

# Εφαρμογή στατιστικής- χημειομετρίας με χρήση του Excel-1<sup>ο</sup> Μέρος

Μάριος Κωστάκης, PhD, MSc

1

## Εισαγωγικά

- Το Microsoft Office αποτελεί την πιο διαδεδομένη σουιτα εφαρμογών γραφείου.
- Μέρος της αποτελεί το Excel που επιτρέπει με τη χρήση υπολογιστικών φύλλων να γίνονται χειρισμοί δεδομένων.
- Πέρα από τον οικονομικό τομέα, μέσω του Excel μπορούν να γίνουν πράξεις στατιστικής που απαιτούνται στη Χημεία και ιδιαίτερα στον τομέα της χημικής ανάλυσης, γνωστές και ως χημειομετρία.
- Πέρα από το βασικό πακέτο του Excel, μέσα στην εφαρμογή υπάρχει πακέτο σχετικό με την Ανάλυση Δεδομένων (Data Analysis Tool) που παρέχεται σε όλους του συνδρομητές, ανεξαρτήτως συνδρομής.

2

# Περιγραφική Στατιστική

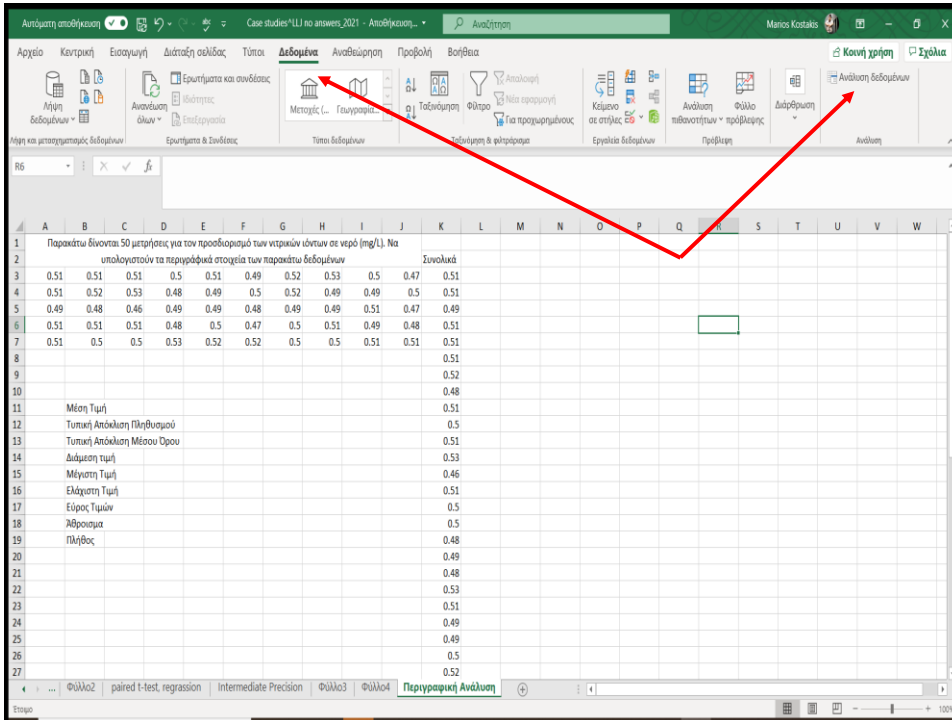
3

## Βασικές Πράξεις – Συναρτήσεις

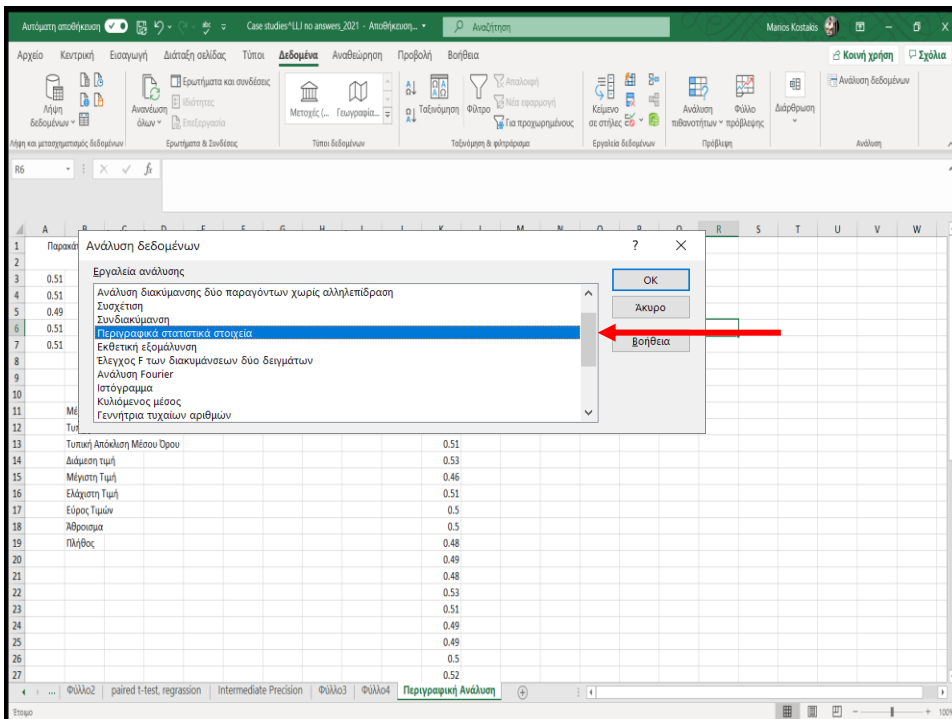
Παράμετρος	Συνάρτηση
Μέσος Όρος	average()
Τυπική απόκλιση πληθυσμού	stdev.P()
Τυπική απόκλιση μέσου όρου	stdev.S() ή stdev()
Διάμεση τιμή (median)	median()
Εύρος Τιμών	range()
Μέγιστη τιμή	max()
Ελάχιστη Τιμή	min()
Άθροισμα Τιμών	sum()
Άθροισμα Τετραγώνων	sumsq()
Ρίζα τιμής	sqrt()

Και συνδυασμός αυτών των συναρτήσεων

4



5



6

Περιγραφικά στατιστικά στοιχεία

Είσοδος  
Περιοχή εισόδου: A1:E7

Ομαδοποίηση κατά:  
