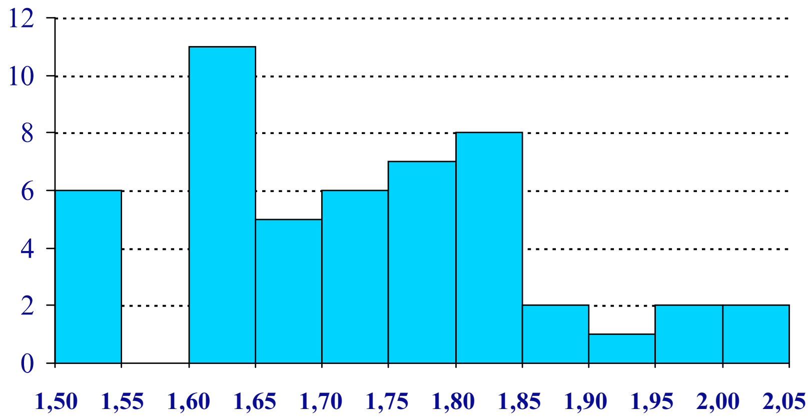
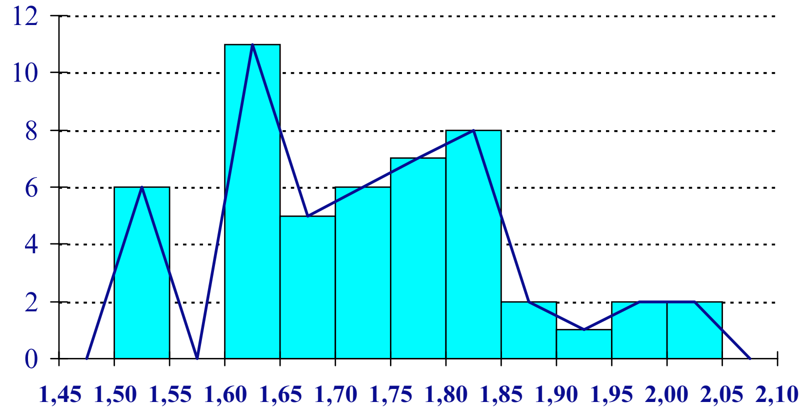
και καταγράφουμε πόσα άτομα ανήκουν σε κάθε διάστημα. Το μέσον κάθε διαστήματος το ονομάζουμε κέντρο της κλάσης.

Διαστήματα	**Συχνότητα**	**Σχετική Συχνότητα**
[1,50, 1,55)	6	0,12
[1,55, 1,60)	0	0
[1,60, 1,75)	11	0,22
[1,65, 1,70)	5	0,10
[1,70, 1,75)	6	0,12
[1,75, 1,80)	7	0,14
[1,80, 1,85)	8	0,16
[1,85, 1,90)	2	0,04
[1,90, 1,95)	1	0,02
[1,95, 2,00)	2	0,04
[2,00, 2,05)	2	0,04
Αθροίσματα	50	1

Το αντίστοιχο ιστόγραμμα της παραπάνω κατανομής είναι το ακόλουθο:

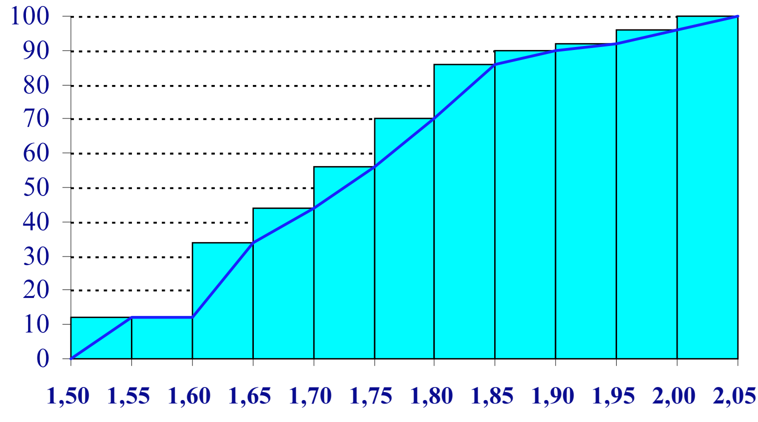


Συνήθως σε κάθε τέτοιο ιστόγραμμα, προσπαθούμε να αντιστοιχίσουμε και ένα πολύγωνο συχνοτήτων. Έτσι, κατασκευάζουμε το ιστόγραμμα κανονικά, εμφανίζοντας όμως και το διάστημα [1,45, 1,50) πριν το [1,50, 1,55) και το [2,05, 2,10) κατόπιν του διαστήματος [2,00, 2,05). Στη συνέχεια, σημειώνουμε το μέσο της πάνω οριζόντιας γραμμής κάθε ορθογωνίου, καθώς και τα μέσα του προηγούμενου διαστήματος [1,45, 1,50) και του επόμενου [2,05, 2,10) και τα ενώνουμε, κατασκευάζοντας έτσι το πολύγωνο συχνοτήτων της κατανομής αυτής:



Το πολύγωνο συχνοτήτων με τον οριζόντιο άξονα σχηματίζει χωρίο ίσου εμβαδού με αυτό που σχηματίζουν τα ορθογώνια αφού παρατηρούμε ότι για κάθε κομμάτι που προσθέτουμε (με φόντο άσπρο) αφαιρείται και ένα ισεμβαδικό κομμάτι (με φόντο γαλάζιο).

Για το ιστόγραμμα της αθροιστικής συχνότητας - σχετικής αθροιστικής συχνότητας ακολουθείται η παραπάνω διαδικασία. Το πολύγωνο αθροιστικών - σχετικών αθροιστικών συχνοτήτων σχηματίζεται ενώνοντας τα δεξιά άκρα (όχι τα μέσα) των άνω βάσεων των ορθογωνίων με ευθύγραμμα τμήματα:



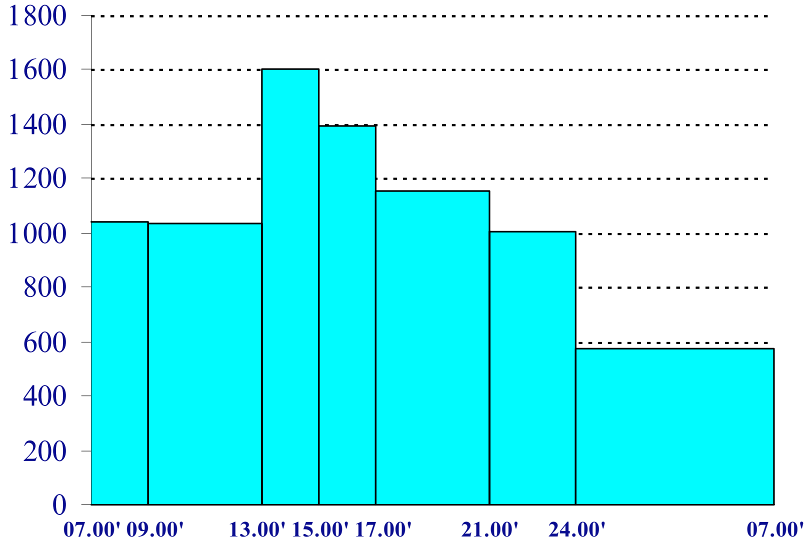
Στο προηγούμενο παράδειγμα επιλέξαμε κλάσεις ίσου πλάτους. Υπάρχουν όμως περιπτώσεις που είναι απαραίτητο να έχουμε κλάσεις με διαφορετικά πλάτη.

'Ωρα	**Ατυχήματα**
07.00′ – 09.00′	2.080
09.00′ – 13.00′	4.141
13.00′ – 15.00′	3.210
15.00′ – 17.00′	2.790
17.00′ – 21.00′	4.625
21.00′ – 24.00′	3.021
24.00′ – 07.00′	4.016
Σύνολο	23.883

Προφανώς η κατασκευή ενός ιστογράμματος με ίσες κλάσεις και ύψη ορθογωνίων τη συχνότητας είναι λάθος. Το ύψος κάθε ορθογωνίου πρέπει να είναι αντίστροφα ανάλογο του πλάτους της κλάσης. Για το λόγο αυτό εκτός από τη συχνότητα ορίζουμε και δύο νέες στήλες: η μία είναι το πλάτος της κλάσης και η άλλη το πηλίκο της συχνότητας προς το πλάτος της κλάσης.

'Ωρα	ν_i	c_i	$\frac{\nu_i}{c_i}$
07.00′ – 09.00′	2080	2	1040
09.00′ – 13.00′	4141	4	1035,25
13.00′ – 15.00′	3210	2	1605
15.00′ – 17.00′	2790	2	1395
17.00′ – 21.00′	4625	4	1156,25
21.00′ – 24.00′	3021	3	1007
24.00′ – 07.00′	4016	7	573,71
Σύνολο	23883		

Κατασκευάζουμε το ιστόγραμμα συχνοτήτων με ύψος κάθε ορθογωνίου το $\frac{v_i}{c_i}$.



