



ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

Έλεγχοι Υποθέσεων Περισσοτέρων των Δύο Ανεξάρτητων Δειγμάτων

Κωνσταντίνος Ι. Μπουγιούκας, PhD

ΘΕΣΣΑΛΟΝΙΚΗ 2025-2026



Ερευνητικό ερώτημα

Έστω ότι μια έρευνα προσπαθεί να απαντήσει αν διαφέρει η απώλεια βάρους ανάλογα με τον τύπο δίαιτας A, B, C, ή D που ακολουθούν οι συμμετέχοντες.

Ανεξάρτητες ομάδες (δείγματα): η δίαιτα με τέσσερις τύπους A, B, C, και D

Μετράμε μια ποσοτική μεταβλητή: η απώλειας βάρους



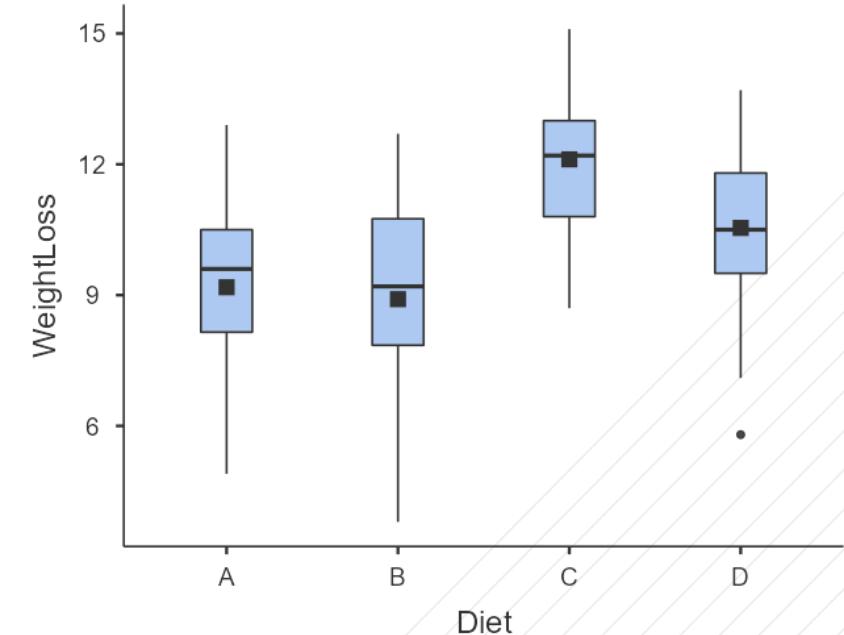
One-way Analysis of Variance (one-way ANOVA)

Για το παράδειγμά μας:

60 υπερβαρά άτομα τυχαιοποιούνται σε 4 διαφορετικούς τύπους δίαιτας A, B, C, D

Μέτρηση της απώλειας βάρους σε kg μετά από 6 μήνες

Diet	Weight Loss	Sample Mean	Sample SD
A	9.9 9.6 8.0 4.9 10.2 9.0 9.8 10.8 6.2 8.3 12.9 11.8 5.5 9.3 11.5	9.18	2.29
B	9.5 3.8 11.5 9.2 4.5 8.9 5.2 8.2 7.5 12.7 10.0 10.0 11.7 8.2 12.7	8.91	2.78
C	10.8 10.2 14.4 8.7 12.2 12.9 13.0 15.1 14.5 13.0 12.5 10.8 10.5 12.2 10.9	12.11	1.79
D	13.2 9.5 10.6 9.9 9.5 10.5 11.8 5.8 10.0 11.6 11.8 7.1 9.4 13.7 13.7	10.54	2.23



Το πρόβλημα των πολλαπλών συγκρίσεων

Αλλά για ποιο λόγο να μην εφαρμοστεί απλά η Student t δοκιμασία για όλες τις δυνατές συγκρίσεις ανά δύο;