 Στήλες  
 Γραμμές

Ετικέτες στην πρώτη γραμμή

Επιλογές εξόδου  
 Περιοχή εξόδου:  
 Νέο φύλλο  
 Νέο βιβλίο εργασίας

Περαιτέρω στατιστικά  
 Βαθμιάς εμπιστοσύνης για το μέσο: 95 %  
 Κ μεγαλύτερο: 1  
 Κ μικρότερο: 1

	A	B	C	D	E
1	Παρακάτω δίνονται 50 μετρήσεις για τον προοιπολογιστούν τα περιγράφοι α				
2					
3	0.51	0.51	0.51	0.5	0.51
4	0.51	0.52	0.53	0.48	0.48
5	0.49	0.48	0.46	0.49	0.48
6	0.51	0.51	0.51	0.48	0.5
7	0.51	0.5	0.5	0.53	0.51
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					0.48
20					0.49
21					0.48
22					0.53
23					0.51
24					0.49
25					0.49
26					0.5
27					0.52

7

Περιγραφική Ανάλυση

- Μέσος
- Τυπικό σφάλμα
- Διάμεσος
- Επικρατούσα τιμή
- Μέση απόκλιση τετραγώνου
- Διακύμανση
- Κέρτυση
- Ασυμμετρία
- Είρος
- Ελάχιστο
- Μέγιστο
- Άθροισμα
- Πλήθος
- Μεγαλύτερο(1)
- Μικρότερο(1)
- Βαθμιάς εμπιστοσύνης(95.0%)

8

The screenshot shows the Microsoft Excel interface with the following content in column A:

Στήλη1
Μέσος
Τυπικό σφάλμα
Διάμεσος
Επικρατούσα τιμή
Μέση απόκλιση τετραγώνου
Διακύμανση
Κύριση
Ασυμμετρία
Εύρος
Ελάχιστο
Μέγιστο
Άθροισμα
Πλήθος
Μεγαλύτεροι[1]
Μικρότεροι[1]
Βαθμύς εμπιστοσύνης(95.0%)

The text **Στη πράξη** is written in large red font across the center of the spreadsheet.

