Στα παρακάτω θεωρούμε ότι το δείγμα έχει επιλεγεί σωστά και δεν θα ασχοληθούμε με τον τρόπο επιλογή του δείγματος.

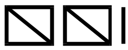
Η στατιστική με τη χρήση των δειγμάτων είναι ουσιαστικά μια επαγωγική επιστήμη αφού από το μέρος ενός συνόλου προσπαθεί να βγάλει συμπεράσματα για το όλον.

2.2. Συχνότητα Παρατηρήσεων

Στο σχολείο μας θέλουμε να εξετάσουμε τους μαθητές ως προς το χαρακτηριστικό «πόσα αδέλφια έχουν». Από τον πληθυσμό του σχολείου σχηματίζουμε ένα δείγμα 50 ατόμων, δηλαδή ν = 50. Από τις απαντήσεις που παίρνουμε σχηματίζουμε ένα πίνακα. Στη μία στήλη του πίνακα βρίσκεται ο αύξων αριθμός του μαθητή και στη δεύτερη ο αριθμός των αδελφών που έχει.

Α/Α Μαθητή	Αδέλφια	Α/Α Μαθητή	Αδέλφια	Α/Α Μαθητή	Αδέλφια
1	0	18	4	35	2
2	1	19	1	36	0
3	0	20	2	37	3
4	1	21	2	38	4
5	0	22	1	39	1
6	1	23	1	40	0
7	2	24	1	41	1
8	0	25	1	42	2
9	1	26	1	43	2
10	0	27	0	44	3
11	2	28	1	45	1
12	3	29	0	46	2
13	0	30	2	47	3
14	1	31	0	48	1
15	2	32	3	49	1
16	1	33	3	50	2
17	0	34	3		

Μία τέτοιου είδους καταγραφή των τιμών είναι μάλλον κακή. Καλύτερα θα ήταν να κάνουμε την εξής καταγραφή: στην πρώτη στήλη να γράψουμε τον αριθμό των αδελφών και στη δεύτερη των αριθμό των μαθητών που έχουν τόσα αδέλφια όσα δηλώνει η πρώτη στήλη. Η παραπάνω εργασία είναι ευκολότερη και έτσι κατασκευάζουμε τον πίνακα:

Αρ. Αδελφών	**Διαλογή**	**Μαθητές**
0		12
1		18
2		11
3		7
4		2

'Ετσι, έχουμε μετρήσει «τον αριθμό των αδελφών» στα άτομα του δείγματος και μπορούμε να πούμε ότι έχουμε την κατανομή του αριθμού των αδελφών του δείγματος.

Αρ. Αδελφών (x_i)	**Μαθητές (ν_i)**
0	12
1	18
2	11
3	7
4	2
Σύνολο	50

Είναι εμφανές ότι στην κατανομή αυτή η τιμή 1 εμφανίζεται συχνότερα από την τιμή 0 και η τιμή 3 εμφανίζεται πιο συχνά από την τιμή 4. Τον αριθμό των εμφανίσεων μιας τιμής ονομάζουμε συχνότητα εμφάνισης της τιμής ή απλά συχνότητα της τιμής και συμβολίζεται με ν_i. Δηλαδή:

- Η τιμή $x_1 = 0$ έχει συχνότητα $\nu_1 = 12$
- Η τιμή $x_2 = 1$ έχει συχνότητα $\nu_2 = 18$

- Η τιμή $x_3 = 2$ έχει συχνότητα $\nu_3 = 11$
- Η τιμή $x_4 = 3$ έχει συχνότητα $\nu_4 = 7$
- Η τιμή $x_5 = 4$ έχει συχνότητα $\nu_5 = 2$

Αν το δείγμα εμφανίζει κ διαφορετικές τιμές της μεταβλητής με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots, \nu_\kappa$ τότε προφανώς ισχύει ότι:

$$\nu_1 + \nu_2 + \cdots + \nu_\kappa = \nu$$

Η συχνότητα ουσιαστικά μας πληροφορεί ότι ο αριθμός 3 εμφανίζεται 7 φορές στις 50. Βέβαια, ο άνθρωπος εν γένει έχει συνηθίσει να εκφράζει τις μεταβολές επί της εκατό (%), δηλαδή αντιλαμβάνεται καλύτερα κάποια αποτελέσματα, όταν αυτά δίνονται κατ' αναλογία του 100 και όχι του 50 ή οποιουδήποτε άλλου αριθμού. Για το λόγο αυτό χρησιμοποιούμε συνήθως τη έννοια της σχετικής συχνότητας, η οποία ορίζεται ως εξής:

Αν ν είναι το μέγεθος ενός δείγματος (πληθυσμού) τότε το πηλίκο $\frac{\nu_i}{\nu}$ ονομάζεται σχετική συχνότητα και συμβολίζεται με f_i δηλαδή:

$$f_i = \frac{\nu_i}{\nu} = \frac{\text{αριθμός εμφανίσεων}}{\text{μέγεθος δείγματος}}$$

Είναι φανερό ότι $f_1 + f_2 + \cdots + f_\kappa = 1$.

Αρ. Αδελφών (x_i)	**Συχνότητα (ν_i)**	**Σχ. συχνότητα (f_i)**	**Σχ. συχνότητα (%)**
0	12	0,24	24
1	18	0,36	36
2	11	0,22	22
3	7	0,14	14
4	2	0,04	4
Αθροίσματα	50	1	100

Συνοψίζοντας, έχουμε ότι:

Ορισμός 1: Συχνότητα τιμής x_i μιας μεταβλητής ονομάζεται το πλήθος των ατόμων του πληθυσμού (ή του δείγματος) για τα οποία η μεταβλητή παίρνει την τιμή x_i και συμβολίζεται με ν_i.

Ορισμός 2: Σχετική συχνότητα τιμής x_i μιας μεταβλητής ονομάζεται ο λόγος της συχνότητας προς το μέγεθος του δείγματος και συμβολίζεται με f_i, δηλαδή $f_i = \frac{\nu_i}{\nu}$.

Η προσεκτική ανάγνωση του πίνακα μας δίνει τις πληροφορίες για τους μαθητές του συγκεκριμένου σχολείου και για το πόσα αδέλφια έχουν. Αν θέλουμε να μάθουμε πόσοι μαθητές έχουν λιγότερα από 2 αδέλφια πρέπει να προσθέσουμε τις συχνότητες των τιμών 0, 1 για λιγότερα από 3 πρέπει να προσθέσουμε τις συχνότητες των τιμών 0, 1, 2 κ.λπ. Για το λόγο αυτό κατασκευάζουμε συνήθως και μια άλλη στήλη η οποία έχει το άθροισμα των συχνοτήτων των τιμών που είναι μικρότερες ή ίσες από αυτήν. Το άθροισμα αυτό ονομάζεται αθροιστική συχνότητα. 'Εχουμε λοιπόν τον εξής πίνακα:

Αριθμός αδελφών (x_i)	**Συχνότητα (ν_i)**	**Αθροιστική συχνότητα**	**Σχετική συχνότητα (f_i)**	**Σχετική αθροιστική συχνότητα**
0	12	12	0,24	0,24
1	18	30	0,36	0,60
2	11	41	0,22	0,82
3	7	48	0,14	0,96
4	2	50	0,04	1
Αθροίσματα	50		1	

Η αθροιστική συχνότητα τιμής x_i μας δίνει την πληροφορία «Πόσα άτομα του πληθυσμού έχουν τιμή της μεταβλητής μικρότερη ή ίση από x_i»

'Οπως φαίνεται από τον παραπάνω πίνακα η σχετική αθροιστική συχνότητα του 3 είναι 0,96 το οποίο εκφράζει ότι «Το 96% των παιδιών έχουν λιγότερα από 4 αδέλφια».

Ορισμός 3: Σε ποσοτική μεταβλητή, αθροιστική συχνότητα μιας τιμής x_i λέγεται το άθροισμα των συχνοτήτων ν_i των τιμών που είναι μικρότερες ή ίσες με την τιμή αυτή.

