# ΕΝΟΤΗΤΑ 2.1 : ΠΛΗΘΥΣΜΟΣ – ΔΕΙΓΜΑ – ΜΕΤΑΒΛΗΤΕΣ

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

Ο δήμαρχος μιας πόλης θέλει να διερευνήσει τα προβλήματα που αντιμετωπίζει η πόλη του ώστε να δώσει έμφαση στην επίλυση αυτών. Για το λόγο αυτό, ανέθεσε σε μια εταιρεία δημοσκοπήσεων μια έρευνα κατά την οποία 500 δημότες κλήθηκαν να δηλώσουν, ποιο πρόβλημα της πόλης θεωρούν το πιο σημαντικό. Στις παρακάτω ερωτήσεις να επιλέξετε τη σωστή απάντηση.

1. Ο πληθυσμός της έρευνας είναι:
   1. Οι 500 δημότες που ρωτήθηκαν.
   2. Όλοι οι δημότες της πόλης.
   3. Όλοι οι Έλληνες πολίτες.
2. Το δείγμα αποτελούν:
   1. Οι 500 δημότες που ρωτήθηκαν.
   2. Όλοι οι δημότες της πόλης.
   3. Όλοι οι Έλληνες πολίτες.
3. Η μεταβλητή της έρευνας είναι:
   1. Τα προβλήματα της πόλης.
   2. Τα προβλήματα των δημοτών.
   3. Το σημαντικότερο πρόβλημα της πόλης.

Λύση

1. Όλοι οι δημότες της πόλης
2. Οι 500 δημότες που ρωτήθηκαν
3. Το σημαντικότερο πρόβλημα της πόλης

##### Άσκηση 2

Στις παρακάτω περιπτώσεις ποιες μπορεί να είναι οι μεταβλητές που μας ενδιαφέρουν; Να γίνει η διάκρισή τους σε ποιοτικές ή ποσοτικές και να αναφερθούν μερικές δυνατές τιμές τους:

* 1. Εξετάζουμε ένα δείγμα υπαλλήλων μιας εταιρείας.
  2. Εξετάζουμε ένα δείγμα προϊόντων από μια παραγωγή.
  3. Εξετάζουμε ένα δείγμα τηλεθεατών.
  4. Εξετάζουμε τους καλαθοσφαιριστές μιας ομάδας σε έναν αγώνα. Λύση

1. Ένα δείγμα υπαλλήλων μιας εταιρείας θα μπορούσε να εξεταστεί ως προς τις αμοιβές

τους. Τότε η μεταβλητή είναι «ύψος αμοιβής», είναι ποσοτική και μερικές τιμές της θα μπορούσε να είναι 580, 750, 1250, 2500 Ευρώ.

Ωστόσο, το ίδιο δείγμα υπαλλήλων θα μπορούσε να εξεταστεί ως προς τις σπουδές. Η μεταβλητή θα μπορούσε να είναι «επίπεδο σπουδών», που είναι ποιοτική και μερικές τιμές της θα μπορούσαν να είναι: «απολυτήριο Γυμνασίου», «απολυτήριο Λυκείου»,

«πτυχίο ανώτατης εκπαίδευσης», «μεταπτυχιακό δίπλωμα», «διδακτορικό δίπλωμα» κ.α. Ή θα μπορούσε να είναι «κατεύθυνση σπουδών» που είναι ποιοτική και μερικές τιμές της είναι «οικονομικές», «μηχανικού», «διοικητικού» κ.α.

1. Ένα δείγμα προϊόντων θα μπορούσε να εξεταστεί με μεταβλητή την ποιότητα (ποιοτική μεταβλητή) και τιμές της θα μπορούσε να είναι: «αποδεκτό» και «απορριπτέο». Αλλά θα μπορούσε να εξεταστεί και με μεταβλητή το βάρος (ποσοτική μεταβλητή).

Ομοίως για τα ερωτήματα 3, 4.

##### Άσκηση 3

Για της ανάγκες μιας έρευνας συγκεντρώσαμε στοιχεία από διερχόμενα οχήματα, σε κάποιο συγκεκριμένο σημείο της πόλης, κατά τη διάρκεια ενός εικοσιτετραώρου. Στον επόμενο πίνακα παρουσιάζονται μερικά από τα στοιχεία αυτά.