2.4. Παράμετροι Θέσης

- **Επικρατούσα τιμή**

> ***Ορισμός 1:*** Επικρατούσα τιμή μιας μεταβλητής ονομάζεται η τιμή με τη μεγαλύτερη συχνότητα. Αν δύο ή περισσότερες τιμές έχουν τη μέγιστη συχνότητα τότε υπάρχουν περισσότερες από μία επικρατούσες τιμές.

Για παράδειγμα:

- Αν η μεταβλητή παίρνει τις τιμές 1, 1, 2, 3, 4, 4, 4, 5, 6, 6, 7, 7, 7, 7, 8, 8, 9 τότε έχουμε τον πίνακα συχνοτήτων:

x_i	1	2	3	4	5	6	7	8	9
v_i	2	1	1	3	1	2	4	2	1

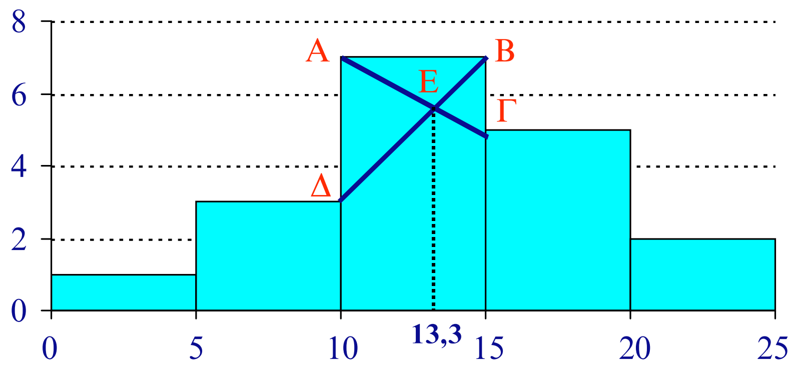
Η επικρατούσα τιμή είναι 7.

- Αν η μεταβλητή παίρνει τις τιμές 1, 1, 2, 2, 2, 3, 4, 5, 6, 6, 6, 7, 7, 8, 9, 9 τότε έχουμε τον πίνακα συχνοτήτων:

x_i	1	2	3	4	5	6	7	8	9
ν_i	2	3	1	1	1	3	2	1	2

Οι επικρατούσες τιμές είναι οι 2, 6.

- Σε συνεχή μεταβλητή ορίζουμε ως επικρατούσα κλάση αυτή με τη μεγαλύτερη συχνότητα. Η επικρατούσα τιμή υπολογίζεται γραφικά όπως φαίνεται στο παρακάτω σχήμα:



Στο παράδειγμα αυτό η επικρατούσα τιμή βρίσκεται από την τετμημένη του σημείου τομής των ευθύγραμμων τμημάτων ΑΓ και ΒΔ και είναι προσεγγιστικά 13,3.

- **Μέση τιμή**

Η μέση τιμή διαφόρων τιμών, είναι γνωστή ως το πηλίκο του αθροίσματος των τιμών προς το πλήθος τους. Για παράδειγμα ο βαθμός του απολυτηρίου ενός μαθητή είναι ο μέσος όρος των βαθμών των μαθημάτων που διδάσκεται. Ο μέσος όρος υπολογίζεται μόνο σε ποσοτικές μεταβλητές. Ο μέσος όρος ποιοτικών χαρακτηριστικών δεν ορίζεται στη στατιστική. Οι εκφράσεις της μορφής «μέσος άνθρωπος» είναι κοινωνιολογικές εκφράσεις και προφανώς περιέχουν μια αοριστία.

- **Διακριτή μεταβλητή**

Ας εξετάσουμε ένα δείγμα πληθυσμού 50 ανθρώπων με μεταβλητή «πόσες πιστωτικές κάρτες έχουν» και έστω ότι έχουμε τα παρακάτω αποτελέσματα.

0 - 0 - 0 - 0 - 0 - 0 - 0 - 0 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -
1 - 1 - 1 - 1 - 1 - 1 - 1 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 -
2 - 3 - 3 - 3 - 3 - 3 - 3 - 3 - 4 - 4

Τότε η μέση τιμή τους ισούται με το άθροισμά τους διά του πλήθους, δηλαδή:

$$0+0+0+0+0+0+0+0+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+2+2+2+2+2+2+2+2+2+2+2+2+2+2+3+3+3+3+3+3+3+4+4=76$$

Οπότε ο μέσος όρος είναι $\frac{76}{50} = 1{,}52$.

Αν όμως έχουμε τον πίνακα των συχνοτήτων τότε έχουμε:

Α/Α	$\mathbf{x_i}$	$\mathbf{\nu_i}$
1	0	8
2	1	19
3	2	14
4	3	7
5	4	2
Αθροίσματα		50

Χρησιμοποιώντας τον πίνακα των συχνοτήτων η μέση τιμή μπορεί να υπολογιστεί από τον τύπο $\frac{\nu_1 x_1 + \nu_2 x_2 + \nu_3 x_3 + \nu_4 x_4 + \nu_5 x_5}{\nu_1 + \nu_2 + \nu_3 + \nu_4 + \nu_5}$, αφού το άθροισμα $\nu_1 x_1 + \nu_2 x_2 + \nu_3 x_3 + \nu_4 x_4 + \nu_5 x_5$ εκφράζει τον αριθμό των πιστωτικών καρτών που υπάρχουν στο δείγμα και το $\nu_1 + \nu_2 + \nu_3 + \nu_4 + \nu_5$ το πλήθος των ατόμων του δείγματος.

Τη μέση τιμή μιας μεταβλητής x την συμβολίζουμε με $\overline{x}$.

Αν η μεταβλητή παρουσιάζει $x_1, x_2, \ldots, x_\kappa$ τιμές με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots, \nu_\kappa$, τότε η μέση τιμή της μεταβλητής δίνεται από τον τύπο:

$$\overline{x} = \frac{\nu_1 x_1 + \nu_2 x_2 + \cdots + \nu_\kappa x_\kappa}{\nu_1 + \nu_2 + \cdots + \nu_\kappa} = \frac{\nu_1 x_1 + \nu_2 x_2 + \cdots + \nu_\kappa x_\kappa}{\nu}$$

Για τον πιο εύκολο υπολογισμό της μέσης τιμής δημιουργούμε μια νέα στήλη στην οποία υπολογίζονται τα $\nu_i x_i$ και κατόπιν το άθροισμά τους.

Α/Α	x_i	v_i	$v_i x_i$
1	0	8	0
2	1	19	19
3	2	14	28
4	3	7	21
5	4	2	8
Αθροίσματα		50	76

Καταλήγουμε λοιπόν πάλι, ότι η μέση τιμή είναι $\frac{76}{50} = 1{,}52$.

- **Συνεχής μεταβλητή**

Ας μετρήσουμε το βάρος 50 αρνιών, οπότε έχουμε τις τιμές:

10,03 - 9,57 - 9,35 - 7,79 - 8,86 - 9,83 - 10,64 - 9,76 - 10,30 - 8,12 - 10,84 - 9,60 - 9,27 - 8,75 - 10,25 - 7,03 - 9,28 - 10,64 - 8,25 - 10,80 - 9,35 - 10,37 - 8,79 - 8,26 - 7,59 - 10,33 - 8,99 - 7,63 - 9,25 - 10,11 - 10,22 - 7,40 - 8,23 - 8,87 - 9,03 - 10,65 - 8,27 - 9,90 - 9,33 - 9,77 - 10,61 - 9,75 - 9,34 - 7,57 - 10,02 - 10,54 - 10,49 - 10,42 - 10,69 - 7,83

Γνωρίζουμε ότι η μέση τιμή του βάρους είναι ίση με το πηλίκο του αθροίσματος των τιμών αυτών διά το πλήθος των τιμών. Αν προσθέσουμε όλες τις τιμές βρίσκουμε 468,56, οπότε έχουμε μέση τιμή $\frac{468{,}74}{50} = 9{,}37$.

Ο υπολογισμός της μέσης τιμής στην παραπάνω κατανομή μέσω της ομαδοποίησης γίνεται ως εξής:

Διάστημα	**Συχνότητα v_i**	**Μέσο διαστήματος O_i**	**$v_i O_i$**
7,0 – 7,5	2	7,25	14,5
7,5 – 8,0	4	7,75	31
8,0 – 8,5	6	8,25	49,5
8,5 – 9,0	5	8,75	43,75
9,0 – 9,5	8	9,25	74
9,5 – 10,0	7	9,75	68,25
10,0 – 10,5	10	10,25	102,5
10,5 – 11,0	8	10,75	86
Αθροίσματα	50		469,5

Σε κάθε διάστημα υπολογίζουμε το μέσον του O_i, για παράδειγμα, το μέσον του διαστήματος από το 7 έως 7,5 είναι $\frac{7+7,5}{2} = 7,25$, οπότε έχουμε $O_1 = 7,25$. 'Ομοια υπολογίζουμε και τα υπόλοιπα μέσα διαστημάτων.

