

Επανάληψη



Γούργουλης Βασίλειος, Ph.D.
Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

1

Τι μετρώ; → Εξαρτημένη μεταβλητή

Ποιους μετρώ; → Ανεξάρτητη μεταβλητή – ανεξάρτητος παράγοντας

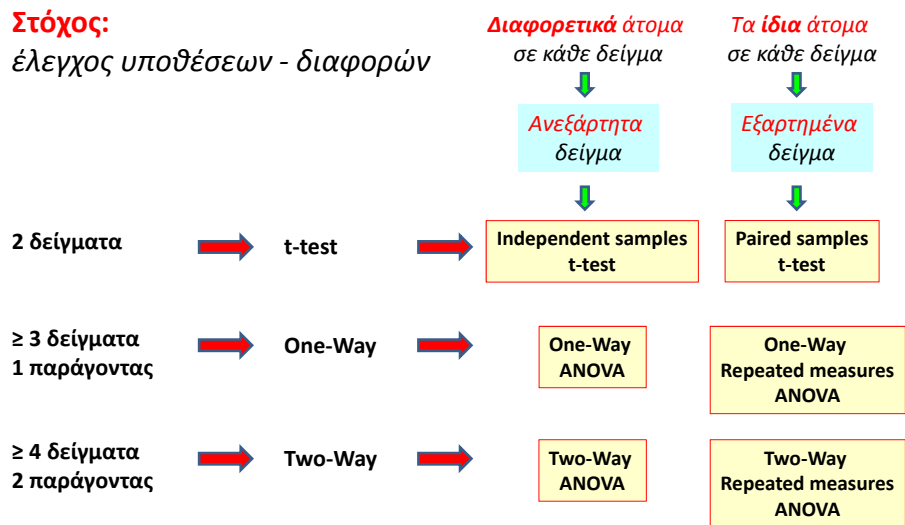
Πότε / πώς μετρώ; → Ανεξάρτητη μεταβλητή – επαναλαμβανόμενος παράγοντας

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

2

Στόχος:

έλεγχος υποθέσεων - διαφορών



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

3

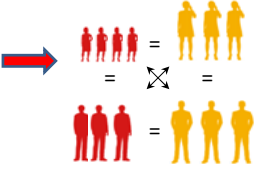
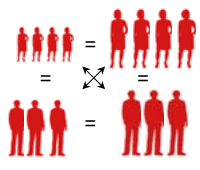
Στόχος:

Ανάλυση:



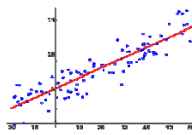
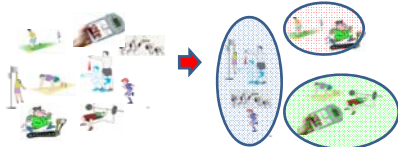
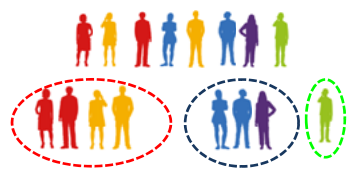
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

4

	Ανεξάρτητα δείγμα	Εξαρτημένα δείγμα
2 δείγματα <i>t-test</i>		
>3 δείγματα 1 παράγοντας <i>One - Way</i>		
>4 δείγματα 2 παράγοντας <i>Two - Way</i>		

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

5

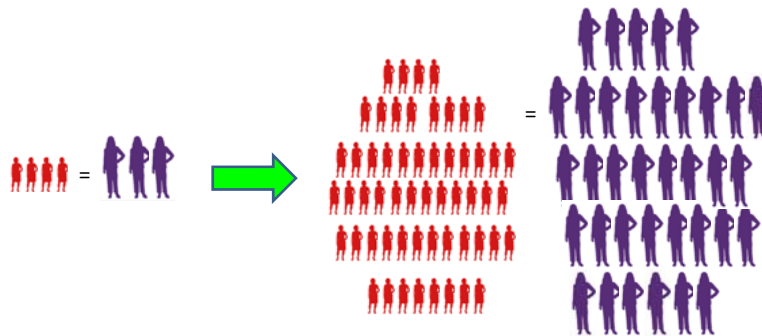
Πρόβλεψη επίδοσης – τιμών <i>Ανάλυση Παλινδρόμησης (Regression)</i>	
Ομαδοποίηση μεταβλητών <i>Παραγοντική ανάλυση (Factor)</i>	
Ομαδοποίηση ατόμων <i>Ανάλυση Ταξινόμησης (Cluster)</i>	
Διάκριση ατόμων σε ομάδες <i>Ανάλυση Διακριτότητας (Discriminant)</i>	

