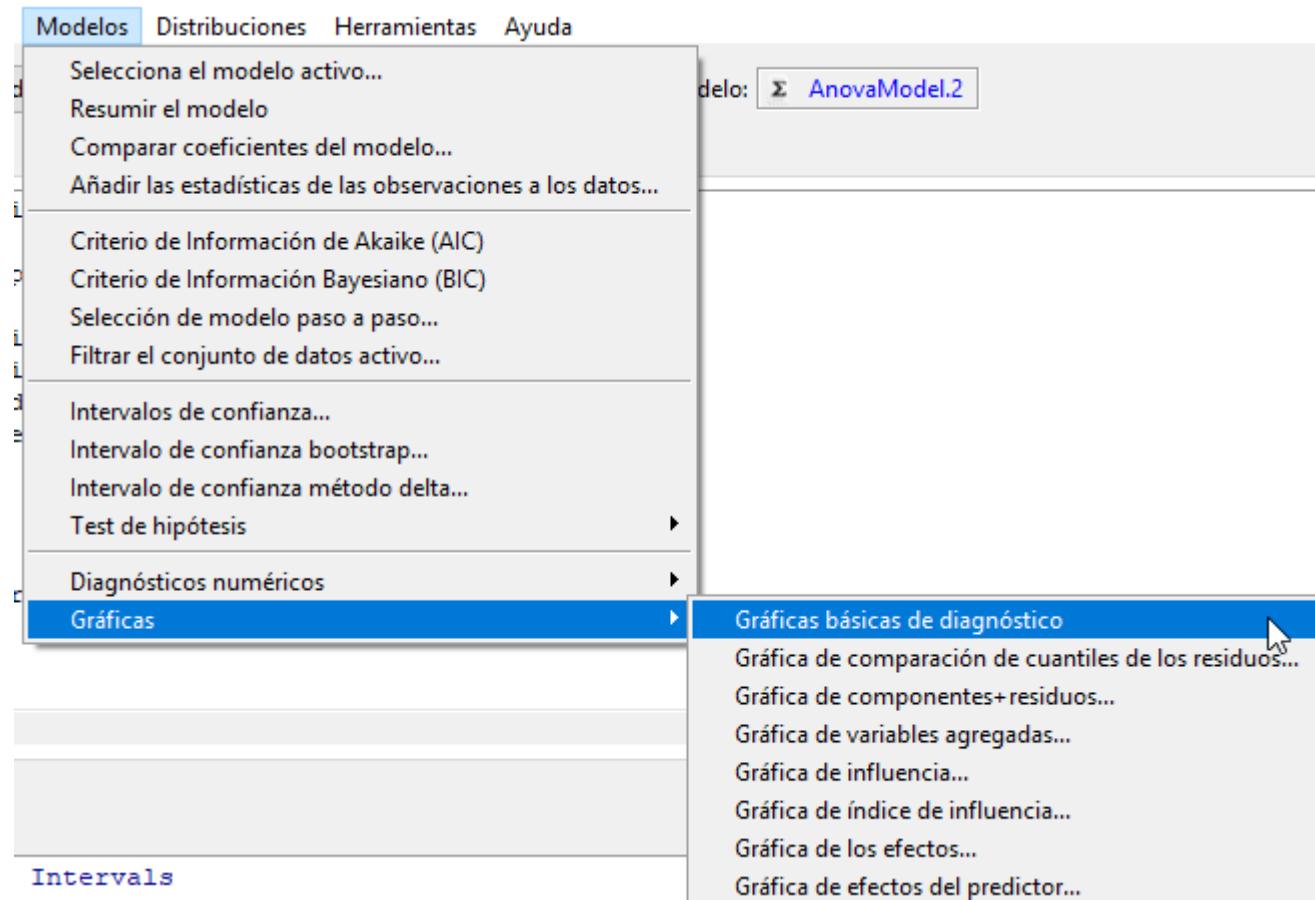
Ayuda Reiniciar Aceptar Cancelar Aplicar

