

ΣΤΑΤΙΣΤΙΚΗ

ΕΠΙΧΕΙΡΗΣΕΩΝ

Β' ΕΠΑ.Λ.



Ασκήσεις - Λύσεις

Δραστηριότητες

ΤΟΜΕΑΣ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

ΣΤΑΤΙΣΤΙΚΗ ΕΠΙΧΕΙΡΗΣΕΩΝ

ΑΣΚΗΣΕΙΣ - ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ

ΛΥΣΕΙΣ

ΣΤΟΙΧΕΙΑ ΑΡΧΙΚΗΣ ΕΚΔΟΣΗΣ

Ομάδα συγγραφής:

Καραγεώργος Δημήτριος, *Λέκτορας Πανεπιστημίου Αθηνών, τ. Σύμβουλος Π.Ι.*
Κόκλα Άννα-Μαρία, *Καθηγήτρια Β/θμιας Εκπαίδευσης*
Παπακωνσταντίνου Ευάγγελος, *Σχολικός Σύμβουλος Β/θμιας Εκπαίδευσης*

Κριτές:

Καραγεώργος Δημήτριος, *Λέκτορας Πανεπιστημίου Αθηνών, τ. Σύμβουλος Π.Ι.*
Κόκλα Άννα-Μαρία, *Καθηγήτρια Β/θμιας Εκπαίδευσης*
Παπακωνσταντίνου Ευάγγελος, *Σχολικός Σύμβουλος Β/θμιας Εκπαίδευσης*

Γλωσσική επιμέλεια:

Βαρδαλάχου Αικατερίνη, *Φιλολόγος*

Συντονιστής:

Καραγεώργος Δημήτριος, *Λέκτορας Πανεπιστημίου Αθηνών, τ. Σύμβουλος Π.Ι.*

ΕΠΙΜΕΛΕΙΑ ΕΚΔΟΣΗΣ: INK DESIGN

ΣΤΟΙΧΕΙΑ ΕΠΑΝΕΚΔΟΣΗΣ

Η επανέκδοση του παρόντος βιβλίου πραγματοποιήθηκε από το Ινστιτούτο Τεχνολογίας Υπολογιστών & Εκδόσεων «Διόφαντος» μέσω ψηφιακής μακέτας.

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΘΡΗΣΚΕΥΜΑΤΩΝ ΚΑΙ ΑΘΛΗΤΙΣΜΟΥ
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

Καραγεώργος Δημήτριος, *Λέκτορας Πανεπιστημίου Αθηνών, τ. Σύμβουλος Π.Ι.*

Κόκλα Άννα-Μαρία, *Καθηγήτρια Β/θμιας Εκπαίδευσης*

Παπακωνσταντίνου Ευάγγελος, *Σχολικός Σύμβουλος Β/θμιας Εκπαίδευσης*

Η συγγραφή και η επιστημονική επιμέλεια του βιβλίου πραγματοποιήθηκε
υπό την αιγίδα του Παιδαγωγικού Ινστιτούτου

ΣΤΑΤΙΣΤΙΚΗ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΑΣΚΗΣΕΙΣ - ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ
ΛΥΣΕΙΣ

Β΄ ΕΠΑ.Λ.

ΤΟΜΕΑΣ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

για το μαθητή

Το τεύχος που κρατάς έχει μια ιδιομορφία: σου δίνεται με τη σύσταση να μην το διαβάσεις τουλάχιστον με την έννοια που διαβάζεις ένα άλλο βιβλίο για να κατανοήσεις το περιεχόμενό του. Πράγματι, οι ασκήσεις που σου δίνει ο καθηγητής σου είναι για να εργαστείς μόνος. Γιατί το να λύσεις μια άσκηση σημαίνει πολλές φορές όχι μόνο ότι έχεις κατανοήσει την αντίστοιχη θεωρητική ύλη αλλά και ότι ξέρεις να τη χρησιμοποιήσεις για να δημιουργείς, να ανακαλύπτεις ή να επιβεβαιώνεις κάτι καινούργιο. Και αυτό έχει ιδιαίτερη σημασία για σένα τον ίδιο. Δεν μπορεί παρά να έχεις και συ τη φιλοδοξία να λύνεις μόνος, χωρίς βοήθεια, τις ασκήσεις για να νιώθεις τη χαρά αυτής της δημιουργίας, της ανακάλυψης.

Πρέπει να ξέρεις ότι, όταν δυσκολεύεσαι στη λύση μιας άσκησης, τις πιο πολλές φορές υπάρχει κάποιο κενό στη γνώση της αντίστοιχης θεωρίας. Πήγαινε πίσω λοιπόν στο διδακτικό βιβλίο κάθε φορά που χρειάζεται να εντοπίσεις και να συμπληρώσεις τέτοια κενά. Οπωσδήποτε πριν καταπιαστείς με τη λύση των ασκήσεων πρέπει να αισθάνεσαι κάτοχος της θεωρίας που διδάχτηκες.

Εκτός από την κατανόηση της θεωρίας μπορεί να βοηθηθείς στη λύση μιας άσκησης από τα παραδείγματα και τις εφαρμογές που περιέχει το διδακτικό σου βιβλίο. Αν παρόλ' αυτά δεν μπορείς να προχωρήσεις, στο τέλος του βιβλίου σου θα βρεις μια σύντομη υπόδειξη που ασφαλώς θα σε διευκολύνει.

Στις ελάχιστες περιπτώσεις που έχοντας εξαντλήσει κάθε περιθώριο προσπάθειας δε βρίσκεται η πορεία που οδηγεί στη λύση της άσκησης, τότε και μόνο τότε μπορείς να καταφύγεις σ' αυτό το τεύχος και μάλιστα για να διαβάσεις εκείνο το τμήμα της λύσης που σου είναι απαραίτητο για να συνεχίσεις μόνος.

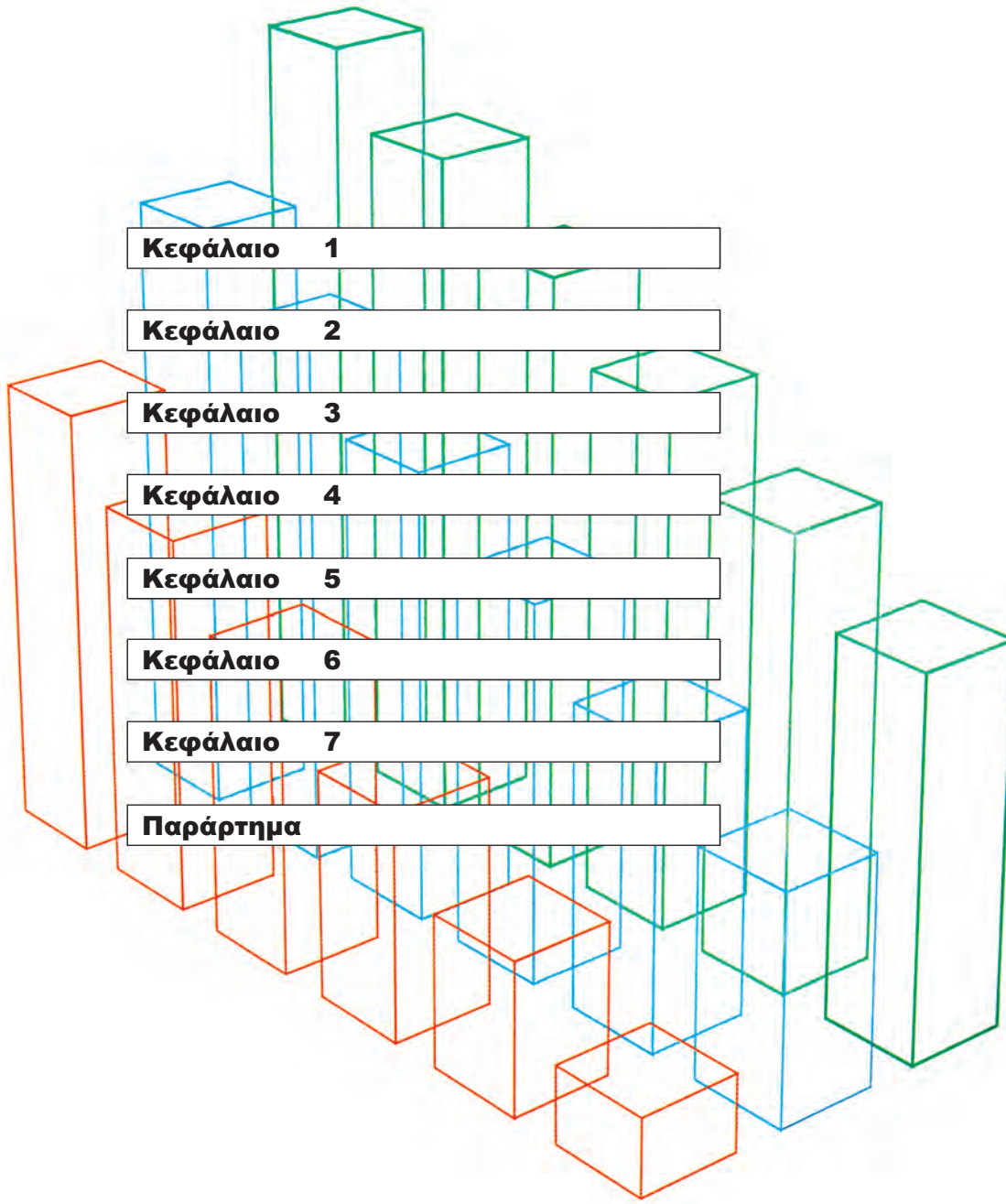
Ουσιαστικά λοιπόν δεν το 'χεις ανάγκη αυτό το τεύχος. Σου παρέχεται όμως για τους εξής λόγους:

- α** Για να μπορείς να συγκρίνεις τις λύσεις που εσύ βρήκες.
- β** Για να σε προφυλάξει από ανεύθυνα "λυσάρια".
- γ** Για να απαλλάξει τους γονείς σου από αντίστοιχη οικονομική επιβάρυνση.
- δ** Για να έχεις εσύ και οι συμμαθητές σου την ίδια συλλογή ασκήσεων που είναι έτσι επιλεγμένες ώστε να εξασφαλίζουν την εμπέδωση της ύλης.
- ε** Για να εργάζεσαι χωρίς το άγχος να εξασφαλίσεις οπωσδήποτε για κάθε μάθημα τις λύσεις των ασκήσεων.
- στ** Για να δεις τις δραστηριότητες που πρέπει να αναπτύξεις.

Το τεύχος που κρατάς είναι λοιπόν φίλος. Να του συμπεριφέρεσαι όπως σ' έναν φίλο που έχει δει πριν από σένα την ταινία που πρόκειται να δεις, μην του επιτρέψεις να σου αποκαλύψει την "υπόθεση" πριν δεις και συ το έργο. Μετά μπορείτε να συζητήσετε. Η σύγκριση των συμπερασμάτων θα είναι ενδιαφέρουσα και προπαντός επωφελής.

Οι συγγραφείς.

Π Ε Ρ Ι Ε Χ Ο Μ Ε Ν Α



Κεφάλαιο 1

Κεφάλαιο 2

Κεφάλαιο 3

Κεφάλαιο 4

Κεφάλαιο 5

Κεφάλαιο 6

Κεφάλαιο 7

Παράρτημα



ΚΕΦΑΛΑΙΟ 1^ο

Να κατανοήσουν οι μαθητές το αντικείμενο της Στατιστικής και τη χρησιμότητά της στη ζωή μας. Να καταλάβουν ότι με την εφαρμογή των στατιστικών μεθόδων σε πραγματικά προβλήματα, με τη σχεδίαση και τη μελέτη ενός φαινομένου, με τη συγκέντρωση, την ταξινόμηση και την επεξεργασία των πληροφοριών και την παρουσίασή τους (με κατανοητό τρόπο) μπορούν να σχεδιάσουν και να μελετήσουν ένα φαινόμενο σωστά και αξιόπιστα.

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1, 2, 3, 4, 7, 8, 9, 12 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του κεφαλαίου.

Οι μαθητές θα πρέπει να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στην τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

- 5** Πιθανώς όχι, γιατί οι περιοχές της ίδιας πόλης, αν αυτή είναι μεγάλη, είναι διαφορετικές μεταξύ τους. Κατά συνέπεια οι απόψεις των κατοίκων μιας περιοχής δεν αντιπροσωπεύουν και τις απόψεις των κατοίκων των άλλων περιοχών.

Ένα “δείγμα” από μια περιοχή δεν αντιπροσωπεύει ολόκληρη την πόλη, αφού οι κάτοικοι ολόκληρης της πόλης είναι ο “πληθυσμός”.

- 6** Τα συμπεράσματα που βγάζουμε για τις αποδοχές των υπαλλήλων της εταιρείας είναι λανθασμένες. Δεν έγινε τυχαία δειγματοληψία από τους υπαλλήλους της εταιρείας, αλλά επιλέχθηκαν άτομα μόνο από μια συγκεκριμένη ομάδα υπαλλήλων? άρα το δείγμα δεν είναι αντιπροσωπευτικό για τον πληθυσμό.

- 10** 1A 2B 3B 4A 5Γ 6B 7A

- 11** Η λιανική τιμή ενός προϊόντος είναι συνεχής μεταβλητή, γιατί “εν δυνάμει” μπορεί να πάρει όλες τις τιμές ενός διαστήματος $[\alpha_0, \alpha_k]$. Η μεταβλητή “αριθμός μονάδων του προϊόντος” είναι διακριτή μεταβλητή, γιατί οι τιμές της αποτελούν μεμονωμένα σημεία των πραγματικών αριθμών.



ΚΕΦΑΛΑΙΟ 2°

Να κατανοηθεί η χρησιμότητα της Στατιστικής στη σύγχρονη επιχείρηση. Η γνώση και η εφαρμογή των στατιστικών μεθόδων στις έρευνες αγοράς, στη διαφήμιση και γενικά στη διοίκηση των επιχειρήσεων είναι απαραίτητα εφόδια στους μαθητές που προορίζονται να εργαστούν σε επιχειρήσεις, ώστε να πετύχουν τη σωστή λειτουργία και την ανάπτυξή τους.

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1-10 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του δευτέρου κεφαλαίου. Οι μαθητές θα πρέπει να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στην τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

11 Λ



ΚΕΦΑΛΑΙΟ 3°

Να κατανοήσουν οι μαθητές τον τρόπο συλλογής και παρουσίασης των στατιστικών δεδομένων και κυρίως να καταλάβουν μια από τις απλές διαδικασίες για την οργάνωση και τη συνοπτική παρουσίασή τους, την κατανομή συχνοτήτων.

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1 - 11 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του τρίτου κεφαλαίου. Επαφίεται στους μαθητές να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στην τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

8 Σ

$$F_i \% = 100 \cdot \sum_{i=1}^K f_i$$

ΑΣΚΗΣΕΙΣ

- 1** Δίνεται η ποσοστιαία σύνθεση (%) του προσωπικού μιας επιχείρησης ως προς το μορφωτικό τους επίπεδο.

Επίπεδο Μόρφωσης	Ποσοστό (%)
Πτυχιούχοι Τριτοβάθμιας Εκπαίδευσης	25
Απόφοιτοι Λυκείου	50
Απόφοιτοι Γυμνασίου	25

- i) Αν ο αριθμός των υπαλλήλων της επιχείρησης είναι 40 άτομα, κατασκευάστε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- ii) Να απεικονίσετε τα παραπάνω δεδομένα με ραβδόγραμμα και με κυκλικό διάγραμμα.

Λύση

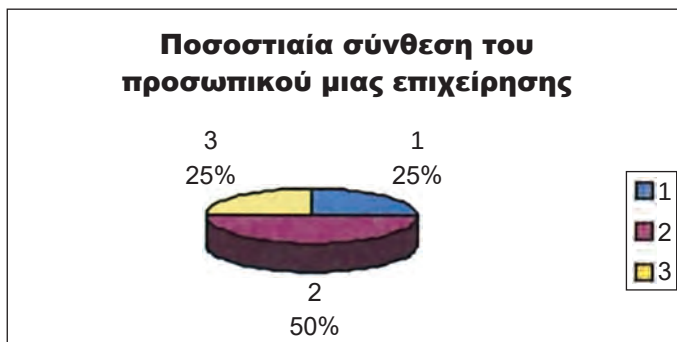
i) Αν ο συνολικός αριθμός των υπαλλήλων μιας επιχείρησης είναι 40 άτομα και το 25% είναι πτυχιούχοι τριτοβάθμιας εκπαίδευσης, τότε ο αριθμός των πτυχιούχων υπαλλήλων θα είναι $0,25 \cdot 40 = 10$.

Αντίστοιχα οι απόφοιτοι λυκείου θα είναι $0,5 \cdot 40 = 20$

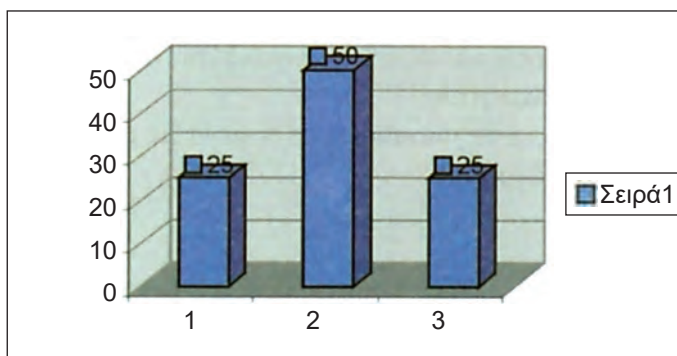
και οι απόφοιτοι γυμνασίου θα είναι $0,25 \cdot 40 = 10$

α/α	Επίπεδο μόρφωσης	Συχνότητες	Σχετικές Συχνότητες
1	Πτυχιούχοι Τριτοβάθμιας Εκπαίδευσης	10	0,25
2	Απόφοιτοι Λυκείου	20	0,50
3	Απόφοιτοι Γυμνασίου	10	0,25
	Σύνολο	40	1,00

ii) Κατασκευάζουμε το παρακάτω κυκλικό διάγραμμα:



Σχ.(1.1): Κυκλικό διάγραμμα



Σχ.(1.2): Ραβδόγραμμα

2 Οι βαθμοί των 32 μαθητών της Γ τάξης ενός Λυκείου είναι: 15, 9, 18, 20, 12, 18, 12, 11, 10, 14, 16, 10, 18, 20, 19, 16, 14, 15, 13, 15, 8, 17, 19, 11, 13, 15, 14, 11, 15, 16, 12, 17.

- i) Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- ii) Να κατασκευάσετε πίνακα αθροιστικών συχνοτήτων και σχετικών αθροιστικών συχνοτήτων.
- iii) Πόσοι μαθητές έχουν βαθμό μικρότερο ή ίσο του 14 και πόσοι μαθητές έχουν βαθμό μεγαλύτερο του 15;

Λύση**i) και ii)**

Σύμφωνα με τα δεδομένα της άσκησης θα κατασκευάσουμε τον πίνακα συχνοτήτων (v_i), σχετικών συχνοτήτων (f_i), αθροιστικών συχνοτήτων (N_i) σχετικών αθροιστικών συχνοτήτων (F_i).

α/α	Βαθμοί	v_i	f_i	N_i	F_i
1	8	1	0,031	1	0,031
2	9	1	0,031	2	0,062
3	10	2	0,063	4	0,125
4	11	3	0,094	7	0,219
5	12	3	0,094	10	0,313
6	13	2	0,063	12	0,376
7	14	3	0,094	15	0,470
8	15	5	0,156	20	0,626
9	16	3	0,094	23	0,720
10	17	2	0,063	25	0,783
11	18	3	0,094	28	0,877
12	19	2	0,063	30	0,940
13	20	2	0,063	32	≈1,00
	Σύνολο	32	≈1,00		

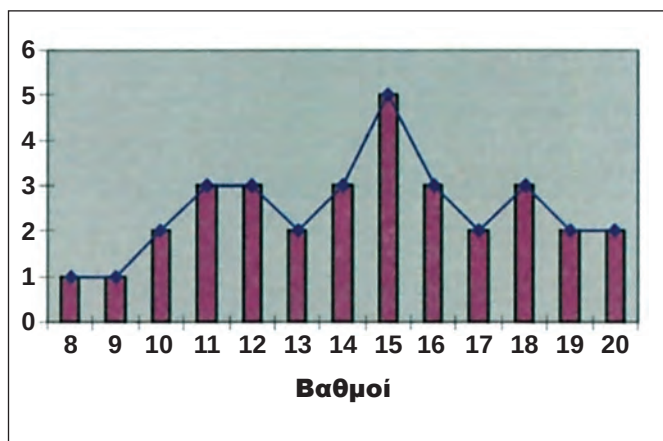
iii) Από την αθροιστική συχνότητα παρατηρούμε ότι 15 μαθητές έχουν βαθμό μικρότερο ή ίσο του 14 και 12 μαθητές βαθμό μεγαλύτερο του 15. Στο δεύτερο συμπέρασμα μπορούμε να καταλήξουμε με δύο τρόπους: $N_{13} - N_8 = 32 - 20 = 12$ ή προσθέτοντας $v_9 + v_{10} + v_{11} + v_{12} + v_{13} = 3 + 2 + 3 + 2 + 2 = 12$

5 Για τα δεδομένα της άσκησης 2 να κατασκευάσετε το διάγραμμα συχνοτήτων, το πολύγωνο συχνοτήτων, καθώς και το διάγραμμα αθροιστικών συχνοτήτων και σχετικών αθροιστικών συχνοτήτων. Από το διάγραμμα ή με κατάλληλους υπολογισμούς να υπολογίσετε:

- i) Το ποσοστό των μαθητών που απέτυχαν (βαθμός < 10).
 - ii) Το ποσοστό των μαθητών που οι βαθμοί τους κυμαίνονται από 13 ως και 16.
 - iii) Το ποσοστό των μαθητών που άριστευσε, δηλαδή που ο βαθμός τους είναι μεγαλύτερος του 18.
- Τι παρατηρείτε για τη βαθμολογία της συγκεκριμένης τάξης;

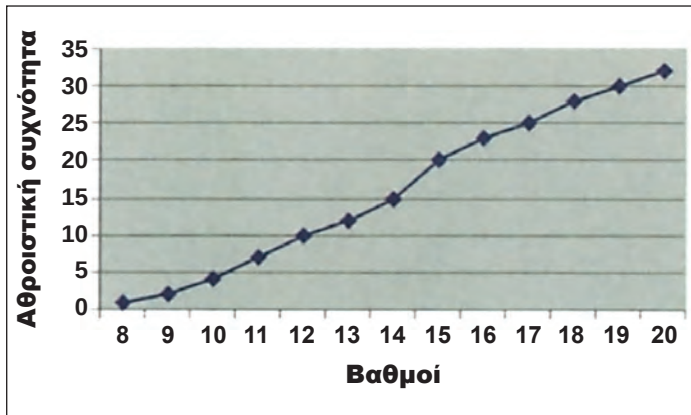
Λύση

Σύμφωνα με τα δεδομένα της άσκησης 2 κατασκευάζουμε το διάγραμμα συχνοτήτων και το πολύγωνο συχνοτήτων

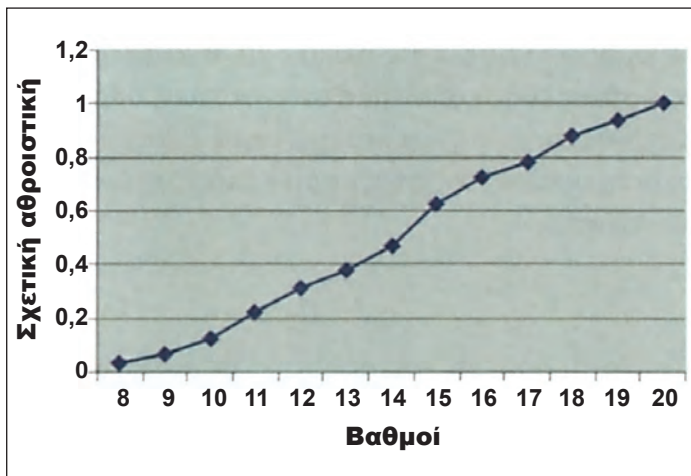


Σχ.(1.3)

Καθώς και το διάγραμμα αθροιστικών συχνοτήτων και σχετικών αθροιστικών συχνοτήτων



Σχ. (1.4)



Σχ. (1.5)

- i) Το ποσοστό των μαθητών που απέτυχαν μπορούμε να το εκτιμήσουμε από το διάγραμμα σχετικών αθροιστικών συχνοτήτων ή να το υπολογίσουμε από τον πίνακα της άσκησης 2.
Αρα το ποσοστό των μαθητών που απέτυχαν (βαθμός < 10) είναι 6,2%
- ii) Το ποσοστό των μαθητών των οποίων οι βαθμοί κυμαίνονται από 13 ως 16 προκύπτει ως εξής: $F_9 - F_5 = 72,0 - 31,3 = 40,7$.
- iii) Το ποσοστό των μαθητών που άριστευσαν ισούται με: $F_{13} - F_{11} = 100,0 - 87,7 = 12,3\%$. Παρατηρούμε ότι η βαθμολογία των περισσότερων μαθητών (40%) κυμαίνεται μεταξύ 13 ως και 16.

4 Οι 50 εργάτες ενός εργοστασίου έχουν τις παρακάτω ηλικίες:

21	43	50	25	55	30	28	40	31	51
18	47	52	34	47	32	27	41	35	54
30	48	36	43	38	33	27	39	41	43
32	22	46	52	29	32	34	34	42	36
35	28	57	56	20	38	35	27	40	35

- Να ομαδοποιήσετε τις ηλικίες σε 8 κλάσεις.
- Να κατασκευάσετε πίνακα συχνοτήτων, αθροιστικών συχνοτήτων, σχετικών συχνοτήτων και σχετικών αθροιστικών συχνοτήτων. Να κατασκευάσετε το ιστόγραμμα συχνοτήτων και το πολύγωνο συχνοτήτων.

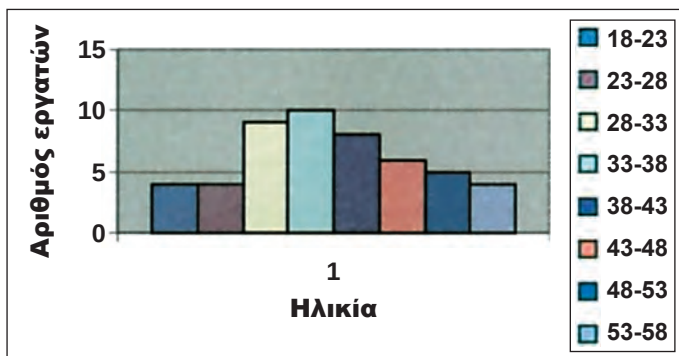
Λύση

i) και ii)

Σύμφωνα με τα δεδομένα της άσκησης θα κατασκευάσουμε τον πίνακα συχνοτήτων και αθροιστικών συχνοτήτων

Ηλικία	Συχνότητα n_i	Σχετική Συχνότητα	Αθροιστική Συχνότητα	Σχετική Αθροιστική Συχνότητα
[18 - 23)	4	8	4	1
[23 - 28)	4	8	8	2
[28 - 33)	9	18	17	4
[33 - 38)	10	20	27	34
[38 - 43)	8	16	35	54
[43 - 48)	6	12	41	70
[48 - 53)	5	10	46	82
[53 - 58)	5	8	50	92
Σύνολο	50	100		100

iii)



Σχ.(1.6)

5 Ο παρακάτω πίνακας δείχνει τον αριθμό των οδικών τροχαίων ατυχημάτων κατά τα έτη 1988 - 1997:

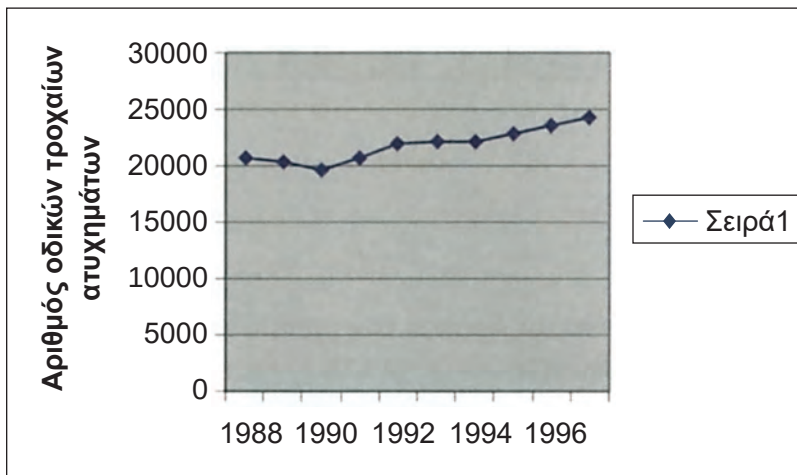
Έτος	Σύνολο Ατυχημάτων
1988	20.753
1989	20.299
1990	19.609
1991	20.764
1992	22.006
1993	22.165
1994	22.222
1995	22.798
1996*	23.606
1997*	24.319

*Προσωρινά στοιχεία Πηγή: Ε.Σ.Υ.Ε.

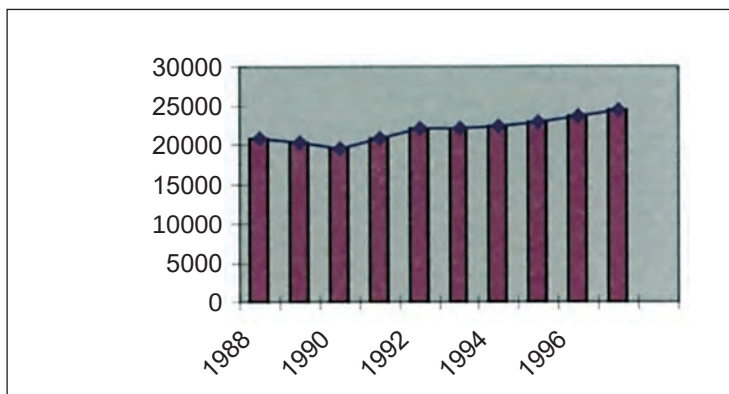
Να παρασταθεί ο αριθμός των οδικών ατυχημάτων κατά τα έτη 1988 - 1997 με χρονολογικό διάγραμμα και να απεικονισθεί η αυξητική εξέλιξη του αριθμού αυτού με χρονολογικό διάγραμμα.

