

Κεφάλαιο 2. ΣΤΑΤΙΣΤΙΚΗ

A. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Στατιστική: ένα σύνολο αρχών και μεθοδολογιών για:

- το σχεδιασμό της διαδικασίας συλλογής δεδομένων
- τη συνοπτική και αποτελεσματική παρουσίασή τους
- την ανάλυση και εξαγωγή αντίστοιχων συμπερασμάτων

Πληθυσμός: Το σύνολο του οποίου τα στοιχεία εξετάζουμε ως προς ένα ή περισσότερα χαρακτηριστικά τους.

Απογραφή: Εξετάζουμε όλα τα άτομα του πληθυσμού.

Δειγματοληψία: Εξετάζουμε ένα αντιπροσωπευτικό μέρος του πληθυσμού, **το δείγμα**, ώστε να βγάλουμε συμπεράσματα για όλον τον πληθυσμό.

Μεταβλητές: Τα χαρακτηριστικά ως προς τα οποία εξετάζουμε έναν πληθυσμό (συμβολίζονται X, Y, Z, ...)

Πλήθος (n) του δείγματος: Ο αριθμός των ατόμων του δείγματος.

Ασκ. 3, 4 (σελ61) Σχολικού Βιβλίου

ΕΙΔΗ ΜΕΤΑΒΛΗΤΩΝ

1. ΠΟΙΟΤΙΚΕΣ Αυτές που δεν μπορούν να μετρηθούν

Χρώμα μαλλιών (μαύρο, καστανό,...), ομάδα αίματος (A,B,0 AB,...), φύλο (άνδρας, γυναίκα).

2. ΠΟΣΟΤΙΚΕΣ Αυτές που μπορούν να μετρηθούν με αριθμούς

ΔΙΑΚΡΙΤΕΣ

Παίρνουν διακεκριμένες τιμές (μετριοούνται με ακέραιους).

ΣΥΝΕΧΕΙΣ

Παίρνουν όλες τις τιμές ενός διαστήματος (μετριοούνται με δεκαδικούς).

Αριθμός παιδιών (0,1,2,3,...),
 Νούμερο παπουτσιού (37, 38, 39, ...)
 Πόντοι στο μπάσκετ (87, 90, 55,)

Βάρος μαθητή, (98,3 76,9 58,7 ...)
 Μήκος ποταμού (3.356.890,8 14.567.832 ...)

Χαρακτηρίστε τις παρακάτω μεταβλητές

ΕΙΔΗ ΜΕΤΑΒΛΗΤΩΝ

- Το μήκος ενός ποταμού
- Το πλήθος των θεατών σε έναν αγώνα ποδοσφαίρου
- Η διάρκεια ζωής ενός ηλεκτρικού λαμπτήρα
- Οι Ιταλοί που έχουν επισκεφθεί την Ελλάδα το καλοκαίρι
- Το πλήθος των σελίδων ενός βιβλίου
- Το επίπεδο μόρφωσης
- Το πλήθος των ορόφων ενός κτιρίου
- Το πλήθος των θεατών σε έναν αγώνα ποδοσφαίρου
- Η διάρκεια μιας κινηματογραφικής ταινίας

- Το βάρος μιας φραντζόλας ψωμιού
- Τα μόρια για την εισαγωγή στα ΑΕΙ και ΤΕΙ
- Η εθνικότητα προσφύγων
- Η ταχύτητα ενός αυτοκινήτου στην κλήση της Τροχαίας
- Το ύψος ενός βουνού
- Το πλήθος των επιβατών λεωφορείου
- Θερμοκρασία ενός δωματίου
- Οι πωλήσεις ενός μοντέλου αυτοκινήτου
- Το βάρος μιας φραντζόλας ψωμιού
- Ο βαθμός Φυσικής στον έλεγχο
- Ο βαθμός απολυτηρίου

Άσκηση 1, 2 (σελ61) Σχολικού Βιβλίου

Β. ΠΙΝΑΚΕΣ ΚΑΤΑΝΟΜΗΣ ΣΥΧΝΟΤΗΤΩΝ

Οι 2 πρώτες στήλες που εμφανίζονται:

X_i : Το όνομα της μεταβλητής

Βρίσκω τις διαφορετικές τιμές που παρουσιάζονται στο δείγμα και βάζω κάθε τιμή σε μια νέα γραμμή του πίνακα. Έτσι, καθορίζεται ο αριθμός των γραμμών. Τον συμβολίζω με k .

v_i : Συχνότητα του X_i

Ο αριθμός των ατόμων που έχουν τιμή μεταβλητής ίση με X_i ,

ΠΑΡΑΔΕΙΓΜΑ

Εξετάσαμε 20 οικογένειες ως προς τον αριθμό των παιδιών που έχουν. Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα:

ΠΙΝΑΚΑΣ ΣΥΧΝΟΤΗΤΩΝ

Αριθμός παιδιών x_i	Οικογένειες Συχνότητα v_i
0	3
1	5
2	8
3	3
4	1
Σύνολο	20

Πληθυσμός: Οι 20 οικογένειες.

Μεταβλητή: Αριθμός παιδιών.

Πλήθος: $v = 20$

Πλήθος γραμμών: 4 (4 διαφορετικές τιμές για X_i)

Διαβάζω: υπάρχουν 3 οικογένειες χωρίς παιδιά, 5 με 1 παιδί, 8 με 2 παιδιά, ...

Πως αναφέρουμε τις θέσεις του πίνακα.

