Έλεγχοι Υποθέσεων Δύο Δειγμάτων

Κωνσταντίνος Μπουγιούκας, MSc, PhD
mpougioukas@auth.gr

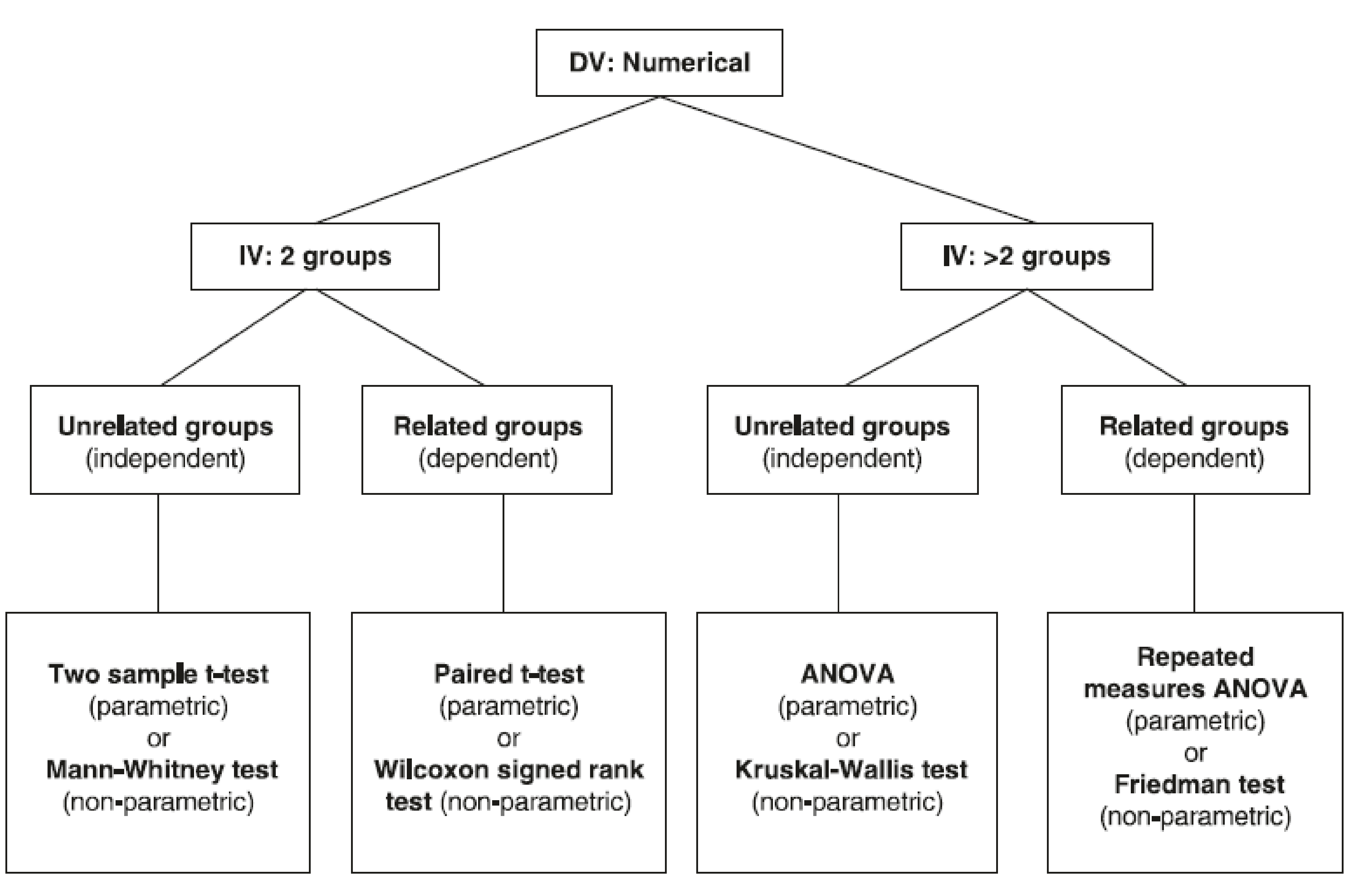
2024-2025

Στόχοι του σημερινού μαθήματος

- ☑ Εφαρμογή του ελέγχου υποθέσεων
- ☑ Έλεγχος δύο ανεξάρτητων δειγμάτων (Student's t-test, Welch's t-test, Mann-Whitney U test)
- ☑ Έλεγχος δυο κατά ζεύγη εξαρτημένων δειγμάτων (Paired t-test, Signed-Ranks Wilcoxon test)

