

ΟΜΟΙΟΤΗΤΑ ΒΙΟΚΟΙΝΟΤΗΤΩΝ

Είναι δυνατόν δύο βιοκοινότητες να έχουν τον ίδιο (ή σχεδόν τον ίδιο) δείκτη ποικιλότητας ειδών αν και τα είδη που συνθέτουν τη μία βιοκοινότητα να είναι -σε μεγάλο βαθμό ή και τελείως- διαφορετικά από τα είδη της άλλης βιοκοινότητας.

Χρειαζόμαστε επομένως μεθόδους που να μας επιτρέπουν να συγκρίνουμε τις ομοιότητες μιας σειράς βιοκοινοτήτων με κριτήριο τα είδη που τις απαρτίζουν, και να ελέγξουμε πιθανή διαφοροποίησή τους ή υποβάθμιση μερικών εξ αυτών λόγω ρύπανσης.

Τα δεδομένα μπορεί να είναι *ποιοτικής ή ποσοτικής φύσεως*.

Ποιοτικής φύσεως λέγονται όταν σε κάθε δείγμα σημειώνουμε απλώς την παρουσία ή την απουσία του κάθε είδους.

Ποσοτικής φύσεως είναι τα δεδομένα όταν έχει καταγραφεί η αφθονία του κάθε είδους σε κάθε δείγμα.

$$X \underset{\text{Είδη}}{\parallel} \begin{matrix} & \text{Δείγματα} \\ \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \end{bmatrix} \end{matrix}$$

Τα στοιχεία ενός τέτοιου πίνακα μπορούν να αναλυθούν με:

(I) **R-mode ανάλυση** στην οποία υπολογίζονται οι συντελεστές ομοιότητας **μεταξύ των ειδών (γραμμών)** αναφορικά με τα δείγματα στα οποία υπάρχουν. Με άλλα λόγια αξιολογούμε το βαθμό συνύπαρξης (ή σχέσης) των ειδών στα δείγματα (δηλ. στον χώρο), όπως είδαμε σε προηγούμενο κεφάλαιο. *(Τα είδη συναντούνται μαζί στα ίδια δείγματα-βιοκοινότητες? Το γεγονός ότι σε ένα δείγμα υπάρχει κάποιο είδος A πόσο μας βεβαιώνει ότι υπάρχει και ένα είδος B)*

(II) **Q-mode ανάλυση** στην οποία υπολογίζονται οι συντελεστές ομοιότητας **μεταξύ των δειγμάτων (στηλών)** αναφορικά με τα είδη που περιέχουν. Οι συντελεστές αυτοί μπορεί να χρησιμοποιηθούν για «ταξινόμηση» (classification) των δειγμάτων (βιοκοινοτήτων) ή για «τακτοποίηση» των δεδομένων ή για «ιεράρχηση» (ordination) (π.χ. ανάλυση κυρίων συνιστωσών) όπως θα δούμε σε επόμενο κεφάλαιο.

(Τα δείγματα-βιοκοινότητες περιέχουν ίδια είδη?)

Δείγματα

Είδη	Δ1	Δ2	Δ3	Δ4	Δ5	Δ6	Δ7	Σύνολο
A	5	0	6	1	1	8	9	30
B	0	10	1	9	3	0	0	23
Γ	6	1	7	0	1	9	10	34
Δ	1	8	0	7	4	0	0	20
E	2	7	1	6	5	0	0	21
Z	1	5	0	4	3	1	1	15
Σύνολο	15	31	15	27	17	18	20	143

Αναδιάταξη γραμμών

Είδη	Δ1	Δ2	Δ3	Δ4	Δ5	Δ6	Δ7	Σύνολο
B	0	10	1	9	3	0	0	23
Δ	1	8	0	7	4	0	0	20
E	2	7	1	6	5	0	0	21
Z	1	5	0	4	3	1	1	15
A	5	0	6	1	1	8	9	30
Γ	6	1	7	0	1	9	10	34
Σύνολο	15	31	15	27	17	18	20	143

Αναδιάταξη στηλών

Είδη	Δ 2	Δ4	Δ5	Δ1	Δ3	Δ6	Δ7	Σύνολο
B	10	9	3	0	1	0	0	23
Δ	8	7	4	1	0	0	0	20
E	7	6	5	2	1	0	0	21
Z	5	4	3	1	0	1	1	15
A	0	1	1	5	6	8	9	30
Γ	1	0	1	6	7	9	10	34
Σύνολο	31	27	17	15	15	18	20	143

Ανάδειξη πληροφορίας

Ποιοτικά Δεδομένα

		Δείγμα Α		
		Παρουσία	Απουσία	Σύνολο
Δείγμα Β	Παρουσία	a	b	m
	Απουσία	c	-	
Σύνολο		r		

$$OI = \frac{a}{\sqrt{a+b} \sqrt{a+c}} \quad \text{Ochiai}$$

Γεωμετρικός μέσος
των a/m και a/r

$$DI = \frac{2a}{2a+b+c} \quad \text{Sorensen ή Dice}$$

Αρμονικός μέσος
των a/m και a/r

$$JI = \frac{a}{a+b+c} \quad \text{Jaccard}$$

<u>Είδος</u>	ΔΕΙΓΜΑ			
	I	II	III	IV
S1	5	6	+	+
S2	3	0		+
S3	0	3	+	+
S4	0	1	+	
S5	3	0	+	+

Να βρεθεί ο δείκτης ομοιότητας Jaccard

μεταξύ των δειγμάτων III και IV

$$JI = \frac{a}{a + b + c}$$

$$JI = \frac{a}{a + b + c} = \frac{3}{3 + 1 + 1}$$

Jaccard	Δείγμα III		
	+	-	
Δείγμα IV	+	3	1
	-	1	
			0.60

<u>Είδος</u>	ΔΕΙΓΜΑ			
	I	II	III	IV
S1	5	6	+	+
S2	3	0		+
S3	0	3	+	+
S4	0	1	+	
S5	3	0	+	+

Να βρεθεί ο δείκτης ομοιότητας Dice

μεταξύ των δειγμάτων I και III

$$DI = \frac{2a}{2a + b + c}$$

$$DI = \frac{2a}{2a + b + c} = \frac{2 \times 2}{2 \times 2 + 2 + 1} = 0.57$$

Dice	Δείγμα III		
	+	-	
Δείγμα I	+	2	1
	-	2	
			0.57

<u>Είδος</u>	ΔΕΙΓΜΑ			
	I	II	III	IV
S1	5	6	+	+
S2	3	0		+
S3	0	3	+	+
S4	0	1	+	
S5	3	0	+	+

Να βρεθεί ο δείκτης ομοιότητας Ochiai

μεταξύ των δειγμάτων II και IV

$$OI = \frac{a}{\sqrt{a+b}\sqrt{a+c}}$$

$$OI = \frac{2}{\sqrt{2+2}\sqrt{2+1}} = 0.58$$

Ochiai	Δείγμα II		
	+	-	
Δείγμα IV	+	2	2
	-	1	
			0.57

	Δείγμα J	Δείγμα k	Σύνολο
Είδος 1	5	10	15
Είδος 2	6	9	15
Είδος 3	12	7	19
Σύνολο	23	26	49

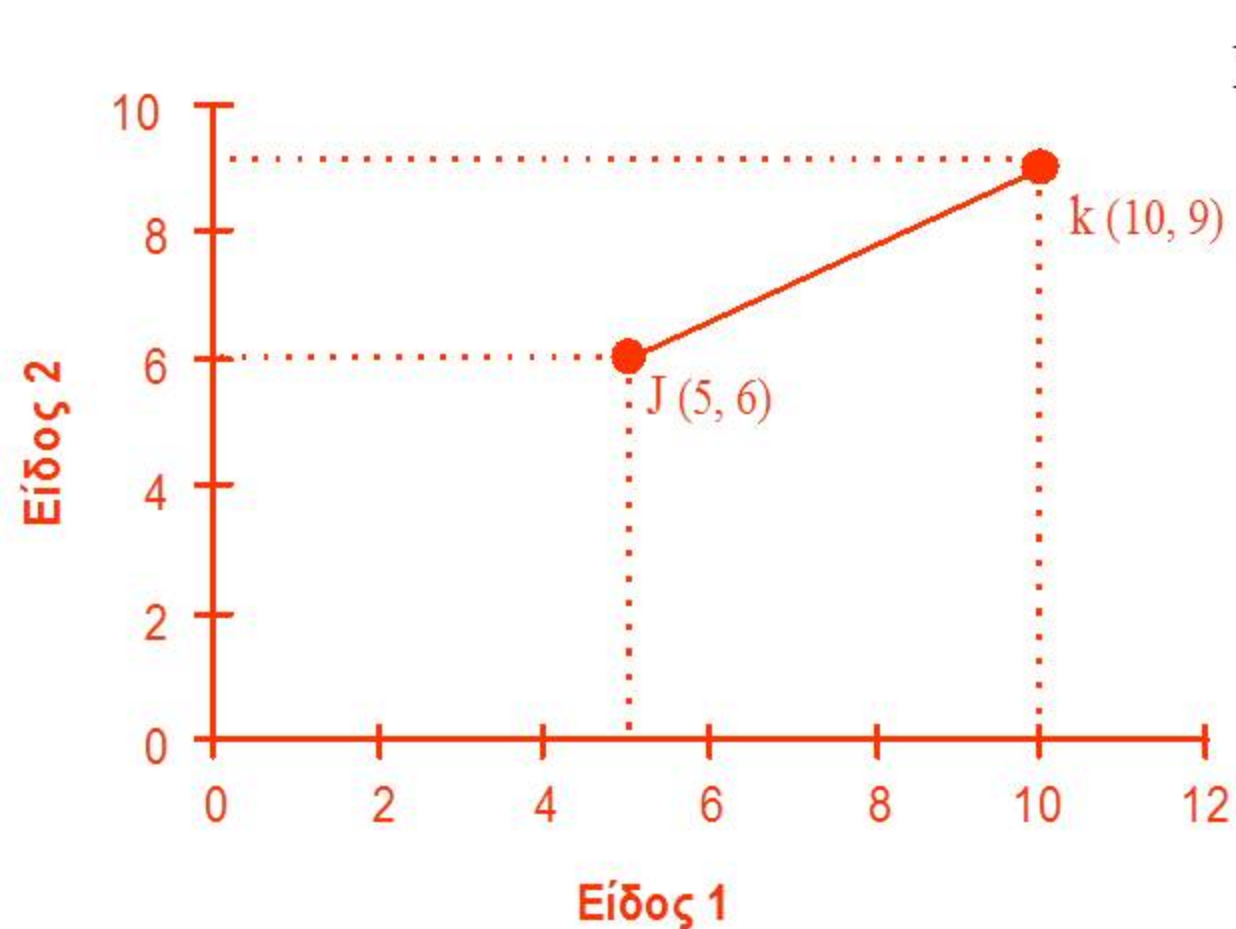
Ποσοτικά Δεδομένα

Ευκλείδεια απόσταση

$$ED_{jk} = \sqrt{\sum_i^s (X_{ij} - X_{ik})^2} \quad (0 \leq ED \leq \infty)$$

Επίσης το τετράγωνο της Ευκλείδειας απόστασης

	Δείγμα J	Δείγμα k	Σύνολο
Είδος 1	5	10	15
Είδος 2	6	9	15
Σύνολο	11	19	30



$$ED_{jk} = \sqrt{(5 - 10)^2 + (6 - 9)^2} = 5,8$$

	Δείγμα J	Δείγμα k	Σύνολο
Είδος 1	5	10	15
Είδος 2	6	9	15
Είδος 3	12	7	19
Σύνολο	23	26	49

$$ED_{jk} = \sqrt{(5-10)^2 + (6-9)^2 + (12-7)^2} = 7,7$$

Το τετράγωνο της Ευκλείδειας Απόστασης είναι 59,29

Η Σχετική Ευκλείδεια Απόσταση

$$\mathbf{RED}_{jk} = \sqrt{\sum_i^s \left(\frac{X_{ij}}{\sum_i^s X_{ij}} - \frac{X_{ik}}{\sum_i^s X_{ik}} \right)^2} \quad (0 \leq \mathbf{RED} \leq \sqrt{2})$$

