

Ανάλυση **διακρίτοτητας**

Discriminant analysis

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

1

Η **ανάλυση διακρίτοτητας**

εφαρμόζεται

στις περιπτώσεις όπου

το συνολικό δείγμα (ή πληθυσμός)

μπορεί να **διακριθεί σε ομάδες – κατηγορίες – κλάσεις**

βάσει μιας εξαρτημένης μεταβλητής,

η οποία έχει διάφορες γνωστές κατηγορίες (κλάσεις)

2

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

Ανάλυση διακριτότητας:

Είναι γνωστός (εκ των προτέρων) ο αριθμός των ομάδων
και ποια άτομα ανήκουν σε κάθε ομάδα



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

3

Μοντέλο ανάλυσης διακριτότητας

άτομο	X_1	X_2	X_3	...	X_m	Y
A_1	X_{11}	X_{12}	X_{13}	...	X_{1m}	Y_1
A_2	X_{21}	X_{22}	X_{23}	...	X_{2m}	Y_2
A_3	X_{31}	X_{32}	X_{33}	...	X_{3m}	Y_3
...	...	...	...	...	...	...
A_N	X_{N1}	X_{N2}	X_{N3}	...	X_{Nm}	Y_N

X_m = ποσοτικές μεταβλητές (χαρακτηριστικά)

Y = ποιοτική μεταβλητή (διαχωρίζει σε κλάσεις)

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

4

Προσπαθεί να εξηγήσει

γιατί τα άτομα διαχωρίζονται σε κλάσεις
(βάσει της μεταβλητής Y –
ποιοτική μεταβλητή – εξαρτημένη μεταβλητή)

με βάση τα χαρακτηριστικά τους
(ποσοτικές μεταβλητές X – ανεξάρτητες μεταβλητές)



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

5

Στόχοι ανάλυσης διακριτότητας



- Περιγραφικός στόχος

Να καθοριστούν οι **σχέσεις** που υπάρχουν
μεταξύ των **χαρακτηριστικών** του πληθυσμού
(ποσοτικές μεταβλητές X - ανεξάρτητες μεταβλητές)
οι οποίες τον διαχωρίζουν σε κλάσεις
(ποιοτική μεταβλητή Y - εξαρτημένη μεταβλητή)

- Προβλεπτικός στόχος

Να μπορεί να **«ενταχθεί»**
σε κάποιες από τις υπάρχουσες κατηγορίες (κλάσεις)
(ποιοτική μεταβλητή Y)
όποιο καινούριο άτομο παρουσιαστεί

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

6

Στόχοι ανάλυσης διακριτότητας



- Εξακρίβωση της ύπαρξης ή όχι **στατιστικά σημαντικών διαφορών** μεταξύ των μέσων όρων των δύο (ή περισσότερων) ήδη προκαθορισμένων ομάδων (κλάσεων)



- Καθορισμός της **σπουδαιότητας** των ποσοτικών μεταβλητών για την εμφάνιση των διαφορών μεταξύ των μέσων όρων των ήδη υπαρχόντων ομάδων (κλάσεων)

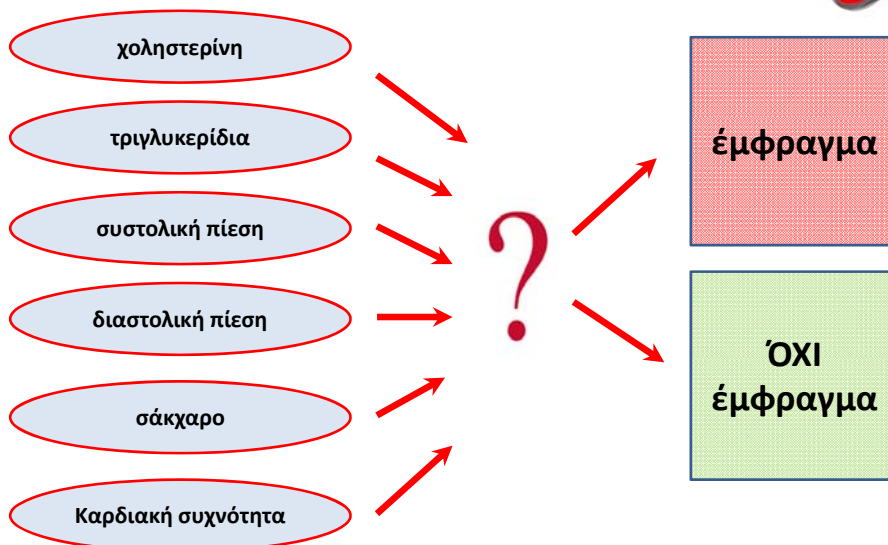


- Δυνατότητα **ένταξης** ενός καινούριου ατόμου σε κάποια από τις ήδη υπάρχουσες ομάδες (κλάσεις) βάσει της επίδοσής του σε κάποιες ποσοτικές μεταβλητές

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

7

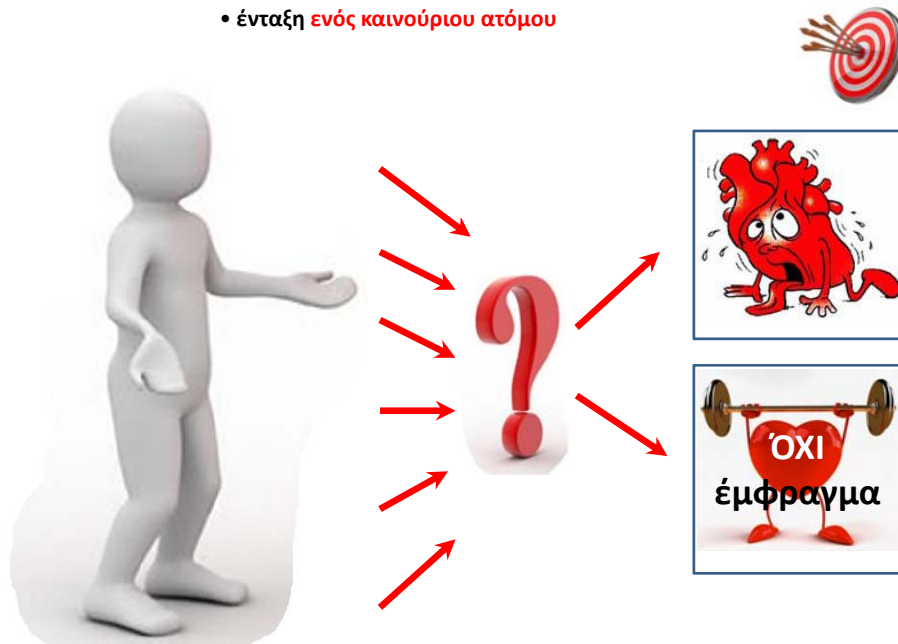
- ύπαρξη ή όχι **στατιστικά σημαντικών διαφορών**
- καθορισμός **σπουδαιότητας**



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

8

- ένταξη ενός καινούριου ατόμου



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

9

*Δεδομένα_06_

	xol	triglik	spres	dpres	sakx	ks	group
1	240	165	140	100	110	75	εμφραγμα
2	235	155	145	95	125	76	εμφραγμα
3	255	170	150	110	130	70	εμφραγμα
4	265	135	160	100	120	80	εμφραγμα
5	270	145	140	90	115	65	εμφραγμα
6	256	170	150	100	100	68	εμφραγμα
7	267	150	155	95	103	70	εμφραγμα
8	289	165	160	100	130	75	εμφραγμα
9	278	170	145	90	140	70	εμφραγμα
10	265	180	150	95	120	80	εμφραγμα
11	220	55	120	80	80	70	εμφραγμα
12	180	50	130	75	90	65	οχι εμφραγμα
13	200	90	125	80	95	75	οχι εμφραγμα
14	210	70	135	85	98	70	οχι εμφραγμα
15	195	68	140	90	100	80	οχι εμφραγμα
16	200	70	120	80	70	80	οχι εμφραγμα
17	210	65	135	85	75	75	οχι εμφραγμα
18	170	80	140	85	80	80	οχι εμφραγμα

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

10

Ανεξάρτητες μεταβλητές (ποσοτικές)

	xol	triglik	spres	dpres	sakx	ks	group
1	240	165	140	100	110	75	εμφραγμα
2	235	155	145	95	125	76	εμφραγμα
3	255	170	150	110	130	70	εμφραγμα
4	265	135	160	100	120	80	εμφραγμα
5	270	145	140	90	115	65	εμφραγμα
6	256	170	150	100	100	68	εμφραγμα
7	267	150	155	95	103	70	εμφραγμα
8	289	165	160	100	130	75	εμφραγμα
9	278	170	145	90	140	70	εμφραγμα
10	265	180	150	95	120	80	εμφραγμα
11	220	55	120	80	80	70	εμφραγμα
12	180	50	130	75	90	65	οχι εμφραγμα
13	200	90	125	80	95	75	οχι εμφραγμα
14	210	70	135	85	98	70	οχι εμφραγμα
15	195	68	140	90	100	80	οχι εμφραγμα
16	200	70	120	80	70	80	οχι εμφραγμα
17	210	65	135	85	75	75	οχι εμφραγμα
18	170	80	140	85	80	80	οχι εμφραγμα

Εξαρτημένη μεταβλητή (ποιοτική – κατηγορική)

1 = εμφραγμα
2= ΌΧΙ εμφραγμα

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

11

Ανεξάρτητες μεταβλητές (ποσοτικές)

	xol	triglik	spres	dpres	sakx	ks	group
1	240	165	140	100	110	75	1
2	235	155	145	95	125	76	1
3	255	170	150	110	130	70	1
4	265	135	160	100	120	80	1
5	270	145	140	90	115	65	1
6	256	170	150	100	100	68	1
7	267	150	155	95	103	70	1
8	289	165	160	100	130	75	1
9	278	170	145	90	140	70	1
10	265	180	150	95	120	80	1
11	220	55	120	80	80	70	1
12	180	50	130	75	90	65	2
13	200	90	125	80	95	75	2
14	210	70	135	85	98	70	2
15	195	68	140	90	100	80	2
16	200	70	120	80	70	80	2
17	210	65	135	85	75	75	2
18	170	80	140	85	80	80	2