Έλεγχος υποθέσεων-Βήματα

1. Καθορίζεται η **μηδενική υπόθεση H0 (=)** και **εναλλακτική υπόθεση H1 (≠) .**
2. Ορίζεται το **επίπεδο σημαντικότητας** α (συνήθως α=0.05).
3. Επιλέγεται μια κατάλληλη **στατιστική δοκιμασία** και υπολογίζεται η τιμή του στατιστικού με βάση τα δεδομένα του δείγματος.
4. Σύγκριση της **πιθανότητας p** να έχουμε την συγκεκριμένη τιμή του στατιστικού (ή κάτι πιο ακραίο) θεωρώντας ότι ισχύει η Ho, με το **επίπεδο σημαντικότητας α** (0.05). Εάν p <0.05, απόρριψη της Ho (στατιστικά σημαντικό αποτέλεσμα).
5. **Ερμηνεία** αποτελεσμάτων.


DV: Numerical
IV: 2 groups
IV: >2 groups
Unrelated groups (independent)
Related groups (dependent)
Unrelated groups (independent)
Related groups (dependent)
Two sample t-test (parametric) or Mann-Whitney test (non-parametric)
Paired t-test (parametric) or Wilcoxon signed rank test (non-parametric)
ANOVA (parametric) or Kruskal-Wallis test (non-parametric)
Repeated measures ANOVA (parametric) or Friedman test (non-parametric)

Student's t-test (ανεξάρτητα δείγματα, παραμετρικός έλεγχος)

Χρησιμοποιείται όταν θέλουμε να συγκρίνουμε τις μέσες τιμές δύο ανεξάρτητων δειγμάτων X1 και X2 μιας συνεχής μεταβλητής X.

Προϋποθέσεις

- Τα τυχαία δείγματα θα πρέπει να είναι **ανεξάρτητα**.
- Τα δείγματα να προέρχονται από **κανονικές κατανομές**.
- **Ισότητα** των πληθυσμιακών **διακυμάνσεων** (άγνωστες αλλά ίσες, $\sigma_1^2=\sigma_2^2$).

Προϋποθέσεις: Έλεγχος κανονικότητας (normality test)

Shapiro-Wilk test

Ho: τα δεδομένα προέρχονται από κανονική κατανομή.

Vs

H1: τα δεδομένα **δεν** προέρχονται από κανονική κατανομή.

Αν $p \geq \alpha$ ➡ **δεν** απορρίπτεται η Ho. Τα δεδομένα προέρχονται από κανονική κατανομή.

Προϋποθέσεις: Έλεγχος ισότητας διακυμάνσεων (equality of variances)

Levene’s test

Ho: Οι διακυμάνσεις των συγκρινόμενων ομάδων είναι ίσες. ($\sigma_1^2 = \sigma_2^2$)

Vs

H1: Οι διακυμάνσεις των συγκρινόμενων ομάδων **δεν** είναι ίσες. (δηλ. είναι άνισες, $\sigma_1^2 \neq \sigma_2^2$)

Αν $p \geq \alpha$ ➡ **δεν** απορρίπτεται η Ho. Οι διακυμάνσεις (variances) είναι ίσες.

Έλεγχος της διαφοράς δύο μέσων τιμών (διακυμάνσεις άγνωστες αλλά ίσες, σ1 = σ2)

Student's t-test

Ho: Οι μέσες τιμές των δύο ομάδων **δεν** διαφέρουν (δηλ. είναι ίσες).