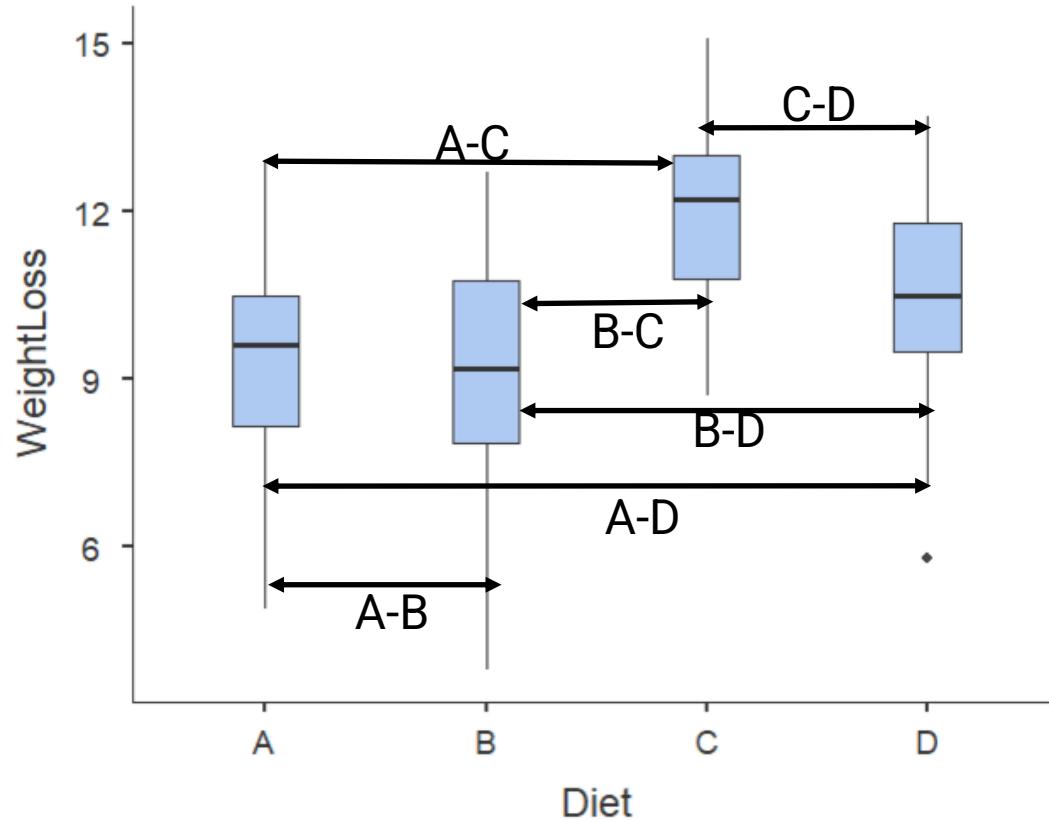
Ο λόγος είναι ότι αυξάνεται το συνολικό σφάλμα τύπου I (FWER):

$$FWER = 1 - (1 - a_{\{\text{για κάθε σύγκριση}\}})^{\text{αριθ. συγκρίσεων}}$$

Π.χ. για 6 συγκρίσεις: $FWER = 1 - (1 - 0.05)^6 = 1 - 0.74 = 0.26$

για 45 συγκρίσεις $FWER = 0.9$, δηλ. 90% πιθανότητα για ψευδώς θετικό αποτέλεσμα

(FWER: family-wise error rate)



One-way ANOVA: Υποθέσεις που ελέγχονται

Μηδενική Υπόθεση H_0 : οι μέσες τιμές της απώλειας βάρους στις τέσσερις ομάδες δίαιτας **δεν** διαφέρουν (δηλ. είναι ίσες $\mu_1 = \mu_2 = \dots = \mu_k$).

Εναλλακτική Υπόθεση H_1 : υπάρχει τουλάχιστον μια ομάδα δίαιτας με μέση τιμή απώλειας βάρους που διαφέρει από τις άλλες.
(ισοδύναμα: τουλάχιστον δυο από τις $\mu_1, \mu_2, \dots, \mu_k$ διαφέρουν μεταξύ τους)