Λύση

Σύμφωνα με τα δεδομένα του παραπάνω πίνακα κατασκευάζουμε τα χρονολογικά διαγράμματα



Σχ.(1.7)



Σχ. (1.8): Αυξητική εξέλιξη του αριθμού των οδικών ατυχημάτων

- 6** Ο παρακάτω πίνακας δείχνει το σύνολο των εξαγωγών μιας επιχείρησης - σε εκατομμύρια δραχμές - σε δύο διαδοχικά χρόνια:

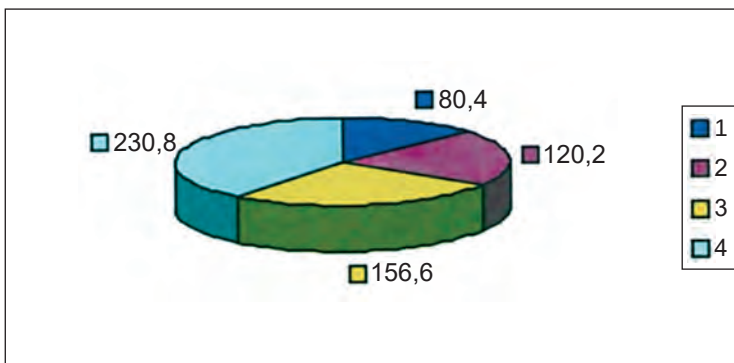
Έτος	Αφρική	Αμερική	Ασία	Ευρώπη
1996	80,4	120,2	150,6	230,8
1997	50,5	60,7	130,2	190,6

- i) Κατασκευάστε δύο κυκλικά διαγράμματα που να απεικονίζουν τις εξαγωγές της επιχείρησης κατά τα έτη 1996 και 1997. Με τη βοήθεια των κυκλικών διαγραμμάτων να συγκρίνετε τις ετήσιες εξαγωγές της επιχείρησης.
- ii) Πόσο τοις εκατό αυξήθηκαν ή μειώθηκαν οι εξαγωγές της επιχείρησης τα έτη 1996-1997 σε κάθε ήπειρο;

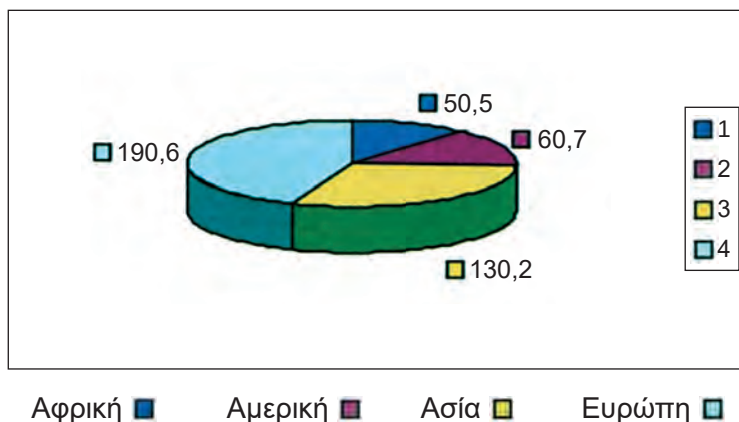
Λύση

- i) Τα παρακάτω κυκλικά διαγράμματα απεικονίζουν το σύνολο των εξαγωγών μιας επιχείρησης σε δύο διαδοχικά χρόνια

Έτος 1996



Έτος 1996



Σχ. (1.9)

Με τη βοήθεια των κυκλικών διαγραμμάτων παρατηρούμε ότι η ποσοστιαία σύνθεση των εξαγωγών παρέμεινε κατά τη διάρκεια των δύο ετών περίπου η ίδια.

ii)

Ποσοστό Μείωσης	Αφρική	Αμερική	Ασία	Ευρώπη
1996 – 1997 (%)	37,2	49,5	13,6	17,4

7 Η στατιστική υπηρεσία της τροχαίας θέλοντας να μελετήσει τους μήνες όπου εμφανίζεται αύξηση των σοβαρών τραυματισμών σε οδικά ατυχήματα, καθώς και τη μεταβολή του αριθμού των σοβαρά τραυματισμένων κατά τα έτη 1996-1997, κατασκεύασε τον παρακάτω πίνακα:

Μήνες	Σοβαρά ατυχήματα	
	1997 v_i	1996 v_i^*
Ιανουάριος	238	190
Φεβρουάριος	201	211
Μάρτιος	278	189
Απρίλιος	326	249
Μάιος	330	294
Ιούνιος	454	319
Ιούλιος	538	344
Αύγουστος	553	433
Σεπτέμβριος	383	317
Οκτώβριος	372	284
Νοέμβριος	319	231
Δεκέμβριος	295	215

*Πηγή: Ε.Σ.Υ.Ε.

- i) Συμπληρώστε τον πίνακα συχνοτήτων με μία στήλη, όπου να φαίνεται η μεταβολή (Δ) του αριθμού των βαριά τραυματιών κατά τα έτη 1996, 1997 (δηλαδή Δ = αριθμός τραυματιών 1997 - αριθμός τραυματιών 1996) στους αντίστοιχους μήνες ενός έτους.
- ii) Υπολογίστε το ποσοστό της μεταβολής αυτής ως προς τον αριθμό των τραυματιών του 1996, δηλαδή $\left(\frac{\Delta}{v_i} \cdot 100 \right)$ με προσέγγιση δέκατου και συμπληρώστε τον πίνακα.
- iii) Κατασκευάστε συγκριτικά μηνιαία ραβδογράμματα για τον αριθμό των σοβαρά τραυματιών σε οδικά ατυχήματα κατά τα έτη 1996 - 1997.
- iv) Κατασκευάστε ένα χρονολογικό διάγραμμα της εξέλιξης της ποσοστιαίας μεταβολής του ερωτήματος (β) για τα έτη 1996, 1997 για τους αντίστοιχους μήνες.

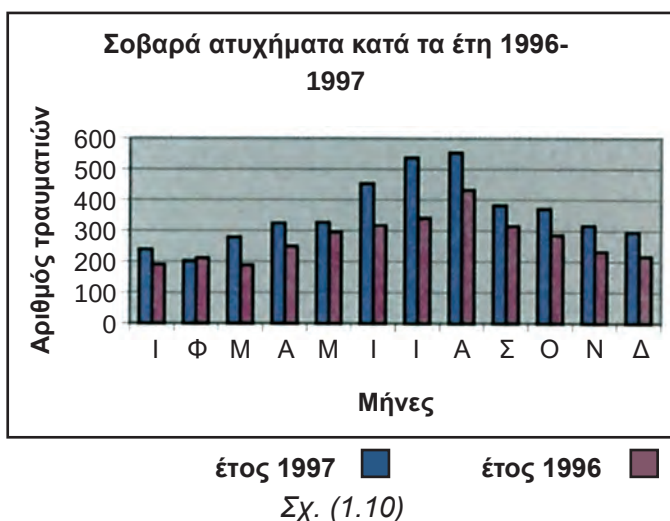
Λύση

i) και ii)

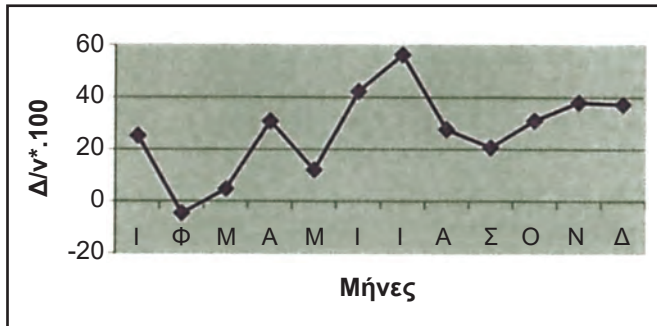
Μήνες	Σοβαρά ατυχήματα			
	1997 v_i	1996 v_i^*	Δ	$\frac{\Delta}{\kappa} \cdot 100$
Ιανουάριος	238	190	48	25,3
Φεβρουάριος	201	211	-10	-4,7
Μάρτιος	278	189	89	4,7
Απρίλιος	326	249	77	30,9
Μάιος	330	294	36	12,2
Ιούνιος	454	319	135	42,3
Ιούλιος	538	344	194	56,4
Αύγουστος	553	433	120	27,7
Σεπτέμβριος	383	317	66	20,8
Οκτώβριος	372	284	88	31
Νοέμβριος	319	231	88	38
Δεκέμβριος	295	215	80	37,2
Σύνολο	4.287	3.276		

iii) και iv)

Κατασκευάζουμε συγκριτικά ραβδογράμματα για τον αριθμό των σοβαρά τραυματιών σε οδικά ατυχήματα τα έτη 1996-1997



Επίσης κατασκευάζουμε ένα χρονολογικό διάγραμμα εξέλιξης της ποσοστιαίας μεταβολής (Δ)-δηλαδή του αριθμού των βαριά τραυματιών κατά τα έτη 1996,1997 ως προς τον αριθμό των τραυματιών του 1996-για τα έτη 1996,1997 στους αντίστοιχους μήνες.



Σχ.(1.11)

- 8 Τα παρακάτω δεδομένα αφορούν την οικογενειακή κατάσταση των υπαλλήλων μιας επιχείρησης.**
Άνδρες ανύπαντροι: 22, Γυναίκες ανύπαντρες: 8, Άνδρες παντρεμένοι: 50, Γυναίκες παντρεμένες: 15, Άνδρες διαζευγμένοι: 3, Γυναίκες διαζευγμένες: 2.

Ζητείται:

- Να παρουσιαστεί η οικογενειακή κατάσταση των υπαλλήλων της επιχείρησης με τη βοήθεια πίνακα διπλής εισόδου.
- Να απεικονισθούν τα παραπάνω δεδομένα της οικογενειακής κατάστασης των υπαλλήλων με ραβδόγραμμα.

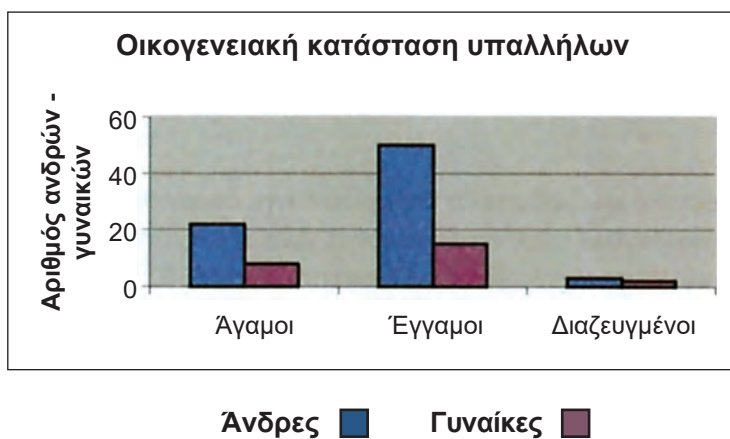
Λύση

i)

Πίνακας διπλής εισόδου

Φύλο	Οικογενειακή Κατάσταση			Σύνολο
	Άγαμοι	Έγγαμοι	Διαζευγμένοι	
Άνδρες	22	50	3	75
Γυναίκες	8	15	2	25
Σύνολο	30	65	5	100

ii)



Σχ.(1.12)

ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ ΚΕΦΑΛΑΙΟΥ 3

ΚΑΤΑΣΚΕΥΗ ΔΙΑΓΡΑΜΜΑΤΩΝ

A Με χρήση υπολογιστή γραφικών (graphics calculator)

Ο σημαντικότερος λόγος για τη χρήση του graphics calculator (υπολογιστής γραφικών τσέπης) είναι το γεγονός ότι γίνεται πιο εύκολη η επεξεργασία μεγάλου αριθμού δεδομένων. Ο πιο συνηθισμένος τρόπος για την παρουσίαση συνεχών δεδομένων είναι η κατασκευή *ιστογράμματος*. Αυτό ακριβώς είναι το αντικείμενο της δραστηριότητας που ακολουθεί:

Οι 30 μαθητές μιας τάξης του Λυκείου αγωνίστηκαν στο άλμα εις μήκος για την επιλογή των 3 πρώτων που θα συμμετέχουν σε διασχολικούς αγώνες. Ο πίνακας που ακολουθεί δείχνει τις επιδόσεις τους (σε μέτρα):

1,8	2,2	2,5	2,7	2,9	3,2	3,3	3,3	3,4	3,8
4,1	4,1	4,3	4,3	4,4	4,5	4,5	4,7	4,7	4,8
4,8	4,9	4,9	4,9	5,0	5,2	5,3	5,4	5,6	6,2

Για την κατασκευή του ιστογράμματος με την βοήθεια του graphics calculator ακολουθούμε τα εξής βήματα:

 Εισαγωγή δεδομένων

AC/ON → MENU

Επιλέγουμε από το MENU το STAT (μαυρισμένο) ακολουθούμενο από το EXE. Τοποθετούμε τις 30 τιμές στον πίνακα (LIST 1), π.χ:

```

1      1.8  EXE
2      2.2  EXE
....
30     6.2  EXE

```

❖ Επεξεργασία δεδομένων

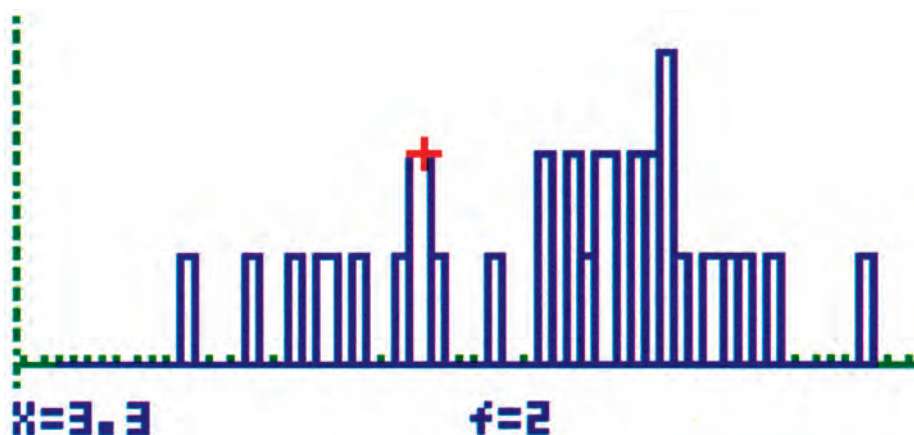
Επιλέγουμε F1 (GRPH) και F6(SET) για να διαλέξουμε το ιστόγραμμα (HIST) EXE και επιστρέφουμε στον πίνακα δεδομένων. Αν θέλουμε να καθορίσουμε τις κλίμακες του ιστογράμματος, αντί να αφήσουμε τον υπολογιστή να το κάνει, πληκτρολογούμε SHIFT MENU και επιλέγουμε Stat Wind : Manual. Προτείνεται επίσης να τοποθετηθεί το Graph Func στη θέση off. Επιστρέφουμε στον πίνακα δεδομένων με το EXE και επιλέγουμε τις κατάλληλες ρυθμίσεις χρησιμοποιώντας το SHIFT F3 (π.χ Xmin=0.0). Επιστρέφουμε στα δεδομένα με το EXE και πληκτρολογούμε F1 καθορίζοντας τις ρυθμίσεις στην οθόνη

Set Interval

Start :

Pitch :

Πληκτρολογώντας F6(DRAW) έχουμε στην οθόνη το ζητούμενο ιστόγραμμα (HIST). Με το SHIFT F1 (trace) και μετακινώντας το σταυρό που εμφανίζεται στο κέντρο της πάνω βάσης κάθε ορθογωνίου του ιστογράμματος, έχουμε τις ενδείξεις x και f (παρατήρηση και αντίστοιχη συχνότητα).



B Με χρήση Η/Υ (πρόγραμμα Microsoft EXCEL)

Ταξινομούμε τα δεδομένα του παραδείγματος σε κλάσεις εύρους 0.5 αρχίζοντας από την τιμή 1.75 μέχρι την 6.25.

❖ Εισαγωγή δεδομένων

Από το MENU επιλέγουμε το πρόγραμμα Microsoft EXCEL
Στο Βιβλίο Ι ,Φύλλο Ι και στη στήλη Α πληκτρολογούμε τις κλάσεις και στη στήλη Β τις συχνότητες π.χ.

A	B
1,75 - 2,25	2
2,25 - 2,75	2
2,75 - 3,25	2
3,25 - 3,75	3
3,75 - 4,25	3
4,25 - 4,75	7
4,75 - 5,25	7
5,25 - 5,75	3
5,75 - 6,25	1

❖ Επεξεργασία δεδομένων

Επιλέγουμε (μαυρίζουμε) τη στήλη Α και Β και από την γραμμή εργαλείων (toolbar) επιλέγουμε τον οδηγό γραφημάτων.

Οι δυνατότητες που έχουμε είναι δύο:

Βασικοί τύποι –Προσαρμοσμένοι τύποι γραφημάτων.

Στο συγκεκριμένο παράδειγμα χρησιμοποιούμε τους “βασικούς τύπους γραφημάτων” και ακολουθούμε τα παρακάτω βήματα.

Βασικοί τύποι

Τύπος γραφήματος “στήλες” (1° γράφημα) -Επόμενο

Δεδομένα προέλευσης

Περιοχή δεδομένων-Σειρά σε “γραμμές” -Επόμενο

Επιλογές γραφήματος

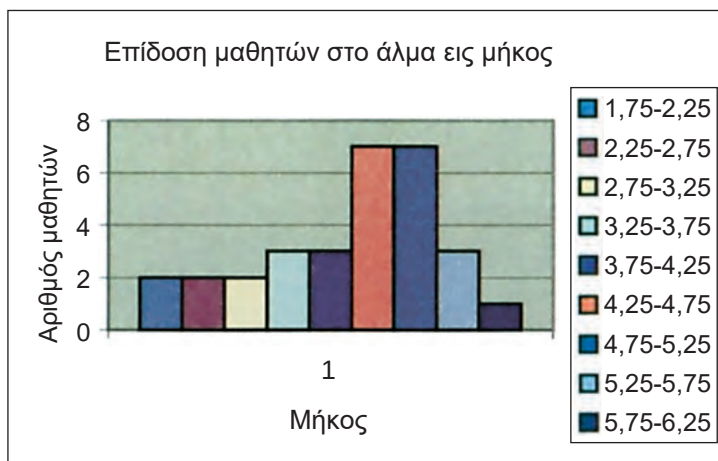
Τίτλοι : Τίτλος γραφήματος

Άξονας κατηγοριών(X)

Άξονας τιμών (Y) -Επόμενο

Τοποθέτηση γραφήματος

Ως αντικείμενο στο Φύλλο Ι -Τέλος





ΚΕΦΑΛΑΙΟ 4°

Να κατανοήσουν οι μαθητές τη σημασία
και τη χρησιμότητα των μέτρων θέσης μιας κατανομής
(μέσο εύρος, αριθμητικός μέσος, διάμεσος, επικρατούσα τιμή
και τεταρτημόρια).

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1, 2, 4, 5, 6, 7, 8 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του τέταρτου κεφαλαίου. Επαφίεται στους μαθητές να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στη τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

- 3 α) μειώνεται κατά δυο μονάδες
β) διπλασιάζεται

9 Γνωρίζουμε ότι $\bar{x} = \frac{\sum x}{v}$ άρα $\frac{\sum x}{v} \Leftrightarrow \sum x = 5v$ (1)

και $6 = \frac{\sum x + 13}{v + 1} \Leftrightarrow \sum x + 13 = 6(v + 1)$ (2)

αντικαθιστώντας τη σχέση (1) στη (2) προκύπτει ότι

$$5v + 13 = 6v + 6 \Leftrightarrow v = 7 \text{ Άρα το πλήθος των αριθμών είναι 7.}$$

- 10 α) Η διάμεσος, διότι συνήθως η κατανομή των μισθών είναι μη συμμετρική.
β) Ο μέσος, θεωρώντας ότι οι χρόνοι που απαιτούνται για να διανυθεί μια δεδομένη απόσταση αναμένεται να κατανέμονται συμμετρικά.
γ) Η επικρατούσα τιμή. Αν τα περισσότερα ρολόγια έδειχναν μια συγκεκριμένη ώρα, το πιθανότερο είναι να είναι η σωστή.
δ) Ο μέσος ή η διάμεσος. Εξαρτάται από το σχήμα της κατανομής.

- 11 Η μέση τιμή των αριθμών είναι 17 άρα θα ισχύει:

$$\frac{12 + 18 + 21 + x + 13}{5} = 17 \Leftrightarrow 12 + 18 + 21 + x + 13 = 85$$

$$64 + x = 85 \Leftrightarrow x = 85 - 64 \Leftrightarrow x = 21$$

ΑΣΚΗΣΕΙΣ

Ομάδα Α.

1 Οι ηλικίες των υπαλλήλων ενός εμπορικού καταστήματος είναι:

25, 22, 20, 20, 23, 25, 25, 30, 35, 40.

Να βρεθούν: i) Η μέση τιμή ii) η διάμεσος ηλικία και iii) το πρώτο και τρίτο τεταρτημόριο.

Λύση**i) Υπολογισμός της μέσης ηλικίας**

Όπως είναι γνωστό όταν δεν έχουμε συχνότητες η μέση τιμή δίνεται από τον τύπο:

$$\bar{x} = \frac{\sum x_i}{n}$$

όπου:

Σx: Το άθροισμα των παρατηρήσεών μας

n : Το πλήθος των παρατηρήσεών μας

Με εφαρμογή λοιπόν του παραπάνω τύπου θα έχουμε:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{25 + 22 + 20 + 20 + 23 + 25 + 25 + 30 + 35 + 40}{10} = \frac{265}{10} = 26,5$$

Δηλαδή, η μέση ηλικία (μέση τιμή) των υπαλλήλων του εμπορικού καταστήματος είναι: 26,5 χρόνια.

ii) Υπολογισμός διαμέσου

Τοποθετούμε κατ' αύξουσα τάξη τις παρατηρήσεις μας .

20 20 22 23 25 25 25 25 25 30 35 40

Το πλήθος των παρατηρήσεων είναι άρτιος αριθμός, άρα η διάμεσος βρίσκεται μεταξύ 5^{ου} και 6^{ου} όρου και συνεπώς η διάμεσος ισούται με:

$$\delta = \frac{25 + 25}{2} = 25$$

iii) Υπολογισμός α' και γ' τεταρτημορίου

Η θέση που κατέχει το πρώτο τεταρτημόριο Q_1 είναι $\frac{v+1}{4} = \frac{11}{4} = 2,75$
 Δηλαδή, το Q_1 είναι η τιμή που αντιστοιχεί στο 75% της απόστασης μεταξύ των μετρήσεων 2^{ης} και 3^{ης} θέσης που είναι 20 και 22.

Δηλαδή, $Q_1 = 20 + 0,75(22 - 20) = 21,5$. Αντίστοιχα η θέση που κατέχει το τρίτο τεταρτημόριο Q_3 είναι $\frac{3+1}{4} = \frac{31}{4} = 7,75$
 άρα η τιμή του Q_3 θα είναι στο 75% της απόστασης μεταξύ των μετρήσεων 7^{ης} και 8^{ης} θέσης που είναι 25 και 25.
 Άρα $Q_3 = 25$.

2 Οι μηνιαίες αποδοχές των υπαλλήλων μιας επιχείρησης σε χιλιάδες δρχ. είναι:

210, 220, 200, 180, 150, 160, 180, 200, 250, 260.

Να βρεθεί: i) Η μέση τιμή ii) η διάμεσος του μηνιαίου μισθού iii) το πρώτο και τρίτο τεταρτημόριο.

Λύση

i) Υπολογισμός της μέσης τιμής του μηνιαίου μισθού

Όταν δεν έχουμε συχνότητες ο μέσος μηνιαίος μισθός δίνεται, όπως γνωρίζουμε, από τον τύπο:

$$\bar{x} = \frac{\sum x}{v}$$

Όπου,

Σχ: Το άθροισμα των παρατηρήσεών μας

v: Το πλήθος των παρατηρήσεών μας.

Εφαρμόζοντας λοιπόν τον παραπάνω τύπο θα έχουμε:

$$\bar{x} = \frac{\sum x}{v} = \frac{210 + 220 + 200 + 180 + 150 + 160 + 180 + 200 + 250 + 260}{10}$$

$$\bar{x} = \frac{2010}{10} = 201$$

Δηλαδή το ύψος του μέσου μηνιαίου μισθού των υπαλλήλων της επιχείρησης είναι 201 χιλιάδες δρχ. ή 201000 δρχ. περίπου.

ii) Υπολογισμός της διάμεσου

Για να βρούμε τη διάμεσο του μηνιαίου μισθού τοποθετούμε κατ' αρχάς τα δεδομένα μας κατ' αύξουσα τάξη, δηλαδή:

150, 160, 180, 180, 200, 200, 210, 220, 250, 260

Το πλήθος των παρατηρήσεών μας είναι $n = 10$ οπότε

η διάμεσος κατέχει τη $\bar{x} = \frac{n+1}{2} = \frac{10+1}{2} = 5,5$

θέση, δηλαδή βρίσκεται ανάμεσα στη 5^η και 6^η παρατήρηση.

Κατά συνέπεια η τιμή της διαμέσου θα είναι:

$$\delta = \frac{200 + 200}{2} = 200 \text{ χιλιάδες } \delta\rho\chi. \text{ ή } \delta = 200.000 \text{ } \delta\rho\chi.$$

iii) Υπολογισμός του πρώτου και τρίτου τεταρτημορίου

Λαμβάνοντας υπόψη μας ότι $n = 10$, έχουμε:

$$\frac{n+1}{4} = \frac{10+1}{4} = 2,75$$

Δηλαδή, η τιμή του πρώτου τεταρτημορίου Q_1 εντοπίζεται ανάμεσα στη 2^η και 3^η θέση, άρα έχουμε $Q_1 = 160 + 0,75(180-160) = 160 + 0,75 \cdot 20 = 160 + 15 = 175$ δηλαδή 175000δρχ.

Αντίστοιχα το Q_3 εντοπίζεται ανάμεσα στη 7^η και 8^η παρατήρηση, άρα έχουμε

$$Q_3 = 210 + 0,75(220-210) = 210 + 7,5 = 217,5 \text{ δηλαδή } 217500 \text{ } \delta\rho\chi.$$

3 Η μέση τιμή του μηνιαίου μισθού των υπαλλήλων μιας επιχείρησης είναι 250 χιλιάδες δρχ.

- i) Αν ο μισθός κάθε υπαλλήλου αυξηθεί κατά 30 χιλιάδες δρχ., ποια μεταβολή θα επέλθει στη μέση τιμή του;
- ii) Αν αυξηθούν οι μισθοί όλων των εργαζομένων κατά 10% ποια μεταβολή θα επέλθει στη μέση τιμή του;

Λύση

- i) Όπως γνωρίζουμε σύμφωνα με την πρώτη ιδιότητα της μέσης τιμής, αν σε όλες τις τιμές μιας μεταβλητής x προσθέσουμε μια σταθερή ποσότητα k , στην περίπτωση μας 30 χιλιάδες δρχ., τότε ο x αυξάνεται κατά τη σταθερή αυτή ποσότητα k .

Δηλαδή, η μέση τιμή του μηνιαίου μισθού των υπαλλήλων αυτής της επιχείρησης θα αυξηθεί κατά 30 χιλιάδες δρχ. και κατά συνέπεια θα είναι:

$$\bar{x} = 250 + 30 = 280 \text{ χιλιάδες δρχ. ή } 280.000 \text{ δρχ.}$$

- ii) Σύμφωνα με τις ιδιότητες της μέσης τιμής θα ισχύει ότι:

$$\bar{x}^* = \alpha \bar{x}$$

Άρα η νέα μέση τιμή θα αυξηθεί κατά $10\% \cdot 250000 = 25000$

Ο παρακάτω πίνακας περιέχει τον αριθμό ωρών υπερωριακής απασχόλησης, που πραγματοποιήθηκε κατά τη διάρκεια μιας ορισμένης χρονικής περιόδου από τους εργάτες μιας βιομηχανίας:

Ωρες (x)	0	1	2	3	4	5	6	7	8	9	10
Εργάτες (v)	5	5	10	15	15	20	15	10	10	10	5

Να βρεθούν και να ερμηνευθούν: i) Η μέση τιμή, ii) η διάμεσος, iii) το πρώτο τεταρτημόριο.

Λύση

i) Υπολογισμός μέσης τιμής

Όπως γνωρίζουμε, όταν έχουμε κατανομή συχνοτήτων η μέση τιμή υπολογίζεται από τον τύπο:

$$\bar{x} = \frac{\sum vx}{\sum v}$$

Κατασκευάζουμε τον παρακάτω πίνακα:

Ώρες x	Αριθμός Εργατών v	vx	Αθροιστική Συχνότητα N
0	5	0	5
1	5	5	10
2	10	20	20
3	15	45	35
4	15	60	50
5	20	100	$70 \leftarrow \frac{v}{2} = 60$
6	15	90	85
7	10	80	105
9	10	90	115
10	5	50	120
Σύνολο	120	610	

Σύμφωνα με τα στοιχεία του παραπάνω πίνακα υπολογίσαμε τη μέση τιμή

$$\bar{x} = \frac{\sum vx}{\sum v} = \frac{610}{120} = 5,083$$

ή $\bar{x} \approx 5$

Δηλαδή, κατά τη διάρκεια της χρονικής περιόδου που εξετάσαμε η υπερωριακή απασχόληση των εργατών ήταν κατά μέσο όρο περίπου πέντε ώρες.

ii) Υπολογισμός της διαμέσου

Για να υπολογίσουμε τη διάμεσο δημιουργούμε τη στήλη των αθροιστικών συχνοτήτων (N). Κατόπιν βρίσκουμε ότι:

$$\frac{\sum v}{2} = \frac{120}{2} = 60$$

Εύκολα βλέπουμε πλέον ότι το $\frac{v}{2} = 60$

βρίσκεται μεταξύ της αθροιστικής συχνότητας

$N_{i-1} = 50$ και $N_i = 70$. Συνεπώς η διάμεσος είναι $\delta = 5$.

Δηλαδή το 50% των εργατών έχει κάνει το πολύ 5 ώρες υπερωρίες και το άλλο 50% έχει κάνει 5 ή και περισσότερες ώρες υπερωρίες.

iii) Υπολογισμός πρώτου τεταρτημορίου

Με τη βοήθεια και πάλι της στήλης των αθροιστικών συχνοτήτων το πρώτο τεταρτημόριο Q_1 εντοπίζεται στη θέση

$$\frac{v}{4} = \frac{120}{4} = 30$$

Ο αριθμός 30 περιέχεται μεταξύ των αθροιστικών συχνοτήτων 20 και 35 και συνεπώς $Q_1 = 3$.

