



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΙΑΤΡΙΚΗ ΣΧΟΛΗ



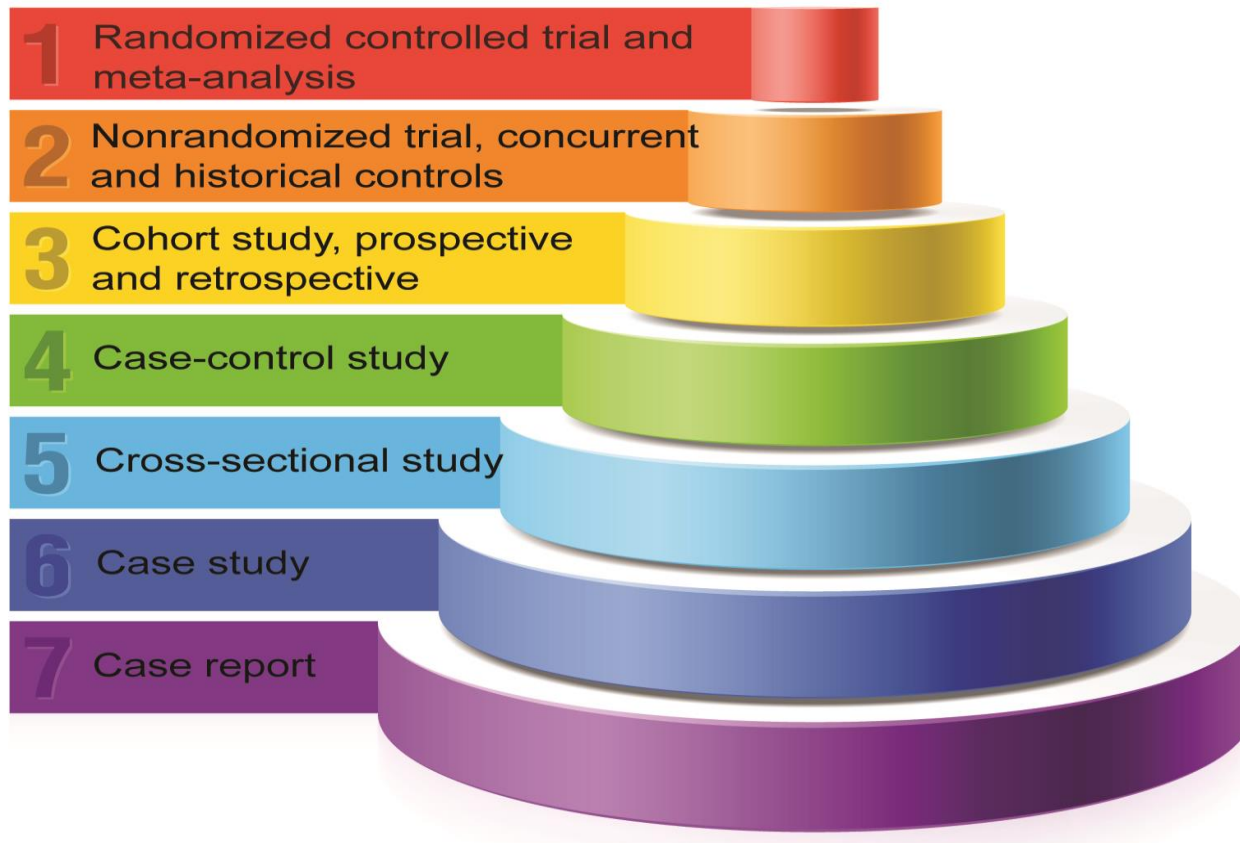
Η ΣΤΑΤΙΣΤΙΚΗ ΣΤΟ ΕΡΓΑΣΤΗΡΙΟ ΣΠΕΡΜΑΤΟΣ

Έλενα Κριτσέλη, MPH PhD

Επιστημονικός Συνεργάτης – Επιδημιολόγος Χρόνιων Παθήσεων,
Α΄ Πανεπιστημιακή Παιδιατρική Κλινική,
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
Ίδρυμα Ιατροβιολογικών Ερευνών, Ακαδημίας Αθηνών

Η θέση της Εργαστηριακής Ανδρολογίας στην Κλινική Πράξη

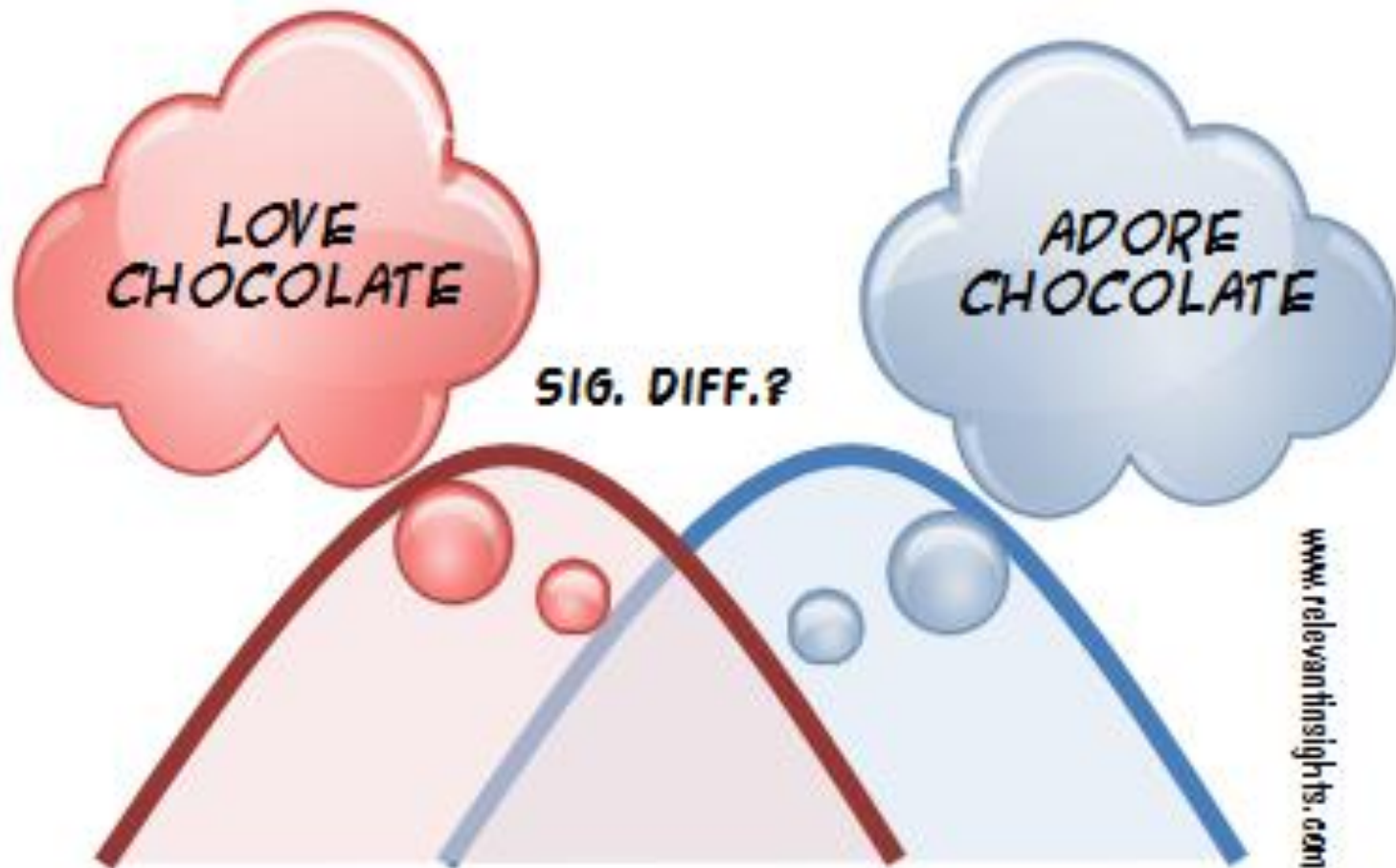
Ιεράρχηση επιδημιολογικών μελετών



Έλεγχος υποθέσεων



Η ερευνητική ιδέα...



Έλεγχος υποθέσεων

- **Στόχος:** Η σύγκριση μεταξύ 2 ή περισσότερων χαρακτηριστικών / μεταβλητών
- Πραγματοποιείται **έλεγχος ερευνητικών υποθέσεων**, όπου:
 - **Μηδενική υπόθεση (H_0):** Απουσία συσχέτισης 2 χαρακτηριστικών / μεταβλητών
 - **Εναλλακτική υπόθεση (H_A):** Ύπαρξη συσχέτισης μεταξύ 2 μεταβλητών



Πιθανά σφάλματα

		<u>Πραγματική Κατάσταση στον πληθυσμό από όπου το δείγμα εξάγεται</u>	
		Η H_0 είναι αληθής (A)	Η H_0 είναι ψευδής (Ψ)
<u>Βασισμένη στο δείγμα η απόφαση είναι:</u>	Η H_0 είναι αληθής (A) (αποδεκτή)	Σωστή απόφαση, (1- α)	Σφάλμα Τύπου II, (β)
	Η H_0 είναι ψευδής (Ψ) (απορρίπτεται)	Σφάλμα Τύπου I, (επίπεδο σημαντικότητας (α))	Σωστή απόφαση (Ισχύς (1- β))