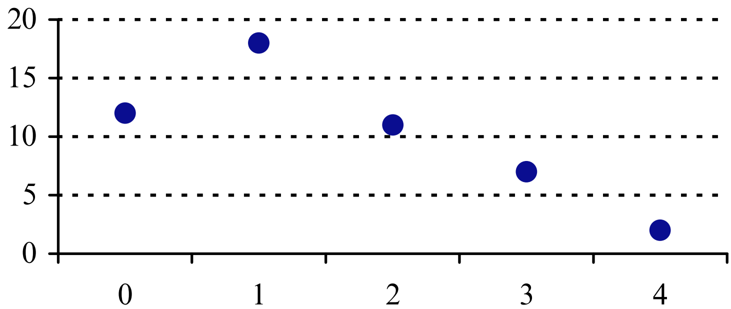
Ορισμός 4: Σε ποσοτική μεταβλητή, σχετική αθροιστική συχνότητα μιας τιμής x_i λέγεται το άθροισμα των σχετικών συχνοτήτων f_i των τιμών που είναι μικρότερες ή ίσες με την τιμή αυτή.

Είναι φανερό ότι η παραπάνω κατασκευή των πινάκων έχει νόημα μόνο αν οι τιμές της μεταβλητής έχουν διαταχθεί σε αύξουσα σειρά.

Για την καλύτερη παρουσίαση των παραπάνω συνήθως χρησιμοποιούμε γραφικές παραστάσεις.

- **Γραφικές Παραστάσεις**

Στον άξονα των τετμημένων (x′x) θεωρούμε τον αριθμό των αδελφών (0, 1, 2, 3, 4) δηλαδή τις τιμές της μεταβλητής x_i και στον άξονα των τεταγμένων (y′y) πόσοι μαθητές έχουν τον αντίστοιχο αριθμό αδελφών, δηλαδή τις τιμές των αντιστοίχων συχνοτήτων.

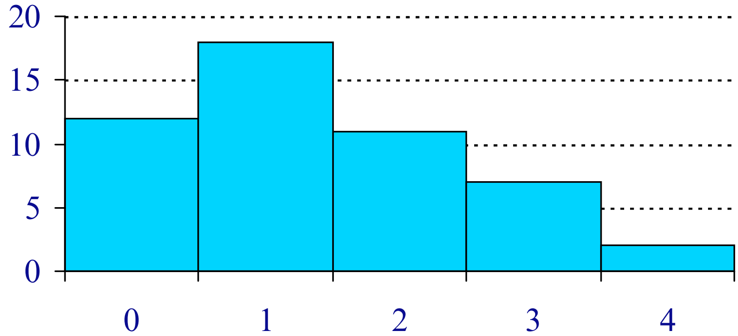


Η εμφάνιση της παραπάνω γραφικής παράστασης δεν μας επιτρέπει να καταλάβουμε γρήγορα και εύκολα την κατανομή. Για το λόγο αυτό υπάρχουν και άλλοι τύποι διαγραμμάτων.

Ραβδογράμματα

- **Κατακόρυφο ραβδόγραμμα**

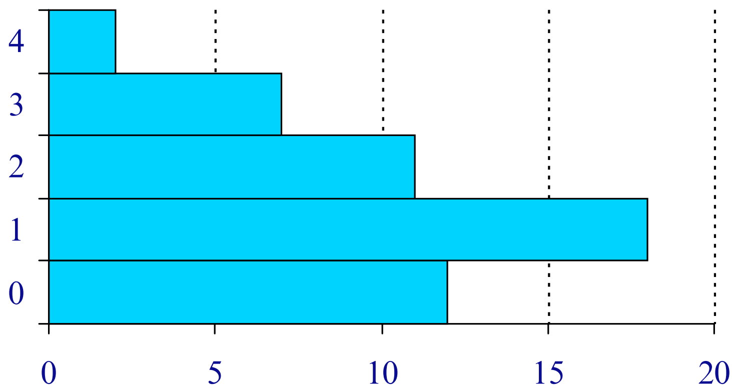
Κατασκευάζουμε όπως και προηγούμενα τον άξονα των τετμημένων, τον οποίο διαιρούμε σε πέντε ισομήκη διαστήματα σε καθένα από αυτά αντιστοιχούμε μια τιμή της μεταβλητής και σε κάθε διάστημα κατασκευάζουμε ένα ορθογώνιο, με ύψος την αντίστοιχη τιμή του.