|  |  |  |  |
| --- | --- | --- | --- |
| **Είδος οχήματος** | **Χρώμα οχήματος** | **Ταχύτητα σε Km/h** | **Πλήθος επιβατών** |
| Φορτηγό | Κόκκινο | 26 | 2 |
| Αυτοκίνητο ΙΧ | Γκρι | 38 | 4 |
| Ποδήλατο | Πράσινο | 13 | 1 |
| Αυτοκίνητο ΙΧ | Κόκκινο | 52 | 2 |
| Λεωφορείο | Λευκό | 34 | 25 |
| Αυτοκίνητο ΙΧ | Λευκό | 45 | 1 |
| Μοτοσικλέτα | Μαύρο | 62 | 2 |

Ποιες είναι οι μεταβλητές της έρευνας και ποιο το είδος τους; Λύση

Πρώτη στήλη: κατηγορία (είδος) οχήματος – ποιοτική μεταβλητή

Δεύτερη στήλη: χρώμα οχήματος – ποιοτική μεταβλητή

Τρίτη στήλη: ταχύτητα οχήματος στο συγκεκριμένο σημείο τη στιγμή της καταγραφής – ποσοτική μεταβλητή

Τέταρτη στήλη: πλήθος επιβατών τη στιγμή της καταγραφής – ποσοτική μεταβλητή

Σχόλιο: η ταχύτητα του οχήματος θα μπορούσε να θεωρηθεί συνεχής ή διακριτή μεταβλητή (ανάλογα με το αν «επιτρέπουμε» ή όχι όλες τις τιμές μεταξύ δύο ακέραιων αριθμών). Αντιθέτως, το πλήθος επιβατών προφανώς δεν μπορεί να θεωρηθεί συνεχής.

##### Άσκηση 4

Σε μια εκπομπή δημόσιας συζήτησης, συγκεκριμένου τηλεοπτικού καναλιού, το κοινό καλείται να ψηφίσει, αν συμφωνεί με την Α άποψη ή τη Β άποψη. Να αναφέρετε δύο λόγους για τους οποίους τα αποτελέσματα της ψηφοφορίας δεν μπορεί να γενικευτούν σε ολόκληρο τον πληθυσμό της χώρας.

Λύση

Το δείγμα (οι τηλεθεατές που απάντησαν) δεν είναι αντιπροσωπευτικό του πληθυσμού της χώρας, αφού καθορίζεται από το ποιοι προτιμούν να βλέπουν τηλεόραση και ποιοι προτιμούν το συγκεκριμένο τηλεοπτικό κανάλι, ποια ώρα διεξάγεται η έρευνα, τον τρόπο ψηφοφορίας (τηλέφωνο, διαδίκτυο, κλπ), κοκ.

# ΕΝΟΤΗΤΑ 2.2: ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

Το επάγγελμα του πατέρα 20 μαθητών καταγράφηκε στον διπλανό πίνακα. Να κάνετε πίνακα σχετικών συχνοτήτων και να κατασκευάσετε το αντίστοιχο ραβδόγραμμα συχνοτήτων καθώς και το κυκλικό διάγραμμα.

|  |  |
| --- | --- |
| **Επάγγελμα πατέρα** | **Αριθμός ατόμων** |
| Ιδιωτικός υπάλληλος | 6 |
| Δημόσιος υπάλληλος | 7 |
| Αυτοαπασχολούμενος | 5 |
| Άλλο | 2 |

Λύση

Οι σχετικές συχνότητες είναι:

f1 = = = 0,3 και f1% = 30% , f2 = = = 0,35 και f2% = 35% ,

f3 = = = 0,25 και f3% = 25% , f4 = = = 0,1 και f4% = 10% .

Έτσι προκύπτει ο παρακάτω πίνακας σχετικών συχνοτήτων

|  |  |  |  |
| --- | --- | --- | --- |
| **Επάγγελμα πατέρα** | **Αριθμός ατόμων** | **σχετική συχνότητα** | **σχετική συχνότητα %** |
| Ιδιωτικός υπάλληλος | 6 | 0,3 | 30% |
| Δημόσιος υπάλληλος | 7 | 0,35 | 35% |
| Αυτοαπασχολούμενος | 5 | 0,25 | 25% |
| Άλλο | 2 | 0,1 | 10% |
| Σύνολο | 20 | 1,0 | 100% |

Το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα φαίνονται παρακάτω:

8

7

6

5

4

3

2

1

0

Ιδιωτικός υπάλληλος

Δημόσιος υπάλληλος

Αυτοαπασχολούμενος

Άλλο

**Επάγγελμα πατέρα**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  | |
|  |  |
|  | | | | |
|  |  |  |  |  |
|  | | | | | | |
|  |  |  |  |  |  |  |
|  | | | | | |
|  |  |  |  |  |  |