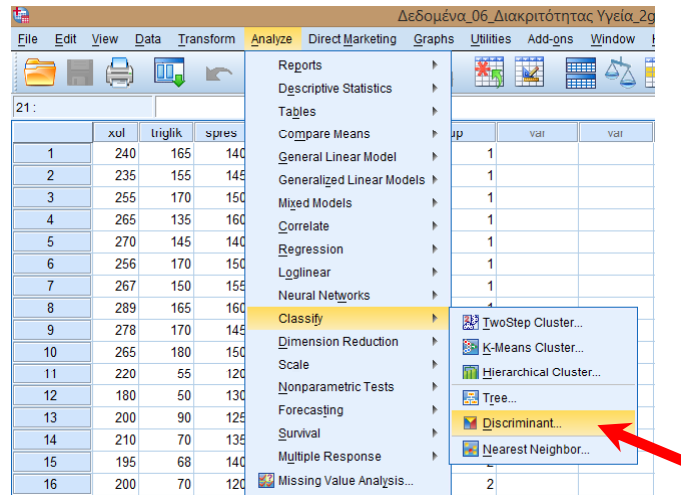
Εξαρτημένη μεταβλητή (ποιοτική – κατηγορική)

1 = εμφραγμα
2= ΌΧΙ εμφραγμα

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

12

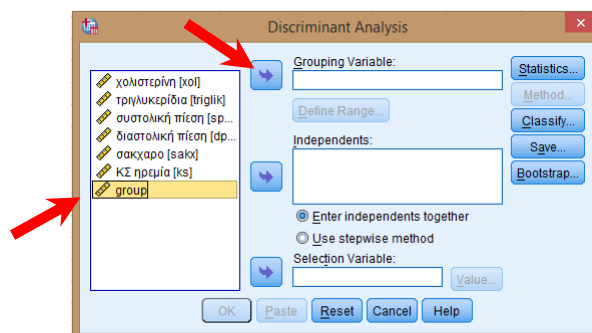
Analyze
Classify
Discriminant ...



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

13

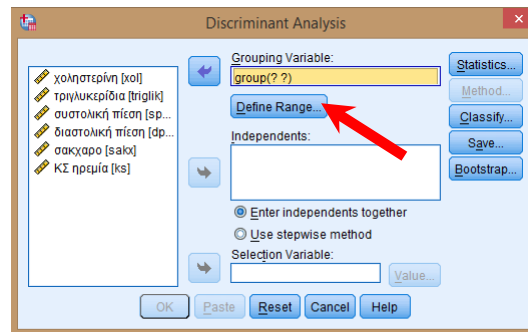
Καθορισμός εξαρτημένης μεταβλητής (Grouping Variable:)



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

14

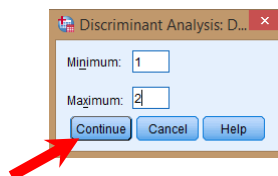
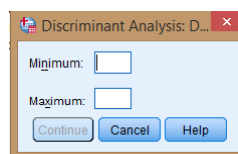
Καθορισμός **βαθμίδων** (κλάσεων – κατηγοριών) **εξαρτημένης** μεταβλητής (Define Range:)



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

15

Καθορισμός **βαθμίδων** (κλάσεων – κατηγοριών) **εξαρτημένης** μεταβλητής (Define Range:)



**Εξαρτημένη
μεταβλητή
(ποιοτική –
κατηγορική)**

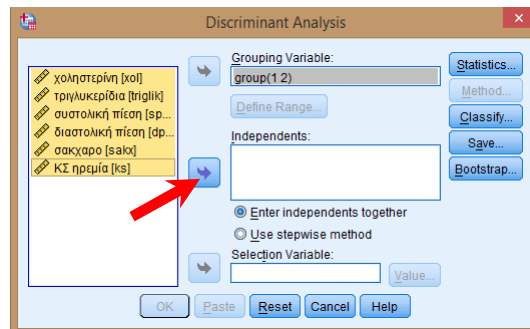
1 = έμφραγμα

2= Όχι έμφραγμα

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

16

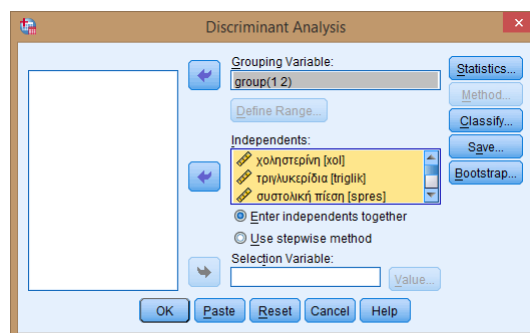
Καθορισμός **ανεξάρτητων** μεταβλητών (Independents:)



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

17

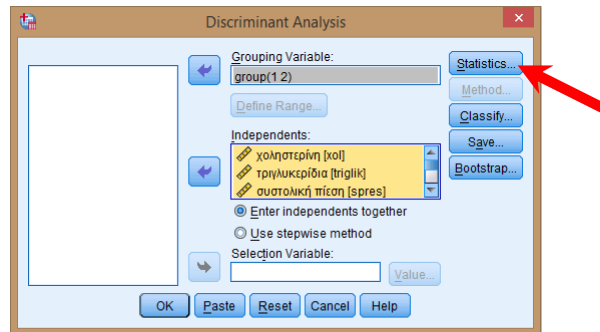
Καθορισμός **ανεξάρτητων** μεταβλητών (Independents:)



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

18

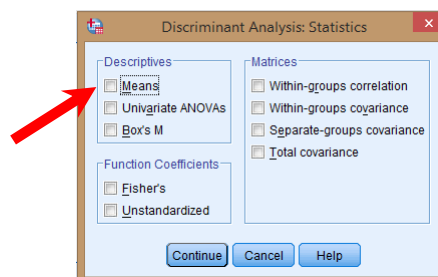
Επιλογή μέσων όρων (Statistics:)



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

19

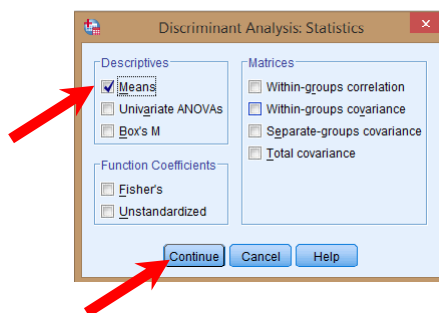
Επιλογή μέσων όρων (Statistics - Means:)



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

20

Επιλογή **μέσων όρων** (Statistics - Means:)



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

21

Μέσοι όροι των **ανεξάρτητων** μεταβλητών σε κάθε βαθμίδα της εξαρτημένης μεταβλητής

Group Statistics				Valid N (listwise)	
group		Mean	Std. Deviation	Unweighted	Weighted
εμφραγμα	χοληστερίνη	258,18	19,984	11	11,000
	τριγλυκερίδια	150,91	34,411	11	11,000
	συστολική πίεση	146,82	11,241	11	11,000
	διαστολική πίεση	95,91	7,687	11	11,000
	σακχαρο	115,73	16,847	11	11,000
	ΚΣ ηρεμία	72,64	4,884	11	11,000
οχι εμφραγμα	χοληστερίνη	197,22	15,635	9	9,000
	τριγλυκερίδια	74,22	13,198	9	9,000
	συστολική πίεση	131,11	7,817	9	9,000
	διαστολική πίεση	81,67	5,000	9	9,000
	σακχαρο	85,89	11,107	9	9,000
	ΚΣ ηρεμία	74,44	5,270	9	9,000
Total	χοληστερίνη	230,75	35,795	20	20,000
	τριγλυκερίδια	116,40	47,209	20	20,000
	συστολική πίεση	139,75	12,511	20	20,000
	διαστολική πίεση	89,50	9,720	20	20,000
	σακχαρο	102,30	20,815	20	20,000
	ΚΣ ηρεμία	73,45	5,010	20	20,000

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

22

Μέσοι όροι των **ανεξάρτητων** μεταβλητών σε κάθε βαθμίδα της εξαρτημένης μεταβλητής

Group Statistics

group		Mean
εμφραγμα	χοληστερίνη	258,18
	τριγλυκερίδια	150,91
	συστολική πίεση	146,82
	διαστολική πίεση	95,91
	σακχαρο	115,73
	ΚΣ ηρεμία	72,64
οχι εμφραγμα	χοληστερίνη	197,22
	τριγλυκερίδια	74,22
	συστολική πίεση	131,11
	διαστολική πίεση	81,67
	σακχαρο	85,89
	ΚΣ ηρεμία	74,44
Total	χοληστερίνη	230,75
	τριγλυκερίδια	116,40
	συστολική πίεση	139,75
	διαστολική πίεση	89,50
	σακχαρο	102,30
	ΚΣ ηρεμία	73,45

**Υπάρχει
στατιστικά
σημαντική διαφορά
μεταξύ τους**

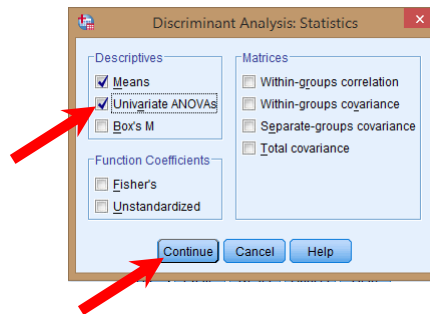


1^{ος} στόχος της ανάλυσης διακριτότητας



Εξακρίβωση της **ύπαρξης ή όχι**
στατιστικά σημαντικών διαφορών
μεταξύ των μέσων όρων των δύο (ή περισσότερων)
ήδη προκαθορισμένων ομάδων (κλάσεων)
(σε κάθε μία ανεξάρτητη μεταβλητή ξεχωριστά)

Έλεγχος στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων (Univariate ANOVAs:)



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

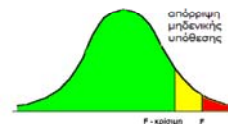
25

Έλεγχος στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων (Univariate ANOVAs:)

	Wilks' Lambda	F	df1	df2	Sig.
χοληστερίνη	,244	55,655	1	18	,000
τριγλυκερίδια	,313	39,593	1	18	,000
συστολική πίεση	,589	12,543	1	18	,002
διαστολική πίεση	,441	22,852	1	18	,000
σακχαρο	,465	20,739	1	18	,000
ΚΣ ηρεμία	,966	,632	1	18	,437

