

Η διάλεξη αυτή βασίζεται σε ενότητες του Κεφαλαίου 2 του Βαγενάς (2019).  
Ωστόσο, οι φοιτητές/-τριες μπορούν να ανατρέχουν και σε άλλα προτεινόμενα συγγράμματα

# ΚΑΤΑΝΟΜΕΣ ΣΥΧΝΟΤΗΤΑΣ



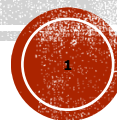
**Αντώνης Κ. Τραυλός (B.A., M.A., Ph.D.)**

**Καθηγητής**

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ**

**Σχολή Επιστημών Ανθρώπινης Κίνησης και Ποιότητας Ζωής**

**Τμήμα Οργάνωσης και Διαχείρισης Αθλητισμού**



1

## Αρχικά Δεδομένα, Περιγραφική Στατιστική και Κατανομές Συχνότητας

Στο κεφάλαιο αυτό γίνεται αναλυτική περιγραφή της ανάπτυξης των αρχικών αριθμητικών δεδομένων (raw data) μιας επιστημονικής έρευνας σε **απλή** και σε **ομαδοποιημένη κατανομή συχνότητων**.

Επιπρόσθετα, επιχειρείται η λεπτομερής περιγραφή των αριθμητικών αναλύσεων στις οποίες μπορεί να υποβληθεί μια κατανομή.

Ειδική έμφαση δίνεται στις εφαρμογές των κατανομών συχνότητας στον τομέα της περιγραφικής στατιστικής.

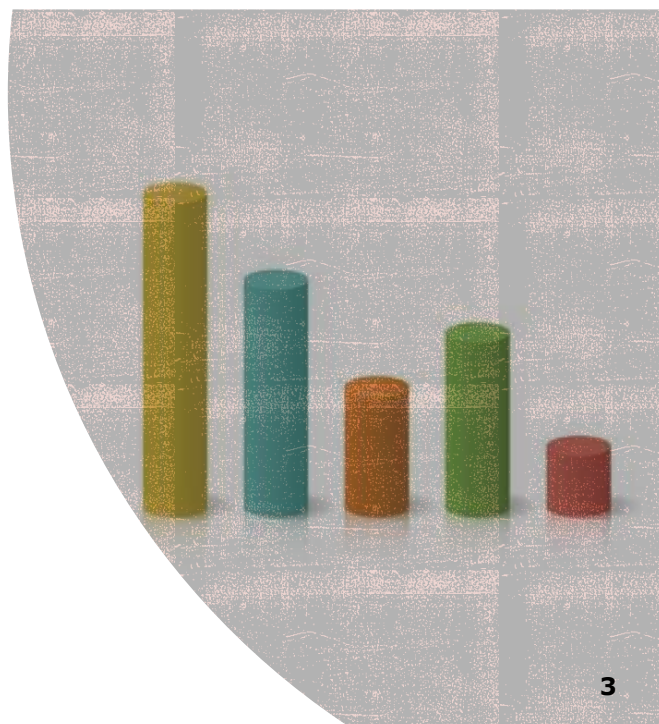


2

## Οργάνωση Αρχικών Δεδομένων

- Το πρώτο βήμα μετά τη συλλογή των αρχικών δεδομένων μιας έρευνας είναι η συστηματική επισκόπησή τους προκειμένου
  - (α) να εντοπιστούν τυχόν σφάλματα καταγραφής,
  - (β) να εντοπιστούν τυχόν ακραίες τιμές και
  - (γ) να αποκτηθεί μια πρώτη εικόνα της κατανομής τους.

Αυτό μπορεί να γίνει με την οργάνωση των δεδομένων ανάλογα με το πλήθος τους (N).



3

Για παράδειγμα, ας πούμε ότι 10 μαθητές αξιολογήθηκαν στο προτιμώμενο άθλημα (βόλεϊ, μπάσκετ, τένις), στον βαθμό γυμναστικής (1, 2, 3, 4, 5) και στη ρίψη (σε 0.1 m), και έδωσαν τα δεδομένα του πίνακα 2.1α.

**Πίνακας 2.1α** - Αρχικά δεδομένα των μεταβλητών (N=10)

Μαθητές	Άθλημα (προτίμηση)	Γυμναστική (βαθμός)	Ρίψη (m)
1	Βόλεϊ	2	20.2
2	Μπάσκετ	1	19.1
3	Μπάσκετ	3	21.0
4	Βόλεϊ	3	22.7
5	Τένις	4	21.3
6	Μπάσκετ	3	23.3
7	Τένις	5	24.4
8	Τένις	5	23.6
9	Τένις	5	20.9
10	Μπάσκετ	2	22.8

4

4

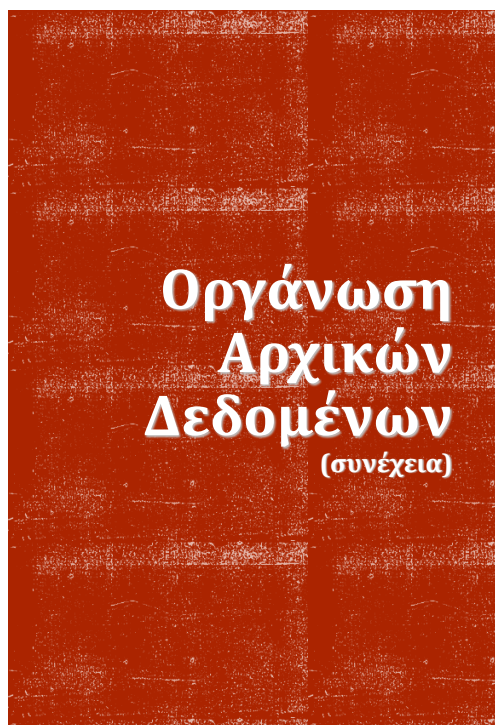
Πίνακας 2.1B - Κατανομή των συχνοτήτων (N=10)

Μεταβλητή	Βαθμίδες σε σειρά	Μαθητές	
		Συχνότητα f	%f
Άθλημα	Βόλει	2	20
	Μπάσκετ	4	40
	Τένις	4	40
Γυμναστική (βαθμός)	1	1	10
	2	2	20
	3	3	30
	4	1	10
	5	3	30
Ρίψη Μπάλας (m)	18.5 - 19.5	1	10
	19.5 - 20.5	1	10
	20.5 - 21.5	3	30
	21.5 - 22.5	-	-
	22.5 - 23.5	3	30
	23.5 - 24.5	2	20

$$\% f = (f / N) * 100 = (\text{συχνότητα} / \text{μέγεθος δείγματος}) * 100$$

Από την ταξινόμηση των τιμών βγήκαν οι συχνότητες (frequencies), δηλαδή πόσοι μαθητές ανήκουν σε κάθε βαθμίδα τής κάθε μεταβλητής:

5



- Όταν τα αρχικά δεδομένα αποτελούνται από λίγες αριθμητικές τιμές (π.χ.  $N < 30$ ) με ή χωρίς επαναλήψεις, τότε αναφέρονται ως **ομάδα αριθμών** ή, γενικότερα, ως **ομάδα τιμών**.
- Η οργάνωση μικρών ομάδων αριθμών ξεκινάει συνήθως από την τοποθέτησή τους σε σειρά μεγέθους, οπότε συνθέτουν συνολικά **σειρά αριθμών** ή, γενικότερα, **σειρά τιμών**.
- Για παράδειγμα,
  - οι αριθμοί 5, 8, 3, 1, 9 αποτελούν **μια απλή ομάδα αριθμών**,
  - που τοποθετούμενοι σε σειρά μεγέθους,
    - δίνουν τη **σειρά αριθμών** 1, 3, 5, 8, 9.

6

6



Η **ομαδική οργάνωση** των αρχικών δεδομένων μιας μέτρησης αποκτά ουσιαστικότερη σημασία, όταν πρόκειται για δεδομένα μεγάλου πλήθους (π.χ.  $N > 30$ ), στα οποία υπάρχουν επαναλήψεις των τιμών.

Στις περιπτώσεις αυτές οι ομάδες των δεδομένων ονομάζονται **κατανομές** και η οργάνωσή τους αρχίζει με τους θεμελιώδεις για την περιγραφική στατιστική **συχνοτικούς πίνακες**.

7

- Οι κατανομές συχνοτήτων αφορούν δεδομένα ποσοτικών μεταβλητών που συνήθως συμβολίζονται με τα λατινικά κεφαλαία γράμματα  $X, Y, Z$ .
- Έτσι, **οι διαφορετικές τιμές** μιας κατανομής συμβολίζονται συνήθως με τα γράμματα  $X_1, X_2, \dots, X_k$  ή  $Y_1, Y_2, \dots, Y_k$  ή  $Z_1, Z_2, \dots, Z_k$  ενώ οι επαναλήψεις τους, δηλαδή το πόσες φορές εμφανίζεται κάθε διαφορετική τιμή στην κατανομή, δηλαδή οι **συχνότητες**, συμβολίζονται με τα γράμματα  $f_1, f_2, \dots, f_k$  αντίστοιχα ( $k$  είναι το πλήθος των διαφορετικών τιμών ή απλά ο αριθμός των κλάσεων που προκύπτουν).



8

8

Age group (Class)	Frequency	Cumulative frequency	% Cumulative frequency	Relative Frequency
20-24	2	2	4.0	0.04
25-29	3	5	10.0	0.06
30-34	13	18	36.0	0.26
35-39	18	36	72.0	0.36
40-44	8	44	88.0	0.16
45-49	4	48	96.0	0.08
50-54	2	50	100.0	0.04
Total	50			

Όπως θα διαπιστώσουμε στα επόμενα κεφάλαια, στην περιγραφική στατιστική γίνεται ευρεία χρήση κατανομών, που για λόγους συντομίας λέγονται απλά **συχνοτικές κατανομές**. Ο τρόπος οργάνωσης (ομαδοποίησης) των κατανομών τις διαχωρίζει σε δύο γενικές κατηγορίες:

- **Απλές κατανομές**, στις οποίες εντοπίζονται οι διαφορετικές τιμές, τοποθετούνται σε σειρά μεγέθους και καταγράφονται οι συχνότητες τους.
- **Ομαδοποιημένες κατανομές**, στις οποίες το πλήθος  $N$  των αριθμών είναι συνήθως τόσο μεγάλο που επιτρέπει τη δημιουργία ομάδων (διαστημάτων) τιμών και την καταγραφή, ως συχνοτήτων, του επί μέρους πλήθους των τιμών που "πέφτουν" σε κάθε τέτοια ομάδα αντίστοιχα.

9

9

Στον πίνακα 2.2 δίνονται τα 2 τυπικά παραδείγματα κατανομής:

1. **1° - Απλή κατανομή.** Στα 45 δεδομένα 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7 υπάρχουν οι  $k=7$  διαφορετικές βαθμίδες 1, 2, 3, 4, 5, 6, 7, με αντίστοιχες συχνότητες ( $f$ ) 4, 8, 7, 10, 6, 6, 4. Οι κατανομές αυτές είναι συνήθεις σε έρευνες με ερωτηματολόγια, όπου οι κλίμακες μέτρησης είναι συνήθως 5-βάθμιες έως 9-βάθμιες Likert. (Likert, 1932).
2. **2° - Ομαδοποιημένη κατανομή.** Στα 121 δεδομένα 1,2, 2, 3, 3, 3, 4, 4,4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 7, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 9, 9, 9, 9, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10, 10, 10, 10, 10, 11, 11, 11, 11, 11, 11, 11, 11, 11, 11, 11, 11, 11, 11, 12, 12, 12, 12, 12, 12, 12, 12, 12, 13, 13, 13, 13, 13, 13, 13, 13, 13, 13, 14, 14, 14, 14, 14, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 16, 16, 16, 16, 17, 17, 17, 17, 17, 17, 18, 18, 18, 18, 18, 19, 19, 19, 20, 20, 21 υπάρχουν 21 διαφορετικές βαθμίδες, που θα μπορούσαν να δώσουν μια απλή κατανομή με 21 κλάσεις, αλλά, για λόγους καλύτερης σύνοψής τους, ομαδοποιήθηκαν σε  $k = 7$  κλάσεις, δηλαδή 7 ισομεγέθη διαστήματα, με αντίστοιχες συχνότητες ( $f$ ) 6, 15, 24, 31, 24, 15, 6.