Στη συνέχεια, πολλαπλασιάζουμε κάθε μέσον διαστήματος με την αντίστοιχη συχνότητα του διαστήματος, δηλαδή σχηματίζουμε μια στήλη με στοιχεία τα $\nu_i O_i$ και υπολογίζουμε το άθροισμα αυτών που είναι 469,5. Διαιρούμε το άθροισμα αυτό με το μέγεθος του δείγματος ($\nu = 50$) και βρίσκουμε τη μέση τιμή του. Οπότε η μέση τιμή είναι $\frac{469,5}{50} = 9,39$.

Παρατηρούμε ότι οι δύο μέσες τιμές που υπολογίσθηκαν διαφέρουν κατά 0,02 και το σφάλμα αυτό παρουσιάζεται λόγω του έμμεσου υπολογισμού του αθροίσματος όλων των τιμών.

Η μέση τιμή είναι μία παράμετρος θέσεως γιατί μας πληροφορεί για τον μέσο όρο του δείγματος. Πολλές φορές όμως δεν είναι αξιόπιστο στοιχείο. Το επόμενο παράδειγμα είναι χαρακτηριστικό.

Παράδειγμα

Σε ένα εργοστάσιο έχουμε 100 εργάτες με καθαρές αποδοχές 175.000 το μήνα, 12 προϊστάμενους με 380.000, 12 συμβούλους με 620.000, 2 γενικούς διευθυντές με 1.200.000 και οι ιδιοκτήτες είναι 4 και αποτελούν το Διοικητικό συμβούλιο με αποδοχές 2.400.000 ο καθένας. 'Εχουμε:

x_i	ν_i	$\nu_i x_i$
165000	100	16500000
380000	12	4560000
620000	12	7440000
1200000	2	2400000
2400000	4	9600000
	130	40500000

Οπότε η μέση τιμή θα είναι $\frac{40500000}{130} = 311538,5$.

Με άλλα λόγια, η μέση τιμή δεν είναι επαρκής πληροφορία για την κατανομή της μεταβλητής. Για το λόγο αυτό ορίζεται στη στατιστική και η διάμεσος.

- **Διάμεσος**

Αν έχουμε τις 9 αριθμητικές τιμές 1, 1, 2, 3, 3, 4, 4, 5, 6, τότε ο μεσαίος όρος αυτών είναι ο $5^{ος}$, δηλαδή ο 3. Αν έχουμε τις 10 αριθμητικές τιμές 1, 1, 2, 2, 3, 7, 7, 8, 9, 9 τότε δεν υπάρχει μεσαίος όρος. Τότε ως μεσαίο όρο θεωρούμε το ημιάθροισμα των δύο μεσαίων, δηλαδή του $5^{ου}$ και του $6^{ου}$ όρου, δηλαδή $\frac{3+7}{2} = 5$. Γενικά έχουμε τον ακόλουθο ορισμό:

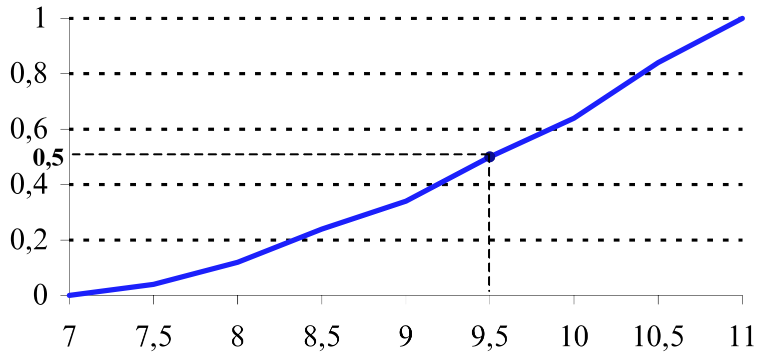
Ορισμός 1: Διάμεσος δ ενός δείγματος ν παρατηρήσεων που έχουν διαταχθεί σε αύξουσα σειρά ονομάζεται:

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

Στο προηγούμενο παράδειγμα, η μεσαία παρατήρηση είναι προφανώς 165.000 (αφού περισσότεροι από τους μισούς εργαζόμενους πληρώνονται με 165.000) και η μεγάλη διαφορά που υπάρχει μεταξύ μέσης τιμής και διαμέσου μας δίνει την πληροφορία ότι η κατανομή δεν είναι «σωστή».
Μπορούμε να υπολογίσουμε τη διάμεσο από την αθροιστική συχνότητα ή τη σχετική αθροιστική συχνότητα. Συνήθως η διαδικασία αυτή εφαρμόζεται στις ομαδοποιημένες παρατηρήσεις.
Στο παράδειγμα που είδαμε προηγούμενα με το βάρος των 50 αρνιών έχουμε τον ακόλουθο πίνακα:

Διάστημα	**Συχνότητα** ν_i	**Αθροιστική**	**Σχετική** f_i	**Σχετική αθροιστική**
7,0 – 7,5	2	2	0,04	0,04
7,5 – 8,0	4	6	0,08	0,12
8,0 – 8,5	6	12	0,12	0,24
8,5 – 9,0	5	17	0,10	0,34
9,0 – 9,5	8	25	0,16	0,50
9,5 – 10,0	7	32	0,14	0,64
10,0 – 10,5	10	42	0,20	0,84
10,5 – 11,0	8	50	0,16	1
Αθροίσματα	50			

Ας χαράξουμε τη γραφική παράσταση της σχετικής αθροιστικής συχνότητας:



Για την τιμή 0,5 του άξονα των σχετικών αθροιστικών συχνοτήτων (yy′) βρίσκουμε την αντίστοιχη τιμή της μεταβλητής η οποία είναι 9,5.

- **Σύγκριση παραμέτρων θέσης**

Για τις τρεις παραμέτρους θέσης έχουμε τα εξής:

- Η μέση τιμή επηρεάζεται από τις ακραίες τιμές και εξαρτάται από όλες τις τιμές της μεταβλητής.
- Η επικρατούσα τιμή εξαρτάται μόνο τη μεγαλύτερη τιμή
- Η διάμεσος δεν επηρεάζεται από τις ακραίες τιμές και εξαρτάται από όλες τις τιμές της μεταβλητής. Ο υπολογισμός της διαμέσου παρουσιάζει δυσκολίες σε ορισμένες περιπτώσεις (π.χ. σε συνεχή μεταβλητή).

2.5. Παράμετροι Διασποράς

Η σύγκριση ερευνών με το ίδιο αντικείμενο σε διαφορετικούς πληθυσμούς είναι ένας από τους σκοπούς της Στατιστικής. Με τις παραμέτρους θέσης έχουμε κάποιες πληροφορίες για τις κατανομές.

Είναι όμως δυνατόν σε δύο κατανομές (σε διαφορετικούς πληθυσμούς) της ίδιας μεταβλητής να έχουμε την ίδια παράμετρο θέσης (π.χ. την ίδια μέση τιμή) και οι τιμές της μιας κατανομής να είναι «κοντά» στη μέση τιμή ενώ στην άλλη κατανομή «μακριά» από αυτήν.

Είναι λογικό λοιπόν να εξετάζουμε τη διασπορά των τιμών της μεταβλητής γύρω από την παράμετρο θέσης και συνήθως αυτό γίνεται για τη μέση τιμή.

Ας υποθέσουμε ότι έχουμε δύο τμήματα της Β′ Λυκείου και εξετάζουμε τους μαθητές ως προς τη μεταβλητή «Βαθμός στα Μαθηματικά».

Τμήμα A_1			**Τμήμα A_2**		
x_i	ν_i	$\nu_i x_i$	x_i	ν_i	$\nu_i x_i$
8	3	24	8	0	0
9	3	27	9	0	0
10	4	40	10	4	40
11	2	22	11	3	33
12	2	24	12	4	48
13	0	0	13	1	13
14	0	0	14	2	28
15	1	15	15	4	60
16	2	32	16	6	96
17	1	17	17	1	17
18	2	36	18	0	0
19	2	38	19	0	0
20	3	60	20	0	0
Αθροίσματα	25	335	Αθροίσματα	25	335

Μέση τιμή $\overline{x} = 13{,}4$ Μέση τιμή $\overline{x} = 13{,}4$

Οι κατανομές αυτές έχουν την ίδια μέση τιμή (παράμετρος θέσης), όμως παρατηρούμε ότι:

- Στο τμήμα A_1 η μέση τιμή επηρεάζεται πολύ από την ύπαρξη πολύ καλών μαθητών όπως και από μαθητές με μικρή βαθμολογία δηλαδή έχουμε μεγάλη διασπορά τιμών.
- Στο τμήμα A_2 οι βαθμολογίες είναι ανάμεσα στο 10 και 17 και έτσι έχουμε μικρότερη διασπορά τιμών.

Οι αριθμοί που μας πληροφορούν για τη διασπορά των τιμών είναι το εύρος, η μέση απόλυτη απόκλιση, η διακύμανση, και η τυπική απόκλιση.