9

The slide contains the following text:

## ΣΤΑΤΙΣΤΙΚΕΣ ΔΟΚΙΜΑΣΙΕΣ ΣΗΜΑΝΤΙΚΟΤΗΤΑΣ

10

## Στατιστικές Δοκιμασίες

- Συχνά ερχόμαστε αντιμέτωποι με τα εξής ερωτήματα:
  - Είναι τα αποτελέσματα εντός ή εκτός ορίων;
  - Η πειραματική τιμή που δίνει CRM είναι ίδια με την τιμή στο πιστοποιητικό;
  - Ποια από τις δύο μεθόδους δίνει καλύτερα αποτελέσματα;
  - Οι δύο τεχνικές δίνουν ισοδύναμα αποτελέσματα;
- Η απάντηση στα παραπάνω ερωτήματα δεν είναι απόλυτη «ναι» ή «όχι».
- Βασίζεται στη στατιστική.
- Η απάντηση που αποδίδεται από τη στατιστική είναι:
  - «εμφανίζει στατιστικά σημαντική διαφορά» ή
  - «δεν εμφανίζει στατιστική σημαντική διαφορά»

11

## Μηδενική Υπόθεση (1)

- Στις στατιστικές δοκιμασίες ελέγχεται η πιθανότητα ισχύος κάποια υπόθεσης.
- **Μηδενική Υπόθεση** (null Hypothesis,  $H_0$ )
  - Υποδηλώνει ότι δεν υπάρχει διαφορά
- **Εναλλακτική Υπόθεση** (alternative Hypothesis,  $H_a$ )
  - Υποδηλώνει την ύπαρξη διαφορά, είναι συμπληρωματική της μηδενικής.
- Η απάντηση βασίζεται στην πιθανότητα το αποτέλεσμα να ανήκει στη μία υπόθεση ή στην άλλη.
- Γι' αυτό το αποτέλεσμα των δοκιμασιών είναι «στατιστικά σημαντικό για στάθμη εμπιστοσύνης» **90% ή 95% ή 99%**.

12

## Μηδενική Υπόθεση (2)

- Πως κρίνεται ότι η ισχύει η μηδενική υπόθεση ή η εναλλακτική;
  - Σφάλμα 1<sup>ου</sup> είδους (Type I ή error  $\alpha$ )
    - Απόρριψη της μηδενική υπόθεσης ενώ αυτή ισχύει.
  - Σφάλμα 2<sup>ου</sup> είδους (Type II ή error  $\beta$ )
    - Αποδοχή της μηδενική υπόθεσης ενώ δεν ισχύει.

Συμπέρασμα στατιστικής δοκιμασίας	Πραγματική κατάσταση	
	Ισχύει η $H_0$	Ισχύει η $H_a$
Απόρριψη $H_0$	<b>Σφάλμα 1ου είδους</b>	Ορθό συμπέρασμα
Απόρριψη $H_a$	Ορθό συμπέρασμα	<b>Σφάλμα 2ου είδους</b>

13

## Μηδενική Υπόθεση (3)

- Πως κρίνεται ότι η ισχύει η μηδενική υπόθεση ή η εναλλακτική;
- 1<sup>ο</sup> Κριτήριο
  - Θεωρητική (υπολογιστική) τιμή > Πειραματικής τιμής
    - Αποδοχή μηδενικής υπόθεσης
  - Θεωρητική (υπολογιστική) τιμή < Πειραματικής τιμής
    - Απόρριψη μηδενικής υπόθεσης
- 2<sup>ο</sup> Κριτήριο
  - p-value: πιθανότητα να συμβεί σφάλμα 1<sup>ου</sup> είδους
  - Εξαρτάται από τη στάθμη εμπιστοσύνης:
    - 90%  $p > 0,10$  αποδοχή  $H_0$ ,  $p < 0,10$  απόρριψη  $H_0$
    - 95%  $p > 0,05$  αποδοχή  $H_0$ ,  $p < 0,05$  απόρριψη  $H_0$
    - 99%  $p > 0,01$  αποδοχή  $H_0$ ,  $p < 0,01$  απόρριψη  $H_0$

14

## t-student test, δοκιμασίες t

15

### Σύγκριση με t-test (ανασκόπηση)

- Σύγκριση μέσης πειραματικής τιμής με την πραγματική.

