Experimental Methods in Computer Science

(Metodologias Experimentais em Informática)

Henrique Madeira

Master in Informatics Engineering
Departamento de Engenharia Informática
Faculdade de Ciências e Tecnologia da Universidade de Coimbra
2023/2024

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 1

1

Hypothesis Testing

ANalysis Of Variance - ANOVA

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 2

2

Problem: what if you have more than two configurations to compare?

Assume you are the database administrator of a big information system. The database has just been installed and you are trying two tuning three configurations: Conf. **A**, Conf. **B** and Conf. **C**.

You use a given SQL package to test the execution time for each configuration.

After running several times the SQL package in each configuration, you want to take a decision.

Questions:
- Are the configurations the same?
- What is the best configuration?

Conf. A exec. time	Conf. B exec. time	Conf. C exec. time
74	69	65
66	71	68
88	80	64
68	88	70
79	64	72
68	65	64
87	74	62
79	76	75
78	89	71
72	68	69
86	67	
85	72	
86		

You cannot use a two-sample T-test (at least directly)

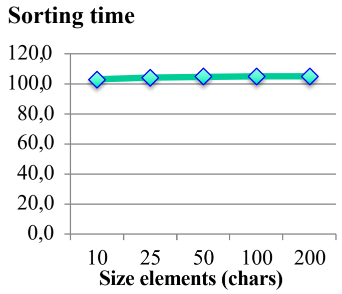
Henrique Madeira, DEI-FCTUC, 2018-2023

Another example: what is the impact of the size of the elements in the sorting time?

How does the size of the elements in an array (each element is a string of chars) affect the time necessary to sort all the elements of the array?

After performing some experiments with Quicksort and an array of 10000 elements, you obtained the following results:

Does the size of elements matter?



Size of elements (no. characters)				
10	**25**	**50**	**100**	**200**
106	106	105	108	104
102	105	101	105	105
99	104	106	108	104
103	102	105	104	103
108	101	106	105	108
102	107	107	107	110
104	105	101	104	104
100	106	107	100	102
	102	105	106	105
	103		102	

Means (milliseconds)

103.0	104.1	104.8	104.9	105.0

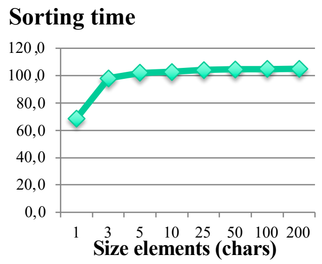
Henrique Madeira, DEI-FCTUC, 2018-2023

Another example: what is the impact of the size of the elements in the sorting time?

How does the size of the elements in an array (each element is a string of chars) affect the time necessary to sort all the elements of the array?

After observing the results, you decided to do **some more experiments**:

Does the size of elements matter?



Size of elements (no. characters)							
1	**3**	**5**	**10**	**25**	**50**	**100**	**200**
71	100	104	106	106	105	108	104
71	95	99	102	105	101	105	105
69	94	103	99	104	106	108	104
69	99	100	103	102	105	104	103
68	99	103	108	101	106	105	108
66	97	100	102	107	107	107	110
71	103	99	104	105	101	104	104
71	98	105	100	106	107	100	102
66	100	103		102	105	106	105
65		105		103		102	

Means (milliseconds)

68.7	98.3	102.1	103.0	104.1	104.8	104.9	105.0

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 5

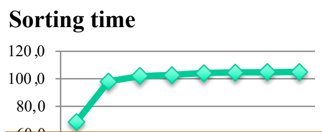
5

Another example: what is the impact of the size of the elements in the sorting time?

How does the size of the elements in an array (each element is a string of chars) affect the time necessary to sort all the elements of the array?

After observing the results, you decided to do **some more experiments**:

Does the size of elements matter?



Size of elements (no. characters)							
1	**3**	**5**	**10**	**25**	**50**	**100**	**200**
71	100	104	106	106	105	108	104
71	95	99	102	105	101	105	105
69	94	103	99	104	106	108	104
69	99	100	103	102	105	104	103
68	99	103	108	101	106	105	108
66	97	100	102	107	107	107	110

But the problem needs more than just looking at the means using an informal approach…

The goal is to test if the difference between multiple sample means is significant.

You need ANOVA to do that!

68.7	98.3	102.1	103.0	104.1	104.8	104.9	105.0

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 6

6

Informal approach

To know if the difference between multiple sample means is significant…

- Use informal graphic approaches (exploratory data analysis)
    - Charts
    - Side-by-side box plots
    - Multiple histograms
- But knowing whether the differences between the groups (factors) are significant depends on…
    - the difference in the means;
    - the standard deviations of each group;
    - the sample sizes.

Another distribution and test like the T or Z

→ **Need ANOVA to determine P-value (from the F statistic)**

Henrique Madeira, DEI-FCTUC, 2018-2023

One-Way ANOVA

- ANalysis Of Variance - **ANOVA**
- The one-way ANOVA is used to test the claim that three or more population means are equal
- This is an extension of the two independent samples tests
- ANOVA tests the following hypotheses:
    - **H_0: $\mu_1 = \mu_2 = \mu_3 \ldots = \mu_k$ (the means of all the groups are equal)**
    - **H_1: Not all the means are equal**
- It does not say how or which ones differ. Need to follow up with multiple comparisons (more tests).