$H_0: \mu_{\text{έμφραγμα}} = \mu_{\text{Όχι έμφραγμα}}$

$F_{1,18} = \dots; p < 0.05$



$H_A: \mu_{\text{έμφραγμα}} \neq \mu_{\text{Όχι έμφραγμα}}$



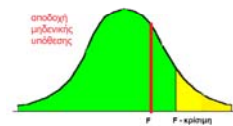
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

26

Έλεγχος στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων (Univariate ANOVAs:)

Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.
χοληστερίνη	,244	55,655	1	18	,000
τριγλυκερίδια	,313	39,593	1	18	,000
συστολική πίεση	,589	12,543	1	18	,002
διαστολική πίεση	,441	22,852	1	18	,000
σακχαρο	,465	20,739	1	18	,000
ΚΣ ηρεμία	,966	,632	1	18	,437

$H_0: \mu_{\text{έμφραγμα}} = \mu_{\text{Όχι έμφραγμα}}$ $F_{1,18} = \dots; p > 0.05$



$H_A: \mu_{\text{έμφραγμα}} \neq \mu_{\text{Όχι έμφραγμα}}$



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

27

Έλεγχος στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων (Univariate ANOVAs:)

Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.
χοληστερίνη	,244	55,655	1	18	,000
τριγλυκερίδια	,313	39,593	1	18	,000
συστολική πίεση	,589	12,543	1	18	,002
διαστολική πίεση	,441	22,852	1	18	,000
σακχαρο	,465	20,739	1	18	,000
ΚΣ ηρεμία	,966	,632	1	18	,437

Οι **F-τιμές**, για την κάθε μεταβλητή ξεχωριστά,
είναι οι ίδιες με τις **F-τιμές** από την
ανάλυση διακύμανσης με έναν παράγοντα (**One-way ANOVA**)

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

28

Οι **F-τιμές**, για την κάθε μεταβλητή ξεχωριστά,
είναι οι ίδιες με τις **F-τιμές** από την
ανάλυση διακύμανσης με έναν παράγοντα (**One-way ANOVA**)

		Sum of Squares	df	Mean Square	F	Sig.
χοληστερίνη	Between Groups	18394,558	1	18394,558	55,655	,000
	Within Groups	5949,192	18	330,511		
	Total	24343,750	19			
τριγλυκερίδια	Between Groups	29110,335	1	29110,335	39,593	,000
	Within Groups	13234,465	18	735,248		
	Total	42344,800	19			
συστολική πίεση	Between Groups	1221,225	1	1221,225	12,543	,002
	Within Groups	1752,525	18	97,363		
	Total	2973,750	19			
διαστολική πίεση	Between Groups	1004,091	1	1004,091	22,852	,000
	Within Groups	790,909	18	43,939		
	Total	1795,000	19			
σάκχαρο	Between Groups	4407,129	1	4407,129	20,739	,000
	Within Groups	3825,071	18	212,504		
	Total	8232,200	19			
ΚΣ ηρεμία	Between Groups	16,182	1	16,182	,632	,437
	Within Groups	460,768	18	25,598		
	Total	476,950	19			

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

29

Οι **F-τιμές**, για την κάθε μεταβλητή ξεχωριστά,
είναι οι ίδιες με τις **F-τιμές** από την
ανάλυση διακύμανσης με έναν παράγοντα (**One-way ANOVA**)

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
χοληστερίνη	Between Groups	18394,558	1	18394,558	55,655	,000
	Within Groups	5949,192	18	330,511		
	Total	24343,750	19			

$$F = \frac{\text{Μέσο τετράγωνο μεταξύ των ομάδων (Mean Square Between Groups)}}{\text{Μέσο τετράγωνο εντός των ομάδων (Mean Square Within Groups)}}$$

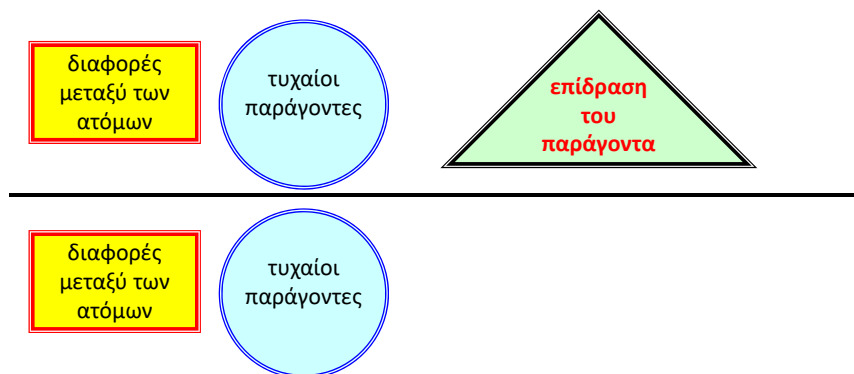
$$\text{Π.χ. } F = \frac{18394.558}{330.511} = 55.655$$

ΚΣ ηρεμία	Between Groups	16,182	1	16,182	,632	,437
	Within Groups	460,768	18	25,598		
	Total	476,950	19			

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

30

$$F = \frac{S_b^2}{S_w^2} = \frac{\text{μέσο άθροισμα τετραγώνων μεταξύ των ομάδων}}{\text{μέσο άθροισμα τετραγώνων εντός των ομάδων}}$$



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

$$F = \frac{S_b^2}{S_w^2} = \frac{\text{μέσο άθροισμα τετραγώνων μεταξύ των ομάδων}}{\text{μέσο άθροισμα τετραγώνων εντός των ομάδων}}$$



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

Έλεγχος **στατιστικά σημαντικών διαφορών** μεταξύ των **μέσων όρων** (Univariate ANOVAs:)

Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.
χοληστερίνη	,244	55,655	1	18	,000
τριγλυκερίδια	,313	39,593	1	18	,000
συστολική πίεση	,589	12,543	1	18	,002
διαστολική πίεση	,441	22,852	1	18	,000
σακχαρο	,465	20,739	1	18	,000
ΚΣ ηρεμία	,966	,632	1	18	,437

$$\text{Wilks' Lambda} = \frac{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}{\text{Συνολικό άθροισμα τετραγώνων (Total Sum of Squares)}}$$

εκφράζει
το ποσοστό της **συνολικής διακύμανσης**
που δεν οφείλεται σε διαφορές μεταξύ των ομάδων

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

33

Η **Wilks' Lambda** τιμή εκφράζει
το ποσοστό της **συνολικής διακύμανσης**
που δεν οφείλεται σε διαφορές μεταξύ των ομάδων

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
χοληστερίνη	Between Groups	18394,558	1	18394,558	55,655	,000
	Within Groups	5949,192	18	330,511		
	Total	24343,750	19			

$$\text{Wilks' Lambda} = \frac{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}{\text{Συνολικό άθροισμα τετραγώνων (Total Sum of Squares)}}$$

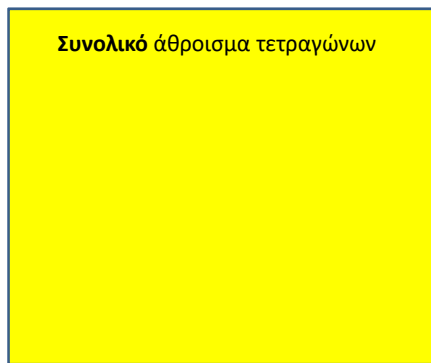
$$\text{Π.χ. Wilks' Lambda} = \frac{5949.192}{24343.750} = 0.244$$

Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.
χοληστερίνη	,244	55,655	1	18	,000

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

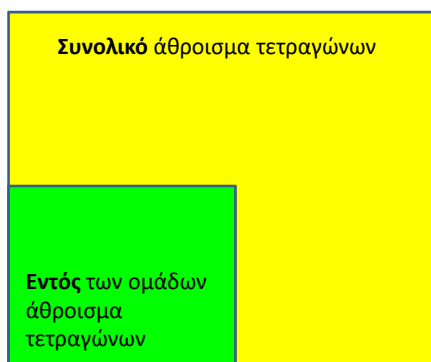
34

$$\text{Wilks' Lambda} = \frac{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}{\text{Συνολικό άθροισμα τετραγώνων (Total Sum of Squares)}}$$



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

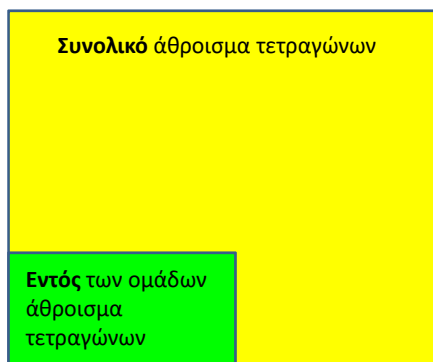
$$\text{Wilks' Lambda} = \frac{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}{\text{Συνολικό άθροισμα τετραγώνων (Total Sum of Squares)}}$$



Η Wilks' Lambda-τιμή
κυμαίνεται από 0 έως 1

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

$$\text{Wilks' Lambda} = \frac{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}{\text{Συνολικό άθροισμα τετραγώνων (Total Sum of Squares)}}$$



Όσο **μικρότερη** είναι η **Wilks' Lambda-τιμή** (και όσο περισσότερο πλησιάζει στο μηδέν) τόσο μεγαλύτερο μέρος της συνολικής διακύμανσης της κάθε μεταβλητής οφείλεται στις διαφορές μεταξύ των ομάδων

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

Συνεπώς ... για να υπάρχουν **στατιστικά σημαντικές διαφορές** μεταξύ των μέσων όρων των ήδη προκαθορισμένων ομάδων (κλάσεων)

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
χοληστερίνη	,244	55,655	1	18	,000
τριγλυκερίδια	,313	39,593	1	18	,000
συστολική πίεση	,589	12,543	1	18	,002
διαστολική πίεση	,441	22,852	1	18	,000
σακχαρο	,465	20,739	1	18	,000
ΚΣ ηρεμία	,966	,632	1	18	,437

- η τιμή **Wilks' Lambda** να είναι **όσο το δυνατόν μικρότερη** (όσο περισσότερο πλησιάζει στο 0, τόσο μεγαλύτερο μέρος της συνολικής διακύμανσης της κάθε μεταβλητής οφείλεται στις διαφορές μεταξύ των ομάδων)
- η **F -τιμή** να είναι **όσο το δυνατόν μεγαλύτερη** (για να υπάρχει στατιστικά σημαντική διαφορά μεταξύ των μέσων όρων των κλάσεων, στην συγκεκριμένη μεταβλητή, θα πρέπει το επίπεδο σημαντικότητας της F-τιμής (Sig.) να είναι μικρότερο από 0.05)