10

10

**Πίνακας 2.2** - Απλή και ομαδοποιημένη κατανομή: μια πρώτη εικόνα

Απλή Κατανομή			Ομαδοποιημένη Κατανομή		
Κλάση (k)	Διαφορετικές Τιμές X	Συχνότητα f	Κλάση (k)	Διάστημα XL - XU	Συχνότητα f
1	1	4	1	1-3	6
2	2	8	2	4-6	15
3	3	7	3	7-9	24
4	4	10	4	10-12	31
5	5	6	5	13-15	24
6	6	6	6	16-18	15
7	7	4	7	19-21	6
Μέγεθος δείγματος N = Σf = 45			Μέγεθος δείγματος N = Σf = 121		

11

11

## Απλή Κατανομή

- Η διαδικασία της πλήρους αριθμητικής ανάπτυξης της απλής συχνοτικής κατανομής περιλαμβάνει τρεις διαδοχικούς πίνακες δεδομένων.
- Πίνακας αρχικών δεδομένων**, στον οποίο καταγράφονται οι αρχικές τιμές στη σειρά με την οποία μετρήθηκαν (πίν. 2.2). Ο πίνακας αυτός χρειάζεται προκειμένου να εντοπίζονται τυχόν ακραίες τιμές ή σφάλματα καταγραφής των δεδομένων πριν υποβληθούν σε στατιστική ανάλυση.
- Αρχικά δεδομένα είναι οι τιμές (συνήθως αριθμητικές) που προκύπτουν από κάθε μέτρηση, δηλαδή καταγραφή των τιμών που παίρνει μια μεταβλητή (ένα μετρήσιμο χαρακτηριστικό) σε κάθε μέλος ενός αντίστοιχου δείγματος. Τα μέλη αυτά δεν έχουν ταξινομηθεί ή οργανωθεί, αλλά παρουσιάζονται με την "τυχαία" σειρά που είχαν στη μέτρηση.

**Πίν. 2.2** - Συμβολικά αρχικά δεδομένα μεταβλητής X (N=30)

X <sub>3</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>5</sub>	X <sub>5</sub>	X <sub>3</sub>
X <sub>2</sub>	X <sub>4</sub>	X <sub>3</sub>	X <sub>3</sub>	X <sub>2</sub>	X <sub>3</sub>
X <sub>5</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>1</sub>
X <sub>1</sub>	X <sub>4</sub>	X <sub>4</sub>	X <sub>1</sub>	X <sub>2</sub>	X <sub>1</sub>
X <sub>1</sub>	X <sub>4</sub>	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>2</sub>

$X_1, X_2, \dots, X_k = k$  διαφορετικές τιμές X.

12

12

Πίνακας 2.3 - Αρχικά δεδομένα της μεταβλητής βάρους (σε kg, N = 115)

70	75	78	82	81	83
71	72	77	76	81	81
65	74	78	76	79	75
74	86	75	71	75	83
73	87	75	77	81	84
76	67	<b>64</b>	80	68	84
76	81	68	69	81	85
75	71	70	77	74	77
70	76	69	69	79	72
72	76	77	77	88	74
74	79	82	77	79	81
74	80	70	71	79	82
83	78	75	<b>90</b>	79	82
85	74	70	77	75	75
79	75	79	77	80	73
84	65	65	68	88	
84	69	72	69	75	
85	72	79	85	75	
71	82	85	83	74	
71	75	76	72	79	

13

13

Πίνακας 2.4 - Διαλογή τιμών (επαναλήψεων) βάρους (N=115)

Κλάση k	Βάρος		Διαλογή	Συχνότητα
	X	Επαναλήψεις (#)	f	
1	<b>64</b>	I	1	1
2	65	III	3	3
3	66	-	0	0
4	67	I	1	1
5	68	III	3	3
6	69	IIIIII	5	5
7	70	IIIIII	5	5
8	71	IIIIII	6	6
9	72	IIIIII	6	6
10	73	II	2	2
11	74	IIIIIIII	8	8
<b>12</b>	<b>75</b>	IIIIIIIIIIIIII	<b>13</b>	<b>13</b>
13	76	IIIIIIII	7	7
14	77	IIIIIIIIII	9	9
15	78	III	3	3
16	79	IIIIIIIIIIII	10	10
17	80	III	3	3
18	81	IIIIIIII	7	7
19	82	IIIIII	5	5
20	83	IIII	4	4
21	84	IIII	4	4
22	85	IIIIII	5	5
23	86	I	1	1
24	87	I	1	1
25	88	II	2	2
26	89	-	0	0
27	<b>90</b>	I	1	1
Μέγεθος δείγματος =			N = Σf =	115

14

14

Πίνακας 2.5 - Απλή κατανομή του βάρους: πλήρης ανάλυση (N = 115)

Κλάση (k)	Βάρος (kg) X	Συχνότητα f	Αθροιστική Συχνότητα c	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	64	1	1	0.9	0.9
2	65	3	4	2.6	3.5
3	66	0	4	0.0	3.5
4	67	1	5	0.9	4.3
5	68	3	8	2.6	7.0
6	69	5	13	4.3	11.3
7	70	5	18	4.3	15.6
8	71	6	24	5.2	20.9
9	72	6	30	5.2	26.1
10	73	2	32	1.7	27.8
11	74	8	40	7.0	34.8
12	75	13	53	11.3	46.1
13	76	7	60	6.1	52.2
14	77	9	69	7.8	60.0
15	78	3	72	2.6	62.6
16	79	10	82	8.7	71.3
17	80	3	85	2.6	73.9
18	81	7	92	6.1	80.0
19	82	5	97	4.3	84.3
20	83	4	101	3.5	87.8
21	84	4	105	3.5	91.3
22	85	5	110	4.3	95.6
23	86	1	111	0.9	96.5
24	87	1	112	0.9	97.4
25	88	2	114	1.7	99.1
26	89	0	114	0.0	99.1
27	90	1	115	0.9	100.0

Οι τιμές στις στήλες %f και c%f στρογγυλοποιήθηκαν στο 1ο δεκαδικό ψηφίο.

15

15



- Οι απλές κατανομές συχνότητας παρουσιάζουν ορισμένα μειονεκτήματα, όπως ότι μερικές τιμές εμφανίζουν συχνότητα μηδέν και ότι ο αριθμός των k διαφορετικών τιμών είναι συνήθως μεγάλος.
- Αυτά τα δύο μειονεκτήματα εμποδίζουν την απόκτηση μιας συνολικής αρχικής εικόνας σχετικά με τα αρχικά δεδομένα και τη μορφή της κατανομής.
- Για τον λόγο αυτό, οι κατανομές αυτού του είδους πινακοποιούνται με τρόπο, ώστε στις θέσεις των διαφορετικών τιμών X να έχουμε ολιγάριθμες ομάδες (κλάσεις) τιμών.
- Η τεχνική αυτή λέγεται ομαδοποίηση και η κατανομή που προκύπτει λέγεται ομαδοποιημένη συχνοτική κατανομή.

