RenR 480/711

Anova assumptions

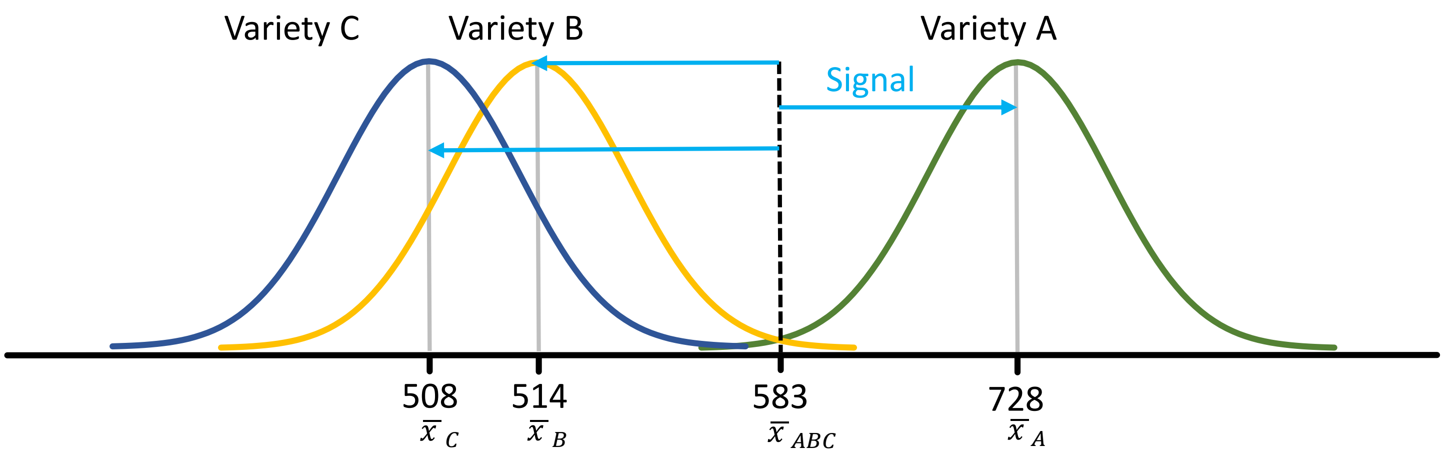
Multi-way Anova

Anova assumptions

1. Normal distribution of experimental errors/residuals
2. Equal variances between treatments
3. Independence of samples

Anova assumptions

1. Normal distribution of experimental errors/residuals

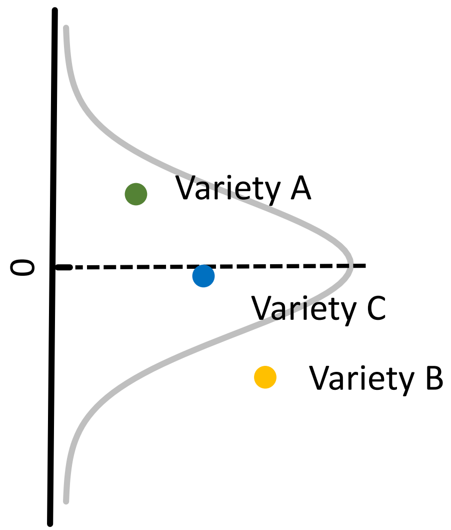


Assessment through residuals of the experiment

Residuals = difference between the experimental mean and treatment means

Anova assumptions

1. Normal distribution of experimental errors



Anova assumptions

1. Normal distribution of experimental errors

Testing for normality: Shapiro-Wilk test

H_0 *Null Hypothesis*: *Distribution of residuals = normal distribution*
H_1 *Alternative Hypothesis*: *Distribution of residuals ≠ normal distribution*