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

38



- Οι **μέσοι όροι των κλάσεων**, στην κάθε μεταβλητή και



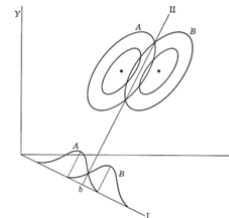
- η **ύπαρξη ή όχι** στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων των κλάσεων στην κάθε μία μεταβλητή ξεχωριστά

παρέχουν μια **πρώτη πληροφορία** για τις **διαφορές μεταξύ των κλάσεων** (ομάδων)

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

39

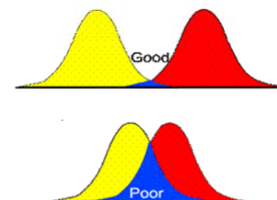
Στην **ανάλυση διακριτότητας** αναζητείται ο **γραμμικός συνδυασμός** μεταξύ όλων των ανεξάρτητων ποσοτικών μεταβλητών X που **διαχωρίζει καλύτερα τις κλάσεις**



$$Z = W_1X_1 + W_2X_2 + W_3X_3 + \dots + W_nX_n$$

Προβάλλοντας τα άτομα σε αυτόν τον **υποχώρο** θα πρέπει να επιτυγχάνεται η **καλύτερη διάκριση των κλάσεων**

(η διακύμανση **μεταξύ** των κλάσεων να 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.

40

Στην **παραγοντική** ανάλυση των κύριων συνιστωσών επιδιώκεται ο καθορισμός του **1^{ου} παραγοντικού άξονα**, στον οποίο η **διακύμανση** όλων των σημείων να είναι η **μέγιστη**.

Γι' αυτό το λόγο επιλέγεται εκείνος ο **άξονας** που έχει την **μεγαλύτερη ιδιοτιμή**.



Βάσει αυτής της **ιδιοτιμής** υπολογίζεται το **ιδιοδιάνυσμα** U_1 που ορίζει τον **1^ο παραγοντικό άξονα** (δηλαδή καθορίζει την διεύθυνσή του)

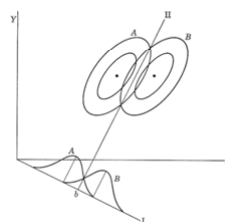
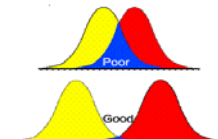
Ιδιοτιμή (το μήκος του ιδιοδιανύσματος)



41

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

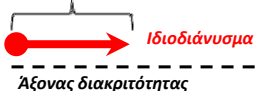
Στην ανάλυση **διακριτότητας** επιδιώκεται ο καθορισμός του γραμμικού συνδυασμού μεταξύ των ποσοτικών μεταβλητών (**άξονας διακριτότητας**) στον οποίο η **διακύμανση μεταξύ των κλάσεων** να είναι η **μέγιστη** (δηλαδή να καθοριστεί ο **άξονας** που διαχωρίζει καλύτερα τις υπάρχουσες κλάσεις).



Γι' αυτό το λόγο επιλέγεται εκείνος ο **άξονας** που έχει την **μεγαλύτερη ιδιοτιμή**.

Βάσει αυτής της **ιδιοτιμής** υπολογίζεται το **ιδιοδιάνυσμα** α που ορίζει τον **άξονα διακριτότητας** (δηλαδή καθορίζει την διεύθυνσή του)

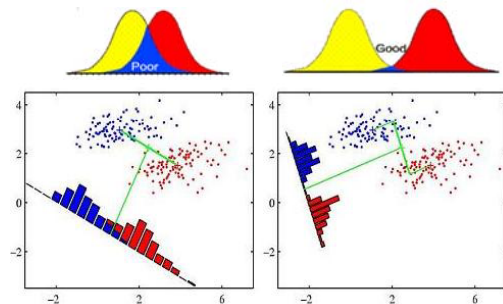
Ιδιοτιμή (το μήκος του ιδιοδιανύσματος)



42

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

Στην ανάλυση **διακριτότητας**
 επιδιώκεται
 ο καθορισμός ενός **υποχώρου**,
 στον χώρο των αρχικών μεταβλητών όπου βρίσκονται τα άτομα,
 έτσι ώστε
προβάλλοντας τα άτομα σ' αυτόν τον υποχώρο
 να επιτυγχάνεται η **καλύτερη διάκριση των κλάσεων**
 (η διακύμανση **μεταξύ των κλάσεων** να είναι η μέγιστη
 σε σχέση με τη διακύμανση **εντός των κλάσεων**).

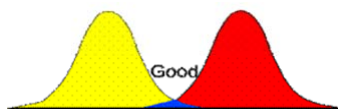


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

43

Κατανομή των τιμών διακριτότητας
 της 1ης ομάδας

Κατανομή των τιμών διακριτότητας
 της 2ης ομάδας



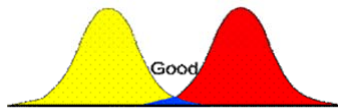
Αν η **αλληλοεπικάλυψη** των δύο κατανομών είναι **μικρή**
 τότε ο **άξονας διακριτότητας διαχωρίζει καλά** τις δύο ομάδες

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

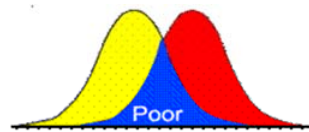
44

Κατανομή των τιμών διακριτότητας
της 1ης ομάδας

Κατανομή των τιμών διακριτότητας
της 2ης ομάδας



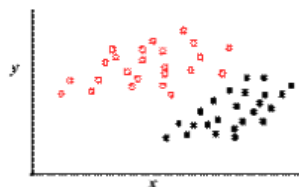
Αν η **αλληλοεπικάλυψη** των δύο κατανομών είναι **μικρή**
τότε ο **άξονας διακριτότητας διαχωρίζει καλά** τις δύο ομάδες



Αν η **αλληλοεπικάλυψη** των δύο κατανομών είναι **μεγάλη**
τότε ο **άξονας διακριτότητας δεν διαχωρίζει καλά** τις δύο ομάδες

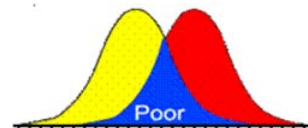
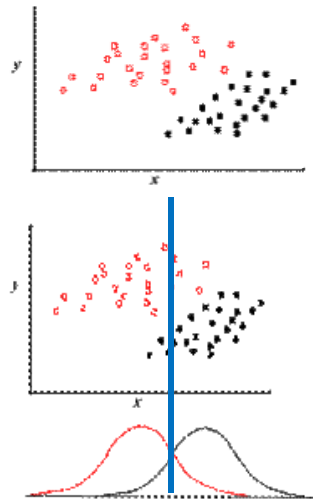
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

45



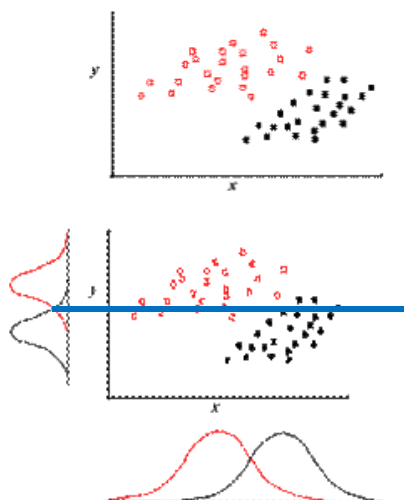
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

46



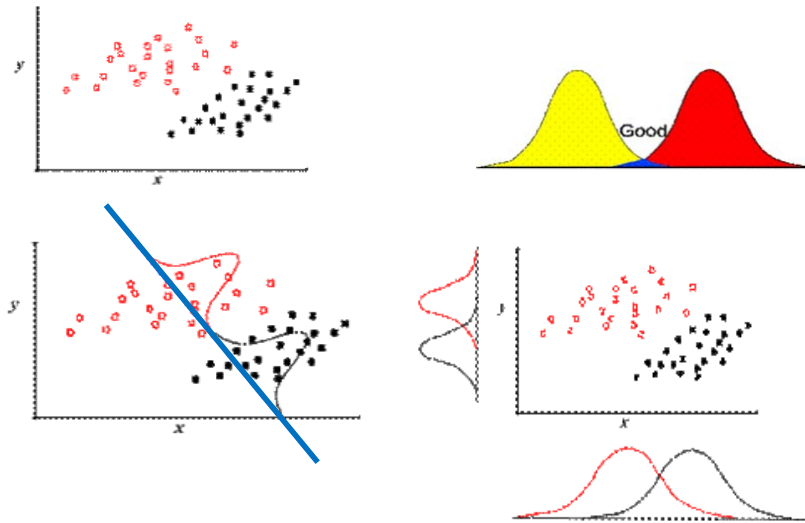
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

47



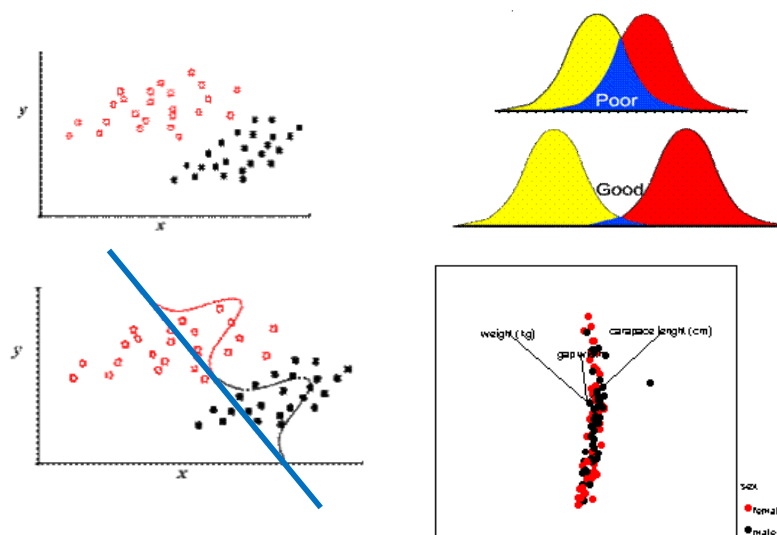
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

48



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

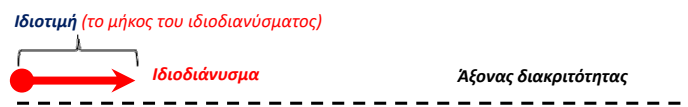
49



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

50

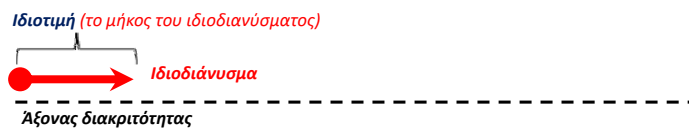
Πως ορίζεται ο νέος **υποχώρος** (**άξονας διακριτότητας**)
 στον χώρο των αρχικών **p** χαρακτηριστικών (ποσοτικών μεταβλητών)
 και ο οποίος διαχωρίζει καλύτερα τις κλάσεις των ατόμων
 όταν προβληθούν επάνω του;



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

51

Το «**μέγεθος**» του **ιδιοδιανύσματος**
 που ορίζει τον άξονα διακριτότητας,
 δηλαδή το **ποσό της συνολικής διακύμανσης**
που ερμηνεύεται από τον άξονα διακριτότητας,
 καθορίζεται από την **ιδιοτιμή** του.



