

Quantitative Tutorials - Session 3

ANOVA

1/12/22 -Week 8

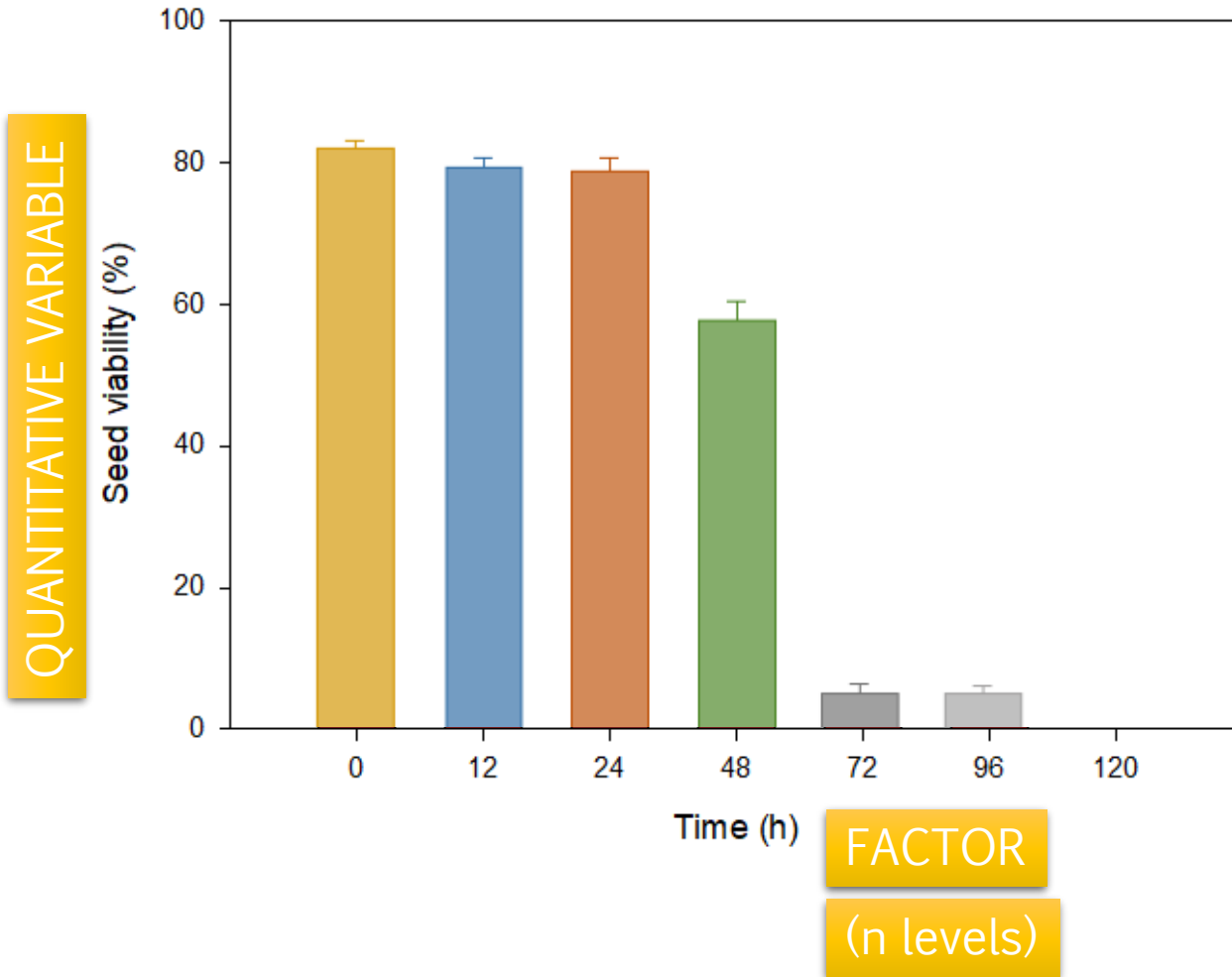


Dr. Erola Fenollosa



Are they different?