Henrique Madeira, DEI-FCTUC, 2018-2023

One-Way ANOVA

- **Dependent variable**: the variable you are comparing
- **Independent variable**: the factor variable being used to define the samples (groups)
- **Levels**: values of the independent variable selected to be studied. Each level will originate a sample (group)
- Example:
  - Dependent variable: **sorting time**
  - Independent variable: **elements size**
  - Levels: 5 levels
    - **10 chars** • **25 chars** • **50 chars**
    - **100 chars** • **200 chars**

Size of elements (no. characters)				
10	**25**	**50**	**100**	**200**
103	108	108	110	108
108	107	109	108	112
105	106	111	111	108
109	104	108	112	113
108	103	109	108	112
103	109	110	110	114
108	107	107	107	108

Henrique Madeira, DEI-FCTUC, 2018-2023

Assumptions of ANOVA

- Each group is approximately normal
  - Can be checked informally by looking at the histogram of the data or using a normality test
- Can cope with some non-normality, but not with severe outliers
- Standard deviations of each group are approximately equal
  - Rule of thumb: ratio of largest to smallest sample standard deviation must be less than 2:1

Henrique Madeira, DEI-FCTUC, 2018-2023

Rationale for ANOVA

- We have at least 3 means to test (each mean is from a sample). For example, H_0: $\mu_1 = \mu_2 = \mu_3$.
- We could use two sample t-test to test them 2 at a time. But with ANOVA we will test them all 3 (or mo[re]) at once.
- Instead of us[ing a mean difference, ANOVA use the variance of the group] means in relation to [the grand mean over all groups].
- The logic is j[ust the same as for the t-test or z-test: compare the observed] variance among mean[s (observed difference in means in the t-test or z-test) to wha]t we would expec[t to get by chance].

The number of comparisons (tests) increase when using two sample t-tests to test successively pairs of samples at a time. Type I error (false positive) also increase dramatically.

Number of samples	**Number of tests**
2	**1**
3	**3**
4	**6**
5	**10**

11

Rationale for ANOVA

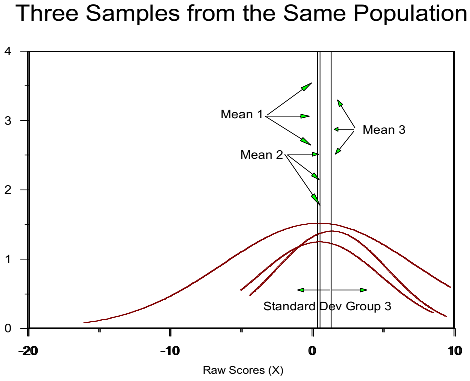
- We have at least 3 means to test (each mean is from a sample). For example, H_0: $\mu_1 = \mu_2 = \mu_3$.
- We could use two sample t-test to test them 2 at a time. But with ANOVA we will test them all 3 (or more) at once.
- Instead of using a mean difference, **ANOVA use the variance of the group means in relation to the grand mean over all groups**.
- The logic is just the same as for the t-test or z-test: compare the observed variance among means (observed difference in means in the t-test or z-test) to what we would expect to get by chance.

12

Rationale for ANOVA (cont.)

Suppose we have 3 samples from the same population. Our results might look like this:

Note that the means from the 3 groups are not exactly the same, but they are close, so the variance among means will be small.

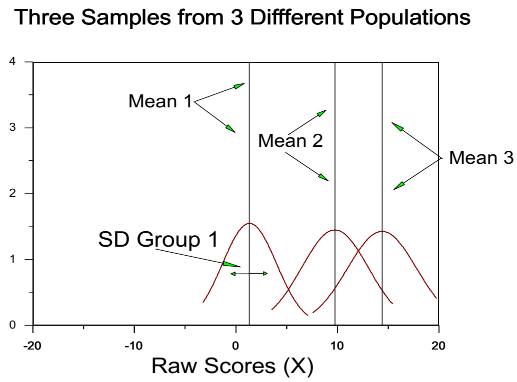


13

Rationale for ANOVA (cont.)

Suppose we sample from 3 different populations. Our results might look like this:

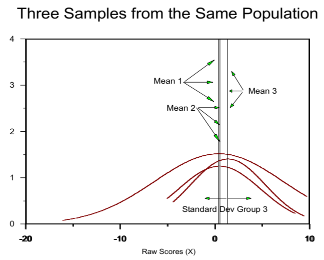
Note that the sample means are far away from one another, so the variance among means will be large.

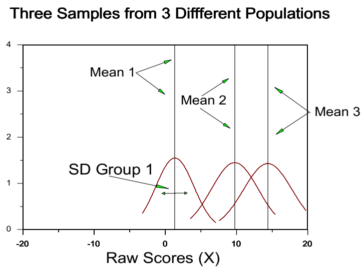


14

Rationale for ANOVA (cont.)