Eigenvalues				
Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	3,975 ^a	100,0	100,0	,894

a. First 1 canonical discriminant functions were used in the analysis.

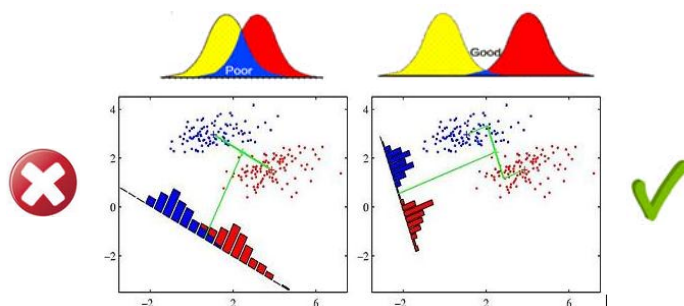
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

52

... Αν υπολογιστούν
 οι **τιμές – προβολές των ατόμων** στον άξονα διακριτότητας

Η **ιδιοτιμή** του ιδιοδιανύσματος
 που καθορίζει τον άξονα διακριτότητας
 θα πρέπει να λαμβάνει τη μέγιστη δυνατή τιμή της
 έτσι ώστε

η **διακύμανση μεταξύ των κλάσεων** (ομάδων) να είναι η **μέγιστη**
 σε σχέση με την διακύμανση **εντός των κλάσεων** (ομάδων)

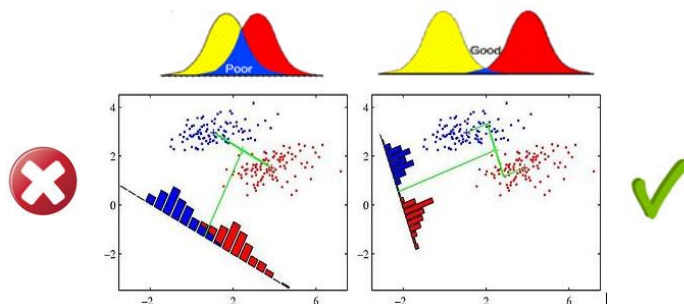


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

53

... Βάσει
 των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας

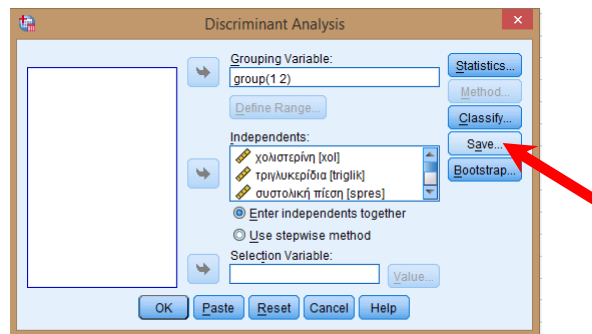
$$\text{Ιδιοτιμή} = \frac{\text{Άθροισμα τετραγώνων μεταξύ των ομάδων (Sum of Squares Between Groups)}}{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}$$



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

54

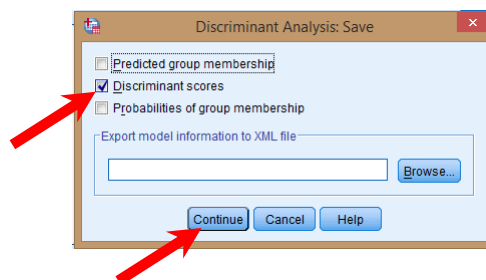
...Υπολογισμός
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας



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

55

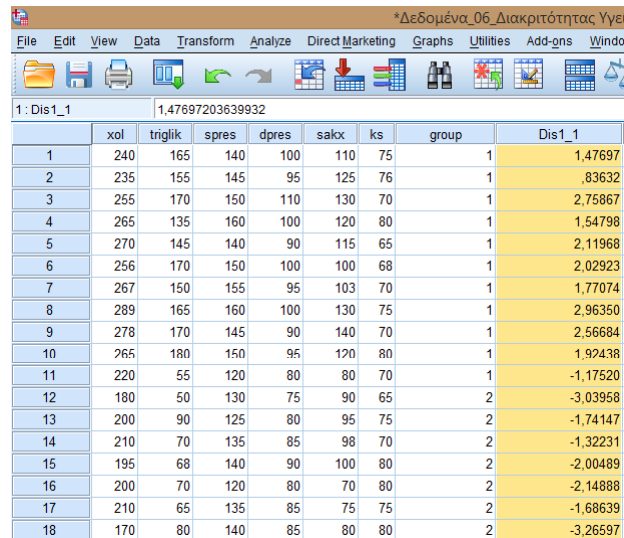
... Υπολογισμός
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας



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

56

... Υπολογισμός
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας

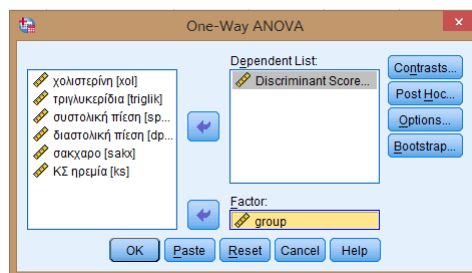


	xol	triglik	spres	dpres	sakx	ks	group	Dis1_1
1	240	165	140	100	110	75	1	1,47697
2	235	155	145	95	125	76	1	,83632
3	255	170	150	110	130	70	1	2,75867
4	265	135	160	100	120	80	1	1,54798
5	270	145	140	90	115	65	1	2,11968
6	256	170	150	100	100	68	1	2,02923
7	267	150	155	95	103	70	1	1,77074
8	289	165	160	100	130	75	1	2,96350
9	278	170	145	90	140	70	1	2,56684
10	265	180	150	95	120	80	1	1,92438
11	220	55	120	80	80	70	1	-1,17520
12	180	50	130	75	90	65	2	-3,03958
13	200	90	125	80	95	75	2	-1,74147
14	210	70	135	85	98	70	2	-1,32231
15	195	68	140	90	100	80	2	-2,00489
16	200	70	120	80	70	80	2	-2,14888
17	210	65	135	85	75	75	2	-1,68639
18	170	80	140	85	80	80	2	-3,26597

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

57

... One Way βάσει
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας



Dependent:
Discriminant score

Factor – independent :
group

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

58

... One Way βάσει
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας

ANOVA

... Από
One Way

Discriminant Scores from Function 1 for Analysis 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71,547	1	71,547	71,547	,000
Within Groups	18,000	18	1,000		
Total	89,547	19			

59

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

... One Way βάσει
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας

ANOVA

... Από
One Way

Discriminant Scores from Function 1 for Analysis 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71,547	1	71,547	71,547	,000
Within Groups	18,000	18	1,000		
Total	89,547	19			

$$\begin{aligned}
 \text{Ιδιοτιμή} &= \frac{\text{Άθροισμα τετραγώνων **μεταξύ** των ομάδων (Sum of Squares Between Groups)}}{\text{Άθροισμα τετραγώνων **εντός** των ομάδων (Sum of Squares Within Groups)}} \\
 &= \frac{71.547}{18.00} = 3.975
 \end{aligned}$$

60

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

... One Way βάσει
των **τιμών – προβολών των ατόμων** στον άξονα διακριτότητας

ANOVA

... Από
One Way

Discriminant Scores from Function 1 for Analysis 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71,547	1	71,547	71,547	,000
Within Groups	18,000	18	1,000		
Total	89,547	19			

Ιδιοτιμή = $\frac{\text{Άθροισμα τετραγώνων μεταξύ των ομάδων (Sum of Squares Between Groups)}}{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}$

$$= \frac{71.547}{18.00} = \mathbf{3.975}$$

Eigenvalues

... Από
Discriminant

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	3,975 ^a	100,0	100,0	,894

a. First 1 canonical discriminant functions were used in the analysis.

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

61

Το ποσό της συνολικής διακύμανσης
που ερμηνεύεται από τον άξονα διακριτότητας
και καθορίζεται από την **ιδιοτιμή** του
είναι στατιστικά σημαντικά διαφορετικό από το μηδέν;



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

62

Το ποσό της συνολικής διακύμανσης
που ερμηνεύεται από τον άξονα διακριτότητας
και καθορίζεται από την ιδιοτιμή του
είναι στατιστικά σημαντικά διαφορετικό από το μηδέν;



Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,201	24,066	6	,001

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

63

Το ποσό της συνολικής διακύμανσης
που ερμηνεύεται από τον άξονα διακριτότητας
και καθορίζεται από την ιδιοτιμή του
είναι στατιστικά σημαντικά διαφορετικό από το μηδέν;

Η **Wilks' Lambda** τιμή (υπολογίζεται βάσει όλων των μεταβλητών)
θα πρέπει να είναι **μικρή** και να πλησιάζει στο μηδέν.