p› 0.05 means that the distribution of residuals is normal

<u>Problems with statistical tests for normality:</u>

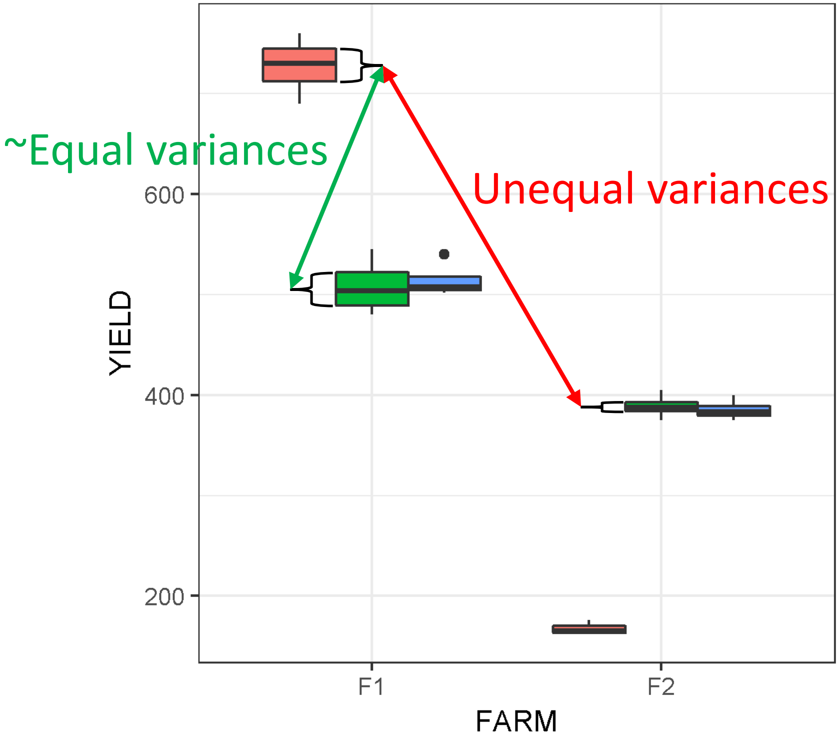
low n: low power → test will fail to detect non-normality