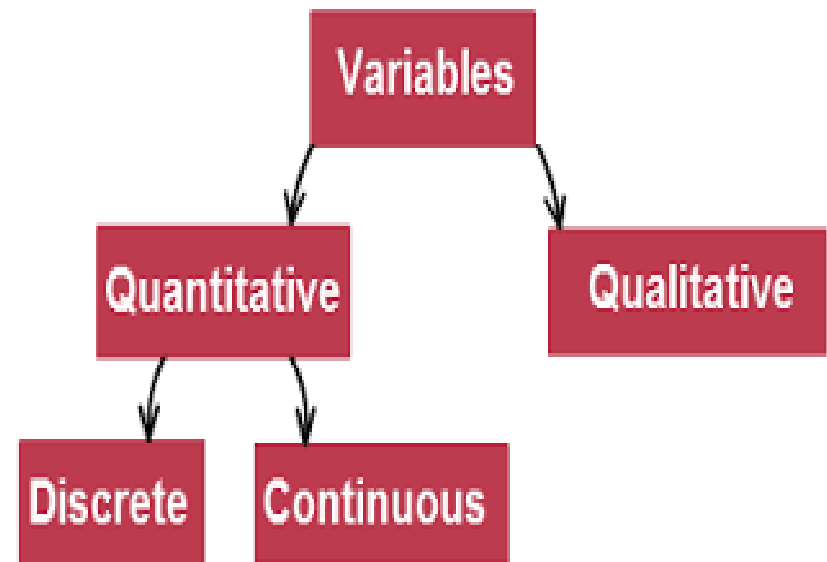
Κριτήριο σημαντικότητας (p-value)

- **Κριτήριο σημαντικότητας:** Η ακριβής τιμή του σφάλματος Τύπου I
 - Βάσει των ερευνητικών ευρημάτων, συμπεραίνεται ότι υπάρχει συσχέτιση μεταξύ των μεταβλητών υπό αξιολόγηση, ενώ στην πραγματικότητα **δεν υφίσταται** πραγματική συσχέτιση
 - Θεμιτό επίπεδο $<5\%$ ($p < 0.05$)



Υλοποίηση ελέγχου υποθέσεων

- Ο έλεγχος υποθέσεων απαιτεί την επιλογή της κατάλληλης στατιστικής δοκιμασίας
- Η επιλογή της δόκιμης στατιστικής μεθόδου εξαρτάται από:
 - Τη **φύση** των ερευνητικών δεδομένων / μεταβλητών
 - Τον **αριθμό των ομάδων** υπό αξιολόγηση



Μονοπαραγοντικές αναλύσεις

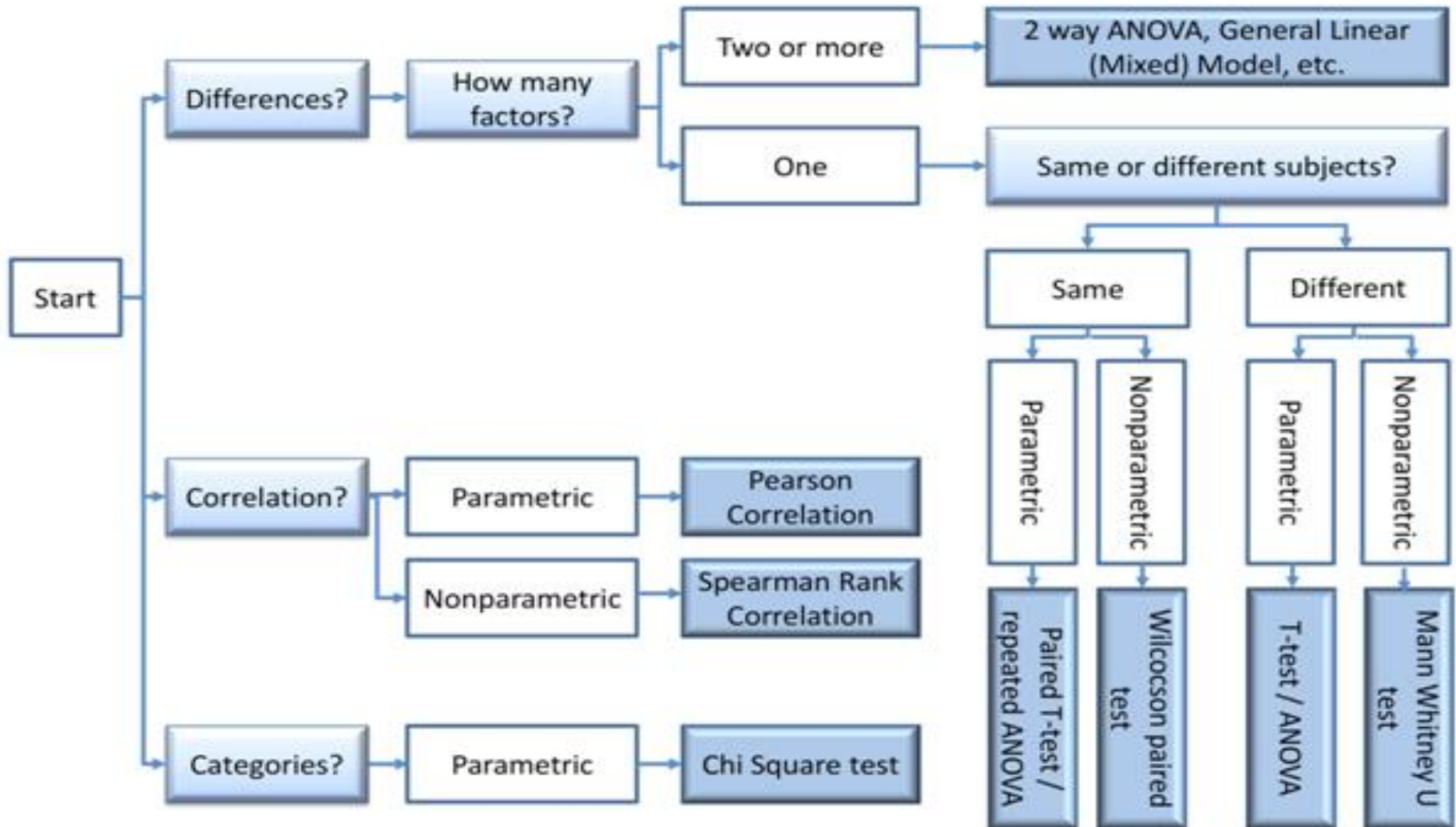


Συγκρίσεις μεταξύ 2 μεταβλητών

- Κατά την σύγκριση 2 χαρακτηριστικών, δύναται να προκύψουν οι ακόλουθοι συνδυασμοί μεταβλητών:
 - ▣ **Ποιοτική με ποιοτική**
 - Φύλο με παχυσαρκία
 - ▣ **Ποιοτική με ποσοτική**
 - Φύλο με ηλικία
 - ▣ **Ποσοτική με ποσοτική**
 - ΔΜΣ με ηλικία
- Οι συγκρίσεις δύνανται να εφαρμοστούν σε **2 ή περισσότερες ομάδες ασθενών**



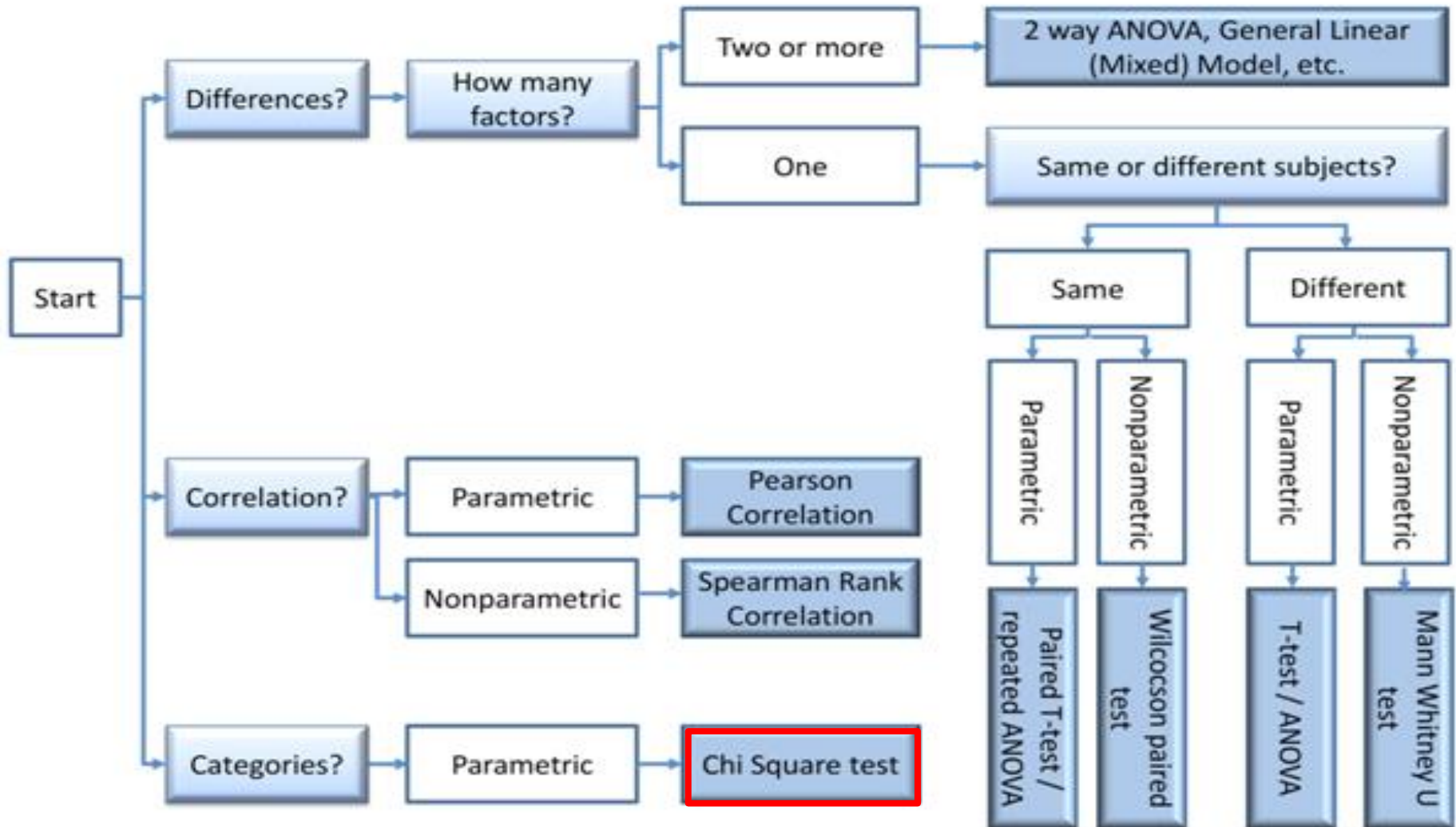
Είδη μονοπαραγοντικών δοκιμασιών



Σύγκριση ποιοτικών μεταβλητών



Είδη μονοπαραγοντικών δοκιμασιών



Δοκιμασία χ^2 (Chi-squared test)