Παράδειγμα: Σύγκριση της πειραματικής τιμής ενός CRM (από επαναλαμβανόμενες μετρήσεις) με αυτή του πιστοποιητικού.

$$t = \frac{(\bar{x} - \mu)\sqrt{n}}{s}$$

- Σύγκριση δύο πειραματικών μέσω τιμών.

Παράδειγμα: Σύγκριση της in-house μεθόδου με την πρότυπη μέθοδο.

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, s^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)}$$

- Σύγκριση μέσω τιμών κατά ζεύγη.

Παράδειγμα: Σύγκριση αποτελεσμάτων δύο μεθόδων για τα ίδια δείγματα

$$t = \frac{\bar{d}\sqrt{n}}{s_d}$$

16



## Σύγκριση μέσης πειραματικής τιμής με την πραγματική.

- Δύο περιπτώσεις:
  - Ενός άκρου (one-tailed),  $H = A > B$ 
    - Σε περιπτώσεις που θέλουμε να δούμε αν μια τιμή είναι στατιστικά σημαντικά πάνω από κάποια άλλη.
  - Δύο άκρων (two-tailed),  $H = A \neq B$ 
    - Σε περιπτώσεις που θέλουμε να δούμε αν μια τιμή διαφέρει στατιστικά από μια άλλη.
- Χρησιμοποιείται συχνά.
- Κανονικά δεν γίνεται με το excel αλλά...
- Πότε το χρησιμοποιούμε:
  - Σύγκριση αν μια τιμή είναι πάνω από το όριο.
  - Σύγκριση αν μια πειραματική τιμή διαφέρει σημαντικά από την αληθή (π.χ. CRM)

17

## Σύγκριση μέσης πειραματικής τιμής με την πραγματική.

- Παράδειγμα: Ένα εργαστήριο ελέγχου αγόρασε ένα CRM για τον προσδιορισμό του Cd σε ιχθυρά. Παρακάτω δίνονται οι τιμές που έλαβε το εργαστήριο και η πιστοποιημένη τιμή του CRM. Να βρεθεί αν υπάρχει συστηματικό σφάλμα στη μέθοδο.

α/α	[Cd] (μg/Kg)
1	267
2	257
3	288
4	255
5	278
6	263
Τιμή πιστοποιητικού	275

18

## Σύγκριση μέσης πειραματικής τιμής με την πραγματική.

- 1<sup>ο</sup> Στάδιο

	A	B	C	D	E	F	G	H
1								
	Τιμές	αληθής τιμή						
	εργαστη							
	ριου							
2	(μg/kg)							
3	267	275						
4	257	275						
5	288							
6	255							
7	278							
8	263							
9								
10								
11								
12								
13								
14								
15								

19

## Σύγκριση μέσης πειραματικής τιμής με την πραγματική.

- 2<sup>ο</sup> Στάδιο

	A	B	C	D	E	F	G	H	I	J	K	L
1												
	Τιμές	αληθής τιμή										
	εργαστη											
	ριου											
2	(μg/kg)											
3	267	275										
4	257	275										
5	288											
6	255											
7	278											
8	263											
9												
10												
11												
12												
13												
14												

20

## Σύγκριση μέσης πειραματικής τιμής με την πραγματική.

- 3<sup>ο</sup> Στάδιο

	A	B	C	D
1	Έλεγχος t δύο δειγμάτων με υποτιθέμενες άνισες διακυμάνσ			
2				
3		Μεταβλητή 1	Μεταβλητή 2	
4	Μέσος	268	275	
5	Διακύμανση	163.2	0	
6	Μέγεθος δείγματος	6	2	
7	Υποτιθέμενη διαφορά μέσω	0		
8	βαθμοί ελευθερίας	5		
9	t	-1.34218873		$t_{exp}$
10	P(T<=t) μονόπλευρη	0.118626992		p-value
11	t κρίσιμο, μονόπλευρο	2.015048372		
12	P(T<=t) δίπλευρη	0.237253984		
13	t κρίσιμο, δίπλευρο	2.570581835		$t_{theor}$
14				
15				