Δηλαδή, το 25% των εργατών είχε το πολύ υπερωρίες 3 ωρών και το άλλο 75% 3 ή και περισσότερες ώρες υπερωρίας.

- 5** Ο παρακάτω πίνακας δείχνει τα ετήσια έξοδα πληρωμής λογαριασμών κατανάλωσης ηλεκτρικού ρεύματος σε χιλιάδες δρχ. 100 νοικοκυριών:

Έξοδα	Αριθμός Νοικοκυριών
40-60	5
60-80	10
80-100	10
100-120	15
120-140	20
140-160	20
160-180	10
180-200	5
200-220	5

Να υπολογιστεί και να ερμηνευθεί: i) Η μέση τιμή και ii) η διάμεσος

Λύση

Κατασκευάζουμε τον παρακάτω πίνακα αριθμητικών υπολογισμών:

Έξοδα	Αριθμός Νοικοκυριών v	Κεντρικοί όροι κλάσεων x	vx
[40-60)	5	50	250
[60-80)	10	70	700
[80-100)	10	90	900
[100-120)	15	110	1650
[120-140)	20	130	2600
[140-160)	20	150	3000
[160-180)	10	170	1700
[180-200)	5	210	1050
	100		12.800

ii) Υπολογισμός μέσου αριθμητικού

Όπως γνωρίζουμε ο μέσος αριθμητικός δίνεται από τον τύπο:

$$\bar{x} = \frac{\sum vx}{\sum v}$$

και επειδή $\sum v = 100$ και $\sum vx = 12.800$ έχουμε

$$\bar{x} = \frac{12.800}{100} = 128$$

Δηλαδή, η μέση ετήσια κατανάλωση ηλεκτρικού ρεύματος των 100 νοικοκυριών ήταν της τάξης των 128.000 δρχ.

ii) Υπολογισμός Διαμέσου

Σχηματίζουμε τον παρακάτω πίνακα:

Τάσεις Εξόδων	Αριθμός Νοικοκυριών v	Αθροιστικές Συχνότητες N
40-60	5	5
60-80	10	15
80-100	10	25
100-120	15	40
120-140	20	$60 \leftarrow \frac{v}{2} = 50$
140-160	20	80
160-180	10	90
180-200	5	95
200-220	5	100
	100	

Παρατηρούμε την 3^η στήλη του πίνακα και βλέπουμε ότι ο αριθμός

$$\frac{v}{2} = \frac{100}{2} = 50$$

βρίσκεται μεταξύ των αθροιστικών συχνοτήτων 40 και 60. Συνεπώς έχουμε:

$$\begin{aligned} \alpha_{i-1} &= 120, & c &= 20 \\ v_i &= 20, & N_{i-1} &= 40 \\ &\text{και} \end{aligned}$$

$$\delta = \alpha_{i-1} + \frac{c}{v_i} \left(\frac{v}{2} - N_{i-1} \right)$$

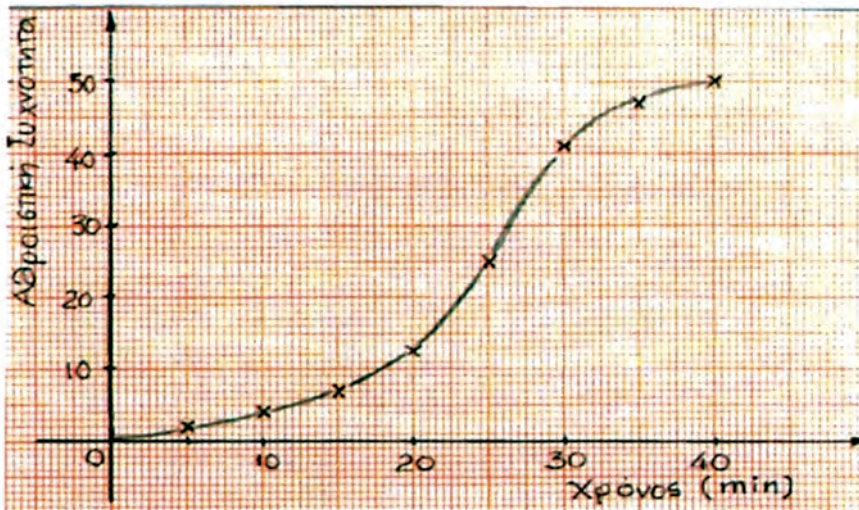
$$\delta = 120 + \frac{20}{20} (50 - 40)$$

$$\delta = 130$$

Δηλαδή, το 50% των νοικοκυριών δαπάνησε ετησίως για κατανάλωση ηλεκτρικού ρεύματος το πολύ 130.000 δρχ.

Ομάδα Β.

- 3** Από το παρακάτω διάγραμμα αθροιστικής συχνότητας παρατηρούμε το χρόνο που πέρασαν 50 άτομα μέσα σε ένα πολυκατάστημα:



- i) Από το διάγραμμα να βρείτε πόσα άτομα έμειναν στο πολυκατάστημα από 17 έως και 27 λεπτά.
- ii) Το 25% των ατόμων έμεινε λιγότερο ή ίσο με t λεπτά.
Να βρείτε το t .
- iii) Το 50% των ατόμων πέρασε περισσότερο από s λεπτά.
Να βρείτε το s .
- iv) Υπολογίστε τη διάμεσο, τι παρατηρείτε;

Λύση

- i) Από το διάγραμμα της αθροιστικής συχνότητας παρατηρούμε ότι 10 άτομα έμειναν στο πολυκατάστημα το πολύ 17 λεπτά, ενώ 34 άτομα έμειναν 27 λεπτά. Άρα από 17 ως 27 λεπτά έμειναν $34 - 10 = 24$ άτομα.
- ii) Από το διάγραμμα παρατηρούμε ότι το 25% των ατόμων παρέμεινε
- iii) Το 50% των ατόμων αντιστοιχεί στην 25 παρατήρηση, δηλαδή τα υπόλοιπα 25 άτομα πέρασαν περισσότερο από 25 λεπτά στο πολυκατάστημα.

iv) Η θέση της διαμέσου θα βρίσκεται στην $\frac{v}{2} = \frac{50}{2} = 25$ παρατήρηση και από το διάγραμμα παρατηρούμε ότι αντιστοιχεί στην τιμή $\delta = 25$ λεπτά, δηλαδή το πολύ 25 άτομα έμειναν στο πολυκατάστημα 25 λεπτά

2 Ο παρακάτω πίνακας δείχνει τους βαθμούς 30 μαθητών στο μάθημα της Στατιστικής της Α΄ τάξης των Τ.Ε.Ε. το 1999.

Βαθμοί	Αριθμός Μαθητών
7	1
8	2
11	3
12	1
13	2
14	5
15	7
16	5
17	3
18	1

- Να βρείτε τη μέση τιμή, τη διάμεσο, το Q_1 και Q_3 . Τι παρατηρείτε για τη βαθμολογία του τμήματος.
- Αν η βαθμολογία όλων των μαθητών αυξανόταν κατά δύο μονάδες πόση θα γινόταν η μέση τιμή;

Λύση

- Για τον υπολογισμό της μέσης τιμής:

$$\begin{aligned}\bar{x} &= \frac{\sum vx}{v} = \frac{1 \cdot 7 + 2 \cdot 8 + 3 \cdot 11 + 1 \cdot 12 + 2 \cdot 13 + 14 \cdot 5 + 15 \cdot 7 + 16 \cdot 5 + 17 \cdot 3 + 18 \cdot 1}{30} = \\ &= \frac{7 + 16 + 33 + 12 + 26 + 70 + 105 + 80 + 51 + 18}{30} = \frac{418}{30} \approx 13,9\end{aligned}$$

Για τον υπολογισμό της διαμέσου:

Στον πίνακα της άσκησης προσθέτουμε τη στήλη της αθροιστικής συχνότητας.

Βαθμοί	Αριθμός Μαθητών n_i	Αθροιστική Συχνότητα N_i
7	1	1
8	2	3
11	3	6
12	1	7
13	2	9
14	5	14
15	7	21
16	5	26
17	3	29
18	1	30

Θεωρούμε ότι η θέση της διαμέσου βρίσκεται στη $\frac{n+1}{2} = \frac{31}{2} = 15,5$

παρατήρηση δηλαδή βρίσκεται μεταξύ 15^{ης} και 16^{ης} παρατήρησης.

Οι παρατηρήσεις αυτές αντιστοιχούν στη $N_7 = 21$ αθροιστική συχνότητα και έχουν τιμές 15 και 15 αντίστοιχα.

Άρα η διάμεσος $\delta = \frac{15+15}{2} = 15$

Για τον υπολογισμό των τεταρτημορίων:

Ομοίως οι τιμές των τεταρτημορίων Q_1 και Q_3 θα υπολογίζονται ως εξής:

Η θέση του πρώτου τεταρτημορίου Q_1 θα βρίσκεται στην $\frac{n+1}{4} = 7,75$

παρατήρηση, δηλαδή μεταξύ 7^{ης} και 8^{ης} παρατήρησης και θα αντιστοιχεί σε αθροιστική συχνότητα $N_w = 9$. Άρα το πολύ το 25% των μαθητών θα έχουν βαθμό Στατιστικής 13.

Ομοίως το πολύ το 75% των μαθητών θα έχει βαθμό 16 στο μάθημα της Στατιστικής.

ii) Από τις ιδιότητες της μέσης τιμής υπολογίζουμε τη νέα μέση τιμή της βαθμολογίας των μαθητών σύμφωνα με τον τύπο $\bar{x}^* = \bar{x} + k$

δηλαδή $\bar{x}^* = 13,9 + 2$

$$\bar{x}^* = 15,9$$

3 Από τους 10 παίκτες μίας ομάδας μπάσκετ, οι 4 είναι Έλληνες και έχουν μέσο ύψος 200 εκ., οι 4 είναι κοινοτικοί και έχουν μέσο ύψος 208 εκ., και οι άλλοι 2 είναι ξένοι και έχουν μέσο ύψος 204 εκ.

i) Ποιο είναι το μέσο ύψος της ομάδας;

ii) Κατά τη διάρκεια της αγωνιστικής περιόδου ο ένας ξένος παίκτης, που είχε ύψος 200εκ. πήρε μεταγραφή σε άλλη ομάδα και ο προπονητής ήθελε στη θέση του να πάρει έναν Έλληνα.

Πόσο πρέπει να είναι το ύψος του Έλληνα παίκτη, ώστε το μέσο ύψος της ομάδας να είναι 205 εκ.;

Λύση

i) Το μέσο ύψος της ομάδας υπολογίζεται ως εξής:

$$\bar{x} = \frac{\sum vx}{v} = \frac{4 \cdot 200 + 2 \cdot 208 + 2 \cdot 204}{10} = \frac{800 + 832 + 408}{10} = \frac{2040}{10} = 204$$

ii) Το μέσο ύψος των δυο ξένων παικτών της ομάδας ήταν 204 εκ.

Ο παίκτης που πήρε μεταγραφή είχε ύψος 200 εκ., άρα ο άλλος που παρέμεινε στην ομάδα θα έχει ύψος :

$$\frac{x + 200}{2} = 204 \Leftrightarrow x + 200 = 408 \Leftrightarrow x = 208 \text{ εκ.}$$

Αν το ύψος του Έλληνα παίκτη που θα πάρει ο προπονητής είναι y , τότε θα ισχύει:

$$\frac{4 \cdot 200 + 4 \cdot 208 + y + 208}{10} = 205 \Leftrightarrow \frac{800 + 832 + 208 + y}{10} = 205$$

$$1840 + y = 2050 \Leftrightarrow y = 210$$

Άρα το ύψος του Έλληνα παίκτη θα είναι 210 εκ.

- 4 Δώδεκα μαθητές μίας τάξης εκτίμησαν το ύψος ενός καμπαναριού μιας εκκλησίας. Οι εκτιμήσεις τους σε μέτρα είναι:
47, 52, 52, 54, 52, 50, 51, 50, 48, 53, 54, 49.

- i) Υπολογίστε τη διάμεσο των εκτιμήσεων.
- ii) Υπολογίστε την επικρατούσα τιμή, δηλαδή την τιμή που έχει την μεγαλύτερη συχνότητα.
- iii) Υπολογίστε τον αριθμητικό μέσο, $\bar{x}$ των 12 εκτιμήσεων.
- iv) Ένας μαθητής από τους δώδεκα αναθεώρησε την εκτίμησή του και ο νέος αριθμητικός μέσος των 12 εκτιμήσεων είναι $\bar{x} + 0,5$

Βρείτε κατά πόσο αυξήθηκε η εκτίμηση αυτού του μαθητή.

- v) Ο καθηγητής της τάξης έκανε και αυτός μία εκτίμηση για το ύψος του καμπαναριού και όταν η εκτίμησή του προστέθηκε στις αρχικές 12, ο αριθμητικός μέσος των 13 πια εκτιμήσεων είναι $x + 0,5$. Βρείτε την εκτίμηση του καθηγητή.

Λύση

- i) Τοποθετούμε τις εκτιμήσεις σε αύξουσα τάξη:
47, 48, 49, 50, 50, 51, 52, 52, 52, 53, 54, 54

Η διάμεσος θα βρίσκεται στην $\frac{12+1}{2} = 6,5$

δηλαδή ανάμεσα στην 6^η και $\frac{51+52}{2} = 51,5$

την 7^η παρατήρηση και έχει τιμή

- ii) Επικρατούσα τιμή είναι η τιμή 52 γιατί εμφανίζει τη μεγαλύτερη συχνότητα.

iii) $\bar{x} = \frac{47 + 48 + 49 + 50 + 50 + 51 + 52 + 52 + 52 + 53 + 54 + 54}{12} = 51$

- iv) Αν συμβολίσουμε με y την αύξηση της εκτίμησης του μαθητή, τότε η νέα μέση τιμή θα ισούται:

$$\bar{x}^* = \frac{x_1 + x_2 + x_3 + \dots + x_k + \dots + x_{12}}{12} = \frac{x_1 + x_2 + x_3 + \dots + (x_k + y) + \dots + x_{12}}{12} =$$

$$= \frac{x_1 + x_2 + x_3 + \dots + x_k + \dots + x_{12} + y}{12} = \frac{\sum x_i + y}{12} \quad \text{και} \quad \bar{x}^* = \bar{x} + 0,5 \quad \text{δηλαδή:}$$

$$\bar{x} + 0,5 = \frac{\sum x_i + y}{12} \Leftrightarrow \bar{x} + 0,5 = \frac{\sum x_i}{12} + \frac{y}{12} \Leftrightarrow \bar{x} + 0,5 = \bar{x} + \frac{y}{12} \Leftrightarrow 0,5 = \frac{y}{12} \Leftrightarrow y = 6$$

ν) Έστω α η εκτίμηση του καθηγτη τότε: $\bar{x}^* = \frac{\sum x_i + \alpha}{13} = \bar{x} + 0,5$

από τα προηγούμενα αποτελέσματα γνωρίζουμε ότι

$$\bar{x} = 51 \text{ και } \sum x_i = 612. \text{ Άρα } \frac{612 + \alpha}{13} = 51 + 0,5 \Leftrightarrow 612 + \alpha = 13 \cdot 51,5$$

$$\alpha = 669,5 - 612 \Leftrightarrow \alpha = 57,5$$

5 Οι μηνιαίες αποδοχές 100 υπαλλήλων μιας επιχείρησης κυμαίνονται μεταξύ 200 και 300 χιλιάδων δρχ.

Ακόμη γνωρίζουμε ότι:

10 υπάλληλοι αμείβονται το πολύ με 220 χιλ. δρχ.

30 υπάλληλοι αμείβονται το πολύ με 240 χιλ. δρχ.

40 υπάλληλοι έχουν μισθούς μεγαλύτερους από 260 χιλ. δρχ.

και 10 υπάλληλοι έχουν μισθούς μεγαλύτερους από 280 χιλ. δρχ.

Ζητείται: i) Να υπολογιστεί η μέση τιμή, ii) να υπολογιστεί η διάμεσος και iii) να υπολογιστεί το πρώτο και το τρίτο τεταρτημόριο.

Λύση

Έχοντας υπόψη μας τα δεδομένα της άσκησης σχηματίζουμε τον παρακάτω πίνακα.

Μηνιαίες Αποδοχές	Αριθμός υπαλλήλων	Κεντρικοί όροι κλάσεων x	vx
200-220	10	210	2100
220-240	20	230	4600
240-260	30	250	7500
260-280	30	270	8100
280-300	10	290	2900
Σύνολο	100		$\sum vx = 25.200$

i) Υπολογισμός μέσης τιμής

Στον τύπο εύρεσης της μέσης τιμής θέτουμε $\Sigma v = 100$ και $\Sigma vx = 25.200$ και βρίσκουμε:

$$\bar{x} = \frac{\sum vx}{\sum i} = \frac{25.200}{100} = 252$$

Δηλαδή, 252 χιλιάδες δραχ. είναι ο μέσος μηνιαίος μισθός των υπαλλήλων αυτής της επιχείρησης.

ii) Υπολογισμός της διάμεσου

Σχηματίζουμε τον παρακάτω πίνακα.

Τάξεις μηνιαίων αποδοχών	Αριθμός υπαλλήλων v	Αθροιστικές Συχνότητες N
200-220	10	10
220-240	20	$30 \searrow \frac{N}{4} = 25$
240-260	30	$60 \searrow \frac{N}{2} = 50$
260-280	30	$90 \searrow \frac{3N}{4} = 75$
280-300	10	100
	$v=100$	

Ο αριθμός $\frac{v}{2} = \frac{100}{2} = 50$, όπως βλέπουμε, βρίσκεται μεταξύ των

αθροιστικών συχνοτήτων 30 και 60, άρα έχουμε:

$$\begin{aligned} \alpha_{i-1} &= 240, & c &= 20 \\ i_{i-1} &= 30, & N_{i-1} &= 30 \end{aligned}$$

και
$$\delta = \alpha_{i-1} + \frac{c}{v_i} \left(\frac{v}{2} - i_{i-1} \right)$$

$$\delta = 240 + \frac{20}{30} (50 - 30)$$

$$\delta \approx 253,33$$

Δηλαδή, ο διάμεσος μηνιαίος μισθός είναι 253.330 δραχ. περίπου.

iii) Υπολογισμός των τεταρτημορίων

α) Υπολογισμός του πρώτου τεταρτημορίου Q_1

Ο αριθμός $\frac{v}{4} = 25$, όπως βλέπουμε, βρίσκεται μεταξύ των αθροιστικών

συχνοτήτων 10 και 30, άρα έχουμε:

$$\begin{aligned} \alpha_{i-1} &= 220 & , & & \delta &= 20 \\ v_i &= 20 & , & & N_{i-1} &= 10 \end{aligned}$$

$$\begin{aligned} \text{και} \quad Q_1 &= \alpha_{i-1} + \frac{c}{v_i} \left(\frac{v}{4} - N_{i-1} \right) \\ Q_1 &= 220 + \frac{20}{20} (25 - 10) \\ Q_1 &= 235 \end{aligned}$$

Δηλαδή, το 25% των υπαλλήλων έχει το πολύ μηνιαίες αποδοχές 235.000 δρχ. και το υπόλοιπο 75% περισσότερες από 235.000 δρχ.

β) Υπολογισμός του τρίτου τεταρτημορίου (Q_3)

Εργαζόμενοι κατά τρόπο παρόμοιο όπως στην εύρεση του Q_1 έχουμε:

$$\begin{aligned} \frac{3v}{4} &= 75 & \alpha_{i-1} &= 260 & , & c &= 20 \\ v_i &= 30 & , & & N_{i-1} &= 60 \end{aligned}$$

$$\begin{aligned} \text{και} \quad Q_3 &= \alpha_{i-1} + \frac{c}{v_i} \left(\frac{3v}{4} - N_{i-1} \right) \\ Q_3 &= 260 + \frac{20}{30} (75 - 60) \\ Q_3 &= 270 \end{aligned}$$

Δηλαδή το 75% των υπαλλήλων έχει το πολύ μηνιαίες αποδοχές 270.000δρχ. και το υπόλοιπο 25% περισσότερες από 270.000 δρχ.



ΚΕΦΑΛΑΙΟ 5°

Να κατανοήσουν οι μαθητές τη σημασία
και τη χρησιμότητα των μέτρων διασποράς
μιας κατανομής (εύρος μεταβολής,
μέση απόλυτη απόκλιση, διακύμανση
και τυπική απόκλιση).

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_1 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v} = \frac{1}{v} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + \dots + (x_v^2 - 2x_v\bar{x} + \bar{x}^2)]$$

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1, 2, 4, 5, 8, 9, 10, 11, 13 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του πέμπτου κεφαλαίου. Οι μαθητές πρέπει να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στην τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

3. Γ: 64

6. Γ: 3,74

Υπολογίζω το α : $\frac{3 + 6 + 7 + \alpha + 14}{5} = 8 \Leftrightarrow 30 + \alpha = 40 \Leftrightarrow \alpha = 10$

άρα οι αριθμοί είναι 3, 6, 7, 10, 14 τότε

$$s^2 = \frac{\sum x^2}{v} - \bar{x}^2 = \frac{3^2 + 6^2 + 7^2 + 10^2 + 14^2}{5} - 64 = \frac{9 + 36 + 49 + 100 + 196}{5} - 64 = \frac{390}{5} - 64 = 14$$

$$\text{και } s = \sqrt{14} = 3,74$$

7. Β: 5,1

$$s^2 = \frac{1}{v} \sum (x - \bar{x})^2 \quad \text{άρα } s^2 = \frac{234}{9} = 26 \quad \text{και } s = 5,099$$

12. Γ:1

$$CV_x = \frac{s_x}{\bar{x}} \Leftrightarrow s_x = \bar{x} \cdot CV_x$$

$$s_x = 2 \cdot 0,5$$

$$s_x = 1$$

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_1 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v} = \frac{1}{v} [(x_1^2 - 2x_1\bar{x}) + \dots + (x_v^2 - 2x_v\bar{x})]$$

ΑΣΚΗΣΕΙΣ

Ομάδα Α'

- 1 Να υπολογίσετε τη μέση τιμή, τη μέση απόλυτη απόκλιση και την τυπική απόκλιση των 4 αριθμών: 2, 3, 6, 9. Δύο αριθμοί α και β προστίθενται στο σύνολο των τεσσάρων αριθμών έτσι, ώστε ο μέσος αυξάνεται κατά 1 και η διακύμανση κατά 2,5. Να βρεθούν τα α και β .

Λύση

Η μέση τιμή για τους αριθμούς 2, 3, 6, 9 είναι:

$$\bar{x} = \frac{2 + 3 + 6 + 9}{4} = \frac{20}{4} = 5$$

$$MAA = \frac{|2-5| + |3-5| + |6-5| + |9-5|}{4} = \frac{3+2+1+4}{4} = 2,50$$

$$s^2 = \frac{\sum x^2}{v} - \bar{x}^2 = \frac{4+9+36+81}{4} - 25 = 32,50 - 25 = 7,50$$

$$s = \sqrt{7,5} = 2,74$$

$$\frac{2+3+6+9+\alpha+\beta}{6} = 6 \Leftrightarrow 20 + \alpha + \beta = 36 \Leftrightarrow \alpha + \beta = 16$$

$$\frac{4+9+36+81+\alpha^2+\beta^2}{6} - 36 = 10 \Leftrightarrow 130 + \alpha^2 + \beta^2 - 216 = 60 \Leftrightarrow$$

$$\Leftrightarrow \alpha^2 + \beta^2 = 60 + 216 - 130 \Leftrightarrow \alpha^2 + \beta^2 = 146$$

Θα λύσουμε το σύστημα:

$$\begin{cases} \alpha + \beta = 16 \\ \alpha^2 + \beta^2 = 146 \end{cases} \Leftrightarrow \begin{cases} \alpha = 16 - \beta \\ (16 - \beta^2) + \beta^2 = 146 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \alpha = 16 - \beta \\ 256 - 32\beta + \beta^2 + \beta^2 = 146 \end{cases} \Leftrightarrow \begin{cases} \alpha = 16 - \beta \\ 2\beta^2 - 32\beta + 110 = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \alpha = 16 - \beta \\ \beta^2 - 16\beta + 55 = 0 \end{cases}$$

Από τον τύπο επίλυσης της β' βαθμίας εξίσωσης έχουμε:

$$\beta_1 = \frac{16 + \sqrt{36}}{2} = \frac{16 + 6}{2} = 11$$

$$\beta_2 = \frac{16 - 6}{2} = 5$$

Άρα για $\beta = 11$ $\alpha = 16 - 11 = 5$

και για $\beta = 5$ $\alpha = 16 - 5 = 11$

- 2 Οι μηνιαίοι μισθοί 5 υπαλλήλων μιας επιχείρησης σε εκατοντάδες χιλιάδες δρχ. είναι: 2, 3, 3, 4, 5.** Να βρεθούν και να ερμηνευθούν i) η μέση τιμή, ii) η μέση απόλυτη απόκλιση και iii) η τυπική απόκλιση.

Λύση

Υπολογίζουμε: **i)** τη μέση τιμή $\bar{x} = \frac{2 + 3 + 3 + 4 + 5}{5} = \frac{17}{5} = 3,4$.

$$\begin{aligned} \text{MAA} &= \frac{|2 - 3,4| + |3 - 3,4| + |3 - 3,4| + |4 - 3,4| + |5 - 3,4|}{5} \\ &= \frac{1,4 + 0,4 + 0,4 + 0,6 + 1,6}{5} = \frac{4,4}{5} = 0,88 \end{aligned}$$

Δηλαδή, οι μηνιαίοι μισθοί των 5 υπαλλήλων της επιχείρησης απέχουν από το μέσο μηνιαίο μισθό $\bar{x}$ κατά μέσο όρο 88000 δρχ. ($0,88 \cdot 100000$).

iii) διακύμανση:

$$s^2 = \frac{\sum x^2}{n} - \bar{x}^2 = \frac{4 + 9 + 9 + 16 + 25}{5} - 3,4^2 = \frac{63}{5} - 11,56 = 12,6 - 11,56 = 1,04$$

$$\text{και } s = \sqrt{1,04} = 1,009.$$

Δηλαδή, η τετραγωνική ρίζα της μέσης τιμής των τετραγώνων των αποκλίσεων των μηνιαίων μισθών των υπαλλήλων από το μέσο μηνιαίο μισθό τους είναι περίπου 10000 δρχ. ($1,009 \cdot 100000$).

- 3 Να δείξετε ότι η τυπική απόκλιση των αριθμών 1, 2, 3, 4, 5, 6, 7 είναι 2. Χρησιμοποιώντας το αποτέλεσμα που βρήκατε υπολογίσετε την τυπική απόκλιση των αριθμών:**

i) 11, 12, 13, 14, 15, 16, 17

ii) 100, 200, 300, 400, 500, 600, 700

Λύση

Υπολογίζουμε τη διακύμανση των αριθμών: 1, 2, 3, 4, 5, 6, 7

Η μέση τιμή είναι: $\bar{x} = \frac{1+2+3+4+5+6+7}{7} = \frac{28}{7} = 4$

και

$$s^2 = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2}{7} - 4^2 = \frac{1+4+9+16+25+36+49}{7} - 16 = 20 - 16 = 4$$

Άρα $s = 2$

Παρατηρούμε ότι

11	12	13	14	15	16	17
↓	↓	↓	↓	↓	↓	↓
(10 + 1)	(10 + 2)	(10 + 3)	(10 + 4)	(10 + 5)	(10 + 6)	(10 + 7)

Δηλαδή οι αριθμοί:

y:	11, 12, 13, 14, 15, 16, 17
----	----------------------------

προκύπτουν αν στους αριθμούς

x:	1, 2, 3, 4, 5, 6, 7
----	---------------------

προσθέσουμε τον αριθμό 10.

Είναι γνωστό όμως από τις ιδιότητες της διακύμανσης ότι: “αν σε όλες τις τιμές μιας μεταβλητής προσθέσουμε μια σταθερή ποσότητα α, τότε η διακύμανση και η τυπική απόκλιση $s_x = s_y$ παραμένουν αμετάβλητες”. Άρα η τυπική απόκλιση των αριθμών:

11, 12, 13, 14, 15, 16, 17 είναι 2.

Επίσης οι αριθμοί

y:	100, 200, 300, 400, 500, 600, 700
----	-----------------------------------

προκύπτουν αν οι αριθμοί

x:	1, 2, 3, 4, 5, 6, 7
----	---------------------

πολλαπλασιασθούν επί 100, δηλαδή:

$1 \times 100, 2 \times 100, 3 \times 100, 4 \times 100, 5 \times 100, 6 \times 100, 7 \times 100$

Και πάλι από τις ιδιότητες της διακύμανσης είναι γνωστό ότι, αν όλες οι τιμές μιας μεταβλητής πολλαπλασιασθούν με μια σταθερή ποσότητα α , τότε και η διακύμανση (β^2) πολλαπλασιάζεται με λ^2 και η τυπική απόκλιση (s) πολλαπλασιάζεται με λ .

$$\begin{aligned} \text{Δηλαδή: } s_y &= 100 \times s_x \\ s_y &= 100 \times 2 \quad \text{Άρα } s_y = 200 \end{aligned}$$

4 Οι εβδομαδιαίες αποδοχές δύο ομάδων εργαζομένων σε χιλιάδες δρχ. δίνονται παρακάτω:

A' ομάδα:	100	120	125	100	120	130
B' ομάδα:	100	110	120	100	125	140

Να βρεθούν:

- Η μέση τιμή των αποδοχών της κάθε ομάδας εργαζομένων.
- Η τυπική απόκλιση κάθε ομάδας.
- Η μέση απόκλιση κάθε ομάδας.
- Να συγκριθούν και να σχολιαστούν τα αποτελέσματα.

Λύση

Υπολογισμός στατιστικών μέτρων πρώτης ομάδας.

- Υπολογισμός της μέσης τιμής των αποδοχών της A' ομάδας.

$$\bar{x} = \frac{\sum x}{v} = \frac{100 + 120 + 125 + 100 + 120 + 130}{6} = \frac{695}{6} = 115,833$$

Δηλαδή 115.833 δρχ.

ii) Υπολογισμός της τυπικής απόκλισης της A' ομάδας.