- **Εύρος**

> ***Ορισμός 1:*** Εύρος είναι η διάφορα της μικρότερης τιμής από τη μεγαλύτερη.

Είναι προφανές ότι το εύρος υπολογίζεται εύκολα αν οι τιμές της μεταβλητής έχουν διαταχθεί. Για παράδειγμα:

- Στο τμήμα $Α_1$ το εύρος είναι $20 - 8 = 12$.
- Στο τμήμα $Α_2$ το εύρος είναι $17 - 10 = 7$.

Το εύρος ως παράμετρος διασποράς έχει το μειονέκτημα ότι χρησιμοποιούμε μόνο τις ακραίες τιμές της μεταβλητής και καθόλου τις υπόλοιπες.

- **Μέση απόλυτη απόκλιση**

> ***Ορισμός 1:*** Αν μια μεταβλητή παίρνει τις ν τιμές $t_1, t_2, \ldots, t_\nu$ που έχουν μέση τιμή $\overline{x}$ τότε μέση (απόλυτη) απόκλιση της μεταβλητής ονομάζεται το πηλίκο:
>
> $$\frac{|\overline{x} - t_1| + |\overline{x} - t_2| + \cdots + |\overline{x} - t_\nu|}{\nu}$$

Ισοδύναμος είναι και ο παρακάτω ορισμός:

> ***Ορισμός 2:*** Αν μια μεταβλητή παίρνει τιμές $x_1, x_2, \ldots, x_\kappa$ με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots, \nu_\kappa$ που έχουν μέση τιμή $\overline{x}$ τότε μέση (απόλυτη) απόκλιση της μεταβλητής ονομάζεται το πηλίκο:
>
> $$\frac{\nu_1|\overline{x} - x_1| + \nu_2|\overline{x} - x_2| + \cdots + \nu_\kappa|\overline{x} - x_\kappa|}{\nu}$$

Αυτός ο τύπος της απόλυτης απόκλισης είναι πιο εύχρηστος για τον υπολογισμό της. Αρκεί να σχηματίσουμε στον πίνακα συχνοτήτων μια νέα στήλη με τα γινόμενα $\nu_i|\overline{x} - x_i|$, να τα προσθέσουμε και να διαιρέσουμε με το μέγεθος του δείγματος.

Τμήμα A_1

x_i	ν_i	$\nu_i x_i$	$\lvert\bar{x} - x_i\rvert$	$\nu_i\lvert\bar{x} - x_i\rvert$
8	3	24	5,4	16,2
9	3	27	4,4	13,2
10	4	40	3,4	13,6
11	2	22	2,4	4,8
12	2	24	1,4	2,8
13	0	0	0,4	0
14	0	0	0,6	0
15	1	15	1,6	1,6
16	2	32	2,6	5,2
17	1	17	3,6	3,6
18	2	36	4,6	9,2
19	2	38	5,6	11,2
20	3	60	6,6	19,8
Αθροίσματα	25	335		101,2

Η μέση τιμή τότε είναι $\bar{x} = \dfrac{335}{25} = 13{,}4$ και η μέση απόλυτη απόκλιση $\dfrac{101{,}2}{25} = 4{,}048$.

Τμήμα A_2

x_i	ν_i	$\nu_i x_i$	$\lvert\bar{x} - x_i\rvert$	$\nu_i\lvert\bar{x} - x_i\rvert$
8	0	0	5,4	0
9	0	0	4,4	0
10	4	40	3,4	13,6
11	3	33	2,4	7,2
12	4	48	1,4	5,6
13	1	13	0,4	0,4
14	2	28	0,6	1,2
15	4	60	1,6	6,4
16	6	96	2,6	15,6
17	1	17	3,6	3,6
18	0	0	4,6	0
19	0	0	5,6	0
20	0	0	6,6	0
Αθροίσματα	25	335		53,6

Στο παράδειγμα αυτό η μέση τιμή είναι $\overline{x} = \frac{335}{25} = 13{,}4$ και η μέση απόλυτη απόκλιση $\frac{53{,}6}{25} = 2{,}144$.

'Ετσι έχουμε ότι η «διασπορά» των βαθμών στο A_1 είναι σχεδόν διπλάσια από τη διασπορά των βαθμών στο A_2.

Η αντίρρηση που μπορεί να διατυπωθεί για την μέση απόλυτη απόκλιση είναι ότι η χρήση των απολύτων τιμών δυσκολεύει τον υπολογισμό της.

- **Διακύμανση – Τυπική απόκλιση**

Για τον υπολογισμό της μέσης απόλυτης απόκλισης χρησιμοποιούνται οι διαφορές $|\overline{x} - x_i|$. Η διασπορά των τιμών γύρω από τη μέση τιμή προσδιορίζεται καλύτερα από τις διαφορές $(\overline{x} - x_i)^2$ και οι οποίες είναι και πιο κατάλληλες για υπολογισμούς. Η μέση τιμή των διαφορών αυτών ονομάζεται διακύμανση και συμβολίζεται με s^2.

Ορισμός 1: Αν μια μεταβλητή παίρνει τις ν τιμές $t_1, t_2, \ldots, t_\nu$ που έχουν μέση τιμή $\overline{x}$ τότε διακύμανση της μεταβλητής ονομάζεται το πηλίκο:

$$s^2 = \frac{(\overline{x} - t_1)^2 + (\overline{x} - t_2)^2 + \cdots + (\overline{x} - t_\nu)^2}{\nu}$$

Ισοδύναμος είναι και ο παρακάτω ορισμός:

Ορισμός 2: Αν μια μεταβλητή παίρνει τιμές $x_1, x_2, \ldots, x_\kappa$ με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots, \nu_\kappa$ που έχουν μέση τιμή $\overline{x}$ τότε διακύμανση της μεταβλητής ονομάζεται το πηλίκο:

$$s^2 = \frac{\nu_1(\overline{x} - x_1)^2 + \nu_2(\overline{x} - x_2)^2 + \cdots + \nu_\kappa(\overline{x} - x_\kappa)^2}{\nu}$$

Ο υπολογισμός της διακύμανσης γίνεται εύκολα αν στον πίνακα συχνοτήτων προσθέσουμε μια στήλη με τα στοιχεία $\nu_i(\overline{x} - x_i)^2$.

Τμήμα A_1

x_i	v_i	$v_i x_i$	$v_i(\overline{x} - x_i)^2$
8	3	24	87,48
9	3	27	58,08
10	4	40	46,24
11	2	22	11,52
12	2	24	3,92
13	0	0	0
14	0	0	0
15	1	15	2,56
16	2	32	13,52
17	1	17	12,96
18	2	36	42,32
19	2	38	62,72
20	3	60	130,68
Αθροίσματα	25	335	472

Η μέση τιμή τότε είναι $\overline{x} = \frac{335}{25} = 13{,}4$ και η διακύμανση $s^2 = \frac{472}{25} = 18{,}88$.

Για το A_2 με τον ίδιο τρόπο έχουμε μέση τιμή $\overline{x} = \frac{335}{25} = 13{,}4$ και η διακύμανση $s^2 = \frac{136}{25} = 5{,}44$.

Η χρήση της διακύμανσης παρουσιάζει ένα σοβαρό πρόβλημα: Οι μονάδες της διακύμανσης είναι τα τετράγωνα των μονάδων της αντίστοιχης μεταβλητής. Για το λόγο αυτό αντί της διακύμανσης χρησιμοποιούμε ως μέτρο διασποράς την τετραγωνική ρίζα της διακύμανσης, που τη συμβολίζουμε με s και την ονομάζουμε τυπική απόκλιση.

Ορισμός 1: Αν μια μεταβλητή παίρνει τις ν τιμές $t_1, t_2, \ldots, t_\nu$ που έχουν μέση τιμή $\overline{x}$ τότε τυπική απόκλιση της μεταβλητής ονομάζεται το:

$$s = \sqrt{\frac{(\overline{x} - t_1)^2 + (\overline{x} - t_2)^2 + \cdots + (\overline{x} - t_\nu)^2}{\nu}}$$

Ισοδύναμος είναι και ο παρακάτω ορισμός:

Ορισμός 2: Αν μια μεταβλητή παίρνει τιμές $x_1, x_2, \ldots, x_\kappa$ με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots, \nu_\kappa$ που έχουν μέση τιμή $\overline{x}$ τότε τυπική απόκλιση της μεταβλητής ονομάζεται το:

$$s = \sqrt{\frac{\nu_1(\overline{x} - x_1)^2 + \nu_2(\overline{x} - x_2)^2 + \cdots + \nu_\kappa(\overline{x} - x_\kappa)^2}{\nu}}$$

Για παράδειγμα το A_1 έχει τυπική απόκλιση $s = \sqrt{18{,}88} = 4{,}354$ ενώ το A_2 $s = \sqrt{5{,}44} = 2{,}33$.