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

6

Αναλύσεις ελέγχου υποθέσεων

t – test, ANOVA, repeated measures ANOVA



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

7

Τι μετρώ; → Εξαρτημένη μεταβλητή

Ποιους μετρώ; → Ανεξάρτητη μεταβλητή – ανεξάρτητος παράγοντας

Πότε / πώς μετρώ; → Ανεξάρτητη μεταβλητή – επαναλαμβανόμενος παράγοντας

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

8

Ποιος είναι ο στόχος
των αναλύσεων ελέγχου υποθέσεων
(t – test, αναλύσεις διακύμανσης);



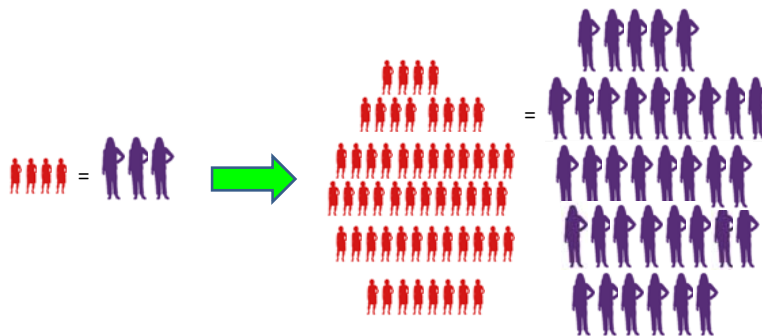
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

9

Ποιος είναι ο στόχος
των αναλύσεων ελέγχου υποθέσεων
(t – test, αναλύσεις διακύμανσης);



Ο **έλεγχος** της ύπαρξης στατιστικά **σημαντικών διαφορών** -
Εξαγωγή συμπερασμάτων **από το δείγμα για τον πληθυσμό**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

10

Πως διατυπώνεται (περιφραστικά και μαθηματικά)
μια μηδενική υπόθεση;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

11

Πως διατυπώνεται (περιφραστικά και μαθηματικά)
μια μηδενική υπόθεση;



- **ΔΕΝ υπάρχει** στατιστικά σημαντική διαφορά μεταξύ των μέσων όρων ...
- Οι μέσοι όροι ... είναι **ίσοι**
- **H₀: $\mu_1 = \mu_2 = \dots$**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

12

Πως γίνεται ο έλεγχος (αποδοχής ή απόρριψης)
μιας μηδενικής υπόθεσης;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

13

Πως γίνεται ο έλεγχος (αποδοχής ή απόρριψης)
μιας μηδενικής υπόθεσης;



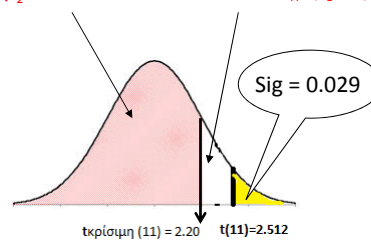
Συγκρίνοντας το εκάστοτε «στατιστικό» (t – τιμή, F – τιμή) με μια «**κρίσιμη**» τιμή

Περιοχή **αποδοχής**
της μηδενικής υπόθεσης

$$H_0 : \mu_1 = \mu_2$$

Περιοχή **απόρριψης**
της μηδενικής υπόθεσης

$$H_A : \mu_1 \neq \mu_2$$



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

14

**Πως εντοπίζονται
οι στατιστικά σημαντικές διαφορές
μεταξύ επιμέρους δειγμάτων**
(όταν υπάρχει κύρια επίδραση ενός παράγοντα);



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

15

**Πως εντοπίζονται
οι στατιστικά σημαντικές διαφορές
μεταξύ επιμέρους δειγμάτων**
(όταν υπάρχει κύρια επίδραση ενός παράγοντα);



Μέσω ενός τεστ «**πολλαπλών συγκρίσεων**» (post hoc test)

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

16

Τι σημαίνει
η ύπαρξη «στατιστικά σημαντικής αλληλεπίδρασης»
μεταξύ δύο παραγόντων ;



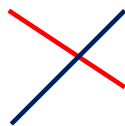
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

17

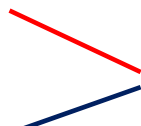
Τι σημαίνει
η ύπαρξη «στατιστικά σημαντικής αλληλεπίδρασης»
μεταξύ δύο παραγόντων ;



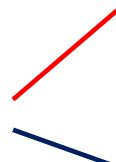
Ότι η επίδραση του ενός παράγοντα
δεν είναι σταθερή (δεν είναι η ίδια)
για όλες τις βαθμίδες του άλλου παράγοντα



Υπάρχει



Υπάρχει



Υπάρχει



ΔΕΝ υπάρχει



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

18

**Τι ΔΕΝ πρέπει
και τι πρέπει να γίνει**



όταν υπάρχει «στατιστικά σημαντική αλληλεπίδραση» ;