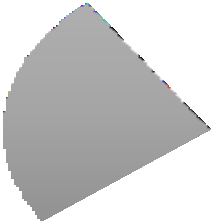


Ιδιωτικός υπάλληλος

Δημόσιος υπάλληλος Αυτοαπασχολούμενος Άλλο

**Επάγγελμα πατέρα**

##### Άσκηση 2



10%

30%

25%

35%

Στον διπλανό πίνακα δίνονται τα καθαρά κέρδη μιας εταιρείας, ανά έτος, από το 2014 έως και το 2017. Να κάνετε χρονόγραμμα όπου να φαίνεται η εξέλιξη των κερδών σε σχέση με το χρόνο.

|  |  |
| --- | --- |
| **Έτος** | **Κέρδη σε ευρώ** |
| 2014 | 180.000 |
| 2015 | 270.000 |
| 2016 | 230.000 |
| 2017 | 210.000 |

Λύση

Το χρονόγραμμα φαίνεται παρακάτω:



##### Άσκηση 3

Η στατιστική υπηρεσία της Πυροσβεστικής μας έδωσε το διπλανό κυκλικό διάγραμμα, που παρουσιάζει τα ποσοστά των κλήσεων ανά κατηγορία. Αν το σύνολο των κλήσεων είναι 60.400, να γίνει πίνακας συχνοτήτων.



Λύση

Οι επεμβάσεις σε ανελκυστήρες είναι το 27% των 60.400 κλήσεων, άρα είναι

27%  60400 

27

100

 60400  16308 , και ομοίως υπολογίζουμε για τις υπόλοιπες κατηγορίες

κλήσεων. Έτσι προκύπτει ο παρακάτω πίνακας συχνοτήτων:

|  |  |  |
| --- | --- | --- |
| **κατηγορίες κλήσεων** | **Αριθμός κλήσεων** | **σχετική συχνότητα %** |
| Επεμβάσεις σε ανελκυστήρες | 16.308 | 27% |
| Παροχές βοήθειας | 12.684 | 21% |
| Πυρκαγιές | 28.388 | 47% |
| Ψευδείς αγγελίες | 3.020 | 5% |
| Σύνολο | 60.400 | 100% |

##### Άσκηση 4

Δίνεται η ποσοστιαία σύνθεση (%) του προσωπικού μιας επιχείρησης, με αριθμό υπαλλήλων 80 άτομα ως προς το μορφωτικό τους επίπεδο.

* 1. Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.

|  |  |
| --- | --- |
| **Τίτλος Σπουδών** | **Ποσοστό (%)** |
| Μεταπτυχιακό Δίπλωμα | 20 |
| Πτυχίο Τριτοβάθμιας Εκπαίδευσης | 50 |
| Απολυτήριο Λυκείου | 30 |

* 1. Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με κυκλικό διάγραμμα.

Λύση

Οι κάτοχοι μεταπτυχιακού διπλώματος είναι το 20% των 80 υπαλλήλων, άρα

20%  80 

20

100

 80  16

υπάλληλοι. Ομοίως βρίσκουμε ότι πτυχίο τριτοβάθμιας

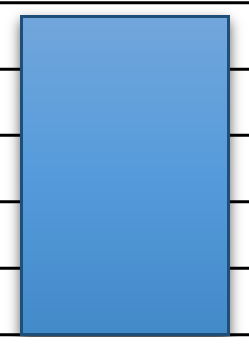
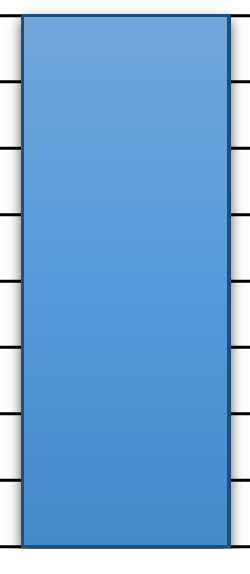
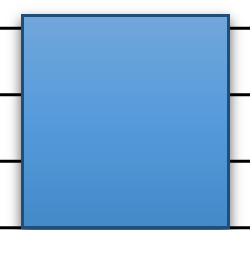
εκπαίδευσης έχουν 40 υπάλληλοι και απόφοιτοι Λυκείου είναι 24 υπάλληλοι. Κατασκευάζουμε έτσι τον παρακάτω πίνακα συχνοτήτων και σχετικών συχνοτήτων:

|  |  |  |  |
| --- | --- | --- | --- |
| **Τίτλος Σπουδών (μεταβλητή)** | **αριθμός υπαλλήλων (συχνότητα)** | **σχετική συχνότητα** | **σχετική συχνότητα %** |
| Μεταπτυχιακό Δίπλωμα | 16 | 0,2 | 20% |
| Πτυχίο Τριτοβάθμιας Εκπαίδευσης | 40 | 0,5 | 50% |
| Απολυτήριο Λυκείου | 24 | 0,3 | 30% |
| Σύνολο υπαλλήλων | 80 | 1,0 | 100% |

Παρακάτω φαίνονται ραβδογράμματα και κυκλικά διαγράμματα που αναπαριστούν αυτά τα δεδομένα. Η επιλογή κάποιου από αυτά εξαρτάται από τους στόχους της παρουσίασης και την προτίμηση αυτού που το παρουσιάζει.