Αρχικά βρίσκουμε το Σx^2

$$\begin{aligned} \Sigma x^2 &= 100^2 + 120^2 + 125^2 + 100^2 + 120^2 + 130^2 = \\ &= 10.000 + 14.400 + 15.625 + 10.000 + 14.400 + 16.900 \end{aligned}$$

και

$$\Sigma x^2 = 81.325$$

Αντικαθιστώντας στο τύπο υπολογισμού της τυπικής απόκλισης όπου $\sum x^2 = 81.325$, $\bar{x} = 115,833$ και $n = 6$, βρίσκουμε:

$$s = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

$$s = \sqrt{\frac{81325}{6} - 115,833^2}$$

$$s = \sqrt{13.554,166 - 13.417,283}$$

$$s = 11,699 \approx 11,7 \text{ και } s \approx 11,7$$

Δηλαδή 11.700 δρχ.

iii) Υπολογισμός Μ.Α.Α

$$\frac{|100 - 115,833| + |120 - 115,833| + |125 - 115,833| + |100 - 115,833| + |120 - 115,833| + |130 - 115,833|}{6}$$

$$= \frac{15,833 + 4,167 + 9,167 + 15,833 + 4,167 + 14,167}{6} = 10,55$$

δηλαδή 10600δρχ.

Υπολογισμός στατιστικών μέτρων δεύτερης ομάδας

i) Υπολογισμός της μέσης τιμής των αποδοχών της Β' ομάδας

$$\bar{x} = \frac{\sum x}{n} = \frac{100 + 110 + 120 + 100 + 125 + 140}{6} = \frac{695}{6} = 115,833$$

ii) Υπολογισμός της τυπικής απόκλισης της δεύτερης ομάδας.

Αρχικά βρίσκουμε το $\sum x^2$

$$\sum x^2 = 100^2 + 110^2 + 120^2 + 100^2 + 125^2 + 140^2 =$$

$$= 10.000 + 12.100 + 14.400 + 10.000 + 15.625 + 19.600$$

και

$$\sum x^2 = 81.725$$

Αντικαθιστώντας στο τύπο υπολογισμού της τυπικής απόκλισης όπου $\sum x^2 = 81.725$, $\bar{x} = 115,833$ και $n = 6$, βρίσκουμε:

$$s = \sqrt{\frac{\sum x^2}{n} - \mu^2}$$

$$s = \sqrt{\frac{81725}{6} - 115,833^2}$$

$$s = \sqrt{13.620,833 - 13.417,283}$$

και

$$s \approx 14,27 \quad \text{δηλαδή } 14270$$

iii) Υπολογισμός M.A.A

$$\frac{|100 - 115,833| + |110 - 115,833| + |120 - 115,833| + |100 - 115,833| + |125 - 115,833| + |140 - 115,833|}{6} =$$

$$= 12,5 \quad \text{δηλαδή } 125.000 \text{ δρχ.}$$

Όπως παρατηρούμε οι συνολικές αποδοχές των δύο ομάδων είναι ίσες. Όμως, συγκρίνοντας την τυπική απόκλιση της δεύτερης ομάδας ($s = 14,26$) με αυτή της πρώτης ομάδας ($s = 11,7$) βλέπουμε ότι οι εβδομαδιαίες αποδοχές της δεύτερης ομάδας παρουσιάζουν μεγαλύτερη διασπορά από αυτές της πρώτης.

3 Το ύψος των ετησίων εξόδων ενοικίου που καταβάλλουν 25 εμπορικές επιχειρήσεις δίνεται από τον ακόλουθο πίνακα σε εκατομμύρια δρχ.

x_i	2	3	5	6	7	8	9	10
v_i	1	2	6	7	4	2	2	1

Να βρεθούν και να ερμηνευθούν:

- i) Η μέση τιμή του ενοικίου
- ii) Η διάμεση τιμή του ενοικίου
- iii) Το εύρος μεταβολής.
- iv) Η διακύμανση και η τυπική απόκλιση.

Λύση

- i) Υπολογισμός της μέσης τιμής
Συνθέτουμε τον παρακάτω πίνακα:

(1) x_i	(2) v_i	(3) $v_i \cdot x_i$
2	1	2
3	2	6
5	6	30
6	7	42
7	4	28
8	2	16
9	2	18
10	1	10
	25	152

Αντικαθιστώντας τα αθροίσματα των στηλών (2) και (3), δηλαδή τα $\Sigma v = 25$ και $\Sigma v \cdot x = 152$, στον τύπο υπολογισμού του μέσου αριθμητικού έχουμε:

$$\bar{x} = \frac{\sum v \cdot x}{\sum v} = \frac{152}{25} = 6,08 \approx 6$$

Δηλαδή η μέση τιμή του ενοικίου είναι ίση με 6 δηλαδή 6 εκ.δρχ

ii) Υπολογισμός της διαμέσου

Συνθέτουμε τον παρακάτω αριθμητικό πίνακα:

x	v	N
2	1	1
3	2	3
5	6	9
6	7	$16 \searrow \frac{v}{2} = \frac{25}{2} = 12,5$
7	4	20
8	2	22
9	2	24
10	1	25

Εργαζόμενοι κατά τα γνωστά βρίσκουμε ότι, ο αριθμός $\frac{v}{2} = \frac{25}{2} = 12,5$

βρίσκεται μεταξύ των αθροιστικών συχνοτήτων 9 και 16 και συνεπώς η διάμεσος είναι $\delta = 6$.

Δηλαδή, οι μισές επιχειρήσεις καταβάλλουν ετησίως ενοίκιο 6 εκατομμύρια και κάτω και οι άλλες μισές 6 εκατομμύρια και πάνω.

iii) Υπολογισμός του εύρους μεταβολής

Η μικρότερη τιμή που έχει πάρει η x είναι το 2 και η μεγαλύτερη, όπως παρατηρούμε στον πίνακα, είναι το 10, συνεπώς έχουμε:

Εύρος μεταβολής: $10 - 2 = 8$

Το εύρος μεταβολής των ετησίων εξόδων ενοικίου είναι πράγματι μεγάλο (ίσο με 8), όμως είναι παρακινδυνευμένο να εξαγάγουμε συμπεράσματα, πριν υπολογίσουμε και τα άλλα μέτρα διασποράς, αφού γνωρίζουμε ότι το εύρος μεταβολής έχει το μειονέκτημα να εξαρτάται αποκλειστικά από τις δύο ακραίες τιμές της x.

- iv) Υπολογισμός διακύμανσης της τυπικής απόκλισης
Συνθέτουμε τον παρακάτω πίνακα αριθμητικών υπολογισμών:

x	v	x ²	v·x	vx ²
2	1	4	2	4
3	2	9	6	18
5	6	25	30	150
6	7	36	42	252
7	4	49	28	196
8	2	64	16	128
9	2	81	18	162
10	1	100	10	100
	25		152	1010

Αντικαθιστώντας τα αθροίσματα $\Sigma v = 25$,
 $\Sigma vx = 152$ και $\Sigma vx^2 = 1010$ στον τύπο υπολογισμού της διακύμανσης έχουμε:

$$s^2 = \frac{\sum vx^2}{N} = \left(\frac{\sum vx}{N} \right)^2$$

και

$$s^2 = \frac{1010}{25} = \left(\frac{152}{25} \right)^2 = 40,4 - 36,966$$

$$s^2 \approx 3,434$$

Για να βρούμε τώρα την τυπική απόκλιση s βρίσκουμε τη θετική τετραγωνική ρίζα της διακύμανσης s^2 , συνεπώς έχουμε

$$s = \sqrt{s^2} \approx \sqrt{3,434}$$

και

$$s \approx 1,853$$

Όπως γνωρίζουμε όσο μικρότερες είναι οι τιμές της διακύμανσης και της τυπικής απόκλισης, τόσο πιο συγκεντρωμένες γύρω από το μ βρίσκονται οι τιμές της x και τόσο πιο αντιπροσωπευτικός είναι ο μ . Στην προκειμένη περίπτωση η τιμή της s μας λέει ότι η τετραγωνική ρίζα του αριθμητικού μέσου των τετραγώνων των αποκλίσεων των δεδομένων μας από τη μέση βαθμολογία s είναι ίση με 1,853.

- 4** Ο παρακάτω πίνακας δίνει τα ύψη των ετησίων διαφημιστικών εξόδων σε εκατ. δρχ. που έγιναν από τις 30 πρώτες σε πωλήσεις επιχειρήσεις, ενός εμπορικού κλάδου.

Διαφημιστικά Έξοδα	Αριθμός Επιχειρήσεων
[0-20)	2
[20-40)	3
[40-60)	8
[60-80)	10
[80-100)	4
[100-120)	2
[120-140)	1
Σύνολο	30

Ζητείται να υπολογισθούν και να ερμηνευθούν:

- Η μέση τιμή
- Η διακύμανση.
- Η τυπική απόκλιση.

Λύση

i) Υπολογισμός της μέσης τιμής

Κατασκευάζουμε τον παρακάτω πίνακα αριθμητικών υπολογισμών:

Κλάσεις	v_i	x_i	$v_i \cdot x_i$
[0-20)	2	10	20
[20-40)	3	30	90
[40-60)	8	50	400
[60-80)	10	70	700
[80-100)	4	90	360
[100-120)	2	110	220
[120-140)	1	130	130
	$v=30$		$\sum v x = 1920$

Στον τύπο υπολογισμού του μέσου αριθμητικού αντικαθιστούμε τα αθροίσματα:

$$\Sigma v = 30, \Sigma v x = 1920$$

και έχουμε:

$$\bar{x} = \frac{\sum v x}{\sum i} = \frac{1920}{30} = 64$$

Άρα

$$\bar{x} = 64$$

Δηλαδή, αυτές οι 30 επιχειρήσεις στη χρονιά που εξετάσαμε δαπάνησαν κατά μέσο όρο 64.000.000 δρχ. για διαφημιστικά έξοδα.

ii) Υπολογισμός της διακύμανσης

Κατασκευάζουμε τον παρακάτω αριθμητικό πίνακα:

Κλάσεις	v_i	x_i	$v_i \cdot x_i$	x_i^2	$v_i x_i^2$
[0-20)	2	10	20	100	200
[20-40)	3	30	90	900	2.700
[40-60)	8	50	400	2.500	20.000
[60-80)	10	70	700	4.900	49.000
[80-100)	4	90	360	8.100	32.400
[100-120)	2	110	220	12.100	24.200
[120-140)	1	130	130	16.900	16.900
	$v=30$		$\sum vx = 1920$		$\sum vx^2 = 145.400$

Αντικαθιστούμε στον τύπο υπολογισμού της διακύμανσης τα αθροίσματα:

$$\sum v = 30, \sum vx = 1920$$

$$\sum vx^2 = 145.400$$

και έχουμε:

$$s^2 = \frac{\sum vx^2}{v} = \left(\frac{\sum vx}{v} \right)^2$$

$$s^2 = \frac{145.400}{30} = \left(\frac{1920}{30} \right)^2 = 4.846,666 - 4.096$$

$$\text{και } s^2 = 750,666$$

iii) Η τυπική απόκλιση είναι:

$$s = \sqrt{s^2} = \sqrt{750,666}$$

$$\text{και } s = 27,398 \approx 27,4$$

Όπως γνωρίζουμε, όσο μικρότερες είναι οι τιμές της s^2 (και της s φυσικά) τόσο πιο συγκεντρωμένες γύρω από τη μέση τιμή βρίσκονται οι παρατηρήσεις μας. Επίσης η τετραγωνική ρίζα του μέσου αριθμητικού των τετραγώνων των αποκλίσεων των δεδομένων μας από τη μέση ετήσια δαπάνη διαφήμισης (64 εκατ. δρχ.) ισούται με 27,4 εκατ. δρχ. περίπου.

Ομάδα Β'

1 Σε εργοστάσιο συσκευασίας αλατιού, ο υπεύθυνος ισχυρίζεται ότι σε κάθε σάκο αλάτι περιέχονται 25 kg. Επιλέχθηκαν 80 σάκοι και μετρήθηκε η μάζα (x σε kg) του κάθε σάκου. Τα αποτελέσματα ήταν τα εξής:

$$\Sigma (x-25) = 27,2 \text{ και } \Sigma (x-25)^2 = 85,1$$

Να βρείτε τη μέση τιμή και την τυπική απόκλιση των μαζών.

Λύση

Από την καταγραφή των μετρήσεων της μάζας (x) κάθε σάκου καταλήγαμε ότι:

$$\Sigma (x - 25) = 27,2 \text{ δηλαδή:}$$

$$(x_1 - 25) + (x_2 - 25) + (x_3 - 25) + (x_4 - 25) + (x_5 - 25) + (x_6 - 25) + (x_7 - 25) + \dots + (x_{80} - 25) = 27,2$$

$$\underbrace{x_1 + x_2 + x_3 + x_4 + \dots + x_{80}}_{\downarrow} - \underbrace{(25 + 25 + 25 + \dots 25)}_{80} = 27,2$$

$$\Sigma x - 80,25 = 27,2$$

$$\Sigma x = 27,2 + 80,25$$

$$\Sigma x = 2027,2$$

$$\text{Άρα } \bar{x} = \frac{\Sigma x}{v} = \frac{2027,2}{80} = 25,34$$

$$\text{Επίσης } \Sigma (x - 25)^2 = 85,1$$

δηλαδή :

$$(x_1 - 25)^2 + (x_2 - 25)^2 + (x_3 - 25)^2 + \dots + (x_{80} - 25)^2 = 85,1$$

$$(x_1^2 - 50x_1 + 625) + (x_2^2 - 50x_2 + 625) + (x_3^2 - 50x_3 + 625) + \dots + (x_{80}^2 - 50x_{80} + 625) = 85,1 \Leftrightarrow$$

$$(x_1^2 + x_2^2 + \dots + x_{80}^2) - 50(x_1 + x_2 + x_3 + \dots + x_{80}) + 80 \cdot 625 = 85,1 \Leftrightarrow$$

$$\sum x^2 - 50 \sum x + 80 \cdot 625 = 85,1$$

$$\sum x^2 - 50 \cdot 2027,2 + 80 \cdot 625 = 85,1$$

$$\sum x^2 = 85,1 + 50 \cdot 2027,2 - 80 \cdot 625$$

$$\sum x^2 = 101.445,1 - 50.000$$

$$\sum x^2 = 51.445,1$$

$$\text{Άρα } s^2 = \frac{\sum x^2}{v} - \bar{x}^2 = \frac{51445,1}{80} - 25,34^2 = 643,06 - 642,12 = 0,94$$

και

$$s \cong 0,97$$

Η μέση τιμή της μάζας των 80 σάκων είναι 25,34 kg και η τυπική απόκλιση είναι 0,97 kg.

2 Το ύψος των ετήσιων κερδών μίας ομάδας 20 επιχειρήσεων σε εκατομμύρια δρχ. είναι:

Κέρδη	25	30	35	40	45	50
Επιχειρήσεις	3	3	6	5	2	1

Ζητείται να υπολογισθούν και να ερμηνευθούν:

- i) Ο μέσος αριθμητικός.
- ii) Η διάμεσος.
- iii) Το εύρος μεταβολής.
- iv) Η τυπική απόκλιση.
- v) Ο συντελεστής μεταβλητότητας.

Λύση

i) Υπολογισμός της μέσης τιμής.

Κατασκευάζουμε τον παρακάτω αριθμητικό πίνακα:

25	3	75
30	3	90
35	6	210
40	5	200
45	2	90
50	1	50
Αθροίσματα	20	715

Αντικαθιστώντας τα αθροίσματα που βρήκαμε $\Sigma n = 20$ και $\Sigma nx = 715$ στον τύπο υπολογισμού του μέσου αριθμητικού έχουμε:

$$\bar{x} = \frac{\sum vx}{\sum v} = \frac{715}{20}$$

και $\bar{x} = 35,75$

Δηλαδή, ο μέσος μηνιαίος μισθός αυτής της ομάδας στελεχών είναι 357.000 δραχ.

(1) x_i	(2) v_i	(3) N
25	3	3
30	3	6
35	6	$12 \sqrt{\frac{v}{2}} = 10$
40	5	17
45	2	19
50	1	20
	$v=20$	

ii) Υπολογισμός της διαμέσου.

Κατασκευάζουμε τον παρακάτω αριθμητικό πίνακα:

(1) x_i	(2) v_i	(3) N
25	3	3
30	3	6
35	6	$12 \cdot \frac{v}{2} = 10$
40	5	17
45	2	19
50	1	20
	$v=20$	

Όπως παρατηρούμε ο αριθμός $\frac{v}{2} = \frac{20}{2} = 10$ βρίσκεται μεταξύ των αθροιστικών συχνοτήτων (στήλη 3) 6 και 12 και συνεπώς η διάμεσος είναι ίση με 35, δηλαδή έχουμε: $\delta=35$

Το 50% λοιπόν των εργαζομένων έχει αποδοχές 350.000 (35 x 10.000) δρχ. και κάτω και το υπόλοιπο 50% 350.000 δρχ. και πάνω.

iii) Το εύρος μεταβολής ως γνωστό είναι η διαφορά μεταξύ των δύο ακραίων τιμών που παίρνει η μεταβλητή x και στη προκειμένη περίπτωση είναι:

Εύρος μεταβολής: $50 - 25 = 25$

Δηλαδή, η μισθολογική σχέση, αναλογία που υπάρχει μεταξύ του πλέον χαμηλόμισθου υπαλλήλου και του πλέον υψηλόμισθου είναι 1 προς 2 ή πιο απλά οι αποδοχές του πλέον υψηλόμισθου στελέχους είναι διπλάσιες από αυτές του πλέον χαμηλόμισθου στελέχους.

iv) Υπολογισμός της τυπικής απόκλισης.

Κατασκευάζουμε τον παρακάτω πίνακα αριθμητικών υπολογισμών.

x_i	v_i	$v_i x_i$	x_i^2	$v_i x_i^2$
25	3	75	625	1875
30	3	90	900	2700
35	6	210	1225	7350
40	5	200	1600	8000
45	2	90	2025	4050
50	1	50	2500	2500
	$v = 20$	$\Sigma vx = 715$		$\Sigma vx^2 = 26475$

Στον τύπο υπολογισμού της τυπικής απόκλισης αντικαθιστούμε τα αθροίσματα.

$$\sum v = v = 20$$

$$\sum v \cdot x = 715$$

$$\sum v \cdot x^2 = 26475$$

και έχουμε

$$s = \sqrt{\frac{\sum vx^2}{v} - \left(\frac{\sum vx}{v}\right)^2}$$

$$s = \sqrt{\frac{26475}{20} - \left(\frac{715}{20}\right)^2}$$

$$s = \sqrt{1323,75 - 1278,0625}$$

$$s \approx 6,76$$

Δηλαδή, η τετραγωνική ρίζα της μέσης τιμής των τετραγώνων των αποκλίσεων των μηνιαίων αποδοχών κάθε στελέχους από το μέσο μηνιαίο μισθό τους (357.500) είναι ίση με 67.600 δρχ. Φυσικά, όσο πιο κοντά στη μέση τιμή βρίσκονται οι παρατηρήσεις μας τόσο πιο μικρή είναι η s .

v) Υπολογισμός συντελεστή μεταβλητότητας.
Όπως γνωρίζουμε ο συντελεστής μεταβλητότητας CV ορίζεται ως ο λόγος της s δια του $\bar{x}$ επί 100, δηλαδή έχουμε:

$$CV = \frac{s}{\bar{x}} \cdot 100$$

$$CV = \frac{6,76}{35,75} \cdot 100$$

και $CV = 18,9\%$

Μπορούμε να πούμε ότι η σχετική διασπορά είναι ικανοποιητική, αφού το αποτέλεσμα που προέκυψε $CV = 18,9\%$, όπως αντιλαμβανόμαστε, δεν είναι ιδιαίτερα υψηλό και φυσικά τώρα μπορούμε να κάνουμε πιο ασφαλείς συγκρίσεις με άλλες κατανομές, σχετικά με τη διασπορά των τιμών τους.

3 Ο παρακάτω πίνακας δείχνει τις μηνιαίες αποδοχές μιας ομάδας 30 υπαλλήλων σε μια επιχείρηση σε δεκάδες χιλιάδες δρχ.

Μηνιαίες Αποδοχές	Αριθμός Υπαλλήλων
20-25	5
25-30	15
30-35	5
35-40	3
40-45	1
45-50	1
Σύνολο	30

Ζητείται:

- i) Η μέση τιμή του μηνιαίου μισθού.
- ii) Ποιος είναι ο μηνιαίος μισθός κάτω από τον οποίο βρίσκεται το 50% των αποδοχών των εργαζομένων.
- iii) Να υπολογισθεί ο συντελεστής μεταβλητότητας.

Λύση

i) Εύρεση της μέσης τιμής

Σχηματίζουμε τον παρακάτω αριθμητικό πίνακα:

Κλάσεις	v_i	x_i	$v_i \cdot x_i$	$v_i \cdot x_i^2$
[20-25)	5	22,5	506,25	2531,25
[25-30)	15	27,5	756,25	11343,75
[30-35)	5	32,5	1056,25	5281,25
[35-40)	3	37,5	1406,25	4218,75
[40-45)	1	42,5	1806,25	1806,25
[45-50)	1	47,5	2256,25	2256,25
Αθροίσματα	30			27437,5

$$\text{και } \bar{x} = \frac{\sum vx}{\sum v} = \frac{890}{30} = 29,666$$

$$\text{άρα } \bar{x} = 29,666$$

Δηλαδή η μέση τιμή του μηνιαίου μισθού είναι περίπου 296.000 δρχ.

ii) Εύρεση της διαμέσου.

Σχηματίζουμε τον παρακάτω αριθμητικό πίνακα:

Κλάσεις	v_i	N_i
[20-25)	5	5
[25-30)	15	20
[30-35)	5	25
[35-40)	3	28
[40-45)	1	29
[45-50)	1	30
Αθροίσματα	$v=30$	

Όπως βλέπουμε, παρατηρώντας τη στήλη των αθροιστικών συχνοτήτων ο αριθμός $\frac{v}{2} = \frac{30}{2} = 15$, βρίσκεται μεταξύ των αθροιστικών συχνοτήτων 5 και 20 συνεπώς η διάμεσος τιμή των μισθών εντοπίζεται στην κλάση [25 - 30), συνεπώς έχουμε:

$$\alpha_{i-1} = 25 \quad , \quad v_i = 15$$

$$N_{i-1} = 5 \quad , \quad c = 5 \quad \text{και} \quad \delta = \alpha_{i-1} + \frac{c}{v_i} - \left(\frac{v}{2} - N_{i-1} \right)$$

$$\delta = 25 + \frac{5}{15} - (15 - 5)$$

$$\delta = 28,333$$

Δηλαδή, η διάμεσος του μισθού είναι περίπου 283.000 δραχμές.

iii) Εύρεση του συντελεστή μεταβλητότητας.

Κατ' αρχάς υπολογίζουμε την τυπική απόκλιση s .
Για το σκοπό αυτό συνθέτουμε τον παρακάτω πίνακα:

Κλάσεις	v_i	x_i	x_i^2	$v_i \cdot x_i^2$
[20-25)	5	22,5	506,25	2531,25
[25-30)	15	27,5	756,25	11343,75
[30-35)	5	32,5	1056,25	5281,25
[35-40)	3	37,5	1406,25	4218,75
[40-45)	1	42,5	1806,25	1806,25
[45-50)	1	47,5	2256,25	2256,25
Αθροίσματα	30			27437,5

Από τον τύπο της διακύμανσης έχουμε:

$$s = \frac{27437,5}{30} - 29,666^2 = 914,58 - 880,071 = 34,50$$

$$\text{άρα } s = 5,87$$

Αντικαθιστώντας στον τύπο εύρεσης του συντελεστή μεταβλητότητας όπου

$$CV = \frac{s}{\bar{x}} 100$$

$$CV = \frac{5,87}{29,666} 100$$

$$\text{και } CV = 19,79\%$$

Σχετικά με τη σύγκριση του CV της παρούσας άσκησης με το CV της προηγούμενης, παρατηρούμε ότι οι αμοιβές των υπαλλήλων της παρούσας άσκησης παρουσιάζουν μια σχετικά μεγαλύτερη διασπορά, όχι σημαντική βέβαια, καθόσον εδώ έχουμε $CV = 19,79$ ενώ ο άλλος ήταν $CV = 19,16$. Εκείνο όμως που είναι αξιοσημείωτο, είναι ότι συμβαίνει το αντίθετο με τις τυπικές αποκλίσεις τους, όπου στην παρούσα άσκηση έχουμε $s = 5,871$ ενώ στην προηγούμενη $s = 6,85$.

- 4 Σε μια βιομηχανία αυτοκινήτων ο αριθμός των εργαζομένων το έτος 1995 ήταν κατανεμημένος σε ομάδες ηλικιών όπως φαίνεται στον παρακάτω πίνακα. Υπολογίστε τη μέση τιμή, τη διάμεσο, τη διακύμανση και την τυπική απόκλιση των ηλικιών των εργαζομένων σ' αυτήν τη βιομηχανία. Εκτιμήστε το ποσοστό των υπαλλήλων των οποίων οι ηλικίες βρίσκονται σε απόσταση μιας τυπικής απόκλισης γύρω από το μέσο (δηλαδή σε διάστημα $\bar{x} - s$, $\bar{x} + s$).

Τι παρατηρείτε;

Εξηγήστε τον ισχυρισμό σας με κατάλληλο διάγραμμα.

Ηλικία	Αριθμός εργαζομένων
15-20	6600
20-25	6500
25-30	5600
30-35	5000
35-40	4200
40-45	3700
45-50	3500
50-55	3000
55-60	2400
60-65	2200
Σύνολο	42700

Λύση

Ηλικία	Αριθμός εργαζομένων v_i	x_i'	$v_i x_i'$	$x_i'^2$	Αθροιστική Συχνότητα N
[15-20)	6600	17,5	115500	2021250	6650
[20-25)	6500	22,5	146250	3290625	13150
[25-30)	5600	27,5	154000	4235000	18700
[30-35)	5000	32,5	162500	5281250	23700
[35-40)	4200	37,5	157500	5906250	27900
[40-45)	3700	42,5	157250	6683125	31600
[45-50)	3500	47,5	166250	7896875	35100
[50-55)	3000	52,5	157500	8268750	38100
[55-60)	2400	57,5	138000	7935000	40500
[60-65)	2200	62,5	137500	8593750	42700
Σύνολο	42700		1492250	60111875	

Για να υπολογίσουμε τη μέση τιμή σε ομαδοποιημένα δεδομένα χρησιμοποιούμε τον τύπο:

$$\bar{x} = \frac{\sum_i^k v_i x_i'}{\sum_i v_i} = \frac{1.492.250}{42.700} \cong 34,95$$

Υπολογισμός διαμέσου

Επειδή η μεταβλητή (ηλικία) είναι συνεχής, η θέση της διαμέσου θα αντιστοιχεί στην $\frac{v}{2} = \frac{42.700}{2} = 21.350$ παρατήρηση η οποία περιέχεται στην κλάση [30, 35), όπως παρατηρούμε από την κατανομή αθροιστικών συχνοτήτων. Στην κλάση αυτή κατανέμονται ομοιόμορφα 5.000 παρατηρήσεις άρα η κάθε μια καταλαμβάνει πλάτος $\frac{5}{5000} = \frac{1}{1000}$.

Άρα οι $21.350 - 18.700 = 2.650$ παρατηρήσεις θα καταλάμβαναν πλάτος $2650 \cdot \frac{1}{1000} = 2,65$ δηλαδή η τιμή της διαμέσου θα είναι:

$$\delta = 30 + 2,65 = 32,65$$

Εναλλακτικά η τιμή της διαμέσου θα μπορούσε να υπολογιστεί αντικαθιστώντας τον τύπο.

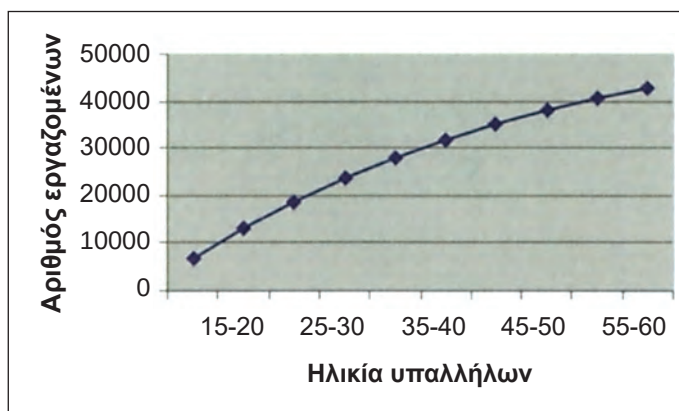
Υπολογισμός διακύμανσης

Στον τύπο $s^2 = \frac{\sum v_i x^2}{\sum v_i} - \bar{x}^2$ αντικαθιστούμε τα στοιχεία του πίνακα.

$$\text{και } s^2 = \frac{60.111.875}{42700} - (34,95)^2 = 186,27$$

και η τυπική απόκλιση $s = 13,65$

Από την αθροιστική συχνότητα θα εκτιμήσουμε το ποσοστό των υπαλλήλων των οποίων οι ηλικίες βρίσκονται σε διάστημα $(34,95 - 13,65, 34,95 + 13,65) = (21,3, 48,59)$



Παρατηρούμε ότι το ποσοστό των υπαλλήλων είναι 67%.

- 4** Οι ετήσιες πωλήσεις (X) των 15 καταστημάτων - παραρτημάτων (αλυσίδα καταστημάτων) που έχει μεγάλη εμπορική επιχείρηση δίνονται σε εκατ. δρχ. από τον παρακάτω πίνακα:

x_i	50	60	70	80	90	100	110
v_i	1	2	4	4	2	1	1

Ζητείται να υπολογισθούν:

- i) Η μέση τιμή των πωλήσεων των 15 καταστημάτων.
- ii) Το εύρος μεταβολής.
- iii) Η διακύμανση και η τυπική απόκλιση.
- iv) Ο συντελεστής μεταβλητότητας.

Μία άλλη επιχείρηση παρουσίασε μέση τιμή πωλήσεων 70 εκατ. δρχ. και τυπική απόκλιση 17 με άλλη επιχείρηση - αλυσίδας καταστημάτων - του ίδιου εμπορικού κλάδου. Να συγκρίνετε και να σχολιάσετε τη δραστηριότητα των δύο επιχειρήσεων.