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

19

**Τι ΔΕΝ πρέπει
και τι πρέπει να γίνει**

όταν υπάρχει «στατιστικά σημαντική αλληλεπίδραση» ;



Δεν πρέπει

να λαμβάνονται υπόψη οι **κύριες επιδράσεις** (main effect) των παραγόντων



Πρέπει

να **αναλύεται η αλληλεπίδραση**

και να ελέγχεται

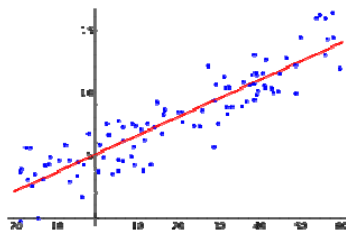
η επίδραση του κάθε παράγοντα σε κάθε βαθμίδα του άλλου παράγοντα (simple effect)

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

20

Ανάλυση Παλινδρόμησης

Regression

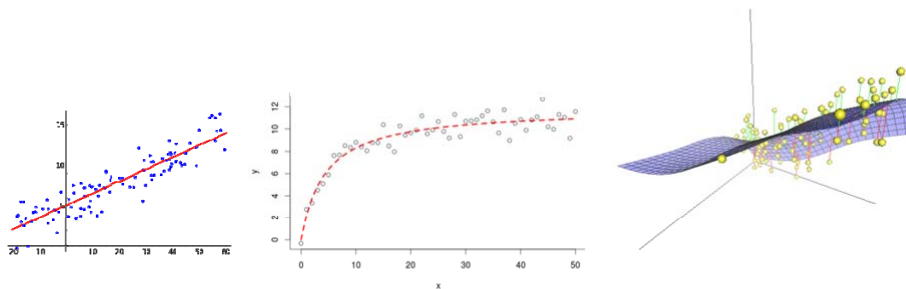


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

21

Ανάλυση Παλινδρόμησης

Regression



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

22

Ποιος είναι ο στόχος της ανάλυσης παλινδρόμησης;



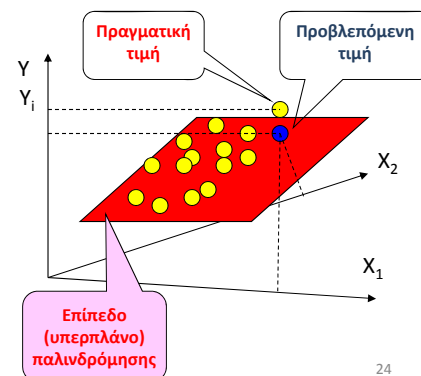
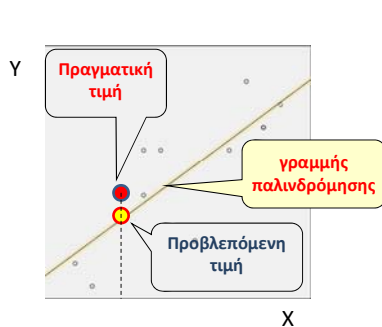
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

23

Ποιος είναι ο στόχος της ανάλυσης παλινδρόμησης;



Η **πρόβλεψη** των τιμών μιας εξαρτημένης μεταβλητής
από τις τιμές **μιας ανεξάρτητης μεταβλητής** (απλή παλινδρόμηση)
ή από τις τιμές **πολλών ανεξάρτητων μεταβλητών** (πολλαπλή παλινδρόμηση)



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

24

Ποια ευθεία (γραμμή) παλινδρόμησης ή επίπεδο (υπερπλάνο) παλινδρόμησης εκφράζει καλύτερα (με μεγαλύτερη ακρίβεια) τα δεδομένα;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

25

Ποια ευθεία (γραμμή) παλινδρόμησης ή επίπεδο (υπερπλάνο) παλινδρόμησης εκφράζει καλύτερα (με μεγαλύτερη ακρίβεια) τα δεδομένα;



Η γραμμή (ή το υπερπλάνο) βάσει της/του οποίας/ου το άθροισμα των τετραγώνων της διαφοράς ΟΛΩΝ των πραγματικών από τις προβλεπόμενες τιμές είναι το **ελάχιστο δυνατό**.

$$\sum (Y - \hat{Y})^2 = \text{ελάχιστο}$$

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

26

Πώς καθορίζεται η διεύθυνση της γραμμής παλινδρόμησης;



Πώς καθορίζεται η διεύθυνση του υπερπλάνου παλινδρόμησης;

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

27

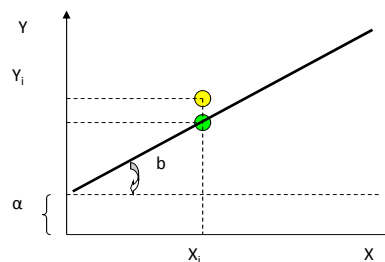
Πώς καθορίζεται η διεύθυνση της γραμμής παλινδρόμησης;