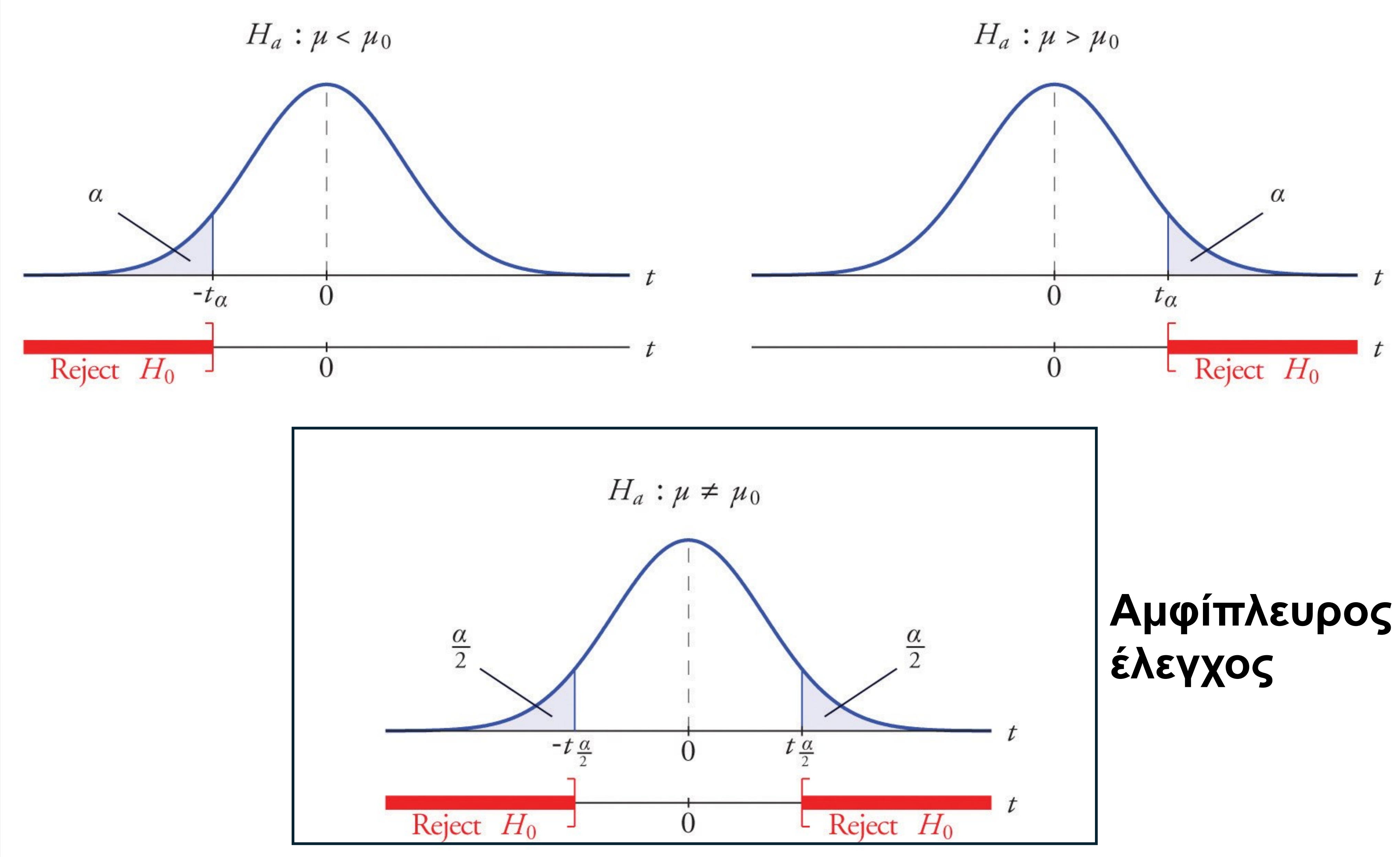
($\mu_1 = \mu_2$ ή $\mu_1 - \mu_2 = 0$)

Vs

H1: Οι μέσες τιμές των δύο ομάδων διαφέρουν.

($\mu_1 \neq \mu_2$ ή $\mu_1 - \mu_2 \neq 0$)

Αν $p < \alpha$ ➡ **απορρίπτεται** η Ho. Οι δύο μέσες τιμές **διαφέρουν**. Στατιστικά σημαντικό αποτέλεσμα.