Αριθμός παιδιών x_i	Συχνότητα v_i
$x_1 = 0$	$v_1 = 3$
$x_2 = 1$	$v_2 = 5$
$x_3 = 2$	$v_3 = 8$
$x_4 = 3$	$v_4 = 3$
$x_5 = 4$	$v_5 = 1$
Σύνολο	$v = 20$

$$v_1 + v_2 + \dots + v_k = v$$

Οι σχετικές συχνότητες:

f_i : Η Σχετική Συχνότητα του X_i

Ο λόγος της συχνότητας προς το μέγεθος του δείγματος (εκφράζει το «πόσα στα πόσα»).

$$f_i = \frac{v_i}{v}$$

$$f_1 + f_2 + \dots + f_k = 1$$

$f_i \%$: Συχνότητα του X_i

Η σχετική συχνότητα σε ποσοστό %.

$$f_i \% = f_i \cdot 100$$

$$f_1 \% + f_2 \% + \dots + f_k \% = 100$$

Το παράδειγμα γίνεται:

Αριθμός παιδιών x_i	Συχνότητα v_i	Σχετική συχνότητα f_i	Σχετική συχνότητα % $f_i \%$
0	3	$f_1 = 3/20 = \mathbf{0,15}$	$f_1 \% = f_1 \cdot 100 = 0,15 \cdot 100 = \mathbf{15}$
1	5	$f_2 = 5/20 = \mathbf{0,25}$	$f_2 \% = f_2 \cdot 100 = 0,25 \cdot 100 = \mathbf{25}$
2	8	$f_3 = 8/20 = \mathbf{0,40}$	$f_3 \% = f_3 \cdot 100 = 0,40 \cdot 100 = \mathbf{40}$
3	3	$f_4 = 3/20 = \mathbf{0,15}$	$f_4 \% = f_4 \cdot 100 = 0,15 \cdot 100 = \mathbf{15}$
4	1	$f_5 = 1/20 = \mathbf{0,05}$	$f_5 \% = f_5 \cdot 100 = 0,05 \cdot 100 = \mathbf{5}$
Σύνολο	20	1	100

Ερωματολόγιο κάψης

Οι αθροιστικές συχνότητες ποσοτικών μεταβλητών

N_i : Η Αθροιστική Συχνότητα του X_i

το άθροισμα των συχνοτήτων v_i των παρατηρήσεων που είναι μικρότερες ή ίσες με αυτή την τιμή του X_i

Απαντάει σε ερωτήσεις του τύπου: «**πόσα** άτομα έχουν X_i μικρότερο ή ίσο του X_i », «...το πολύ ίσο με...», «...μέχρι και...» και παίρνοντας την προηγούμενη τιμή «...λιγότερο από...», «...κάτω από...»

F_i : Η Σχετική Αθροιστική Συχνότητα του X_i

το άθροισμα των σχετικών συχνοτήτων f_i των παρατηρήσεων που είναι μικρότερες ή ίσες με την τιμή X_i

$F_i \%$: Η Σχετική Αθροιστική Συχνότητα % του X_i

το άθροισμα των σχετικών % συχνοτήτων $f_i \%$ των παρατηρήσεων που είναι μικρότερες ή ίσες με την τιμή X_i

Απαντάει σε ερωτήσεις του τύπου: «**τι ποσοστό** ατόμων έχουν X_i μικρότερο ή ίσο του X_i », «...το πολύ ίσο με...», «...μέχρι και...» και παίρνοντας την προηγούμενη τιμή «...λιγότερο από...», «...κάτω από...»

Το παράδειγμα γίνεται:

Αριθμός παιδιών X_i	Συχνότητα v_i	Σχετική συχνότητα f_i	Σχετική συχνότ.% $f_i \%$	Αθροιστική Συχνότητα N_i	Αθροιστική Σχετική συχνότ. F_i	Αθροιστική Σχετ. συχνότ.% $F_i \%$
0	3	0,15	15	3	0,15	15
1	5	0,25	25	8	0,40	40
2	8	0,40	40	16	0,80	80
3	3	0,15	15	19	0,95	95
4	1	0,05	5	20	1	100
Σύνολο	20	1	100			



- Οι ποιοτικές μεταβλητές δεν έχουν Αθροιστικές συχνότητες.
- Τα X_i πρέπει να είναι ταξινομημένα κατά μέγεθος για έχουν νόημα οι Αθροιστικές συχνότητες.
- οι Αθροιστικές συχνότητες δεν έχουν σύνολα. Η τελευταία γραμμή τους είναι ίδια με αυτή των συνόλων των απλών συχνοτήτων.

• Η σχέση των N_i με τις F_i και $F_i\%$ είναι ή ίδια με την σχέση των v_i με τις f_i και $f_i\%$.

$$F_i = \frac{N_i}{v} \quad , \quad F_i \% = F_i \cdot 100$$

Παραγωγή ιδίωνα εννοσήτων, Πίνακας εννοσήτων Φοιτητές, Ερωτήσεις αδρυστικίων, Ελλειψείς ιδίωνακες

Άσκηση 1 (σελ78), 4, 5 (σελ79) Σχολικού Βιβλίου

Γ. ΟΜΑΔΟΠΟΙΗΣΗ ΤΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ

Είναι απαραίτητη κυρίως στην περίπτωση μιας συνεχούς μεταβλητής, αλλά στην περίπτωση μιας διακριτής μεταβλητής με μεγάλο αριθμό διαφορετικών τιμών της X_i .

Ομαδοποιούμε τα δεδομένα σε ομάδες (**Κλάσεις**)

που περιλαμβάνουν τιμές X_i από μια τιμή μέχρι μια άλλη (**όρια της κλάσης**)

με ίδια απόσταση η μια από την άλλη (**πλάτος κλάσης**).

Οι κλάσεις έχουν μορφή διαστήματος [...], δηλαδή περιλαμβάνουν το άνω άκρο, όχι το κάτω.