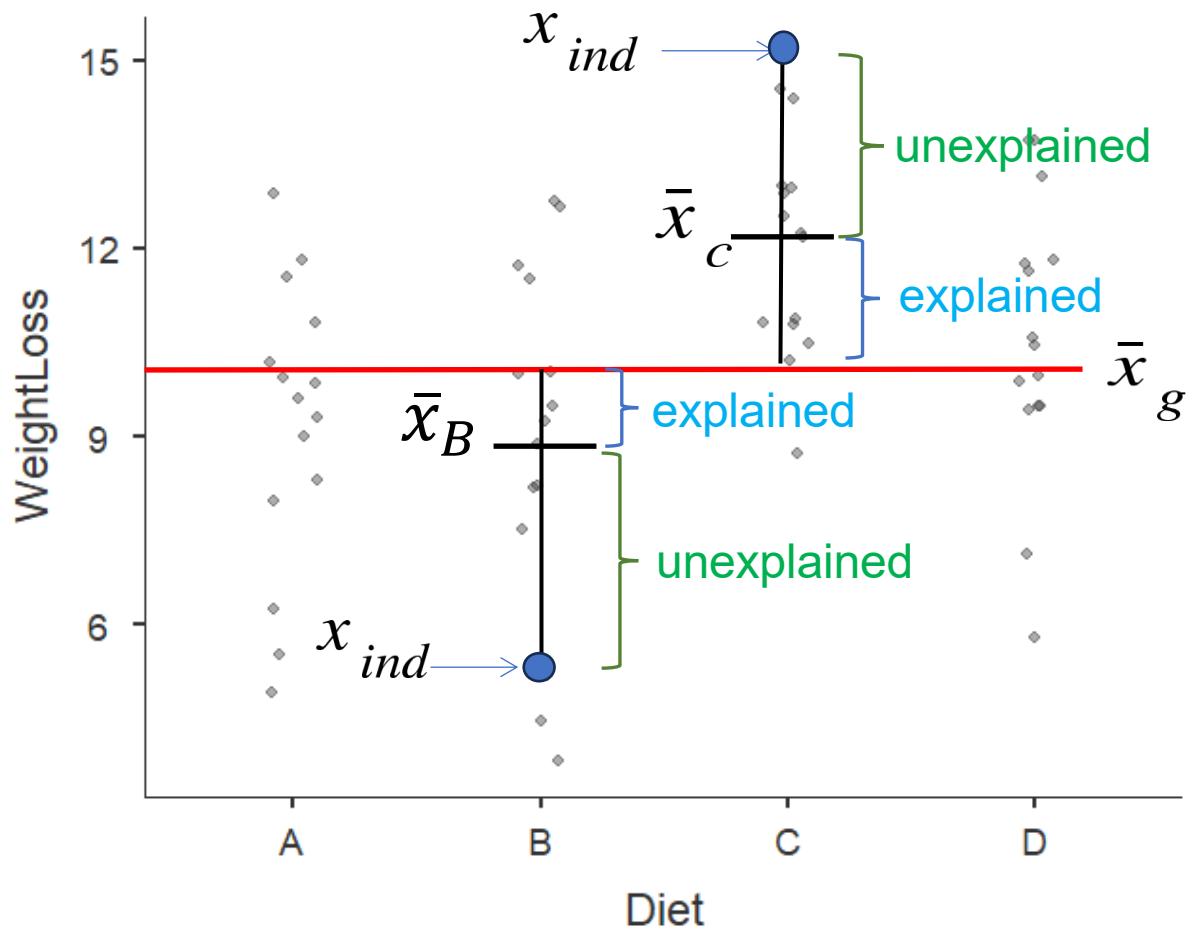


Βασικές προϋποθέσεις της one-way ANOVA

- **Ανεξάρτητα δείγματα**
- **Κανονική κατανομή** της ποσοτικής μεταβλητής (π.χ. της απώλειας βάρους) σε όλες τις ομάδες (π.χ. ομάδες δίαιτας)
 - ⇒ έλεγχος κανονικότητας (π.χ. ιστογράμματα, Shapiro-Wilk test)
- **Ισότητα των διακυμάνσεων** μεταξύ των ομάδων
 - ⇒ έλεγχος ομοιογένειας των διακυμάνσεων (Levene's test)
 - Εάν στο Levene's test, $P > 0.05$ τότε ANOVA (Fisher's)
 - Εάν Levene's test, $P < 0.05$ τότε ANOVA with corrected df (Welch's)



Γιατί δεν έχουν όλοι την ίδια απώλεια βάρους;



Οι δίαιτες μπορεί να επιδρούν διαφορετικά στα άτομα (explained by diet “X” or between group variation)

Οι άνθρωποι είναι διαφορετικοί (unexplained by diet “X” or within group variation)