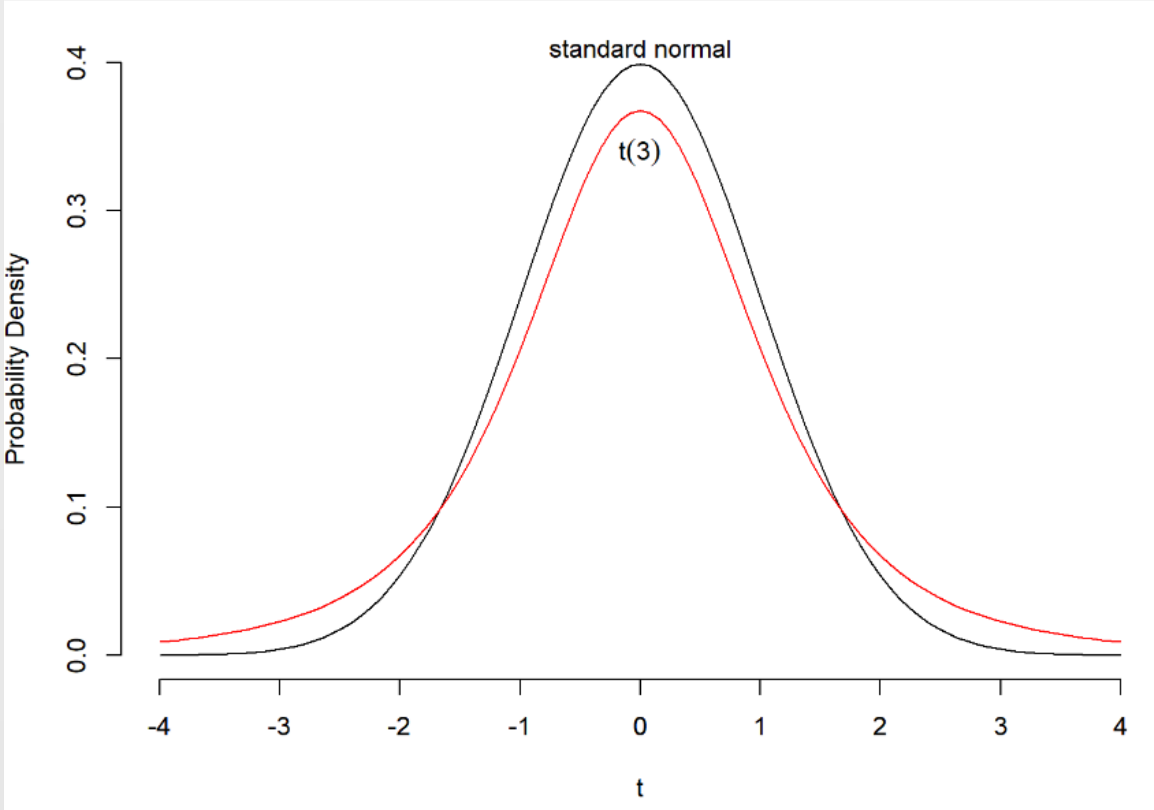


Αμφίπλευρος έλεγχος

- **Το στατιστικό t:**

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

όπου s_p η συνολική τυπική απόκλιση της διαφοράς



- **Βαθμοί ελευθερίας:** n_1+n_2-2

- **Το διάστημα εμπιστοσύνης της διαφοράς των δυο μέσων τιμών είναι:**

$$(1-\alpha)100\%\ CI:\ (\bar{x}_1 - \bar{x}_2) \pm t_{\nu,1-\alpha/2} * SE_{\bar{x}_1 - \bar{x}_2}$$

Π.χ. αν βρεθεί 95%ΔΕ: (-2.8, 0.47) σημαίνει ότι το αποτέλεσμα **δεν** είναι στατιστικά σημαντικό επειδή περιλαμβάνεται το **μηδέν** (δηλ. η τιμή της Ho).