Π.χ. Έστω 20 δεδομένα: 5, 18, 2, 17, 8, 19, 19, 7, 6, 10, 14, 12, 13, 9, 16, 5, 8, 12, 11, 13

1.	Βρίσκουμε την ελάχιστη min και την μέγιστη max τιμή των δεδομένων	$\min = 2 \quad \max = 19$
2.	Το εύρος δείγματος βρίσκεται ως $\max - \min$	$\text{Max} - \min = 19 - 2 = 17$
3.	Για το πλάτος της κλάσης διαιρούμε το εύρος δια του αριθμού των κλάσεων π.χ 3.	$17/3 = 5,67$ δηλ $\cong 6$ (για δεκαδικό στρογγυλεύουμε στον επόμενο ακέραιο)
4.	Τώρα έχουμε κλάσεις με πλάτος 6. Αρχίζω από το $\min=2$ και προσθέτω 6. <u>Πρώτη κλάση</u> 2,8 προσθέτω πάλι 6 <u>Δεύτερη</u> 8,14 <u>Τρίτη</u> 14, 20	Περιλαμβάνω την αρχή κάθε κλάσης και αφήνω το τέλος για την επόμενη [2, 8) [8,14) [14, 20) Δηλ το «8» θα ανήκει στην δεύτερη κλάση

Τώρα μοιράζουμε τα δεδομένα στις κλάσεις και φτιάχνουμε τον πίνακα.
Το X_i θα αντιπροσωπεύεται από τα μέσα της κλάσης δηλ. από τις **Κεντρικές Τιμές**.

Κλάσεις		Κεντρικές τιμές k_i	Διαλογή	v_i
[2, 8)	$(2+8)/2 = 10/2 = 5$	5	5,2,7,6,5	5
[8,14)	$(8+14)/2=22/2 = 11$	11	8,10,12,13,9,8,12,11,13	9
[14, 20)	$(14+20)/2=34/2=17$	17	18,17,19,19,14,16	6
			Σύνολο	$v = 20$

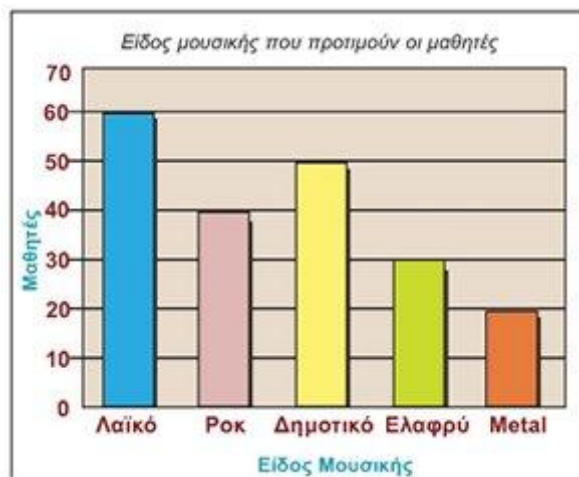
Ομαδοποίηση 80 αριθμών.

Άσκηση 7 (σελ 83) μόνο το α), β)

Δ. ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΚΑΤΑΝΟΜΗΣ ΣΥΧΝΟΤΗΤΩΝ

A) ΡΑΒΔΟΓΡΑΜΜΑ (για ποιοτικές μεταβλητές)

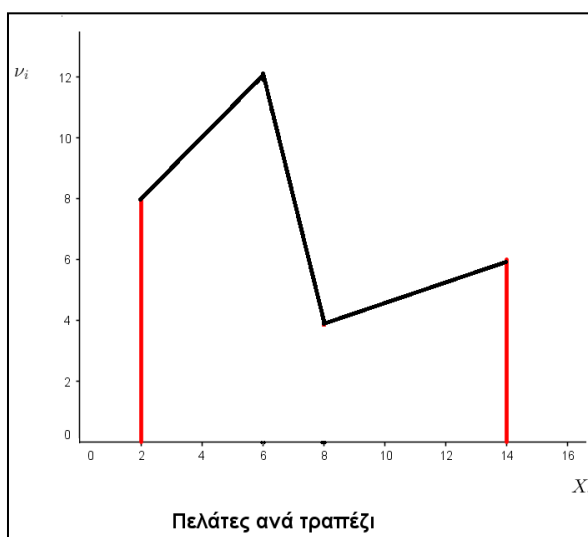
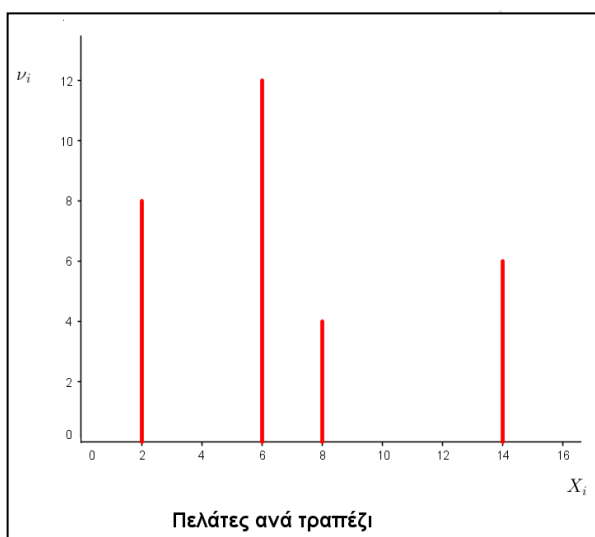
- Στον οριζόντιο άξονα τα X_i και στον κάθετο οι τιμές των συχνοτήτων.
- Σε κάθε τιμή της μεταβλητής X αντιστοιχεί μια ορθογώνια στήλη της οποίας το ύψος είναι ίσο με την αντίστοιχη συχνότητα ή σχετική συχνότητα.
- Τόσο η απόσταση μεταξύ των στηλών όσο και το μήκος των βάσεων τους καθορίζονται αυθαίρετα.
- Η ετικέτα βρίσκεται στο μέσο της στήλης.
- Μπορεί να έχω ραβδόγραμμα συχνοτήτων (v_i), αλλά και σχετικών συχνοτήτων (f_i) ή σχετικών % συχνοτήτων ($f_i\%$). Τότε έχω όμοια εικόνα στηλών.
- Αν εναλλάξουμε τους άξονες από κατακόρυφο το ραβδόγραμμα γίνεται οριζόντιο.