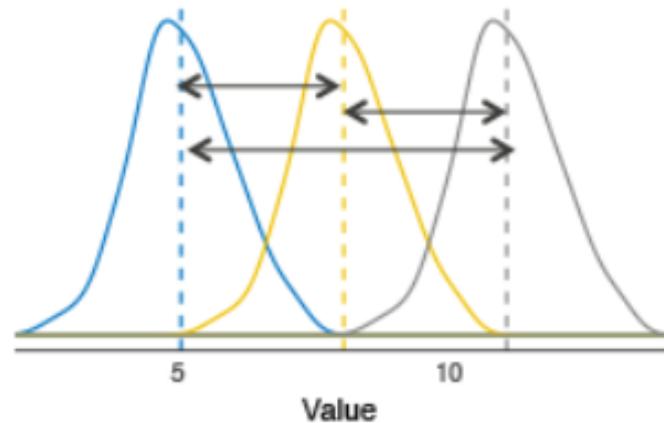
Η ANOVA χωρίζει την συνολική διακύμανση σε:

$$SS_{total} = SS_{between} + SS_{within}$$

- ⇒ Μεταξύ των ομάδων διακύμανση (Between-group variation): οφείλεται στις διαφορές μεταξύ των ομάδων

$$SS_B = \sum_{j=1}^k n_j (\bar{x}_j - \bar{x}_g)^2$$

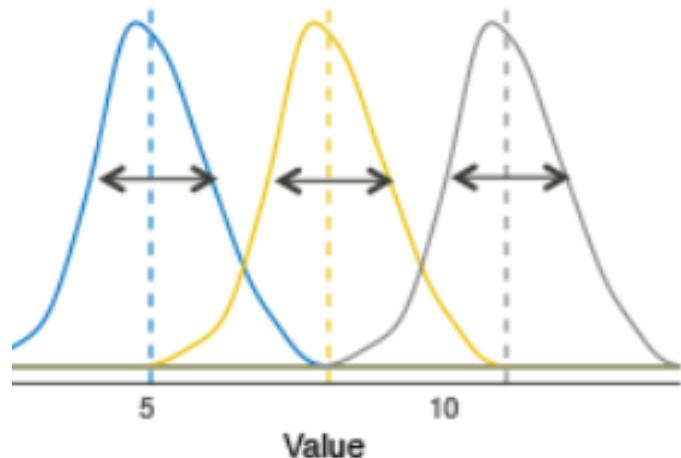
Between-group variation
(i.e. Differences among group means)



- ⇒ Εντός των ομάδων διακύμανση (Within-group variation): οφείλεται στην τυχαία διακύμανση μεταξύ των παρατηρήσεων εντός της κάθε ομάδας

$$SS_W = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

Within-group variation
(i.e. Variability within each group)



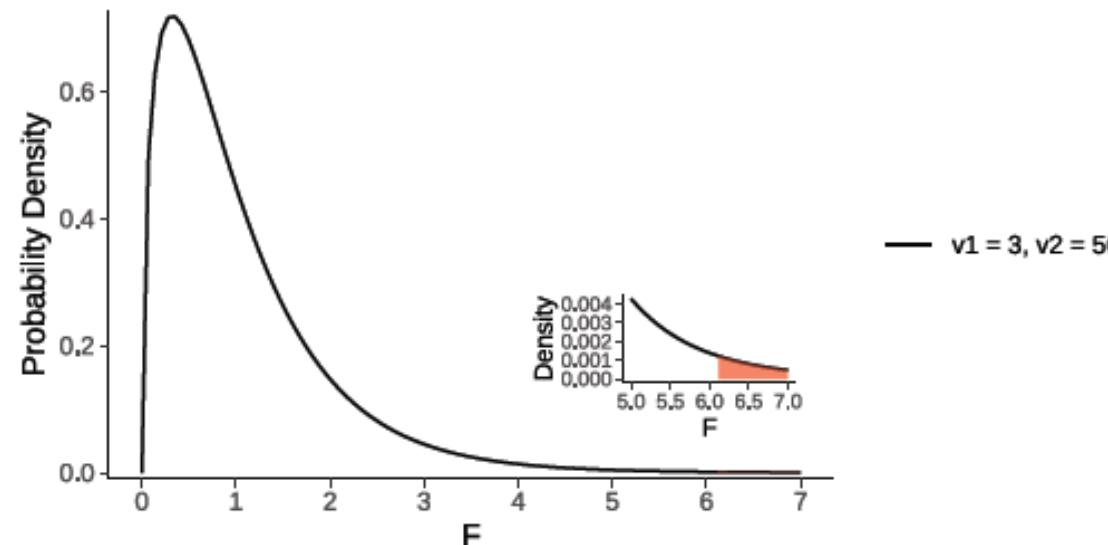
F-test

H One-Way ANOVA χρησιμοποιεί το F στατιστικό:

$$F = \frac{\text{Μεταξύ των ομάδων διακύμανση}}{\text{Εντός των ομάδων διακύμανση}}$$

To F στατιστικό ακολουθεί την F-κατανομή με $df_1=k-1=4-1=3$, $df_2 = n-k=60-4= 56$

- k: ο αριθμός των ομάδων
- n: ο συνολικός αριθμός των συμμετεχόντων στην έρευνα



Σημείωση: Εάν μόνο δύο μέσες τιμές συγκρίνονται, τότε το F-test συμπίπτει με ένα independent samples t-test και ισχύει $F = t^2$.