Συνοπτικά βήματα επιλογής ελέγχου:

1. Καθορισμός μηδενικής και εναλλακτικής υπόθεσης
   - **H_0: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 \neq \mu_2$ or $\mu_1 - \mu_2 \neq 0$**

2. Έλεγχος κανονικότητας της ποσοτικής μεταβλητής Y και για τις δύο ομάδες.
   - Εάν η κατανομή είναι κανονική και στις δύο ομάδες τότε εφαρμόζω T-test:
     - Ελέγχω για ισότητα διακυμάνσεων (Levene's test)
       - Εάν το Levene's test $p>0.05$, τότε εφαρμόζω Student's t-test
       - Εάν το Levene's test $p<0.05$ τότε εφαρμόζω Welch's t-test
   - Εάν δεν πληρείται η προϋπόθεση της κανονικότητας στις δύο ομάδες τότε εφαρμόζω Mann-Whitney U test (μη παραμετρικός έλεγχος)
     - H_0: $md_1 = md_2$ H_1: $md_1 \neq md_2$

Paired samples t-test (εξαρτημένα κατά ζεύγη δείγματα, παραμετρική δοκιμασία)

Το Paired t-test μπορεί να χρησιμοποιηθεί εάν έχουμε δύο δείγματα που **συσχετίζονται** μεταξύ τους και μία ποσοτική μεταβλητή ενδιαφέροντος. (Προϋπόθεση η κανονική κατανομή των **διαφορών**)

Παραδείγματα:

- Μετρήσεις που συλλέχθηκαν πριν και μετά από μια παρέμβαση σε μια πειραματική μελέτη (repeated measures designs).

- Δίδυμα αδέρφια, σύζυγοι, ζευγαρωμένα όργανα όπως τα μάτια.

- Μια διασταυρούμενη δοκιμή (cross-over trial) στην οποία κάθε ασθενής έχει δύο μετρήσεις στη μεταβλητή, μία κατά τη λήψη θεραπείας και μία κατά τη λήψη εικονικού φαρμάκου.

Έλεγχος της μέσης τιμής των διαφορών εξαρτημένων κατά ζεύγη δειγμάτων

Paired t-test

Ho: Η μέση τιμή των διαφορών όλων των ζευγών είναι μηδέν ($\boldsymbol{\mu_\delta = 0}$)

Vs

H1: Η μέση τιμή των διαφορών όλων των ζευγών είναι διάφορη από το μηδέν ($\boldsymbol{\mu_\delta \neq 0}$)

Αν $p < α$ ➡ **απορρίπτεται** η Ho. Η μέση τιμή των διαφορών είναι διάφορη από το μηδέν. Στατιστικά σημαντικό αποτέλεσμα.

Υπολογισμός της διαφοράς

X_1	X_2	$D = X_1 - X_2$
x_{11}	x_{12}	$d_1 = x_{11} - x_{12}$
x_{21}	x_{22}	$d_2 = x_{21} - x_{22}$
x_{31}	x_{32}	$d_3 = x_{31} - x_{32}$
.	.	.
x_{n1}	x_{n2}	$d_n = x_{n1} - x_{n2}$

- **Το στατιστικό t:**

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

- **Βαθμοί ελευθερίας: n-1**

- **Το διάστημα εμπιστοσύνης της διαφοράς είναι:**

$$(1-\alpha)100\%\ CI:\ \bar{d} \pm t_{\nu;1-a/2} * SE_{\bar{d}}$$

Π.χ. αν βρεθεί 95%ΔΕ: (-52, -26) σημαίνει ότι το αποτέλεσμα είναι στατιστικά σημαντικό επειδή **δεν** περιλαμβάνεται το **μηδέν** (δηλ. η τιμή της Ho)**.**