high n: high power → even tiny deviations from normality will be significant

→ Use residual plots to assess this assumption

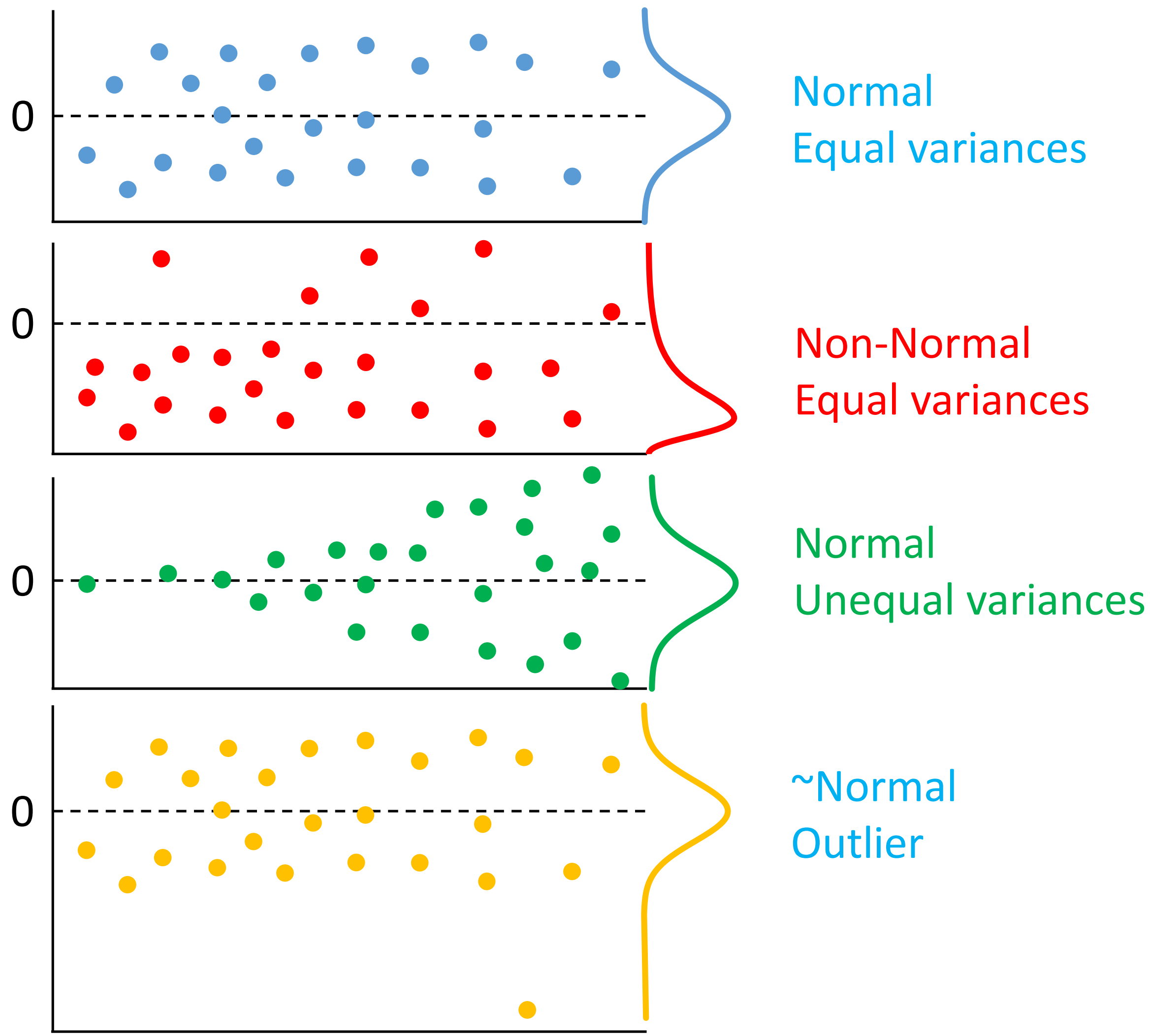
Anova assumptions

2. Equal variances between treatments



Anova assumptions

Residual plots



Transformations

Square root:
$x_T = \sqrt{x + c}$

All values have to be >1;
add a constant if needed

Log transformation:
$x_T = \log(x)$

Increasingly powerful

Inverse transformation
$x_T = \frac{1}{x}$

Anova assumptions

3. Independence of samples

Samples:

→ have been randomly samples ***or***

→ come from a randomized design

This means: rows in your table do not influence each other

→This assumption has to be addressed with the experimental design

<u>Samples may not be independent if:</u>

Multiple measurements on one subject (repeated measures)