Λύση

- i) Υπολογισμός της μέσης τιμής
Κατασκευάζουμε τον πίνακα:

x_i	v_i	$v_i x_i$
50	1	50
60	2	120
70	4	280
80	4	320
90	2	180
100	1	100
110	1	110
	15	1160

$$\text{και } \bar{x} = \frac{\sum v_i x_i}{v} = \frac{1160}{15} \approx 77,333$$

Συνεπώς η μέση τιμή των πωλήσεων των 15 καταστημάτων είναι 77.330.000 δρχ. περίπου.

ii) Το εύρος μεταβολής, ως γνωστό, είναι η διαφορά μεταξύ των δυο ακραίων τιμών της x , δηλαδή έχουμε:

Εύρος μεταβολής: $110 - 50 = 60$

Δηλαδή το εύρος μεταβολής είναι ίσο με 60.000.000 δρχ.

ii) Υπολογισμός διακύμανσης και τυπικής απόκλισης.

Συνθέτουμε τον παρακάτω πίνακα αριθμητικών υπολογισμών:

x_i	v_i	$v_i x_i$	x_i^2	$v_i \cdot x_i^2$
50	1	50	2.500	2.500
60	2	120	3.600	7.200
70	4	280	4.900	19.600
80	4	320	6.400	25.600
90	2	180	8.100	16.200
100	1	100	10.000	10.000
110	1	110	12.100	12.100
Σύνολο	15	1.160		93.200

$$\text{και } s^2 = \frac{\sum v_i x_i^2}{v} - \left(\frac{\sum v_i x_i}{v} \right)^2$$

$$s^2 = \frac{93.200}{15} - \left(\frac{1.160}{15} \right)^2$$

$$s^2 = 6.213,333 - 77,333^2$$

$$s^2 \approx 233$$

Η δε τυπική απόκλιση θα είναι:

$$s = \sqrt{s^2} = \sqrt{233} = 15,264$$

$$\text{άρα : } s \approx 15,264$$

iv) Υπολογισμός του συντελεστή μεταβλητότητας

$$CV = \frac{s}{\bar{x}} \cdot 100$$

$$CV = \frac{15,264}{77,333} 100 = 19,738$$

$$\text{και } CV = 19,738\%$$

Για να προχωρήσουμε σε μια ασφαλή σύγκριση με άλλη επιχείρηση του ίδιου εμπορικού κλάδου είναι προφανές ότι απαιτείται ο υπολογισμός του CV της άλλης επιχείρησης έτσι θα έχουμε:

$$CV = \frac{s}{\bar{x}} \cdot 100$$

$$CV = \frac{17}{70} 100 = 24,286$$

$$\text{και } CV = 24,286\%$$

Σχετικά λοιπόν με τη διασπορά των πωλήσεων αυτών των δυο επιχειρήσεων μπορούμε να πούμε ότι οι πωλήσεις της επιχείρησης που εξετάσαμε παρουσιάζουν μικρότερη σχετικά διασπορά και καλύτερη ομοιογένεια σε σχέση με τις πωλήσεις της άλλης επιχείρησης, καθόσον έχουμε μικρότερο συντελεστή μεταβλητότητας ($19,738 < 24,286$).

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_1 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v} = \frac{1}{v} [(x_1^2 - 2x_1\bar{x}) + \dots + (x_v^2 - 2x_v\bar{x})]$$

ΦΥΛΛΟ ΑΥΤΟΕΛΕΓΧΟΥ ΣΤΑ ΜΕΤΡΑ ΘΕΣΕΩΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ

- 1** Η τιμή της μεταβλητής, η οποία χωρίζει το σύνολο των παρατηρήσεών μας σε δυο ισοπληθείς ομάδες ονομάζεται:
- α) Μέσος αριθμητικός.
 - ☐ β) Διάμεσος.
 - γ) Πρώτο τεταρτημόριο.
 - δ) Επικρατούσα τιμή.
- 2** Η τιμή που παρουσιάζεται περισσότερες φορές σε μια ομάδα στατιστικών στοιχείων, ονομάζεται:
- α) Μέσος αριθμητικός.
 - β) Διάμεσος.
 - ☐ γ) Επικρατούσα τιμή.
 - δ) Πρώτο τεταρτημόριο.
- 3** Το δεύτερο τεταρτημόριο είναι:
- α) Μέσος αριθμητικός.
 - ☐ β) Η διάμεσος.
 - γ) Η επικρατούσα τιμή.
 - δ) Τίποτα από τα παραπάνω.
- 4** Ο μέσος αριθμητικός και η διάμεσος:
- α) Έχουν πάντα την ίδια τιμή.
 - β) Ποτέ δεν έχουν την ίδια τιμή.
 - ☐ γ) Μπορεί να έχουν είτε την ίδια είτε διαφορετική τιμή.
- 5** Το πηλίκο του αθροίσματος των τιμών που παίρνει μια μεταβλητή x δια του συνόλου των τιμών της ονομάζεται:
- α) Επικρατούσα τιμή.
 - β) Διάμεσος.
 - ☐ γ) Μέσος αριθμητικός.
 - δ) Πρώτο τεταρτημόριο.

- 6** Στις ίδιες μονάδες που εκφράζονται οι τιμές της μεταβλητής x εκφράζεται και:
- ☐ α) Η τυπική απόκλιση.
 - β) Η διακύμανση.
 - γ) Ο συντελεστής μεταβλητότητας.
- 7** Δεν προσφέρεται για αλγεβρικούς υπολογισμούς:
- α) Η διακύμανση.
 - β) Η τυπική απόκλιση.
 - ☐ γ) Η μέση απόκλιση.
 - δ) Ο συντελεστής μεταβλητότητας.
- 8** Τα μέτρα θέσης ανήκουν:
- ☐ α) Η διάμεσος και ο μέσος αριθμητικός.
 - β) Η διάμεσος και η διακύμανση.
 - γ) Η διακύμανση και τυπική απόκλιση.
 - δ) Το εύρος μεταβολής και η επικρατούσα τιμή.
- 9** Αν όλες οι τιμές μιας μεταβλητής πολλαπλασιασθούν επί μια σταθερή ποσότητα λ , τότε:
- α) Η διακύμανση πολλαπλασιάζεται επί λ .
 - ☐ β) Η διακύμανση πολλαπλασιάζεται επί λ^2 .
 - γ) Η διακύμανση παραμένει αμετάβλητη.
- 10** Η τυπική απόκλιση είναι η τετραγωνική ρίζα:
- α) Της διαμέσου.
 - ☐ β) Της διακύμανσης.
 - γ) Του μέσου αριθμητικού.
- 11** Η διαφορά μεταξύ της μεγαλύτερης και της μικρότερης τιμής των δεδομένων μας δίνεται:
- α) Από την επικρατούσα τιμή.
 - ☐ β) Από το εύρος μεταβολής.
 - γ) Από το συντελεστή μεταβλητότητας.
 - δ) Από την τυπική απόκλιση.

- 12** Ο συντελεστής μεταβλητότητας δίνεται από το πηλίκο:
- α) Της διακύμανσης δια του μέσου αριθμητικού.
 - β) Της διακύμανσης δια της τυπικής απόκλισης.
 - ☐ γ) Της τυπικής απόκλισης δια του μέσου αριθμητικού.
 - δ) Από την τυπική απόκλιση.
- 13** Από τα παρακάτω στατιστικά μέτρα καθαρός αριθμός είναι:
- α) Η διακύμανση.
 - β) Η τυπική απόκλιση.
 - γ) Ο μέσος αριθμητικός.
 - ☐ δ) Ο συντελεστής μεταβλητότητας.



$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_1 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v} = \frac{1}{v} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + \dots + (x_v^2 - 2x_v\bar{x} + \bar{x}^2)]$$

ΥΠΟΛΟΓΙΣΜΟΣ ΜΕΤΡΩΝ ΘΕΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ

1 Με χρήση υπολογιστή γραφικών (graphics calculator)

❖ Εισαγωγή δεδομένων.

AC/ON MENU

Εισάγουμε τα δεδομένα του πίνακα Δ. 1 όπως ακριβώς στις δραστηριότητες του κεφαλαίου 3.

❖ Επεξεργασία δεδομένων

Επιλέγουμε F2 (CALC) και στη συνέχεια FI(IVAR)

☐ HIST1

```

1-Variable
 $\bar{x}$  = 4.19
 $\Sigma x$  = 125.7
 $\Sigma x^2$  = 559.89
 $s^2$  = 1.05209315
 $s$  = 1.07007895
n = 30
1VAR 2VAR REG SET

```

Παρατηρούμε ότι στην οθόνη έχουμε τη μέση τιμή $\bar{x}=4.19$ καθώς και την τυπική απόκλιση $s = 1,05209315$

☐ HIST2

```

1-Variable
n      =30
minX   =1.8
Q1     =3.3
Med    =4.45
Q3     =4.9
 $\bar{x}$ -xσn=3.13790684
1VAR  2VAR  REG  SET

```

Στην οθόνη έχουμε τις τιμές των τεταρτημορίων Q1 και Q3, την ελάχιστη τιμή των δεδομένων(min X) καθώς και την τιμή της διαμέσου (Med)

☐ HIST3

```

1-Variable
Med    =4.45
Q3     =4.9
 $\bar{x}$ -xσn=3.13790684
 $\bar{x}$ +xσn=5.24209315
maxX   =6.2
Mod    =4.9
1VAR  2VAR  REG  SET

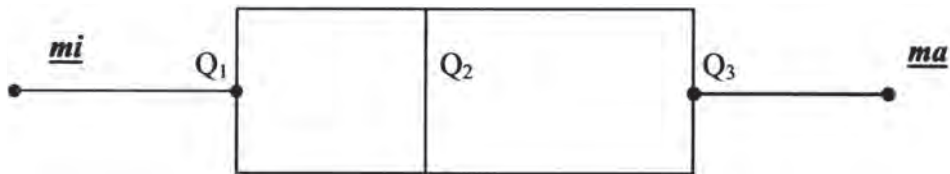
```

Στην τρίτη οθόνη εμφανίζεται η μέγιστη τιμή των δεδομένων(max X) καθώς και η επικρατούσα τιμή (Mod)

EXIT
EXIT

Ένα επίσης πολύ σημαντικό γράφημα μιας κατανομής, το οποίο όμως απαιτεί χρήση Η/Υ, είναι το *θηκόγραμμα*.

Το θηκόγραμμα αποτελεί γραφικό τρόπο παρουσίασης πέντε περιληπτικών μέτρων μιας κατανομής ομαδοποιημένων δεδομένων, με συνδυασμό των οποίων είναι δυνατή η άντληση περισσότερων πληροφοριών από αυτήν που περιέχεται στα πέντε αυτά μέτρα



Τυπική μορφή θηκογράμματος

Όπου $\min\{x_i\}$, $\max\{x_i\}$ η ελάχιστη και η μέγιστη τιμή του συνόλου των ομαδοποιημένων δεδομένων και Q_1 , Q_2 , Q_3 το πρώτο, δεύτερο και τρίτο τεταρτημόριο αντιστοίχως.

Με τη βοήθεια των *graphics calculator* μπορούμε εύκολα να κατασκευάσουμε το θηκόγραμμα μιας κατανομής.

Για παράδειγμα, θα χρησιμοποιήσουμε τις κλάσεις των δεδομένων του πίνακα Δ. 1 και θα ακολουθήσουμε την παρακάτω διαδικασία.

F1(GRAPH)

F6 (SET)



F6(>)

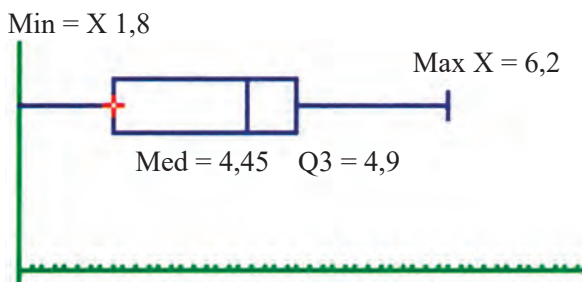
F2(BOX)

EXIT

F1(GRAPH 1)

SHIFT F1

▶▶▶ HIST 4



EXIT

EXIT

2 Με χρήση Η/Υ (πρόγραμμα Microsoft EXCEL)

❖ Εισαγωγή δεδομένων

Από το MENU επιλέγουμε το πρόγραμμα Microsoft EXCEL
Στο Βιβλίο 1 ,Φύλλο 1 και στη στήλη Α πληκτρολογούμε τα δεδομένα του πίνακα Δ. 1

❖ Επεξεργασία δεδομένων.

Από την γραμμή εργαλείων (toolbar) επιλέγουμε την “επικόλληση συνάρτησης” (f_x).

Στο παράθυρο που ανοίγει έχουμε δύο στήλες:
“Κατηγορία συνάρτησης” απ’ όπου επιλέγουμε “στατιστικές” και
“Όνομα συνάρτησης” απ’ όπου μπορούμε να επιλέξουμε:

- AVERAGE - μέση τιμή
- MEDIAN - διάμεσος
- MAX - μέγιστη τιμή
- MIN - ελάχιστη τιμή
- AVEDEV - μέση απόλυτη απόκλιση
- VAR - απόκλιση
- STDEV - τυπική απόκλιση
- MODE - επικρατούσα τιμή
- QUARTILE - τεταρτημόριο

Αν επιλέξουμε τον υπολογισμό τεταρτημορίων (QUARTILE), τότε στην επιλογή ARRAY μαυρίζουμε την στήλη Α και τοποθετούμε το τεταρτημόριο που θέλουμε στην επιλογή QUART.



ΚΕΦΑΛΑΙΟ 6°

Να κατανοήσουν οι μαθητές την αλληλεξάρτηση που υπάρχει μεταξύ δύο μεταβλητών και τον τρόπο με τον οποίο βρίσκεται ο βαθμός συσχέτισης, εφόσον υπάρχει συσχέτιση μεταξύ τους.

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1-12 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του έκτου κεφαλαίου. Οι μαθητές πρέπει να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στην τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \frac{1}{n} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + (x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \dots + (x_n^2 - 2x_n\bar{x} + \bar{x}^2)]$$

ΑΣΚΗΣΕΙΣ

Ομάδα Α'

- 1 Η αξία των επενδυμένων κεφαλαίων (y) και τα αντίστοιχα κέρδη (x) ομάδας 10 ομοειδών βιομηχανικών επιχειρήσεων περιέχονται στον παρακάτω πίνακα (τα δεδομένα εκφράζουν εκατομμύρια δρχ.).

Κεφάλαιο (x)	80	90	100	120	150	180	200	250	300	350
Κέρδη (y)	15	20	25	25	30	30	40	40	60	80

Ζητείται:

- i) Να βρεθεί ο συντελεστής γραμμικής συσχέτισης r , να σχολιασθεί το αποτέλεσμα και ii) να ερμηνευθεί η παράμετρος $\hat{\beta}$ της $\hat{y} = \hat{\alpha} + \hat{\beta}x$.

Λύση

Σχηματίζουμε τον παρακάτω πίνακα:

Κεφάλαια x	Κέρδη y	xy	x^2	y^2
80	15	1.200	6.400	225
90	20	1.800	8.100	400
100	25	2.500	10.000	625
100	25	3.000	14.400	625
150	30	4.500	22.500	900
180	30	5.400	32.400	900
200	40	8.000	40.000	1600
250	40	10.000	62.500	1600
300	60	18.000	90.000	3600
350	80	28.000	122.500	6400
1820	365	82.400	408.800	16875

- i) Υπολογισμός του συντελεστή παλινδρόμησης
Κατ' αρχήν υπολογίζουμε τα αθροίσματα των στηλών των δεδομένων:

$$\sum x = 1820 \text{ και } \sum y = 365$$

Καθώς και τα αθροίσματα:

$$\sum xy = 82.400 \text{ και } \sum y^2 = 16.875 \text{ και } \sum x^2 = 408.800$$

Τέλος αντικαθιστούμε τα αθροίσματα στο τύπο υπολογισμού του r

$$r = \frac{v \sum xy - \sum x \cdot \sum y}{\sqrt{v \sum x^2 - (\sum x)^2} \cdot \sqrt{v \sum y^2 - (\sum y)^2}} =$$

$$= \frac{10 \cdot 82400 - 1820 \cdot 365}{\sqrt{10 \cdot 408800 - (1820)^2} \cdot \sqrt{10 \cdot 16875 - (365)^2}} = 0,962$$

Παρατηρούμε ότι σ' αυτή την ομάδα των βιομηχανικών επενδύσεων υπάρχει μια θετική και πολύ έντονη συσχέτιση μεταξύ των επενδυμένων κεφαλαίων και των αντίστοιχων κερδών

- ii) Εύρεση της εξίσωσης παλινδρόμησης

Υπολογίζουμε το $\bar{x}$, $\bar{y}$

$$\bar{x} = \frac{\sum x}{v} = \frac{1820}{10} = 182$$

$$\bar{y} = \frac{\sum y}{v} = \frac{365}{10} = 36,5$$

και αντικαθιστούμε στον τύπο υπολογισμού του $\hat{\beta}$

$$\hat{\beta} = \frac{v \sum xy - (\sum x)(\sum y)}{v \sum x^2 - (\sum x)^2}$$

$$\hat{\beta} = \frac{10 \cdot 82.400 - 1820 \cdot 365}{10 \cdot 408.800 - (1820)^2}$$

$$\hat{\beta} = \frac{824.000 - 664.300}{4.088.000 - 3.312.400} = \frac{159.700}{775.600}$$

$$\hat{\beta} = 0,206 \approx 0,2$$

και

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\alpha} = 36,5 - 0,206 \cdot 182$$

$$\hat{\alpha} \approx -1$$

Άρα η ζητούμενη εξίσωση παλινδρόμησης θα είναι: $\hat{y} = -1 + 0,2x$

Για παράδειγμα, αν:

α) $x = 35$ και β) $x = 70$, βρίσκουμε:

α) για $x = 35$, $\hat{y} = -1 + 0,2 \cdot 35 = 6$

β) για $x = 70$, $\hat{y} = -1 + 0,2 \cdot 70 = 13$

Σε μια αύξηση του επενδυμένου κεφαλαίου κατά μια μονάδα, δηλαδή ένα εκατομμύριο δρχ. τα κέρδη αναμένεται να αυξηθούν κατά 206000δρχ., αφού το $\hat{\beta} = 0,206$ (206000)

2 Η από κοινού μελέτη των μεταβλητών X και Y μας έδωσε μετά από υπολογισμούς τα παρακάτω αθροίσματα:

$$\begin{array}{lll} \sum x = 30 & \sum y = 33 & \sum x^2 = 196 \\ \sum xy = 210 & \sum y^2 = 227 & N = 6 \end{array}$$

Ζητείται:

- Να βρεθεί ο συντελεστής γραμμικής συσχέτισης r .
- Να βρεθεί η εξίσωση γραμμικής παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta}x$.

Λύση

- Αντικαθιστούμε τα δεδομένα αθροίσματα στον τύπο υπολογισμού του συντελεστή συσχέτισης και λαμβάνουμε:

$$r = \frac{v \sum xy - (\sum x)(\sum y)}{\sqrt{v \sum x^2 - (\sum x)^2} \cdot \sqrt{v \sum y^2 - (\sum y)^2}}$$

$$r = \frac{6 \cdot 210 - 30 \cdot 33}{\sqrt{6 \cdot 196 - (30)^2} \cdot \sqrt{6 \cdot 227 - (33)^2}}$$

$$r = \frac{270}{16,613 \cdot 16,523}$$

$$r = \frac{270}{274,497} = 0,984$$

ii) Προσδιορισμός της εξίσωσης παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta}x$.

Υπολογισμός του $\hat{\beta}$

$$\hat{\beta} = \frac{v \sum xy - (\sum x) \cdot (\sum y)}{v \sum x^2 - (\sum x)^2}$$

$$\hat{\beta} = \frac{6 \cdot 210 - 30 \cdot 33}{6 \cdot 196 - (30)^2}$$

$$\hat{\beta} = \frac{270}{276}$$

$$\hat{\beta} = 0,978$$

Υπολογισμός του $\hat{\alpha}$

Αρχικά βρίσκουμε τους μέσους αριθμητικούς:

$$\bar{x} = \frac{\sum x}{v} = \frac{30}{6} = 5$$

$$\bar{y} = \frac{\sum y}{v} = \frac{33}{6} = 5,5$$

έχουμε λοιπόν ότι :

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\alpha} = 5,5 - 0,978 \times 5$$

$$\hat{\alpha} = 5,5 - 4,89$$

$$\text{και } \hat{\alpha} = 0,61$$

Άρα η ζητούμενη εξίσωση παλινδρόμησης θα είναι: $\hat{y} = 0,61 + 0,978x$

- 3** Ο παρακάτω πίνακας δεδομένων δείχνει τα ετήσια εισοδήματα x σε δεκάδες χιλιάδες δρχ. και τα αντίστοιχα ετήσια έξοδα τηλεφωνικών λογαριασμών y σε δεκάδες χιλιάδες δρχ. μίας ομάδας 8 νοικοκυριών:

x	300	350	400	450	500	550	600
y	9	8	10	13	13	16	15

Ζητείται:

- Να βρεθεί ο συντελεστής γραμμικής συσχέτισης r .
- Να υπολογισθούν οι παράμετροι $\hat{\alpha}$ και $\hat{\beta}$ της εξίσωσης παλινδρόμησης.

Λύση

- i)** Εύρεση του συντελεστή συσχέτισης r .

Κατασκευάζουμε κατά τα γνωστά τον παρακάτω πίνακα αριθμητικών υπολογισμών:

x	y	xy	x^2	y^2
300	9	2700	90000	81
350	8	2800	122500	64
400	10	4000	160000	100
450	13	5850	202500	169
500	13	6500	250000	169
550	16	8800	302500	256
600	15	9000	360000	225
3150	84	39650	1487500	1064

Αντικαθιστούμε στον τύπο υπολογισμού του συντελεστή συσχέτισης τα αθροίσματα:

$$\begin{aligned} \sum x &= 3150, & \sum y &= 84 \\ \sum xy &= 39650, & \sum y^2 &= 1064 \\ \sum x^2 &= 1487500 \end{aligned}$$

και έχουμε:

$$r = \frac{v \sum xy - (\sum x) \cdot (\sum y)}{\sqrt{v \sum x^2 - (\sum x)^2} \cdot \sqrt{v \sum y^2 - (\sum y)^2}}$$

$$r = \frac{7 \cdot 39650 - 3150 \cdot 84}{\sqrt{7 \cdot 1487500 - (3150)^2} \cdot \sqrt{7 \cdot 1064 - (84)^2}}$$

$$r = \frac{277550 - 264600}{\sqrt{10412500 - 9922500} \cdot \sqrt{7448 - 7056}}$$

$$r = \frac{12950}{\sqrt{490000} \cdot \sqrt{392}} \quad r = 0,934$$

ii) Προσδιορισμός των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$ της εξίσωσης παλινδρόμησης

$$\hat{y} = \hat{\alpha} + \hat{\beta}x.$$

Αρχικά βρίσκουμε τους μέσους αριθμητικούς:

$$\bar{x} = \frac{\sum x}{v} = \frac{3150}{7} = 450$$

$$\text{και } \bar{y} = \frac{\sum y}{v} = \frac{84}{7} = 12$$

Κατόπιν βρίσκουμε το $\hat{\beta}$

$$\hat{\beta} = \frac{v \sum xy - (\sum x)(\sum y)}{v \sum x^2 - (\sum x)^2}$$

$$\hat{\beta} = \frac{7 \cdot 39650 - 3150 \cdot 84}{7 \cdot 1487500 - (3150)^2}$$

$$\hat{\beta} = \frac{277550 - 264600}{10412500 - 9922500}$$

$$\hat{\beta} = 0,026$$

Τέλος βρίσκουμε το $\hat{\alpha}$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\alpha} = 12 - 0,026 \cdot 450$$

$$\hat{\alpha} = 12 - 11,7$$

και $\hat{\alpha} = 0,3$

Συνεπώς η εξίσωση παλινδρόμησης θα είναι: $\hat{y} = 0,3 + 0,026x$

- 4** Ο παρακάτω πίνακας αναφέρεται σε 10 ομοειδείς εμπορικές επιχειρήσεις και δείχνει τις ετήσιες πωλήσεις τους (y) και τις αντίστοιχες ετήσιες διαφημιστικές τους δαπάνες (x), (τα δεδομένα εκφράζονται σε εκατομμύρια δρχ.):

x	20	22	25	30	32	35	40	50	60	70
y	250	270	260	350	330	300	370	480	450	500

Ζητείται:

- Να βρεθεί ο συντελεστής γραμμικής συσχέτισης r και να σχολιασθεί το αποτέλεσμα.
- Να προσδιορισθεί η εξίσωση γραμμικής παλινδρόμησης.

Λύση

Κατασκευάζουμε τον παρακάτω αριθμητικό πίνακα:

x	y	xy	x^2	y^2
20	250	5000	400	62500
22	270	5940	484	72900
25	260	6500	625	67600
30	350	10500	900	122500
32	330	10560	1024	108900
35	300	10500	1225	90000
40	370	14800	1600	136900
50	480	24000	2500	230400
60	450	27000	3600	202500
70	500	35000	4900	250000
384	3560	149800	17258	1344200

Αντικαθιστούμε στον τύπο υπολογισμού του συντελεστή συσχέτισης τα αθροίσματα:

$$\begin{aligned}\sum x &= 384 & , & \quad \sum y = 3560 \\ \sum xy &= 149800 & , & \quad \sum y^2 = 1344200 \\ \sum x^2 &= 17258\end{aligned}$$

$$\begin{aligned}r &= \frac{v \sum xy - (\sum x)(\sum y)}{\sqrt{v \sum x^2 - (\sum x)^2} \cdot \sqrt{v \sum y^2 - (\sum y)^2}} \\ r &= \frac{10 \cdot 149800 - 384 \cdot 3560}{\sqrt{10 \cdot 17258 - (384)^2} \sqrt{10 \cdot 1344200 - (3560)^2}} \\ r &= \frac{1498000 - 1367040}{\sqrt{172580 - 147456} \cdot \sqrt{13442000 - 12673600}} \\ r &= \frac{130960}{158,506 \cdot 876,584} = 0,943\end{aligned}$$

άρα: $r = 0,943$

Από το παραπάνω αποτέλεσμα που βρήκαμε ($\rho = 0,943$) συμπεραίνουμε, ότι υπάρχει θετική συσχέτιση μεταξύ των διαφημιστικών δαπανών (x) και των πωλήσεων (y). Επίσης παρατηρούμε ότι αυτή η γραμμική συσχέτιση μεταξύ διαφημιστικών δαπανών και πωλήσεων είναι έντονη, γιατί το ρ εκτός από θετικό είναι πολύ κοντά στη μονάδα.

- ii)** Προσδιορισμός της εξίσωσης γραμμικής παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta}x$
Υπολογισμός της παραμέτρου $\hat{\beta}$:

$$\begin{aligned}\hat{\beta} &= \frac{v \sum xy - (\sum x)(\sum y)}{v \sum x^2 - (\sum x)^2} \\ \hat{\beta} &= \frac{10 \cdot 149800 - 384 \cdot 3560}{10 \cdot 17258 - (384)^2} \\ \hat{\beta} &= \frac{1498000 - 1367040}{172580 - 147456} = \frac{130960}{25124} \\ \text{και } \hat{\beta} &= 5,213\end{aligned}$$

Υπολογισμός της παραμέτρου $\hat{\alpha}$.

Αρχικά βρίσκουμε τους μέσους αριθμητικούς:

$$\bar{x} = \frac{\sum x}{v} = \frac{384}{10} = 38,4$$

$$\text{και } \bar{y} = \frac{\sum y}{v} = \frac{3560}{10} = 356$$

Συνεπώς έχουμε:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\alpha} = 356 - 5,213 \cdot 38,4$$

$$\hat{\alpha} = 356 - 200,1792 = 155,82$$

Δηλαδή η ζητούμενη εξίσωση παλινδρόμησης θα είναι:

$$\hat{y} = 155,82 + 5,213x$$

■ Ερμηνεία των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$ της εξίσωσης παλινδρόμησης.

Κατ' αρχήν η παράμετρος $\hat{\alpha} = 155,82$ μας δηλώνει ότι το ύψος των πωλήσεων μιας επιχείρησης, η οποία δεν προβαίνει σε καμιά διαφημιστική δαπάνη, θα ήταν 155,82 εκατ. δρχ. δηλαδή 155.820.000 - περίπου 156.000.000 δρχ.

Η παράμετρος $\hat{\beta}$ μας δηλώνει ότι σ' αυτήν την ομάδα των επιχειρήσεων με αύξηση των διαφημιστικών δαπανών κατά μία μονάδα, δηλαδή κατά ένα εκατομμύριο δρχ., αναμένεται να σημειωθεί αύξηση των πωλήσεων κατά 5,213 εκατ. δρχ. (5.213.000 δρχ.).

5 Ο παρακάτω πίνακας μας δείχνει τη μέση ετήσια κατά κεφαλή κατανάλωση y σε κιλά ενός αγαθού σε σχέση με το κατά κεφαλή εισόδημα x σε δεκάδες χιλιάδες δρχ. μιας ομάδας ατόμων:

x	65	68	70	73	75	78	85	90
y	40	40	41	43	45	48	50	50

Ζητείται:

- i) Να βρεθεί και να σχολιασθεί η τιμή του συντελεστή γραμμικής συσχέτισης r .
- ii) Να προσδιορισθεί η εξίσωση γραμμικής παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta}x$.
- iii) Να ερμηνευθούν οι τιμές των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$ της εξίσωσης παλινδρόμησης.
- iv) Να βρεθούν οι διαφορές μεταξύ των πραγματικών τιμών της εξαρτημένης μεταβλητής y και των αντίστοιχων τιμών που μας δίνει η $\hat{y} = \hat{\alpha} + \hat{\beta}x$

Λύση

i) Υπολογισμός του συντελεστή συσχέτισης.