21

## Σύγκριση δύο πειραματικών μέσω τιμών

- Παράδειγμα: Σύγκριση των αποτελεσμάτων της in-house μεθόδου με την πρότυπη ISO

in-house	πρότυπη
3,23	3,51
3,45	3,71
3,48	3,79
3,67	3,98
3,71	

22

	A	B	C	D	E	F	G	H	I	J
1	Έλεγχος t δύο δειγμάτων με υποτιθέμενες ίσες διακυμάνσεις									
2										
3		Μεταβλητή 1	Μεταβλητή 2							
4	Μέσος	3,508	3,7475							
5	Διακύμανση	0,03712	0,037891667							
6	Μέγεθος δείγματος	5	4							
7	Διάμεση διακύμανση	0,037450714								
8	Υποτιθέμενη διαφορά μέσων	0								
9	βαθμοί ελευθερίας	7								
10	t	-1,844884611				t <sub>exp</sub>				
11	P(T<=t) μονόπλευρη	0,053785719						p-value		
12	t κρίσιμο, μονόπλευρο	1,894578605								
13	P(T<=t) δίπλευρη	0,107571439								
14	t κρίσιμο, δίπλευρο	2,364624252								
15										

23

## Σύγκριση δύο πειραματικών μέσω τιμών *F-test*

- Απαραίτητο για την σύγκριση δύο πειραματικών είναι οι δύο πληθυσμοί να έχουν στατιστικά ίδιες διακυμάνσεις
- Ο Έλεγχος γίνεται με F-test.
- Υπενθύμιση:
  - $F = \frac{s_1^2}{s_2^2}$ , τέτοιο ώστε  $F > 1$
  - $F_{theor} > F_{exp}$
- Προσοχή ποια διακύμανση είναι στον αριθμητή και ποια στον παρονομαστή.

24

Excel Χημειομετρία [Κατάσταση λειτουργίας συμβατότητας] - Microsoft Excel

Αρχείο Κεντρική Εισαγωγή Διάταξη σελίδας Τύποι Δεδομένα Αναθεώρηση Προβολή

Από την Access Από το Web Από κείμενο Λήψη εξωτερικών δεδομένων Από άλλες προελεύσεις Υπάρχουσες συνδέσεις Ανανέωση όλων Συνδέσεις Επεξεργασία Συνδέσεις Ταξινόμηση Φίλτρο Απαλοιφή Νέα εφαρμογή Για προχωρημένους Κείμενο σε στήλες Κατάργηση διπλότυπων Ανάλυση πιθανότητας Συνολική εικόνα Ανάλυση πιθανότητας Εργαλεία δεδομένων

D17

	A	B	C	D	E	F	G	H
1	Έλεγχος F των διακυμάνσεων δύο δειγμάτων							
2								
3	<i>Μεταβλητή 1εταβλητή 2</i>							
4	Μέσος	3,7475	3,508					
5	Διακύμανση	0,03789	0,03712					
6	Μέγεθος δείγματος	4	5					
7	βαθμοί ελευθερίας	3	4					
8	F	1,02079						
9	P(F<=f) μονόπλευρη	0,47188						
10	F κρίσιμο, μονόπλευρο	6,59138						
11								
12								
13								
14								
15								
16								

αflatoxin regression Zn ακαρούμενα Άσκηση Zn F-test Zn ANOVA linearity Cu-regr Φύλλο3

25

## Σύγκριση δύο πειραματικών μέσω τιμών *equal vs unequal*

Excel Χημειομετρία - Κατάσταση λειτουργίας συμβατότητας - Excel

Αρχείο Κεντρική Εισαγωγή Διάταξη σελίδας Τύποι Δεδομένα Αναθεώρηση Προβολή Προγραμματισμός Office2012 Τις μου τι θέτε να κάνετε Κουλή γραφή

Λήψη δεδομένων Ανανέωση όλων Επεξεργασία Συνδέσεις Ταξινόμηση Φίλτρο Για προχωρημένους Κείμενο σε στήλες Ανάλυση πιθανότητας προβλεψής Φύλλο Διαφύση Επύση Ανάλυση δεδομένων

Λήψη & δεδομένων μετασημασμού Ερωτήματα & Συνδέσεις Ταξινόμηση & φίλτρα Εργαλεία δεδομένων Πρόβλεψη Ανάλυση

B11

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	Σύγκριση των αποτελεσμάτων της in-house μεθόδου με την πρότυπη ISO																					
2																						
3	in-house πρότυπη																					
4	3.23	3.51																				
5	3.45	3.71																				
6	3.48	3.79																				
7	3.67	3.98																				
8	3.71																					
9																						
10																						
11																						
12																						