Suppose we do a study and find the following results (either graph). How would we know or decide whether there is a real effect or not?





To decide, we can compare our observed variance in means to what we would expect to get on the basis of chance, if H_0 is true (i.e., no difference in the means).

Henrique Madeira, DEI-FCTUC, 2018-2023

15

Definitions of terms in ANOVA

Break the analysis of the variance into meaningful pieces that correspond to independent variable (IV) effects and errors.

$\bar{X}_G$ The Grand Mean, taken over all observations.

$\bar{X}_A$ The mean of any level of the IV (group).

$\bar{X}_{A_1}$ The mean of a specific level (1 in this case).

X_i The observation or raw data for the i^{th} measurement.

Variation is the sum of the squares of the deviations between a value and the mean of the sample (group)

$$\boxed{SS_{(T)} = \sum (X_i - \bar{X}_G)^2}$$

Sum of Squares (SS) is often followed by a variable in parentheses such as $SS_{(B)}$ or $SS_{(W)}$ that indicates which sum of squares is referred to:

Henrique Madeira, DEI-FCTUC, 2018-2023

16

Sources of variations

- Sums of squares measure three sources of variation:
  - Groups (variation between group means)
  - Error (variation within groups)
  - Total (SST = SSG + SSE)

- Degrees of freedom (n observations, k sample means)
  - df (total) = $n - 1$
  - df (within) = $n - k$
  - df (between) = $k - 1$

  df (total) = df (between) + df (within)

Size of elements (no. characters)				
10	**25**	**50**	**100**	**200**
103	108	108	110	108
108	107	109	108	112
105	106	111	111	108
109	104	108	112	113
108	103	109	108	112
103	109	110	110	114

Henrique Madeira, DEI-FCTUC, 2018-2023

17

ANOVA calculations

$$GM = \frac{\sum n_i \bar{x}_i}{\sum n_i}$$

The grand mean is the weighted average of the individual sample means

$$SS_{(T)} = \sum (X_i - \bar{X}_G)^2$$

The total sum of squares comes from the distance of all the scores to the grand mean. This is the big total.

$$SS_{(W)} = \sum df\, s^2$$

The within-group sum of squares comes from the distance of the scores to the sample means. The degrees of freedom are equal to the sum of the individual *df* for each sample. This indicates error.

$$SS_{(B)} = \sum n_A (\bar{X}_A - \bar{X}_G)^2$$

The between-groups sum of squares represents the distance of the sample means from the grand mean. This indicates IV effects.

Henrique Madeira, DEI-FCTUC, 2018-2023

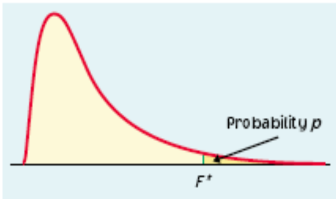
18

ANOVA: the **F statistics**

- The test statistic for ANOVA is called from the **F statistic**, which we obtain from the F Test.
- The **F statistic determines if the variation between sample means is significant**:

$$\frac{\text{Variation Among Sample Means}}{\text{Variation Among Individuals In Each Sample}}$$

$$F = \frac{SS_{(B)}/(k-1)}{SS_{(w)}/(n-k)}$$



Henrique Madeira, DEI-FCTUC, 2018-2023

19

Obtaining the critical from F table

How to obtain the critical value from F tables for a given α? For example, for α = 0.05?

$$F = \frac{SS_{(B)}/(k-1)}{SS_{(w)}/(n-k)}$$

***df* for the numerator**

Example: For 4 levels (i.e., 4 values of the IV), *df* = k - 1 = 4 - 1 = 3

***df* for the denominator**

Example: For 13 observations *df* = n - k = 13 - 4 = 9

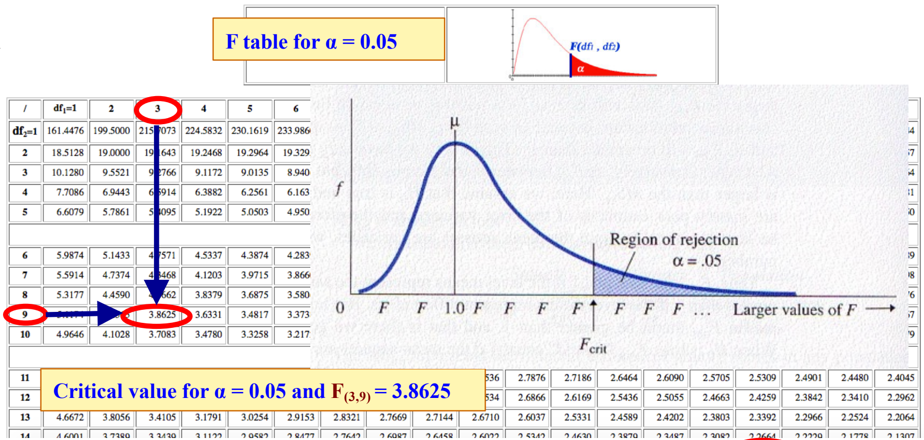
In the F table for α = 0.05, look for the critical value for $F_{(3,9)}$

Henrique Madeira, DEI-FCTUC, 2018-2023

20

Obtaining the critical from F table

How ?