One-way ANOVA table

ANOVA Test Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Value
Between Groups	$SSB = \sum n_j (\bar{X}_j - \bar{X})^2$	$df_1 = k - 1$	$MSB = SSB / (k - 1)$	$f = MSB / MSW$
Within Groups	$SSW = \sum \sum (X - \bar{X}_j)^2$	$df_2 = N - k$	$MSW = SSW / (N - k)$	
Total	$SST = SSB + SSW$	$df_3 = N - 1$		



Διόρθωση για πολλαπλές συγκρίσεις (post-hoc tests)

- Tukey (όταν οι διακυμάνσεις των ομάδων είναι ίσες)
- Games-Howell (όταν οι διακυμάνσεις των ομάδων είναι άνισες)
- Bonferroni



Kruskal-Wallis test

- Μη παραμετρικός έλεγχος εναλλακτικός της one-way ANOVA (στηρίζεται στην διάταξη των μετρήσεων)
- Συνήθως χρησιμοποιείται σε **μικρά δείγματα και μη-κανονικές κατανομές**
- Η ερμηνεία των αποτελεσμάτων θα πρέπει να γίνει με βάση τις **διάμεσες τιμές** των ομάδων και το **IQR**

Διόρθωση για πολλαπλές συγκρίσεις (post-hoc tests)



Ερώτηση 1

Στην ανάλυση διακύμανσης (ANOVA) η εναλλακτική υπόθεση είναι:

- (α) Όλες οι μέσες τιμές είναι απαραίτητα άνισες μεταξύ τους
- (β) Όλες οι μέσες τιμές είναι ίσες μεταξύ τους
- (γ) Τουλάχιστον τρεις μέσες τιμές να είναι άνισες μεταξύ τους
- (δ) Τουλάχιστον μια μέση τιμή να διαφέρει από τις άλλες



Ερώτηση 2

Στην ανάλυση διακύμανσης (ANOVA) :

- (α) Πρέπει η διακύμανση μεταξύ των ομάδων να είναι ίση με την διακύμανση εντός των ομάδων
- (β) Χρησιμοποιείται ο λόγος F
- (γ) Χρησιμοποιείται η t-κατανομή
- (δ) Θα πρέπει η ποσοτική μεταβλητή να μην ακολουθεί την κανονική κατανομή στις ομάδες



Ερώτηση 3

Για δυο ανεξάρτητα δείγματα που ακολουθούν την κανονική κατανομή αν εφαρμοστεί τ δοκιμασία και ανάλυση διακύμανσης θα ισχύει:

- (α) $F = t^2$
- (β) $F = t^3$
- (γ) $F = t^2/2$
- (δ) $F = \sqrt{t}$



Ερώτηση 4

Η μη παραμετρική δοκιμασία αντίστοιχη της παραμετρικής one-way ANOVA είναι:

- (α) Η δοκιμασία Mann-Whitney
- (β) Δεν υπάρχει αντίστοιχος μη παραμετρικός έλεγχος
- (γ) Η δοκιμασία Wilcoxon
- (δ) Η δοκιμασία Kruskal-Wallis



Ερώτηση 5

Σε μια έρευνα συγκρίθηκε η αποτελεσματικότητα 3 θεραπειών με συνολικό μέγεθος δείγματος 300 άτομα. Θεωρώντας ότι πληρούνται οι προϋποθέσεις της ανάλυσης διακύμανσης (ANOVA) βρέθηκε ότι η p -τιμή του ελέγχου είναι ίση με **0.3**. Το συμπέρασμα που προκύπτει για επίπεδο σημαντικότητας $\alpha=0.05$ είναι:

- (α) Όλες οι μέσες τιμές που αφορούν την αποτελεσματικότητα των θεραπειών διαφέρουν μεταξύ τους.
- (β) Πρέπει απαραίτητα να ακολουθήσει έλεγχος post-hoc.
- (γ) Στατιστικά μη σημαντικό αποτέλεσμα, δεν απορρίπτεται η H_0 .
- (δ) Στατιστικά σημαντικό αποτέλεσμα, απορρίπτεται η H_0 .