Are there significant differences in viability between times?



Hypothesis contrast

$H_0 \rightarrow$ All are the same

$H_1 \rightarrow$ At least one is different

Presumption of innocence

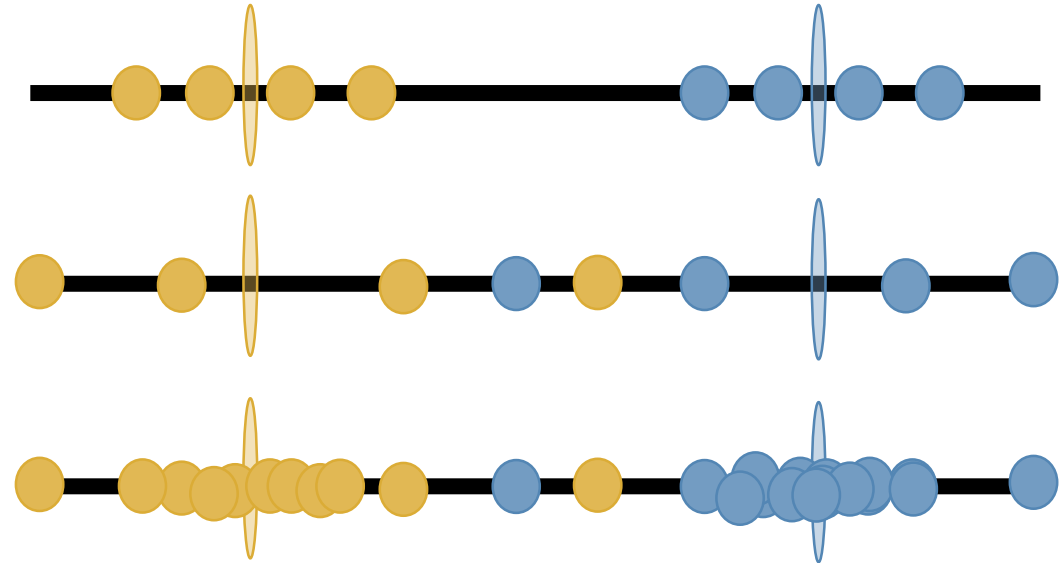


QUICK TASK: Draw in a paper an approximated plot of your factor and variable

ANOVA = Analysis of the variance

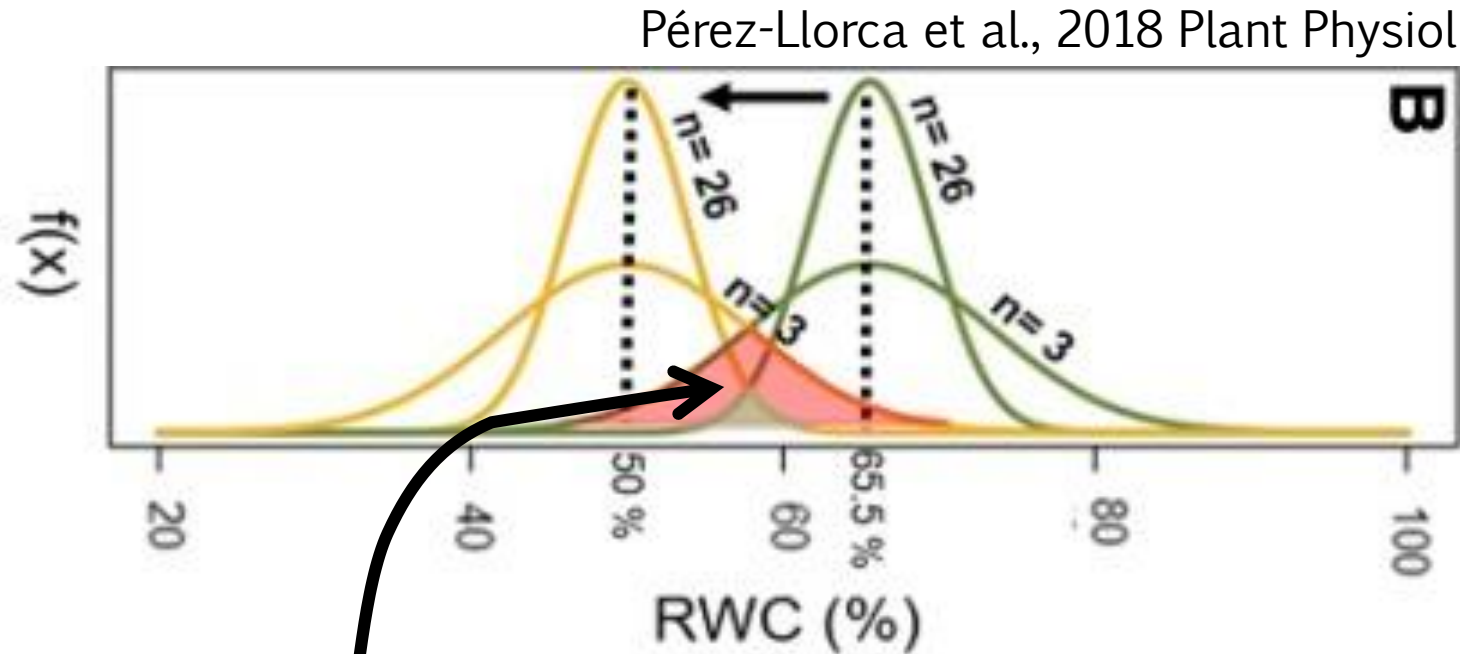