## 2. Run the model

```
> AnovaModel.5 <- aov(bua ~ grupedad, data=osteo)

> summary(AnovaModel.5)
      Df Sum Sq Mean Sq F value Pr(>F)
grupedad     4 35122   8780   35.35 <2e-16 ***
Residuals  995 247149     248
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

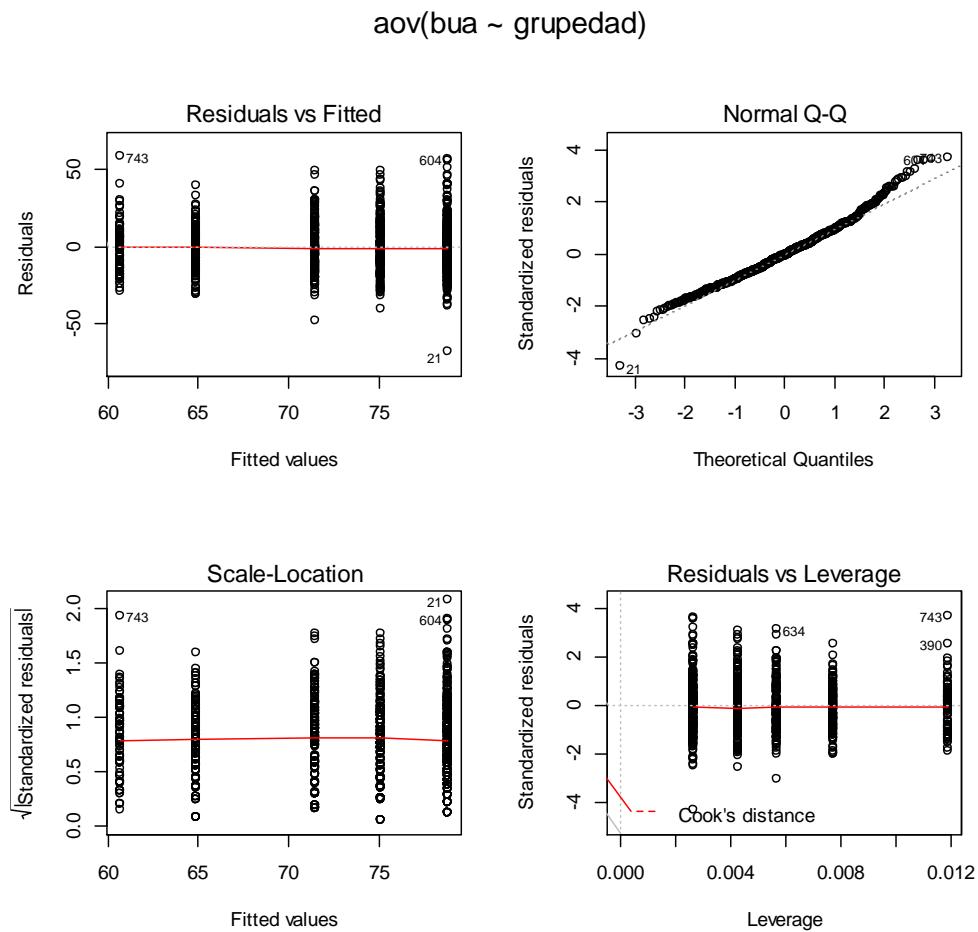
### 3. Check for assumptions of the model (graphically)



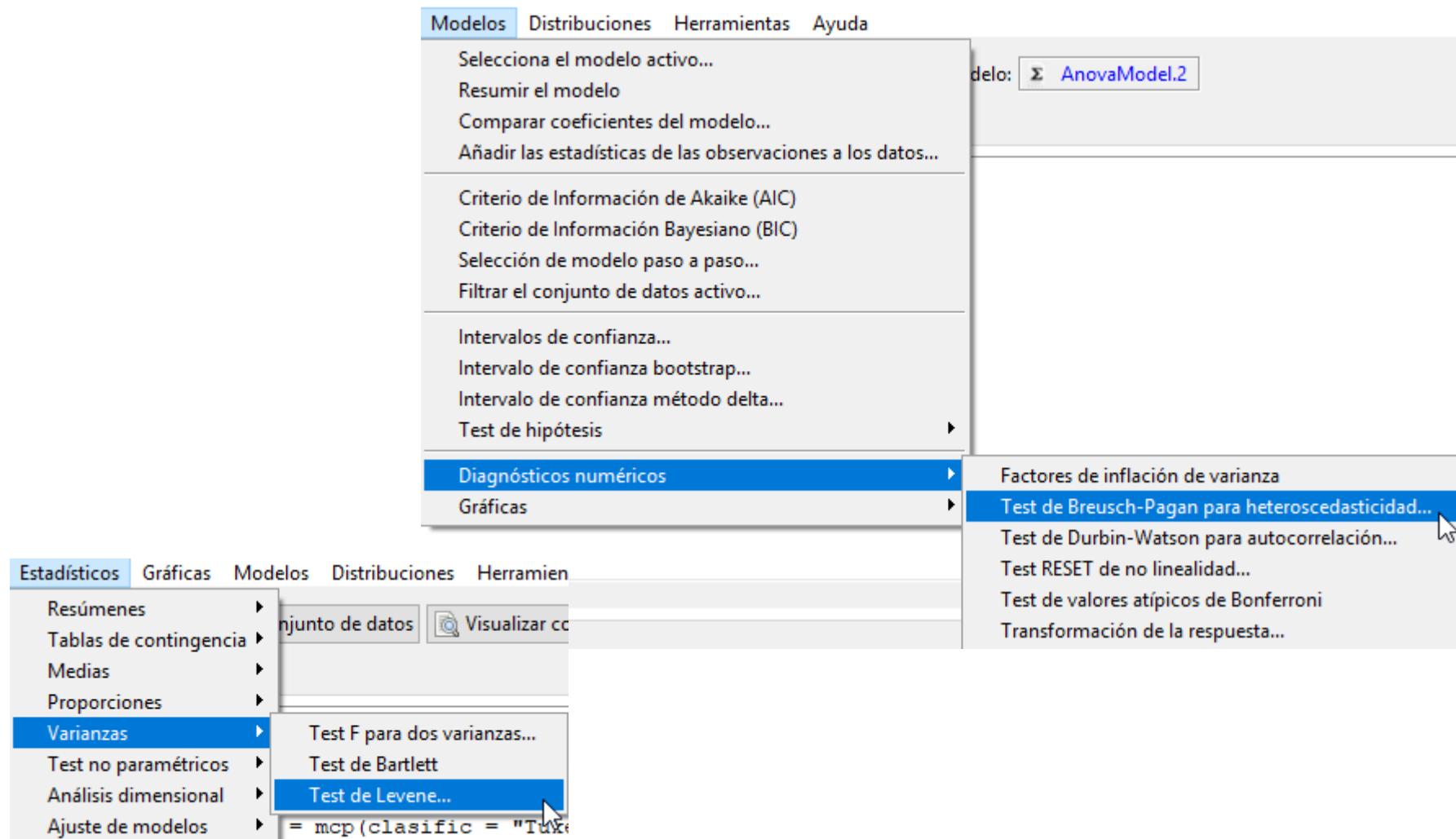
The screenshot shows the SPSS Modeler interface with the 'Modelos' tab selected in the top menu bar. A context menu is open over a model named 'AnovaModel.2'. The menu path 'Modelos > Gráficas > Gráficas básicas de diagnóstico' is highlighted. The following options are visible in the 'Gráficas básicas de diagnóstico' submenu:

- Gráfica de comparación de cuantiles de los residuos...
- Gráfica de componentes+residuos...
- Gráfica de variables agregadas...
- Gráfica de influencia...
- Gráfica de índice de influencia...
- Gráfica de los efectos...
- Gráfica de efectos del predictor...

### 3. Check for assumptions of the model (graphically)



### 3. Check for assumptions of the model (graphically)



The screenshot shows the SPSS software interface with the 'Modelos' (Models) menu open. The 'Diagnósticos numéricos' (Numerical