B) ΔΙΑΓΡΑΜΜΑ ΣΥΧΝΟΤΗΤΩΝ (για ποσοτικές μεταβλητές)

- Όπως το ραβδόγραμμα αλλά με κάθετες γραμμές αντί για στήλες, ακριβώς στην ετικέτα.
- Ο άξονας των X είναι αριθμητικός.
- Μπορεί να έχω ραβδόγραμμα συχνοτήτων (v_i), αλλά και όλων των άλλων συχνοτήτων (και αθροιστικών).
- Όταν αναπαριστώ αθροιστικά μεγέθη τα ύψη είναι αύξοντα.

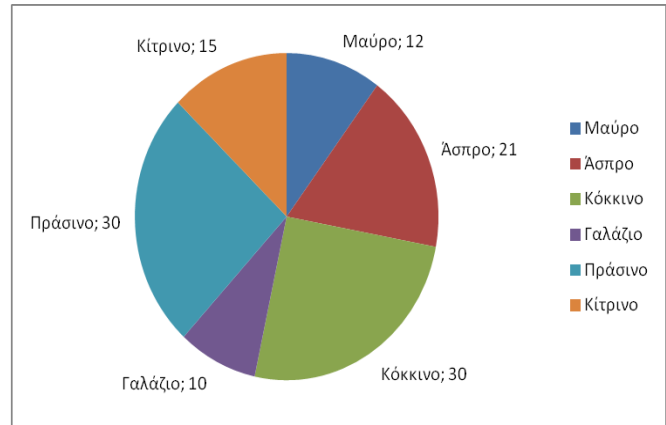
Αν ενώσω τα σημεία (x_i, v_i) προκύπτει το πολύγωνο συχνοτήτων.



Γ) ΚΥΚΛΙΚΟ ΔΙΑΓΡΑΜΜΑ (για ποιοτικές και ποσοτικές μεταβλητές)

Γωνιά σε μοίρες που αναλογεί

X_i	v_i	f_i	α_i
Μαύρο	12	0,102	$0,102 * 360 = 37$
Άσπρο	21	0,178	$0,178 * 360 = 64$
Κόκκινο	30	0,254	$0,254 * 360 = 91$
Γαλάζιο	10	0,085	$0,085 * 360 = 31$
Πράσινο	30	0,254	$0,254 * 360 = 91$
Κίτρινο	15	0,127	$0,127 * 360 = 46$
	118	1	360



Χωρίζουμε έναν κύκλο σε κυκλικούς τομείς των οποίων οι γωνίες είναι ανάλογες με τις συχνότητες που αναπαριστούμε.

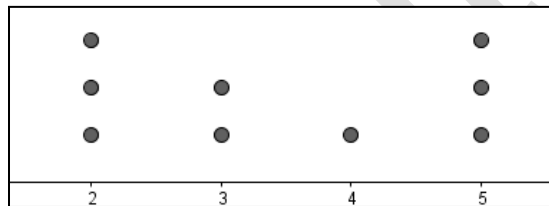
$$\alpha_i = 360 (v_i/v) = 360 f_i$$

Πρακτικά, παίρνω γωνίες α_i σαν ένα κομμάτι των 360° του κύκλου, ανάλογο της f_i .

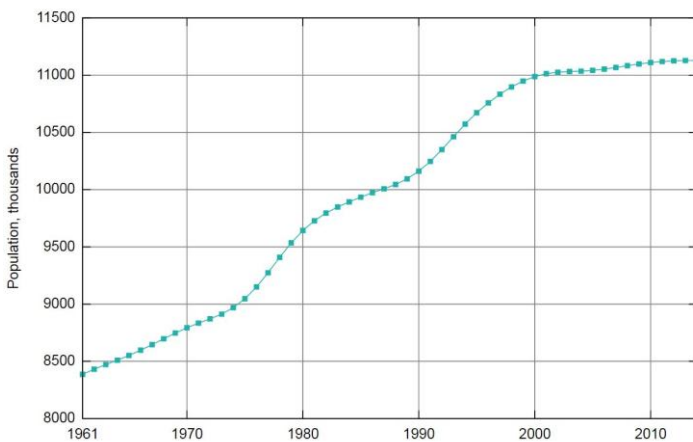
Δ) ΣΗΜΕΙΟΓΡΑΜΜΑ (για λίγα δεδομένα)

Εξαγωγές δοντιών 9 ζαχαροπλαστών:

5	2	3	2	5	2	5	3	4
---	---	---	---	---	---	---	---	---



στ) ΧΡΟΝΟΓΡΑΜΜΑ (Για την διαχρονική αναπαράσταση φαινομένων.)

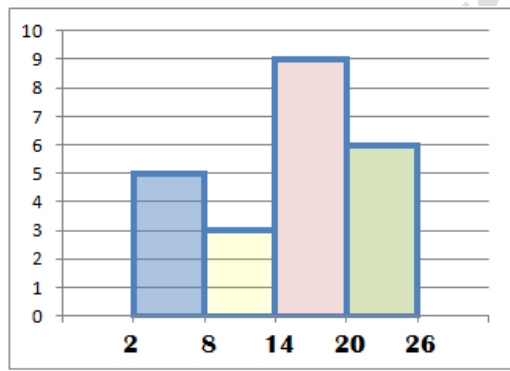


Στα διαγράμματα αυτά ο άξονας x παριστάνει χρόνο.