2.6. Συντελεστής Μεταβλητότητας (CV)

'Οταν εξετάζουμε δύο δείγματα ως προς την ίδια μεταβλητή παρουσιάζονται διαφορετικές τιμές στις παραμέτρους θέσης και διασποράς.

- Η μέση τιμή των υψών των μαθητών ενός Γυμνασίου είναι $\overline{x} = 162$ cm και η τυπική απόκλιση είναι $s = 5{,}7$ cm.
- Η μέση τιμή των υψών των μαθητών ενός Γυμνασίου είναι $\overline{x} = 181$ cm και η τυπική απόκλιση είναι $s = 5{,}8$ cm.

Είναι φανερό ότι δε μπορούμε να πούμε ότι οι δύο πληθυσμοί παρουσιάζουν τον ίδιο βαθμό μεταβλητότητας. Το παραπάνω πρόβλημα παρουσιάζεται αν μετράμε σε διαφορετικές μονάδες ή κλίμακες.

- Οι μαθητές ενός Λυκείου στην Ελλάδα εξεταζόμενοι ως προς την επίδοσή τους παρουσιάζουν μέση τιμή $\overline{x} = 14{,}5$ και τυπική απόκλιση $s = 2{,}5$ (κλίμακα βαθμολογίας 0 – 20).
- Οι μαθητές ενός Λυκείου στις Η.Π.Α (U.S.A) εξεταζόμενοι ως προς την επίδοσή τους παρουσιάζουν μέση τιμή $\overline{x} = 62$ και τυπική απόκλιση $s = 16$ (κλίμακα βαθμολογίας 0 – 100).

Στην περίπτωση αυτή δεν είναι δυνατόν να συγκριθούν τα δύο δείγματα λόγω της διαφορετικής κλίμακας βαθμολογίας.

'Ενα μέτρο με το οποίο μπορούμε να ξεπεράσουμε τις παραπάνω δυσκολίες είναι ο συντελεστής μεταβολής ή συντελεστής μεταβλητότητάς.

Ορισμός 1: Αν ένα δείγμα εξεταζόμενο ως προς μία ποσοτική μεταβλητή του, παρουσιάζει μέση τιμή $\bar{x}$ και τυπική απόκλιση s, συντελεστής μεταβολής ή συντελεστής μεταβλητότητας ονομάζεται το πηλίκο:

$$CV = \frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}} = \frac{s}{\bar{x}}\, 100\%$$

Στο πρώτο παράδειγμα:

- Οι μαθητές στο Γυμνάσιο παρουσιάζουν συντελεστή μεταβλητότητας $CV = \frac{5{,}7 \text{ cm}}{162 \text{ cm}}\, 100\% = 3{,}51\%$.
- Οι μαθητές στο Λύκειο παρουσιάζουν συντελεστή μεταβλητότητας $CV = \frac{5{,}8 \text{ cm}}{181 \text{ cm}}\, 100\% = 3{,}2\%$.

Στο δεύτερο παράδειγμα:

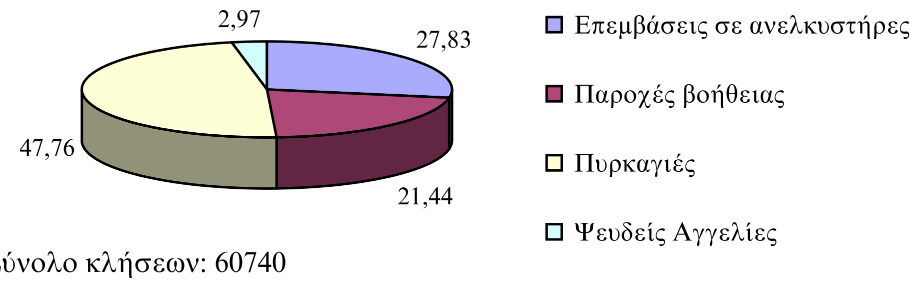
- Οι μαθητές στην Ελλάδα παρουσιάζουν συντελεστή μεταβλητότητας $CV = \frac{2{,}5}{14{,}5}\, 100\% = 17{,}24\%$.
- Οι μαθητές στις Η.Π.Α (U.S.A) παρουσιάζουν συντελεστή μεταβλητότητας $CV = \frac{16}{62}\, 100\% = 38{,}9\%$.

Ο συντελεστής μεταβλητότητας ουσιαστικά μετράει την ομοιογένεια του πληθυσμού. Εάν η τιμή του συντελεστή μεταβλητότητας είναι κάτω του 10% ο πληθυσμός του δείγματος θεωρείται ομοιογενής.

Εφαρμογές

Εφαρμογή 1

Από τη Στατιστική υπηρεσία της Πυροσβεστικής έχουμε το παρακάτω διάγραμμα:



Να βρεθούν οι συχνότητες των διαφόρων ποιοτικών χαρακτηριστικών.

Λύση

Ο τύπος που μας δίνει τη σχετική συχνότητα είναι $f_i = \frac{\nu_i}{\nu}$, οπότε έχουμε ότι $\nu_i = f_i \cdot \nu$ και εφόσον $\nu = 60740$ έχουμε:

$$\nu_1 = f_1 \cdot \nu = 0{,}2783 \cdot 60740 = 16904$$

$$\nu_2 = f_2 \cdot \nu = 0{,}2144 \cdot 60740 = 13024 \text{ κ.λπ.}$$

οπότε έχουμε τον πίνακα συχνοτήτων:

Είδος κλήσης	**f_i%**	**f_i**	**ν_i**
Επεμβάσεις σε ανελκυστήρες	27,83	0,2783	16904
Παροχές βοήθειας	21,44	0,2144	13024
Πυρκαγιές	47,76	0,4776	29010
Ψευδείς Αγγελίες	2,97	0,0297	1802
Σύνολο	100	1	

Εφαρμογή 2

Οι βαθμοί των 25 μαθητών της Γ′ τάξης ενός Λυκείου είναι:

15, 9, 17, 8, 11, 12, 18, 11, 10, 14, 17, 18,
16, 10, 11, 12, 15, 9, 17, 11, 13, 12, 13, 12, 11

(α) Να κατασκευάσετε πίνακα συχνοτήτων και αθροιστικών συχνοτήτων.

(β) Πόσοι μαθητές έχουν βαθμό μικρότερο ή ίσο από 12;

(γ) Πόσοι μαθητές έχουν βαθμό μεγαλύτερο του 14;

Λύση

(α) Κατασκευάζουμε τον πίνακα συχνοτήτων με διαλογή:

x_i	Διαλογή	ν_i
8		1
9		2
10		2
11		5
12		4
13		2
14		1
15		2
16		1
17		3
18		2
'Αθροισμα		25

Συμπληρώνουμε τον πίνακα με την αθροιστική συχνότητα και έχουμε:

x_i	ν_i	Αθροιστική
8	1	1
9	2	3
10	2	5
11	5	10
12	4	14
13	2	16
14	1	17
15	2	19
16	1	20
17	3	23
18	2	25
'Αθροισμα	25	

(β) Ο αριθμός των μαθητών που έχουν βαθμό μικρότερο ή ίσο του 12 είναι η αθροιστική συχνότητα του 12 δηλ. 14.

(γ) Ο αριθμός των μαθητών που έχουν βαθμό μικρότερο ή ίσο του 14 είναι η αθροιστική συχνότητα του 14 δηλαδή 17. Οι μαθητές που έχουν βαθμό μεγαλύτερο του 14 είναι οι υπόλοιποι, δηλαδή $25 - 17 = 8$.

Εφαρμογή 3

Για τα δεδομένα της εφαρμογής 2 να βρείτε τη μέση τιμή, τη διάμεσο, την επικρατούσα τιμή και την τυπική απόκλιση.

Λύση

Συμπληρώνουμε τον πίνακα της εφαρμογής 2 με τις στήλες $\nu_i x_i$, $\nu_i(\overline{x} - x_i)^2$:

x_i	ν_i	Αθροιστική	$\nu_i x_i$	$\nu_i(\overline{x} - x_i)^2$
8	1	1	8	27,8784
9	2	3	18	36,6368
10	2	5	20	21,5168
11	5	10	55	25,992
12	4	14	48	6,5536
13	2	16	36	0,1568
14	1	17	14	0,5184
15	2	19	30	5,9168
16	1	20	16	7,3984
17	3	23	51	41,5152
18	2	25	36	44,5568
Αθροίσματα	25		332	218,64

Οπότε έχουμε ότι:

- Η μέση τιμή $\overline{x}$ ισούται με το κλάσμα που έχει αριθμητή το άθροισμα των στοιχείων της στήλης $\nu_i x_i$ και παρανομαστή το πλήθος των στοιχείων, άρα $\overline{x} = \frac{332}{25} = 13,28$.