16

16

Πίνακας 2.6 - Ομαδοποιημένη κατανομή του βάρους: ανάλυση (N = 115 )

Κλάση	Διάστημα Τιμών	Μεσαία Τιμή	Συχνότητα	Αθροιστική Συχνότητα	Σχετική Συχνότητα	Αθροιστική Σχετική Συχνότητα
k	$X_L - X_U$	X	f	cf	%f	c%f
1	64-66	65	4	4	3.5	3.5
2	67-69	68	9	13	7.8	11.3
3	70-72	71	17	30	14.8	26.1
4	73-75	74	23	53	20.0	46.1
5	76-78	77	19	72	16.5	62.6
6	79-81	80	20	92	17.4	80.0
7	82-84	83	13	105	11.3	91.3
8	85-87	86	7	112	6.1	97.4
9	88-90	89	3	115	2.6	100.0
		Σf =	115	Σ%f =	100	--

- Οι ομαδοποιήσεις αυτές ποικίλλουν από κατανομή σε κατανομή ως προς το πλήθος k των κλάσεων και το εύρος κάθε διαστήματος. Στον πίνακα 2.6 παρουσιάζεται η ομαδοποίηση της συχνοτικής κατανομής του πίνακα 2.5 με αριθμό κλάσεων  $k=9$  και εύρος διαστήματος 3 kg.
- Ο αριθμός των κλάσεων προέκυψε από τη διαίρεση του εύρους της κατανομής ( $90-64 = 26$ ) με ένα επιθυμητό πλάτος (width) διαστήματος 3 kg:  $26 / 3 = 8.7 \sim 9$  κλάσεις.

17

17

- Στη διάταξη του πίνακα 2.5 υπάρχει διατήρηση των αρχικών δεδομένων, τα οποία μπορούν να αναπαραχθούν από την κατανομή των συχνοτήτων χωρίς απώλεια ακρίβειας. Στην ομαδοποιημένη, όμως, κατανομή (πίν. 2.6) υπεισέρχεται αναπόφευκτα σφάλμα λόγω της ομαδοποίησης.
- Γενικά, στις ομαδοποιημένες κατανομές οι αρχικές τιμές που πέφτουν σε κάθε ομάδα (κλάση) εκπροσωπούνται από την αντίστοιχη μεσοτιμή του διαστήματος. Η τελική μορφή της ομαδοποίησης μιας κατανομής πρέπει να καθορίζεται από τον βέλτιστο συνδυασμό του αριθμού των ομάδων (k) και του πλήθους της (N) και να ικανοποιεί τους στόχους της ανάλυσης.

Πίνακας 2.6 - Ομαδοποιημένη κατανομή του βάρους: ανάλυση (N = 115 )

Κλάση	Διάστημα Τιμών	Μεσαία Τιμή	Συχνότητα	Αθροιστική Συχνότητα	Σχετική Συχνότητα	Αθροιστική Σχετική Συχνότητα
k	$X_L - X_U$	X	f	cf	%f	c%f
1	64-66	65	4	4	3.5	3.5
2	67-69	68	9	13	7.8	11.3
3	70-72	71	17	30	14.8	26.1
4	73-75	74	23	53	20.0	46.1
5	76-78	77	19	72	16.5	62.6
6	79-81	80	20	92	17.4	80.0
7	82-84	83	13	105	11.3	91.3
8	85-87	86	7	112	6.1	97.4
9	88-90	89	3	115	2.6	100.0
		Σf =	115	Σ%f =	100	--

18

18

## Οδηγίες για τη συστηματοποίηση της ομαδοποίησης

- Οι ακόλουθες **γενικές οδηγίες** συστηματοποιούν τη διαδικασία της ομαδοποίησης, δίνοντας μια πρώτη βέλτιστη λύση, την οποία ο ερευνητής μπορεί να αποδεχτεί ή να τροποποιήσει κατάλληλα, ώστε να καταλήξει σε μια ομαδοποίηση που να εκπληρώνει το σκοπό της:
  - Υπολογίζουμε το εύρος (E) της κατανομής αφαιρώντας από τη μεγαλύτερη αρχική τιμή τη μικρότερη  
Εύρος Κατανομής:  $E = X_{\max} - X_{\min}$
  - Χρησιμοποιώντας τον τύπο του Sturges υπολογίζουμε τον αριθμό των k ομάδων (κλάσεων) με βάση το πλήθος N της κατανομής:  
Αριθμός ομάδων:  $k = 1 + 3.32(\log_{10}N)$
  - Υπολογίζουμε το εύρος (E<sub>k</sub>) των ομάδων από τη σχέση  
Εύρος ομάδας:  $E_k = E / k$

19

19

## Ο αριθμός των κλάσεων μπορεί να υπολογισθεί με τους εξής τρόπους:

1. ως  $k = 1 + 3.322(\log_{10}N)$ , σύμφωνα με τον αλγόριθμο του Sturges (1926).
  - π.χ. με εύρος τιμών 90 - 64 = 26 και αριθμό κλάσεων  $k = 1 + 3.322 (\log_{10} 115) = 7.84 \sim 8$ , το ακριβές πλάτος διαστήματος είναι  $E = 26 / 8 = 3.25$ , που με βέλτιστο πλάτος διαστήματος 3 kg οι κλάσεις είναι τελικά  $26 / 3 = 8.7 = 9$ .
2. ως  $k = \sqrt{N}$  που χρησιμοποιεί το excel ή ως  $k = 2 * N^{1/3}$  (Rice University Rule).
  - Το πλάτος του διαστήματος μπορεί να υπολογιστεί με τους εξής τρόπους:
    - $E = 3.5 * s / N^{1/3}$ ,  $s =$  τυπική απόκλιση (τυχαία κανονικά δείγματα, Scott, 1979), ή
    - $E = 2 * (Q_1 / N^{1/3})$ , όπου  $Q_1$  το ενδοτεταρτημοριακό εύρος (Freedman και Diaconis, 1981).

20

20

## Εκατοστημοριακές αναλύσεις

- Μια από τις εφαρμογές της κατανομής των αθροιστικών εκατοστηαίων συχνοτήτων (c% f, πίν. 2.7 & 2.8) χρησιμοποιείται για τον προσδιορισμό των αρχικών τιμών της κατανομής που **αντιστοιχούν σε επιλεγμένες εκατοστηαίες θέσεις** και αντίστροφα για τον προσδιορισμό των εκατοστηαίων θέσεων ή απλά της εκατοστηαίας σειράς επιλεγμένων αρχικών τιμών.
  - Σε χώρες με παράδοση στη συγκέντρωση και χρήση στατιστικών στοιχείων, οι εκατοστημοριακές καθώς και άλλες συναφείς αναλύσεις εφαρμόζονται σε βάσεις δεδομένων και με τον τρόπο αυτό καθιερώνονται νόρμες (πληθυσμιακά ή άλλα ειδικά κριτήρια). Οι νόρμες αυτές αφορούν διάφορους κρίσιμους τομείς της ζωής, όπως η εκπαίδευση, η υγεία, ο αθλητισμός, η οικονομία, η ασφάλιση, το περιβάλλον, η βιομηχανία κ.λπ.
- Οι εκατοστηαίες αυτές αντιστοιχίες είναι τα γνωστά από την αριθμητική **εκατοστημόρια** και οι **εκατοστημοριακές θέσεις ή σειρές**.