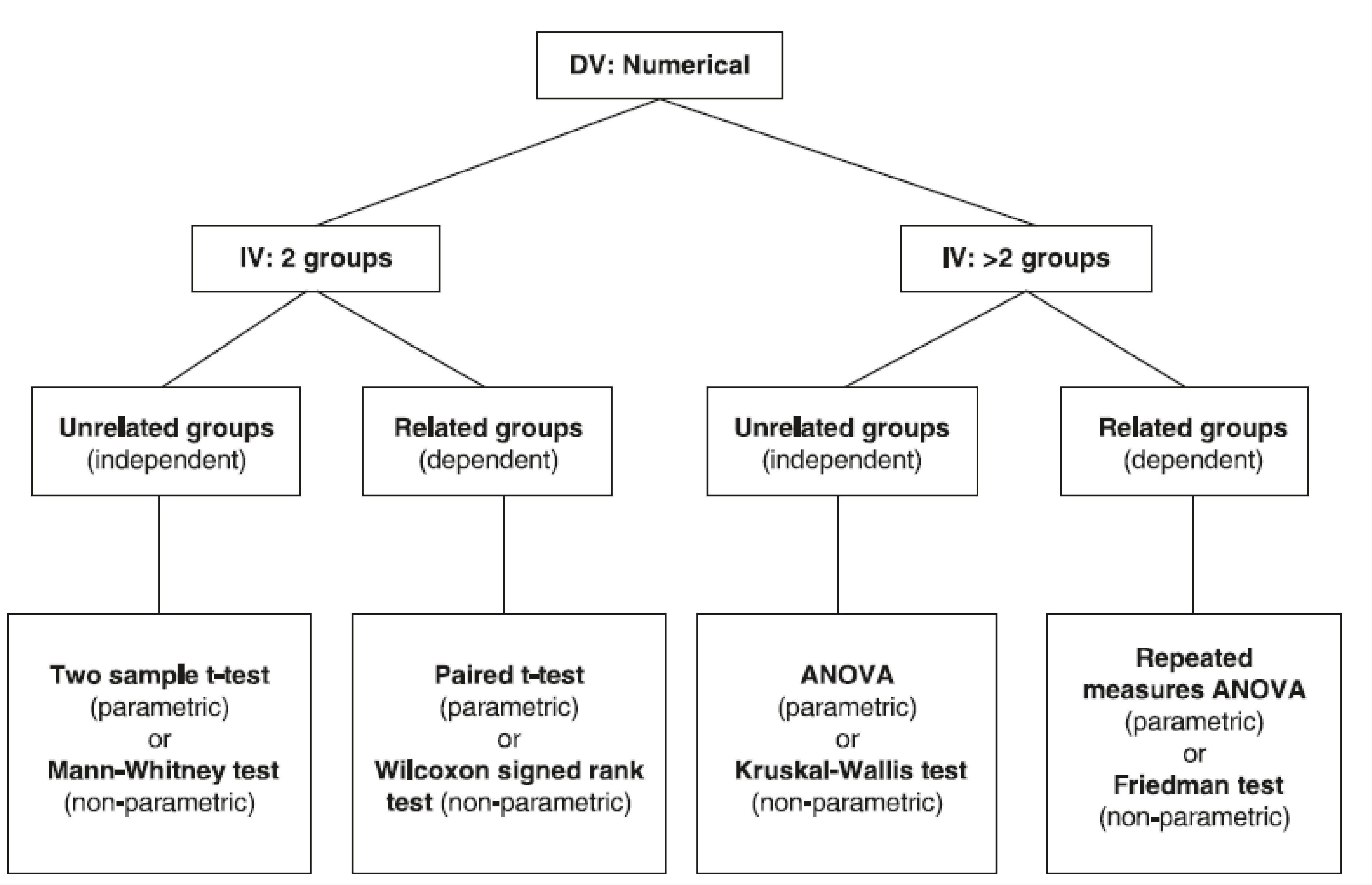
Συνοπτικά βήματα επιλογής ελέγχου:

1. Καθορισμός μηδενικής και εναλλακτικής υπόθεσης
   - **H0: μδ = 0   H1: μδ ≠ 0**

2. Υπολογισμός της διαφοράς των μετρήσεων στα δύο δείγματα και έλεγχος κανονικότητας.

- Εάν η διαφορά είναι κανονική τότε **Paired t-test**

- Εάν η διαφορά **δεν** είναι κανονική τότε **Signed-Ranks Wilcoxon test** (μη παραμετρικός έλεγχος)
  - $\mathbf{H_0}$: $md_1 = md_2$ $\mathbf{H_1}$: $md_1 \neq md_2$


DV: Numerical
IV: 2 groups
IV: >2 groups
Unrelated groups (independent)
Related groups (dependent)
Unrelated groups (independent)
Related groups (dependent)
Two sample t-test (parametric) or Mann-Whitney test (non-parametric)
Paired t-test (parametric) or Wilcoxon signed rank test (non-parametric)
ANOVA (parametric) or Kruskal-Wallis test (non-parametric)
Repeated measures ANOVA (parametric) or Friedman test (non-parametric)