- Η διάμεσος είναι η μεσαία τιμή των 25 τιμών δηλαδή η $13^η$. Η $13^η$ είναι στις πρώτες δεκατέσσερις αλλά όχι στις δέκα πρώτες, άρα είναι η τιμή που αντιστοιχεί στη τιμή με σχετική συχνότητα 14 άρα είναι 12.
- Η επικρατούσα τιμή προφανώς είναι 11 (Η τιμή της μεταβλητής με τη μεγαλύτερη συχνότητα).
- Η διακύμανση s^2 ισούται με το κλάσμα που έχει αριθμητή το άθροισμα των στοιχείων της στήλης $v_i(\overline{x} - x_i)^2$ και παρανομαστή το πλήθος των στοιχείων, άρα $s^2 = \dfrac{218{,}64}{25} = 8{,}7456$, οπότε τυπική απόκλιση $s = \sqrt{8{,}7456} = 2{,}9572$.

Εφαρμογή 4

Ο επόμενος πίνακας παρουσιάζει την ηλικία 50 εργαζομένων ενός εργοστασίου.

[18,23)	[23,28)	[28,33)	[33,38)	[38,43)	[43,48)	[48,53)	[53,58)
2	5	6	11	9	7	6	4

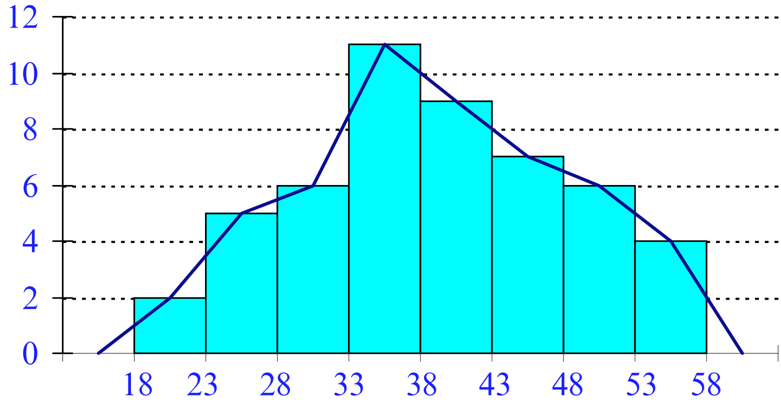
(α) Να γίνει ιστόγραμμα συχνοτήτων.
(β) Να γίνει πολύγωνο αθροιστικών συχνοτήτων.
(γ) Να βρεθεί η διάμεσος.

Λύση

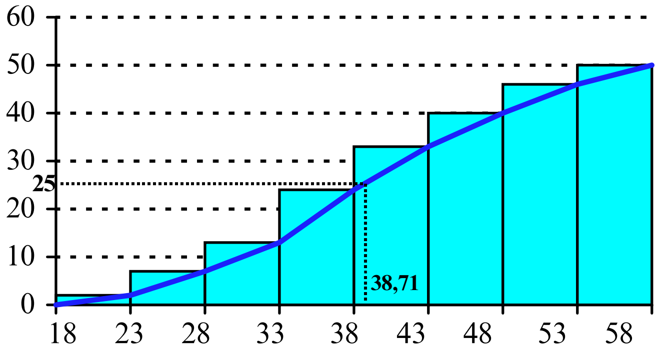
Κατασκευάζουμε τον πίνακα συχνοτήτων και αθροιστικών συχνοτήτων.

x	v_i	**Αθροιστική συχνότητα**
[18,23)	2	2
[23,28)	5	7
[28,33)	6	13
[33,38)	11	24
[38,43)	9	33
[43,48)	7	40
[48,53)	6	46
[53,58)	4	50
'Αθροισμα	50	

(α) Το ιστόγραμμα συχνοτήτων και το πολύγωνο συχνοτήτων είναι:



(β) Το ιστόγραμμα – πολύγωνο αθροιστικών συχνοτήτων είναι:



(γ) Η διάμεσος αντιστοιχεί στο 25 του $y'y$. Από το παραπάνω γράφημα παρατηρούμε ότι αυτή είναι 38,71.

Εφαρμογή 5

Το μέσο βάρος 30 μαθητών και μαθητριών ενός Γυμνασίου είναι 62 Kg. Ποιό θα είναι το μέσο βάρος του Γυμνασίου:

(α) Αν έρθει μια μαθήτρια με βάρος 64 Kg.

(β) Αν φύγει ένας μαθητής με βάρος 66 Kg.

Λύση

Είναι $\overline{x} = 62$ Kg, αλλά $\overline{x} = \dfrac{t_1 + t_2 + \cdots + t_{30}}{30}$, οπότε:

$$62 \cdot 30 = t_1 + t_2 + \cdots + t_{30}$$

(α) Η νέα μέση τιμή θα είναι:

$$\overline{y} = \frac{t_1 + t_2 + \cdots + t_{30} + 64}{31} = \frac{62 \cdot 30 + 64}{31} = 62{,}06 \text{ Kg}$$

(β) Η νέα μέση τιμή θα είναι:

$$\overline{y} = \frac{t_1 + t_2 + \cdots + t_{30} - 66}{29} = \frac{62 \cdot 30 - 66}{29} = 61{,}86 \text{ Kg}$$

Εφαρμογή 6

Δίνεται ο πίνακας:

x_i	ν_i
4	9
5	;
6	3
7	2
8	4

με τις τιμές μιας μεταβλητής και τις αντίστοιχες συχνότητες. Αν η συχνότητα του 5 έχει "χαθεί", να βρεθεί αυτή στις παρακάτω περιπτώσεις:

(α) Υπάρχουν δύο επικρατέστερες τιμές.

(β) Η μέση τιμή είναι 5,5.

Λύση

(α) Οι γνωστές συχνότητες είναι μικρότερες ή ίσες από τη συχνότητα του 4 που είναι 9. Για να υπάρχουν δύο επικρατέστερες τιμές πρέπει η συχνότητα του 5 να είναι ίση με τη συχνότητα του 4 δηλαδή να είναι 9.

(β) Αν η συχνότητα του 5 είναι κ τότε έχουμε:

x_i	ν_i	$\nu_i x_i$
4	9	36
5	κ	5κ
6	3	18
7	2	14
8	4	32
Αθροίσματα	18 + κ	100 + 5κ

'Αρα η μέση τιμή είναι $\overline{x} = \dfrac{100+5\kappa}{18+\kappa}$, επομένως:

$$5{,}5 = \frac{100+5\kappa}{18+\kappa} \quad \text{ή} \quad \frac{55}{10} = \frac{100+5\kappa}{18+\kappa} \quad \text{ή} \quad 55(18+\kappa) = 10(100+5\kappa) \quad \text{ή}$$

$$990 + 55\kappa = 1000 + 50\kappa \quad \text{ή} \quad 55\kappa - 50\kappa = 1000 - 990 \quad \text{ή}$$

$$5\kappa = 10 \quad \text{ή} \quad \kappa = 2$$

Εφαρμογή 7

Η μέση τιμή 40 παρατηρήσεων είναι 20. Αν από αυτές οι 7 μειώνονται κατά 2 και οι 9 αυξάνονται κατά 6, να βρεθεί η νέα μέση τιμή.

Λύση

Αν $\overline{x} = 20$ τότε $20 = \dfrac{t_1 + t_2 + \cdots + t_{40}}{40}$, οπότε $t_1 + t_2 + \cdots + t_{40} = 800$.

Για τη νέα μέση τιμή θα έχουμε $\overline{y} = \dfrac{t_1' + t_2' + \cdots + t_{40}'}{40}$, αλλά:

$$t_1' + t_2' + \cdots + t_{40}' = t_1 + t_2 + \cdots + t_{40} - 7\cdot 2 + 9\cdot 6 = 800 - 14 + 54 = 840$$

'Αρα $\overline{y} = \dfrac{840}{40} = 21$.

Εφαρμογή 8

Ελέγχονται 35 αυτοκίνητα ως προς το χρώμα τους και μας παραδίδεται ο παρακάτω πίνακας:

Χρώμα	v_i	$f_i\%$
Κόκκινο	3	
Κίτρινο	9	
Πράσινο	14	
Μπλε		20
Μαύρο		
Σύνολα	35	

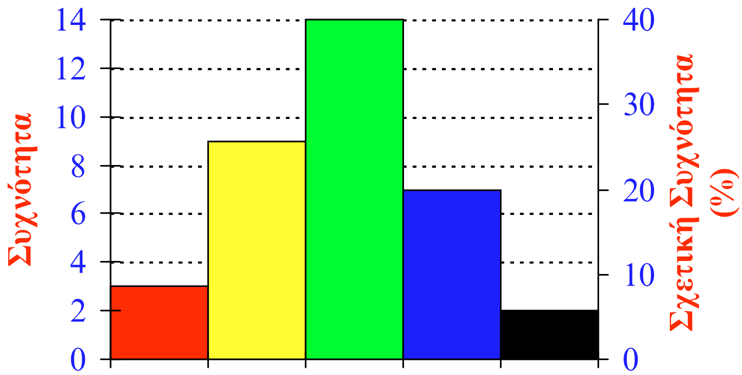
Να υπολογιστούν όλες οι συχνότητες και οι σχετικές συχνότητες και να γίνει το ραβδόγραμμα των συχνοτήτων – σχετικών συχνοτήτων.