Ανάλυση δεδομένων

Εργαλεία ανάλυσης

- Στάθμισμα
- Κλίμακος μέτρ
- Γενίωμα τιμών αριθμών
- Τάξη και ακατοπυρία
- Παινερόμυρία
- Δεγαμολογία
- Ελεγχος t του μέσου δύο δειγμάτων συσχετισμένων ζευγών
- Ελεγχος t δύο δειγμάτων με υποτιθέμενες ίσες διακυμάνσεις
- Ελεγχος t δύο δειγμάτων με υποτιθέμενες άνισες διακυμάνσεις
- Ελεγχος z του μέσου δύο δειγμάτων

OK Ακύρω Εξήθηα

26

## Σύγκριση δύο πειραματικών μέσω τιμών *equal vs unequal*

- *Equal Variances*

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}$$

$$d.f. = n_1 + n_2 - 2$$

- *Unequal Variances*

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$d.f. = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1 - 1)} + \frac{s_2^4}{n_2^2(n_2 - 1)}}$$

27

## Σύγκριση δύο πειραματικών μέσω τιμών *equal vs unequal*

- *Equal Variances*

	Μεταβλητή 1	Μεταβλητή 2
Μέσος	3.508	3.7475
Διακύμανση	0.03712	0.037891667
Μέγεθος δείγματος	5	4
Διάμεση διακύμανση	0.037450714	
Υποτιθέμενη διαφορά μέσων	0	
βαθμοί ελευθερίας	7	
t	-1.84488461	
P(T<=t) μονόπλευρη	0.053785719	
t κρίσιμο, μονόπλευρο	1.894578605	
P(T<=t) δίπλευρη	0.107571439	
t κρίσιμο, δίπλευρο	2.364624252	

- *Unequal Variances*

	Μεταβλητή 1	Μεταβλητή 2
Μέσος	3.508	3.7475
Διακύμανση	0.03712	0.037891667
Μέγεθος δείγματος	5	4
Υποτιθέμενη διαφορά μέσων	0	
βαθμοί ελευθερίας	7	
t	-1.84247578	
P(T<=t) μονόπλευρη	0.053976115	
t κρίσιμο, μονόπλευρο	1.894578605	
P(T<=t) δίπλευρη	0.107952231	
t κρίσιμο, δίπλευρο	2.364624252	

28

## Σύγκριση μέσων τιμών κατά ζεύγη

Παράδειγμα: Σύγκριση πηγών ιοντισμού (ESI-APPI) για τον προσδιορισμό της αφλατοξίνης M1 σε γάλα με LC-MS/MS. Συγκρίθηκαν διαφορετικά δείγματα γάλακτος και με τις δύο πηγές ιοντισμού

α/α	ESI-MS/MS (ng/Kg)	APPI- MS/MS (ng/Kg)
1	12,2	10,8
2	13,2	14
3	18,8	17,9
4	23,5	25
5	18,3	17,9
6	12,4	11,3
7	13,4	15,2
8	14	13,2
9	25,1	23

29

Excel Χημειομετρία [Κατάσταση λειτουργίας συμβατότητας] - Microsoft Excel

	A	B	C	D	E	F
1	t-Test: Paired Two Sample for Means					
2						
3		<i>Variable 1</i>	<i>Variable 2</i>			
4	Mean	16,76667	16,47778			
5	Variance	24,1125	24,59694			
6	Observations	9	9			
7	Pearson Correlation	0,96266				
8	Hypothesized Mean Difference	0				
9	df	8				
10	t Stat	0,64222				
11	P(T<=t) one-tail	0,269347				
12	t Critical one-tail	1,859548				
13	P(T<=t) two-tail	0,538694				
14	t Critical two-tail	2,306004				
15						

two exp values / ISO- inhouse / F-test / Aflatoxin ESI-APPI / aflatoxin paired t-test / aflatoxin regression / Ζη ακρωπ

30

## Άσκηση

- Ο προσδιορισμός του Zn σε δείγματα από αιωρούμενα σωματίδια έγινε με ICP-AES και Ανοδική Αναδιάλυτη Βολταμμετρία. Να μελετηθεί αν διαφέρουν οι δύο τεχνικές σημαντικά.