Καθορίζοντας τους συντελεστές παλινδρόμησης (a , b)

$$Y = a + b \cdot X_1 + e$$

$$\hat{Y} = a + b \cdot X_1$$



Πώς καθορίζεται η διεύθυνση του υπερπλάνου παλινδρόμησης;



Καθορίζοντας τους συντελεστές παλινδρόμησης (a , b_1 , b_2 , ..., b_i)

$$\hat{Y} = a + b_1 \cdot X_1 + b_2 \cdot X_2 + \dots + b_i \cdot X_i$$

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

28

Πότε υπάρχει πραγματικά σχέση
μεταξύ της εξαρτημένης (Y) και της ανεξάρτητης (X) (μεταβλητής
ή των ανεξάρτητων μεταβλητών);



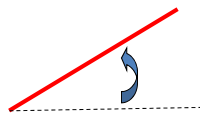
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

29

Πότε υπάρχει πραγματικά σχέση
μεταξύ της εξαρτημένης (Y) και της ανεξάρτητης (X) (μεταβλητής
ή των ανεξάρτητων μεταβλητών);



Όταν ο αντίστοιχος **συντελεστής B** (Unstandardized Coefficient)
(που π.χ. στην γραμμική παλινδρόμηση
δηλώνει την κλίση της γραμμής παλινδρόμησης)
είναι στατιστικά σημαντικά **διαφορετικός από το 0**.



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

30

Πως καθορίζεται το ποσοστό της διακύμανσης της εξαρτημένης μεταβλητής (Y), που ερμηνεύεται από την ανεξάρτητη μεταβλητή X (ή τις ανεξάρτητες μεταβλητές Xi);



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

31

Πως καθορίζεται το ποσοστό της διακύμανσης της εξαρτημένης μεταβλητής (Y), που ερμηνεύεται από την ανεξάρτητη μεταβλητή X (ή τις ανεξάρτητες μεταβλητές Xi);



Βάσει του **συντελεστή προσδιορισμού R^2** (coefficient of determination)

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

32

Πως καθορίζεται ποιού βαθμού πολυώνυμο (στη μη-γραμμική παλινδρόμηση)
εκφράζει καλύτερα τη σχέση
μεταξύ της εξαρτημένης μεταβλητής Y και της ανεξάρτητης μεταβλητής X;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

33

Πως καθορίζεται ποιού βαθμού πολυώνυμο (στη μη-γραμμική παλινδρόμηση)
εκφράζει καλύτερα τη σχέση
μεταξύ της εξαρτημένης μεταβλητής Y και της ανεξάρτητης μεταβλητής X;



Βάσει του μερικού συντελεστή προσδιορισμού (π.χ. 2^{ου} βαθμού)
(ή του R Square Change),
που δηλώνει **το ποσοστό της ανεξήγητης διακύμανσης** της Y,
όταν χρησιμοποιείται η γραμμική σχέση ($1-R^2_{Y,X}$),
το οποίο πλέον εξηγείται με τη χρήση της μη γραμμικής σχέσης ($R^2_{Y,X, X^2} - R^2_{Y,X}$)
(δηλαδή
το ποσοστό της διακύμανσης **που ερμηνεύεται επιπλέον**
από το μη-γραμμικό μοντέλο (π.χ. 2^{ου} βαθμού),
σε σχέση με το αμέσως μικρότερο (π.χ. 1^{ου} βαθμού – γραμμικό μοντέλο).

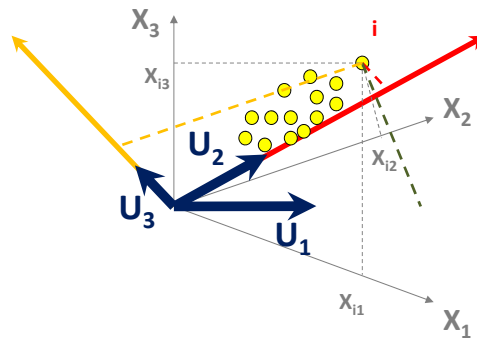
$$R^2_{Y,X,X^2} - R^2_{Y,X}$$

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

34

Παραγοντική Ανάλυση

Factor

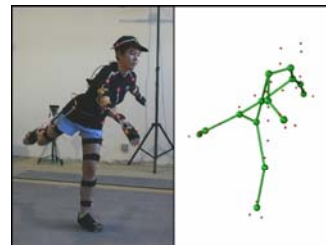
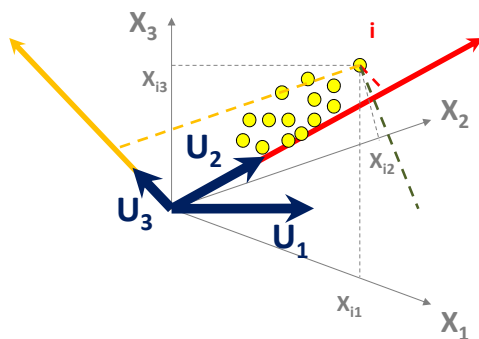