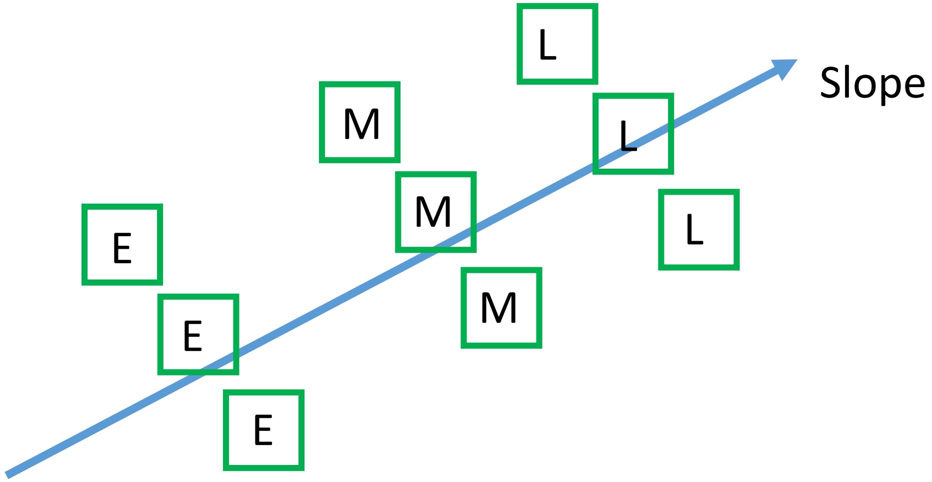
Observations are correlated in space

Observations are correlated in time

Anova assumptions

Spatial correlation

Example: Species richness in forest stands



Successional stage
E= Early
M= Mid
L= Late

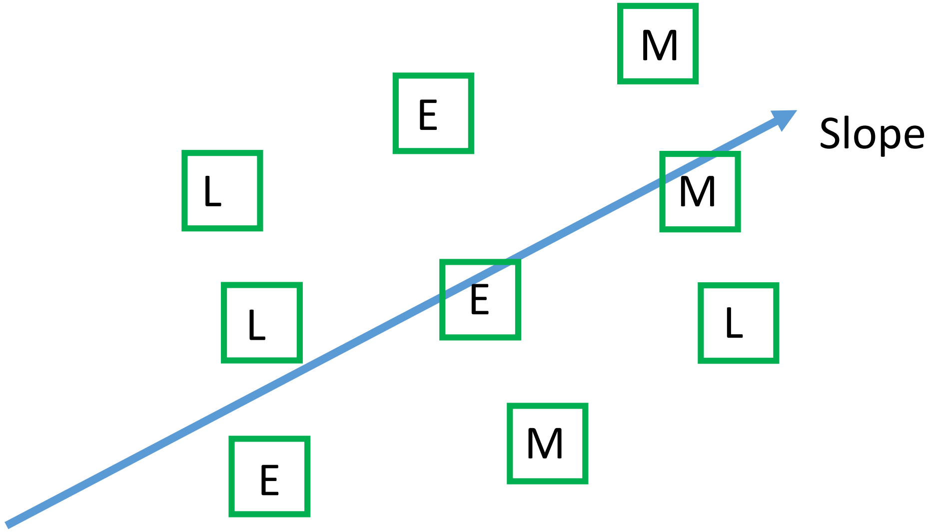
Stage	Richness
E	24
E	20
E	22
M	15
M	14
M	17
L	9
L	12
L	11

- Pseudoreplication because samples are spatially not independent
- Observed differences cannot be attributed to stage because slope is confounding & and there is no way of teasing this apart

Anova assumptions

Spatial correlation

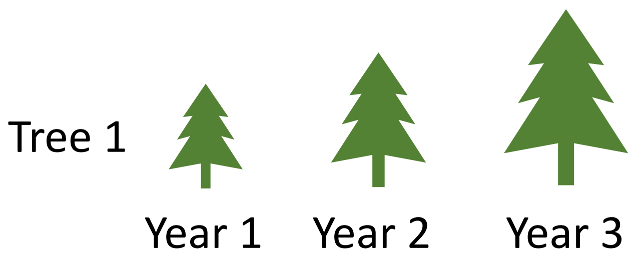
Better:



Creating a randomized design avoids pseudoreplication

Anova assumptions

Temporal correlation



Tree	Year	Height
T1	1	5
T1	2	6
T1	3	7
T2	1	6
T2	2	6
T2	3	7
T3	1	15
T3	2	16
T3	3	17

Rows are not independent – standard ANOVA not applicable
→ repeated measures ANOVA or regression analysis

Multi-way ANOVA

Multi-way-ANOVA

Same as One-way-Anova but with more than one treatment

anova(lm(YIELD~VARIETY*FARM))
anova(lm(YIELD~VARIETY+FARM+FARM:VARIETY))

Source of variation	df	Sum of squares	Mean Squares	F-value	p-value
Variety	2	Analog to One-way Anova			
Farm	1				
Variety:Farm	2				
Residuals	18				