ΙΣΤΟΓΡΑΜΜΑ ΣΥΧΝΟΤΗΤΩΝ (όταν έχω ομαδοποιημένα δεδομένα συνεχούς μεταβλητής)

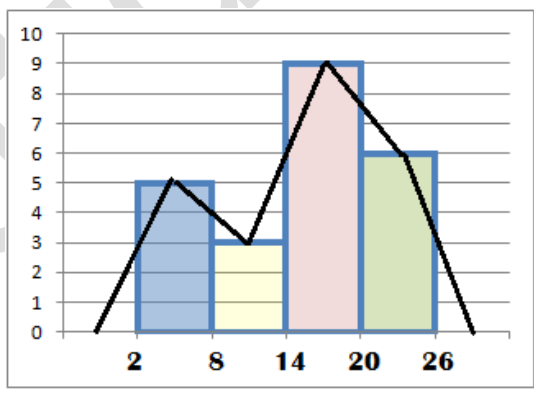
- Φυσικά πρόκειται για ποσοτική μεταβλητή και μάλιστα με κλάσεις ίσου πλάτους c .
- Χωρίζω τον άξονα x σύμφωνα με με το πλάτος c της κλάσης. Άρα έχω ίσα διαστήματα.
- Η θέση των αριθμών είναι στο σημείο αρίθμησης και όχι στο κέντρο της ράβδου.
- Προσθέτω μια κλάση μηδενικής συχνότητας πριν και μετά τα δεδομένα.
- Τα ορθογώνια κατασκευάζονται με βάση τα πλάτη, όποτε είναι διαδοχικά χωρίς κενά ανάμεσά τους.
- Το ύψος τους αντιστοιχεί στην συχνότητα. Επειδή η μονάδα μέτρησης των x είναι ίσο με ένα, το εμβαδό τους ισούται με την συχνότητα.

Κλάσεις	v_i
[2, 8)	5
[8,14)	3
[14, 20)	9
[20, 26)	6



Κατασκευή **πολυγώνου συχνοτήτων**.

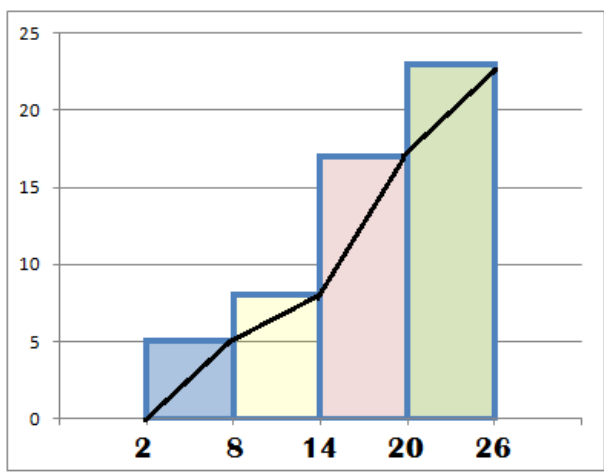
- Ενώνω τα μέσα τα μέσα των άνω βάσεων των στηλών.
- Αν στον κάθετο άξονα θέσω τις σχετικές συχνότητες θα έχω σε όμοιο σχήμα το ιστόγραμμα σχετικών συχνοτήτων και το πολύγωνο σχετικών συχνοτήτων.
- Το εμβαδό του πολυγώνου είναι ίσο με το εμβαδό των στηλών του ιστογράμματος.



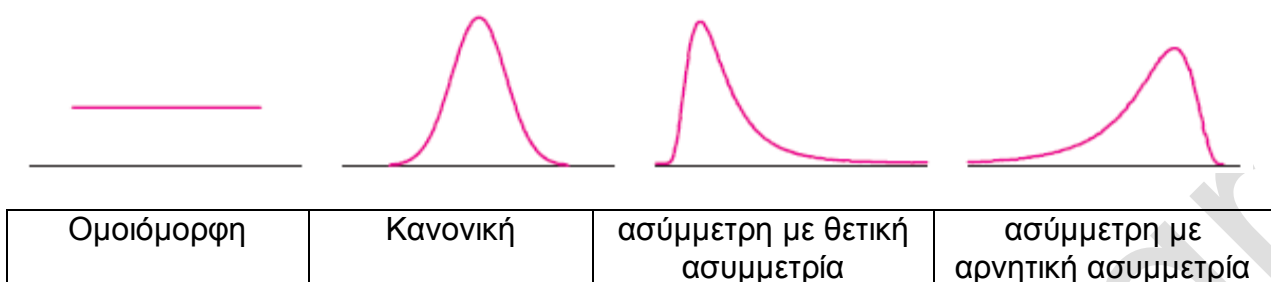
Με το παρόμοιο τρόπο κατασκευάζω **ιστόγραμμα και πολύγωνο αθροιστικών συχνοτήτων**.

- Ενώνω τα δεξιά πάνω άκρα των άνω βάσεων των στηλών.
- Αν στον κάθετο άξονα θέσω τις σχετικές συχνότητες θα έχω σε όμοιο σχήμα το ιστόγραμμα σχετικών αθροιστικών συχνοτήτων και το πολύγωνο σχετικών αθροιστικών συχνοτήτων.