- **Στόχος:** Σύγκριση της συχνότητας παρουσίας ενός ποιοτικού χαρακτηριστικού μεταξύ 2 ομάδων ασθενών
 - Αμοιβαία εξαιρετές κατηγορίες
- **Έλεγχος υποθέσεων:**
 - H_0 : Η συχνότητα παρουσίας του ποιοτικού χαρακτηριστικού είναι ίδια στις 2 ομάδες
 - H_A : Η συχνότητα του ποιοτικού χαρακτηριστικού ανευρίσκεται πιο συχνά στην ομάδα υπό αξιολόγηση
- Δύναται να εφαρμοστεί **ανεξάρτητα** από τον **αριθμό των κατηγοριών** των 2 ποιοτικών μεταβλητών

Παράδειγμα: Pearson Chi-Square

Gender * Knowledge Category Crosstabulation

		Knowledge Category		Total	
		High	Low		
Gender	Female	Count	20	14	34
		% within Gender	58.8%	41.2%	100.0%
	Male	Count	20	19	39
		% within Gender	51.3%	48.7%	100.0%
Total		Count	40	33	73
		% within Gender	54.8%	45.2%	100.0%

Chi-square statistic = 0.417
df = 1; P-value = 0.518

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.417 ^a	1	.518		
Continuity Correction ^b	.168	1	.682		
Likelihood Ratio	.418	1	.518		
Fisher's Exact Test				.638	.341
Linear-by-Linear Association	.411	1	.521		
N of Valid Cases	73				

Must be ≥ 2

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 15.37

b. Computed only for a 2x2 table

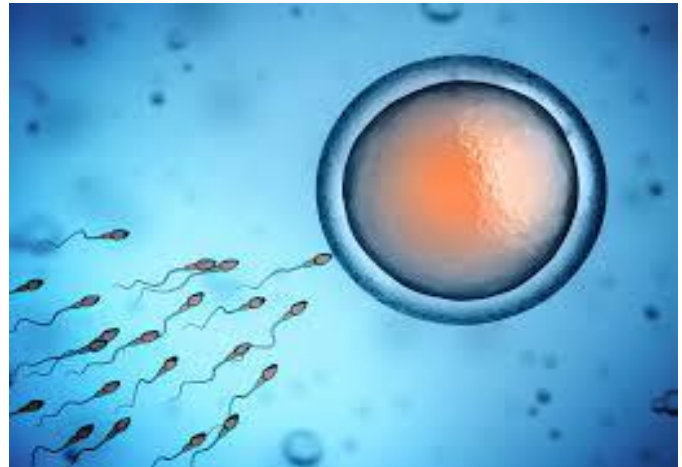
Must be $< 20\%$

2 EC assumptions is met

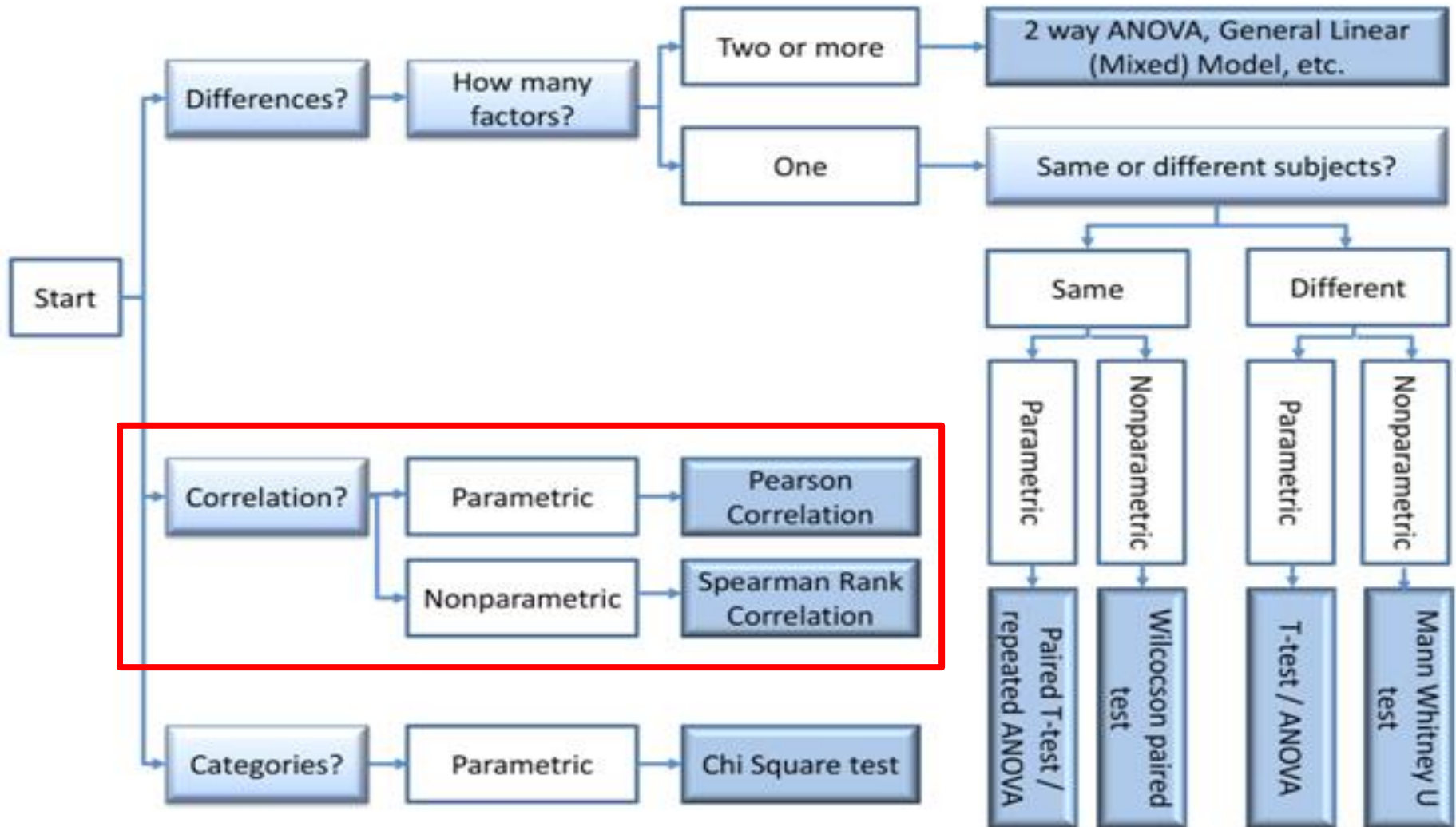
Δοκιμασία του Fisher

- Σε περίπτωση που η αναμενόμενη συχνότητα, σε οποιαδήποτε υποομάδα, είναι < 5 , τότε εφαρμόζεται η μη παραμετρική **δοκιμασία του Fisher**
- Αποδίδει την **τιμή p**, η οποία δεν βασίζεται στην χ^2 -κατανομή
- Ισχύει αντίστοιχη ερμηνεία των αποτελεσμάτων

Συσχέτιση μεταβλητών



Είδη μονοπαραγοντικών δοκιμασιών

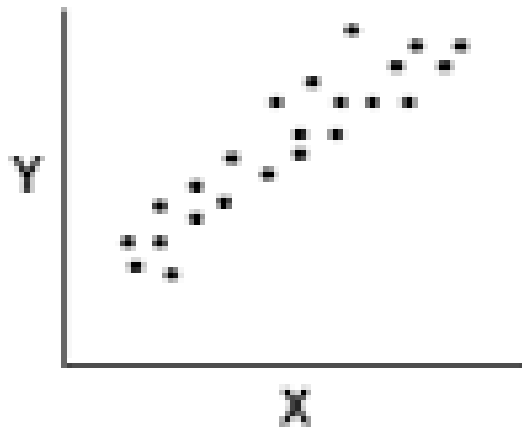


Ανάλυση συσχέτισης ποσοτικών μεταβλητών

- Υπολογισμός του βαθμού συσχέτισης μεταξύ 2 ποσοτικών μεταβλητών
 - Σε ποιο βαθμό η ποσοτική μεταβλητή X επηρεάζεται γραμμικά από την ποσοτική μεταβλητή Y
 - Π.χ. Ανάστημα και βάρος
- **Στικτόγραμμα (Scatter diagram):**
Απεικόνιση των ζευγών των μεταβλητών X και Y στην ομάδα υπό αξιολόγηση

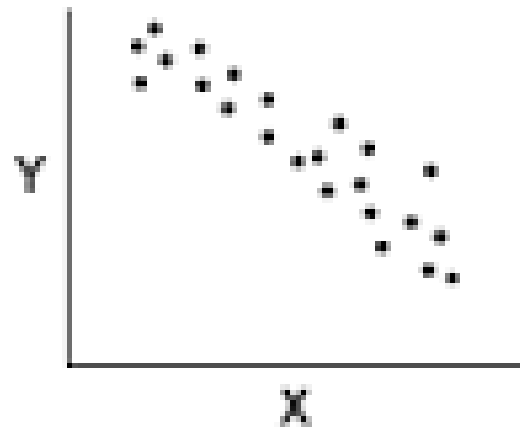


Scatter diagrams



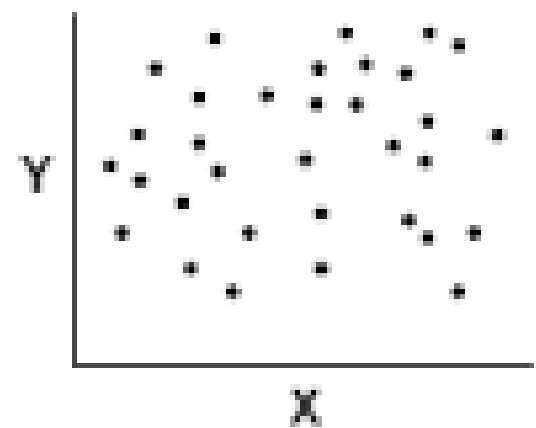
Positive Correlation

For an increase in "X" there is a corresponding increase in "Y"