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

35

Παραγοντική Ανάλυση

Factor



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

36

Ποιος είναι ο στόχος της παραγοντικής ανάλυσης;



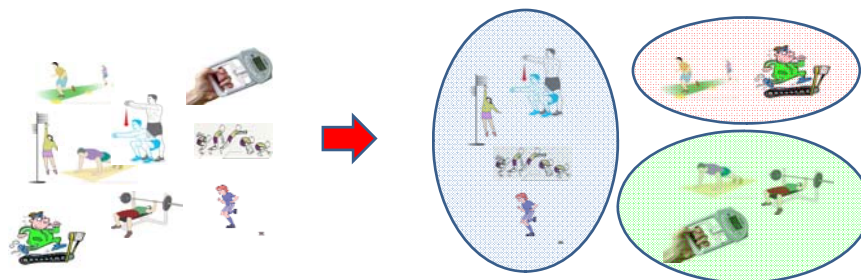
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

37

Ποιος είναι ο στόχος της παραγοντικής ανάλυσης;



Να **ομαδοποιηθούν** οι αρχικές μεταβλητές και να καταδειχθεί η **σχέση** που υπάρχει μεταξύ τους, αναδεικνύοντας τις λανθάνουσες μεταβλητές – **παράγοντες** που τις εκφράζουν με τον καλύτερο τρόπο.



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

38

Πότε έχει νόημα η εφαρμογή της παραγοντικής ανάλυσης;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

39

Πότε έχει νόημα η εφαρμογή της παραγοντικής ανάλυσης;



Όταν οι αρχικές μεταβλητές **συσχετίζονται** μεταξύ τους .



υψηλοί συντελεστές συσχέτισης



KMO > 0.70 = καλές



Bartlett's Test of Sphericity Sig < 0.05



Partial correlation coefficients = μικρά (κοντά στο 0)



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

40

Πως ορίζεται η διεύθυνση ενός παραγοντικού άξονα;



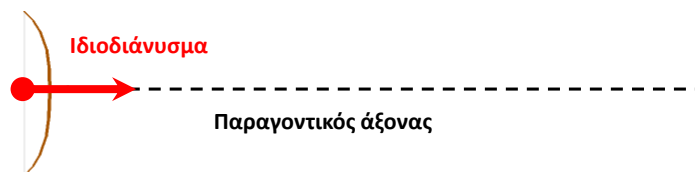
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

41

Πως ορίζεται η διεύθυνση ενός παραγοντικού άξονα;



Από το **ιδιοδιάνυσμά** του.



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

42

Ποιο είναι το μέτρο (μήκος) του ιδιοδιανύσματος ενός παραγοντικού άξονα και τι εκφράζει;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

43

Ποιο είναι το μέτρο (μήκος) του ιδιοδιανύσματος ενός παραγοντικού άξονα και τι εκφράζει;



Είναι η **ιδιοτιμή** του
και **εκφράζει** το ποσό της διακύμανσης όλων των μεταβλητών
που ερμηνεύεται από το παραγοντικό άξονα

Ιδιοτιμή (το μήκος του ιδιοδιανύσματος)



Παραγοντικός άξονας

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

44

Ποιος παραγοντικός άξονας ερμηνεύει
το μεγαλύτερο ποσό της διακύμανσης όλων των μεταβλητών ;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

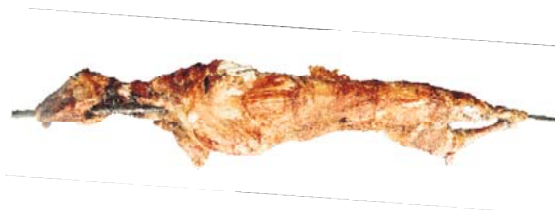
45

Ποιος παραγοντικός άξονας ερμηνεύει
το μεγαλύτερο ποσό της διακύμανσης όλων των μεταβλητών ;



Ο πρώτος (1^{ος}) παραγοντικός άξονας

1^{ος} παραγοντικός άξονας



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

46



Τι επιτυγχάνεται με την περιστροφή των παραγοντικών αξόνων;