Κλάσεις	v_i	N_i
[2, 8)	5	5
[8,14)	3	8
[14, 20)	9	17
[20, 26)	6	23



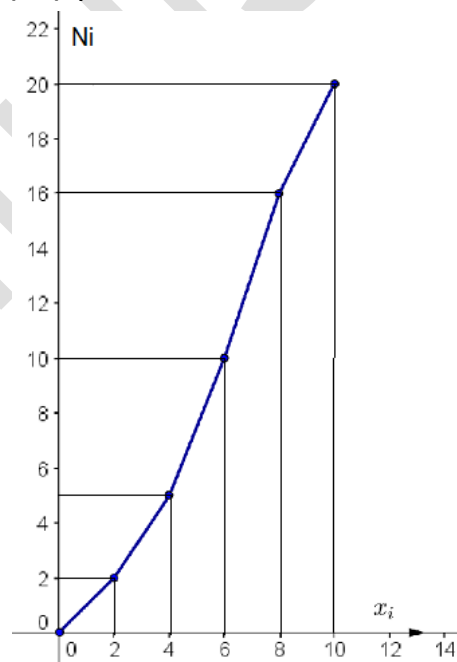
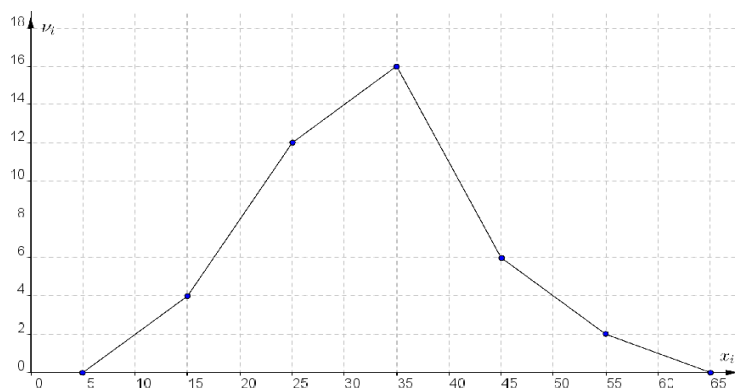
Εάν υποθέσουμε ότι ο αριθμός των κλάσεων για μια συνεχή μεταβλητή είναι αρκετά μεγάλος (τείνει στο άπειρο) και ότι το πλάτος των κλάσεων είναι αρκετά μικρό (τείνει στο μηδέν), τότε η πολυγωνική γραμμή συχνοτήτων τείνει να πάρει τη μορφή μιας ομαλής καμπύλης, η οποία ονομάζεται **καμπύλη συχνοτήτων**.



Ομαδοποίηση 80 αριθμών. Ελληνοί, Σε ιστόγραμμα πολύγωνο και πολύγωνο αθροιστικών

Θέμα 2011. Και όλα τα θέματα σε διαγράμματα. Αντίστροφη διαδικασία.

Δ1. Στο πρώτο σχήμα φαίνεται το πολύγωνο των συχνοτήτων των ηλικιών ενός χωριού.



Δ2. στο δεύτερο το πολύγωνο των αθροιστικών συχνοτήτων των ετών εργασίας, των καθηγητών σε ένα σχολείο

Να κατασκευάσετε πίνακα κατανομής συχνοτήτων με τις στήλες των: κλάσεων, v_i , N_i , k_i , f_i , $F_i\%$

Ε. ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ

Αριθμητικοί δείκτες για να περιγράψουμε σύντομα μια κατανομή.

ΜΕΤΡΑ ΘΕΣΗΣ	ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ:
η θέση του κέντρου των παρατηρήσεων στον οριζόντιο άξονα.	πόσο εκτείνονται οι παρατηρήσεις γύρω από το «κέντρο» τους
<ul style="list-style-type: none"> • Επικρατούσα τιμή • Μέση τιμή \bar{x} • Σταθμικός μέσος • Διάμεσος δ 	<ul style="list-style-type: none"> • Εύρος R • Διακύμανση s^2 • Τυπική απόκλιση s • Συντελεστής μεταβολής CV

ΠΡΟΣΟΧΗ ΤΑ ΜΕΤΡΑ ΑΦΟΡΟΥΝ ΜΟΝΟ ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ
(Μόνο η Επικρατούσα τιμή μπορεί να αναφέρεται και σε ποιοτική μεταβλητή)

ΜΕΤΡΑ ΘΕΣΗΣ

ΕΠΙΚΡΑΤΟΥΣΑ ΤΙΜΗ

Η τιμή του X_i με την μεγαλύτερη συχνότητα. (το X_i με το μεγαλύτερο v_i)

- Μπορεί να υπάρχουν περισσότερες της μιας.
- Είναι το μόνο μέτρο που μπορεί να οριστεί και για ποιοτική μεταβλητή.

** εκτός ύλης, αλλά αναφερόμενο

ΜΕΣΗ ΤΙΜΗ

<p>Αν έχω v παρατηρήσεις t_1, t_2, \dots, t_v</p>	<p>Ρωτήσαμε, τυχαία, 10 καθηγητές πόσες φορές την εβδομάδα γυμνάζονται και πήραμε τις παρακάτω απαντήσεις: 1, 0, 2, 1, 1, 4, 2, 3, 5, 3</p>
$\bar{x} = \frac{t_1 + t_2 + \dots + t_v}{v}$	$\bar{x} = \frac{\sum_{i=1}^v t_i}{v}$
$\bar{x} = \frac{1+0+2+1+1+4+2+3+5+3}{10} = \frac{23}{10} = 2,3$	