Κατά τα γνωστά προχωρούμε στην κατασκευή του παρακάτω αριθμητικού πίνακα:

x	y	xy	x ²	y ²
65	40	2600	4225	1600
68	40	2720	4624	1600
70	41	2870	4900	1681
73	43	3139	5329	1849
75	45	3375	5625	2025
78	48	3744	6084	2304
85	50	4250	7225	2500
90	50	4500	8100	2500
604	357	27198	46112	16059

Αντικαθιστούμε στον τύπο υπολογισμού του συντελεστή συσχέτισης τα αθροίσματα:

$$\begin{aligned}\Sigma x &= 604, & \Sigma y &= 357 \\ \Sigma xy &= 27198, & \Sigma y^2 &= 16059 \\ \Sigma x^2 &= 46112\end{aligned}$$

$$r = \frac{n \Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n \Sigma x^2 - (\Sigma x)^2} \cdot \sqrt{n \Sigma y^2 - (\Sigma y)^2}}$$

$$r = \frac{8 \cdot 27198 - 604 \cdot 357}{\sqrt{8 \cdot 46112 - (604)^2} \cdot \sqrt{8 \cdot 16059 - (357)^2}}$$

$$r = \frac{217584 - 215628}{63,875 \cdot 31,984}$$

$$r = \frac{1956}{2042,978} = 0,957$$

άρα $r = 0,957$

Παρατηρούμε λοιπόν αφ' ενός ότι υπάρχει μια θετική συσχέτιση ανάμεσα στην κατανάλωση αυτού του αγαθού και στο εισόδημα αφ' ετέρου ότι αυτή η θετική συσχέτιση είναι πολύ έντονη, επειδή όπως βλέπουμε ο συντελεστής συσχέτισης είναι πολύ κοντά στη μονάδα, $\rho = 0,957$.

ii) Προσδιορισμός της εξίσωσης γραμμικής παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta}x$.

Υπολογισμός της παραμέτρου $\hat{\beta}$

$$\hat{\beta} = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

$$\hat{\beta} = \frac{8 \cdot 27198 - 604 \cdot 357}{8 \cdot 46112 - (604)^2}$$

$$\hat{\beta} = \frac{217584 - 215628}{368896 - 364816} = \frac{1956}{4080}$$

άρα $\hat{\beta} = 0,479$

Υπολογισμός της παραμέτρου $\hat{\alpha}$

Αρχικά βρίσκουμε τους μέσους αριθμητικούς $\bar{x}$ και $\bar{y}$:

$$\bar{x} = \frac{\sum x}{n} = \frac{604}{8} = 75,5$$

$$\bar{y} = \frac{\sum y}{n} = \frac{357}{8} = 44,625$$

και $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

$$\hat{\alpha} = 44,625 - 0,479 \cdot 75,5$$

$$\hat{\alpha} = 44,625 - 36,1645 = 8,46$$

Άρα η εξίσωση παλινδρόμησης θα είναι :

$$\hat{y} = 8,46 + 0,479x$$

iii) Ερμηνεία των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$ της εξίσωσης παλινδρόμησης.

Η παράμετρος $\hat{\beta} = 0,479$ δηλώνει ότι σε αύξηση του κατά κεφαλή εισοδήματος κατά μια μονάδα, δηλαδή στην προκειμένη περίπτωση κατά δέκα χιλιάδες δρχ., αναμένεται να σημειωθεί αύξηση στη μέση ετήσια κατανάλωση της τάξης του $0,479 \approx 0,5$ κιλού, δηλαδή περίπου μισό κιλό. Τώρα στην υποθετική περίπτωση όπου το x θα έπαιρνε τιμή ίση με το μηδέν, τότε η κατανάλωση θα ελαχιστοποιούνταν, φθάνοντας στο όριο των 8,46 κιλών δηλαδή, οκτώμισι κιλά περίπου ($\alpha = 8,46 \approx 8,5$ κιλά).

iv) Εύρεση των διαφορών $y - \hat{y}$

Κατ' αρχάς πηγαίνουμε στην εξίσωση παλινδρόμησης $\hat{y} = 8,46 + 0,479x$ και θέτουμε όπου x τις τιμές της μεταβλητής X που μας δίνει ο αρχικός πίνακας των δεδομένων μας και λαμβάνουμε έτσι τις αντίστοιχες τιμές.

$$\begin{aligned} \text{Για } x = 65, \quad \hat{y} &= 8,46 + 0,479 \cdot 65 \\ \text{και } \hat{y} &= 39,5965 \end{aligned}$$

$$\begin{aligned} \text{Για } x = 68 \quad \hat{y} &= 8,46 + 0,479 \cdot 68 \\ \text{και } \hat{y} &= 41,032 \end{aligned}$$

$$\begin{aligned} \text{Για } x = 70 \quad \hat{y} &= 8,46 + 0,479 \cdot 70 \\ \text{και } \hat{y} &= 41,99 \end{aligned}$$

$$\begin{aligned} \text{Για } x = 73 \quad \hat{y} &= 8,46 + 0,479 \cdot 73 \\ \text{και } \hat{y} &= 43,427 \end{aligned}$$

Για $x = 75$ $\hat{y} = 8,46 + 0,479 \cdot 75$
και $\hat{y} = 44,385$

Για $x = 78$ $\hat{y} = 8,46 + 0,479 \cdot 78$
και $\hat{y} = 45,822$

Για $x = 85$ $\hat{y} = 8,46 + 0,479 \cdot 85$
και $\hat{y} = 49,175$

Για $x = 90$ $\hat{y} = 8,46 + 0,479 \cdot 90$
και $\hat{y} = 51,57$

Στη συνέχεια από τις πραγματικές τιμές y αφαιρούμε τις αντίστοιχες τιμές $\hat{y}$ και κατασκευάζουμε την τελευταία στήλη του παρακάτω αριθμητικού πίνακα:

x	y	$\hat{y}$	$y - \hat{y}$
65	40	36,595	0,40
68	40	41,032	-1,03
70	41	41,990	-0,99
73	43	43,427	-0,43
75	45	44,385	0,62
78	48	45,822	2,18
85	50	49,175	0,83
90	50	51,570	-1,57

- 6 Πέντε άτομα μετέχουν σε πρόγραμμα ελάττωσης βάρους. Ο αριθμός των εβδομάδων που κάθε άτομο συμμετέχει στο πρόγραμμα (x), καθώς και τα κιλά που έχασε στο αντίστοιχο διάστημα (y) εμφανίζονται στον παρακάτω πίνακα:

Αριθμός Εβδομάδων (x)	3	2	1	4	5
Ελάττωση Βάρους σε κιλά (y)	3	2,5	2	4,5	5,5

Ζητείται:

- Να βρείτε την εξίσωση της ευθείας των ελαχίστων τετραγώνων.
- Πόσες εβδομάδες πρέπει κάποιος να συμμετέχει στο πρόγραμμα για να χάσει 10 κιλά;

Λύση

i)

x	y	x ²	y ²	xy
3	3	9	9	9
2	2,5	4	6,25	5
1	2	1	4	2
4	4,5	16	20,25	18
5	5,5	25	30,25	27,5
$\sum x=15$	$\sum y=17,5$	$\sum x^2=55$	$\sum y^2=69,75$	$\sum xy=61,5$

$$\hat{\beta} = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

$$\hat{\beta} = \frac{5 \cdot 61,5 - 15 \cdot 17,5}{5 \cdot 55 - (15)^2}$$

$$\hat{\beta} = \frac{307,5 - 262,5}{275 - 225} = \frac{45}{50}$$

και $\hat{\beta} = 0,9$

Αρχικά βρίσκουμε τους μέσους αριθμητικούς $\hat{y}$ και $\bar{x}$:

$$\bar{y} = \frac{\sum y}{v} = \frac{17,5}{5} = 3,5$$

$$\bar{x} = \frac{\sum x}{v} = \frac{15}{5} = 3$$

$$\text{και } \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\alpha} = 3,5 - 0,9 \cdot 3$$

$$\hat{\alpha} = 3,5 - 2,7$$

$$\hat{\alpha} = 0,8$$

Άρα η εξίσωση παλινδρόμησης θα είναι:

$$\hat{y} = 0,8 + 0,9x$$

ii) για $y = 10$ αντικαθιστούμε:

$$\hat{y} = 0,8 + 0,9x \text{ και έχουμε}$$

$$10 = 0,8 + 0,9x \Leftrightarrow 0,9x = 10 - 0,8 \Leftrightarrow 0,9x = 9,2 \Leftrightarrow$$

$$\Leftrightarrow x \approx 10,22 \text{ εβδομάδες}$$

Ομάδα Β'

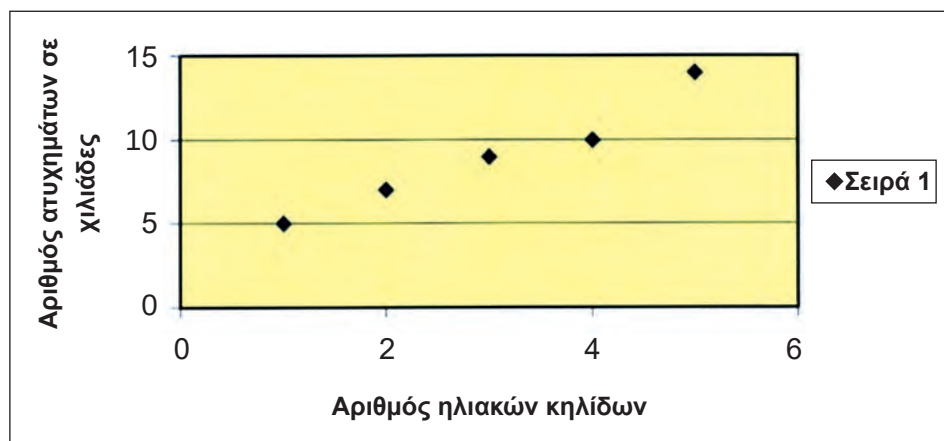
- 1 Ένας ερευνητής μελετά τον αριθμό των ηλιακών κηλίδων (x) και τον αριθμό των ατυχημάτων (y) για πέντε συνεχή χρόνια. Σύμφωνα με τη θεωρία του ο αριθμός των ηλιακών κηλίδων συνδέεται με κάποιο τρόπο με τον αριθμό των ατυχημάτων. Οι παρατηρήσεις του φαίνονται στον πίνακα που ακολουθεί:

Αριθμός ηλιακών κηλίδων (x)	3	1	2	5	4
Αριθμός ατυχημάτων σε χιλιάδες (y)	9	5	7	14	10

Ζητείται:

- Να κατασκευαστεί το διάγραμμα διασποράς των δύο μεταβλητών.
- Μπορεί στο διάγραμμα αυτό να προσαρμοστεί ευθεία γραμμή; Αν ναι, να προσδιοριστεί η εξίσωσή της.
- Να βρεθεί ο συντελεστής συσχέτισης. Τι παρατηρείτε;
- Αν δεν παρατηρηθούν ηλιακές κηλίδες, πόσα ατυχήματα προβλέπεται ότι θα συμβούν;

Λύση



ii) Το νέφος σημείων μας οδηγεί στην αναζήτηση γραμμικής σχέσης μεταξύ των μεταβλητών X και Y . Άρα θα υπολογίσουμε τις εκτιμήτριες ελαχίστων τετραγώνων $\hat{\alpha}$ και $\hat{\beta}$ και θα προσδιορίσουμε την ευθεία παλινδρόμησης.

Συμπληρώνουμε τον παρακάτω πίνακα:

x	y	x ²	y ²	xy
3	9	9	81	27
1	5	1	25	5
2	7	4	49	14
5	14	25	196	70
4	10	16	100	40
$\sum x=15$	$\sum y=45$	$\sum x^2=55$	$\sum y^2=451$	$\sum xy=156$

Υπολογίζουμε τη μέση τιμή των τιμών των μεταβλητών X και Y :

$$\bar{x} = \frac{\sum x}{v} = \frac{15}{5} = 3$$

$$\bar{y} = \frac{\sum y}{v} = \frac{45}{5} = 9$$

και υπολογίζουμε τις εκτιμήτριες των ελαχίστων τετραγώνων $\hat{\alpha}$ και $\hat{\beta}$ ως εξής:

$$\hat{\beta} = \frac{\frac{\sum xy}{v} - \bar{x} \cdot \bar{y}}{\frac{\sum x^2}{v} - \bar{x}^2} = \frac{\frac{156}{5} - 3 \cdot 9}{\frac{55}{5} - 9} = \frac{31,2 - 27}{11 - 9} = \frac{4,2}{2} = 2,1$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 9 - 2,1 \cdot 3 = 9 - 6,3 = 2,7$$

Άρα η ευθεία παλινδρόμησης είναι $\hat{y} = 2,7 + 2,1x$

- iii) Ο συντελεστής συσχέτισης μετρά το βαθμό εξάρτησης των δύο μεταβλητών. Στη συγκεκριμένη περίπτωση ο συντελεστής συσχέτισης είναι:

$$r = \frac{v \sum xy - (\sum x)(\sum y)}{\sqrt{v \sum x^2 - (\sum x)^2} \cdot \sqrt{v \sum y^2 - (\sum y)^2}}$$

$$r = \frac{5 \cdot 156 - 15 \cdot 45}{\sqrt{5 \cdot 55 - 15^2} \cdot \sqrt{5 \cdot 451 - 45^2}}$$

$$r = \frac{780 - 675}{\sqrt{275 - 225} \cdot \sqrt{2255 - 2025}}$$

$$\text{και } r = \frac{105}{\sqrt{50} \cdot \sqrt{230}} \quad \text{άρα } r = 0,98$$

Παρατηρούμε ότι η συσχέτιση μεταξύ του αριθμού των ηλιακών κηλίδων και του αριθμού των ατυχημάτων είναι πολύ ισχυρή και μάλιστα θετική. Δηλαδή, αν αυξηθεί ο αριθμός των ηλιακών κηλίδων, αυξάνεται και ο αριθμός των ατυχημάτων.

- iv) Αν δεν παρατηρηθούν ηλιακές κηλίδες τότε μπορούμε να προβλέψουμε ότι θα γίνουν 2.700 ατυχήματα (για $x = 0$, $\hat{y} = 2,7$ χιλιάδες ατυχήματα).

2 Μία εταιρεία έκανε μελέτη της ποσότητας καλαμποκιού που παράγεται από μια συγκεκριμένη καλλιεργήσιμη έκταση σε σχέση με την ποσότητα λιπάσματος που χρησιμοποιήθηκε για την παραγωγή. Τα αποτελέσματα φαίνονται στον πίνακα που ακολουθεί:

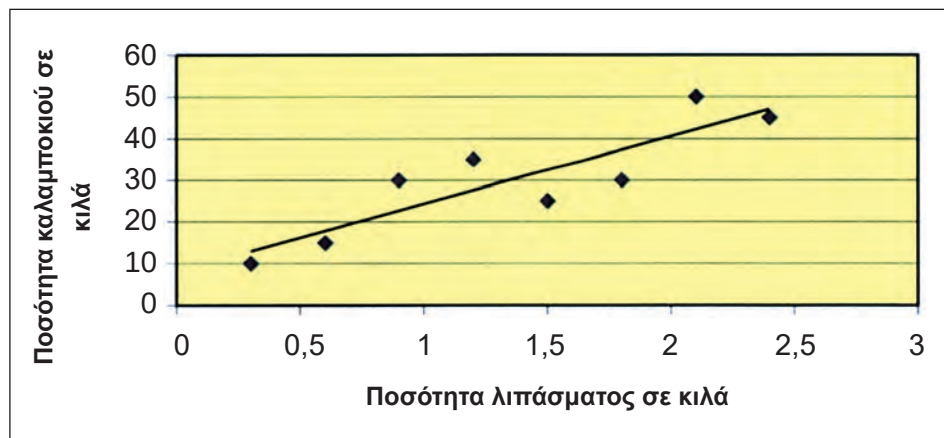
Ποσότητα λιπάσματος (x) σε κιλά	0,3	0,6	0,9	1,2	1,5	1,8	2,1	2,4
Ποσότητα καλαμποκιού (y) σε κιλά	10	15	30	35	25	30	50	45

Ζητείται:

- Να παρασταθεί γραφικά η εξαρτημένη μεταβλητή σαν συνάρτηση της ανεξάρτητης μεταβλητής.
- Να βρεθεί η ευθεία ελαχίστων τετραγώνων για τα συγκεκριμένα δεδομένα.
- Να βρεθεί η ποσότητα καλαμποκιού, αν δεν χρησιμοποιηθεί λίπασμα.
- Αν είχαν χρησιμοποιηθεί 3 κιλά λίπασμα, ποια θα είναι η καλύτερη εκτίμηση της ποσότητας καλαμποκιού που θα είχε παραχθεί;

Λύση

i)



ii) Συμπληρώνουμε τον παρακάτω πίνακα:

x	y	x ²	xy
0,3	10	0,09	3
0,6	15	0,36	9
0,9	30	0,81	27
1,2	35	1,44	42
1,5	25	2,25	37,5
1,8	30	3,24	54
2,1	50	4,41	105
2,4	45	5,76	108
10,8	240	18,36	385,5

Υπολογίζουμε τη μέση τιμή των τιμών των μεταβλητών X και Y:

$$\bar{x} = \frac{\sum x}{v} = \frac{10,8}{8} = 1,35$$

$$\bar{y} = \frac{\sum y}{v} = \frac{240}{8} = 30$$

και υπολογίζουμε τις εκτιμήτριες των ελαχίστων τετραγώνων $\hat{\alpha}$ και $\hat{\beta}$ ως εξής:

$$\hat{\beta} = \frac{8 \cdot 385,5 - 10,8 \cdot 240}{8 \cdot 18,36 - (10,8)^2} = \frac{3084 - 2592}{146,88 - 116,64} = \frac{492}{30,24} = 16,27$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} = 30 - 16,27 \cdot 1,35 = 30 - 21,97 = 8,03$$

άρα η ευθεία παλινδρόμησης είναι $\hat{y} = 8,03 + 16,27x$

iii) Αν δε χρησιμοποιηθεί λίπασμα η ποσότητα καλαμποκιού που αναμένεται να παραχθεί από αυτή την έκταση είναι 8,03 κιλά (γιατί για $x = 0$, $\hat{y} = 8,03$).

- iv) Αν είχαν χρησιμοποιηθεί 3 κιλά λίπασμα, τότε η αναμενόμενη παραγωγή καλαμποκιού θα ήταν:

$$\hat{y} = 8,03 + 16,27 \cdot 3 \quad \hat{y} = 56,84 \text{ kg}$$

- 3 Μία εταιρεία θέλει να προβλέψει την απόδοση των πωλητών που εκπαιδεύει. Στην αρχή της διμηνιαίας εκπαίδευσης 10 υποψήφιοι πωλητές απάντησαν σε έλεγχο δεξιοτήτων. Οι επιδόσεις τους παρουσιάζονται στον παρακάτω πίνακα. Η εταιρεία κράτησε αρχεία των πωλήσεων κάθε υπαλλήλου (μεταβλητή Y) σε ορισμένο χρονικό διάστημα.

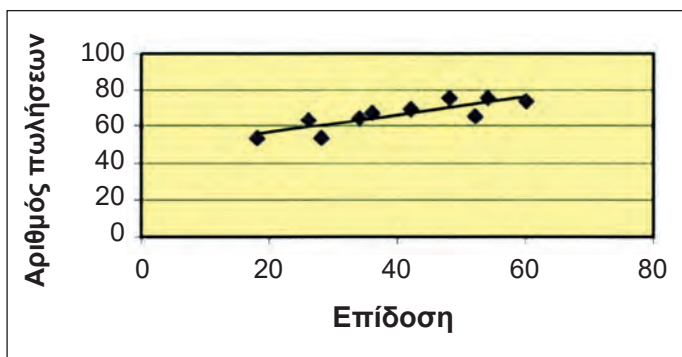
Επίδοση στον έλεγχο (x)	18	26	28	34	36	42	48	52	54	60
Αριθμός πωλήσεων (y)	54	64	54	62	68	70	76	66	76	74

Ζητείται:

- Να παρασταθούν γραφικά τα δεδομένα του πίνακα και να βρεθεί η ευθεία ελαχίστων τετραγώνων.
- Ποιος ο αριθμός των πωλήσεων που περιμένετε από 3 πωλητές των οποίων τα αποτελέσματα στον έλεγχο δεξιοτήτων ήταν 40, 50 και 70 αντίστοιχα;

Λύση

i)



Συμπληρώνουμε τον παρακάτω πίνακα:

x	y	x ²	xy
18	54	324	972
26	64	676	1664
28	54	784	1512
34	62	1156	2103
36	68	1296	2448
42	70	1764	2940
48	76	2304	3643
52	66	2704	3432
54	76	2916	4104
60	74	3600	4440
$\sum x=398$	$\sum y=664$	$\sum x^2=17524$	$\sum xy=27268$

Υπολογίζουμε τη μέση τιμή των τιμών των μεταβλητών X και Y:

$$\bar{x} = \frac{\sum x}{v} = \frac{398}{10} = 39,8$$

$$\bar{y} = \frac{\sum y}{v} = \frac{664}{10} = 66,4$$

και υπολογίζουμε τις εκτιμήτριες $\hat{\alpha}$ και $\hat{\beta}$ ως εξής:

$$\hat{\beta} = \frac{v \sum xy - (\sum x)(\sum y)}{v \sum x^2 - (\sum x)^2} = \frac{10 \cdot 27268 - 398 \cdot 664}{10 \cdot 17524 - (398)^2} = \frac{8408}{16836} \approx 0,5$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} = 66,4 - 0,5 \cdot 39,8 = 66,4 - 19,9 = 46,5$$

άρα η ευθεία παλινδρόμησης είναι $\hat{y} = 46,5 + 0,5x$

- ii) Αν τα αποτελέσματα στον έλεγχο δεξιοτήτων ήταν 40, 50, και 70, τότε ο αριθμός των πωλήσεων που αναμένεται από τους τρεις πωλητές θα ήταν αντίστοιχα:

$$x = 40 \quad \hat{y} = 46,5 + 0,5 \cdot 40 = 66,5$$

$$x = 50 \quad \hat{y} = 46,5 + 0,5 \cdot 50 = 71,5$$

$$x = 70 \quad \hat{y} = 46,5 + 0,5 \cdot 70 = 81,5$$

ΠΑΛΙΝΔΡΟΜΗΣΗ - ΣΥΣΧΕΤΙΣΗ ΦΥΛΛΟ ΑΥΤΟΕΛΕΓΧΟΥ

- 1 Στη γραμμική παλινδρόμηση η σχέση μεταξύ της ανεξάρτητης μεταβλητής της εξαρτημένης είναι αυτή της:
 - ☐ α) Ευθείας γραμμής.
 - ☐ β) Καμπύλης γραμμής.

- 2 Ο συντελεστής γραμμικής παλινδρόμησης $\hat{\beta}$ μας δείχνει:
 - ☐ α) Την ποσοστιαία μεταβολή που επέρχεται στην εξαρτημένη μεταβλητή από μια μεταβολή της ανεξάρτητης κατά 1%.
 - ☐ β) Την ποσοστιαία μεταβολή που επέρχεται στην εξαρτημένη μεταβλητή από μια μεταβολή της ανεξάρτητης κατά 10%.
 - ☐ γ) Τη μεταβολή που επέρχεται στην εξαρτημένη μεταβλητή από μια μοναδιαία μεταβολή της ανεξάρτητης μεταβλητής.

- 3 Ο συντελεστής γραμμικής παλινδρόμησης $\hat{\beta}$ έχει πρόσημο:
 - ☐ α) Πάντα θετικό.
 - ☐ β) Πάντα αρνητικό.
 - ☐ γ) Είτε θετικό είτε αρνητικό.

- 4 Στην εξίσωση παλινδρόμησης με $\hat{y}$.
 - ☐ α) Συμβολίζουμε τις πραγματικές τιμές της εξαρτημένης μεταβλητής.
 - ☐ β) Συμβολίζουμε τις “θεωρητικές” τιμές της εξαρτημένης μεταβλητής, που μας δίνει η εξίσωση παλινδρόμησης.
 - ☐ γ) Συμβολίζουμε και τις πραγματικές και τις θεωρητικές τιμές της εξαρτημένης μεταβλητής.
 - ☐ δ) Συμβολίζουμε τις τιμές της ανεξάρτητης μεταβλητής.

5 Η παράμετρος $\hat{\alpha}$ της γραμμικής παλινδρόμησης δίνεται από τον τύπο:

α) $\hat{\alpha} = y - \hat{\beta}x$

☐ β) $\hat{\alpha} = \mu_y - \hat{\beta}\mu_x$

γ) $\hat{\alpha} = \hat{y} - \hat{\beta}x$

δ) $\hat{\alpha} = \mu_y - \hat{\beta}x$

6 Αν στην εξίσωση γραμμικής παλινδρόμησης είναι το $x = 0$, τότε θα έχουμε:

α) $\hat{y} = \hat{\beta}$

☐ β) $\hat{y} = \hat{\alpha}$

γ) $\hat{y} = -\hat{\beta}$

7 Η παλινδρόμηση και η συσχέτιση χρησιμοποιούνται για τη μελέτη:

α) Μόνο μονομεταβλητών πληθυσμών.

β) Μόνο διμεταβλητών πληθυσμών.

γ) Μόνο πολυμεταβλητών πληθυσμών

☐ δ) Είτε διμεταβλητών είτε πολυμεταβλητών πληθυσμών.

8 Ο συντελεστής γραμμικής συσχέτισης παίρνει τιμές:

α) Από 0 έως 1.

β) Από - 1 έως 0.

☐ γ) Από - 1 έως 1.

δ) Από - ∞ έως + ∞ .

9 Ο συντελεστής γραμμικής συσχέτισης ρ και ο συντελεστής γραμμικής παλινδρόμησης $\hat{\beta}$ έχουν:

☐ α) Πάντοτε το ίδιο πρόσημο.

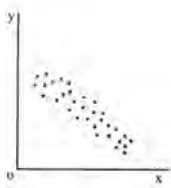
β) Πάντοτε διαφορετικό πρόσημο.

γ) Άλλοτε το ίδιο και άλλοτε διαφορετικό πρόσημο.

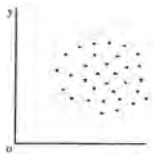
10 Η μεταξύ των μεταβλητών υπάρχουσα συσχέτιση είναι:



- ☐ α) Θετική
☐ β) Αρνητική
☐ γ) Δεν υπάρχει συσχέτιση



- ☐ α) Θετική
☐ β) Αρνητική
☐ γ) Δεν υπάρχει συσχέτιση



- ☐ α) Θετική
☐ β) Αρνητική
☐ γ) Δεν υπάρχει συσχέτιση.

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v} = \frac{1}{v} [(x_1^2 - 2x_1\bar{x}) + (x_2^2 - 2x_2\bar{x}) + \dots + (x_v^2 - 2x_v\bar{x})]$$

ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ ΚΕΦΑΛΑΙΟΥ 6

ΠΡΟΣΑΡΜΟΓΗ ΕΥΘΕΙΑΣ ΓΡΑΜΜΗΣ ΣΕ ΣΥΝΟΛΟ
ΠΕΙΡΑΜΑΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Για την επαλήθευση του νόμου των ιδανικών αερίων μετρήθηκε η πίεση (y) μιας συγκεκριμένης ποσότητας αερίου με την θερμοκρασία του (x). Τα πειραματικά δεδομένα παρουσιάζονται στον πίνακα Δ2 που ακολουθεί:

Θερμοκρασία ($^{\circ}\text{C}$)	20	41	60	79	98
Πίεση (Nt/cm^2)	11.0	12.0	12.5	13.5	14.0

Ζητείται η εξίσωση της ευθείας παλινδρόμησης για τα παραπάνω δεδομένα.

A Με χρήση υπολογιστή γραφικών (*graphics calculator*)

MENU

STAT (μαυρισμένο)

EXE

LIST4 | 1 20

| 2 41

|.....

LIST5 | 1 11

| 2 12

|.....

SHIFT F3 (V-WIN)

0 EXE

100 EXE

10 EXE

0 EXE

15 EXE

1 EXE

EXIT

FI(GRAPH)

F6 (SET)

F3(GRAPH3)

FI(SCAT)

F4(LIST4)

F5(LIST5)

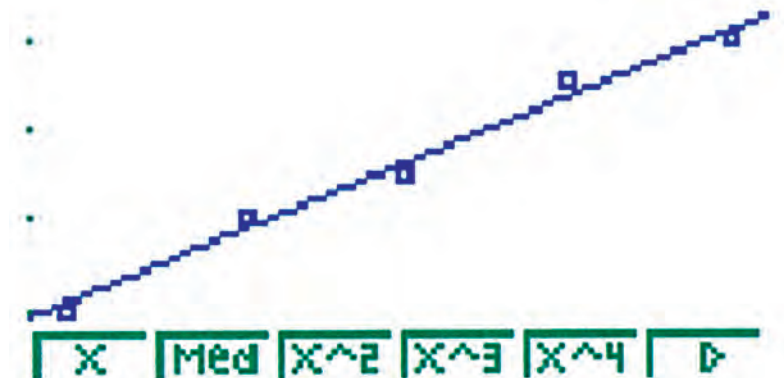
EXIT

F3(GRAPH3)



F1(X)

F6(DRAW)



EXIT

EXIT

MENU

SHIFT AC (OFF)

Σημ. Η εξίσωση της ευθείας είναι:

$$\text{Πίεση} = 0.039 (\text{Θερμοκρασία}) + 10.3$$

Δηλαδή $\hat{y} = 0.039x + 10.3$

B Με χρήση Η/Υ (πρόγραμμα Microsoft EXCEL)

Εισαγωγή δεδομένων

Από το MENU επιλέγουμε το πρόγραμμα Microsoft EXCEL
Στο Βιβλίο 1 ,Φύλλο 1 και στις στήλες A και B πληκτρολογούμε τα δεδομένα του πίνακα Δ.2

❖ Επεξεργασία δεδομένων.

Επιλέγουμε (μαυρίζουμε) τη στήλη A και B και από την γραμμή εργαλείων (toolbar) επιλέγουμε τον οδηγό γραφημάτων. Οι δυνατότητες που έχουμε είναι δύο:

Βασικοί τύποι –Προσαρμοσμένοι τύποι γραφημάτων.

Στο συγκεκριμένο παράδειγμα χρησιμοποιούμε τους “βασικούς τύπους γραφημάτων” και ακολουθούμε τα παρακάτω βήματα.