Είναι φανερό ότι η συχνότητα της τιμής x_i ισούται με το εμβαδόν του ορθογωνίου που αντιστοιχεί σε αυτήν, αν θεωρήσουμε ως μονάδα μέτρησης το πλάτος του ορθογωνίου. Ακόμη, ισχύει ότι το άθροισμα των εμβαδών των ορθογωνίων είναι ίσο με το μέγεθος του δείγματος.

- **Οριζόντιο ραβδόγραμμα**

Κατασκευάζεται όπως και το κατακόρυφο, αλλά με εναλλαγή του ρόλου των αξόνων.



Κυκλικό διάγραμμα

Πολλές φορές εμφανίζουμε τα αποτελέσματα μιας μέτρησης με τη μορφή κυκλικού διαγράμματος. Αυτό γίνεται ως εξής:

1. Διαιρούμε το κύκλο σε μέρη ανάλογα των συχνοτήτων των τιμών της μεταβλητής .
2. Κατασκευάζουμε τους αντίστοιχους κυκλικούς τομείς και χρωματίζουμε τον καθένα με ένα διαφορετικό χρώμα. Δηλαδή:

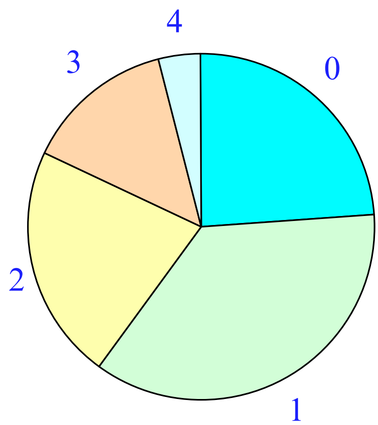
– Στο 12 αντιστοιχούμε το $M_0 = \frac{12}{50} 360 = 86° \, 24'$.

– Στο 18 αντιστοιχούμε το $M_1 = \frac{18}{50} 360 = 129° \, 36'$.

– Στο 11 αντιστοιχούμε το $M_2 = \frac{11}{50} 360 = 79° \, 12'$.

– Στο 7 αντιστοιχούμε το $M_3 = \frac{7}{50} 360 = 50° \, 24'$.

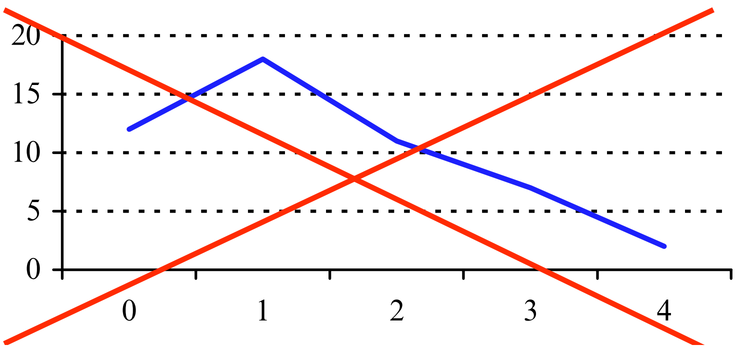
– Στο 4 αντιστοιχούμε το $M_4 = \frac{4}{50} 360 = 14° \, 24'$.



Σημείωση: *Η μονάδα μέτρησης των γωνιών ως γνωστόν είναι η μοίρα η οποία όμως υποδιαιρείται σε 60 πρώτα λεπτά και το κάθε λεπτό σε 60 δεύτερα. 'Ετσι δεν είναι σωστό να γράφουμε 45,4 μοίρες αλλά 45 μοίρες και* $\frac{4}{10} 60 = 24$ *πρώτα λεπτά δηλαδή* $45° \, 24'$*.*

ΠΡΟΣΟΧΗ !!!

Σε μία κλασική γραφική παράσταση σημείων (σε διακριτή μεταβλητή), είναι λάθος να συνδέσουμε τα μεμονωμένα σημεία μεταξύ τους με ευθύγραμμα τμήματα.