<p>Αν οι τιμές βρίσκονται σε πίνακα κατανομής συχνοτήτων με k διαφορετικές τιμές X_i (δηλ. ο πίνακας έχει k γραμμές)</p> <p>Σχηματίζω την νέα στήλη $x_i \cdot v_i$ η οποία έχει χρήση και στα μέτρα διασποράς</p>	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr style="background-color: #a6a6a6;"> <th>X_i</th> <th>V_i</th> <th>$X_i \cdot V_i$</th> </tr> </thead> <tbody> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>3</td><td>3</td></tr> <tr><td>2</td><td>2</td><td>4</td></tr> <tr><td>3</td><td>2</td><td>6</td></tr> <tr><td>4</td><td>1</td><td>4</td></tr> <tr><td>5</td><td>1</td><td>5</td></tr> <tr style="background-color: #a6a6a6;"> <td>Σύνολα</td> <td>$v = 10$</td> <td>22</td> </tr> </tbody> </table>	X_i	V_i	$X_i \cdot V_i$	0	1	0	1	3	3	2	2	4	3	2	6	4	1	4	5	1	5	Σύνολα	$v = 10$	22
X_i	V_i	$X_i \cdot V_i$																							
0	1	0																							
1	3	3																							
2	2	4																							
3	2	6																							
4	1	4																							
5	1	5																							
Σύνολα	$v = 10$	22																							

$\bar{x} = \frac{x_1 v_1 + x_2 v_2 + \dots + x_k v_k}{v}$	$\bar{x} = \frac{\sum_{i=1}^k x_i v_i}{v}$	$\bar{x} = \frac{0 \cdot 1 + 1 \cdot 3 + 2 \cdot 2 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 1}{10} = \frac{22}{10} = 2,2$
---	--	--

Ιδιότητα:

$$\bar{x} = \frac{\sum_{i=1}^k x_i v_i}{v} = \sum_{i=1}^k x_i \frac{v_i}{v} = \sum_{i=1}^k x_i f_i$$

Σε δεδομένα ομαδοποιημένα σε κλάσεις, αντί X_i χρησιμοποιούμε τις κεντρικές τιμές k_i

Κλάσεις	Κεντρικές τιμές k_i	v_i	$k_i \cdot v_i$
[2, 8)	5	5	25
[8, 14)	11	9	99
[14, 20)	17	6	102
Σύνολα		$v = 20$	226

$$\bar{x} = \frac{\sum_{i=1}^k x_i v_i}{v} = \frac{226}{20} = 11,3$$

ΣΤΑΘΜΙΚΟΣ ΜΕΣΟΣ

Είναι η μέση τιμή στην περίπτωση που έχω τιμές x_1, x_2, \dots, x_v που έχουν διαφορετικούς συντελεστές στάθμισης (βαρύτητας) w_1, w_2, \dots, w_v π.χ βαθμός πρόσβασης

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \dots + x_v w_v}{w_1 + w_2 + \dots + w_v}$$

$$\bar{x} = \frac{\sum_{i=1}^v x_i w_i}{\sum_{i=1}^v w_i}$$

ΑΣΚΗΣΕΙΣ ΠΑΝΩ ΣΤΟΝ ΤΥΠΟ ΜΕΣΗΣ ΤΙΜΗΣ

- E1.** Δίνονται οι παρατηρήσεις 2, 6, 7, 3, 7, x με μέση τιμή ίση με 6. Υπολογίστε το x.
- E2.** Αν η Μέση Τιμή των παρακάτω δεδομένων είναι 8, βρείτε το ω
8, 2, 6, 2, 8, ω + 6, 12, 8, 8, 3ω, 12, 6, 6
- E3.** Σε μια επιχείρηση, ο μέσος μηνιαίος μισθός των 9 υπαλλήλων ενός τμήματος, είναι 850 € .
 - α.** Αν προσληφθεί ένας ακόμη υπάλληλος με μισθό 650 € , ποιος θα είναι τότε ο μέσος μηνιαίος μισθός ;

- β. Μετά την πρόσληψη ενός ακόμα υπαλλήλου ο μέσος μηνιαίος μισθός ανέρχεται στα 860 € . Ποιος είναι ο μισθός του;
- γ. Για λογιστικούς λόγους, η επιχείρηση θα πρέπει να κρατήσει το μέσο μηνιαίο μισθό το πολύ έως 880 € . Πόσο χρήματα μπορεί να δώσει, το πολύ, σε ένα νέο εργαζόμενο;
- δ. Αν στο μέσο μηνιαίο μισθό των βασικών 9 υπαλλήλων, συμπεριλάβουμε και το μισθό τριών διευθυντικών στελεχών, τότε ο μέσος μισθός ανέρχεται στα 1100 €. Ποιος είναι, τότε, ο μέσος μηνιαίος μισθός των διευθυντών;

Ε4. Η Μέση Τιμή 25 παρατηρήσεων είναι 80. Αν από αυτές οι 2 μειώνονται κατά 8 και οι 3 αυξάνονται κατά 12 , τότε να βρεθεί η νέα Μέση Τιμή .

Ε5. Η Μέση Τιμή 20 παρατηρήσεων είναι 40. Ποια θα είναι η νέα Μέση Τιμή αν οι παρατηρήσεις :

- α. αυξηθούν κατά 10% ;
- β. μειωθούν κατά 10% ;

ΔΙΑΜΕΣΟΣ δ

Διάμεσος δ ενός δείγματος n παρατηρήσεων που έχουν διαταχθεί σε αύξουσα σειρά ονομάζεται:

- Η **μεσαία** παρατήρηση αν το πλήθος των παρατηρήσεων είναι περιπτό.
- Το **ημιάθροισμα** των μεσαίων παρατηρήσεων αν το πλήθος των παρατηρήσεων είναι άρτιο.