	Δείγμα J	Δείγμα k	Σύνολο
Είδος 1	5	10	15
Είδος 2	6	9	15
Είδος 3	12	7	19
Σύνολο	23	26	49

$$RED_{jk} = \sqrt{\left(\frac{5}{23} - \frac{10}{26}\right)^2 + \left(\frac{6}{23} - \frac{9}{26}\right)^2 + \left(\frac{12}{23} - \frac{7}{26}\right)^2} = 0,31$$

Ποσοτικά Δεδομένα

Συντελεστής ομοιότητας
Bray και Curtis

$$W_{jk} = \sum_i^s [\min(X_{ij}, X_{ik})]$$

$$PS_{jk} = \frac{2W}{A+B} \quad (0 \leq PS \leq 1)$$

$$A = \sum_i^s X_{ij}$$

$$B = \sum_i^s X_{ik}$$

Συντελεστής ανομοιότητας
Bray και Curtis

$$PD = 1 - PS$$

	Δείγμα J	Δείγμα k	Σύνολο
Είδος 1	5	10	15
Είδος 2	6	9	15
Είδος 3	12	7	19
Σύνολο	23	26	49

$$PS_{jk} = \frac{(2)(18)}{(23) + (26)} = \frac{36}{49} = 0.7347$$

Είδος	ΔΕΙΓΜΑ		Min
	I	II	
S1	3	4	3
S2	1	1	1
S3	7	2	2
S4	1	6	1
S5	3	5	3
	15	18	10
	A	B	W
Bray-Curtis			
PS= 0.6061			

$$PS = \frac{2 \times 10}{15 + 18} = \frac{20}{33} = 0.6061$$

<u>Είδος</u>	ΔΕΙΓΜΑ		ED
	I	II	
S1	5	3	4
S2	3	1	4
S3	2	3	1
S4	1	1	0
S5	3	5	4
			13

Ευκλείδεια
Απόσταση **3.606**

$$ED = \sqrt{(5-3)^2 + (3-1)^2 + (2-3)^2 + (1-1)^2 + (3-5)^2} = 3.606$$

<u>Είδος</u>	ΔΕΙΓΜΑ		RED		
	I	II			
S1	3	2	0.158	0.167	0.000
S2	4	1	0.211	0.083	0.016
S3	3	3	0.158	0.250	0.008
S4	6	1	0.316	0.083	0.054
S5	3	5	0.158	0.417	0.067
	19	12			

Σχετική
Ευκλείδεια
Απόσταση **0.382**

	A	B	Γ	Δ	E
Είδος 1	2	3	7	20	25
Είδος 2	4	7	18	22	21

$$RI_{jk} = \frac{\sum_i^s \min(X_{ij}, X_{ik})}{\sum_i^s \max(X_{ij}, X_{ik})}$$

Δείκτης Ομοιότητας του Ruzicka

$$AD_{jk} = \sum_i^s |X_{ij} - X_{ik}|$$

Απόλυτος Απόσταση (Manhattan δείκτης)
 $0 \leq AD \leq \infty$

$$RAD_{jk} = \sum_i^s \left| \left(\frac{X_{ij}}{\sum_i^s X_{ij}} - \frac{X_{ik}}{\sum_i^s X_{ik}} \right) \right|$$

Σχετική Απόλυτος Απόσταση
 $0 \leq RAD \leq 2$

$$MAD_{jk} = \frac{\sum_i^s |X_{ij} - X_{ik}|}{S}$$

Μέση Απόλυτος Απόσταση
 $0 \leq MAD \leq \infty$

Απόσταση Χορδής

$$CRD_{jk} = \sqrt{2(1 - \cos\theta_{jk})} = 2\eta\mu\frac{\theta_{jk}}{2} \quad \cos\theta_{jk} = \frac{\sum_i^s (X_{ij} \cdot X_{ik})}{\sqrt{\sum_i^s X_{ij}^2 \cdot \sum_i^s X_{ik}^2}} \quad 0 \leq CRD \leq \sqrt{2}$$

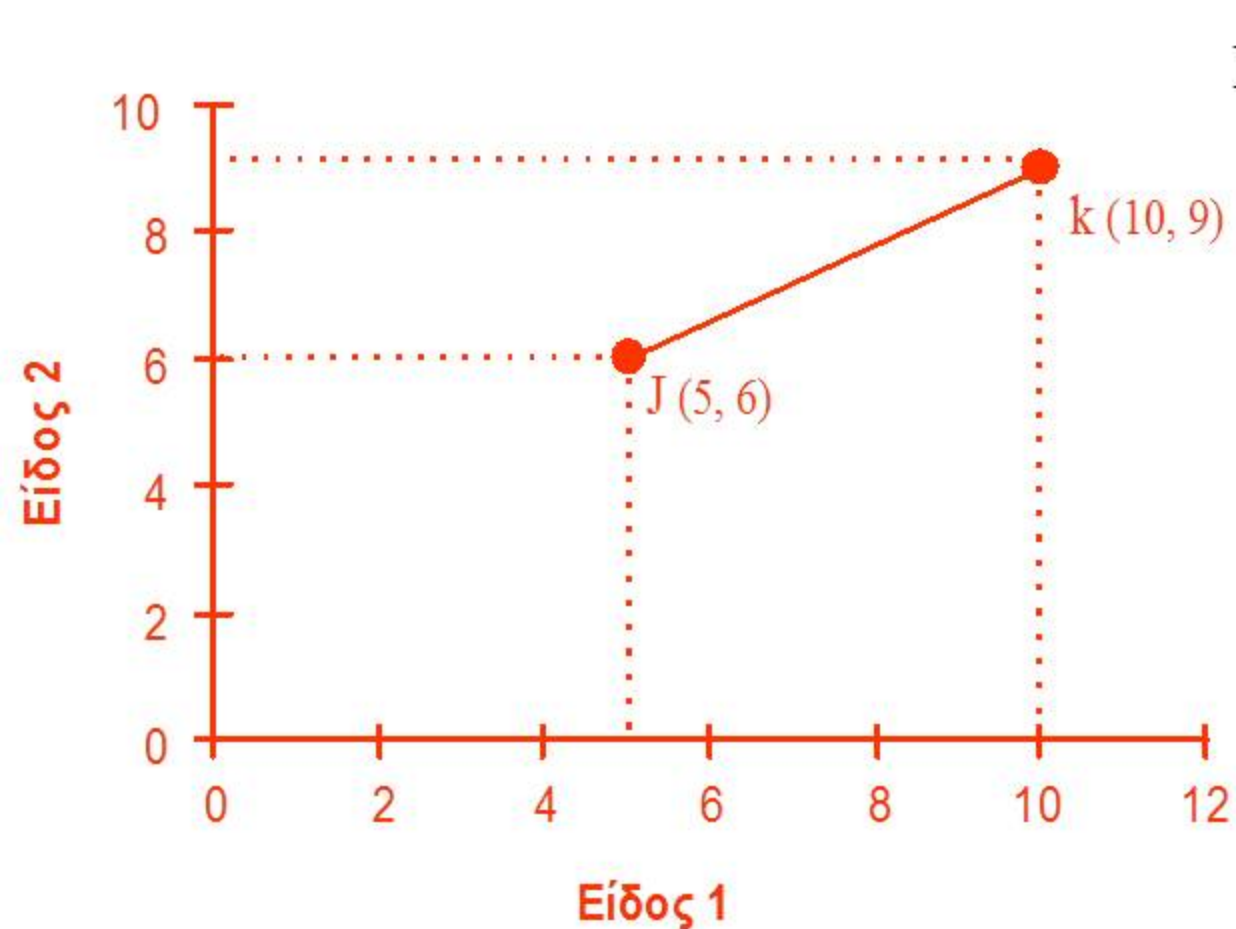
Γεωδαιτική Απόσταση (τόξου)

Είναι επίσης ένας σχετικός δείκτης που ορίζεται ως το μήκος τόξου κύκλου μοναδιαίας ακτίνας, που αντιστοιχεί στην γωνία θ_{jk} (σε ακτίνια) την οποία σχηματίζουν τα δύο δείγματα (j και k) στον χώρο των S διαστάσεων. Το μήκος του τόξου είναι απλά το μέτρο της γωνίας θ_{jk} σε ακτίνια.

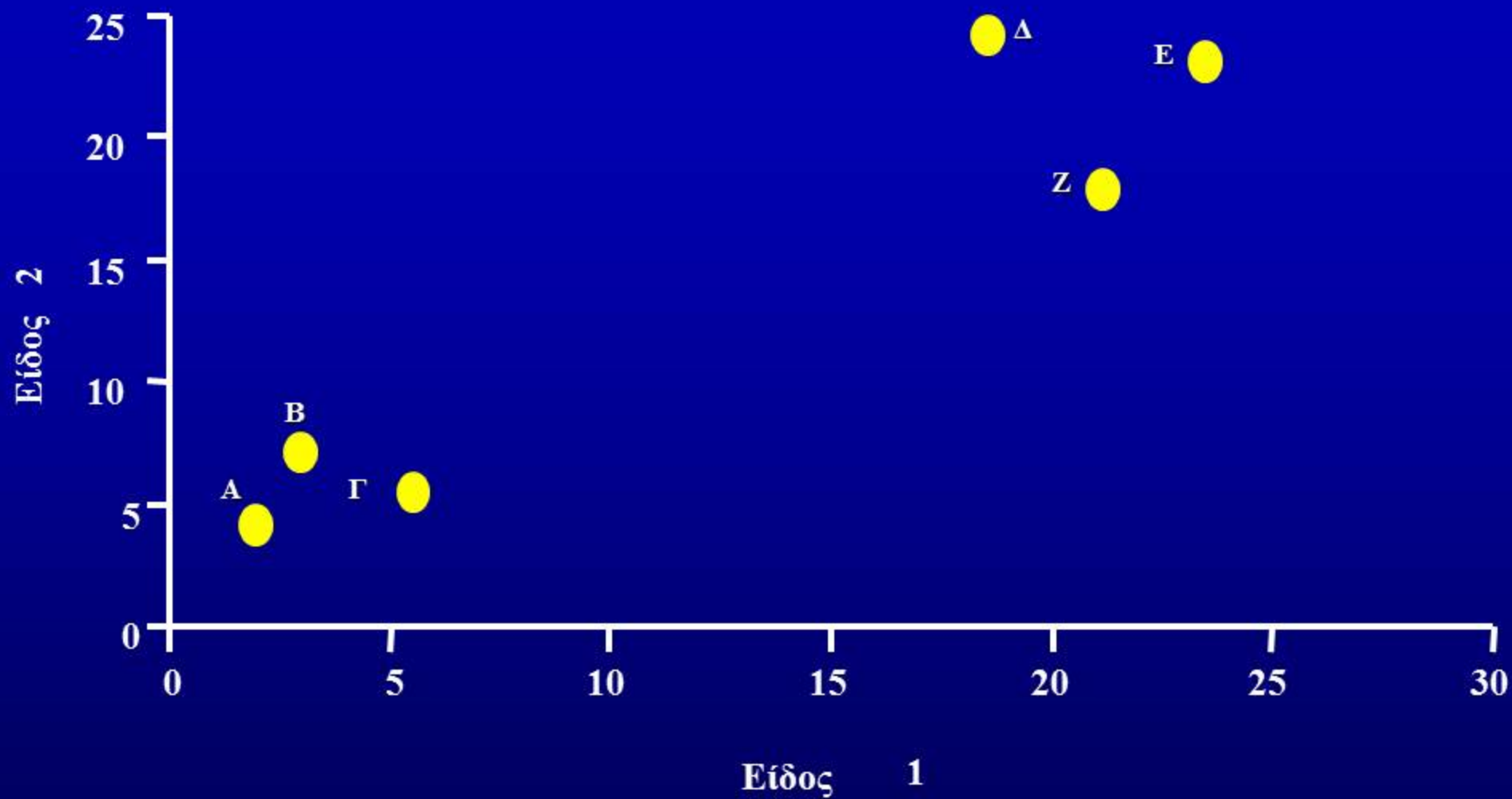
$$GDD_{jk} = \theta_{jk} \quad 0 \leq GDD \leq 1,576$$