Βασικοί τύποι

Τύπος γραφήματος “Διασπορά ΧΥ”(1° γράφημα) -Επόμενο

Δεδομένα προέλευσης

Περιοχή δεδομένων-Σειρά σε “στήλες”-Επόμενο

Επιλογές γραφήματος

Τίτλοι: Τίτλος γραφήματος

Άξονας κατηγοριών (X)

Άξονας τιμών (Y)

-Επόμενο

Τοποθέτηση γραφήματος

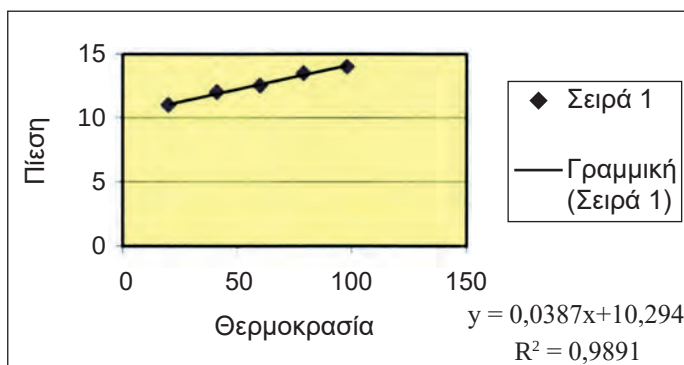
Θέση γραφήματος

Ως αντικείμενο στο Φύλλο 1 -Τέλος

Για να εμφανιστεί η εξίσωση της ευθείας παλινδρόμησης ακολουθούμε την εξής διαδικασία:

Δεξί κλικ πάνω στα σημεία του διαγράμματος διασποράς και από τις επιλογές του παραθύρου που εμφανίζεται επιλέγουμε την “προσθήκη τάσης”.

Στη συνέχεια από τη στήλη με τίτλο “τύπος” επιλέγουμε το “τύπος τάσης γραμμικής”, ενώ από τη στήλη με τίτλο “επιλογές” επιλέγουμε το “προβολή εξίσωσης στο γράφημα” και “προβολή τιμής R-τετράγωνο στο γράφημα”



$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \frac{1}{n} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + (x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \dots + (x_n^2 - 2x_n\bar{x} + \bar{x}^2)]$$



ΚΕΦΑΛΑΙΟ 7°

Να κατανοήσουν οι μαθητές τη σημασία και το ρόλο των χρονολογικών σειρών στη σύγχρονη επιχειρηματική δραστηριότητα και να γνωρίσουν τα βασικά στοιχεία ανάλυσής τους.

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \frac{1}{n} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + (x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \dots + (x_n^2 - 2x_n\bar{x} + \bar{x}^2)]$$

ΕΡΩΤΗΣΕΙΣ ΚΑΤΑΝΟΗΣΗΣ

Οι ερωτήσεις 1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13 είναι ασκήσεις συμπλήρωσης θεωρίας και οι απαντήσεις τους βρίσκονται στις παραγράφους του έβδομου κεφαλαίου. Οι μαθητές πρέπει να τις εντοπίσουν και να συμπληρώσουν τις ερωτήσεις αυτές είτε μέσα στην τάξη, για να ελεγχθεί αν έχει γίνει η σωστή εμπέδωση της θεωρίας, είτε σαν εργασία στο σπίτι.

11.

14. α

15. γ

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \frac{1}{n} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + (x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \dots + (x_n^2 - 2x_n\bar{x} + \bar{x}^2)]$$

ΑΣΚΗΣΕΙΣ

Ομάδα Α'

- 1 Ο παρακάτω πίνακας δίνει την ετήσια παραγωγή ενός προϊόντος από μία επιχείρηση σε τόνους κατά τη διάρκεια των ετών 1996-2000.

Έτη	Προϊόν
1996	5
1997	10
1998	10
1999	15
2000	20

Ζητείται με τη βοήθεια της μεθόδου των ελαχίστων τετραγώνων να προσδιοριστεί η ευθύγραμμη τάση.

Λύση

ΒΗΜΑ I:

Δημιουργούμε τη στήλη (t), αντιστοιχίζοντας το μεσαίο έτος 1998 στην υποδιαίρεση μηδέν του άξονα t (επειδή έχουμε περιττό αριθμό ετών).

ΒΗΜΑ II:

Υπολογισμός των Σy , Σyt , και Σt^2 . Αθροίζοντας τους όρους της στήλης (2) βρίσκουμε $\Sigma y = 60$. Πολλαπλασιάζοντας τους όρους της στήλης (3) με τους αντίστοιχους όρους της στήλης (2), δημιουργούμε τη στήλη (4). Η άθροιση των όρων της (4) μας δίνει $\Sigma yt = 35$.

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Προϊόν y	t	yt	t^2
1996	5	-2	-10	4
1997	10	-1	-10	1
1998	10	0	0	0
1999	15	1	15	1
2000	20	2	40	4
	$\sum y = 60$		$\sum yt = 35$	$\sum t^2 = 10$

ΒΗΜΑ III:

Υψώνοντας στο τετράγωνο κάθε όρο της στήλης t δημιουργούμε τους αντίστοιχους όρους της στήλης t^2 , στήλη (5). Αθροίζοντας τους όρους της (5) βρίσκουμε: $\sum t^2 = 10$.

ΒΗΜΑ IV:

Υπολογισμός των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$. Αντικαθιστούμε στους τύπους υπολογισμού των $\hat{\alpha}$ και $\hat{\beta}$ έχουμε:

$$\hat{\alpha} = \frac{\sum y}{v} = \frac{60}{5} = 12, \quad \hat{\alpha} = 12$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{35}{10} = 3,5, \quad \hat{\beta} = 3,5$$

Συνεπώς η ζητούμενη εξίσωση ευθύγραμμης τάσης είναι:

$$y = 12 + 3,5t$$

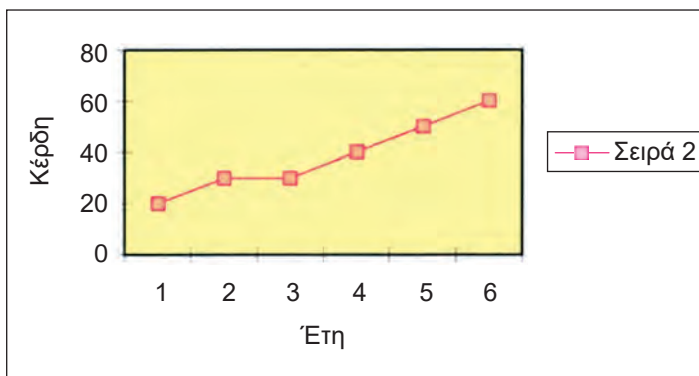
- 2 Τα κέρδη προσφορών μιας επιχείρησης σε εκατ. δρχ. κατά τα έτη 1994-2000 δίνονται από τον παρακάτω πίνακα:

Έτη	Κέρδη
1995	20
1996	30
1997	30
1998	40
1999	50
2000	60

Ζητείται να γίνει το σχετικό χρονοδιάγραμμα.

Λύση

Κατασκευάζουμε το σχετικό χρονοδιάγραμμα έχοντας υπόψη ότι ο άξονας των τετμημένων (οριζόντιος) συμβολίζει τη μεταβλητή του χρόνου (t) ο δε άξονας των τεταγμένων (κάθετος) τη μεταβλητή του κέρδους (Υ).



Είναι προφανές ότι, αν θελήσω να αντικαταστήσω την παραπάνω “ακανόνιστη, μη ομαλή καμπύλη”, με μια εξομαλυνθείσα καμπύλη τάσης, θα πρέπει να προσδιορίσω την αντίστοιχη εξίσωση ευθύγραμμης τάσης (βλέπε άσκηση 3).

- 3 Με βάση τα δεδομένα της άσκησης (2) να προσδιοριστεί με τη μέθοδο των ελαχίστων τετραγώνων η ευθύγραμμη τάση $\hat{y} = \hat{\alpha} + \hat{\beta}t$**

Λύση

Συμπληρώνουμε τον πίνακα δεδομένων της άσκησης 2:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Κέρδη y	t	yt	t^2
1995	20	-2,5	-50	6,25
1996	30	-1,5	-45	2,25
1997	30	-0,5	-15	0,25
1998	40	0,5	20	0,25
1999	50	1,5	75	2,25
2000	60	2,5	150	6,25
	$\sum y = 230$		$\sum yt = 135$	$\sum t^2 = 17,5$

ΒΗΜΑ I:

Σχηματισμός της στήλης t , (3), λαμβάνοντας ως διάμεσο έτος το ημιάθροισμα των ετών 1997 και 1998 (1997,5), επειδή εδώ έχουμε άρτιο αριθμό ετών (6 έτη).

ΒΗΜΑ II:

Σχηματίζουμε τη στήλη yt , (4) του πίνακα, πολλαπλασιάζοντας τους όρους της στήλης y , (2), επί τους αντίστοιχους όρους της στήλης t , (3).

Κατόπιν αθροίζοντας τους όρους της στήλης yt βρίσκουμε $\sum yt = 135$. Επίσης αθροίζοντας τους όρους της στήλης (2), βρίσκουμε $\sum y = 230$.

ΒΗΜΑ III:

Υψώνοντας στο τετράγωνο κάθε όρο της στήλης t , δημιουργούμε κατόπιν τους όρους της στήλης (5) και βρίσκουμε $\sum t^2 = 17,5$.

ΒΗΜΑ IV:

Υπολογισμός των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$. Αντικαθιστούμε στους τύπους υπολογισμού των $\hat{\alpha}$ και $\hat{\beta}$ έχουμε:

$$\hat{\alpha} = \frac{\sum y}{N} = \frac{230}{6} = 38,33, \quad \hat{\alpha} = 38,33$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{135}{17,5} = 7,71, \quad \hat{\beta} = 7,71$$

Κατά συνέπεια η ζητούμενη εξίσωση ευθύγραμμης τάσης είναι η:

$$\hat{y} = 38,33 + 7,71t$$

- 4 Ο αριθμός των ετήσιων τροχαίων ατυχημάτων μιας διοικητικής περιφέρειας κατά τα έτη 1994 έως και το 2000 δίνεται από τον πίνακα που ακολουθεί:**

Έτη	Αριθμός Ατυχημάτων
1994	350
1995	400
1996	400
1997	450
1998	500
1999	550
2000	600

Να προσαρμοστεί στα παραπάνω δεδομένα η ευθύγραμμη τάση $\hat{y} = \hat{\alpha} + \hat{\beta}t$. Επίσης να γίνει μια πρόβλεψη για τον αριθμό ατυχημάτων το έτος 2002.

Λύση

Συμπληρώνουμε τον πίνακα δεδομένων με τις στήλες υπολογισμών:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Αριθμός Ατυχημάτων y	t	yt	t^2
1994	350	-3	-1050	9
1995	400	-2	-800	4
1996	400	-1	-400	1
1997	450	0	0	0
1998	500	1	500	1
1999	550	2	1100	4
2000	600	3	1800	9
	$\sum y = 3250$		$\sum yt = 1150$	$\sum t^2 = 28$

ΒΗΜΑ I:

Σχηματισμός της στήλης t , (3), λαμβάνοντας ως διάμεσο έτος το 1997.

ΒΗΜΑ II:

Σχηματίζουμε τη στήλη yt , (4) του πίνακα, και αθροίζοντας τους όρους της στήλης βρίσκουμε $\sum yt = 1150$. Επίσης από την άθροιση των όρων της στήλης y βρίσκουμε $\sum y = 3250$.

ΒΗΜΑ III:

Σχηματίζουμε, υψώνοντας στο τετράγωνο κάθε όρο της στήλης t , τη στήλη t^2 , (5). Αθροίζοντας τους όρους της στήλης (5) βρίσκουμε $\sum t^2 = 28$.

ΒΗΜΑ IV:

Αντικαθιστούμε στους τύπους υπολογισμού των $\hat{\alpha}$ και $\hat{\beta}$, και προκύπτει ότι:

$$\hat{\alpha} = \frac{\sum y}{v} = \frac{3250}{7} = 464,285, \quad \hat{\alpha} = 464,28$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{1150}{28} = 41,07, \quad \hat{\beta} = 41,07$$

Συνεπώς η ζητούμενη εξίσωση ευθύγραμμης τάσης είναι:

$$\hat{y} = 464,28 + 41,07t$$

Ο προβλεπόμενος αριθμός ατυχημάτων για το έτος 2002 θα είναι:

$$\hat{y} = 464,28 + 41,07 \cdot 5$$

και $\hat{y} = 669,63$. Δηλαδή περίπου 670 ατυχήματα.

* Σ.Σ. Επειδή είχαμε αντιστοιχήσει για το έτος 1997 $t = 0$, η αρχή μέτρησης των ετών γίνεται από το έτος 1997 ($t = 0$), οπότε στο έτος 2002 έχουμε $t = 5$ το οποίο θέσαμε στην εξίσωση που βρήκαμε.

5 Τα γενικά έξοδα σε εκατ. δρχ. ετησίως της εταιρείας Αστήρ-3Α ήταν κατά τα έτη 1992 μέχρι και 1999 όπως παρακάτω:

Έτη	Έξοδα
1992	0,5
1993	0,5
1994	1
1995	1
1996	1,5
1997	2
1998	2,5
1999	3

Να προσδιοριστεί: i) η ευθύγραμμη τάση $\hat{y} = \hat{\alpha} + \hat{\beta}t$,
ii) να γίνει πρόβλεψη για το 2002 και το 2004.

Λύση

Δημιουργούμε τον παρακάτω πίνακα:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Έξοδα y	t	yt	t ²
1992	0,5	-3,5	-1,75	12,25
1993	0,5	-2,5	-1,25	6,25
1994	1	-1,5	-1,5	2,25
1995	1	-0,5	-0,5	0,25
1996	1,5	0,5	0,75	0,25
1997	2	1,5	3	2,25
1998	2,5	2,5	6,25	6,25
1999	3	3,5	10,5	12,25
Αθροίσματα	12		15,5	42

ΒΗΜΑ I:

Σχηματισμός της στήλης t, (3), λαμβάνοντας ως διάμεσο έτος το ημιάθροισμα των ετών 1995 και 1996 (1995,5).

ΒΗΜΑ II:

Σχηματίζουμε κατά τα γνωστά τη στήλη yt, και βρίσκουμε $\sum yt = 15,5$. Επίσης από τη στήλη (2) βρίσκουμε, $\sum y = 12$.

ΒΗΜΑ III:

Κατά τα γνωστά σχηματίζουμε τη στήλη t² και βρίσκουμε $\sum t^2 = 42$.

ΒΗΜΑ IV:

Αντικαθιστούμε στους τύπους υπολογισμού των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$ και έχουμε:

$$\hat{\alpha} = \frac{\sum y}{v} = \frac{12}{8} = 1,5, \quad \hat{\alpha} = 1,5$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{15,5}{42} = 0,37 \quad , \quad \hat{\beta} = 0,37$$

Συνεπώς η προσδιορισθείσα γραμμική τάση είναι:

$$\hat{y} = 1,5 + 0,37t$$

Ο προβλεπόμενος αριθμός ατυχημάτων για το έτος 2002 και το έτος 2004 είναι:

$$\text{Έτος 2002: } \hat{y} = 1,5 + 0,37 \cdot 6,5 = 3,9$$

$$\text{Έτος 2004: } \hat{y} = 1,5 + 0,37 \cdot 8,5 = 4,645$$

Δηλαδή, προβλέπεται το ύψος των γενικών εξόδων για το έτος 2002 να είναι 3.900.000 δρχ. και για το έτος 2004 να είναι 4.645.000 δρχ.

Ομάδα Β'

- 1 Οι διαφημιστικές δαπάνες της εταιρείας Ελλάς-Κόσμος κατά τα έτη 1993 μέχρι 1999 δίνονται από τον παρακάτω πίνακα (σε εκατομ. δρχ.):**

Έτη	Δαπάνες
1993	8
1994	10
1995	10
1996	12
1997	14
1998	16
1999	20

Ζητείται:

- i) Να προσδιοριστεί η ευθύγραμμη τάση $\hat{y} = \hat{\alpha} + \hat{\beta}t$.
- ii) Με τη βοήθεια της εξίσωσης $\hat{y} = \hat{\alpha} + \hat{\beta}t$ να εκτιμήσετε την τάση για τα έτη 1993-1999.
- iii) Να βρείτε τις διαφορές μεταξύ των πραγματικών τιμών y και των τιμών $\hat{y}$ που μας έδωσε η $\hat{y} = \hat{\alpha} + \hat{\beta}t$ για τα έτη 1993 μέχρι και 1999.
- iv) Να σχολιαστεί το αποτέλεσμα του ερωτήματος (γ) , οι διαφορές $y - \hat{y}$

Λύση

Δημιουργούμε τον παρακάτω πίνακα:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Δαπάνες y	t	yt	t^2
1993	8	-3	-24	9
1994	10	-2	-20	4
1995	10	-1	-10	1
1996	12	0	0	0
1997	14	1	14	1
1998	16	2	32	4
1999	20	3	60	9
	$\sum_y = 90$		$\sum_{yt} = 52$	$\sum_t^2 = 28$

i) Προσδιορισμός της γραμμικής τάσης

ΒΗΜΑ I:

Σχηματισμός της στήλης (3). Διάμεσο έτος είναι το 1996.

ΒΗΜΑ II:

Εύρεση του $\sum yt = 52$ από τη στήλη (4). Επίσης από τη (2) βρίσκουμε $\sum y = 90$.

ΒΗΜΑ III:

Από τη στήλη (5) έχουμε $\sum t^2 = 28$.

ΒΗΜΑ IV:

Εύρεση των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$:

$$\hat{\alpha} = \frac{\sum y}{v} = \frac{90}{7} = 12,85, \quad \hat{\alpha} = 12,85$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{52}{28} = 1,85, \quad \hat{\beta} = 1,85$$

Συνεπώς έχουμε: $\hat{y} = 12,85 + 1,85t$

ii) Εκτίμηση για τα έτη 1993 μέχρι 1999.

Έτος

1993	$\hat{y} = 12,85 + 1,85 \cdot (-3) = 7,3$
1994	$\hat{y} = 12,85 + 1,85 \cdot (-2) = 9,15$
1995	$\hat{y} = 12,85 + 1,85 \cdot (-1) = 11$
1996	$\hat{y} = 12,85 + 1,85 \cdot 0 = 12,85$
1997	$\hat{y} = 12,85 + 1,85 \cdot 1 = 23,772$
1998	$\hat{y} = 12,85 + 1,85 \cdot 2 = 47,545$
1999	$\hat{y} = 12,85 + 1,85 \cdot 3 = 71,317$

iii) Εύρεση των διαφορών $(y - \hat{y})$ για τα έτη 1993 - 1999.

(1) Έτη	(2) y	(3) $\hat{y}$	(4) (y- $\hat{y}$)
1993	8	7,3	0,7
1994	10	9,15	0,85
1995	10	11	-1
1996	12	12,85	-0,85
1997	14	14,7	-0,7
1998	16	16,55	-0,55
1999	20	18,45	1,55

iv) Όπως παρατηρούμε από τους όρους της στήλης (4) οι διαφορές μεταξύ των πραγματικών τιμών και των αντίστοιχων τιμών που μας έδωσε η εξίσωση γραμμική τάσης $\hat{y} = 12,85 + 1,85x$ είναι πολύ μικρές. Αυτό σημαίνει ότι η εξίσωση γραμμικής τάσης που προσδιορίσαμε κάνει πολύ καλή προσαρμογή στα δεδομένα μας, δηλαδή η καμπύλη της $\hat{y} = 12,85 + 1,85x$ περιγράφει πολύ καλά τη διαχρονική εξέλιξη (1993 - 1999) των διαφημιστικών μας δαπανών. (Φυσικά υπάρχουν και άλλες μέθοδοι ακόμα και μεγαλύτερης ακρίβειας για την ανωτέρω επισήμανση που ξεφεύγουν όμως των ορίων της παρούσας παρουσίασης).

2 Ο παρακάτω πίνακας παρουσιάζει τις αγορές που έγιναν με πιστωτικές κάρτες στην Ελλάδα κατά τα έτη 1991 μέχρι και 1997. Τα ποσά εκφράζουν εκατομμύρια δρχ. και έχουν στρογγυλοποιηθεί.

Έτη	Ποσά
1991	8
1992	10
1993	21
1994	44
1995	29
1996	61
1997	45

Πηγή: Δελτίο Ελληνικών Τραπεζών

Ζητείται:

- i) Να προσδιοριστεί η ευθεία τάσης $\hat{y} = \hat{\alpha} + \hat{\beta}t$.
- ii) Με βάση την $\hat{y} = \hat{\alpha} + \hat{\beta}t$ να εκτιμηθούν οι τιμές της $\hat{y}$ για το 1998 και το 2000.
- iii) Ποιος ήταν ο ρυθμός της μέσης ετήσιας μεταβολής των αγορών με πιστωτικές κάρτες κατά την εξεταζόμενη περίοδο;
- iv) Να γίνει το σχετικό διάγραμμα και να χαραχθεί πάνω σ' αυτό η ευθεία τάσης.

Λύση

Κατασκευάζουμε τον παρακάτω πίνακα:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Ποσά y	t	yt	t^2
1991	8	-3	-24	9
1992	19	-2	-38	4
1993	21	-1	-21	1
1994	44	0	0	0
1995	29	1	29	1
1996	61	2	122	4
1997	45	3	135	9
	$\sum y=227$		$\sum yt=203$	$\sum t^2=28$

- i) Προσδιορισμός της $\hat{y} = \hat{\alpha} + \hat{\beta}t$

ΒΗΜΑ I:

Επιλογή ως διαμέσου έτους το 1994 και σχηματισμός της στήλης t , (3).

ΒΗΜΑ II:

Εύρεση του $\sum yt = 203$ από τη στήλη (4). Επίσης από τη (2) βρίσκουμε $\sum y = 227$.

ΒΗΜΑ III:

Από τη στήλη t^2 έχουμε $\sum t^2 = 28$.

ΒΗΜΑ IV:

Εύρεση των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$:

$$\hat{\alpha} = \frac{\sum y}{v} = \frac{227}{7} = 32,43, \quad \hat{\alpha} = 32,43$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{203}{28} = 7,25, \quad \hat{\beta} = 7,25$$

Συνεπώς έχουμε:

$$\hat{y} = 32,43 + 7,25t$$

- ii)** Οι τιμές της $\hat{y}$ για τα έτη 1998 και 2000 είναι:
Έτος

$$\hat{y} = 32,43 + 7,25 \cdot 4 = 61,43$$

1998 και

$$\hat{y} = 61,43$$

$$\hat{y} = 32,43 + 7,25 \cdot 6 = 75,93$$

2000 και

$$\hat{y} = 75,93 \text{ εκατομμύρια δρχ.}$$

- iii)** Ο ρυθμός της μέσης ετήσιας μεταβολής, ως γνωστό, δίνεται, εκφράζεται από το συντελεστή $\hat{\beta}$, δηλαδή ήταν ίσος με 7,25 εκατομμύρια δρχ. Πιο απλά από έτος σε έτος οι αγορές με πιστωτικές κάρτες αυξάνονται κατά μέσο όρο κατά 7,25 εκατομμύρια δρχ.

- iv)** Για τη χάραξη της ευθείας τάσης πρέπει να λάβουμε υπόψη μας ότι η ευθεία τάσης και ο κάθετος άξονας y τέμνονται στο σημείο ($t = -3$, $y = 10,68$). Αυτό γίνεται, διότι έχουμε αντιστοιχίσει το έτος 1994 στην υποδιαίρεση 0 (μηδέν) του οριζόντιου άξονα του χρόνου (t). Δηλαδή $\hat{y} = 32,43 + 7,25 \cdot (-3) = 10,68$.

- 3 Ο παρακάτω πίνακας παρουσιάζει το σύνολο ασφαλιστρων, ως ποσοστό επί του Α.Ε.Π, κατά τα έτη 1990 μέχρι και το 1997 που καταβλήθηκαν στα διάφορα ταμεία κοινωνικής ασφάλισης:

Έτη	Σύνολο ασφαλιστρων (%)
1989	1,10
1990	1,25
1991	1,30
1992	1,46
1993	1,59
1994	1,61
1995	1,65
1996	1,72

Πηγή: 8^η Συν/ση για την Υγεία και τις πολ. υγείας - 1999

Ζητείται:

- Να γίνει ο προσδιορισμός της ευθύγραμμης τάσης $\hat{y} = \hat{\alpha} + \hat{\beta}t$
- Να βρεθεί η μέση ετήσια μεταβολή των ασφαλιστρων κατά την περίοδο 1989-1996.
- Να γίνει μια πρόβλεψη για το 2000.
- Να εκτιμηθούν οι τιμές της y για τα έτη 1989 μέχρι 1996 με βάση την εξίσωση $\hat{y} = \hat{\alpha} + \hat{\beta}t$
- Να σχολιασθεί το αποτέλεσμα του ερωτήματος (δ), δηλαδή οι διαφορές $(y - \hat{y})$.

Λύση

- i) Προσδιορισμός της γραμμικής τάσης

Δημιουργούμε τον παρακάτω πίνακα:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	Σύνολο ασφαλιστρών (%)			
	y	t	yt	t ²
1989	1,10	-3,5	-3,85	12,25
1990	1,25	-2,5	-3,12	6,25
1991	1,30	-1,5	-1,95	2,25
1992	1,46	-0,5	-0,73	0,25
1993	1,59	0,5	0,79	0,25
1994	1,61	1,5	2,41	2,25
1995	1,65	2,5	4,12	6,25
1996	1,72	3,5	6,02	12,25
	$\sum y = 11,68$		$\sum yt = 3,69$	$\sum t^2 = 42$

ΒΗΜΑ I:

Επειδή έχουμε άρτιο αριθμό ετών λάβαμε ως διάμεσο έτος το ημιάθροισμα των ετών 1992 και 1993 και δημιουργήσαμε τη στήλη t, (3).

ΒΗΜΑ II:

Υπολογισμός των αθροισμάτων των όρων των στηλών (4) και (2)

$$\sum yt = 3,69$$

$$\sum y = 11,68$$

ΒΗΜΑ III:

Από την άθροιση των όρων της στήλης t² έχουμε $\sum t^2 = 42$.

ΒΗΜΑ IV:

Εύρεση των $\hat{\alpha}$ και $\hat{\beta}$:

$$\hat{\alpha} = \frac{\sum y}{n} = \frac{11,68}{8} = 1,46, \quad \hat{\alpha} = 1,46$$

και

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{3,69}{42} = 0,08, \quad \hat{\beta} = 0,08$$

Επομένως η εξίσωση της γραμμικής τάσης είναι: $\hat{y} = 1,46 + 0,08t$

ii) Η μέση ετήσια μεταβολή των ασφαλιστρών κατά την περίοδο 1989 με 1996 δίνεται από το συντελεστή $\hat{\beta}$ και είναι ίση με 0,08. Δηλαδή είχαμε μια μέση αύξηση από έτος σε έτος αρκετά μικρή, προφανώς επειδή τα ετήσια ασφαλιστρα εκφράζονται ως ποσοστό από του Α.Ε.Π. και δεν ήταν καθαροί αριθμοί.

iii) Η πρόβλεψη για το έτος 2000 είναι:

$$\hat{y} = 1,46 + 0,08 \cdot 7,5 = 2,06$$

2000 και

$$\hat{y} = 2,06.$$

iv) Τιμές της $\hat{y}$ για τα έτη 1989 ως και 1996.

Με βάση την $\hat{y} = 1,46 + 0,08t$ βρίσκουμε:

Για το

$$1989 \quad \hat{y} = 1,46 + 0,08 \cdot (-3,5) = 1,18$$

$$1990 \quad \hat{y} = 1,46 + 0,08 \cdot (-2,5) = 1,26$$

$$1991 \quad \hat{y} = 1,46 + 0,08 \cdot (-1,5) = 1,34$$

$$1992 \quad \hat{y} = 1,46 + 0,08 \cdot (-0,5) = 1,42$$

$$1993 \quad \hat{y} = 1,46 + 0,08 \cdot (0,5) = 1,50$$

$$1994 \quad \hat{y} = 1,46 + 0,08 \cdot (1,5) = 1,58$$

$$1995 \quad \hat{y} = 1,46 + 0,08 \cdot (2,5) = 1,66$$

$$1996 \quad \hat{y} = 1,46 + 0,08 \cdot (3,5) = 1,74$$

οπότε έχουμε την τρίτη στήλη του παρακάτω πίνακα:

(1) Έτη	(2) y	(3) $\hat{y}$	(4) $(y-\hat{y})$
1989	1,10	1,18	-0,08
1990	1,25	1,26	-0,01
1991	1,30	1,34	-0,04
1992	1,46	1,42	0,04
1993	1,59	1,50	0,09
1994	1,61	1,58	0,03
1995	1,65	1,66	-0,01
1996	1,72	1,74	-0,02

- v) Όπως παρατηρούμε οι τιμές $\hat{y}$ που βρέθηκαν από την εξίσωση γραμμικής τάσης $\hat{y} = 1,46 + 0,08t$ είναι πολύ κοντά στις πραγματικές τιμές y . Αυτό γίνεται ακόμη πιο φανερό από τις μικρές διαφορές των πραγματικών τιμών y από τις αντίστοιχες $\hat{y}$, τέταρτη στήλη $(y - \hat{y})$. Αυτή η διαπίστωση μας δείχνει ότι η προσδιορισθείσα εξίσωση τάσης εκφράζει πολύ ικανοποιητικά τη διαχρονική εξέλιξη της μεταβλητής Y .

4 Η συνολική ιδιωτική δαπάνη υγείας κατά τη χρονική περίοδο 1990 μέχρι και 1997 δίνεται από τον παρακάτω πίνακα:

Έτη	Δαπάνες υγείας
1990	370
1991	446
1992	546
1993	702
1994	978
1995	1.043
1996	1.120
1997	1.170

Πηγή: 8^η Συν/ση για την υγεία και τις πολ. υγείας - 1999.

* Τα ποσά εκφράζουν εκατ. δρχ. και έχουν στρογγυλοποιηθεί.

Ζητείται:

- i) Να γίνει ο προσδιορισμός της ευθείας τάσης $\hat{y} = \hat{\alpha} + \hat{\beta}t$
- ii) Να γίνει πρόβλεψη για τα έτη 1999 και 2002
- iii) Να γίνει με βάση την $\hat{y} = \hat{\alpha} + \hat{\beta}t$ εκτίμηση των τιμών της Y για τα έτη 1990 μέχρι και 1997
- iv) Να βρεθούν οι διαφορές των πραγματικών τιμών y της χρονολογικής σειράς από τις αντίστοιχες $\hat{y}$ που μας δίνει η $\hat{y} = \hat{\alpha} + \hat{\beta}t$
- v) Να γίνει το σχετικό διάγραμμα και να χαραχθεί πάνω σ' αυτό η ευθεία τάσης.