Diagnostics) option is selected, and its submenu is visible. The 'Test de Levene...' option is highlighted with a blue selection bar. The top status bar displays the message 'Modelo: AnovaModel.2'.

Modelos Distribuciones Herramientas Ayuda

Selecciona el modelo activo...  
Resumir el modelo  
Comparar coeficientes del modelo...  
Añadir las estadísticas de las observaciones a los datos...

Criterio de Información de Akaike (AIC)  
Criterio de Información Bayesiano (BIC)  
Selección de modelo paso a paso...  
Filtrar el conjunto de datos activo...

Intervalos de confianza...  
Intervalo de confianza bootstrap...  
Intervalo de confianza método delta...  
Test de hipótesis

Diagnósticos numéricos

Gráficas

Factores de inflación de varianza  
Test de Breusch-Pagan para heteroscedasticidad...  
Test de Durbin-Watson para autocorrelación...  
Test RESET de no linealidad...  
Test de valores atípicos de Bonferroni  
Transformación de la respuesta...

Estadísticos Gráficas Modelos Distribuciones Herramientas

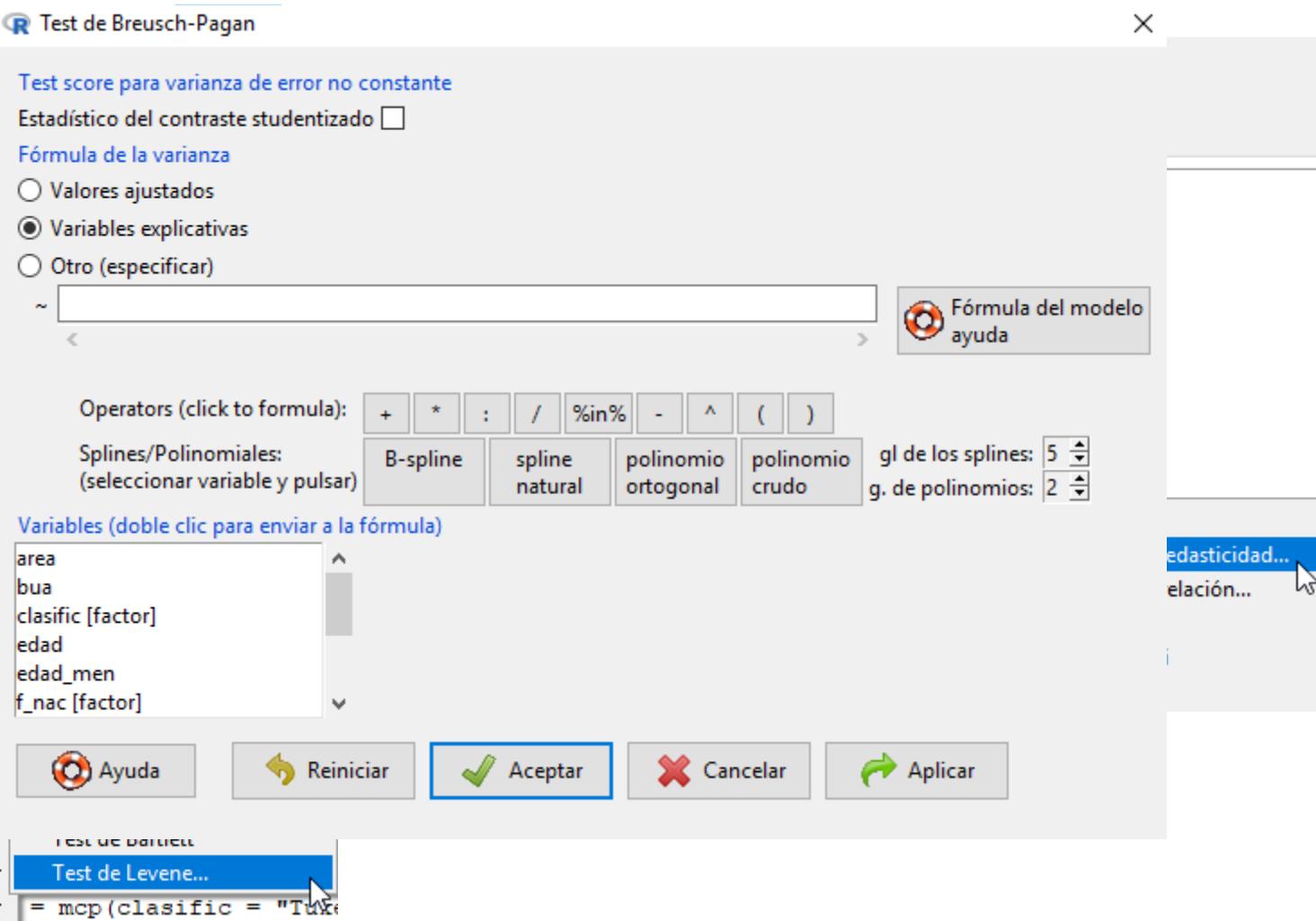
Resúmenes Tablas de contingencia Medias Proporciones Varianzas Test no paramétricos Análisis dimensional Ajuste de modelos

conjunto de datos Visualizar co

Test F para dos varianzas...  