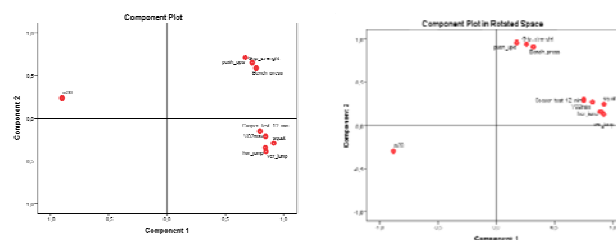
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

47

Τι επιτυγχάνεται με την περιστροφή των παραγοντικών αξόνων;



Μετά την περιστροφή
αλλάζει η σχετική θέση των παραγοντικών αξόνων ως προς τις μεταβλητές
οπότε
αλλάζει η συσχέτιση της κάθε μεταβλητής με τον κάθε παραγοντικό άξονα
έτσι ώστε **να συσχετίζεται κυρίως με έναν μόνο παραγοντικό άξονα**



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

48

Πως καθορίζονται – «ονοματίζονται» οι παράγοντες ;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

49

Πως καθορίζονται – «ονοματίζονται» οι παράγοντες ;



Βάσει των μεταβλητών που «φορτίζονται» - συσχετίζονται ξεκάθαρα με τον συγκεκριμένο παράγοντα.

Rotated Component Matrix^a

	Component		
	1	2	3
vertical jump (cm)	.955		
horizontal jump (m)	.896		
squat (number)	.848		
30 m sprint (sec)	-.814		
push ups (number)		.954	
Grip Strength (kg)		.940	
Bench Press (kg)		.908	
Cooper Test 12-min (m)			.904
VO2max (ml/min/kg)			.860

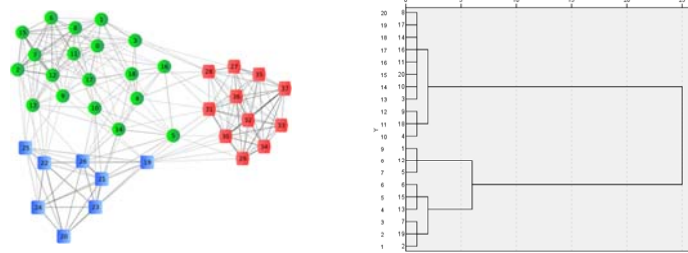
Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

50

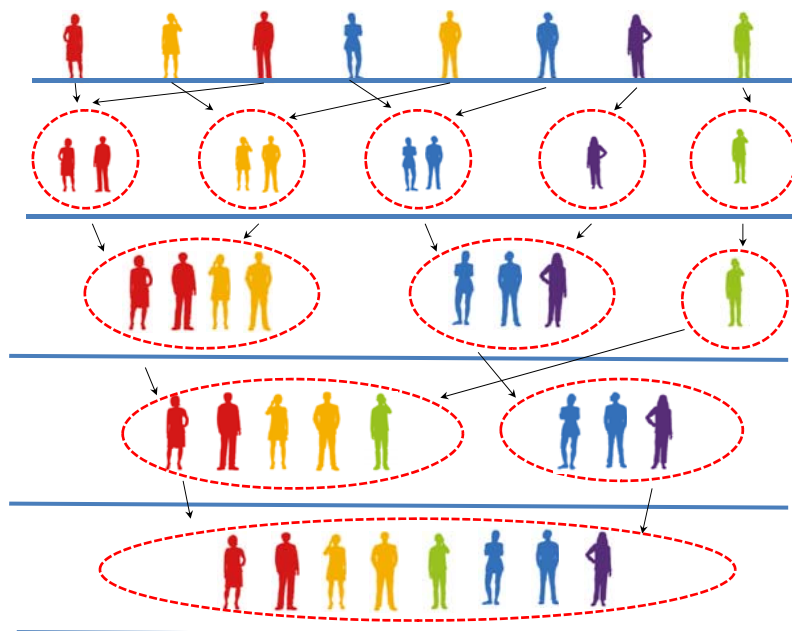
Ανάλυση ταξινόμησης

Cluster



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

51



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

52

Ποιος είναι ο στόχος της ανάλυσης ταξινόμησης;



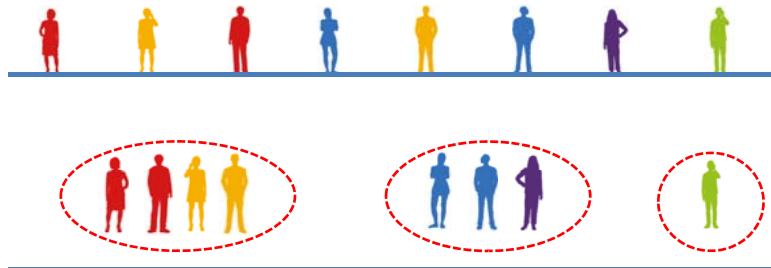
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