ΠΑΡΑΔΕΙΓΜΑΤΑ

έχω το δείγμα 9, 8, 3, 6, 4, 9, 8	Ή το δείγμα 9, 8, 3, 6, 4, 9, 8, 3
Μονός αριθμός παρατηρήσεων (v = 7) 3, 4, 6, 8 , 8, 9, 9 δ = 8	Ζυγός αριθμός παρατηρήσεων (v = 8) 3, 3, 4, 6 , 8 , 8, 9, 9 δ = (6+8)/2 = 7
Ή αλλιώς $7 / 2 = 3,5 \rightarrow$ η επόμενη 4^{η} παρατήρηση	$8 / 2 = 4 \rightarrow (4^{\eta} + 5^{\eta} \text{ παρατήρηση}) / 2$

ΠΑΡΑΔΕΙΓΜΑ ΜΕ ΠΙΝΑΚΑ: Οι θερμοκρασίες που μετρήσαμε το πρώτο εικοσαήμερο του Σεπτεμβρίου στην πόλη μας είναι οι παρακάτω:

20	19	17	18	19	18	18	19	17	19
17	19	20	19	20	19	17	18	18	18

Βρείτε την διάμεσο

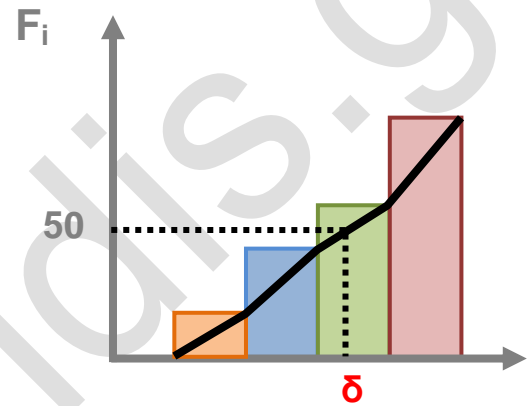
Έτσι και αλλιώς θα χρειαστεί να ταξινομήσουμε το δείγμα κατά αύξουσα σειρά. Κάνουμε πίνακα.

Θερμοκρασίες X_i	Συχνότητα v_i	Έχω 20 παρατηρήσεις δηλ ζυγός αριθμός Οι μισές είναι $20/2 = 10$. Άρα $\delta = (10^{ος} + 11^{ος})/2$ Ο $10^{ος} = 18$ Ο $11^{ος} = 19$ $\delta = (18+19) / 2 = 18,5$	Αθροιστική Συχνότητα N_i
17	4		3
18	6		8
19	7		16
20	3		19
Σύνολο	20		

Φαίνεται καλύτερα όταν έχω και στήλη N_i

Η διάμεσος σε πίνακα με κλάσεις υπολογίζεται γραφικά, από το **ιστόγραμμα σχετικών αθροιστικών συχνοτήτων %** και το **πολύγωνο συχνοτήτων**

Επειδή θέλουμε το 50% των δεδομένων κάτω από την διάμεσο και 50% πάνω από την διάμεσο, φέρω την κάθετο στο 50% και όπου συναντήσει το πολύγωνο φέρω την κάθετο στον οριζόντιο άξονα.



ΑΣΚΗΣΕΙΣ ΒΙΒΛΙΟΥ (σελ. 100)

	\bar{x}	σταθ.μεσος	δ
1			<input checked="" type="checkbox"/>
2	<input checked="" type="checkbox"/>		
4	<input checked="" type="checkbox"/>		
7			<input checked="" type="checkbox"/>
8	<input checked="" type="checkbox"/>		
9	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
11	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
12	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
13	<input checked="" type="checkbox"/>		
18	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

Άσκηση 14, σελ 101

ΕΥΡΟΣ R

Ονομάζεται η διαφορά της μικρότερης τιμής από τη μεγαλύτερη και συμβολίζεται συνήθως με το γράμμα R.

$$R = x_{\max} - x_{\min}$$

ΔΙΑΚΥΜΑΝΣΗ s^2

Διακύμανση s^2 μιας μεταβλητής X, που παίρνει ν το πλήθος τιμές t_i , $i = 1, 2, \dots, \nu$ με μέση τιμή \bar{x} ονομάζεται το πηλίκο:

$$s^2 = \frac{\sum_{i=1}^{\nu} (\bar{x} - t_i)^2}{\nu} = \frac{(\bar{x} - t_1)^2 + (\bar{x} - t_2)^2 + \dots + (\bar{x} - t_{\nu})^2}{\nu}$$

Αν οι τιμές της μεταβλητής X είναι ταξινομημένες σε πίνακα συχνοτήτων, με κ διαφορετικές τιμές, τότε:

$$s^2 = \frac{\sum_{i=1}^{\nu} v_i (\bar{X} - x_i)^2}{\nu} = \frac{v_1 (\bar{X} - x_1)^2 + v_2 (\bar{X} - x_2)^2 + \dots + v_k (\bar{X} - x_k)^2}{\nu}$$

ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ s

Τυπική απόκλιση s μιας μεταβλητής X, που παίρνει ν το πλήθος ονομάζεται η τετραγωνική ρίζα της διακύμανσης, δηλαδή:

$$s = \sqrt{s^2}$$

* * Οι άλλοι τύποι δίνονται, αρα δεν προτιμούνται.

ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΤΑΒΟΛΗΣ

Συντελεστής μεταβολής μιας ποσοτικής μεταβλητής X, που παρουσιάζει μέση τιμή \bar{x} και τυπική απόκλιση s, ονομάζεται το πηλίκο:

$$CV = \frac{s}{\bar{x}} \quad \text{ή} \quad CV = \frac{s}{\bar{x}} \cdot 100\%$$

Ένας πληθυσμός θα ονομάζεται **ομοιογενής** αν $CV \leq 10\%$

Εξάσκηση σε παραμέτρους διασποράς.

ΕΦΑΡΜΟΓΗ σελ 97. Με το προηγούμενο φύλλο.
ΑΣΚΗΣΕΙΣ σελ 101-102

15	Όλα
16	Για εξάσκηση με κομπιουτεράκι
19	Για εξάσκηση με κομπιουτεράκι
20	Τύποι διασποράς (μόνο την s^2)

www.kutidis.gr