21

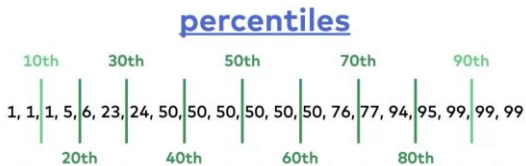
21



- Η χρησιμοποίηση των **εκατοστημορίων** είναι συχνή σε διδακτικές εφαρμογές και έρευνες στην εκπαίδευση, όπου πολλές φορές χρειάζεται να προσδιοριστεί η θέση της ατομικής απόδοσης (π.χ. του βαθμού) των μαθητών σε κάποιο μάθημα σε σχέση με την απόδοση ολόκληρης της τάξης του σχολείου ή / και με αντίστοιχες εθνικές νόρμες.
- Για **παράδειγμα**, ας υποθέσουμε ότι ένας μαθητής πήρε στις πανελλήνιες εξετάσεις στο μάθημα της φυσικής το βαθμό 17.5.
  - Αν υπήρχαν εθνικές νόρμες, δηλαδή αθροιστικά δεδομένα από όλη την Ελλάδα για το μάθημα της φυσικής, θα μπορούσαμε να προσδιορίσουμε σε ποιο ακριβώς εκατοστημόριο αντιστοιχεί ο βαθμός αυτός.
  - Έτσι θα είχαμε άμεση εκτίμηση-σύγκριση της επίδοσης του συγκεκριμένου μαθητή στο μάθημα αυτό σε εθνικό επίπεδο, προσδιορίζοντας λ.χ. ότι ο βαθμός 17.5 αντιστοιχεί στο 89<sup>ο</sup> εκατοστημόριο, πράγμα που θα σήμαινε ότι αυτή η επίδοση σε εθνική κλίμακα βρίσκεται στο 89%.

22

22

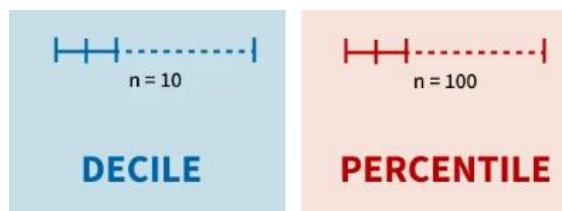


- Τα εκατοστημόρια δείχνουν απλά **την αρχική τιμή που αντιστοιχεί σε κάποια εκατοστιαία σειρά (θέση) της κατανομής.**
- Η κλίμακα τους είναι διακριτή και περιλαμβάνει μονάδες από το 0 μέχρι και το 100, ενώ ο συμβολισμός τους γίνεται συνήθως με το  $C_{\%}$  όπου στη θέση του % μπαίνει η εκατοστιαία σειρά.
  - Για παράδειγμα, οι συμβολισμοί  $C_{10}$  και  $C_{90}$  δείχνουν απλά το 10<sup>ο</sup> και το 90<sup>ο</sup> εκατοστημόριο.
  - Οι τιμές που αντιστοιχούν στα εκατοστημόρια είναι κάποιες από τις αρχικές τιμές της κατανομής.

23

23

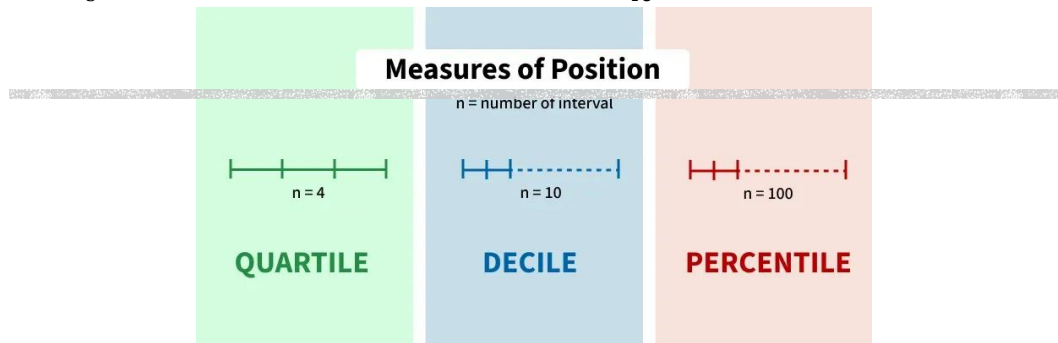
- Οι **εκατοστημοριακές τιμές** ολόκληρων δεκάδων αντιστοιχούν στις θέσεις μιας σειράς δευτερογενών εκατοστημορίων τιμών που ονομάζονται **δεκατημόρια**.
- Αυτά συμβολίζονται συνήθως με το γενικό συμβολισμό  $D_k$ , όπου στη θέση του  $k$  μπαίνει η δεκατημοριακή τιμή, που μπορεί να πάρει τιμές από το 1 μέχρι και το 10.
- Έτσι, **το δεκατημόριο  $D_1$**  αντιστοιχεί στο **εκατοστημόριο  $C_{10}$**  και το δεκατημόριο  $D_{10}$  αντιστοιχεί στο εκατοστημόριο  $C_{100}$ .
- Τα δεκατημόρια χρησιμοποιούνται για κατανομές με μεγάλο πλήθος  $N$ , κύρια όταν ο διαχωρισμός της κλίμακας της κατανομής για την εκατοστιαία σειρά των τιμών δεν απαιτεί ακρίβεια εκατοστού.



24

24

- Με την ίδια λογική, η κατανομή των αθροιστικών εκατοστιαίων συχνοτήτων μπορεί να θεωρηθεί ότι διαιρείται σε 4 ποσοστιαία διαστήματα που χωρίζονται μεταξύ τους από 3 αντίστοιχες τιμές, τα **τεταρτημόρια**.
- Αυτές οι τιμές συμβολίζονται με το γενικό συμβολισμό  $Q_k$ , όπου το  $k$  μπορεί να πάρει μόνο τις τιμές 1, 2, & 3.
- Έτσι έχουμε 3 **τεταρτημοριακές τιμές**:
  - την  $Q_1$  που αντιστοιχεί στο εκατοστημόριο  $C_{25}$ ,
  - την  $Q_2$  που αντιστοιχεί στο εκατοστημόριο  $C_{50}$  και
  - την  $Q_3$  που αντιστοιχεί στο εκατοστημόριο  $C_{75}$ .