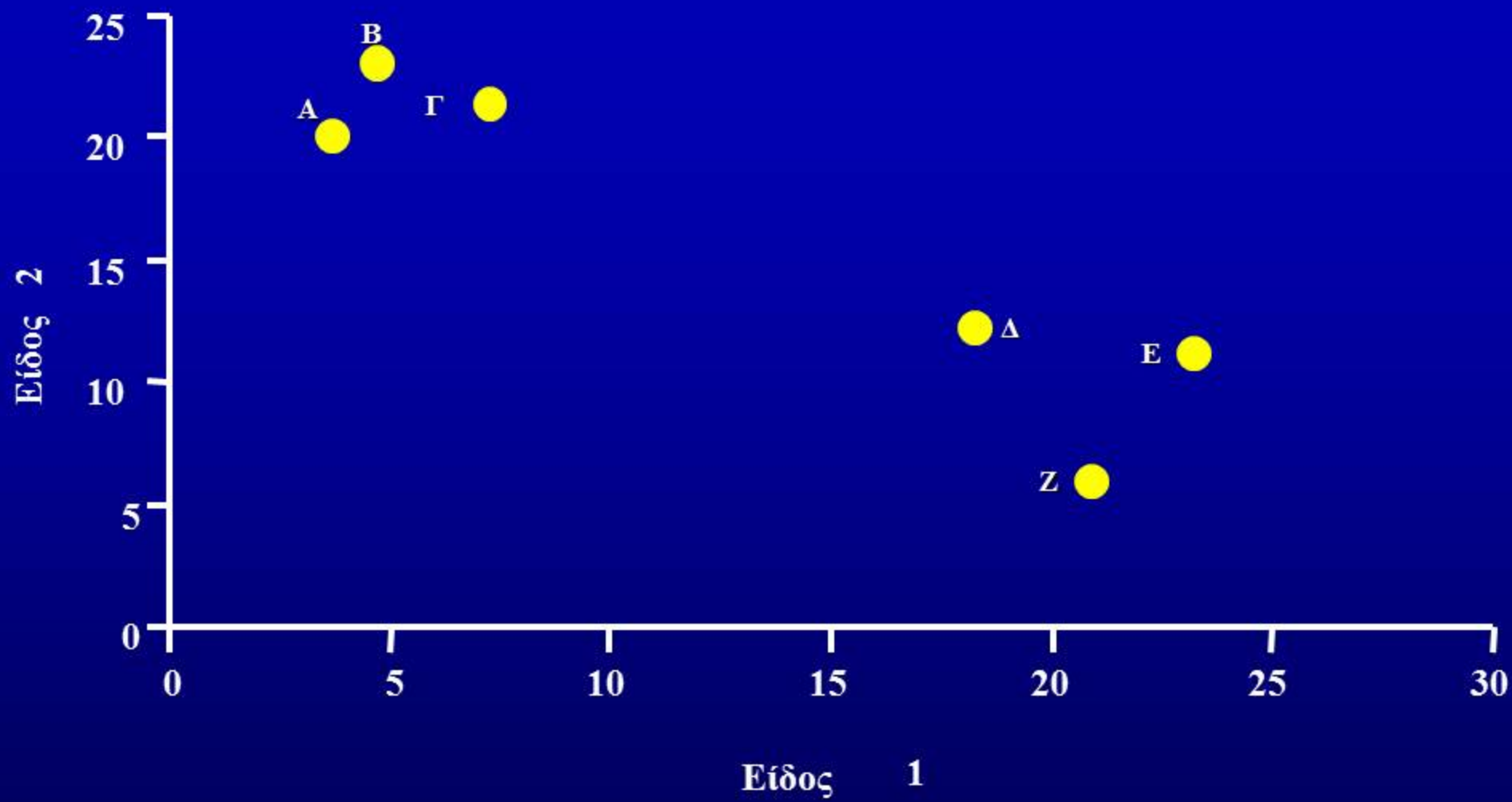
Είναι εύκολο να δούμε ότι για δύο τελείως διαφορετικά (j και k) δείγματα (χωρίς κανένα είδος κοινό και στα δύο) το $\text{συν}\theta_{jk} = 0$ άρα $\theta_{jk} = 1,576 = \frac{\pi}{2}$. Για δύο τελείως όμοια δείγματα ο τύπος δίνει $\text{συν}\theta_{jk} = 1$, άρα $\theta_{jk} = 0$.

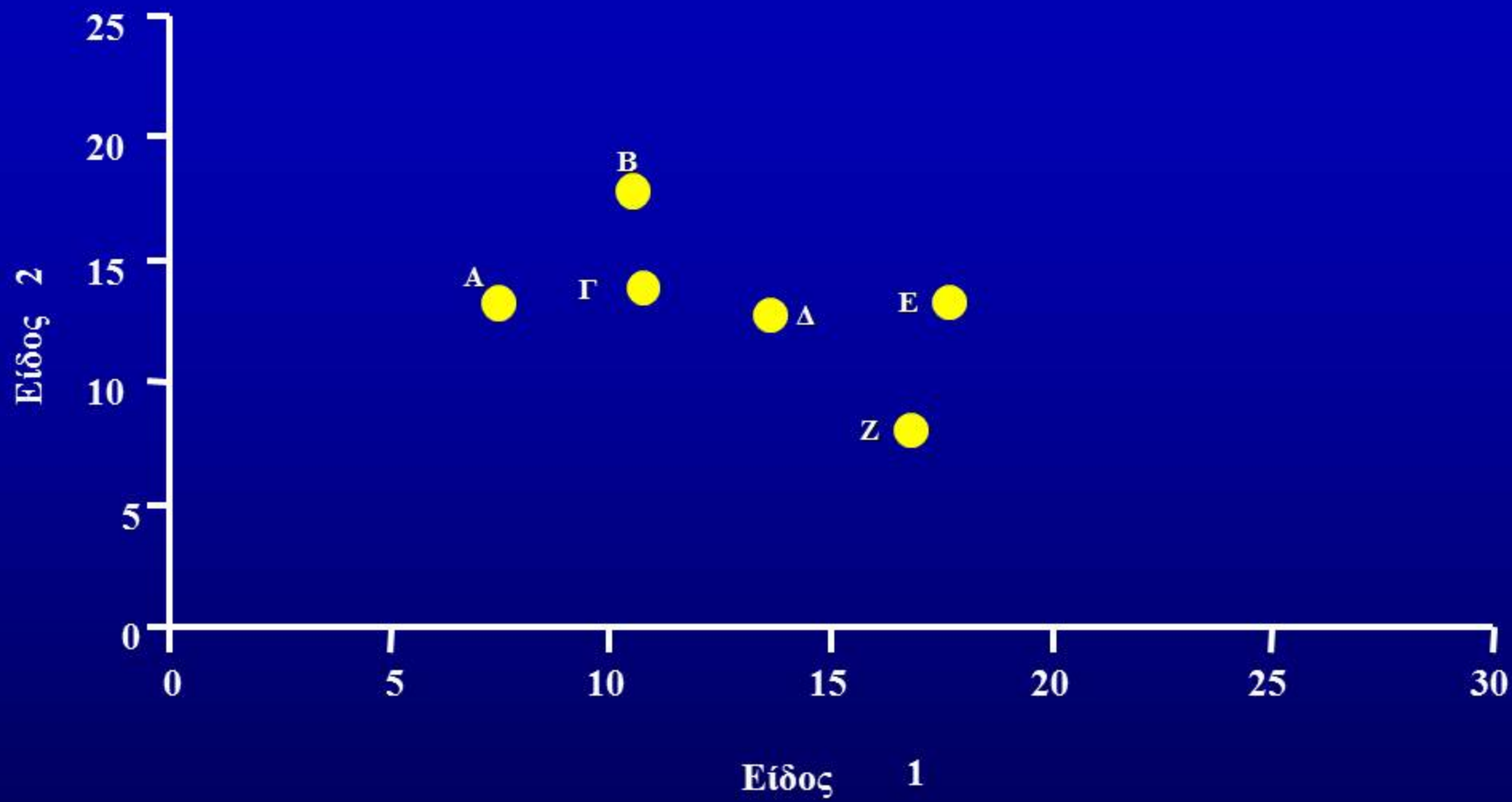
	Δείγμα J	Δείγμα k	Σύνολο
Είδος 1	5	10	15
Είδος 2	6	9	15
Σύνολο	11	19	30