But it's more than that!

- Difference between means
- Variance
- Sample size



A **one-way ANOVA** calculates the variance associated with the levels of the factor and compares the mean squares among groups versus within groups

Visual help:



Alpha or confidence level 0.05

Why do we call it one-way?

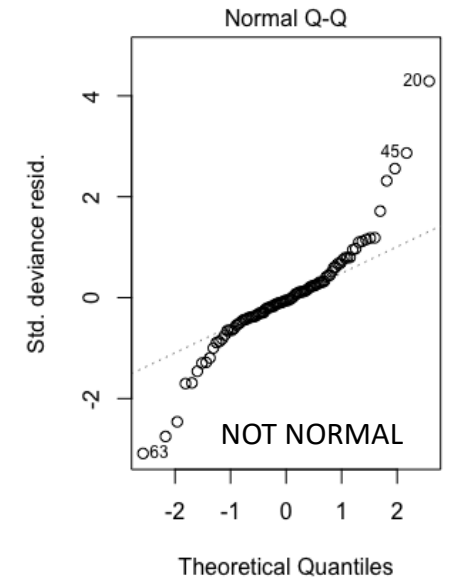
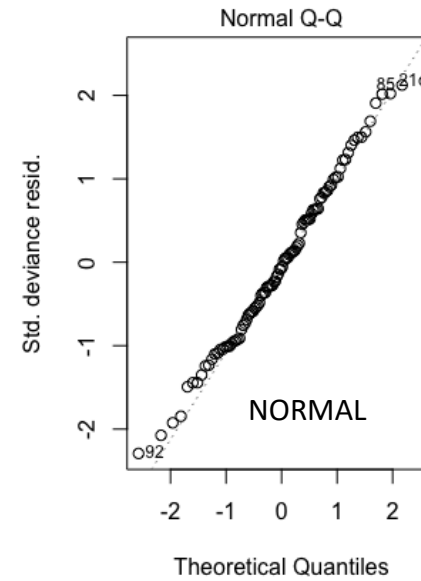
ANOVA requisites