53

Ποιος είναι ο στόχος της ανάλυσης ταξινόμησης;



η ομαδοποίηση των **ατόμων** (βάσει κοινών χαρακτηριστικών – μεταβλητών)



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

54



**Ποια είναι η κυρίαρχη διαφορά μεταξύ
«ιεραρχημένης» και «μη ιεραρχημένης» ταξινόμησης;**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

55

**Ποια είναι η κυρίαρχη διαφορά μεταξύ
«ιεραρχημένης» και «μη ιεραρχημένης» ταξινόμησης;**

«ιεραρχημένη» ... (*hierarchical*):



- δεν είναι προαποφασισμένος ο αριθμός των κλάσεων



«μη ιεραρχημένη - διαμερισματική» ... (*k - means*):



- είναι προαποφασισμένος ο αριθμός των κλάσεων

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

56



Ποιες είναι

**δύο (2) από τις βασικές μεθόδους ελέγχου της απόστασης
(του βαθμού συγγένειας – γειτνίασης) μεταξύ 2 ατόμων
στην ανάλυση ταξινόμησης;**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

57

Ποιες είναι

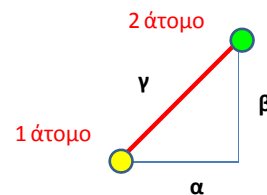
**δύο (2) από τις βασικές μεθόδους ελέγχου της απόστασης
(του βαθμού συγγένειας – γειτνίασης) μεταξύ 2 ατόμων
στην ανάλυση ταξινόμησης;**



Ευκλείδεια απόσταση: $\gamma = \sqrt{\alpha^2 + \beta^2}$
(*Euclidean distance*)



Τετράγωνο Ευκλείδειας απόστασης: $\gamma^2 = \alpha^2 + \beta^2$
(*Squared Euclidean distance*)



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

58



Βάσει ποιού κριτηρίου

πραγματοποιείται η ένωση 2 ατόμων (ή κλάσεων) ;

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

59

Βάσει ποιού κριτηρίου

πραγματοποιείται η ένωση 2 ατόμων (ή κλάσεων) ;



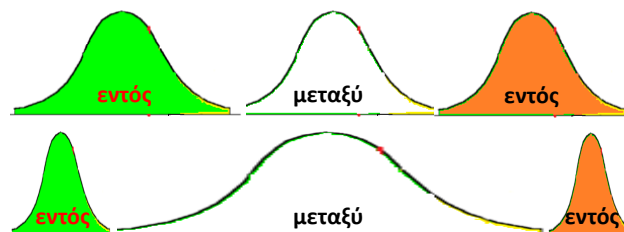
Βάσει του κριτηρίου του Ward



ελαχιστοποίηση της «εντός» της κλάσης διακύμανσης
και



αύξηση της «μεταξύ» των κλάσεων διακύμανσης



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

60



**Με τι ισούται η συνολική διακύμανση όλων των ατόμων
στην ΑΡΧΗ και στο ΤΕΛΟΣ της ανάλυσης ταξινόμησης**
(σε σχέση με τη διακύμανση «εντός» και «μεταξύ» των κλάσεων);

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

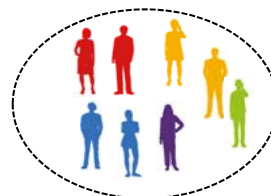
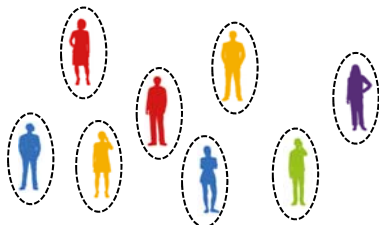
61

**Με τι ισούται η συνολική διακύμανση όλων των ατόμων
στην ΑΡΧΗ και στο ΤΕΛΟΣ της ανάλυσης ταξινόμησης**
(σε σχέση με τη διακύμανση «εντός» και «μεταξύ» των κλάσεων);

Στην ΑΡΧΗ
ισούται με την
«μεταξύ» των κλάσεων διακύμανση.



Στο ΤΕΛΟΣ
ισούται με την
«εντός» των κλάσεων διακύμανση.