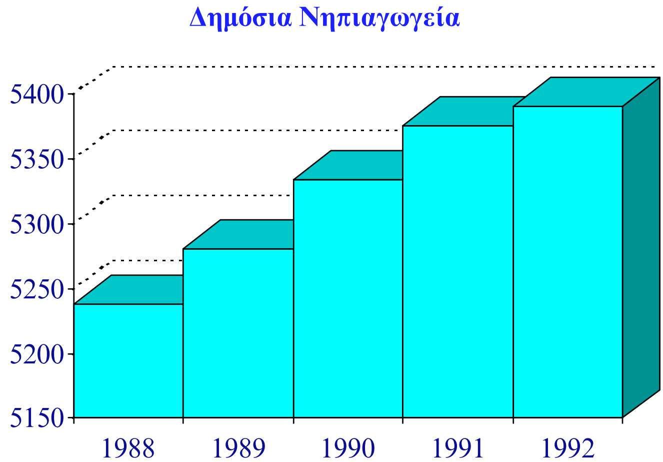


Παράδειγμα 1

Από τη στατιστική επετηρίδα της Εθνικής Στατιστικής Υπηρεσίας της Ελλάδος παίρνουμε τον παρακάτω πίνακα για τα δημόσια νηπιαγωγεία που λειτούργησαν κατά τα έτη 1988 - 1992.

Δημόσια Νηπιαγωγεία	
'Ετος	**Σχολικές μονάδες**
1988	5238
1989	5281
1990	5334
1991	5376
1992	5391

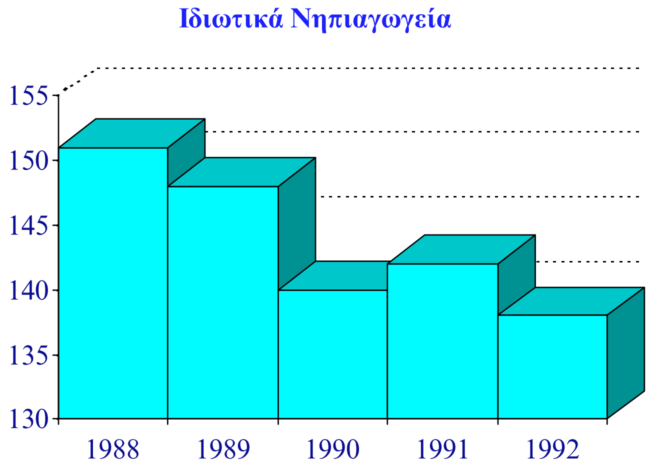
Τα ραβδογράμματα πολλές φορές τα κατασκευάζουμε τρισδιάστατα:



Αν κάνουμε την ίδια μελέτη και για τα ιδιωτικά νηπιαγωγεία θα διαπιστώσουμε ότι:

Ιδιωτικά Νηπιαγωγεία	
'Ετος	**Σχολικές μονάδες**
1988	151
1989	148
1990	140
1991	142
1992	138

Οπότε το αντίστοιχο ραβδόγραμμα θα είναι:



Από τη σύγκριση αυτών των δύο διαγραμμάτων συμπεραίνουμε ότι υπάρχει μία αύξηση (μικρή βέβαια σε σχέση με τα υπάρχοντα) στα δημόσια και συγχρόνως μία μικρή μείωση των ιδιωτικών νηπιαγωγείων.

Πολλές φορές για τη μελέτη της ίδιας μεταβλητής σε δύο πληθυσμούς (π.χ άρρενες, θήλεις) κατασκευάζουμε για κάθε τιμή της μεταβλητής δύο ορθογώνια, ένα για κάθε πληθυσμό, με σκοπό να έχουμε εύκολη σύγκριση.

Παράδειγμα 2

Από το Κέντρο Ερευνών για Θέματα Ισότητας (Κ.Ε.Θ.Ι) έχουμε τον παρακάτω πίνακα:

Εργατικό δυναμικό, απασχολούμενοι, άνεργοι και μη ενεργοί κατά φύλλο το 1996 στην Κρήτη

Κατηγορία	'Ανδρες	Γυναίκες
Πληθυσμός Εργάσιμης Ηλικίας	196960	215134
Εργατικό Δυναμικό	135432	95823
Απασχολούμενοι	133138	89181
Ενεργοί	2294	6642
Ενεργοί Μακράς Διάρκειας	785	3140
Μη Ενεργοί	61527	119311