	ASV (mg/L)	ICP-AES (mg/L)
1	105	90,3
2	88,2	102
3	112	110
4	110	100
5	136	102
6	108	127
7	132	192
8	218	261
9	242	295
10	296	325

31

## ANOVA : ANalysis Of VAriance

- Ελέγχει αν οι διακυμάνσεις (μεταξύ ομάδων & μέσα στις ομάδες) αν διαφέρουν στατιστικά.
- Πιο απλή μορφή η μονόδρομη (ως προς ένα παράγοντα)
- Ο Έλεγχος γίνεται με F-test.
- Όμως:
  - $F = \frac{s_{between}^2}{s_{within}^2}$ , **μπορεί να μην είναι πάντα  $F > 1$**
  - $F_{theor} > F_{exp}$
- Στην ANOVA πάντα στο αριθμητή είναι η διακύμανση μεταξύ των ομάδων και στον παρονομαστή η διακύμανση μέσα στις ομάδες

32



# ANOVA : ANalysis Of VAriance

- Υπολογισμοί (σημειώσεις κ. Ευσταθίου):

ΓΙΑ ΚΑΘΕ ΟΜΑΔΑ  $n_j$  ΜΕΤΡΗΣΕΩΝ  
( $1 \leq j \leq m$ )

- 1) Το άθροισμα (sum) των μετρήσεων:

$$A_j = \sum_{i=1}^{n_j} X_i$$

- 2) Η μέση τιμή της ομάδας (group average):

$$\bar{x}_j = A_j/n_j$$

- 3) Η διακύμανση της ομάδας (group variance):

$$v_j = s_j^2 = \sum_{i=1}^{n_j} (\bar{x}_j - x_i)^2 / (n_j - 1)$$

- 4) Το άθροισμα των τετραγώνων των υπολοίπων (sum of squared residuals):

$$\sum_{i=1}^{n_j} r_i^2 = \sum_{i=1}^{n_j} (\bar{x}_j - x_i)^2 = v_j (n_j - 1)$$

ΓΙΑ ΤΟ ΣΥΝΟΛΟ ΤΩΝ ΜΕΤΡΗΣΕΩΝ

- 1) Ο συνολικός αριθμός μετρήσεων:

$$N = n_1 + n_2 + \dots + n_m$$

- 2) Το μεγάλο άθροισμα (grand sum):

$$A = A_1 + A_2 + \dots + A_m$$

- 3) Η μεγάλη μέση τιμή (grand average):

$$\bar{x} = A/N$$

- 4) Το άθροισμα των τετραγώνων των υπολοίπων (sum of squared residuals):

$$S_1 = \sum_{i=1}^{n_1} r_1^2 + \sum_{i=1}^{n_2} r_2^2 + \dots + \sum_{i=1}^{n_m} r_m^2 = \sum_{j=1}^m \left( \sum_{i=1}^{n_j} r_i^2 \right)$$

- 5) Το άθροισμα των ζυγισμένων τετραγώνων (sum of weighted squares):

$$S_2 = n_1 (\bar{x} - \bar{x}_1)^2 + n_2 (\bar{x} - \bar{x}_2)^2 + \dots + n_m (\bar{x} - \bar{x}_m)^2 = \sum_{k=1}^m (n_k - \bar{x}_k)^2$$

- 6) Το ολικό άθροισμα τετραγώνων (total sum of squares)

$$S_T = S_1 + S_2$$

33

## ANOVA : Σύγκριση Διακυμάνσεων

- Πίνακας Ανατομίας της ANOVA

Source of variation	Sum of squares (SS)	$\nu$	Mean square (MS)	$F$	$P$	$F_{crit}$
Between groups	$SS_b$	$p-1$	$MS_b = SS_b/(p-1)$	$MS_b/MS_w$		
Within group (residuals)	$SS_w$	$N-p$	$MS_w = SS_w/(N-p)$			
Total	$SS_{tot} = SS_b + SS_w$	$N-1$				

- $F_{crit} > F_{(exp)}$  ή ( $p\text{-value} > 0,05$ ) ισχύει η μηδενική υπόθεση, οι ομάδες προέρχονται από τον ίδιο πληθυσμό.
- $F_{crit} < F_{(exp)}$  ή ( $p\text{-value} < 0,05$ ) ισχύει η εναλλακτική υπόθεση, υπάρχει στατιστικά σημαντική διαφορά στους πληθυσμούς.