Τίτλος Σπουδών

ραβδογράμματα συχνοτήτων:



45

40

35

30

25

20

15

10

5

0

Μεταπτυχιακό Δίπλωμα Πτυχίο Τριτοβάθμιας Απολυτήριο Λυκείου

Εκπαίδευσης

Τίτλος Σπουδών

Αριθμός Υπαλλήλων

Αριθμός Υπαλλήλων

50

40

30

20

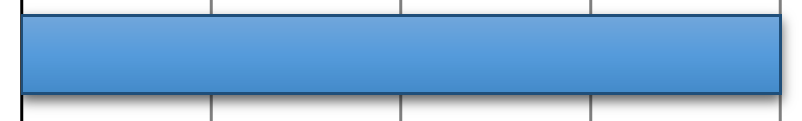
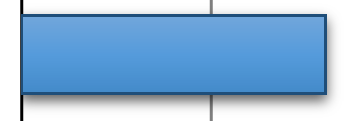
10

0

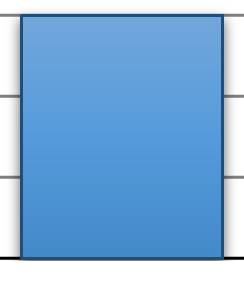
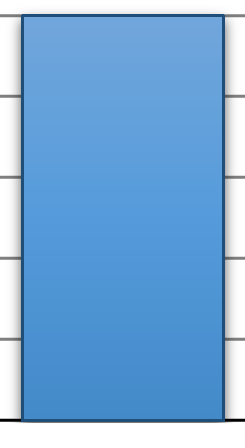
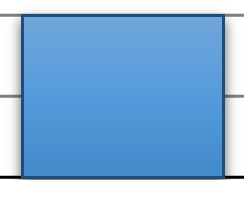
Μεταπτυχιακό Δίπλωμα

Πτυχίο Τριτοβάθμιας Εκπαίδευσης

Απολυτήριο Λυκείου



ραβδογράμματα σχετικών συχνοτήτων:



60%

50%

40%

30%

20%

10%

0%

Μεταπτυχιακό Δίπλωμα

Πτυχίο Τριτοβάθμιας

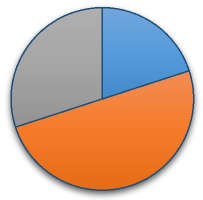
Εκπαίδευσης

Τίτλος Σπουδών

Απολυτήριο Λυκείου

Ποσοστό Υπαλλήλων

Κυκλικά διαγράμματα συχνοτήτων και σχετικών συχνοτήτων:



**τίτλος σπουδών**

**υπαλλήλων**

30%

20%

Μεταπτυχιακό

Δίπλωμα

50%

Πτυχίο

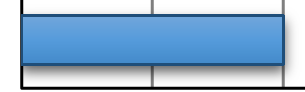
Τριτοβάθμιας Εκπαίδευσης

Μεταπτυχιακό Δίπλωμα

Πτυχίο Τριτοβάθμιας Εκπαίδευσης Απολυτήριο Λυκείου

τίτλος σπουδών

υπαλλήλων



Απολυτήριο Λυκείου

Πτυχίο Τριτοβάθμιας Εκπαίδευσης

Μεταπτυχιακό Δίπλωμα

0% 10% 20% 30% 40% 50% 60%

Ποσοστό Υπαλλήλων

16

24

40

Τίτλος Σπουδών

##### Άσκηση 5

Οι βαθμοί στην Ιστορία 25 μαθητών, ενός τμήματος της Β΄ τάξης ΓΕΛ, είναι:

1. Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.