- Normal distribution of the residuals

Kolmogórov-Smirnov 

Shapiro Wilk 

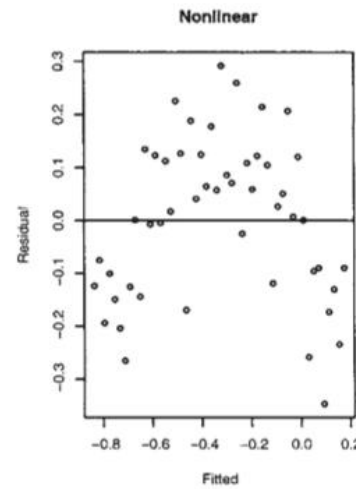
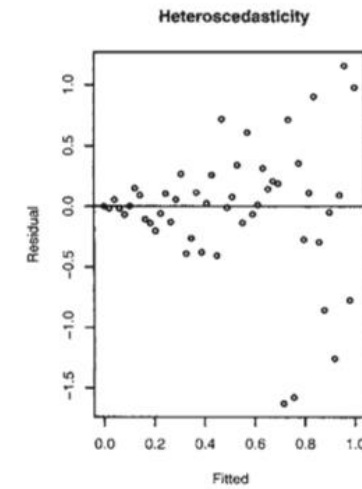
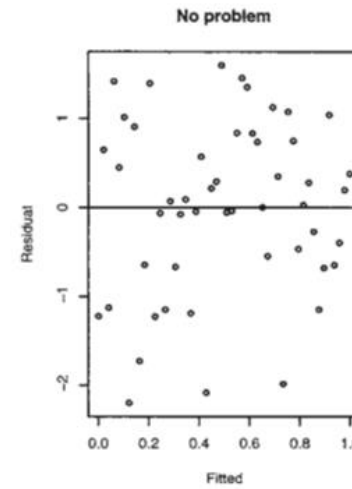
Q-Q plot 



- Homoscedasticity (equal variance of the residuals)

Levene 

Residuals vs. Fitted plot 

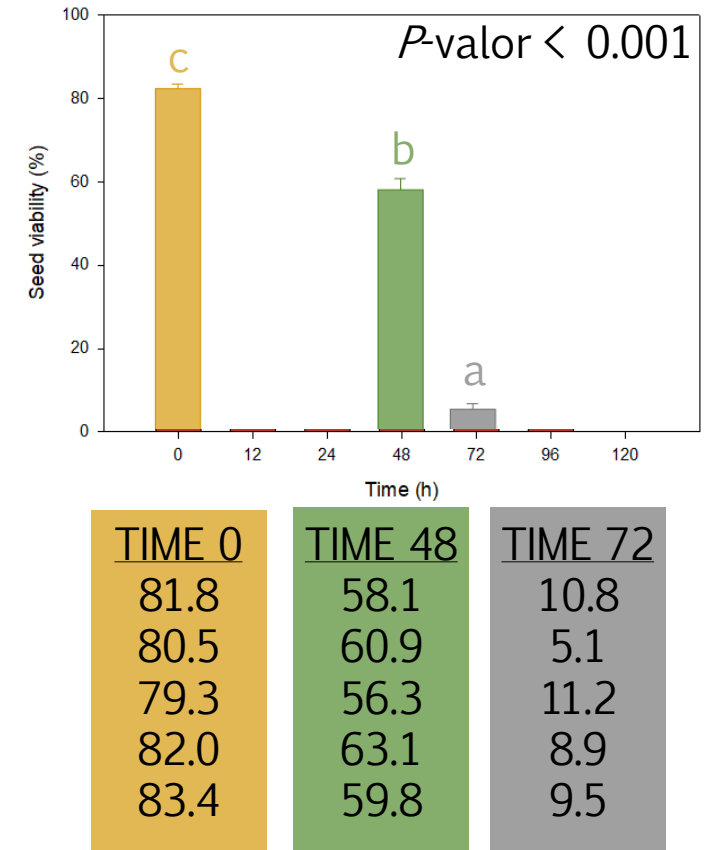


<https://stats.stackexchange.com/questions/76226/interpreting-the-residuals-vs-fitted-values-plot-for-verifying-the-assumptions>



The ANOVA workflow

- 1) Prepare data and examine it first
- 2) Test ANOVA requisites
 - 1) Normal distribution (P -valor > 0.05)
 - 2) Homocedasticity (P -valor > 0.05)
- 3) Test ANOVA (P -valor < 0.05)
- 4) Test post-hoc (opcional)



Post-hoc??