25

25

Ανάλυση συχνοτήτων κατανομής : ύψη (N=115)

Κλάσεις k	Τιμές X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	164	1	1	0.9	0.9
2	165	3	4	2.6	3.5
3	166	0	4	0.0	3.5
4	167	1	5	0.9	4.3
5	168	3	8	2.6	7.0
6	169	5	13	4.3	11.3
7	170	5	18	4.3	15.6
8	171	6	24	5.2	20.9
9	172	6	30	5.2	26.1
10	173	2	32	1.7	27.8
11	174	8	40	7.0	34.8
12	175	13	53	11.3	46.1
13	176	7	60	6.1	52.2
14	177	9	69	7.8	60.0
15	178	3	72	2.6	62.6
16	179	10	82	8.7	71.3
17	180	3	85	2.6	73.9
18	181	7	92	6.1	80.0
19	182	5	97	4.3	84.3
20	183	4	101	3.5	87.8
21	184	4	105	3.5	91.3
22	185	5	110	4.3	95.6
23	186	1	111	0.9	96.5
24	187	1	112	0.9	97.4
25	188	2	114	1.7	99.1
26	189	0	114	0.0	99.1
27	190	1	115	0.9	100.0
--	Σf =	115	Σcf =	100	--

Τα αποτελέσματα στις στήλες %f &amp; c%f στρογγυλοποιήθηκαν στο ένα δεκαδικό ψηφίο.

## Υπολογισμός τιμής (X) από εκατοστημόριο (C%)

- Στις απλές κατανομές η εύρεση της τιμής X που αντιστοιχεί σε μια εκατοστημοριακή τιμή  $C_{\%}$  είναι σχετικά άμεση.
- Για παράδειγμα από τα δεδομένα του πίνακα μπορούμε εύκολα να εντοπίσουμε ορισμένες "έτοιμες" εκατοστιαίες θέσεις από τη στήλη των αθροιστικών εκατοστιαίων συχνοτήτων ( $c\%f$ ).
- Έτσι το 26<sup>ο</sup> εκατοστημόριο είναι  $C_{26}=172$  και αντιστοιχεί στην 6η τιμή 172 της κλάσης  $k=9$ . ενώ το 60<sup>ο</sup> εκατοστημόριο είναι  $C_{60}=177$  και αντιστοιχεί στην 9η τιμή 177 της κλάσης  $k=14$ .

25

26

## Υπολογισμός Τιμής (X) Από Εκατοστημόριο (C%)

- Για τιμές που δεν είναι εμφανείς από τη στήλη c%f χρησιμοποιούμε απλή διαδικασία, που αφορά στον προσδιορισμό της συγκεκριμένης σε σειρά (μέσα στην κλάση της) αρχικής τιμής που αντιστοιχεί σε κάποιο εκατοστημόριο.

Ανάλυση συχνοτήτων κατανομής : ύψη (N=115)

Κλάσεις k	Τιμές X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	164	1	1	0.9	0.9
2	165	3	4	2.6	3.5
3	166	0	4	0.0	3.5
4	167	1	5	0.9	4.3
5	168	3	8	2.6	7.0
6	169	5	13	4.3	11.3
7	170	5	18	4.3	15.6
8	171	6	24	5.2	20.9
9	172	6	30	5.2	26.1
10	173	2	32	1.7	27.8
11	174	8	40	7.0	34.8
12	175	13	53	11.3	46.1
13	176	7	60	6.1	52.2
14	177	9	69	7.8	60.0
15	178	3	72	2.6	62.6
16	179	10	82	8.7	71.3
17	180	3	85	2.6	73.9
18	181	7	92	6.1	80.0
19	182	5	97	4.3	84.3
20	183	4	101	3.5	87.8
21	184	4	105	3.5	91.3
22	185	5	110	4.3	95.6
23	186	1	111	0.9	96.5
24	187	1	112	0.9	97.4
25	188	2	114	1.7	99.1
26	189	0	114	0.0	99.1
27	190	1	115	0.9	100.0
--	Σf =	115	Σ%f =	100	--

Τα αποτελέσματα στις στήλες %f & c%f στρογγυλοποιήθηκαν στο ένα δεκαδικό ψηφίο.

27

Ανάλυση συχνοτήτων κατανομής : ύψη (N=115)

Κλάσεις k	Τιμές X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	164	1	1	0.9	0.9
2	165	3	4	2.6	3.5
3	166	0	4	0.0	3.5
4	167	1	5	0.9	4.3
5	168	3	8	2.6	7.0
6	169	5	13	4.3	11.3
7	170	5	18	4.3	15.6
8	171	6	24	5.2	20.9
9	172	6	30	5.2	26.1
10	173	2	32	1.7	27.8
11	174	8	40	7.0	34.8
12	175	13	53	11.3	46.1
13	176	7	60	6.1	52.2
14	177	9	69	7.8	60.0
15	178	3	72	2.6	62.6
16	179	10	82	8.7	71.3
17	180	3	85	2.6	73.9
18	181	7	92	6.1	80.0
19	182	5	97	4.3	84.3
20	183	4	101	3.5	87.8
21	184	4	105	3.5	91.3
22	185	5	110	4.3	95.6
23	186	1	111	0.9	96.5
24	187	1	112	0.9	97.4
25	188	2	114	1.7	99.1
26	189	0	114	0.0	99.1
27	190	1	115	0.9	100.0
--	Σf =	115	Σ%f =	100	--

Τα αποτελέσματα στις στήλες %f & c%f στρογγυλοποιήθηκαν στο ένα δεκαδικό ψηφίο.

### Για παράδειγμα, για να βρούμε την εκατοστημοριακή τιμή $C_5$ κάνουμε τους εξής απλούς υπολογισμούς:

- Η τιμή X που αντιστοιχεί στο εκατοστημόριο αυτό είναι η τιμή της οποίας η θέση ορίζει το πρώτο 5% των  $N=115$  αρχικών τιμών της κατανομής. Αυτή η τιμή έχει σειρά  $(5\%)115 = 5.75 \approx 6$ , είναι δηλαδή η 6<sup>η</sup> σε σειρά τιμή της κατανομής.
- Όπως διαπιστώνουμε από τα δεδομένα της στήλης cf του πίνακα, η 6<sup>η</sup> σε σειρά τιμή της κατανομής πέφτει στην κλάση  $k=5$ , είναι η  $X=168$  και αφορά την 1<sup>η</sup> από τις 3 τιμές 168 της κλάσης αυτής.

28

28

Με τον ίδιο τρόπο για να βρούμε την εκατοστημοριακή τιμή  $C_{53}$  κάνουμε τους εξής απλούς υπολογισμούς:

- Η τιμή  $X$  που αντιστοιχεί στο εκατοστημόριο αυτό είναι η τιμή της οποίας η θέση ορίζει το πρώτο 53% των  $N=115$  αρχικών τιμών της κατανομής. Αυτή η τιμή έχει σειρά (53%) $115 = 60.95 \approx 61$ , είναι δηλαδή η 61η σε σειρά τιμή της κατανομής.
- Όπως διαπιστώνουμε από τα δεδομένα της στήλης  $cf$  του πίνακα 2.7, η 61η σε σειρά τιμή της κατανομής πέφτει στην κλάση  $k=14$ , είναι η  $X=177$  και αφορά την 1η από τις 9 τιμές 177 της κλάσης αυτής.