34

Excel Χημειομετρία [Κατάσταση λειτουργίας συμβατότητας] - Microsoft Excel

Αρχείο Κεντρική Εισαγωγή Διάταξη σελίδας Τύποι Δεδομένα Αναθεώρηση Προβολή

Calibri 11 Α<sup>+</sup> Α<sup>-</sup> Αναδίπλωση κειμένου Γενική

Επικόλληση Β U Σ Συγχώνευση και στοιχισή στο κέντρο Συγχώνευση και στοιχισή στο κέντρο

Πρόχειρο Γραμματοσειρά Στοιχισή Αριθμός Μορφοποίηση υπό όρους Μορφοποίηση ως πίνακας

F24

	A	B	C	D	E	F	G	H	I
1	Anova: Single Factor								
2									
3	SUMMARY								
4		<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>			
5	Column 1	10	1547,2	154,72	5039,273				
6	Column 2	10	1704,3	170,43	8262,471				
7	Column 3	10	1695	169,5	4747,833				
8									
9									
10	ANOVA								
11	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>		
12	Between Groups	1553,725	2	776,8623	0,129121	0,879407	3,354131		
13	Within Groups	162446,2	27	6016,526					
14									
15	Total	163999,9	29						
16									
17									
18									

αflatoxin regression Zn ακαρούμενα Άσκηση Zn F-test Zn ANOVA linearity Cu-regr Φύλλο3

35

## ANOVA : Ομοιογένεια

$MS_{between} > MS_{within}$

$$S_{bb} = \sqrt{\frac{MS_{between\ groups} - MS_{within\ groups}}{n}}$$

$MS_{between} < MS_{within}$

$$u_{bb} = \sqrt{\frac{MS_{within}}{n} \cdot \frac{2}{N(n-1)}}$$

36

Excel Χημειομετρία [Κατάσταση λειτουργίας συμβατότητας] - Microsoft Excel

Αρχείο Κεντρική Εισαγωγή Διάταξη σελίδας Τύποι Δεδομένα Αναθεώρηση Προβολή

Calibri 11 A<sup>+</sup> A<sup>-</sup> Αναδίπλωση κειμένου Γενική

Επικόλληση Προχείρο Γραμματοσειρά Στοιχισμός Συγχώνευση και στοιχισμός στο κέντρο Αριθμός Μορφοποίηση ως πίνακας

F24

	A	B	C	D	E	F	G	H	I
1	Anova: Single Factor								
2									
3	SUMMARY								
4	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>				
5	Column 1	10	1547,2	154,72	5039,273				
6	Column 2	10	1704,3	170,43	8262,471				
7	Column 3	10	1695	169,5	4747,833				
8									
9									
10	ANOVA								
11	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>		
12	Between Groups	1553,725	2	776,8623	0,129121	0,879407	3,354131		
13	Within Groups	162446,2	27	6016,526					
14									
15	Total	163999,9	29						
16									
17									
18									

Πίν

afiatoxin regression Zn ακαρούμενα Άσκηση Zn F-test Zn ANOVA linearity Cu-regr Φύλλο3

37

## ANOVA : Σύγκριση Διακυμάνσεων

- Πίνακας Ανατομίας της ANOVA

Source of variation	Sum of squares (SS)	$\nu$	Mean square (MS)	$F$	$P$	$F_{crit}$
Between groups	$SS_b$	$p-1$	$MS_b = SS_b/(p-1)$	$MS_b/MS_w$		
Within group (residuals)	$SS_w$	$N-p$	$MS_w = SS_w/(N-p)$			
Total	$SS_{tot} = SS_b + SS_w$	$N-1$				

- Επαναληψιμότητα:  $s_r = \sqrt{MS_{within}}$
- $s_{between} = \sqrt{\frac{MS_{between} - MS_{within}}{n}}$
- Ενδιάμεση Πιστότητα:  $s_l = \sqrt{s_{between}^2 + s_r^2}$

38

## Υπολογισμός Ενδιάμεσης Πιστότητας με την ANOVA

- Εναλλακτικός τρόπος υπολογισμού της ενδιάμεσης πιστότητας.
- Όταν υπάρχει ανόμοιος αριθμός μετρήσεων για τις διαφορετικές ομάδες μετρήσεων τότε χρειάζεται ένας συντελεστής διόρθωσης
- EURACHEM Guide Fitness For Purpose, <https://eurachem.org/>