Interaction

$$df_{variety} = n_{levels\ of\ Variety} - 1$$

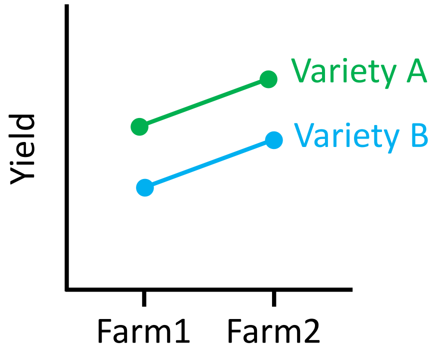
$$df_{Farm} = n_{levels\ of\ Farm} - 1$$

$$df_{Variety:Farm} = (n_{levels\ of\ Variety} - 1)(n_{levels\ of\ Farm} - 1)$$

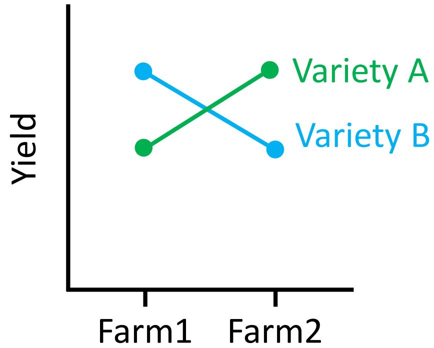
$$df_{Residuals} = n_{observations} - n_{levels\ of\ Variety} \times n_{levels\ of\ Farm}$$

Multi-way-ANOVA

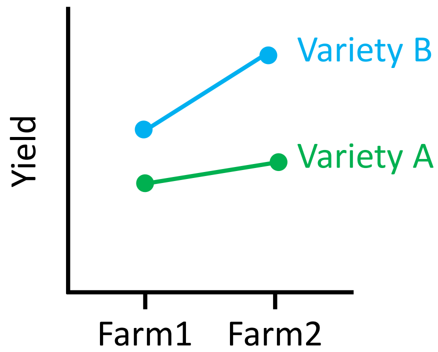
Interaction: Different story under different conditions



Variety A better than B
Farm 2 better than Farm 1
→ no interaction



Main effects n.s.
(Variety and Farm)
However, VarA is better on Farm 2 and VarB is better on Farm 1
→ significant interaction



Main effects significant
(but difficult to interpret with overall means)
→ significant interaction

In case of significant interactions:

→ Disregard n.s. main effects
→ Avoid to interpret significant main effects
→ Show the different stories for different conditions with graphs
→ Follow up with separate analyses for subsets (e.g. Farm1 and Farm2)

Pairwise comparisons

Pairwise comparisons

Experiment-wise Type I-Error

Number of possible comparisons (3 varieties: A,B,C)

A-B

A-C

B-C

$$c = \frac{treatments(treatments - 1)}{2}$$

$$c = \frac{3 \times (3 - 1)}{2} = 3$$

Probability of making a Type-I error in at least one comparison
=1-probability of making no type-I error at all
$=1 - 0.95^c$
$=1 - 0.95^3$
=0.86

Need for adjustment!

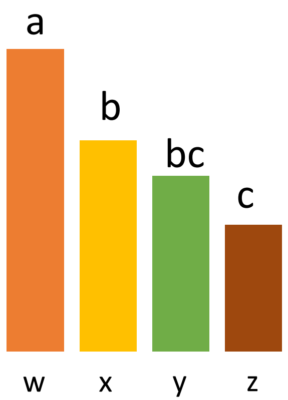
Pairwise comparisons

Bonferroni-Adjustment

$$\frac{\alpha}{c} = \frac{0.05}{3} = 0.017$$

Reporting pairwise comparisons

Graphics



	w	x	y	z
w		*	*	*
x			n.s.	*
y				n.s.
z				

Tables

	Var A	Var B	Var C
Farm1	728 a A	508 b A	514 b A
Farm2	168 a B	389 b B	385 b B