In the F table for α = 0.05, look for the critical value for $F_{(3,9)}$

Online F critical value calculator: https://www.danielsoper.com/statcalc/calculator.aspx?id=4

Henrique Madeira, DEI-FCTUC, 2018-2023

The ANOVA table

- The ANOVA table is a summary of all the elements needed for the calculation of the P-value

	SS	df	MS	F	P
Between	SS(B)	k-1	$\frac{SS(B)}{k-1}$	$\frac{MS(B)}{MS(W)}$	Tail area above F
Within	SS(W)	n-k	$\frac{SS(W)}{n-k}$		
Total	SS(W) + SS(B)	n-1			

Henrique Madeira, DEI-FCTUC, 2018-2023

The ANOVA table

- The ANOVA table is a summary of all the elements needed for the calculation of the P-value

$$SS_{(B)} = \sum n_A(\bar{X}_A - \bar{X}_G)^2$$

	SS		MS		P
Between	SS(B)	k-1	$\frac{SS(B)}{k-1}$	$\frac{MS(B)}{MS(W)}$	Tail area above F
Within	SS(W)	n-k	$\frac{SS(W)}{n-k}$		
Total	SS(W) + SS(B)				

$$SS_{(W)} = \sum df\, s^2$$

$$GM = \frac{\sum n_i \bar{x}_i}{\sum n_i}$$

Henrique Madeira, DEI-FCTUC, 2018-2023

23

One-way ANOVA steps

Use the same general steps of hypothesis testing

1. State the hypothesis
2. Calculations (to compute the test statistic)
3. Obtain p value
4. Make a decision

$$F = \frac{SS_{(B)}/(k-1)}{SS_{(w)}/(n-k)}$$

Henrique Madeira, DEI-FCTUC, 2018-2023

24

Example one-way ANOVA

How does the size of the elements in an array affect the time necessary to sort all the elements of the array?

For economy of space, we will consider only 3 level of the independent variable size of elements: 3, 5 and 10 characters.

The table on the right shows the sorting times in milliseconds obtained with the Quicksort and an array of 10000 elements.

Size of elements (no. characters)		
3	**5**	**10**
100	104	106
95	99	102
94	103	99
99	100	103
99	103	108
97	100	102
103	99	104
98	105	100
100	103	
	105	

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: one-way ANOVA
Step 1- State the hypothesis

Hypothesis

- **H0: $\mu 3 = \mu 5 = \mu 10$** → All samples have equal means
- **H1: $\mu 3 \neq \mu 5 \neq \mu 10$** → Not all the means are equal
  - Does not say how or which ones differ
  - Can follow up with multiple comparisons

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: one-way ANOVA
Step 2 - Calculations

Determine characteristics of the samples under comparison.

	3 chars	5 chars	10 chars
Sample size	9	10	8
Mean	98.3	102.1	103.0
Std. Dev.	2.74	2.38	2.98
Variance	7.50	5.66	8.86

$n = 27$
$k = 3$

Henrique Madeira, DEI-FCTUC, 2018-2023

27

Example: one-way ANOVA
Step 2 – Calculations (cont.)

Grand Mean calculation

– Weighted average of the individual sample means

$$GM = \frac{\sum n_i \bar{x}_i}{\sum n_i}$$

	3 chars	5 chars	10 chars
Sample size	9	10	8
Mean	98.3	102.1	103.0
Std. Dev.	2.74	2.38	2.98
Variance	7.50	5.66	8.86

$$GM = \frac{9(98.3) + 10(102.1) + 8(103.0)}{9 + 10 + 8} = 101.10$$

Henrique Madeira, DEI-FCTUC, 2018-2023

28

Example: one-way ANOVA
Step 2 – Calculations (cont.)

Between Group Variation

- Variation between each sample mean and the grand mean
- Each group variation is weighted by the sample size

$$SS_{(B)} = \sum n_A (\bar{X}_A - \bar{X}_G)^2$$

$$SS(B) = 9(98.3 - 101.10)^2 + 10(102.1 - 101.10)^2 + 8(103.0 - 101.10)^2 = 107.77$$

$$GM = 101.10$$

	3 chars	5 chars	10 chars
Sample size	9	10	8
Mean	98.3	102.1	103.0
Std. Dev.	2.74	2.38	2.98
Variance	7.50	5.66	8.86

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: one-way ANOVA
Step 2 – Calculations (cont.)

Within Group Variation

- Is the weighted total of the individual variations
- The weighting is done with the degrees of freedom. For each sample df is one less than the sample size

$$SS_{(W)} = \sum df\, s^2$$

	3 chars	5 chars	10 chars
Sample size	9	10	8
Mean	98.3	102.1	103.0
Std. Dev.	2.74	2.38	2.98
Variance	7.50	5.66	8.86

$$SS(W) = 8(2.74)^2 + 9(2.38)^2 + 7(2.98)^2 = 172.90$$

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: one-way ANOVA
Step 2 – Calculations (cont.)