Λύση

- i) Προσδιορισμός της $\hat{y} = \hat{\alpha} + \hat{\beta}t$
Κατασκευάζουμε τον παρακάτω πίνακα:

Στήλες δεδομένων		Στήλες Υπολογισμών		
(1)	(2)	(3)	(4)	(5)
Έτη	y	t	yt	t ²
1991	446	-3	-1338	9
1992	546	-2	-1092	4
1993	702	-1	-702	1
1994	978	0	0	0
1995	1043	1	1043	1
1996	1.120	2	2240	4
1997	1.170	3	3510	9
	$\sum y = 6.005$		$\sum yt = 3661$	$\sum t^2 = 28$

Ακολουθώντας τη γνωστή διαδικασία των 4 βημάτων βρίσκουμε τα παρακάτω αθροίσματα -ως Κεντρικό Έτος (Κ.Ε.) επιλέξαμε το 1994.

Κ.Ε. : 1994

$$\sum y = 6.005$$

$$\sum yt = 3.661$$

$$\sum t^2 = 28$$

$$n = 7$$

$$\hat{\alpha} = \frac{\sum y}{v} = \frac{6005}{7} = 857,85, \quad \hat{\alpha} = 857,85$$

$$\hat{\beta} = \frac{\sum yt}{\sum t^2} = \frac{3661}{28} = 130,75, \quad \hat{\beta} = 130,75$$

Επομένως η εξίσωση της γραμμικής τάσης είναι:

$$\hat{y} = 857,85 + 130,75t$$

ii) Πρόβλεψη για τα έτη 1999 και 2002 είναι:

Έτος

$$\hat{y} = 857,85 + 130,75 \cdot 5 = 1511,6$$

1999

και

$$\hat{y} = 1511,6 \cdot 10^6 \text{ δρχ.}$$

$$\hat{y} = 857,85 + 130,75 \cdot 8 = 1903,85$$

2002

και

$$\hat{y} = 1903,85 \cdot 10^6 \text{ δρχ.}$$

iii) Εκτίμηση τιμών της $\hat{y}$ για τα έτη 1991 ως και 1997.

Έτος

$$\hat{y} = 857,85 + 130,75 \cdot (-3) = 465,6$$

1991

$$\text{και } \hat{y} = 465,6 \cdot 10^6 \text{ δρχ.}$$

$$\hat{y} = 857,85 + 130,75 \cdot (-2) = 596,35$$

1992

$$\text{και } \hat{y} = 596,35 \cdot 10^6 \text{ δρχ.}$$

$$\hat{y} = 857,85 + 130,75 \cdot (-1) = 727,1$$

1993

$$\text{και } \hat{y} = 727,1 \cdot 10^6 \text{ δρχ.}$$

$$\hat{y} = 857,85 + 130,75 \cdot 0 = 857,85$$

1994

$$\text{και } \hat{y} = 857,85 \cdot 10^6 \text{ δρχ.}$$

1995 $\hat{y} = 857,85 + 130,75 \cdot 1 = 987,75$
 και $\hat{y} = 987,75 \cdot 10^6$ δρχ.

1996 $\hat{y} = 857,85 + 130,75 \cdot 2 = 1.119,35$
 και
 $\hat{y} = 1.119,35 \cdot 10^6$ δρχ.

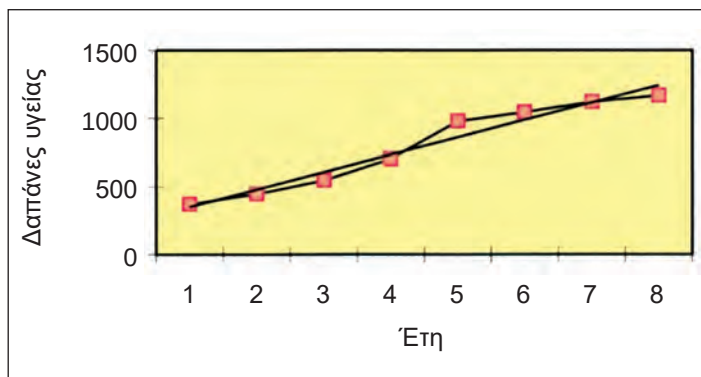
1997 $\hat{y} = 857,85 + 130,75 \cdot 3 = 1250,1$
 και
 $\hat{y} = 1250,1 \cdot 10^6$ δρχ.

οπότε έχουμε την τρίτη στήλη του παρακάτω πίνακα:

iv) Εύρεση των διαφορών ($y - \hat{y}$)

(1) Έτη	(2) y	(3) $\hat{y}$	(4) $(y - \hat{y})$
1991	446	465,6	-19,6
1992	546	596,35	-50,35
1993	702	727,1	-25,1
1994	978	857,85	120,15
1995	1.043	987,75	55,25
1996	1.120	1.119,85	0,15
1997	1.170	1.250,1	-80,1

v) Κατασκευή χρονοδιαγράμματος



Κατασκευάζουμε το σχετικό χρονοδιάγραμμα κατά το γνωστό μας τρόπο (βλέπε άσκηση 7) και παρατηρούμε ότι η ευθεία τάσης $\hat{y} = 857,85 + 130,75t$ διέρχεται ανάμεσα από τα πραγματικά ζεύγη των παρατηρήσεων μας (t_i, y_i) . Παρατηρούμε επίσης ότι η $\hat{y} = 857,85 + 130,75t$ προσαρμόζεται πολύ καλά στα πραγματικά μας δεδομένα και αναπαριστά πολύ καλά τη διαχρονική εξέλιξη της ιδιωτικής δαπάνης για την υγεία.

ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ ΚΕΦΑΛΑΙΟΥ 6

Η ευθύγραμμη τάση χρησιμοποιείται για δεδομένα που σχετίζονται με πωλήσεις, τιμές μετοχών, τιμές προϊόντων και άλλα είδη δεδομένων που συλλέγονται περιοδικά με στόχο να αναλυθούν με τη βοήθεια χρονολογικών σειρών.

Στη δραστηριότητα που ακολουθεί τα δεδομένα (πίνακας Δ3) αφορούν την εγκατάσταση πισίνων από μια εταιρεία ανά τρίμηνο.

Τρίμηνο	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Αριθμός πισίνων	15	18	27	23	20	22	33	29	24	27	38	33	28	30

Να γίνει το σχετικό διάγραμμα και να χαραχθεί σ' αυτό η ευθεία τάσης.

Λύση



Εισαγωγή δεδομένων.

AC/ON → MENU

Εισάγουμε τα δεδομένα του πίνακα Δ.3 όπως ακριβώς στις δραστηριότητες των προηγούμενων κεφαλαίων.

```

➤ → LIST2      1  1 EXE
                 2  2 EXE
                 3  3 EXE
                 4  4 EXE
                 .....

```

```

14              14 EXE

```

```

➤➤ LIST 3      1 15 EXE
                 2 18 EXE
                 3 27 EXE
                 4 23 EXE
                 .....

```

```

14 30 EXE

```




Επεξεργασία δεδομένων

Επιλέγουμε:

F1 (GRAPH)

F6 (SET)

F2 (GRAPH)

F2 (XY)

F2 (LIST2)

F3 (LIST3)

F1(□)

EXIT

F2 (GRAPH2) XY



F1 (X)

XY1

LinearReg

a =1.1010989

b =17.9560439

r =0.72666335

r²=0.52803963

y=ax+b

[COPY] [DRAW]

$$s^2 = \frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2] = \frac{1}{n} [(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + (x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \dots + (x_n^2 - 2x_n\bar{x} + \bar{x}^2)]$$

F6 (DRAW)

XY2



EXIT

MENU

RUN (μαυρισμένο)

EXE

OPTN

F5 (STAT)

15 F2 ($\hat{y}$) EXE

50 F1 ($\hat{x}$) EXE

XY3

15

50

34.47252747

29.10179641

Σ Σ

EXIT

EXIT

B Με χρήση Η/Υ (πρόγραμμα Microsoft EXCEL)

❖ Εισαγωγή δεδομένων

Από το MENU επιλέγουμε το πρόγραμμα Microsoft EXCEL
Στο Βιβλίο 1 ,Φύλλο 1 και στις στήλες Α και Β πληκτρολογούμε τα δεδομένα του πίνακα Δ.3

❖ Επεξεργασία δεδομένων.

Επιλέγουμε (μαυρίζουμε) τη στήλη Β (αριθμός πισίνων) και από την γραμμή εργαλείων (toolbar) επιλέγουμε τον οδηγό γραφημάτων. Οι δυνατότητες που έχουμε είναι δύο:

Βασικοί τύποι –Προσαρμοσμένοι τύποι γραφημάτων.

Στο συγκεκριμένο παράδειγμα χρησιμοποιήσαμε τους “βασικούς τύπους γραφημάτων” και ακολουθήσαμε τα παρακάτω βήματα.

Βασικοί τύποι

Τύπος γραφήματος “Γραμμές” (4° γράφημα) -Επόμενο

Δεδομένα προέλευσης

Περιοχή δεδομένων-Σειρά σε “στήλες” -Επόμενο

Επιλογές γραφήματος

Τίτλοι : Τίτλος γραφήματος

Άξονας κατηγοριών (X)

Άξονας τιμών (Y) -Επόμενο

Τοποθέτηση γραφήματος

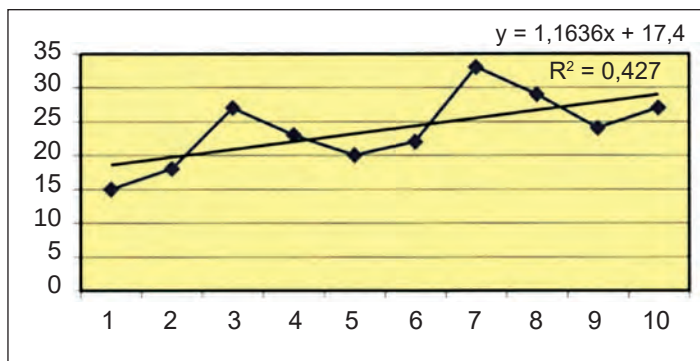
Θέση γραφήματος

Ως αντικείμενο στο Φύλλο 1 -Τέλος

Για να εμφανιστεί η εξίσωση της γραμμικής τάσης ακολουθούμε την εξής διαδικασία:

Δεξί κλικ πάνω στα σημεία του γραφήματος και από τις επιλογές του παραθύρου που εμφανίζεται επιλέγουμε την “προσθήκη τάσης”.

Στη συνέχεια από τη στήλη με τίτλο “τύπος” επιλέγουμε το “τύπος τάσης γραμμικής” ενώ από τη στήλη με τίτλο “επιλογές” επιλέγουμε το “προβολή εξίσωσης στο γράφημα” και “προβολή τιμής R-τετράγωνο στο γράφημα”



Οι παραπάνω δραστηριότητες έγιναν με τη χρήση του γραφικού υπολογιστή CASIO CFX - 9850GB PLUS

Επίσης θα μπορούσαν να χρησιμοποιηθούν οι γραφικοί υπολογιστές CASIO CFX - 9700, 9750, 9800, 9950, 9970.

Καθώς και οι γραφικοί υπολογιστές TI - 82, 83, 85, 86, 89.

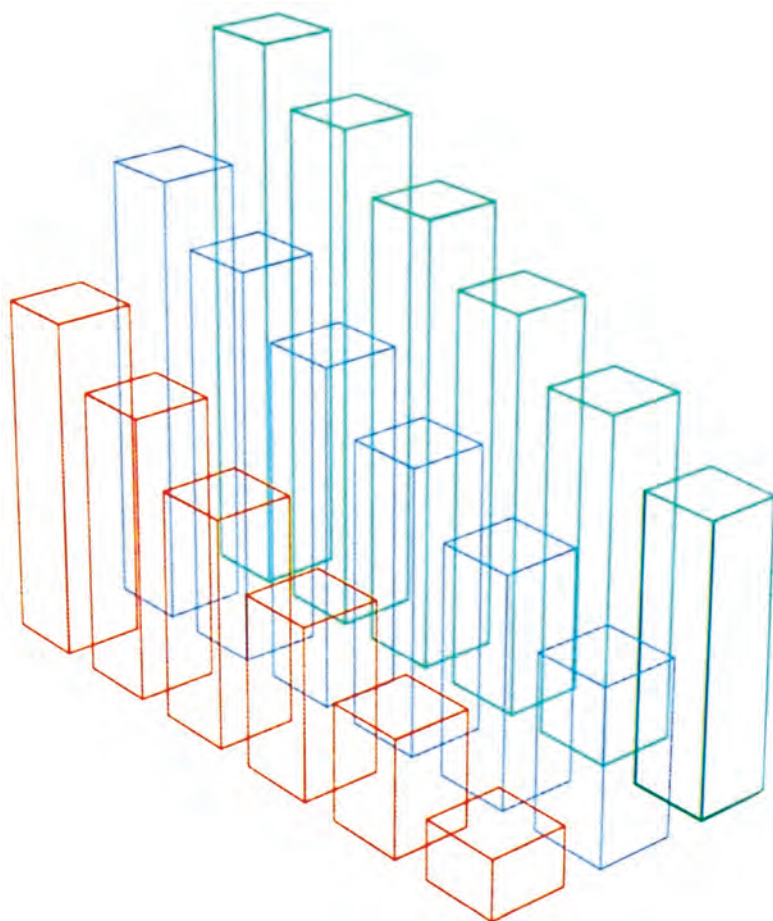
Επίσης HP 28, 38G, 48

SHARP 9200, 9300, 9600 κ.λ.π.

Από το πρόγραμμα Microsoft Excel χρησιμοποιήθηκαν μόνο οι βασικές δυνατότητες του προγράμματος.

ΠΑΡΑΡΤΗΜΑ

ΛΕΞΙΛΟΓΙΟ



Ο Ρ Ο Λ Ο Γ Ι Α

Η εξέλιξη της σύγχρονης τεχνολογίας καθώς και η χρήση των ηλεκτρονικών υπολογιστών επιβάλλουν τη γνώση ξένης (κυρίως αγγλικής) ορολογίας. Για το λόγο αυτό παραθέτουμε ένα μικρό απόσπασμα λεξικού (αγγλοελληνικού) μαθηματικών και οικονομικών όρων. Πιστεύουμε ότι ο μαθητής θα βοηθηθεί στα πρώτα του βήματα στη μελέτη και επεξεργασία στατιστικών δεδομένων.

Λ Ε Ξ Ι Λ Ο Γ Ι Ο

Α Γ Γ Λ Ι Κ Η

Ε Λ Λ Η Ν Ι Κ Η

A

A_i	= κατάταξη είδους που βρίσκεται σε άριστη κατάταξη
Ability	= ικανότητα
abscissa	= τετμημένη
absolute	= απόλυτος
aggregative index	= γενικός δείκτης
analysis	= ανάλυση
analysis of variance	= ανάλυση διακύμανσης
annual mean	= μέσος ετήσιος
approximation	= προσέγγιση
arithmetic mean	= αριθμητικός μέσος όρος
ascribable to x	= αποδιδόμενη στη x
asymptotic	= ασυμπτωτικός
attribute data	= ποιοτικά δεδομένα
average	= μέσος όρος
average deviation	= μέση απόκλιση
average production	= μέση παραγωγή
average tax rate	= μέσος φορολογικός συντελεστής

B

Balance	= υπόλοιπο, ισοζύγιο
bar charts of bar diagrams	= ακιδωτά διαγράμματα (ραβδογράμματα)
base period of base year	= περίοδος βάσης ή έτος βάσης
block	= κομμάτι, τμήμα
boundary	= σύνορο, όριο
boundary, lower, upper	= όριο τάξης, κατώτερο, ανώτερο
business cycle	= επιχειρηματικός κύκλος
business forecasting	= επιχειρηματική πρόγνωση
business statistics	= στατιστική επιχειρήσεων

C

cac 40	= σημαντικός χρηματιστηριακός δείκτης τιμών μετοχών του Χ.Α. Παρισίων
calculation	= υπολογισμός
call	= απαίτηση, κλήση
category	= κατηγορία, κλάση
central tendency	= κεντρική τάση
charts	= διαγράμματα
class	= τάξη (κλάση)
class boundaries	= όρια τάξης
class frequency	= συχνότητα τάξης
class interval	= διάστημα τάξης
class limits	= όρια διαστήματος τάξης
classification	= ταξινόμηση
classification of data	= κατάταξη δεδομένων
classification table	= πίνακας ταξινόμησης
closed interval	= κλειστό διάστημα
coding method	= κωδικοποιημένη μέθοδος

coefficient	= συντελεστής
coefficient of correlation	= συντελεστής συσχέτισης
coefficient of regression	= συντελεστής παλινδρόμησης
coefficient of variability	= συντελεστής μεταβλητικότητας
collection	= συλλογή
column	= στήλη
complex	= πολύπλοκος, σύνθετος
components	= συνιστώσες
computation	= υπολογισμός
condition	= όρος, συνθήκη
constant	= σταθερά
continuous	= συνεχής
continuous variable	= συνεχής μεταβλητή
coordinates	= συντεταγμένες
correlation	= συσχέτιση
correlation coefficient	= συντελεστής συσχέτισης
cumulant diagrams	= αθροιστικά διαγράμματα
cumulative	= αθροιστικός
cumulative frequency distribution	= αθροιστική κατανομή συχνότητας
cumulative frequency polygon	= πολύγωνο αθροιστικής συχνότητας
curve	= καμπύλη
curvilinear regression	= καμπυλόγραμμος παλινδρόμηση
curve fitting	= προσαρμογή καμπύλης
curve of yield	= καμπύλη απόδοσης

D

daily interest	= ημερήσιος τόκος
data	= δεδομένα
data bank	= τράπεζα πληροφοριών
decile	= δεκατημόριο
dependence	= εξάρτηση

dependence variables	= εξαρτημένες μεταβλητές
descriptive	= περιγραφικός
descriptive statistics	= περιγραφική στατιστική
deviation	= απόκλιση
diagram	= διάγραμμα
difference	= διαφορά
dimension	= διάσταση
direct taxation	= άμεση φορολογία
discontinuous distribution	= ασυνεχής κατανομή
discrete	= διακεκριμένος, διακριτός
discrete variable	= διακριτή (απαριθμητή) μεταβλητή
dispersion	= διασπορά
dispersion parameters	= παράμετρος διασποράς
distribution curve	= καμπύλη κατανομής
distribution plan	= σχέδιο κατανομής

E

E and O (Errors and Omissions)	= σφάλματα και παραλείψεις στην αναγραφή στοιχείων σε καταστάσεις
economic statistic	= οικονομική στατιστική
economic indicators	= οικονομικοί δείκτες
effect	= αποτέλεσμα
efficiency	= αποτελεσματικότητα, αποδοτικότητα
elementary	= στοιχειώδης, βασικός
elementary event	= στοιχειώδες γεγονός
elements of statistics	= στοιχεία στατιστικής
error	= σφάλμα
error of sampling	= σφάλμα δειγματοληψίας
estimated regression equation	= εκτιμηθείσα εξίσωση παλινδρόμησης
estimation	= εκτίμηση
evaluator	= εκτιμητής

even number	= άρτιος αριθμός
Eurostat	= η Στατιστική Υπηρεσία της Ευρωπαϊκής Ένωσης
explanatory variable	= ερμηνευτική μεταβλητή
exponential	= εκθετικός
exponential curve	= εκθετική καμπύλη
exponential regression	= εκθετική παλινδρόμηση
external audit	= εξωτερικός έλεγχος

F

factor	= παράγοντας
finance	= διαχείριση χρηματικών πόρων επιχείρησης, κράτους ή ιδιώτη
finite	= πεπερασμένος
finite population	= πεπερασμένος αριθμός
fit	= προσαρμόζω - γη
fitting	= προσαρμογή
fitted line	= προσαρμοσθείσα γραμμή
fluctuation	= διακύμανση
formula	= τύπος
frequency	= συχνότητα
frequency curve	= καμπύλη συχνοτήτων
frequency diagrams	= διαγράμματα συχνοτήτων
frequency distribution	= κατανομή συχνότητας
frequency polygon	= πολύγωνο συχνότητας
frequency table	= πίνακας συχνότητας
function	= συνάρτηση
fundamental	= θεμελιώδης, βασικός
future value	= μελλοντική αξία

G

gain	= κέρδος
game theory	= θεωρία παιγνίων
garnering of data	= συλλογή δεδομένων
general	= γενικός
general wholesale price index	= γενικός τιμάριθμος χονδρικής πώλησης
generalization	= γενίκευση
goodness of fit	= προσαρμοστική ικανότητα
grade	= βαθμός
graph	= γραφική παράσταση
graph paper	= χαρτί σχεδίασης
graphic method	= γραφική μέθοδος
group	= ομάδα
grouped data	= ταξινομημένα σε ομάδες δεδομένα

H

half	= μισός
highly significant result	= πολύ σημαντικό αποτέλεσμα
histogram	= ιστόγραμμα
historical series or time series	= ιστορικές ή χρονολογικές σειρές
hyper inflation	= υπερπληθωρισμός
hypothesis	= υπόθεση

I

import duty	= εισαγωγικός δασμός
independence	= ανεξαρτησία
independent events	= ανεξάρτητα γεγονότα
independent variables	= ανεξάρτητες μεταβλητές
index number	= αριθμοδείκτες

index numbers Laspeyres	= αριθμοδείκτης του Laspeyres
index numbers Paasche	= αριθμοδείκτης του Paasche
index numbers unweighted	= αστάθμητοι αριθμοδείκτες
indexes or indices	= αριθμοδείκτες
inductive statistics	= επαγωγική στατιστική
inefficient	= μη αποτελεσματικός, μη αποδοτικός
inequality	= ανισότητα
inference	= συμπέρασμα
infinite	= άπειρος
infinite population	= άπειρος πληθυσμός
infinity	= άπειρο
inflation	= πληθωρισμός
intercep (on x-axis)	= τομή (με τον άξονα x)
interpretation	= ερμηνεία
interval	= διάστημα
inverse	= αντίστροφος
inversion	= αντιστροφή
investor	= επενδυτής

L

Lagrangian multiplier	= μέθοδος των πολλαπλασιαστών (πολλαπλασιαστής του Lagrange)
last quarter	= τελευταίο τρίμηνο
least square estimate	= εκτίμηση με τη μέθοδο ελαχίστων τετραγώνων
least squares line	= ευθεία ελαχίστων τετραγώνων
least squares method	= μέθοδος ελαχίστων τετραγώνων
limit	= όριο
line	= ευθεία
linear	= γραμμικός

linear correlation	= γραμμική συσχέτιση
linear correlation coefficient	= συντελεστή γραμμικής συσχέτισης
linear function	= γραμμική συνάρτηση
linear model	= γραμμικό υπόδειγμα
linear relationship	= γραμμική σχέση εξάρτησης
linearity	= γραμμικότητα
loan	= δάνειο
long term	= χρονική περίοδος > 12 μηνών

M

macroeconomics	= μακροοικονομική
map charts	= χαρτοδιαγράμματα
maximizing	= μεγιστοποίηση
maximum	= μέγιστος
mean	= μέση τιμή
mean deviation	= μέση απόκλιση
means	= μέσο, μέσα
measurable	= μετρήσιμος
measurable set	= μετρήσιμο σύνολο
measure	= μέτρο
measurement	= μέτρηση
measures of location	= στατιστικά μέτρα θέσης
measures of central tendency	= μέτρα κεντρικής τάσης
measures of dispersion	= μέτρα διασποράς
median	= διάμεσος
method	= μέθοδος
minimum	= ελάχιστος
mode	= επικρατούσα τομή της μεταβλητής ή τύπος
model	= υπόδειγμα
monthly index	= μηνιαίος δείκτης
multiple	= πολλαπλός

N

national income accounts	= εθνικοί λογαριασμοί εισοδήματος
national statistical service	= εθνική στατιστική υπηρεσία
negative	= αρνητικός
nonlinear relationship	= μη γραμμική σχέση
nonlinear correlation	= μη ευθύγραμμη συσχέτιση
normal	= κανονικός, κάθετος
normal curve	= κανονική καμπύλη
normal equations	= κανονικές εξισώσεις
notation	= συμβολισμός
numbers	= αριθμοί

O

objectivity	= αντικειμενικότητα
observation	= παρατήρηση
observed frequency	= παρατηρούμενη συχνότητα
odd	= περιττός
odd number	= περιττός αριθμός
open interval	= ανοικτό διάστημα
operation	= επιχείρηση, πράξη
origin	= αρχή
original	= αρχικός
original data	= αρχικά αριθμητικά δεδομένα
orthogonal	= ορθογώνιος
output units	= μονάδες εξόδου

P

pair	= ζεύγος
puper currency	= χαρτονόμισμα
parabola	= παραβολή

parabolic regression	= παραβολική παλινδρόμηση
parameter	= παράμετρος
parity	= ισοτιμία
parity index	= δείκτης ισοτιμίας
percentages	= ποσοστά
percentage change	= ποσοστιαία μεταβολή
percentile	= εκατοστημόριο
percentiles or percentile values	= εκατοστιαία σημεία
perfect linear correlation	= πλήρους γραμμική συσχέτιση
permutation	= διάταξη
personal interviews	= προσωπικές συνεντεύξεις
pictograms or pictographs	= ειδοδιαγράμματα
pie diagram	= κυκλικό διάγραμμα
plane	= επίπεδο
point	= σημείο
polygon	= πολύγωνο
polygon graph	
or frequency polygon	= πολύγωνο συχνότητας
population	= πληθυσμός
population size	= μέγεθος πληθυσμού
precision	= ακρίβεια
prediction	= πρόβλεψη
price index	= δείκτης τιμών
price relative	= σχετική τιμή
price consumer index	= δείκτης τιμών καταναλωτή
principle	= αρχή
probability	= πιθανότητα
probable	= πιθανός
production	= παραγωγή
program	= πρόγραμμα
provisional data	= προσωρινά στοιχεία
provisions	= προβλέψεις

Q

quality	= ποιότητα
qualitative distribution	= ποιοτική κατανομή
quality control	= έλεγχος ποιότητας
quantity	= ποσότητα
quantity index number	= αριθμοδείκτης ποσότητας
quantity price relation	= σχέση ποσότητας και τιμής
quantum index	= δείκτης όγκου
quartile	= τεταρτημόριο
quotient	= πηλίκο

R

raise	= αύξηση (τιμών, μισθών κ.λ.π.)
random	= τυχαίος
randomization	= τυχαιότητα
random number	= τυχαίος αριθμός
range	= κλίμακα
ratio	= λόγος
raw data	= ακατέργαστα δεδομένα
real number	= πραγματικός αριθμός
regression	= παλινδρόμηση
regression coefficient	= συντελεστής παλινδρόμησης
regression curve	= καμπύλη παλινδρόμησης
regression equation	= εξίσωση παλινδρόμησης
regression line	= ευθεία παλινδρόμησης
relation	= σχέση
relationship	= σχέση
relative	= σχετικός
relative frequency	= σχετική συχνότητα
reliability	= αξιοπιστία
runaway inflation	= υψηλότατος πληθωρισμός

representative	= αντιπροσωπευτικός
retail price index	= δείκτης τιμών λιανικής πώλησης
rule	= κανόνας

S

salary scale	= κλίμακα μισθών
sample	= δείγμα
sample mean	= μέση τιμή δείγματος
sample size	= μέγεθος δείγματος
sampling	= δειγματοληψία
sampling technique	= τεχνική δειγματοληψία
scale	= κλίμακα
sampling technique	= τεχνική δειγματοληψία
scale	= κλίμακα
seasonal variation	= εποχική μεταβολή
shape	= μορφή, σχήμα
slope	= κλίση
standard deviation	= σταθερή (τυπική) απόκλιση
standards score	= τυπικό αποτέλεσμα
standard unit	= τυπική μονάδα
statistical analysis	= στατιστική ανάλυση
statistical data	= στατιστικά δεδομένα
statistical decision	= στατιστική απόφαση
statistical hypothesis	= στατιστική υπόθεση
statistical inference	= στατιστική συμπερασματολογία
statistics	= στατιστική
sum	= άθροισμα
summation signs	= σύμβολα άθροισης
sure event	= βέβαιο γεγονός
systematic error	= συστηματικό λάθος

T

table	= πίνακας
tabulation	= τακτοποίηση και παρουσίαση στοιχείων σε πίνακα
tendency	= τάση
term	= όρος
theoretical	= θεωρητικός
time	= χρόνος
time chart or time series chart	= χρονοδιάγραμμα
time series	= χρονολογικές σειρές
tolerance	= ανοχή
total	= σύνολο, ολικός
transformation	= μετασχηματισμός
triangle	= τρίγωνο
type	= τύπος
turnover	= κύκλος εργασιών ή ρυθμός αντικατάστασης

U

unadjusted data	= αρχικά δεδομένα μη εξομαλυθέντα
uncorrelated variables	= ασυσχέτιστες μεταβλητές
unequal	= άνισος
ungrouped data	= αταξινόμητα δεδομένα
uniform distribution	= ομοιόμορφη κατανομή
unimodal distribution	= μονοκόρυφος κατανομή
unit	= μονάδα
unit value index	= δείκτης μέσης αξίας
universe set	= βασικό σύνολο
unweighted price indices	= αστάθμητος αριθμοδείκτης τιμών

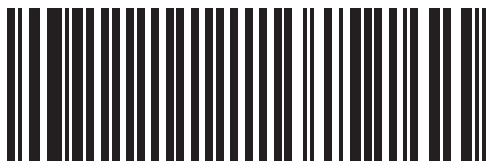
V

valuation	= αξιολόγηση, εκτίμηση
value	= αξία, τιμή
value indices	= δείκτης αξίας
variable	= μεταβλητή
variance	= διακύμανση
variance analysis	= ανάλυση διακύμανσης
variation	= μεταβλητότητα (μεταβολή)
variation, coefficient of	= συντελεστής μεταβλητότητας
volatility	= μεταβλητότητα
vertical scale	= κατακόρυφος κλίμακα
volume indices	= δείκτης όγκου

W

wage	= μισθός, ημερομίσθιο
weights	= συντελεστές στάθμισης
weighted aggregative price index	= σταθμικός δείκτης τιμών
weighted price indices	= σταθμικοί αριθμοδείκτες τιμών
weighted relative price index	= σταθμικός δείκτης των σχετικών τιμών
width or class interval	= πλάτος ή διάστημα τάξης
write off	= διαγραφή

Κωδικός βιβλίου: 0-24-0200
ISBN 978-960-06-2978-1



(01) 000000 0 24 0200 0