Negative Correlation

For an increase in "X" there is a corresponding decrease in "Y"



No Correlation

For an increase in "X" there is no corresponding reaction in "Y"

Συντελεστές συσχέτισης (Correlation coefficients)

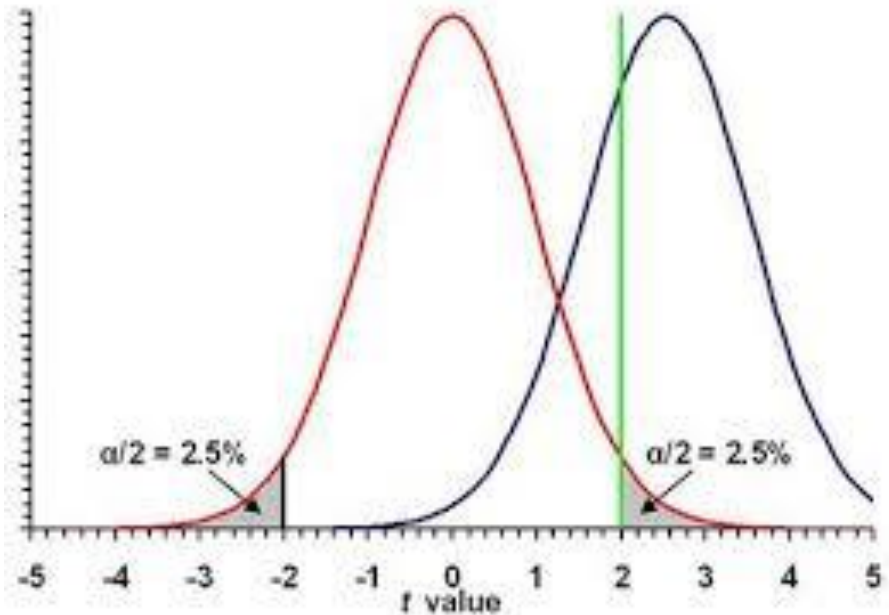
- Συντελεστής συσχέτισης:
 - ▣ του Pearson (r)
 - ▣ του Spearman (ρ)

- Ενδεικνυόμενες εφαρμογές:

		Ποσοτικές συνεχείς μεταβλητές		Ποσοτικές διακριτές μεταβλητές
		Κανονική κατανομή	Μη κανονική κατανομή	
Ποσοτικές συνεχείς μεταβλητές	Κανονική κατανομή	Pearson	Spearman	Spearman
	Μη κανονική κατανομή	Spearman	Spearman	Spearman
Ποσοτικές διακριτές μεταβλητές		Spearman	Spearman	Spearman

Φυσιολογικές κατανομές ποσοτικών μεταβλητών

- Η κανονικότητα κατανομής μπορεί να αξιολογηθεί με την χρήση των ακόλουθων απεικονίσεων και δοκιμασιών:
 - Ιστόγραμμα
 - ή
 - Δοκιμασία Kolmogorov-Smirnov
 - Δοκιμασία Shapiro Wilks



Συντελεστές συσχέτισης

□ Εύρος τιμών:

- Από **-1** (αρνητική συσχέτιση) έως **+1** (θετική συσχέτιση)
- Όσο ο συντελεστής πλησιάζει τις ακραίες τιμές (-1 ή +1), τόσο πιο ισχυρή είναι η συσχέτιση.

Ισχυρή αρνητική συσχέτιση	Μέτρια αρνητικά συσχετισμένες	Ελαφρά αρνητικά συσχετισμένες	Ασυσχέτιστες ή πολύ ελαφρά συσχετισμένες	Ελαφρά θετικά συσχετισμένες	Μέτρια θετικά συσχετισμένες	Ισχυρή θετική συσχέτιση
<-0,80	-0,80...-0,60	-0,60...-0,30	-0,20...+0,20	+0,30...+0,60	+0,60...+0,80	> +0,80

Συντελεστές συσχέτισης

- Οι συντελεστές συσχέτισης εκφράζουν τον βαθμό κατά τον οποίο 2 ποσοτικές μεταβλητές συσχετίζονται **γραμμικά** μεταξύ τους
 - ▣ H_0 : Δεν υπάρχει **γραμμική** συσχέτιση μεταξύ των μεταβλητών

- **Προσοχή!**

- ▣ Τυχόν μηδενικός συντελεστής ενδέχεται να αντικατοπτρίζει **λογαριθμική** συσχέτιση



- Ως εκ τούτου, οι συντελεστές εκφράζουν **μόνο την ισχύ της γραμμικής συσχέτισης** των μεταβλητών

Παράδειγμα: Συντελεστής Pearson

“There was a positive correlation between the two variables, $r = 0.985$, $n = 5$, $p = 0.002$.”

Correlations

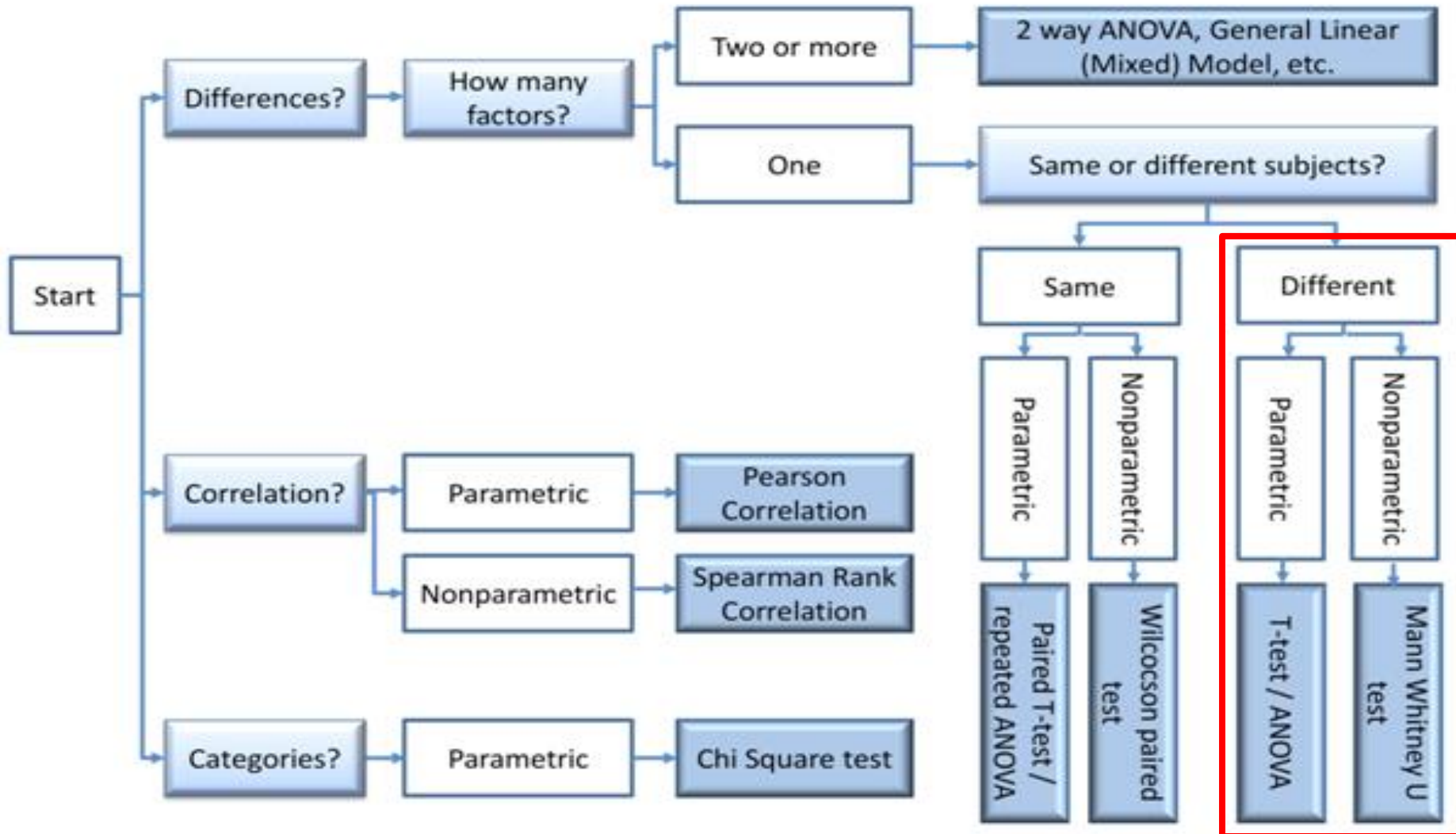
		WATER	SKIN
WATER	Pearson Correlation	1.000	.985
	Sig. (2-tailed)	.	.002
	N	5	5
SKIN	Pearson Correlation	.985**	1.000
	Sig. (2-tailed)	.002	.
	N	5	5

** . Correlation is significant at the 0.01 level

Σύγκριση ποσοτικών μεταβλητών σε ανεξάρτητες ομάδες ασθενών



Είδη μονοπαραγοντικών δοκιμασιών



Καθορισμός δόκιμης δοκιμασίας

Ποιοτική με 2 κατηγορίες (π.χ. φύλο, παρουσία ή απουσία νόσου)

Έλεγχος κανονικότητας για την ποσοτική μεταβλητή με στατιστικό κριτήριο (π.χ. Kolmogorov-Smirnov test ή Shapiro Wilk) ή γραφικά με ιστόγραμμα.

t- test

Αν η ποσοτική μεταβλητή ακολουθεί την κανονική κατανομή.

Mann – Whitney test

Αν η ποσοτική μεταβλητή δεν ακολουθεί την κανονική κατανομή

Ποιοτική με >2 κατηγορίες (π.χ. τύπος σάματος βάση του ΔΜΣ: φυσιολογικό βάρος, υπέρβαροι παχύσαρκοι).