We can now fill the one-way ANOVA table

	SS	df	MS	F	P
Between	SS(B)	k-1	$\frac{SS(B)}{k-1}$	$\frac{MS(B)}{MS(W)}$	Tail area above F
Within	SS(W)	n-k	$\frac{SS(W)}{n-k}$		
Total	SS(W) + SS(B)	n-1			

Example: one-way ANOVA
Step 2 – Calculations (cont.)

We can now fill the one-way ANOVA table

	SS	df	MS	F	P
Between	107.77	2	53.88	7.48	≈ 0.003
Within	172.90	24	7.20		
Total	280.67	26			

Example: one-way ANOVA

Step 3 – Obtain P-value

We can now fill the one-way ANOVA table

	SS	df	MS	F	P
Between	107.77	2	53.88	7.48	≈ 0.003
Within	172.90	24	7.20		
Total	280.67	26			

The P-Value for $F_{(2,24)} = 7.48$ is 0.002986 (using an online calculator)

Example: one-way ANOVA

Step 4 – Make a decision

	SS	df	MS	F	P
Between	107.77	2	53.88	7.48	≈ 0.003
Within	172.90	24	7.20		
Total	280.67	26			

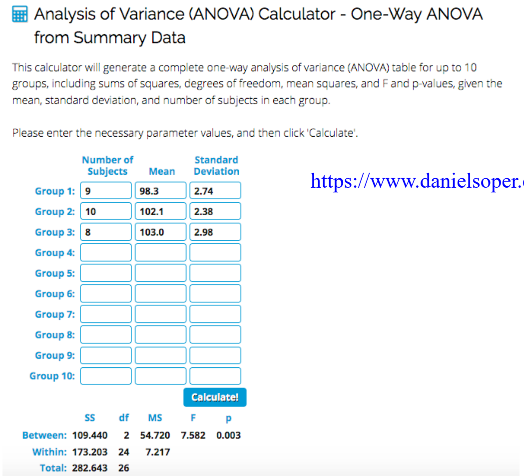
ANOVA

Copy Settings

	Sum of Squares	df	Mean Squares	F	p
Factor	107.77	2	53.88	7.48	.003
Residual	172.9	24	7.2		
Total	280.67	26			

qual

Using an online one-way ANOVA calculator



https://www.danielsoper.com/statcalc/calculator.aspx?id=43

Henrique Madeira, DEI-FCTUC, 2018-2023

Two-Way ANOVA

- Tests the equality of two or more population means when two independent variables are used: factor A and factor B (more than two factors: multi-way ANOVA).
- Each independent variables (factors) may have any number of levels.
- Same results as separate one-way ANOVA on each variable. **But interaction can be tested**.
- Saves time and effort, compared to consecutive one-way ANOVA tests.

Henrique Madeira, DEI-FCTUC, 2018-2023

Two-way ANOVA assumptions

- Normality
  Populations are normally distributed
- Homogeneity of variance
  Populations have similar variances
- Independence of errors
  Independent random samples

Two-way ANOVA: Null Hypotheses

Tests 3 hypothesis simultaneously:

- No difference in means due to factor A
  - H_0: $\mu_{1.} = \mu_{2.} = \ldots = \mu_{a.}$
- No difference in means due to factor B
  - H_0: $\mu_{.1} = \mu_{.2} = \ldots = \mu_{.b}$
- No interaction of factors A and B
  - H_0: $AB_{ij} = 0$

Two-way ANOVA vs. one-way ANOVA

- The test follows the same steps in both cases.

1. **State the hypothesis**
2. **Calculations (to compute the test statistic)**
3. **Obtain p value**
4. **Make a decision**

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 39

39

Two-way ANOVA vs. one-way ANOVA

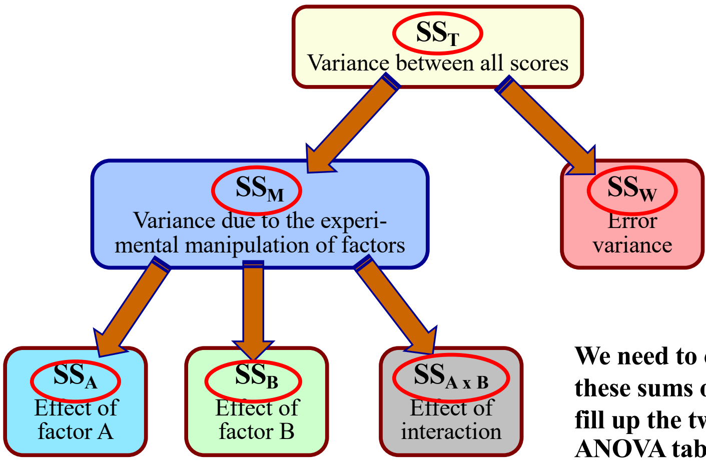
- The test follows the same steps in both cases.
- The two-way ANOVA table is similar (but has more lines than the one-way ANOVA table as there are more sums of squares and more degrees of freedom to calculate).
- The formulas for the calculations in two-way ANOVA are similar to the ones used in one-way ANOVA.
- The F metric is the same.
- The interpretation of the results (two-way ANOVA table filled up with the calculation is more complex, but still quite straightforward).