Λύση

Το 20% του 35 είναι 35·0,2 = 7, δηλαδή η συχνότητα του μπλε χρώματος είναι 7 και έχουμε ότι η συχνότητα του μαύρου χρώματος θα είναι 35 – (3 + 9 + 14 + 7) = 35 – 33 = 2. 'Αρα ο πίνακας γίνεται:

Χρώμα	ν_i	$f_i\%$
Κόκκινο	3	8,57
Κίτρινο	9	25,71
Πράσινο	14	40
Μπλε	7	20
Μαύρο	2	5,72
Σύνολα	35	100

και το ραβδόγραμμα συχνοτήτων – σχετικών συχνοτήτων είναι:



Ασκήσεις Εμπέδωσης

'Ασκηση 1

Οι εισαγωγές – εξαγωγές της χώρας από το 1974 έως το 1993 δίνονται από τον παρακάτω πίνακα. Να γίνει η γραφική παράσταση αυτών με πολυγωνικές γραμμές στην ίδια γραφική παράσταση.

Εισαγωγές και Εξαγωγές 1974 – 1993

'Ετος	Εισαγωγές	Εξαγωγές	'Ετος	Εισαγωγές	Εξαγωγές
1974	132181	60980	1984	1083880	542676
1975	172041	74441	1985	1412797	629085
1976	223159	93811	1986	1582298	789995
1977	252150	101330	1987	1867354	955070
1978	287729	123727	1988	1756998	776434
1979	356820	144238	1989	2625714	1230942
1980	452881	221108	1990	3137524	1267507
1981	493764	237928	1991	3921522	1579967
1982	665920	286281	1992	4484059	1880763
1983	848295	392652	1993	5050531	1933422

Πηγή: Εθνική Στατιστική Υπηρεσία της Ελλάδος

'Ασκηση 2

Να αποδείξετε ότι η μέση τιμή είναι το άθροισμα των γινομένων των τιμών της μεταβλητής επί τις αντίστοιχες σχετικές συχνότητες.

'Ασκηση 3

Το μέσο ύψος των 80 κατοίκων ενός νησιού είναι 172 cm. Στις καλοκαιρινές διακοπές φθάνουν στο νησί 60 τουρίστες με μέσο ύψος 187 cm. Ποιό είναι το μέσο ύψος των ατόμων που βρίσκονται στο νησί στις καλοκαιρινές διακοπές;

'Ασκηση 4

Σε ένα εργοστάσιο με 100 εργαζόμενους η μέση τιμή των αμοιβών είναι 185000. Το 45% των εργαζομένων στο εργοστάσιο πληρώνονται με μισθό μικρότερο της μέσης τιμής και οι μισθοί τους έχουν μέση τιμή 165000. Αν οι αποδοχές των εργαζομένων με μισθό μικρότερο της μέσης τιμής αυξηθούν στη μέση τιμή, ποιά θα είναι η νέα μέση τιμή;

'Ασκηση 5

Ο παρακάτω πίνακας μας δίνει τη σχετική συχνότητα μιας μεταβλητής με μέση τιμή 2,1. Να συμπληρωθεί ο πίνακας και να βρεθεί η διάμεσος.

X_i	1	2	3	4
$f_i\%$	26		23	

Σύνθετες Ασκήσεις

'Ασκηση 1

Μια μεταβλητή παίρνει τις τιμές x, x + 2ω, x + 3ω, x + 4ω, x + 5ω με ω > 0. Να βρεθεί η τυπική απόκλιση.

'Ασκηση 2

Μετράμε το βάρος των κατοίκων ενός χωριού και έχουμε μέση τιμή 72Kg και τυπική απόκλιση 5Kg. Η μέτρηση του βάρους έγινε με μια ζυγαριά που όπως αποδείχθηκε μετά τις μετρήσεις έδειχνε 2Kg λιγότερο σε κάθε μέτρηση. Ποιά είναι η πραγματική μέση τιμή και ποιά η πραγματική τυπική απόκλιση;

'Ασκηση 3

Οι απουσίες των μαθητών μιας τάξης Λυκείου κατά τον μήνα Σεπτέμβριο είναι 0, 1, 2, 2, 0, 3, 2, 7, 2, 1, 0, 5, 6, 7, 4, 5, 2, 1, 0, 6, 7, 8, 9, 0, 0, 2, 1, 2, 1, 0, 9. Να βρεθεί το ποσοστό των μαθητών που έχουν περισσότερες από 6 απουσίες.

'Ασκηση 4

Αν οι απουσίες δύο Λυκείων Α, Β δίνονται από τον παρακάτω πίνακα:

(α) Να βρεθεί η μέση τιμή και η τυπική απόκλιση των δύο κατανομών

(β) Να σχεδιαστεί το πολύγωνο συχνοτήτων σε κοινό διάγραμμα.

(γ) Από τα Λύκεια Α, Β ποιό είναι πιο ομοιογενές ως προς τις απουσίες;

Απουσίες	Α′ Λύκειο	Β′ Λύκειο
[0, 10)	20	25
[10, 20)	16	11
[20, 30)	25	17
[30, 40)	26	22
[40, 50)	7	8
[50, 60)	8	20
[60, 80)	9	13
[80, 100)	12	4
[100, 125)	3	4

'Ασκηση 5

Οι 8 από τους 20 εργαζόμενους σε ένα εργοστάσιο έχουν μέσο μηνιαίο μισθό 175000 και οι υπόλοιποι 215000. Ποιός είναι ο μέσος μηνιαίος μισθός όλων των εργαζόμενων;

'Ασκηση 6

Η ευστοχία μιας ομάδας Μπάσκετ στις ελεύθερες βολές είναι 70%. Οι 5 βασικοί παίκτες έχουν ευστοχία 75% και οι 3 επιθετικοί αναπληρωματικοί 60%. Πόση ευστοχία έχουν οι 2 αμυντικοί αναπληρωματικοί;

Πρακτικές Εφαρμογές

'Ασκηση 1

Να γίνει ο πίνακας συχνοτήτων – σχετικών συχνοτήτων – αθροιστικών – σχετικών αθροιστικών συχνοτήτων της παρακάτω κατανομής:

ΠΙΝΑΚΑΣ ΗΛΙΚΙΑΣ ΠΑΘΟΝΤΩΝ (οδηγοί – επιβάτες – πεζοί)
κατά το 1998 έως 50 ετών

Ηλικία (έτη)	Μέχρι 5	6-12	13-17	18-20	21-24	25-29	30-34	35-44	45-50
'Ατομα	354	704	2.124	3.739	5.043	4.783	3.824	4.654	2.259

Πηγή: Υπουργείο Δημόσιας Τάξης

Κατόπιν να κατασκευαστεί το διάγραμμα συχνοτήτων.

'Ασκηση 2

Ο παρακάτω πίνακας μας παρουσιάζει την κατανομή των μηχανοκίνητων αλιευτικών σκαφών ως προς την ιπποδύναμη τους:

Μηχανοκίνητα αλιευτικά σκάφη, κατά κλιμάκια ιπποδυνάμεως μέχρι 200 HP κατά το έτος 1991

Κλιμάκια ιπποδυνάμεως	ν_i
[0, 10)	10
[10, 20)	52
[20, 30)	1777
[30, 40)	1317
[40, 50)	856
[50, 60)	863
[60, 70)	434
[70, 80)	517
[80, 90)	486
[90, 100)	452
[100, 150)	1243
[150, 200)	410

Πηγή: Εθνική Στατιστική Υπηρεσία της Ελλάδος

(α) Να γίνει το ιστόγραμμα συχνοτήτων.
(β) Να βρεθεί η μέση τιμή.
(γ) Να βρεθεί η διακύμανση και τυπική απόκλιση.

'Ασκηση 3

Ο παρακάτω πίνακας μας δίνει την κατανομή των κατοίκων της Ελλάδας ως προς την ηλικία κατά το έτος 1992:

Κατανομή ηλικίας κατοίκων Ελλάδας κατά το 1992

Ηλικία	'Ατομα	Ηλικία	'Ατομα	Ηλικία	'Ατομα
[0, 5)	529501	[30, 35)	731310	[60, 65)	634154
[5, 10)	620540	[35, 40)	698124	[65, 70)	511206
[10, 15)	741210	[40, 45)	666503	[70, 75)	358530
[15, 20)	768765	[45, 50)	624437	[75, 80)	285435
[20, 25)	788764	[50, 55)	609587	[80, 85)	203942
[25, 30)	759301	[55, 60)	659035	85 και άνω	131539

Πηγή: Εθνική Στατιστική Υπηρεσία της Ελλάδος (α) Να γίνει γραφική παράσταση των συχνοτήτων.