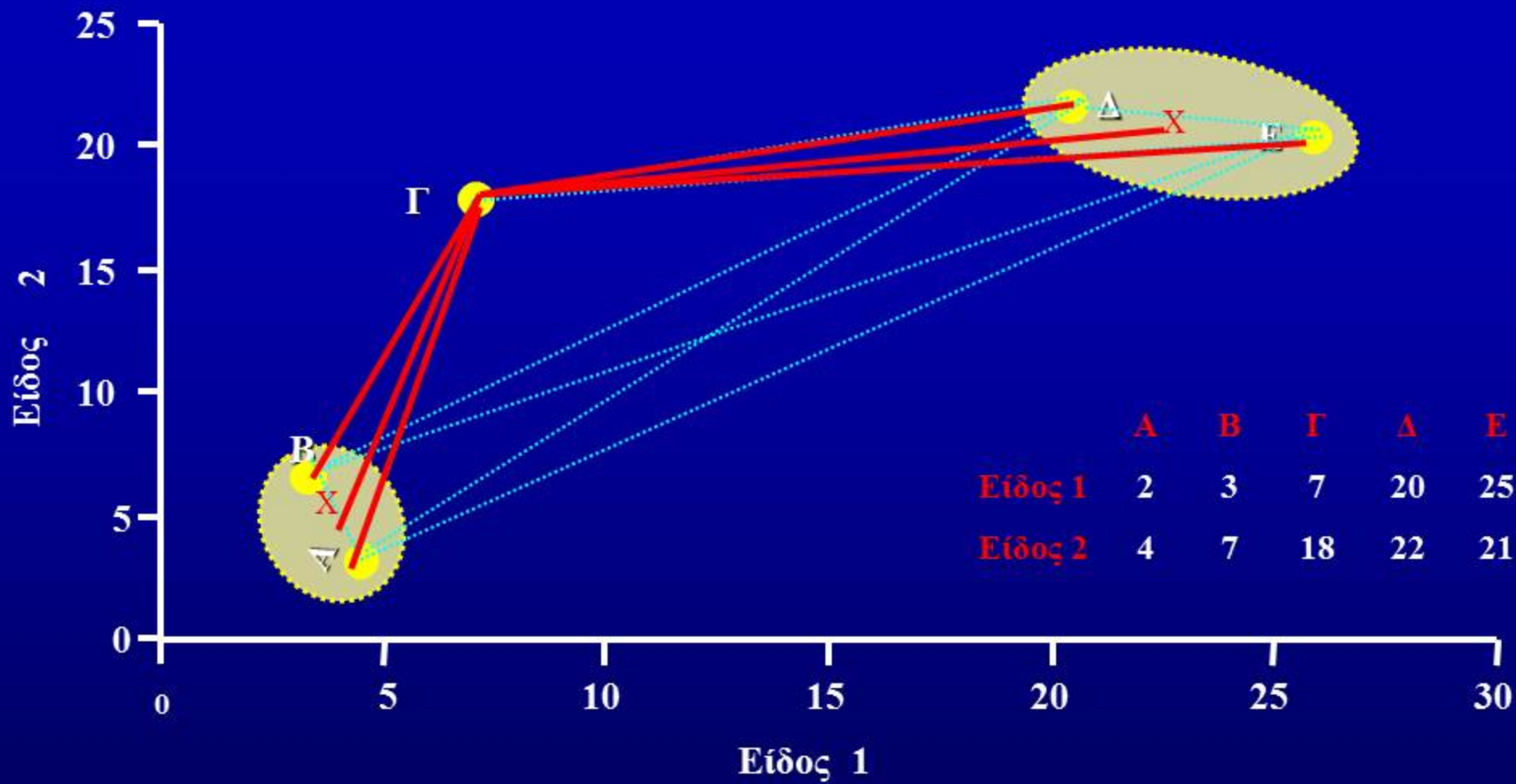
$$ED_{jk} = \sqrt{(5 - 10)^2 + (6 - 9)^2} = 5,8$$







	A	B	Γ	Δ	E
Είδος 1	2	3	7	20	25
Είδος 2	4	7	18	22	21



	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
<i>A</i>	1	0.074	0.113	0.095	0.103
<i>B</i>		1	0.136	0.144	0.117
<i>C</i>			1	0.135	0.119
<i>D</i>				1	0.068
<i>E</i>					1

Συντελεστής ομοιότητας (Bray-Curtis)

	<i>A</i>	<i>BD</i>	<i>C</i>	<i>E</i>
<i>A</i>	1			
<i>BD</i>		1		
<i>C</i>			1	
<i>E</i>				1

$$PS(BD : A) = \frac{PS(B : A) + PS(D : A)}{2}$$

$$PS(BD : C) = \frac{PS(B : C) + PS(D : C)}{2}$$

$$PS(BD : E) = \frac{PS(B : E) + PS(D : E)}{2}$$

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
<i>A</i>	1	0.074	0.113	0.095	0.103
<i>B</i>		1	0.136	0.144	0.117
<i>C</i>			1	0.135	0.119
<i>D</i>				1	0.068
<i>E</i>					1

$$\text{PS}(\mathbf{BD} : \mathbf{A}) = \frac{\text{PS}(\mathbf{B} : \mathbf{A}) + \text{PS}(\mathbf{D} : \mathbf{A})}{2} = \frac{0.074 + 0.095}{2} = 0.0845$$

$$\text{PS}(\mathbf{BD} : \mathbf{C}) = \frac{\text{PS}(\mathbf{B} : \mathbf{C}) + \text{PS}(\mathbf{D} : \mathbf{C})}{2} = \frac{0.136 + 0.135}{2} = 0.1355$$

$$\text{PS}(\mathbf{BD} : \mathbf{E}) = \frac{\text{PS}(\mathbf{B} : \mathbf{E}) + \text{PS}(\mathbf{D} : \mathbf{E})}{2} = \frac{0.117 + 0.068}{2} = 0.0925$$

Συντελεστής ομοιότητας (Bray-Curtis)

	<i>A</i>	<i>BD</i>	<i>C</i>	<i>E</i>
<i>A</i>	1	0.0845	0.113	0.103
<i>BD</i>		1	0.1355	0.0925
<i>C</i>			1	0.119
<i>E</i>				1

Συντελεστής ομοιότητας (Bray-Curtis)

	<i>A</i>	<i>BD</i>	<i>C</i>	<i>E</i>
<i>A</i>	1	0.0845	0.113	0.103
<i>BD</i>		1	0.1355	0.925
<i>C</i>			1	0.119
<i>E</i>				1

$$PS(BCD : A) = \frac{PS(B : A) + PS(C : A) + PS(D : A)}{3} = \frac{0.074 + 0.113 + 0.095}{3} = 0.094$$

$$PS(BCD : E) = \frac{PS(B : E) + PS(C : E) + PS(D : E)}{3} = 0.101$$

Συντελεστής ομοιότητας (Bray-Curtis)

	<i>A</i>	<i>BCD</i>	<i>E</i>
<i>A</i>	1	0.094	0.103
<i>BCD</i>		1	0.101
<i>E</i>			1

$$PS(BCD : A) = \frac{PS(B : A) + PS(C : A) + PS(D : A)}{3} = \frac{0.074 + 0.113 + 0.095}{3} = 0.094$$

$$PS(BCD : E) = \frac{PS(B : E) + PS(C : E) + PS(D : E)}{3} = 0.101$$

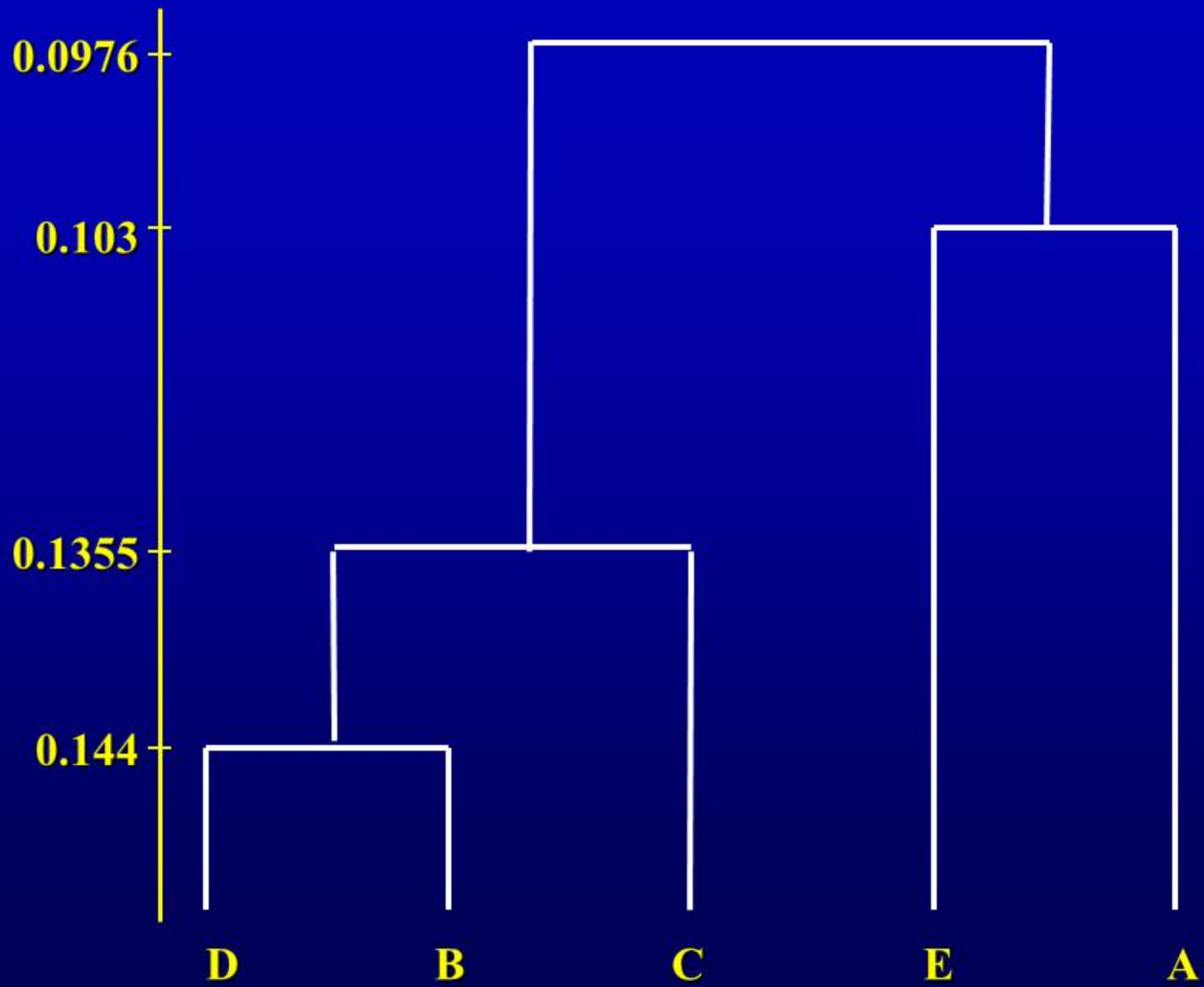
Συντελεστής ομοιότητας (Bray-Curtis)

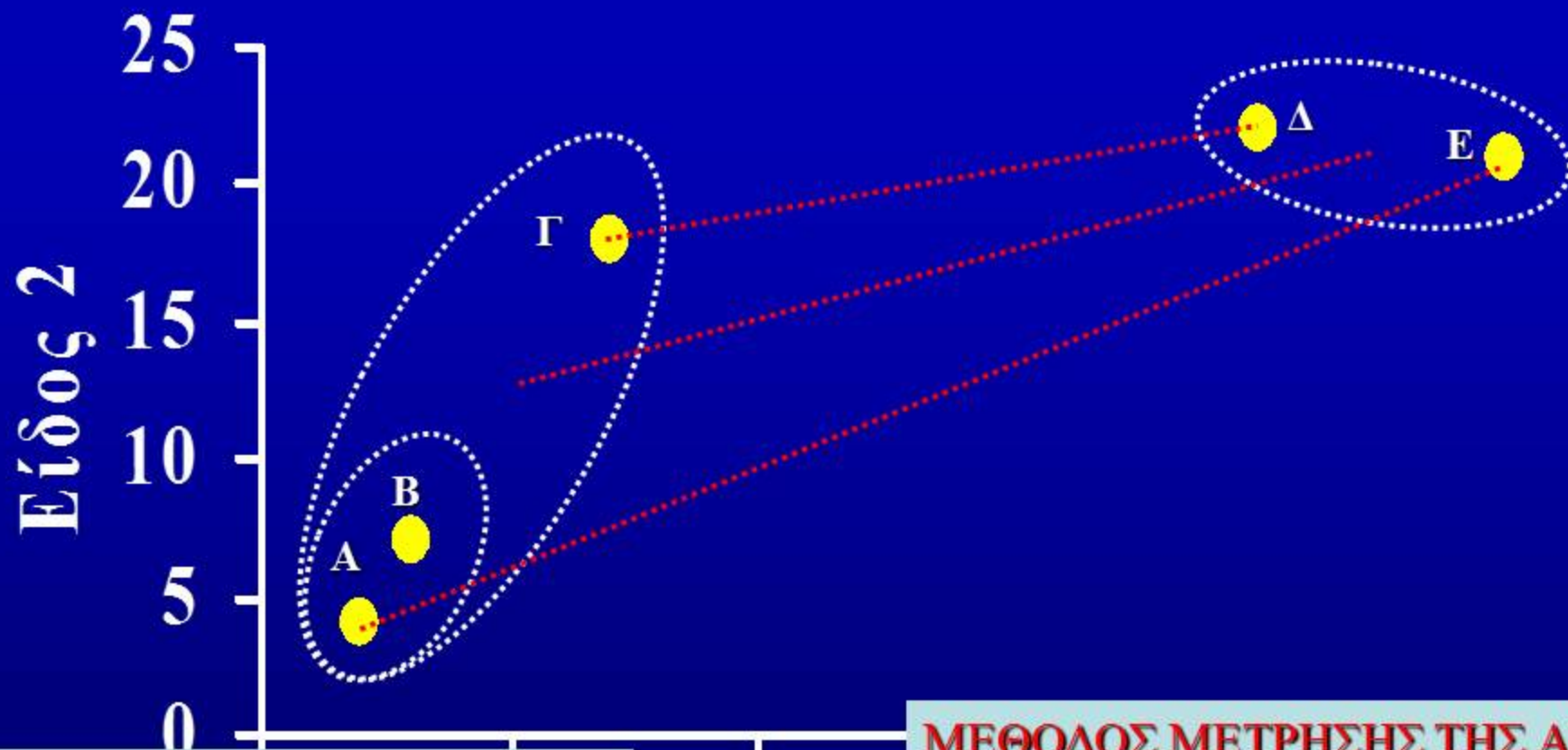
	<i>A</i>	<i>BCD</i>	<i>E</i>
<i>A</i>	1	0.094	0.103
<i>BCD</i>		1	0.101
<i>E</i>			1