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 40

40

Two-way ANOVA sums of squares (SS)



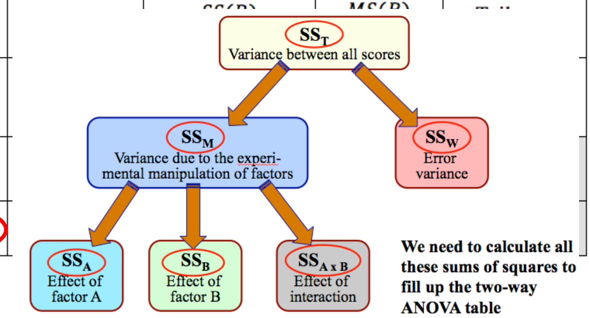
We need to calculate all these sums of squares to fill up the two-way ANOVA table

Two-way ANOVA table

	SS	df	MS	F	P
Factor A	$SS_{(A)}$	a-1	$\frac{SS(A)}{a-1}$	$\frac{MS(A)}{MS(W)}$	Tail area above F
Factor B	$SS_{(B)}$	b-1	$\frac{SS(B)}{b-1}$	$\frac{MS(B)}{MS(W)}$	Tail area above F
Interaction A x B	$SS_{(AxB)}$	(a-1)(b-1)	$\frac{SS(A)}{(a-1)(b-1)}$	$\frac{MS(AxB)}{MS(W)}$	Tail area above F
Within (error)	$SS_{(W)}$	n-a-b	$\frac{SS(W)}{n-a-b}$		
Total	$SS_{(M)} + SS_{(W)}$	n-1			

Two-way ANOVA table

	SS	df	MS	F	P
Factor A	$SS_{(A)}$	a-1	$\frac{SS(A)}{a-1}$	$\frac{MS(A)}{MS(W)}$	Tail area above F
Factor B	$SS_{(B)}$				
Interaction A x B	$SS_{(AxB)}$				
Within (error)	$SS_{(W)}$				
Total	$SS_{(M)} + SS_{(W)}$				



We need to calculate all these sums of squares to fill up the two-way ANOVA table

Henrique Madeira, DEI-FCTUC, 2018-2023

Example two-way ANOVA

How does the **number of elements in an array** and **the size of each element** affect the sorting time of the array?

For economy of space, we will consider only 3 level of the independent variable **size of element** (10, 50 and 100 characters) and 2 levels of the independent variable **number of elements** (10000 and 50000).

The table on the right shows the sorting times in milli-seconds obtained with a Quicksort implementation.

10 chars		50 chars		100 chars	
10K	50K	10K	50K	10K	50K
106	313	105	308	108	312
102	307	101	309	105	307
103	308	106	307	108	307
103	309	105	307	104	311
108	306	106	311	105	312
102	308	107	310	107	308
104	310	101	311	104	308
105	312	107	311	100	309
108		105		106	312
				102	

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: two-way ANOVA
Step 1- State the hypothesis

Factor A – Size (in chars) of the elements to be sorted
Factor B – Number of elements to sort

Hypothesis

- **H_0:** $\mu_{a.10} = \mu_{a.50} = \mu_{a.100}$ → The size of elements to be sorted is not relevant for the sorting time (means are equal)
- **H_0:** $\mu_{b.10k} = \mu_{b.100K}$ → The number of elements to be sorted is not relevant for the sorting time (means are equal)
- **H_0:** $AB_{ij} = 0$ → There are no interaction between the size and the number of elements
- H_1: → Not all the means are equal and there is interaction (only right tail is possible).

Henrique Madeira, DEI-FCTUC, 2018-2023

45

Example: two-way ANOVA
Step 2 - Calculations

Determine characteristics of the samples under comparison.

	10 chars		50 chars		100 chars	
	10K	50K	10K	50K	10K	50K
Sample size	9	8	9	8	10	9
Mean	104.1	308.8	104.8	309.3	104.9	309.6
Std. Dev.	1.96	2.12	2.28	1.75	2.56	2.19
Variance	3.86	4.50	5.19	3.07	6.54	4.78

$$GM = \frac{\sum n_i \bar{x}_i}{\sum n_i}$$ Weighted average of the individual sample means

$$GM = \frac{9(104.1) + 8(308.8) + 9(104.8) + 8(309.3) + 10(104.9) + 9(309.6)}{9 + 8 + 9 + 8 + 10 + 9} = 201.11$$

Henrique Madeira, DEI-FCTUC, 2018-2023

46

Example: two-way ANOVA

Step 2 – Calculations (cont.)

$SS_{(T)}$ - The total sum of squares comes from the distance of all the scores from the grand mean. This is the big total

$$SS_{(T)} = \sum (X_i - \bar{X}_G)^2$$

$\bar{X}_G$ is the GM (Grand Mean), taken over all observations/scores.