(β) Να γίνει γραφική παράσταση των σχετικών συχνοτήτων

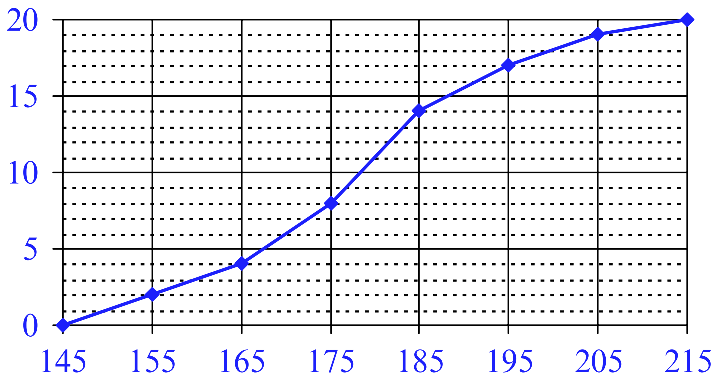
(γ) Να βρεθεί τι ποσοστό του πληθυσμού είχε ηλικία μικρότερη ή ίση των 23 ετών.

(δ) Να βρεθεί τι ποσοστό του πληθυσμού είχε ηλικία μεγαλύτερη ή ίση των 50 ετών.

Γενικές Ασκήσεις Κεφαλαίου

'Ασκηση 1

Σε σχολείο πραγματοποιήθηκε έρευνα για το ύψος των μαθητών. Κατόπιν όμως χάθηκαν τα στοιχεία μετρήσεων και βρέθηκε μόνο το πολύγωνο αθροιστικών συχνοτήτων.



(α) Να κατασκευάσετε το πολύγωνο συχνοτήτων.

(β) Να βρείτε τη μέση τιμή.

(γ) Να βρείτε τη διάμεσο.

(δ) Να βρείτε τη διακύμανση και την τυπική απόκλιση.

'Ασκηση 2

Αν οι τιμές μιας μεταβλητής αυξηθούν κατά L να δειχθεί ότι η μέση τιμή αυξάνεται κατά L ενώ η τυπική απόκλιση παραμένει σταθερή.

'Ασκηση 3

(α) Ποιά είναι η μέση τιμή των γωνιών ενός τριγώνου;

(β) Ποιά είναι η μέση τιμή των γωνιών ενός ν-γωνου;

'Ασκηση 4

Αν οι τιμές των παρατηρήσεων μεταβλητής x πολλαπλασιασθούν με έναν αριθμό $\alpha > 0$, πώς μεταβάλλεται η μέση τιμή και η τυπική απόκλιση;

'Ασκηση 5

Οι οκτώ εργαζόμενοι μιας επιχείρησης έχουν μέσο μισθό 125000 και ο μέσος μισθός των υπολοίπων είναι 145000. Αν ο μέσος μισθός όλων των εργαζομένων στην επιχείρηση είναι 137000, να βρεθεί πόσοι εργάζονται στην επιχείρηση.

Ανακεφαλαίωση

- Στατιστική είναι ο κλάδος των μαθηματικών ο οποίος ως έργο έχει την *συγκέντρωση* στοιχείων, την *ταξινόμησή* τους και την παρουσίασή τους σε κατάλληλη μορφή ώστε να μπορούν να αναλυθούν και να ερμηνευθούν για την εξυπηρέτηση διαφόρων σκοπών.
- Οι μεταβλητές διαχωρίζονται
  - Ποιοτικές μεταβλητές
  - Ποσοτικές μεταβλητές και αυτές σε:
    - Διακριτές
    - Συνεχείς
- Το δείγμα του πληθυσμού πρέπει να είναι μια ακριβής μικρογραφία του και να αποδίδει σωστά την εικόνα του πληθυσμού.
- Τον αριθμό των εμφανίσεων μιας τιμής στο σύνολο τιμών μιας μεταβλητής την ονομάζουμε συχνότητα εμφάνισης της τιμής ή απλά συχνότητα της τιμής.
- Σχετική συχνότητα είναι ο λόγος της συχνότητας προς το μέγεθος του δείγματος.
- Σε ποσοτική μεταβλητή, αθροιστική συχνότητα μιας τιμής x_i λέγεται το άθροισμα των συχνοτήτων v_i των τιμών που είναι μικρότερες ή ίσες με την τιμή αυτή.
- 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.
- Επικρατούσα τιμή μιας μεταβλητής ονομάζεται η τιμή με τη μεγαλύτερη συχνότητα.
- Η μέση τιμή διαφόρων τιμών είναι το πηλίκο του αθροίσματος των τιμών προς το πλήθος τους.
- Διάμεσος ενός δείγματος ν παρατηρήσεων που έχουν διαταχθεί σε αύξουσα σειρά ονομάζεται:
  - Η *μεσαία παρατήρηση* αν το πλήθος των παρατηρήσεων είναι περιττό.
  - Το *ημιάθροισμα* των μεσαίων παρατηρήσεων αν το πλήθος των παρατηρήσεων είναι άρτιο.

- Εύρος τιμών μεταβλητής είναι η διάφορα της μικρότερης τιμής από τη μεγαλύτερη.
- Αν μια μεταβλητή παίρνει τις ν τιμές $t_1, t_2, \ldots, t_\nu$ που έχουν μέση τιμή $\overline{x}$ τότε μέση (απόλυτη) απόκλιση της μεταβλητής ονομάζεται το πηλίκο:

$$\frac{|\overline{x}-t_1| + |\overline{x}-t_2| + \cdots + |\overline{x}-t_\nu|}{\nu}$$

- Αν μια μεταβλητή παίρνει τις ν τιμές $t_1, t_2, \ldots, t_\nu$ που έχουν μέση τιμή $\overline{x}$ τότε διακύμανση της μεταβλητής ονομάζεται το πηλίκο:

$$s^2 = \frac{(\overline{x}-t_1)^2 + (\overline{x}-t_2)^2 + \cdots + (\overline{x}-t_\nu)^2}{\nu}$$

- Αν μια μεταβλητή παίρνει τις ν τιμές $t_1, t_2, \ldots, t_\nu$ που έχουν μέση τιμή $\overline{x}$ τότε τυπική απόκλιση της μεταβλητής ονομάζεται το:

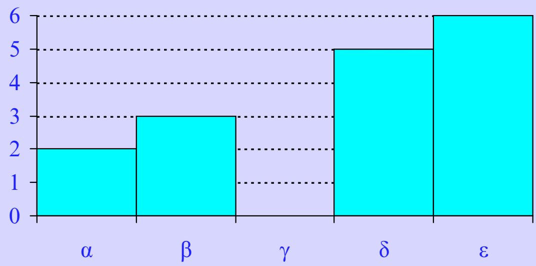
$$s = \sqrt{\frac{(\overline{x}-t_1)^2 + (\overline{x}-t_2)^2 + \cdots + (\overline{x}-t_\nu)^2}{\nu}}$$

- Συντελεστής μεταβλητότητας ονομάζεται το πηλίκο:

$$CV = \frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}} = \frac{s}{\overline{x}}\, 100\%$$

Ερωτήσεις Κατανόησης

1. Τα νεογέννητα σκυλιά ανοίγουν τα μάτια τους κατά μέσο όρο τη $15^{η}$ μέρα. 'Ενα νεογέννητο δεν ανοίγει τα μάτια του τη $15^{η}$ μέρα. Ο ιδιοκτήτης του πρέπει να ανησυχεί;
2. Αν σε μια κατανομή συχνοτήτων απαλείψουμε τη μέγιστη και την ελάχιστη τιμή, η διάμεσος μεταβάλλεται;
3. Σε δείγμα 5 ατόμων για τον έλεγχο του αριθμού των αυτοκινήτων που κατέχουν παρουσιάζεται το παρακάτω διάγραμμα. Αν ο συνολικός αριθμός αυτοκινήτων είναι 20, μπορείτε να το συμπληρώσετε;



4. Είναι ομογενές το δείγμα 50 ατόμων που εξετάζονται ως προς το ύψος τους αν έχουν μέση τιμή ύψους 1,80 m και διακύμανση 36 cm;
5. Η διακύμανση εκφράζεται με τις μονάδες μέτρησης της μεταβλητής;
6. Ο λόγος της διακύμανσης προς την τυπική απόκλιση είναι μία παράμετρος διασποράς;
7. Αν από τις τιμές μιας μεταβλητής εξαιρέσουμε μια τιμή της που είναι "κοντά" στη μέση τιμή, θα έχουμε μικρή μεταβολή της μέσης τιμής;
8. Η σχετική συχνότητα μας δίνει το ποσοστό επί τις χιλίοις ‰ των ατόμων του δείγματος που έχουν την αντίστοιχη τιμή της μεταβλητής;
9. Είναι ποιοτική μεταβλητή του πληθυσμού των γυναικών το χρώμα των ματιών τους;
10. Η σύμπτωση διαμέσου και μέσης τιμής μας δίνει την πληροφορία ότι είναι ομογενές το δείγμα ως προς το χαρακτηριστικό μέγεθος που εξετάζεται;