$$PS(AE : BCD) = \frac{PS(A : B) + PS(A : C) + PS(A : D) + PS(E : B) + PS(E : C) + PS(E : D)}{6} = 0.0976$$

Συντελεστής ομοιότητας (Bray-Curtis)

	<i>AE</i>	<i>BCD</i>
<i>AE</i>	1	0.0976
<i>BCD</i>		1





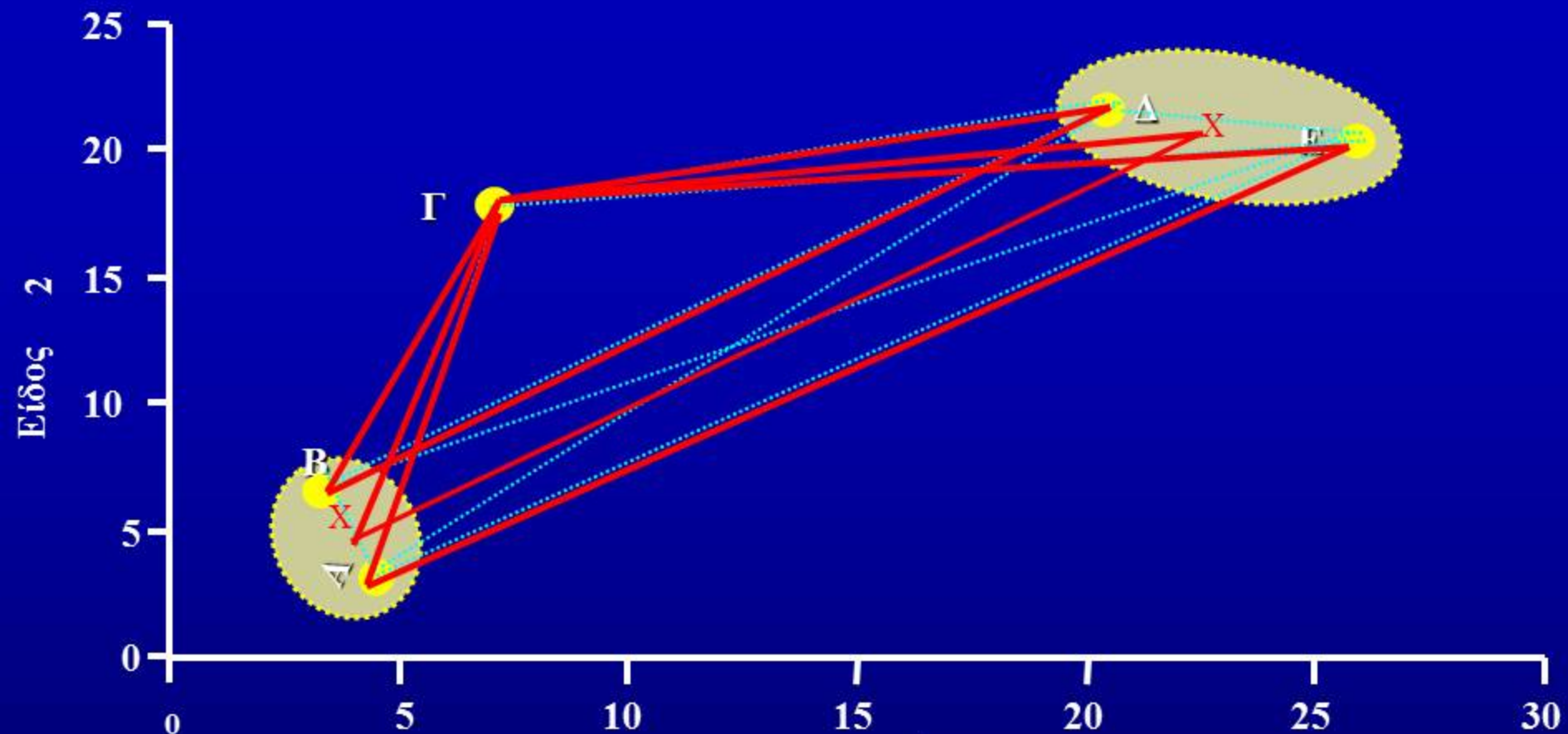
ΜΕΤΡΟ ΑΠΟΣΤΑΣΗΣ

- Euclidean distance.
- Squared Euclidean distance.
- City-block (Manhattan) distance.
- Chebyshev distance.
- Power distance.
- Percent disagreement.

10

ΜΕΘΟΔΟΣ ΜΕΤΡΗΣΗΣ ΤΗΣ ΑΠΟΣΤΑΣΗΣ

- Single linkage (nearest neighbor).
- Complete linkage (furthest neighbor).
- Unweighted pair-group average.
- Weighted pair-group average.
- Unweighted pair-group centroid.
- Weighted pair-group centroid (median).
- Ward's method.



ΜΕΤΡΟ ΑΠΟΣΤΑΣΗΣ

Bray-Curtis

Euclidean distance.

Squared Euclidean distance.

City-block (Manhattan) distance.

Chebychev distance.

Power distance.

Percent disagreement.

Είδος 1

ΜΕΘΟΔΟΣ ΜΕΤΡΗΣΗΣ ΤΗΣ ΑΠΟΣΤΑΣΗΣ

Single linkage (nearest neighbor).

Complete linkage (furthest neighbor).

Unweighted pair-group average.

Weighted pair-group average.

Unweighted pair-group centroid.

Weighted pair-group centroid (median).

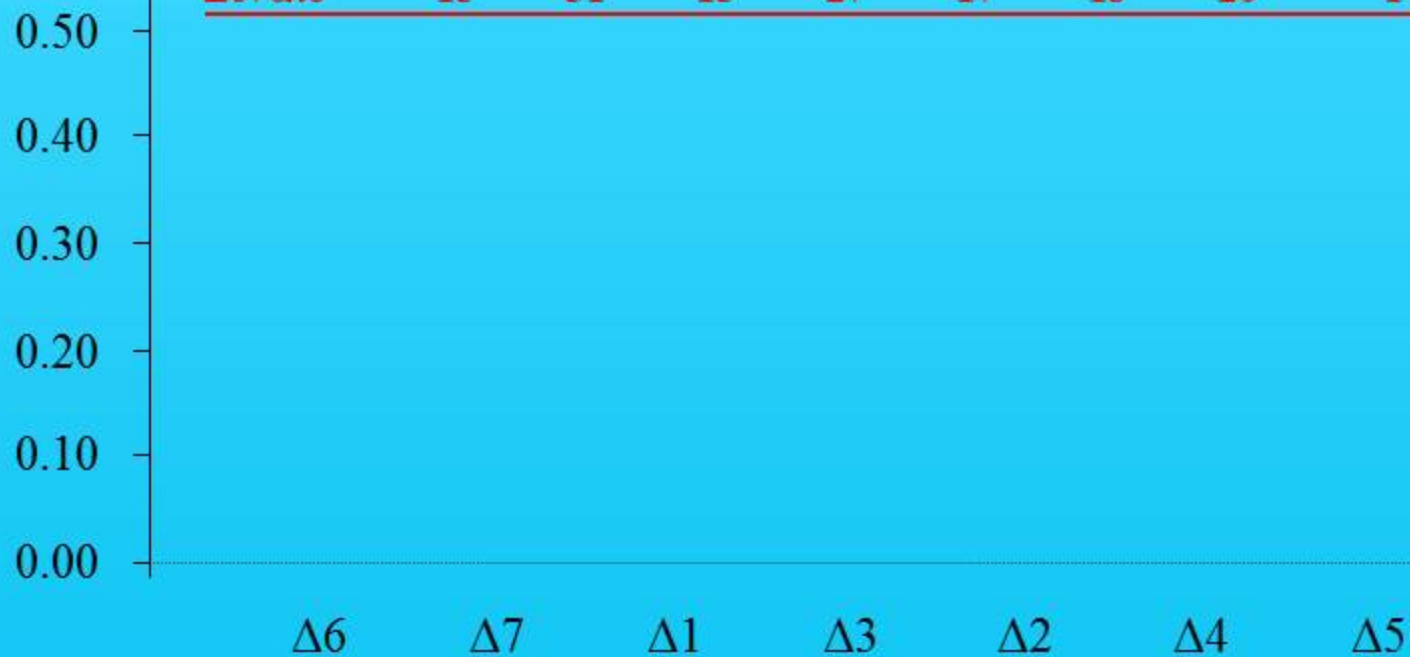
Ward's method.