· Ανάλυση συχνοτήτων κατανομής : ύψη ( $N=115$ )

Κλάσεις k	Τιμές X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	164	1	1	0.9	0.9
2	165	3	4	2.6	3.5
3	166	0	4	0.0	3.5
4	167	1	5	0.9	4.3
5	168	3	8	2.6	7.0
6	169	5	13	4.3	11.3
7	170	5	18	4.3	15.6
8	171	6	24	5.2	20.9
9	172	6	30	5.2	26.1
10	173	2	32	1.7	27.8
11	174	8	40	7.0	34.8
12	175	13	53	11.3	46.1
13	176	7	60	6.1	52.2
14	177	9	69	7.8	60.0
15	178	3	72	2.6	62.6
16	179	10	82	8.7	71.3
17	180	3	85	2.6	73.9
18	181	7	92	6.1	80.0
19	182	5	97	4.3	84.3
20	183	4	101	3.5	87.8
21	184	4	105	3.5	91.3
22	185	5	110	4.3	95.6
23	186	1	111	0.9	96.5
24	187	1	112	0.9	97.4
25	188	2	114	1.7	99.1
26	189	0	114	0.0	99.1
27	190	1	115	0.9	100.0
--	Σf =	115	Σ%f =	100	--

Τα αποτελέσματα στις στήλες %f & c%f στρογγυλοποιήθηκαν στο ένα δεκαδικό ψηφίο.

29

- Στις ομαδοποιημένες κατανομές η διαδικασία προσδιορισμού των εκατοστημορίων γίνεται με τη βοήθεια της **αλγοριθμικής μεθόδου**.
- Για να κατανοήσουμε καλύτερα τα διαδοχικά βήματα αυτής της μεθόδου, ας δούμε το εξής **παράδειγμα**:
  - να υπολογίσουμε την εκατοστημοριακή τιμή  $C_{41}=?$  της ομαδοποιημένης κατανομής των σωματικών υψών που δίνεται στον πίνακα 2.8.

## Ομαδοποιημένες κατανομές και αλγοριθμική μέθοδος

Ανάλυση ομαδοποιημένης κατανομής : ύψος ( $N=115$ )

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική ή Σχετική Συχνότητα
						c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
	Σf =		115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

30

1. Εντοπίζουμε την ομάδα (κλάση) μέσα στην οποία πέφτει η εκατοστημοριακή τιμή  $C_{41}$ :

- Η τιμή  $X$  που αντιστοιχεί στο εκατοστημόριο αυτό χωρίζει την κατανομή σε ένα σημείο τέτοιο, ώστε το πρώτο **41%** των αρχικών τιμών να πέφτουν πριν την τιμή  $X$ .
- Το **41%** των 115 αρχικών τιμών πέφτουν πριν την 47η τιμή, καθότι  **$(0.41)(115)=47.15$** .
- Από τη στήλη  $cf$  του πίνακα 2.8 φαίνεται πως η 47η τιμή της κατανομής ανήκει στην ομάδα  **$k=4$** .

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική	Σχετική	Αθροιστικ
				Συχνότητα cf	Συχνότητα %f	ή Σχετική Συχνότητα c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

31

31

2. Υπολογίζουμε τον αριθμό των τιμών από την αρχή της κατανομής μέχρι το **κατώτερο αριθμητικό όριο** του διαστήματος της ομάδας  $k=4$  και τον βρίσκουμε **30** (στήλη  $cf$ ).
3. Υπολογίζουμε πόσες ακόμα τιμές χρειάζονται για να συμπληρωθεί ο αριθμός των 47 τιμών που προσδιορίσαμε και βρίσκουμε  **$47-30 = 17$  τιμές**.

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική	Σχετική	Αθροιστικ
				Συχνότητα cf	Συχνότητα %f	ή Σχετική Συχνότητα c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

32

32

4. Θεωρούμε ότι οι **23** τιμές του διαστήματος της ομάδας  $k=4$  διαιρούν το εύρος του διαστήματος αυτού σε **23** ισαπέχουσες θέσεις.
5. Υπολογίζουμε το πρόσθετο τμήμα, μέσα στα όρια του διαστήματος, που υπολείπεται για να φτάσουμε στη θέση της ζητούμενης τιμής που αντιστοιχεί στο  $C_{41}$  και το βρίσκουμε να είναι ίσο με  $17/23$  του εύρους 3 του διαστήματος, δηλαδή  $(17/23)*3 = 2.22$ .

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική	Σχετική	Αθροιστικ ή Σχετική
				Συχνότητα cf	Συχνότητα %f	Συχνότητα c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

33

33

6. Προσθέτουμε την τιμή αυτή στο **κατώτερο όριο 172.5** του συγκεκριμένου διαστήματος και βρίσκουμε την εκατοστημοριακή τιμή να είναι:

$$C_{41} = 172.5 + 2.22 = 174.7 \approx 175$$

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική	Σχετική	Αθροιστικ ή Σχετική
				Συχνότητα cf	Συχνότητα %f	Συχνότητα c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

34

34

- Ο γενικός αλγόριθμος της διαδικασίας αυτής εκφράζεται με τον τύπο:

$$C_{\%} = X = X_L + (\%N - cf_{k-1}) \frac{E_k}{f_k}$$

- όπου

$X_L$ , η τιμή του κατώτερου ορίου του διαστήματος της κλάσης  $k$ ,

$\%$  η εκατοστιαία σειρά,

$N$  το πλήθος της κατανομής,

$cf_{k-1}$  η αθροιστική συχνότητα της κλάσης  $k-1$ ,

$E_k$  το εύρος (των πραγματικών ορίων) του διαστήματος της κλάσης  $k$  και

$f_k$  η συχνότητα της κλάσης  $k$ .

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική ή Σχετική Συχνότητα c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

35

35

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική ή Σχετική Συχνότητα c%f
1	164-166	165	4	4	3.5	3.5
2	167-169	168	9	13	7.8	11.3
3	170-172	171	17	30	14.8	26.1
4	173-175	174	23	53	20.0	46.1
5	176-178	177	19	72	16.5	62.6
6	179-181	180	20	92	17.4	80.0
7	182-184	183	13	105	11.3	91.3
8	185-187	186	7	112	6.1	97.4
9	188-190	189	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

- Για να επιβεβαιώσουμε την ισχύ του τύπου αυτού, ως επαληθεύσουμε την εκατοστημοριακή τιμή  $C_{41}=175$ , που λογαριάσαμε λίγο πριν.
- Στην περίπτωση αυτή η ομάδα (κλάση) που περιέχει τη ζητούμενη τιμή  $X$  είναι η  $k=4$  και η τιμή του κατώτερου πραγματικού ορίου του αντίστοιχου διαστήματος είναι η  $X_L=172.5$ .