Also called: multiple comparisons.

Only when we have more than two levels in a factor, to know actually which ones are different

Too much options...

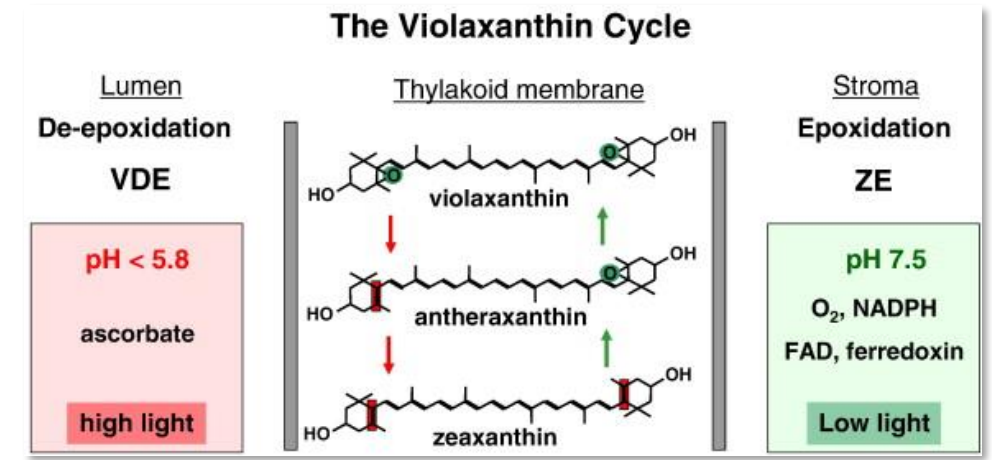
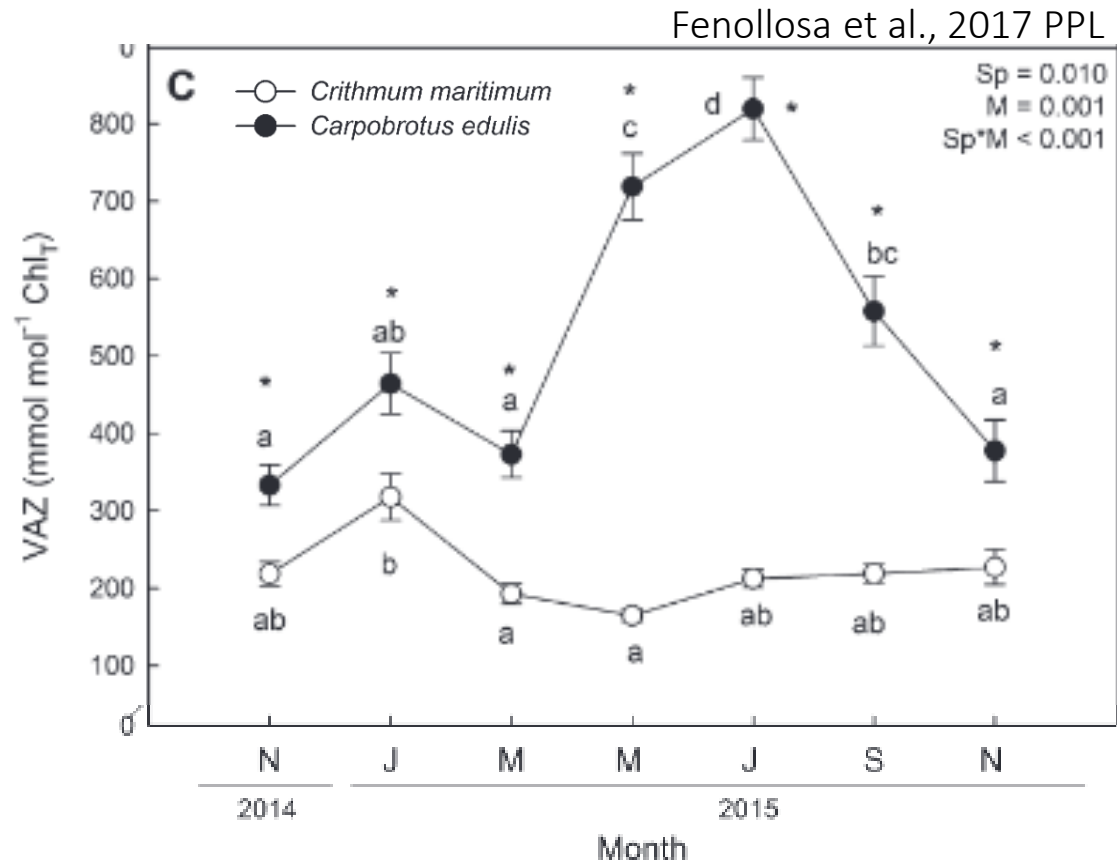
- Bonferroni Procedure
- Duncan's new multiple range test (MRT)
- Dunn's Multiple Comparison Test
- Fisher's Least Significant Difference (LSD)
- Holm-Bonferroni Procedure
- Newman-Keuls
- Rodger's Method
- Scheffé's Method
- Tukey's Test (see also: Studentized Range Distribution)
- Dunnett's correction
- Benjamini-Hochberg (BH) procedure