Ελέγχει αν υπάρχουν **διαφορές**
μεταξύ των μέσων όρων των discriminant scores των ομάδων
(τιμές απόμων στον άξονα διακριτότητας)
και εκφράζει
το **τμήμα της συνολικής διακύμανσης των discriminant scores**
που δεν οφείλεται σε διαφορές μεταξύ των ομάδων

Wilks' Lambda

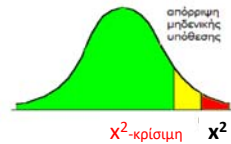
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,201	24,066	6	,001

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

64

Το ποσό της συνολικής διακύμανσης
που ερμηνεύεται από τον άξονα διακριτότητας
και καθορίζεται από την ιδιοτιμή του
είναι στατιστικά σημαντικά διαφορετικό από το μηδέν;

$$H_0: \mu_{\text{έμφραγμα}} = \mu_{\text{Όχι έμφραγμα}} \quad \chi^2_6 = 24.066; p < 0.05$$



$$H_A: \mu_{\text{έμφραγμα}} \neq \mu_{\text{Όχι έμφραγμα}}$$



Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,201	24,066	6	,001

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

65

**Επαλήθευση ... One Way βάσει
των τιμών – προβολών των ατόμων** στον άξονα διακριτότητας

ANOVA

... Από
One Way

Discriminant Scores from Function 1 for Analysis 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71,547	1	71,547	71,547	,000
Within Groups	18,000	18	1,000		
Total	89,547	19			

$$\text{Wilks' Lambda} = \frac{\text{Άθροισμα τετραγώνων εντός των ομάδων (Sum of Squares Within Groups)}}{\text{Συνολικό άθροισμα τετραγώνων (Total Sum of Squares)}}$$

$$= \frac{18.000}{89.547} = 0.201$$

Wilks' Lambda

... Από
Discriminant

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,201	24,066	6	,001

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

66

Τι ποσοστό της συνολικής διακύμανσης όλων των μεταβλητών
ερμηνεύεται από τις διαφορές μεταξύ των ομάδων



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

67

Τι ποσοστό της συνολικής διακύμανσης όλων των μεταβλητών
ερμηνεύεται από τις διαφορές μεταξύ των ομάδων

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	3,975 ^a	100,0	100,0	,894

a. First 1 canonical discriminant functions were used in the analysis.

$$(0.894)^2 = 0.80$$

80% της συνολικής διακύμανσης όλων των μεταβλητών
ερμηνεύεται από τις διαφορές μεταξύ των ομάδων

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

68

Πως ορίζεται ο **άξονας διακριτότητας** (το **ιδιοδιάνυσμά** του)
στον χώρο των αρχικών **p** χαρακτηριστικών (ποσοτικών μεταβλητών)



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

69

Πως ορίζεται ο **άξονας διακριτότητας** (το **ιδιοδιάνυσμά** του)
στον χώρο των αρχικών **p** χαρακτηριστικών (ποσοτικών μεταβλητών)

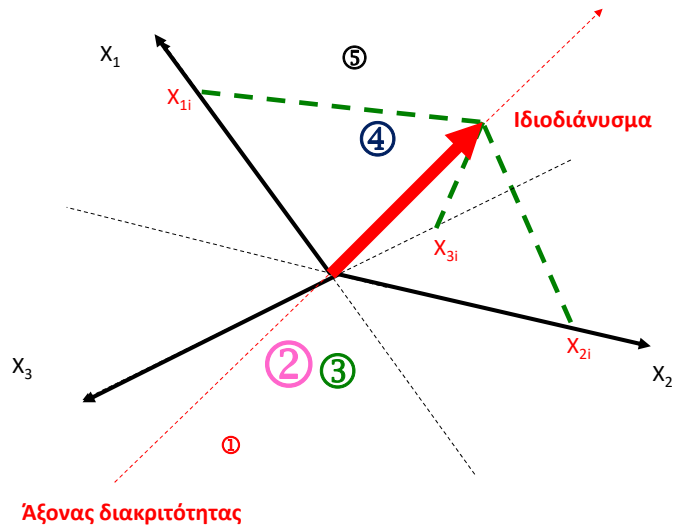


Μέσω των p συντεταγμένων του,
δηλαδή των προβολών του **ιδιοδιανύσματος**
που καθορίζει τη **διεύθυνση του άξονα διακριτότητας**
πάνω στις αρχικές μεταβλητές

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

70

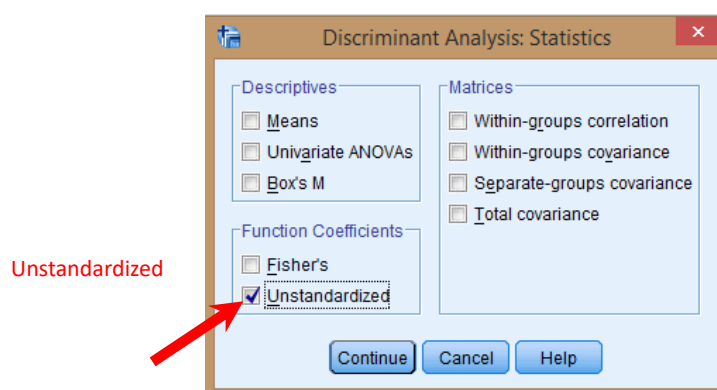
Άξονας διακρίτοτητας
(συντεταγμένες του ιδιοδιανύσματος που ορίζει τη διεύθυνση του άξονα διακρίτοτητας)



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

71

Συντεταγμένες **ιδιοδιανύσματος** άξονα διακρίτοτητας
στις αρχικές μεταβλητές



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

72

Οι **συντεταγμένες του ιδιοδιανύσματος**
που καθορίζει τη διεύθυνση του άξονα διακριτότητας
αποτελούν τους **συντελεστές του γραμμικού συνδυασμού**
(*unstandardized coefficients*)
των ανεξάρτητων (ποσοτικών) μεταβλητών
(όταν εκφράζονται σε πραγματικές μονάδες μέτρησης)

Canonical Discriminant Function Coefficients	
	Function
	1
χοληστερίνη	,037
τριγλυκερίδια	,010
συστολική πίεση	-,024
διαστολική πίεση	,062
σακχαρο	,007
ΚΣ ηρεμία	-,031
(Constant)	-10,401

Unstandardized coefficients

$$Z = W_0 + W_1X_1 + W_2X_2 + W_3X_3 + \dots$$

73

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

Βάσει αυτών των συντελεστών
(**unstandardized coefficients**)
του γραμμικού συνδυασμού των ανεξάρτητων (ποσοτικών) μεταβλητών
(όταν εκφράζονται **σε πραγματικές μονάδες μέτρησης**)
μπορούν να υπολογιστούν
οι **προβολές των ατόμων στον άξονα διακριτότητας** (Discriminant scores)

$$Z_1 = (0.037 * \text{χολ}) + (0.010 * \text{τριγ}) + (-0.024 * \text{ΣΠ}) + (0.062 * \text{ΔΠ}) + (0.007 * \text{Σακχ}) + (-0.031 * \text{ΚΣ}) + (-10.401) = 1.47697$$

Canonical Discriminant Function Coefficients	
	Function
	1
χοληστερίνη	,037
τριγλυκερίδια	,010
συστολική πίεση	-,024
διαστολική πίεση	,062
σακχαρο	,007
ΚΣ ηρεμία	-,031
(Constant)	-10,401

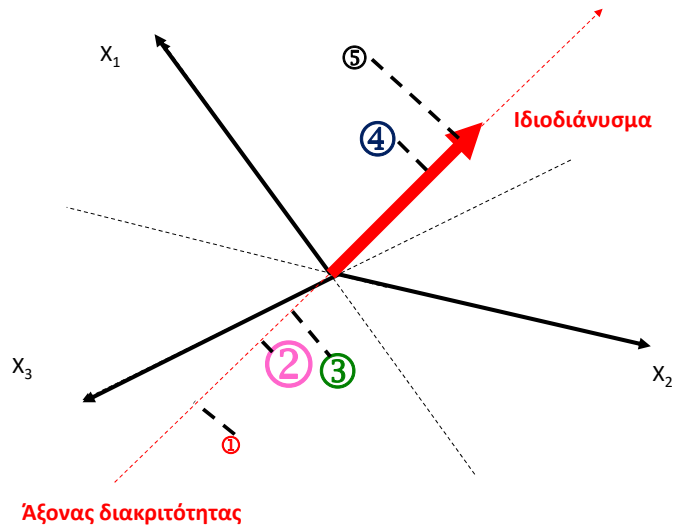
Unstandardized coefficients

	xol	triglik	spres	dpres	sakx	ks	group	Dis1_1
1	240	165	140	100	110	75	1	1,47697
2	235	155	145	95	125	76	1	,83632
3	255	170	150	110	130	70	1	2,75867

74

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

Άξονας διακριτότητας
(προβολές των ατόμων - *discriminant scores* - στον άξονα διακριτότητας)



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

75

Μέσος όρος
των τιμών – προβολών των ατόμων - *discriminant scores* –
στον άξονα διακριτότητας)

Functions at Group Centroids

group	Function
εμφραγμα	1,711
οχι εμφραγμα	-2,091

Unstandardized
canonical discriminant
functions evaluated at
group means



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

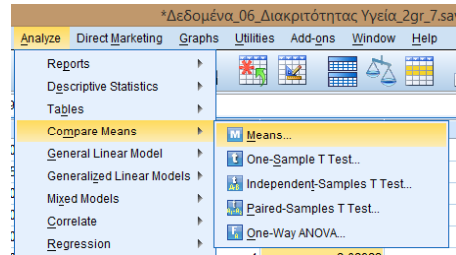
76

Επαλήθευση !!!

discriminant scores

	xol	triglik	spres	dpres	sakx	ks	group	Dist_1
1	240	165	140	100	110	75	1	1,47697
2	235	155	145	95	125	76	1	,83632
3	255	170	150	110	130	70	1	2,75967
4	265	135	160	100	120	80	1	1,54798
5	270	145	140	90	115	65	1	2,11968
6	256	170	150	100	100	68	1	2,02923
7	267	150	155	95	103	70	1	1,77074
8	289	165	160	100	130	75	1	2,96350
9	278	170	145	90	140	70	1	2,56684
10	265	180	150	95	120	80	1	1,52438
11	220	55	120	80	80	70	1	-1,17520
12	180	50	130	75	90	65	2	-3,03958
13	200	90	125	80	95	75	2	-1,74147
14	210	70	135	85	98	70	2	-1,32231
15	195	68	140	90	100	80	2	-2,00489
16	200	70	120	80	70	80	2	-2,14888
17	210	65	135	85	75	75	2	-1,68639
18	170	80	140	85	80	80	2	-3,26597
19	190	85	135	75	90	75	2	-2,74703
20	220	90	120	80	75	70	2	-.86259