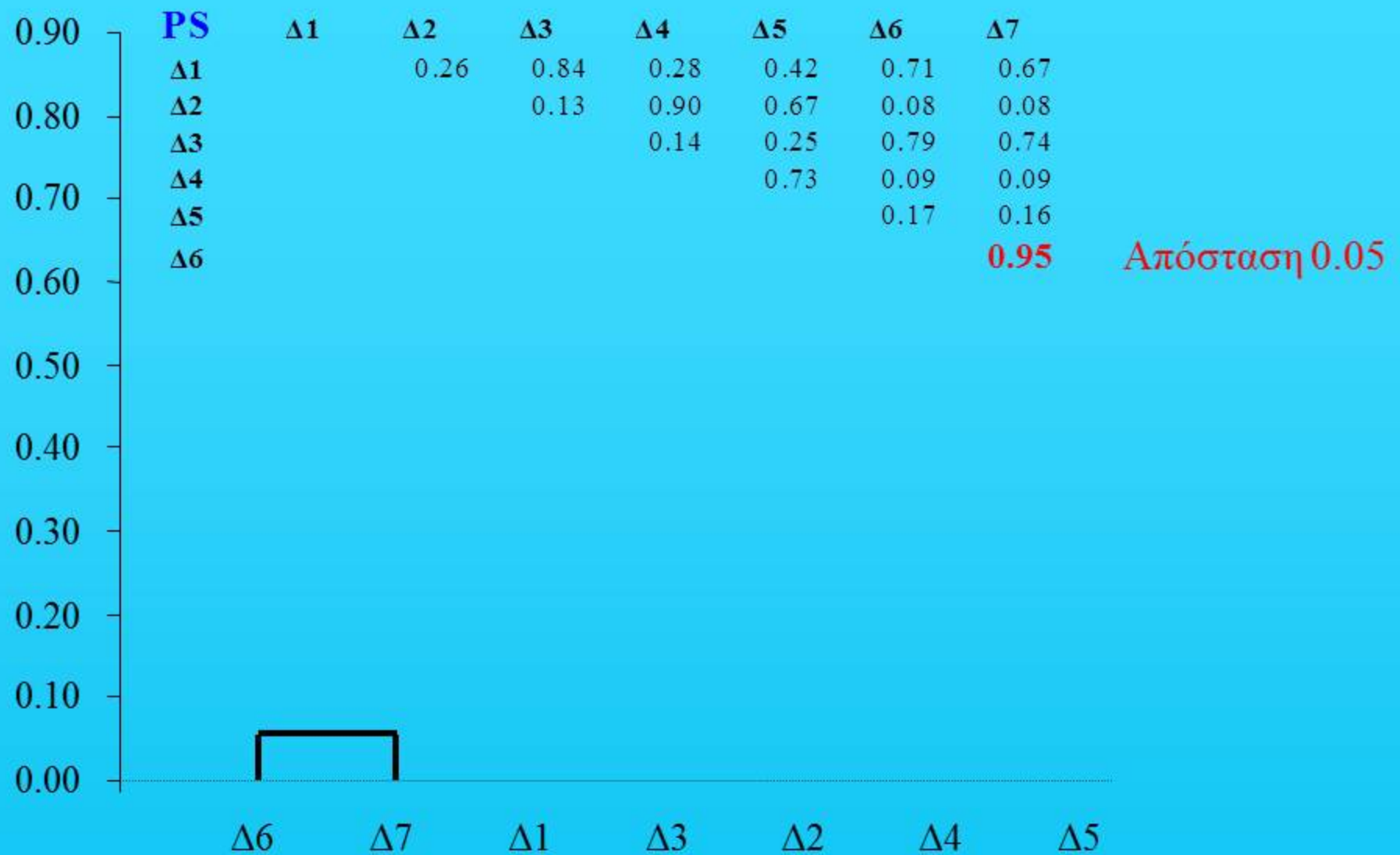
Δείγματα

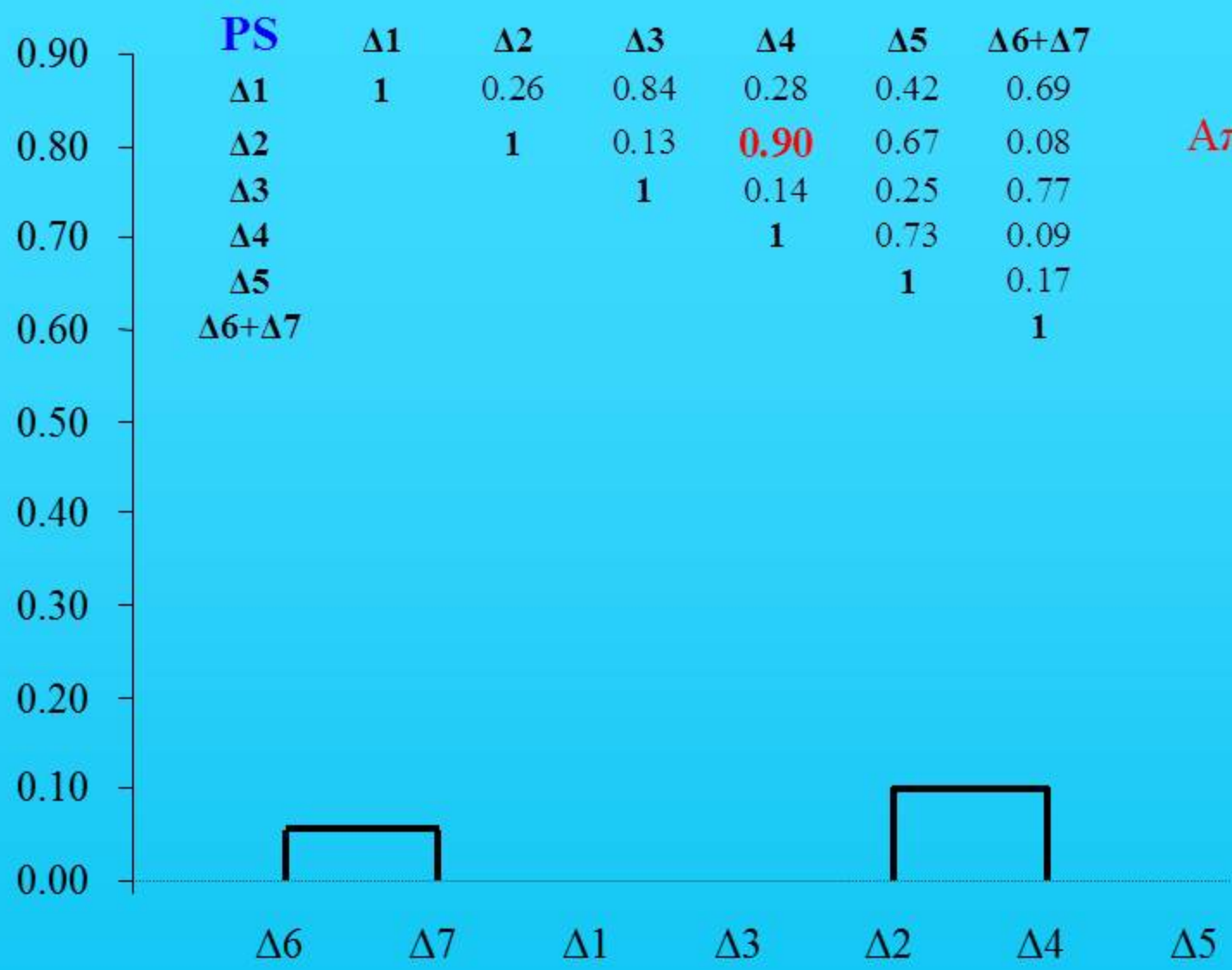
Είδη	Δ1	Δ2	Δ3	Δ4	Δ5	Δ6	Δ7	Σύνολο
A	5	0	6	1	1	8	9	30
B	0	10	1	9	3	0	0	23
Γ	6	1	7	0	1	9	10	34
Δ	1	8	0	7	4	0	0	20
E	2	7	1	6	5	0	0	21
Z	1	5	0	4	3	1	1	15
Σύνολο	15	31	15	27	17	18	20	143

Δείγματα

Είδη	Δ1	Δ2	Δ3	Δ4	Δ5	Δ6	Δ7	Σύνολο
A	5	0	6	1	1	8	9	30
B	0	10	1	9	3	0	0	23
Γ	6	1	7	0	1	9	10	34
Δ	1	8	0	7	4	0	0	20
E	2	7	1	6	5	0	0	21
Z	1	5	0	4	3	1	1	15
Σύνολο	15	31	15	27	17	18	20	143