1. Έλεγχος κανονικότητας για την ποσοτική μεταβλητή με στατιστικό κριτήριο (π.χ. Kolmogorov-Smirnov ή Shapiro Wilk) ή γραφικά με ιστόγραμμα.
2. Έλεγχος ισότητας των διακυμάνσεων με το Levene test

ANOVA

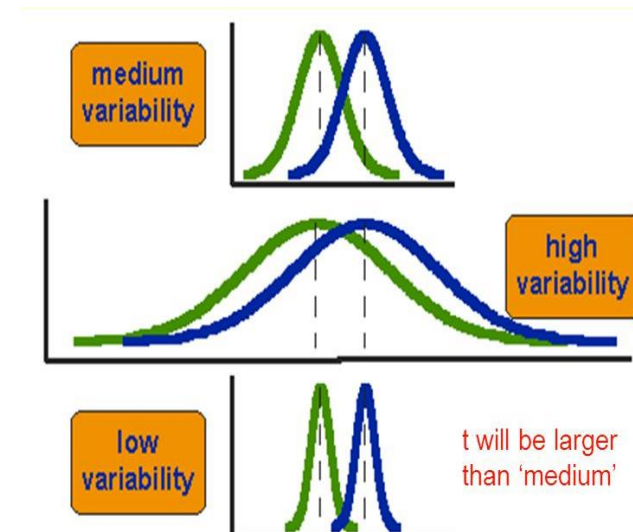
Αν η ποσοτική μεταβλητή ακολουθεί την κανονική κατανομή & οι διακυμάνσεις είναι ίδιες.

Kruskal-Wallis

Αν η ποσοτική μεταβλητή δεν ακολουθεί την κανονική κατανομή ή οι διακυμάνσεις δεν είναι ίδιες ή & τα 2.

Student's t-test

- Ελέγχεται εάν η **μέση τιμή** μιας μεταβλητής στην ομάδα A διαφέρει από την αντίστοιχη τιμή στην ομάδα B.
- **Προϋπόθεση:** Η μεταβλητή πρέπει να ακολουθεί την **κανονική κατανομή** και στις 2 ομάδες υπό αξιολόγηση
 - Οπότε, πρέπει πρώτα να εφαρμόζεται **έλεγχος για την ισότητα των διακυμάνσεων**
 - Δοκιμασίες Kolmogorov-Smirnov ή Shapiro Wilks
- Η μηδενική και η εναλλακτική υπόθεση:
 - $H_0: \mu_1 = \mu_2$
 - $H_A: \mu_1 \neq \mu_2$



Παράδειγμα: Student's t-test

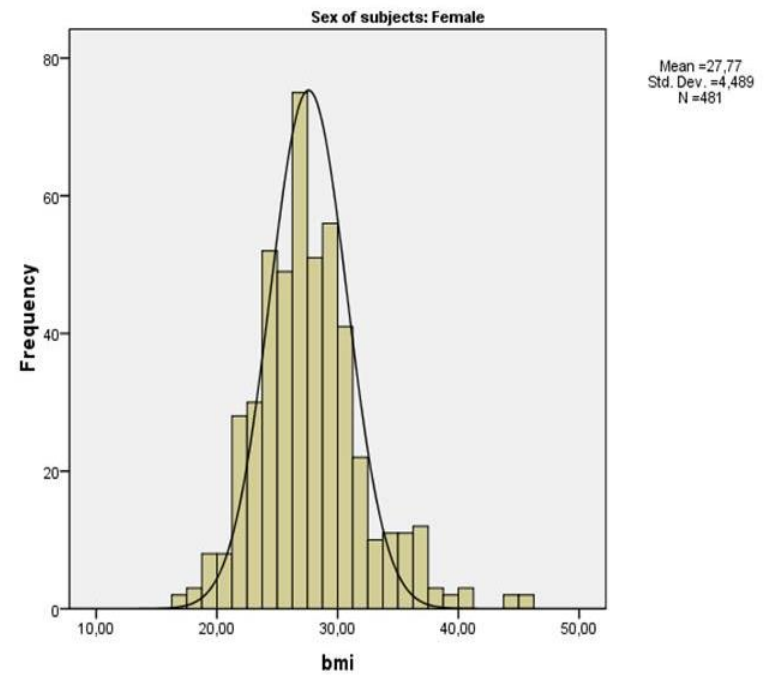
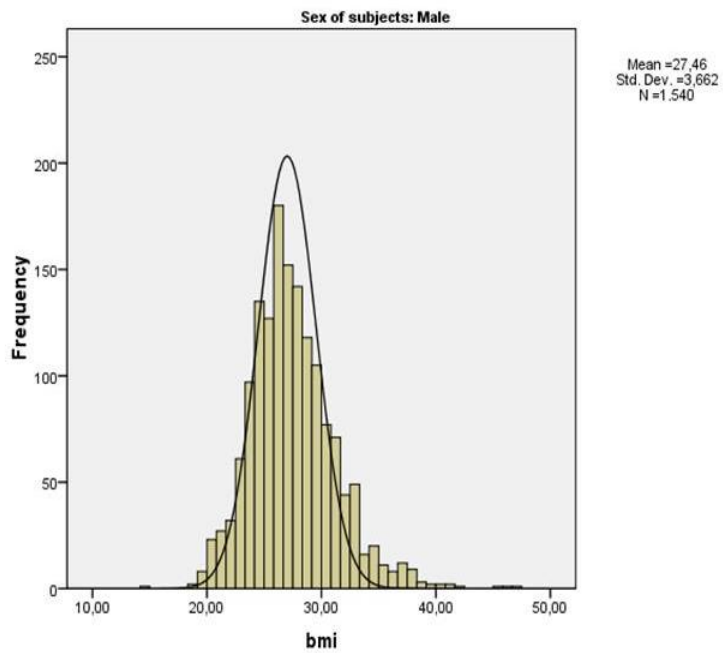
Group Statistics

	sex	N	Mean	Std. Deviation	Std. Error Mean
bmi	male	1540	27,4628	3,66187	,09331
	female	481	27,7715	4,48867	,20467
SBPmmHg	male	1312	136,29	25,707	,710
	female	408	141,57	26,940	1,334

Independent Samples Test

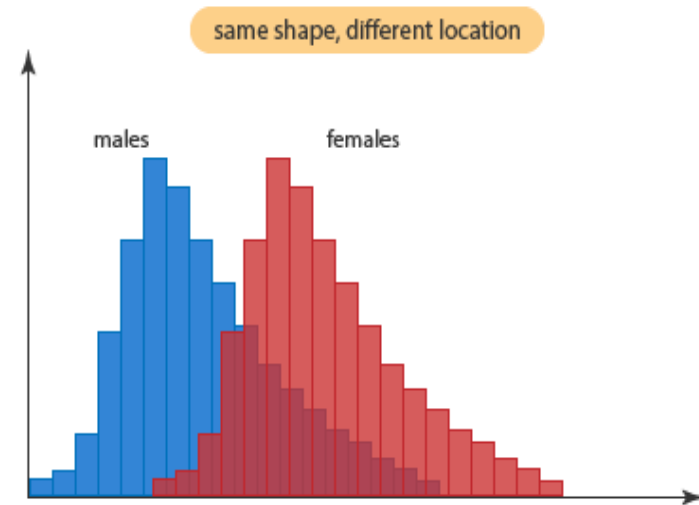
		Levene's Test for Equality of Variance		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
bmi	Equal variances assumed	20,101	,000	-1,526	2019	,127	-,30876	,20238	-,70565	,08813
	Equal variances not assumed			-1,373	690,985	,170	-,30876	,22493	-,75040	,13287
SBPmmHg	Equal variances assumed	,980	,322	-3,587	1718	,000	-5,288	1,474	-8,179	-2,397
	Equal variances not assumed			-3,500	653,849	,000	-5,288	1,511	-8,254	-2,321

Όμως...



Δοκιμασία Mann-Whitney

- Αποτελεί την αντίστοιχη μη παραμετρική δοκιμασία του t-test για 2 ανεξάρτητες ομάδες
 - ▣ Δεν απαιτεί οι τιμές των ομάδων υπό αξιολόγηση να ακολουθούν την κανονική κατανομή
- Η μηδενική υπόθεση:
 - ▣ Οι **διάμεσες τιμές** των 2 ομάδων είναι ισότιμες, ή,
 - ▣ Οι τιμές των 2 ομάδων **κατανέμονται** με τον ίδιο τρόπο



Παράδειγμα: Δοκιμασία Mann-Whitney

N	Valid	1435
	Missing	214
Mean		77,69
Std. Deviation		122,696
Percentiles	25	15,00
	50	34,00
	75	85,00

a. Sex of subjects = Male

N	Valid	449
	Missing	74
Mean		57,35
Std. Deviation		92,175
Percentiles	25	13,00
	50	28,00
	75	61,00

a. Sex of subjects = Female

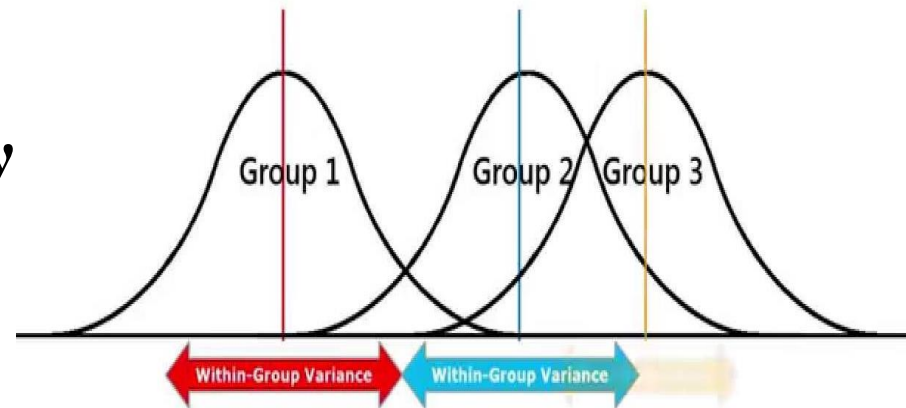
Test Statistics^a

	cpkmbmax
Mann-Whitney U	286623,50
Wilcoxon W	387648,50
Z	-3,532
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: sex

Ανάλυση Διακύμανσης (One-way analysis of variance, ANOVA)

- Έλεγχος για την ύπαρξη διαφορών στις μέσες τιμές ενός ποσοτικού χαρακτηριστικού μεταξύ **>2 ομάδων**
- **Προϋπόθεση:** Η ποσοτική μεταβλητή πρέπει να κατανέμεται **κανονικά** σε όλες τις ομάδες
- Έλεγχος για την **ισότητα των διακυμάνσεων**



Παράδειγμα: ANOVA

“There was a significant effect of amount of sugar on words remembered at the $p < .05$ level for the three conditions [$F(2, 12) = 4.94, p = 0.027$].”