<https://www.laverdad.es/sociedad/trabajo-oscura-20190315105523-ntvo.html>

Some are more restrictive than others
So expose always which one you've used.

Explain these results:



Jahns et al., 2009 BBA Bioenergetics

Crithmum maritimum



Carpobrotus edulis

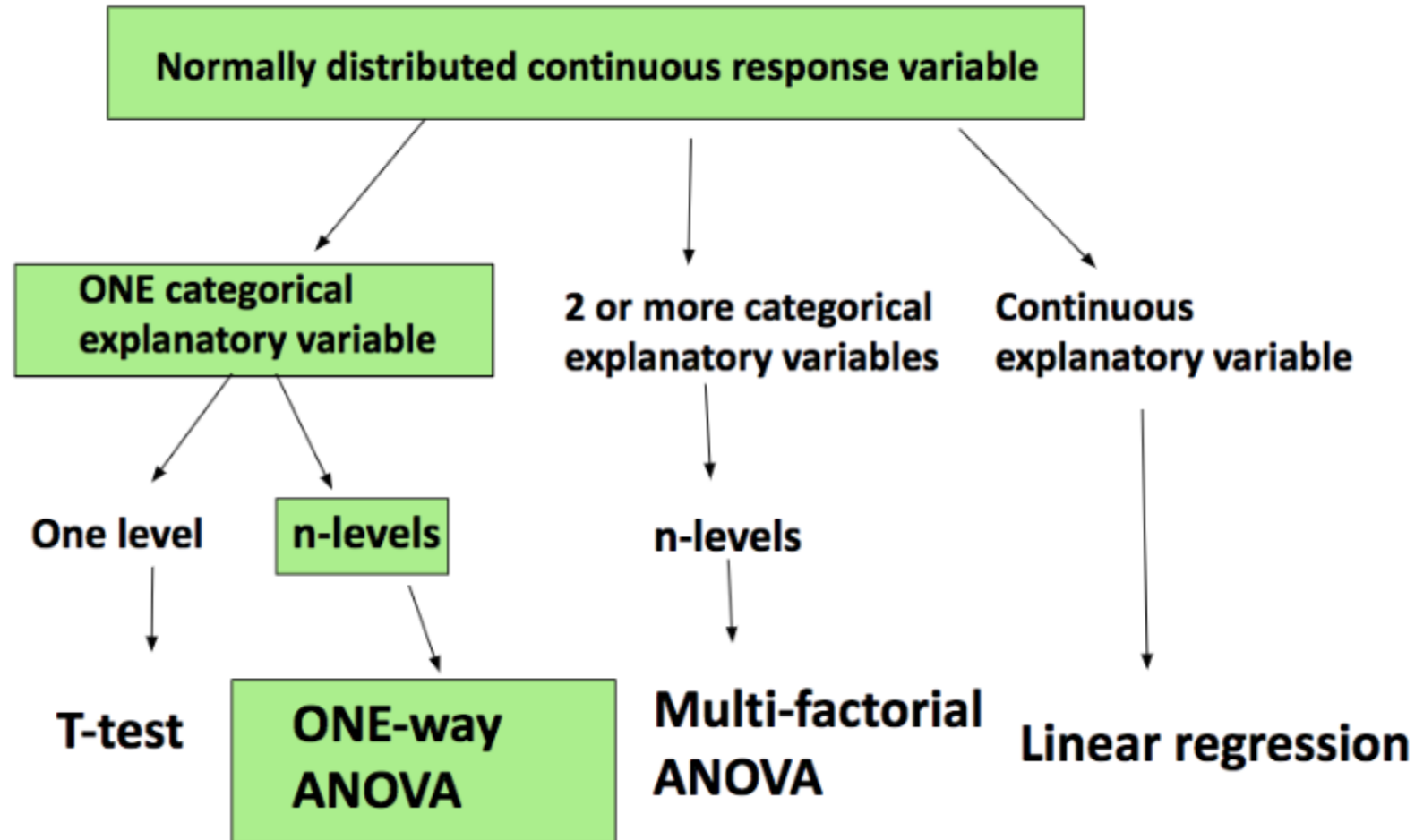


What may the **asterisks** mean?

What may the **letters** mean?

What may the **top right numbers** mean?

Which is the appropriate stats test for my data?



What if...

- Repeated measures
- More factors: two-way ANOVA, three-way ANOVA
- Random factors
- Nested factors
- Interaction
- Covariates (ANCOVA)
- Multidimensional (PERMANOVA)

Linear models and ANOVA, are they the same?

We are trying to predict a variable value from the group it belongs. This is very similar to what we do in regression.

Linear regression and linear model are complete synonyms

We usually talk about an ANOVA when we're quantifying the effect of a **discrete, or categorical** explanatory variable on a **continuous** response variable. It is also a linear model, but instead of getting a slope we get an estimate of the variable value for each category.

Your turn:



Perform and interpret your first ANOVA by this steps (Add them at your Rmd file)