The observation or raw data for the i[th] measurement/score

Performing the calculation: $SS_{(T)} = 552982$

(actually, this value is not needed in the calculations)

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: two-way ANOVA

Step 2 – Calculations (cont.)

$SS_{(M)}$ – The sum of squares that gives the variance due to the experimental manipulation of factors (all factors are considered here)

$$SS_{(M)} = \sum n_i (\bar{X}_i - \bar{X}_G)^2$$

$\bar{X}_G$ is the GM (Grand Mean), taken over all observations/scores.

Number of samples of level ***i***

Average of the scores of level ***i***

Performing the calculation:

$$SS_{(M)} = 9(104.1 - 201.25)^2 + 8(308.8 - 201.25)^2 + 9(104.8 - 201.25)^2 + 8(309.3 - 201.25)^2 + 10(104.9 - 201.25)^2 + 9(309.8 - 201.25)^2 = \mathbf{552736.78}$$

Henrique Madeira, DEI-FCTUC, 2018-2023

Example: two-way ANOVA
Step 2 – Calculations (cont.)

SS$_{(A)}$ – The sum of squares for factor A (effect of factor A)

$$SS_{(A)} = \sum n_i(\bar{X}_i - \bar{X}_G)^2$$

First, organize the data according factor A observations

10 chars	
10K	50K
106	313
102	307
103	308
103	309
108	306
102	308
104	310
105	312
108	

50 chars	
10K	50K
105	308
101	309
106	307
105	307
106	311
107	310
101	311
107	311
105	

100 chars	
10K	50K
108	312
105	307
108	307
104	311
105	312
107	308
104	308
100	309
106	312
102	

Example: two-way ANOVA
Step 2 – Calculations (cont.)

SS$_{(A)}$ – The sum of squares for factor A (effect of factor A, size of each element to be sorted)

$$SS_{(A)} = \sum n_i(\bar{X}_i - \bar{X}_G)^2$$

$$SS_{(A)} = 17(200.82 - 201.25)^2 + 17(201.00 - 201.25)^2 + 19(201.84 - 201.25)^2 = \mathbf{10.81}$$

Example: two-way ANOVA
Step 2 – Calculations (cont.)

SS(B) – The sum of squares for factor B (effect of factor B)

$$SS_{(B)} = \sum n_i(\bar{X}_i - \bar{X}_G)^2$$

Organize the data according factor B observations

10K		
10 chars	50 chars	100 chars
106	105	108
102	101	105
103	106	108
103	105	104
108	106	105
102	107	107
104	101	104
105	107	100
108	105	106
		102

50K		
10 chars	50 chars	100 chars
313	308	312
307	309	307
308	307	307
309	307	311
306	311	312
308	310	308
310	311	308
312	311	309
		312

Compute averages

Henrique Madeira, DEI-FCTUC, 2018-2023

51

Example: two-way ANOVA
Step 2 – Calculations (cont.)

SS(B) – The sum of squares for factor B (effect of factor B, number of elements to be sorted)

$$SS_{(B)} = \sum n_i(\bar{X}_i - \bar{X}_G)^2$$

$$SS_{(B)} = 28(104.75 - 201.25)^2 + 25(309.32 - 201.25)^2 = 552721.12$$

Henrique Madeira, DEI-FCTUC, 2018-2023

52

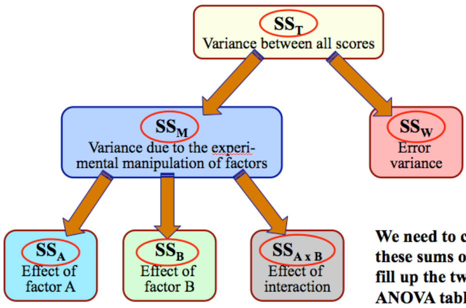
Example: two-way ANOVA
Step 2 – Calculations (cont.)

$SS_{(AxB)}$ – The sum of squares for the interaction between factor A and factor B

$$SS_{(A\times B)} = SS_M - SS_A - SS_B$$

Performing the calculatic

$SS_{(AxB)} =$ 552736.78 -



We need to calculate all these sums of squares to fill up the two-way ANOVA table

Henrique Madeira, DEI-FCTUC, 2018-2023

53

Example: two-way ANOVA
Step 2 – Calculations (cont.)

Within Group Variation

- Is the weighted total of the individual variations (error)
- The weighting is done with the degrees of freedom. For each sample df is one less than the sample size

$$SS_{(W)} = \sum df\, s^2$$

$$SS_{(W)} = 1.96^2\,(9-1) + 2.12^2\,(8-1) + 2.28^2\,(9-1) + 1.75^2\,(8-1) + 2.56^2\,(10-1) + 2.19^2\,(9-1) = 245.28$$

Henrique Madeira, DEI-FCTUC, 2018-2023

54

Example: two-way ANOVA
Step 2 – Calculations (cont.)