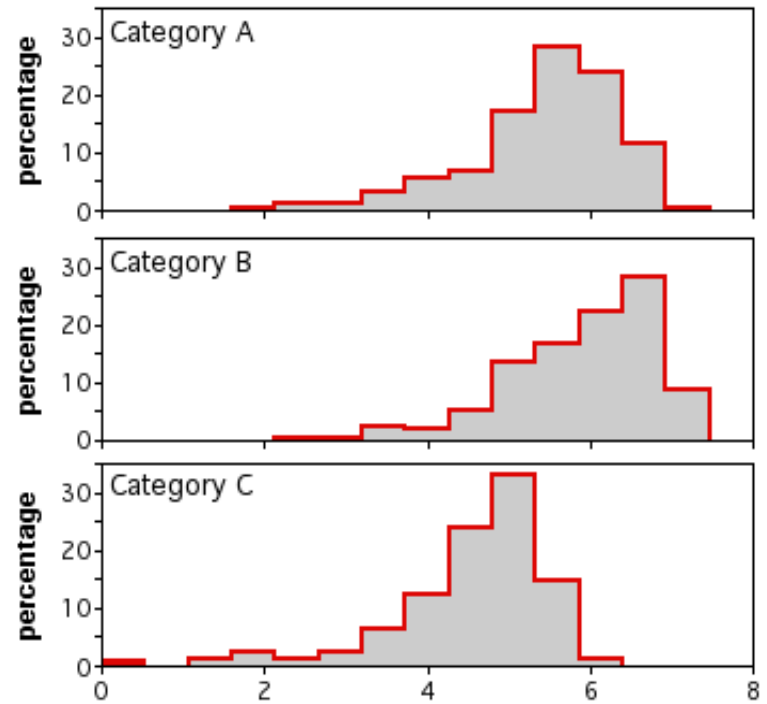
ANOVA

DVWORDS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.533	2	5.267	4.938	.027
Within Groups	12.800	12	1.067		
Total	23.333	14			

Δοκιμασία Kruskal-Wallis

- Αντίστοιχη μη παραμετρική δοκιμασία της ANOVA
- Αποτελεί επέκταση της δοκιμασίας Mann-Whitney σε >2 ομάδες
- Έλεγχος για την ισότητα κατανομής σε >2 ομάδων



Παράδειγμα: Δοκιμασία Kruskal-Wallis

Kruskal-Wallis Test

Drug Treatment Group		N	Mean Rank
Pain_Score	Drug A	20	35.33
	Drug B	20	34.83
	Drug C	20	21.35
	Total	60	

Test Statistics^{a,b}

	Pain_Score
Chi-square	8.520
df	2
Asymp. Sig.	.014

a. Kruskal Wallis Test

b. Grouping Variable:
Drug Treatment
Group

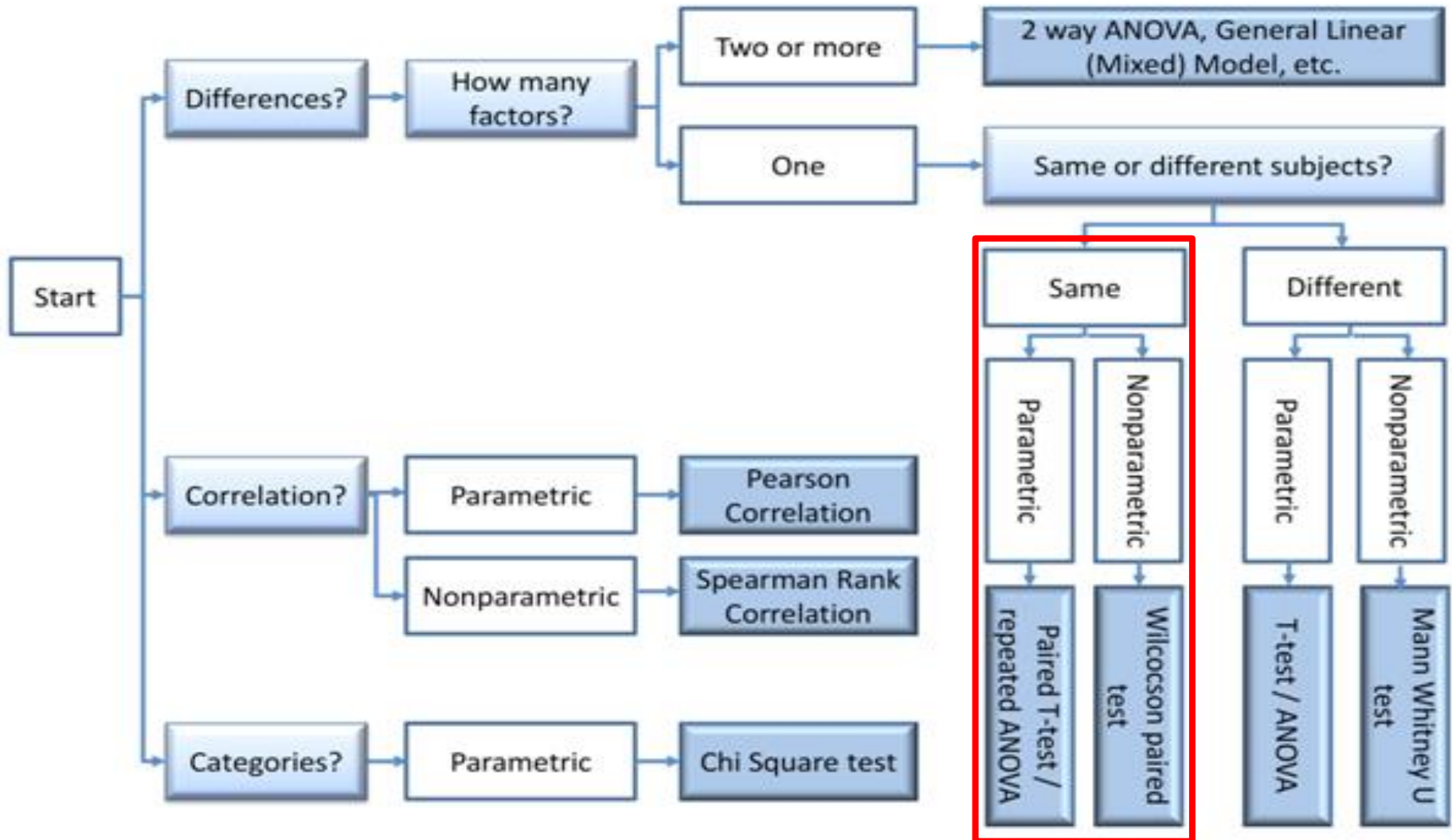
OUTPUT INTERPRETATION

There was a statistically significant difference in **pain score** between the **different drug treatments**, ($\chi^2(2) = 8.520, p = 0.014$)

Σύγκριση κατά ζεύγη ποσοτικών μεταβλητών



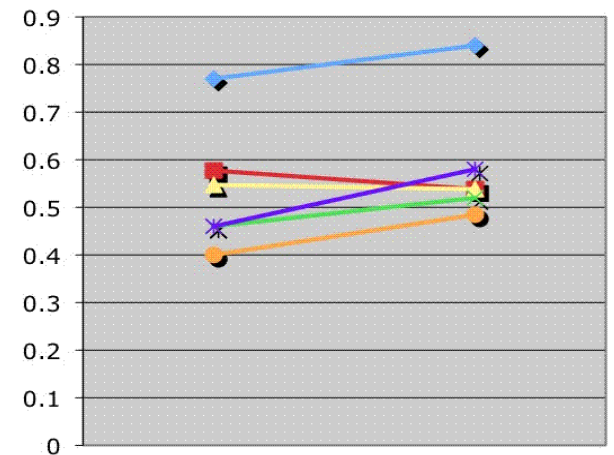
Είδη μονοπαραγοντικών δοκιμασιών



Συγκρίσεις παρατηρήσεων κατά ζεύγη

- **Παρατηρήσεις κατά ζεύγη:** η μεταβλητή έχει μετρηθεί σε κάθε άτομο σε **δύο περιπτώσεις**
 - ▣ Πχ Πριν και μετά την εφαρμογή θεραπευτικής παρέμβασης, ώστε να αξιολογηθεί η αποτελεσματικότητα της παρέμβασης

- Έλεγχος της διαφοράς της **μέσης μεταβολής** των τιμών των 2 εξαρτημένων δειγμάτων



- ▣ H_0 : Η μέση διαφορά στον πληθυσμό ισούται με 0

- ▣ H_A : Η μέση διαφορά στον πληθυσμό δεν ισούται με 0

Συγκρίσεις παρατηρήσεων κατά ζεύγη



Παράδειγμα: T-test κατά ζεύγη

1

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-Knowledge Test	7.80	100	4.731	.473
	Final Exam Score	71.10	100	15.821	1.582

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pre-Knowledge Test & Final Exam Score	100	.664	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pre-Knowledge Test - Final Exam Score	-63.30	13.166	1.317	-65.91	-60.69	-48.078	99	.000

2

Παράδειγμα: T-test κατά ζεύγη

	Control Group	Experimental Group
Signal R1	3700	4900
Signal R2	4000	5200
Signal R3	4200	4900
Signal R4	3900	5000
Signal R5	4100	4800
Signal R6	4000	4750

Paired T test

$$\text{Mean1} = 3983 \quad v_1 = 5$$

$$\text{Mean2} = 4925 \quad v_2 = 5$$

$$\frac{SE(d_1 - d_2)}{5} = \sqrt{228.065} = 45.61$$

$$t = \frac{941.67}{45.61} = 20.65$$

$t_{0.05,10} = 2.228$ as $20.65 > 2.228$ then reject H_0 :

$P < 0.0001$. The differences between the means is greater than 0.