1. Prepare the data
2. Write down your research question
3. Plot the data of interest with the groups you will be comparing
4. Check ANOVA assumptions
5. Perform the ANOVA analysis + posthoc
6. Interpret the results
7. Writte down your conclusions

16 January send me the code that you used in a knitted Rmd file (html)

The code

To build the ANOVA table:

- `res_aov <- aov (quantitativeVariableName ~ FactorName, data = datasetName)`
- `summary(res_aov)`

To check ANOVA assumptions:

- `plot(res_aov)`

To do Tukey HSD post-hoc tests:

- `library(multcomp)`
- `post_test <- glht(res_aov, linfct = mcp(FactorName = "Tukey"))`
- `summary(post_test)`

A bit of help coding:

<https://ourcodingclub.github.io/tutorials/anova/#model>

<https://statsandr.com/blog/anova-in-r/#anova>

Take home messages

1. What is an ANOVA?
2. Why do we use an ANOVA?
3. What type of variables are we using in an ANOVA? What is the difference between a factor and a level?
4. What is the difference between ANOVA and t-test?
5. Can you describe the H_0 and H_1 of a one-way ANOVA?
6. What does the “P-value” represents?
7. What requisites do we have to meet to use ANOVA?
8. Is an ANOVA a linear model?
9. What is a posthoc test? Can you name one?
10. What is a one-way ANOVA?