36

36

- Ακόμα το πλήθος της κατανομής είναι  $N=115$ , η εκατοστιαία θέση της ζητούμενης τιμής  $X$  είναι η  $\%=0.41$ , η αθροιστική συχνότητα για την ομάδα της κλάσης  $k-1=4-1=3$  είναι  $cf_{k-1}=30$ , η συχνότητα της κλάσης  $k=4$  είναι  $f_k=23$  και το εύρος του διαστήματος είναι

$$E_k=175.5-172.5=3.$$

- Έτσι έχουμε:

$$C_{\%} = X = X_L + (\%N - cf_{k-1}) \frac{E_k}{f_k}$$

$$\begin{aligned} C_{\%} = X &= 172.5 + [(0.41)(115) - 30] (3 / 23) \\ &= 172.5 + (47.15 - 30)(0.13) \\ &= 172.5 + 2.23 \\ &= 174.7 \approx 175 \end{aligned}$$

Ανάλυση ομαδοποιημένης κατανομής : ύψος (N=115)

k	Όρια Τιμών $X_L - X_U$	Μεσοτιμή X	Συχνότητα f	Αθροιστική		Αθροιστική ή Σχετική	
				Συχνότητα cf	Συχνότητα %f	Συχνότητα cf	Συχνότητα %f
1	164-166	165	4	4	3.5	3.5	
2	167-169	168	9	13	7.8	11.3	
3	170-172	171	17	30	14.8	26.1	
4	173-175	174	23	53	20.0	46.1	
5	176-178	177	19	72	16.5	62.6	
6	179-181	180	20	92	17.4	80.0	
7	182-184	183	13	105	11.3	91.3	
8	185-187	186	7	112	6.1	97.4	
9	188-190	189	3	115	2.6	100.0	
			Σf =	115	Σ%f =	100	--

L = κατώτερο όριο διαστήματος, U = ανώτερο όριο διαστήματος.

37

37

## Υπολογισμός εκατοστημορίου ( $C_{\%}$ ) από τιμή (X)

- Όπως είδαμε παραπάνω, η αντίστροφη διαδικασία παράγει τις λεγόμενες εκατοστιαίες σειρές, δηλαδή τις εκατοστιαίες αντιστοιχίες (θέσεις) των αρχικών τιμών  $X$ .
- Στην πράξη, η εύρεση της εκατοστιαίας σειράς ή θέσης μιας αρχικής τιμής είναι μερικές φορές χρησιμότερη από αυτή της εύρεσης του εκατοστημορίου.
- Ο προσδιορισμός των εκατοστιαίων αντιστοιχιών επιλεγμένων αρχικών τιμών έχει μεγάλη πρακτική αξία σε αναλύσεις νορμών, με βάση τις οποίες πρέπει να πάρουμε αποφάσεις σχετικά με τη συμπερίληψη ή τον αποκλεισμό κάποιων ακραίων ποσοστών από ένα εκπαιδευτικό ή προπονητικό πρόγραμμα (π.χ. επιλογή ταλέντων).

38

38

· Ανάλυση συχνότητας κατανομής : ύψη (N=115)

- Στις απλές κατανομές η εύρεση της εκατοστημοριακής θέσης ( $C\%$ ) κάποιας τιμής ( $X$ ) είναι σχετικά εύκολη.
- Για παράδειγμα, από τα δεδομένα του πίνακα 2.7 μπορούμε εύκολα να εντοπίσουμε την εκατοστημοριακή θέση οποιασδήποτε τιμής  $X$ , μέσω της στήλης των αθροιστικών εκατοστιαίων συχνοτήτων ( $c\%f$ ).
- Έτσι, η τιμή 167 cm έχει εκατοστημοριακή θέση 4.3% και η τιμή 187 έχει εκατοστημοριακή θέση 97.4%.
- Η ερμηνεία αυτών των δύο % σειρών είναι ότι το σύνολο των ατόμων της κατανομής που έχουν ύψος μικρότερο ή ίσο με το 167 cm αποτελεί περίπου το 4% του πλήθους της, ενώ το σύνολο των ατόμων της κατανομής που έχουν ύψος μικρότερο ή ίσο με 187 cm αποτελεί το 97.4% του πλήθους της.

Κλάσεις k	Τιμές X	Συχνότητα f	Αθροιστική Συχνότητα cf	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	164	1	1	0.9	0.9
2	165	3	4	2.6	3.5
3	166	0	4	0.0	3.5
4	167	1	5	0.9	4.3
5	168	3	8	2.6	7.0
6	169	5	13	4.3	11.3
7	170	5	18	4.3	15.6
8	171	6	24	5.2	20.9
9	172	6	30	5.2	26.1
10	173	2	32	1.7	27.8
11	174	8	40	7.0	34.8
12	175	13	53	11.3	46.1
13	176	7	60	6.1	52.2
14	177	9	69	7.8	60.0
15	178	3	72	2.6	62.6
16	179	10	82	8.7	71.3
17	180	3	85	2.6	73.9
18	181	7	92	6.1	80.0
19	182	5	97	4.3	84.3
20	183	4	101	3.5	87.8
21	184	4	105	3.5	91.3
22	185	5	110	4.3	95.6
23	186	1	111	0.9	96.5
24	187	1	112	0.9	97.4
25	188	2	114	1.7	99.1
26	189	0	114	0.0	99.1
27	190	1	115	0.9	100.0
--	Σf =	115	Σ%f =	100	--

Τα αποτελέσματα στις στήλες %f & c%f στρογγυλοποιήθηκαν στο ένα δεκαδικό ψηφίο.

39

αριθμητική τιμή - numerical value

δεκατημόριο - centile

εκατοστημόριο - percentile

εύρος - range

κατανομή απλή - simple distribution

κατανομή ομαδοποιημένη - grouped distribution

κατανομή συχνότητας - frequency distribution

κλάσεις κατανομής - frequency classes

ομάδα αριθμών - group of numbers

ομάδα τιμών - group of values

πίνακας συχνοτήτων - frequency table

σειρά αριθμών - set of numbers

συχνότητα αθροιστική - cumulative frequency

συχνότητα σχετική - relative frequency

τεταρτημόριο - quartile

Ελληνο-Αγγλική  
Ορολογία

40

**Σας ευχαριστώ για  
την προσοχή σας**



**Αντώνης Κ. Τραυλός**