Means



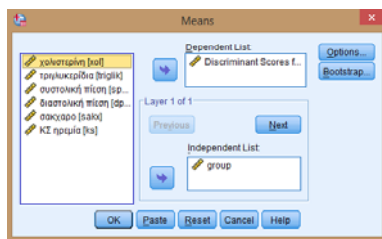
Report

Discriminant Scores from Function 1 for Analysis 1

group	Mean	N	Std. Deviation
εμφραγμα	1,7108279	11	1,13546230
οχι εμφραγμα	-2,0910118	9	,79900358
Total	0E-7	20	2,17094723

Functions at Group Centroids

group	Function
εμφραγμα	1,711
οχι εμφραγμα	-2,091



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

77

Συντελεστές διακριτότητας

που χρησιμοποιούνται όταν

οι τιμές των μεταβλητών έχουν μετατραπεί σε **z-τιμές**

(**Standardized Canonical Discriminant Function Coefficients**)

Standardized Canonical Discriminant Function Coefficients

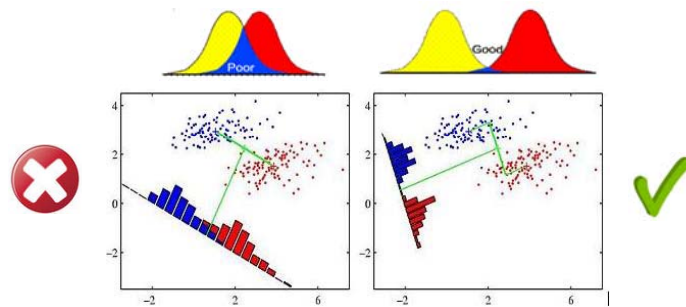
	Function
	1
χολιστερίνη	,677
τριγλυκερίδια	,269
συστολική πίεση	-,236
διαστολική πίεση	,410
σάκχαρο	,102
ΚΣ ηρεμία	-,156

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

78

Για να είναι οι κλάσεις καλά διακεκριμένες
θα πρέπει
οι αποστάσεις **μεταξύ** των κλάσεων να είναι **μέγιστες**

δηλαδή
η διακύμανση **μεταξύ των κλάσεων** να λαμβάνει την μέγιστη τιμή της
σε σχέση με την διακύμανση **εντός των κλάσεων**



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

79

2^{ος} στόχος της ανάλυσης διακριτότητας



Καθορισμός της σπουδαιότητας των ποσοτικών μεταβλητών
για την εμφάνιση των διαφορών
μεταξύ των μέσων όρων των ήδη υπάρχοντων ομάδων (κλάσεων)

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

80



**Ποιες μεταβλητές είναι σημαντικές
για τη διάκριση – διαχωρισμό των ομάδων (κλάσεων)**



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

81

Structure Matrix

	Function
	1
χοληστερίνη	.882
τριγλυκερίδια	.744
διαστολική πίεση	.565
σακχαρο	.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.

**Συντελεστές συσχέτισης που δηλώνουν
τη σχέση της κάθε μεταβλητής με την εξίσωση διακριτότητας.**

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

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

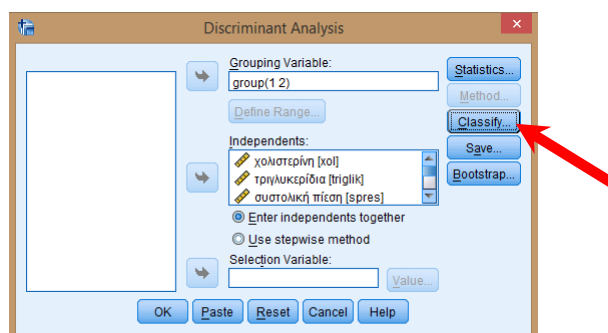
82

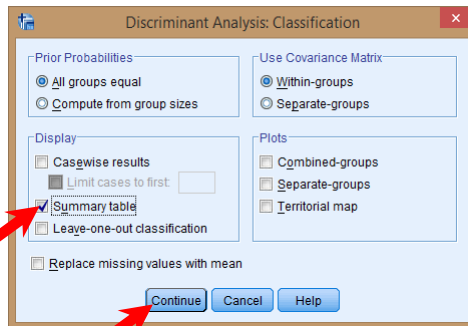
Ποιότητα της διακριτότητας

Η ποιότητα της διακριτότητας μετριέται με το **ποσοστό των ατόμων που ομαδοποιήθηκαν «καλά»**.

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

ισχύς διακριτότητας =
Ποσοστό των «ομαδοποιημένων καλώς ατόμων»





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

85

ισχύς διακριτότητας =
Ποσοστό των «ομαδοποιημένων καλώς ατόμων»

Classification Results^a

			Predicted Group Membership		Total
			εμφραγμα	οχι εμφραγμα	
Original	Count	εμφραγμα	10	1	11
		οχι εμφραγμα	0	9	9
	%	εμφραγμα	90,9	9,1	100,0
		οχι εμφραγμα	,0	100,0	100,0

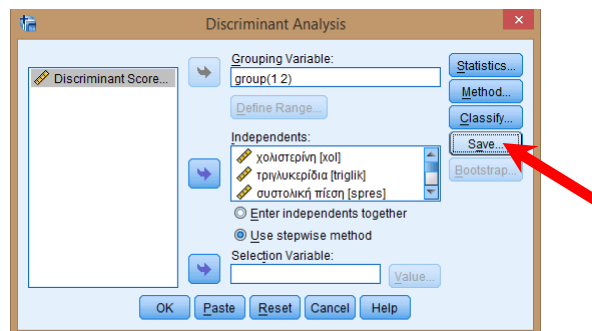
a. 95,0% of original grouped cases correctly classified.

95% των ατόμων
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.

86

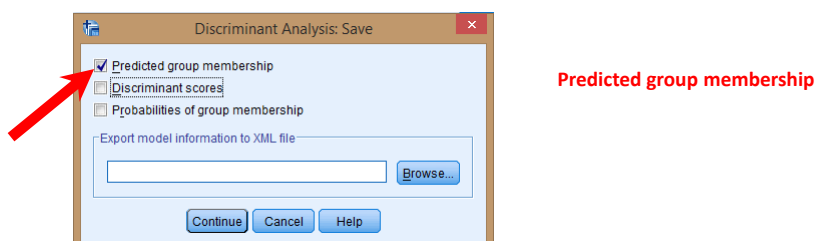
Ομαδοποίηση των ατόμων βάσει της εξίσωσης διακριτότητας



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

87

Ομαδοποίηση των ατόμων βάσει της εξίσωσης διακριτότητας



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

88

Ομαδοποίηση των ατόμων βάσει της εξίσωσης διακριτότητας

	xol	triglik	spres	dpres	sakx	ks	group	Dis_1
1	240	165	140	100	110	75	1	1
2	235	155	145	95	125	76	1	1
3	255	170	150	110	130	70	1	1
4	265	135	160	100	120	80	1	1
5	270	145	140	90	115	65	1	1
6	256	170	150	100	100	68	1	1
7	267	150	155	95	103	70	1	1
8	289	165	160	100	130	75	1	1
9	278	170	145	90	140	70	1	1
10	265	180	150	95	120	80	1	1
11	220	55	120	80	80	70	1	2
12	180	50	130	75	90	65	2	2
13	200	90	125	80	95	75	2	2
14	210	70	135	85	98	70	2	2
15	195	68	140	90	100	80	2	2
16	200	70	120	80	70	80	2	2
17	210	65	135	85	75	75	2	2
18	170	80	140	85	80	80	2	2
19	190	85	135	75	90	75	2	2
20	220	90	120	80	75	70	2	2

Predicted
group
membership

89

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

Ομαδοποίηση των ατόμων βάσει της εξίσωσης διακριτότητας

	xol	triglik	spres	dpres	sakx	ks	group	Dis_1
1	240	165	140	100	110	75	1	1
2	235	155	145	95	125	76	1	1
3	255	170	150	110	130	70	1	1
4	265	135	160	100	120	80	1	1
5	270	145	140	90	115	65	1	1
6	256	170	150	100	100	68	1	1
7	267	150	155	95	103	70	1	1
8	289	165	160	100	130	75	1	1
9	278	170	145	90	140	70	1	1
10	265	180	150	95	120	80	1	1
11	220	55	120	80	80	70	1	2
12	180	50	130	75	90	65	2	2
13	200	90	125	80	95	75	2	2
14	210	70	135	85	98	70	2	2
15	195	68	140	90	100	80	2	2
16	200	70	120	80	70	80	2	2
17	210	65	135	85	75	75	2	2
18	170	80	140	85	80	80	2	2
19	190	85	135	75	90	75	2	2
20	220	90	120	80	75	70	2	2

Predicted
group
membership



95%

90

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

3^{ος} στόχος της ανάλυσης διακριτότητας



Ένταξη ενός καινούριου ατόμου

σε κάποια από τις ήδη υπάρχουσες ομάδες (κλάσεις)
βάσει της επίδοσής του στις ποσοτικές μεταβλητές

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

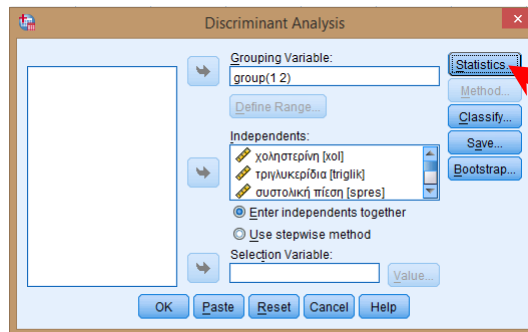
91

Για την **ένταξη ενός καινούριου ατόμου**
σε κάποια από τις ήδη υπάρχουσες ομάδες (κλάσεις)
βάσει της επίδοσής του στις ποσοτικές μεταβλητές
υπολογίζεται
η «τιμή» του νέου ατόμου
βάσει των συντελεστών κατάταξης Fisher.