Unpaired T test

$$\text{Mean1} = 3983 \quad \text{Sum of Squares 1} = 148334 \quad v_1 = 5$$

$$\text{Mean2} = 4925 \quad \text{Sum of Squares 2} = 128750 \quad v_2 = 5$$

$$S^2_p = \frac{148334 + 128750}{5 + 5} = 29666.8$$

$$S_{\text{mean1}-\text{mean2}} = \sqrt{\frac{29666.8}{6}} + \sqrt{\frac{29666.8}{6}} = 140.63$$

$$t = \frac{3983 - 4925}{140.63} = -6.70$$

$t_{0.05,10} = 2.228$ as $6.7 > 2.228$ then reject H_0 :

$P < 0.0001$. The two means are not the same.

Παράδειγμα: Wilcoxon Signed Ranks

Ranks

	N	Mean Rank	Sum of Ranks
AFTER - BEFORE Negative Ranks	3 ^a	3.67	11.00
Positive Ranks	8 ^b	6.88	55.00
Ties	1 ^c		
Total	12		

a. AFTER < BEFORE

b. AFTER > BEFORE

c. BEFORE = AFTER

Test Statistics^b

	AFTER - BEFORE
Z	-1.971 ^a
Asymp. Sig. (2-tailed)	.049

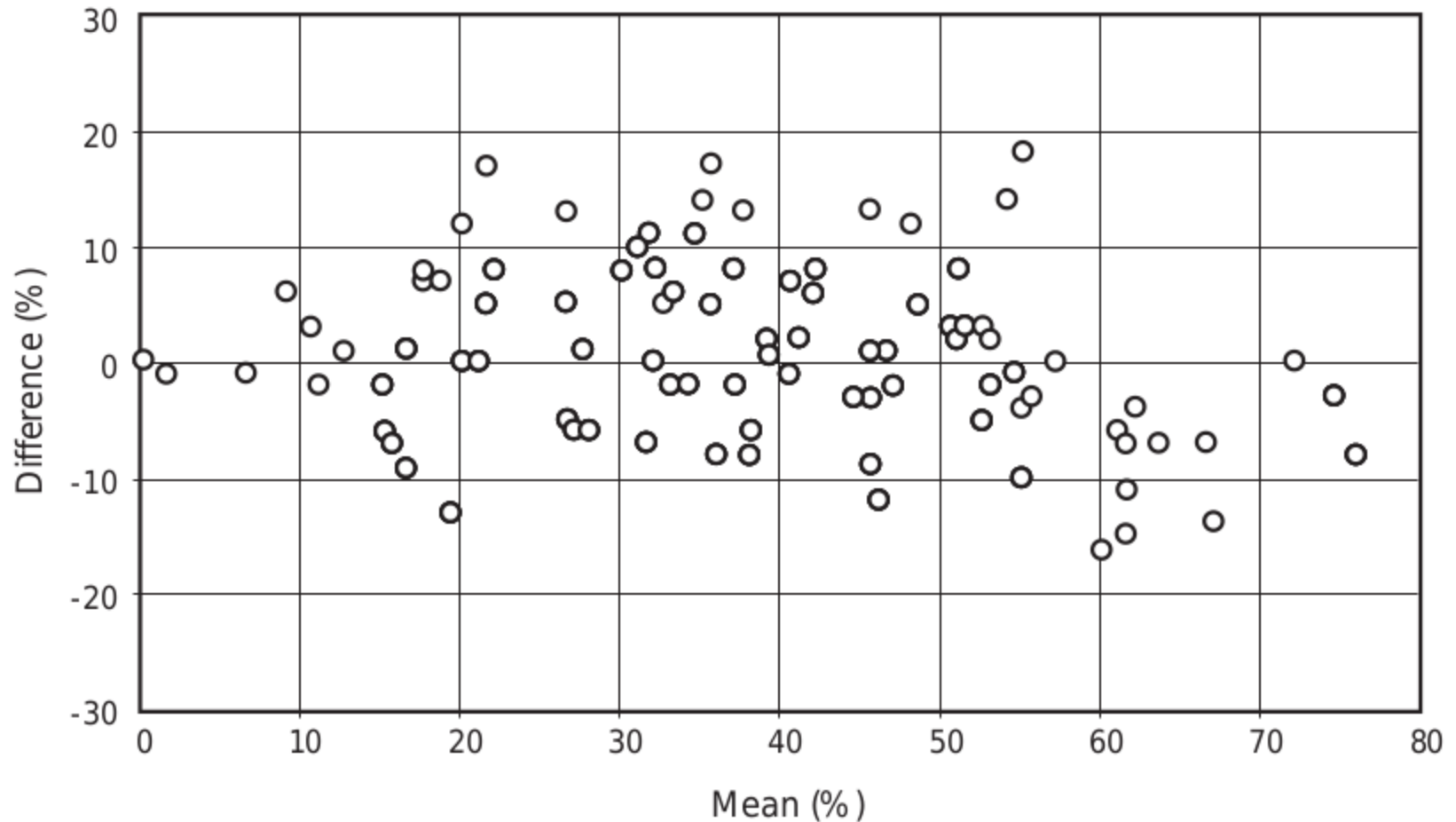
a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

Αξιολόγηση εργαστηριακών μετρήσεων

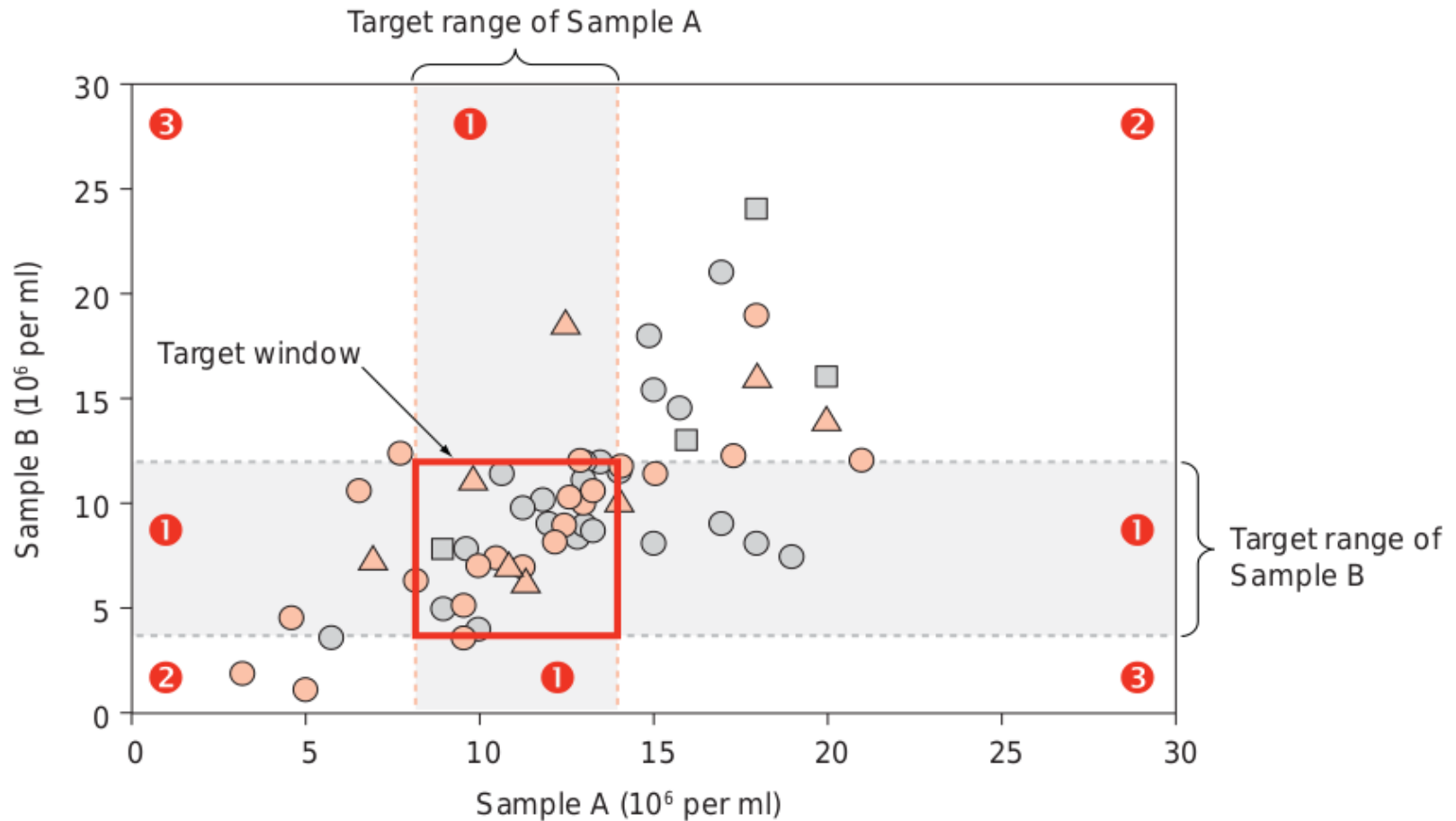


Έλεγχος 2 τεχνικών: Bland Altman plot

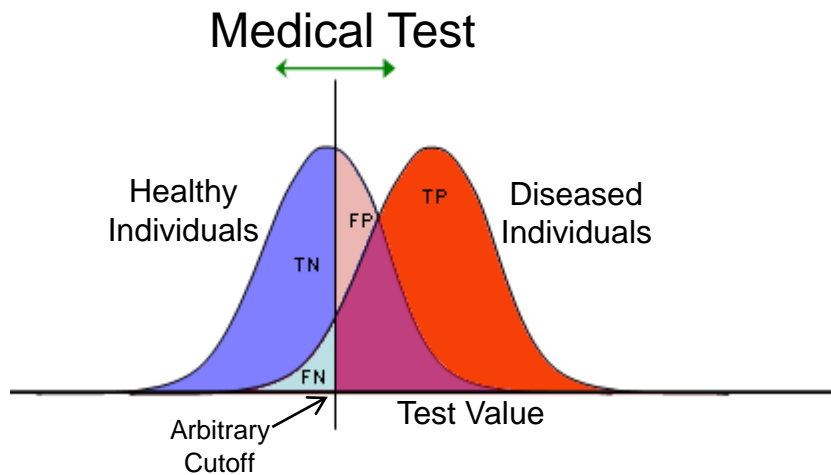


• Έλεγχος με paired t-test

Έλεγχος >2 τεχνικών: Youden plot



Καμπύλες ROC



Confusion Matrix

	Predicted +	Predicted -
Actual +	TP	FP
Actual -	FN	TN
	1	1

true positive (TP)