Summary table:

	SS	df	MS	F	P
Factor A (size elem.)	SS(A)	a-1	$\frac{SS(A)}{a-1}$	$\frac{MS(A)}{MS(W)}$	Tail area above F
Factor B (no. elem.)	SS(B)	b-1	$\frac{SS(B)}{b-1}$	$\frac{MS(B)}{MS(W)}$	Tail area above F
Interaction A x B	SS(AxB)	(a-1)(b-1)	$\frac{SS(A)}{(a-1)(b-1)}$	$\frac{MS(AxB)}{MS(W)}$	Tail area above F
Within (error)	SS(W)	n-a-b	$\frac{SS(W)}{n-a-b}$		
Total	SS(M) + SS(W)	n-1			

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 55

55

Example: two-way ANOVA
Step 2 – Calculations (cont.)

Summary table:

	SS	df	MS	F	P
Factor A (size elem.)	10.81	2	5.41	1.06	Tail area above F
Factor B (no. elem.)	552721.12	1	552721.12	108166.81	Tail area above F
Interaction A x B	4.84	2	2.42	0.47	Tail area above F
Within (error)	245.28	48	5.11		
Total	552982.05	52			

Henrique Madeira, DEI-FCTUC, 2018-2023

Experimental Methods in Computer Science, Master in Informatics Engineering , DEI-FCTUC, 2023/2024 56

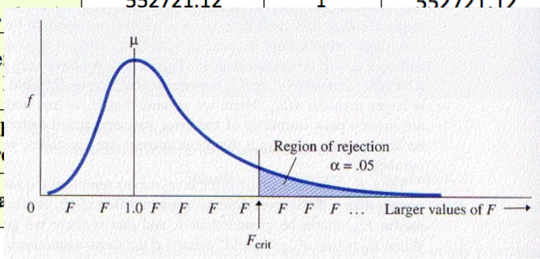
56

Example: two-way ANOVA

Step 2 – Calculations (cont.)

Summary table:

	SS	df	MS	F	P
Factor A (size elem.)	10.81	2	5.41	1.06 $F_{(2,48)}$	Tail area above F
Factor B (no.	552721.12	1	552721.12	108166.81	Tail area above F
Inte A x				0.47 $F_{(2,48)}$	Tail area above F
Witl (err					
Tota					



Henrique Madeira, DEI-FCTUC, 2018-2023

57

Example: two-way ANOVA

Step 3 – Obtain P value

Summary table:

	SS	df	MS	F	P
Factor A (size elem.)	10.81	2	5.41	1.06 $F_{(2,48)}$	0.3589
Factor B (no. elem.)	552721.12	1	552721.12	108166.81	≈ 0
Interaction A x B	4.84	2	2.42	0.47 $F_{(2,48)}$	0.6252
Within (error)	245.28	48	5.11		
Total	552982.05	52			

Henrique Madeira, DEI-FCTUC, 2018-2023

58

Example: two-way ANOVA

Step 4 – Make a decision

Summary table:

Hypothesis

- **H_0:** $\mu_{b.10} = \mu_{b.50} = \mu_{b.100}$ → The size of elements to be sorted is not relevant for the sorting time (means are equal)
- **H_0:** $\mu_{10k.a} = \mu_{50K.}$ → The number of elements to be sorted is not relevant for the sorting time (means are equal)
- **H_0:** $AB_{ij} = 0$ → There are no interaction between the size and the number of elements
- **H_1:** → Not all the means are equal and there is interaction (only right tail is possible).

Fa... (si...					
Fa... (no...					
Int... A x...					
Wi... (error)					
Total	552982.05	52			

59

Example: two-way ANOVA

Step 4 – Make a decision

Summary table:

	SS			F	P
Factor A (size elem.)		2			0.3589
Factor B (no. elem.)					≈ 0
Interaction A x B	4.84	2	2.42	0.47	0.6252
Within (error)					
Total					

P value = 36% for the **factor A**, size of elements to be sorted → H0 is retained for this factor: the size of the elements is not relevant for the sorting time

P value ≈ 0 for the **factor B**, number of elements to be sorted → H0 is rejected: the number of elements to be sorted is significant and determines the sorting time.

P value = 63% for the interactions between **factor A and B** → H0 is retained: there is no interactions between the size of the elements and the number of elements

H_0 as whole is rejected

60

Examples of online calculators and etc.

- **Example of F critical value calculator:**
  http://www.danielsoper.com/statcalc3/calc.aspx?id=4
- **F critical tables:**
  http://www.socr.ucla.edu/applets.dir/f_table.html
- **ANOVA calculator:**
  http://www.danielsoper.com/statcalc3/calc.aspx?id=43
- **Two-way ANOVA calculator:**
  http://scistatcalc.blogspot.pt/2013/11/two-factor-anova-test-calculator.html#
- **P calculator for several distributions:**
  http://vassarstats.net/tabs.html
- **Videos two-way ANOVA**
  https://www.youtube.com/watch?v=cNIIn9bConY&list=PLWtoq-EhUJe2TjJYfZUQtuq7a0dQCnOWp