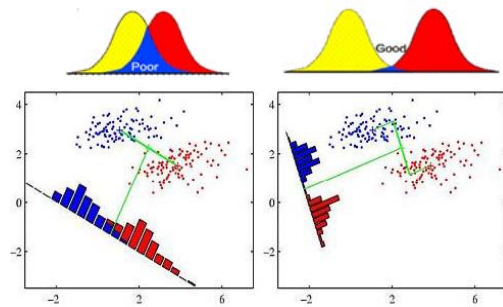


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

62

Ανάλυση διακριτότητας

Discriminant



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

63



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

64

Ποιος είναι ο κύριος στόχος
της ανάλυσης διακριτότητας:



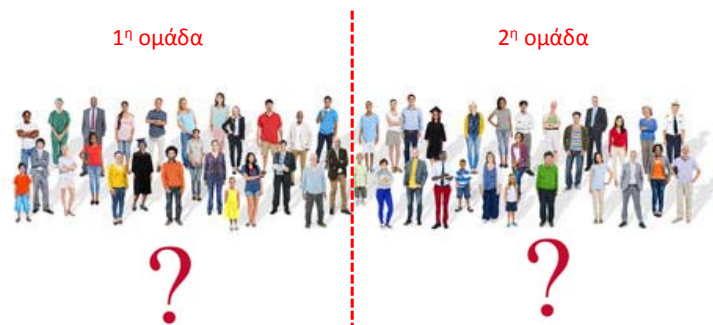
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

65

Ποιος είναι ο κύριος στόχος
της ανάλυσης διακριτότητας:



Να εξηγήσει
γιατί τα άτομα διαχωρίζονται σε (ήδη καθορισμένες) κλάσεις – ομάδες



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

66



Ποιοι είναι οι επιμέρους στόχοι της ανάλυσης διακριτότητας:

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

67

Ποιοι είναι οι επιμέρους στόχοι της ανάλυσης διακριτότητας:



Ποιες μεταβλητές είναι **σημαντικές**
για τη διάκριση των ατόμων σε **(ήδη καθορισμένες)** κλάσεις – ομάδες



Ποια είναι η **σειρά σπουδαιότητας** των μεταβλητών
για τη διάκριση των ατόμων σε **(ήδη καθορισμένες)** κλάσεις – ομάδες



Που θα **ενταχθεί** (σε ποια κλάση – ομάδα) **ένα νέο άτομο**
βάσει των τιμών – επιδόσεών του στις αρχικές μεταβλητές

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

68

Πότε επιτυγχάνεται η καλύτερη διάκριση μεταξύ των κλάσεων:



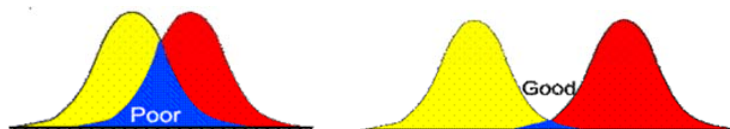
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

69

Πότε επιτυγχάνεται η καλύτερη διάκριση μεταξύ των κλάσεων:



Όταν
η διακύμανση **μεταξύ των κλάσεων** είναι η μέγιστη
σε σχέση με τη διακύμανση **εντός των κλάσεων**.



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

70



**Ποιος άξονας διακριτότητας
επιτυγχάνει την καλύτερη διάκριση
μεταξύ των κλάσεων:**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

71

**Ποιος άξονας διακριτότητας
επιτυγχάνει την καλύτερη διάκριση
μεταξύ των κλάσεων:**



Ο άξονας διακριτότητας
που έχει την **μεγαλύτερη ιδιοτιμή**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

72



**Τι δηλώνει η ιδιοτιμή
του άξονα διακριτότητας:**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

73

**Τι δηλώνει η ιδιοτιμή
του άξονα διακριτότητας:**



Το ποσό της συνολικής διακύμανσης
που ερμηνεύεται από τον άξονα διακριτότητας.

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

74



**Πως καθορίζεται
η σειρά σπουδαιότητας των μεταβλητών
για τη διάκριση μεταξύ των κλάσεων;**

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

75

**Πως καθορίζεται
η σειρά σπουδαιότητας των μεταβλητών
για τη διάκριση μεταξύ των κλάσεων;**



Η μεταβλητή
με τον μεγαλύτερο
συντελεστή συσχέτισης
με τον άξονα διακριτότητας
συμβάλει περισσότερο
στη δημιουργία
του άξονα διακριτότητας

Structure Matrix

	Function 1
χοληστερίνη	,882
τριγλυκερίδια	,744
διαστολική πίεση	,565
σάκχαρο	,538
συστολική πίεση	,419
ΚΣ ηρεμία	-,094

Pooled within-groups
correlations between
discriminating variables and
standardized canonical
discriminant functions
Variables ordered by
absolute size of correlation
within function.

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

76



Πως καθορίζεται η ποιότητα της διακριτότητας;

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

77

Πως καθορίζεται η ποιότητα της διακριτότητας;



Με το **ποσοστό των ατόμων που ομαδοποιήθηκαν «καλά».**

Δηλαδή,
το ποσοστό των ατόμων που μετά την ανάλυση
εντάχθηκαν στην κλάση που ανήκαν εξ αρχής.

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

78

Ερωτήσεις;



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

79