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

92

Ένταξη ενός καινούριου ατόμου

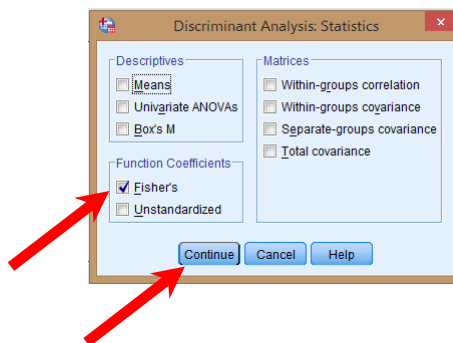


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

93

Ένταξη ενός καινούριου ατόμου

Fisher's



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

94

Ένταξη ενός καινούριου ατόμου

Classification Function Coefficients		
	group	
	εμφραγμα	οχι εμφραγμα
χοληστερίνη	,991	,849
τριγλυκερίδια	-,529	-,567
συστολική πίεση	,359	,450
διαστολική πίεση	1,586	1,351
σακχαρο	,096	,069
ΚΣ ηρεμία	2,706	2,823
(Constant)	-294,948	-256,127

Fisher's linear discriminant functions

$$Z = W_0 + W_1X_1 + W_2X_2 + W_3X_3 + \dots + W_nX_n$$

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

95

Ένταξη ενός καινούριου ατόμου

Για κάθε group (1= έμφραγμα, 2= όχι έμφραγμα) υπολογίζεται μια τιμή.

$$Z = W_0 + W_1X_1 + W_2X_2 + W_3X_3 + \dots + W_nX_n$$

Το άτομο **εντάσσεται στην ομάδα όπου έχει τη μεγαλύτερη τιμή**

	χολ	τριγ	ΣΠ	ΔΠ	σακχ	ΚΣ		
1 άτομο	240	165	140	100	110	75		271,493
12 άτομο	180	50	130	75	90	65		217,873
							Συντελεστής	Συντελεστής
							Fisher	Fisher
							1ου group	2ου group
							"έμφραγμα"	"όχι έμφραγμα"
							0,991	0,849
							-0,529	-0,567
							0,359	0,450
							1,586	1,351
							0,096	0,069
							2,706	2,823
							-294,948	-256,127
							Fisher's linear discriminant functions	

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

96

Ένταξη ενός καινούριου ατόμου

Π.χ. Το 1^ο άτομο

έχει μεγαλύτερη τιμή στην 1^η ομάδα (277,977) απ' ότι στην 2^η ομάδα (271.493)
οπότε εντάσσεται **στην 1^η ομάδα**

$$Z_1 = (0.991 * \text{χολ}) + (-0.529 * \text{τριγ}) + (0.359 * \text{ΣΠ}) + (1.586 * \text{ΔΠ}) + (0.096 * \text{σακχ}) + (2.706 * \text{ΚΣ}) + (-294.948) = 277.977$$

$$Z_2 = (0.849 * \text{χολ}) + (-0.567 * \text{τριγ}) + (0.450 * \text{ΣΠ}) + (1.351 * \text{ΔΠ}) + (0.069 * \text{σακχ}) + (2.823 * \text{ΚΣ}) + (-256.127) = 271.493$$

	χολ	τριγ	ΣΠ	ΔΠ	σακχ	ΚΣ		
1 άτομο	240	165	140	100	110	75	277,977	271,493
12 άτομο	180	50	130	75	90	65	207,132	217,873

Classification Function Coefficients			Συντελεστές Fisher	Συντελεστές Fisher
	group		1ου group	2ου group
	εμφραγμα	οχι εμφραγμα	"εμφραγμα"	"οχι εμφραγμα"
χοληστερίνη	,991	,849	0,991	0,849
τριγλυκερίδια	-,529	-,567	-0,529	-0,567
συστολική πίεση	,359	,450	0,359	0,45
διαστολική πίεση	1,586	1,351	1,586	1,351
σακχαρο	,096	,069	0,096	0,069
ΚΣ ηρεμία	2,706	2,823	2,706	2,823
(Constant)	-294,948	-256,127	-294,948	-256,127

Fisher's linear discriminant functions

97

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

Ένταξη ενός καινούριου ατόμου

Π.χ. Το 12^ο άτομο

έχει μεγαλύτερη τιμή στην 2^η ομάδα (217,873) απ' ότι στην 1^η ομάδα (207.132)
οπότε εντάσσεται **στην 2^η ομάδα**

$$Z_1 = (0.991 * \text{χολ}) + (-0.529 * \text{τριγ}) + (0.359 * \text{ΣΠ}) + (1.586 * \text{ΔΠ}) + (0.096 * \text{σακχ}) + (2.706 * \text{ΚΣ}) + (-294.948) = 207.132$$

$$Z_2 = (0.849 * \text{χολ}) + (-0.567 * \text{τριγ}) + (0.450 * \text{ΣΠ}) + (1.351 * \text{ΔΠ}) + (0.069 * \text{σακχ}) + (2.823 * \text{ΚΣ}) + (-256.127) = 217.873$$

	χολ	τριγ	ΣΠ	ΔΠ	σακχ	ΚΣ		
1 άτομο	240	165	140	100	110	75	277,977	271,493
12 άτομο	180	50	130	75	90	65	207,132	217,873

Classification Function Coefficients			Συντελεστές Fisher	Συντελεστές Fisher
	group		1ου group	2ου group
	εμφραγμα	οχι εμφραγμα	"εμφραγμα"	"οχι εμφραγμα"
χοληστερίνη	,991	,849	0,991	0,849
τριγλυκερίδια	-,529	-,567	-0,529	-0,567
συστολική πίεση	,359	,450	0,359	0,45
διαστολική πίεση	1,586	1,351	1,586	1,351
σακχαρο	,096	,069	0,096	0,069
ΚΣ ηρεμία	2,706	2,823	2,706	2,823
(Constant)	-294,948	-256,127	-294,948	-256,127

Fisher's linear discriminant functions

98

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

Συνοψίζοντας ...

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

99

Συνοψίζοντας ...

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.

100

Συνοψίζοντας ...

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.

101

Συνοψίζοντας ...

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.

102

Συνοψίζοντας ...

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.

103

Συνοψίζοντας ...

Πότε επιτυγχάνεται η καλύτερη διάκριση μεταξύ των κλάσεων:



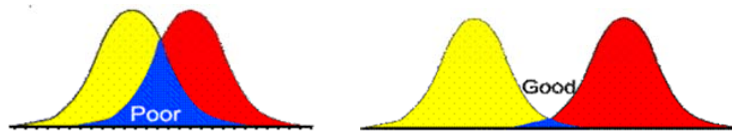
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

104

Συνοψίζοντας ...

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.

105

Συνοψίζοντας ...

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.

106

Συνοψίζοντας ...

**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.

107

Συνοψίζοντας ...

**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.

108

Συνοψίζοντας ...

**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.

109

Συνοψίζοντας ...

**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.

110

Συνοψίζοντας ...

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
τριηλυκερίδια	,744
διαστολική πίεση	,565
σακχαρο	,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.

111

Συνοψίζοντας ...

Πως καθορίζεται η ποιότητα της διακριτότητας;



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

112

Συνοψίζοντας ...

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.

113

Συγγραφή αποτελεσμάτων ...

Από την εφαρμογή της **ανάλυσης διακριτότητας**,
για τον εντοπισμό των σημαντικότερων μεταβλητών που συμβάλλουν στη **διάκριση μεταξύ των δύο (2)**
προκαθορισμένων ομάδων,

προέκυψε

ένας άξονας διακριτότητας με ιδιοτιμή 3.975 (Wilks' Lambda= -0.201; $p < 0.05$)

και

οι μεταβλητές που συσχετίζονται με τον άξονα διακριτότητας

είναι **κατά σειρά προτεραιότητας**

η χοληστερίνη (0.882), τα τριγλυκερίδια (0.744), η διαστολική πίεση (0.585), το σάκχαρο (0.538), η
συστολική πίεση (0.419) και τέλος η καρδιακή συχνότητα (-0.094).

Ωστόσο,

στατιστικά σημαντικές διαφορές μεταξύ των δύο ομάδων, εντοπίστηκαν σε όλες τις μεταβλητές
(χοληστερίνη: $F_{1,18} = 55.655$; $p < 0.05$; τριγλυκερίδια: $F_{1,18} = 39.593$; $p < 0.05$; συστολική πίεση: $F_{1,18} = 12.543$;
 $p < 0.05$; διαστολική πίεση: $F_{1,18} = 22.852$; $p < 0.05$; σάκχαρο: $F_{1,18} = 20.739$; $p < 0.05$), εκτός από την
καρδιακή συχνότητα ($F_{1,18} = 0.632$; $p = 0.437$).

Σε όλες τις μεταβλητές, εκτός από την καρδιακή συχνότητα, τα άτομα της πρώτης ομάδας (έμφραγμα)
παρουσίαζαν στατιστικά σημαντικά υψηλότερες τιμές, σε σύγκριση με τα άτομα της δεύτερης ομάδας
(όχι έμφραγμα).

Η ισχύς της διακριτότητας,

βάσει του ποσοστού των καλώς ομαδοποιημένων ατόμων στις ήδη υπάρχουσες ομάδες, ανέρχεται στο
95%.

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

114

... συγγραφή αποτελεσμάτων

Οι εξισώσεις, βάσει των οποίων ένα νέο άτομο (έχοντας τη μεγαλύτερη τιμή), θα ενταχθεί στην 1^η (έμφραγμα) ή στην 2^η ομάδα (όχι έμφραγμα) είναι η ακόλουθη :

$$Z_1 = (-294.948) + (0.991 * \text{χοληστερίνη}) + (-0.529 * \text{τριγλυκερίδια}) + (0.359 * \text{συστολική πίεση}) + (1.586 * \text{διαστολική πίεση}) + (0.096 * \text{σάκχαρο}) + (2.706 * \text{καρδιακή συχνότητα})$$

$$Z_2 = (-256.127) + (0.849 * \text{χοληστερίνη}) + (-0.567 * \text{τριγλυκερίδια}) + (0.450 * \text{συστολική πίεση}) + (1.351 * \text{διαστολική πίεση}) + (0.069 * \text{σάκχαρο}) + (2.823 * \text{καρδιακή συχνότητα})$$

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

115

Ερωτήσεις;



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

116

Youtube video_ Discriminant Analysis

Discriminant Analysis – SPSS : https://www.youtube.com/watch?v=vm1OLoL_VuY

Discriminant Analysis – SPSS : <https://www.youtube.com/watch?v=7zYcMZ-61c4>