Test de Bartlett  
Test de Levene...

= mcp (clasific = "Tukey")

### 3. Check for assumptions of the model (graphically)

 Test de Breusch-Pagan

Test score para varianza de error no constante

Estadístico del contraste studentizado

Fórmula de la varianza

Valores ajustados

Variables explicativas

Otro (especificar) ~

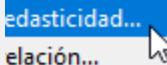
 Fórmula del modelo  
ayuda

Operators (click to formula): + \* : / %in% - ^ ( )

Splines/Polinomiales: gl de los splines: 5  
(seleccionar variable y pulsar) g. de polinomios: 2

Variables (doble clic para enviar a la fórmula)

- area
- bua
- clasic [factor]
- edad
- edad\_men
- f\_nac [factor]

 elasticidad...  
elación...

Estadísticos Gráficas Mo

- Resúmenes
- Tablas de contingencia
- Medias
- Proporciones
- Varianzas**
- Test no paramétricos
- Análisis dimensional
- Ajuste de modelos

 Ayuda Reiniciar Aceptar Cancelar Aplicar

TEST DE BREUSCH-PAGAN

Test de Levene...

= mcp (clasific = "Tux")

### 3. Check for assumptions of the model (graphically)

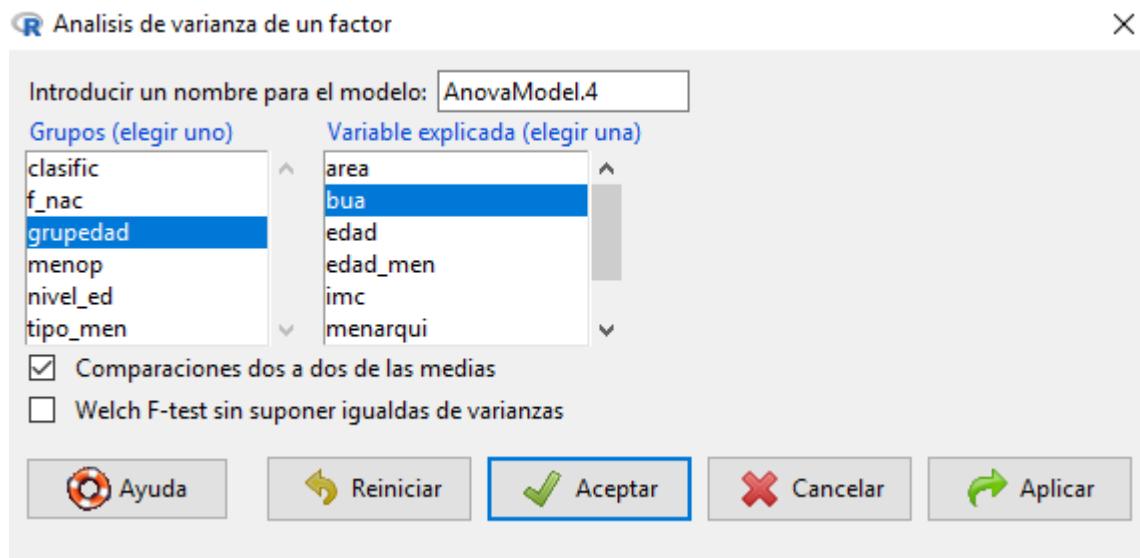
```
> leveneTest(bua ~ grupedad, data=osteo, center="median")
Levene's Test for Homogeneity of Variance (center = "median")
  Df F value Pr(>F)
group   4  0.4487 0.7734
      995
```