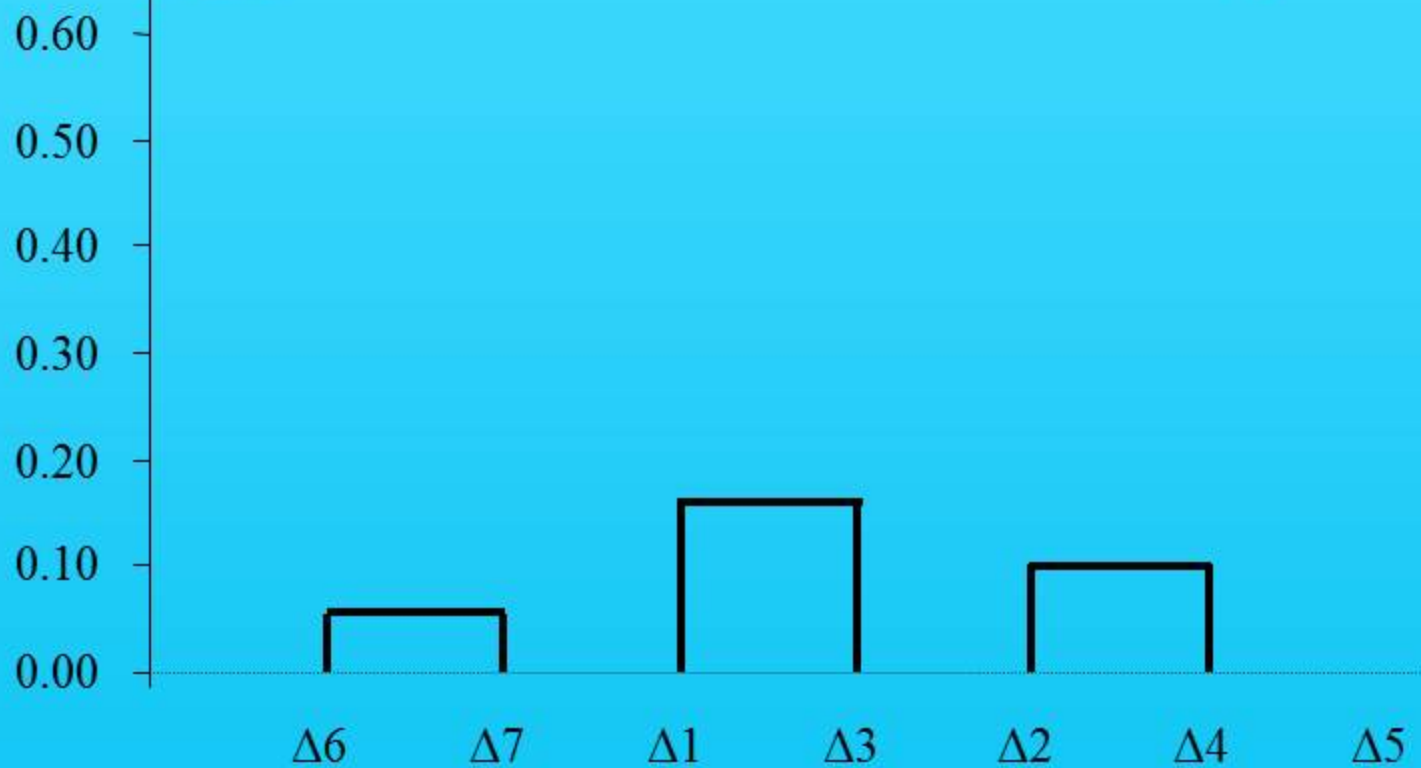


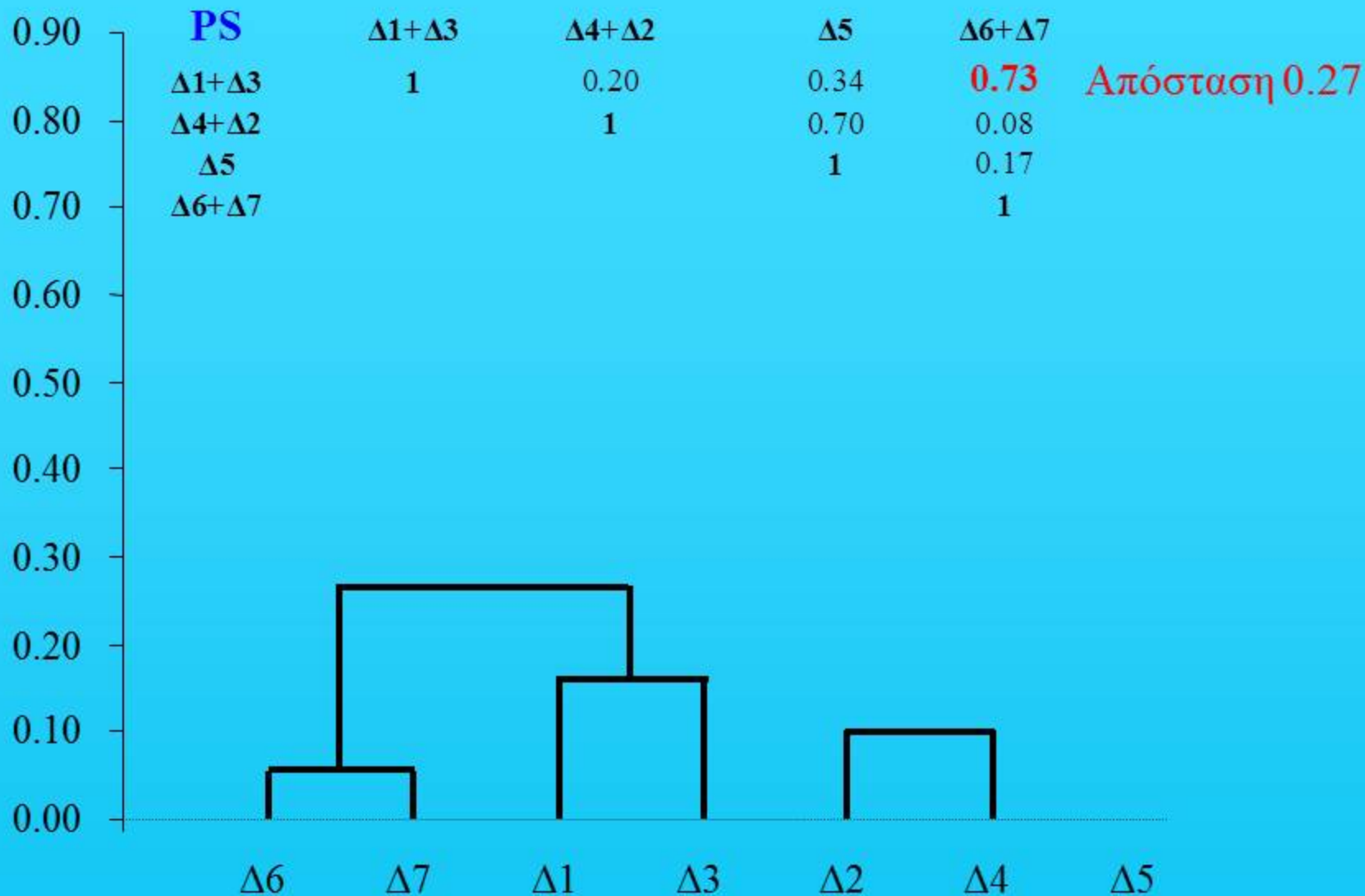


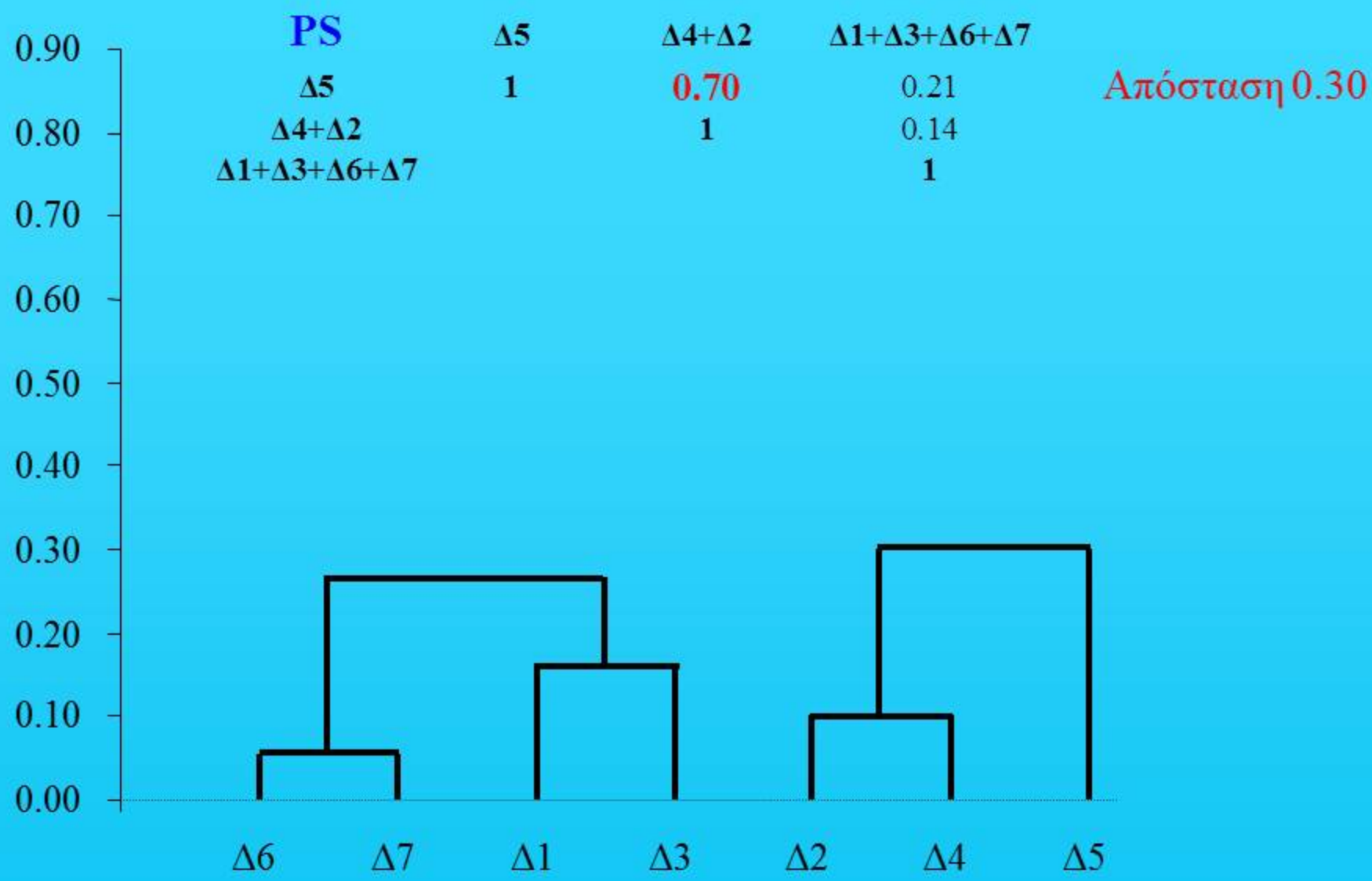
Απόσταση 0.10

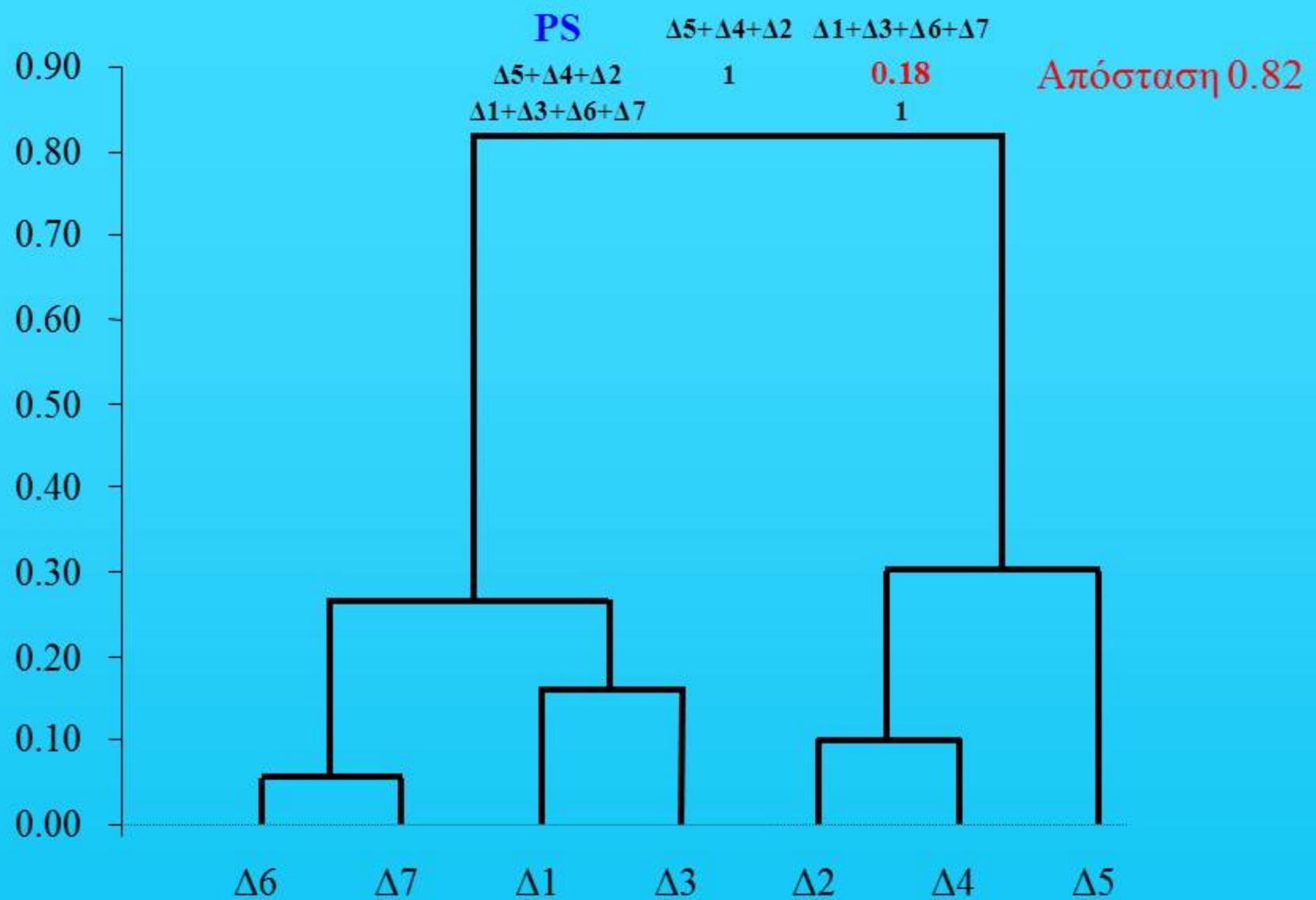
PS	$\Delta 1$	$\Delta 3$	$\Delta 4+\Delta 2$	$\Delta 5$	$\Delta 6+\Delta 7$
$\Delta 1$	1	0.84	0.27	0.42	0.69
$\Delta 3$		1	0.14	0.25	0.77
$\Delta 4+\Delta 2$			1	0.70	0.08
$\Delta 5$				1	0.17
$\Delta 6+\Delta 7$					1