eqv. with hit

true negative (TN)

eqv. with correct rejection

false positive (FP)

eqv. with false alarm, Type I error

false negative (FN)

eqv. with miss, Type II error

sensitivity or true positive rate (TPR)

eqv. with hit rate, recall

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$

specificity (SPC) or true negative rate (TNR)

$$SPC = \frac{TN}{N} = \frac{TN}{FP + TN}$$

precision or positive predictive value (PPV)

$$PPV = \frac{TP}{TP + FP}$$

negative predictive value (NPV)

$$NPV = \frac{TN}{TN + FN}$$

fall-out or false positive rate (FPR)

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} = 1 - SPC$$

false discovery rate (FDR)

$$FDR = \frac{FP}{FP + TP} = 1 - PPV$$

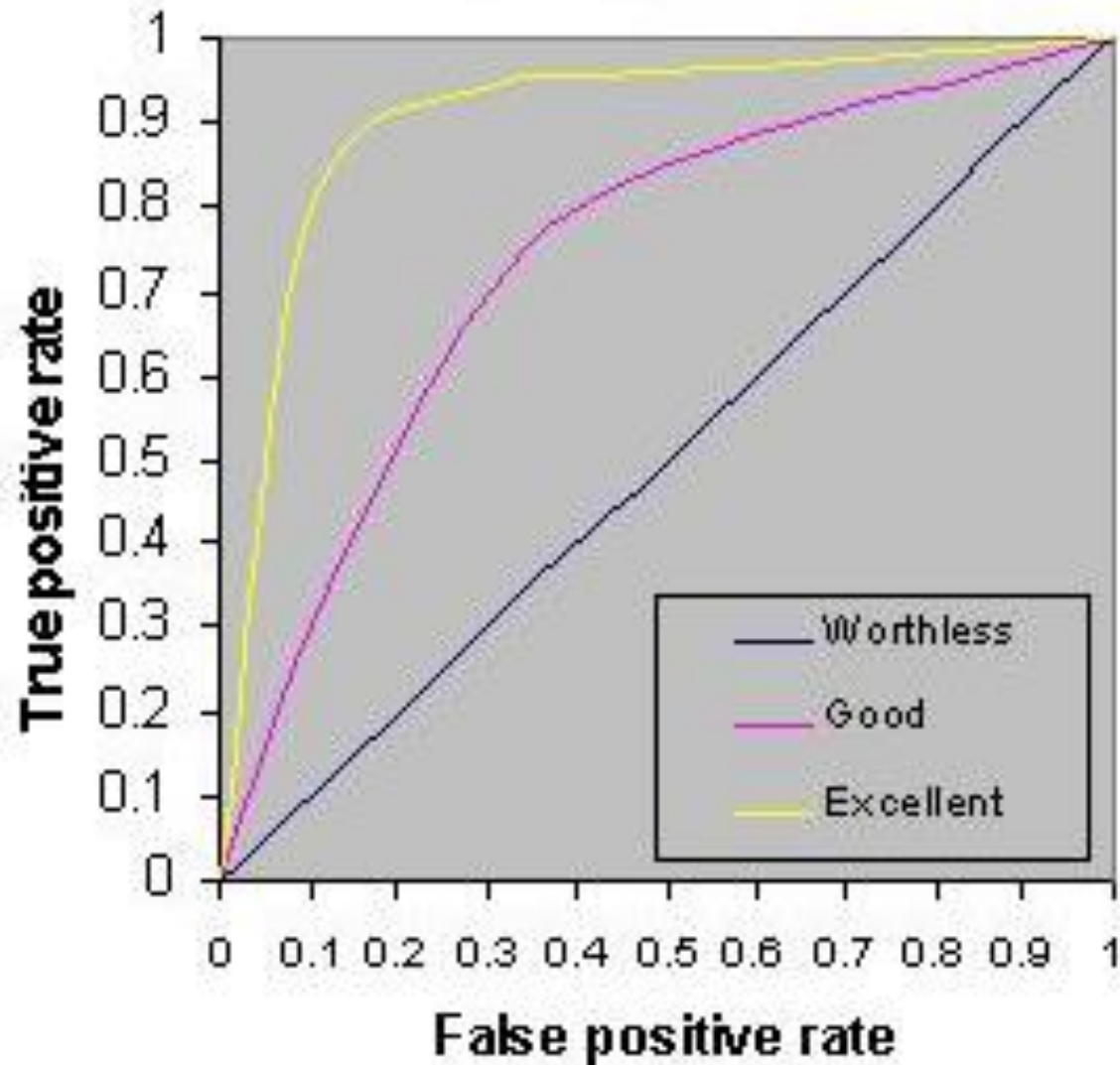
miss rate or false negative rate (FNR)

$$FNR = \frac{FN}{P} = \frac{FN}{FN + TP}$$

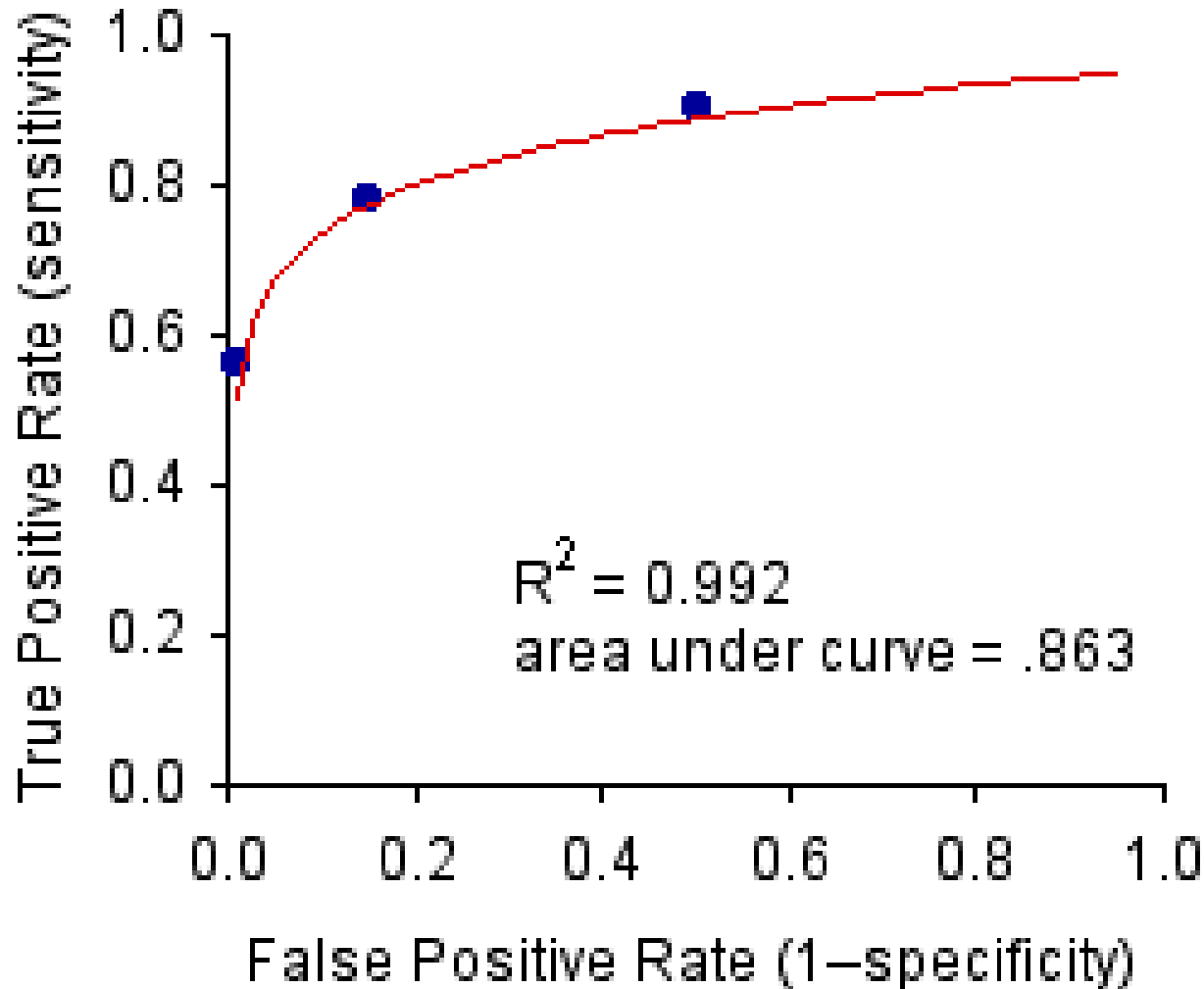
accuracy (ACC)

$$ACC = \frac{TP + TN}{P + N}$$

Καμπύλες ROC



Παράδειγμα: Καμπύλη ROC



Σας ευχαριστώ για την προσοχή σας!



Για περισσότερες πληροφορίες: ecritselis@bioacademy.gr



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΙΑΤΡΙΚΗ ΣΧΟΛΗ