```
> bptest(bua ~ grupedad, studentize=FALSE, data=osteo)

Breusch-Pagan test

data: bua ~ grupedad
BP = 4.0163, df = 4, p-value = 0.4038
```

#### 4. Run post hoc comparison if necessary



## 4. Run post hoc comparison if necessary

```
Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: aov(formula = bua ~ clasific, data = osteo)

Linear Hypotheses:
Estimate Std. Error t value Pr(>|t|)
OSTEOPENIA - NORMAL == 0     -23.4279    0.6185  -37.88  <2e-16 ***
OSTEOPOROSIS - NORMAL == 0   -44.5249    1.2608  -35.31  <2e-16 ***
OSTEOPOROSIS - OSTEOPENIA == 0 -21.0970    1.2611  -16.73  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

## Practical exercise of Biostatistics. Master VHIR 2018-19

In the database information of 149 diabetics patients followed up during 17 years is available. On the table below time to follow up after a baseline exam, vital status and other potential prognostic factors are provided. Age at study entry(EDAT), Body Mass Index (BMI), Age at diagnosis(EDATDIAGNO), Smoking habits( TABAC), systolic (SBP) and diastolic(DBP) blood pressure, Electrocardiogram result (ECG) and Coronary Heart Disease(CHD) are those prognosis factors.

- Is there a relationship among blood pressure and the result of ECG