Απόσταση 0.16

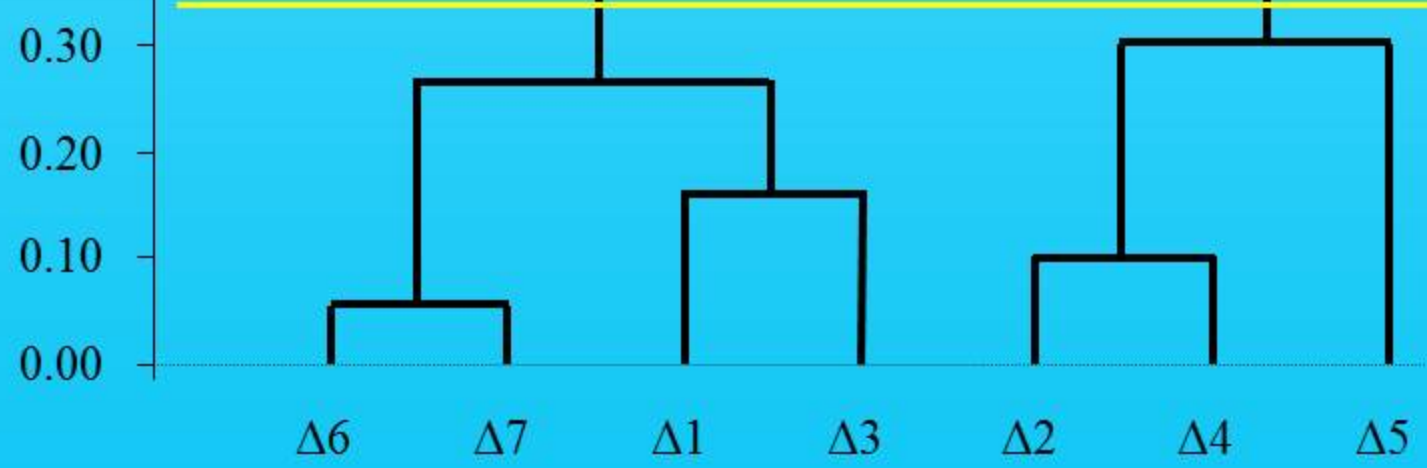




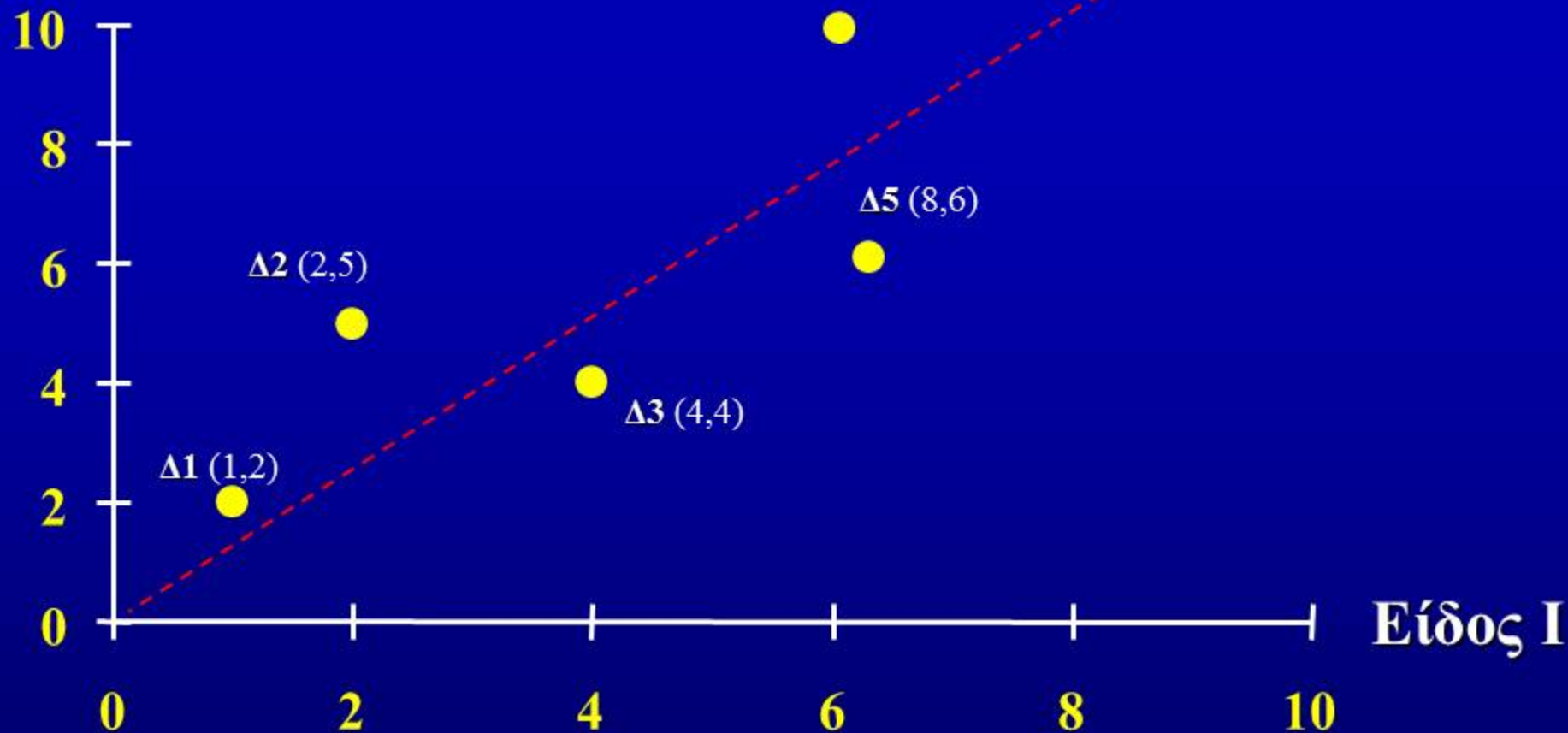




Είδη	Δ2	Δ4	Δ5	Δ1	Δ3	Δ6	Δ7	Σύνολο
B	10	9	3	0	1	0	0	23
Δ	8	7	4	1	0	0	0	20
E	7	6	5	2	1	0	0	21
Z	5	4	3	1	0	1	1	15
A	0	1	1	5	6	8	9	30
Γ	1	0	1	6	7	9	10	34
Σύνολο	31	27	17	15	15	18	20	143

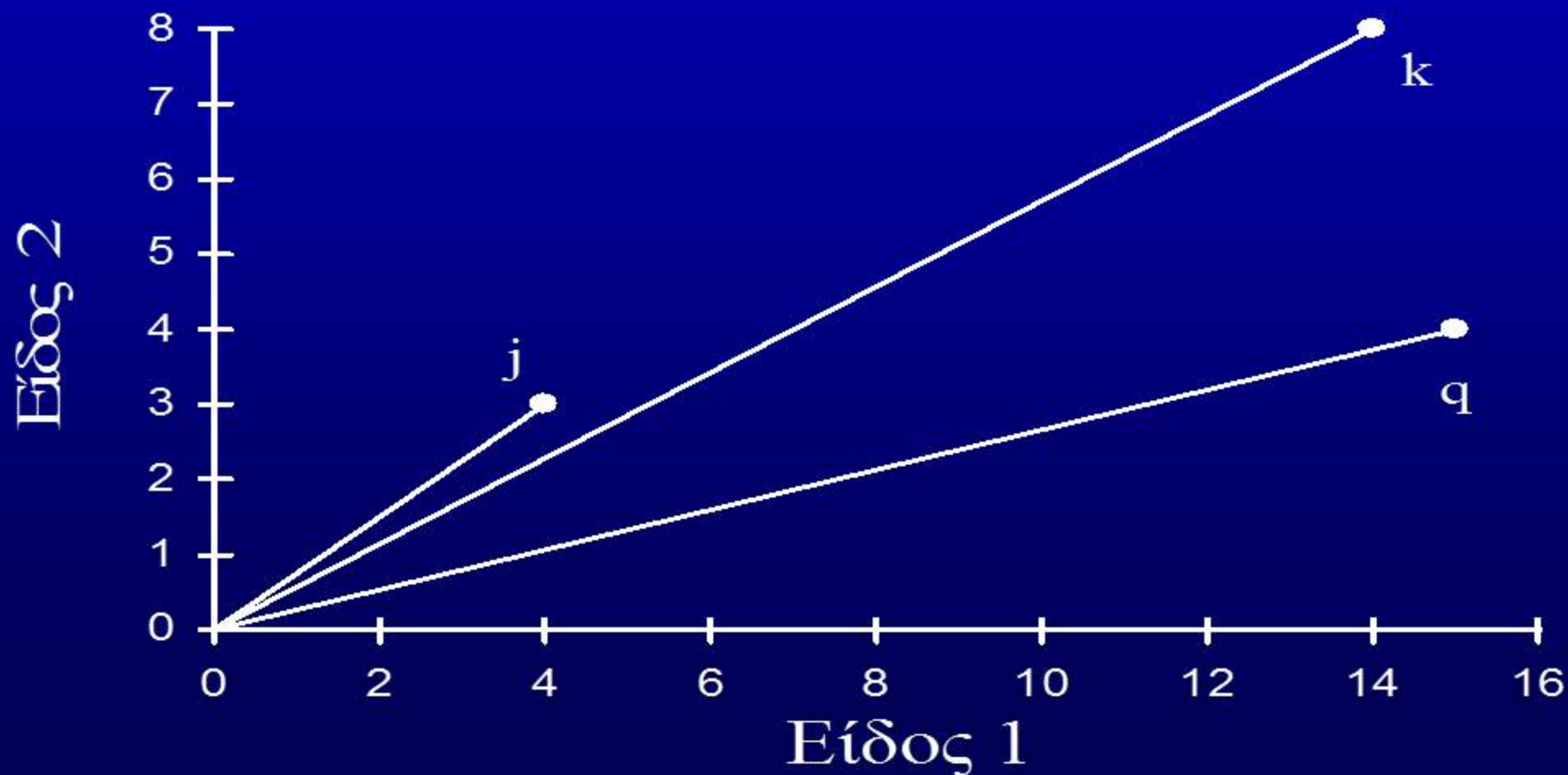


Είδος II



Ιεράρχηση με βάση τον είδος I:	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 4$	$\Delta 5$
Ιεράρχηση με βάση τον είδος II:	$\Delta 1$	$\Delta 3$	$\Delta 2$	$\Delta 5$	$\Delta 4$
Ιεράρχηση με βάση τον Άξονα III:	$\Delta 1$	$\Delta 2$	$\Delta 3$	$\Delta 5$	$\Delta 4$

	Δείγμα J	Δείγμα k	Δείγμα q	Σύνολο
Είδος 1	4	14	15	33
Είδος 2	3	8	4	15
Σύνολο	7	22	19	48



Δείκτες	Ζεύγη Δειγμάτων		
	<i>jk</i>	<i>jq</i>	<i>kq</i>
ED	11,18	11,05	4,12
AD	15,00	12,00	5,00
GDD	0,12	0,38	0,26
RED	0,09	0,31	0,22
RAD	0,13	0,44	0,31

Η σημαντική παρατήρηση στο παράδειγμα αυτό είναι ότι οι δείκτες ED και PD συμπεριφέρονται παρόμοια. Από το άλλο μέρος, ο δείκτης GDD και οι σχετικές αποστάσεις RED και RAD συμπεριφέρονται επίσης παρόμοια γιατί ενσωματώνουν κάποια στάθμιση των δεδομένων.