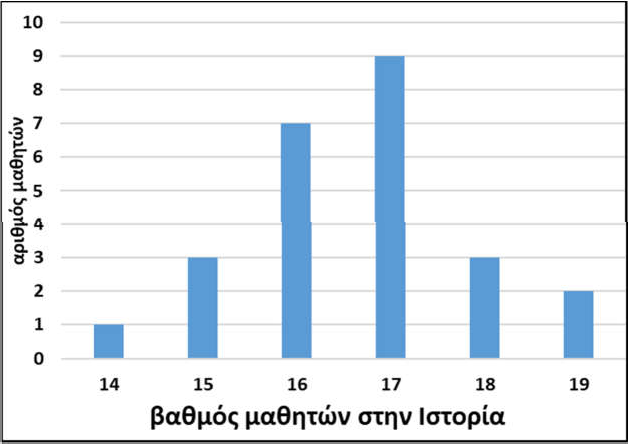
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 16 | 15 | 17 | 16 | 17 |
| 18 | 17 | 16 | 17 | 18 |
| 16 | 19 | 17 | 15 | 16 |
| 17 | 16 | 15 | 17 | 18 |
| 17 | 14 | 17 | 16 | 19 |

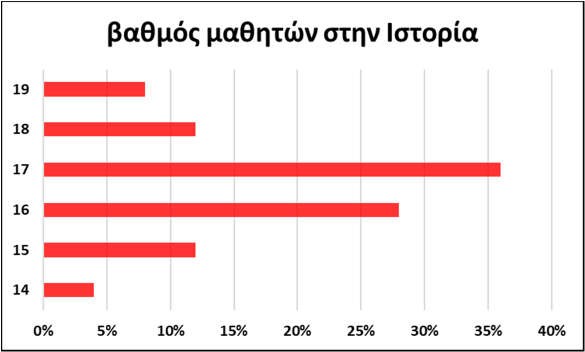
1. Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με σημειόγραμμα.

Λύση

1. Μετά την καταμέτρηση και τους υπολογισμούς, ο πίνακας συχνοτήτων και σχετικών συχνοτήτων είναι ο εξής:

|  |  |  |  |
| --- | --- | --- | --- |
| **βαθμός**  **(μεταβλητή)** | **αριθμός μαθητών**  **(συχνότητα)** | **σχετική**  **συχνότητα** | **σχετική**  **συχνότητα %** |
| 14 | 1 | 0,04 | 4% |
| 15 | 3 | 0,12 | 12% |
| 16 | 7 | 0,28 | 28% |
| 17 | 9 | 0,36 | 36% |
| 18 | 3 | 0,12 | 12% |
| 19 | 2 | 0,08 | 8% |
| Σύνολο | 25 | 1,0 | 100% |

1. Ραβδόγραμμα συχνοτήτων:

Ραβδόγραμμα σχετικών συχνοτήτων:

Σημειόγραμμα:

##### Άσκηση 6

Οι πιο κάτω αριθμοί παρουσιάζουν τις ενδείξεις της άνω έδρας ενός ζαριού το οποίο ρίξαμε 30 φορές.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2 | 5 | 6 | 1 | 2 | 5 | 4 | 3 | 2 | 5 |
| 1 | 3 | 5 | 4 | 1 | 3 | 2 | 6 | 5 | 4 |
| 1 | 2 | 6 | 2 | 4 | 3 | 1 | 6 | 4 | 5 |

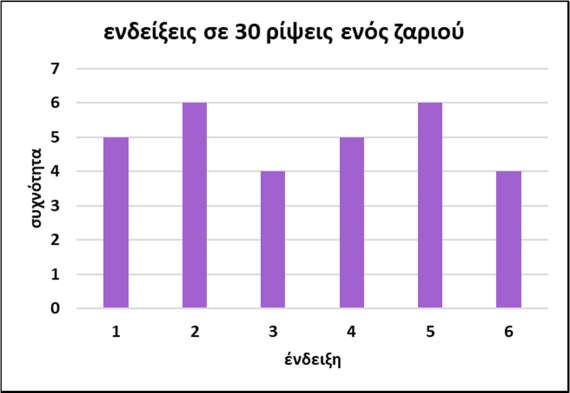
* 1. Να κατασκευάσετε πίνακα συχνοτήτων.
  2. Να κατασκευάσετε ραβδόγραμμα συχνοτήτων. Λύση

Μετά την καταμέτρηση και τους υπολογισμούς, ο πίνακας συχνοτήτων και σχετικών

συχνοτήτων είναι ο εξής:

|  |  |  |  |
| --- | --- | --- | --- |
| **ένδειξη ζαριού**  **(μεταβλητή)** | **συχνότητα** | **σχετική**  **συχνότητα** | **σχετική**  **συχνότητα %** |
| 1 | 5 | 0,167 | 16,7% |
| 2 | 6 | 0,2 | 20% |
| 3 | 4 | 0,133 | 13,3% |
| 4 | 5 | 0,167 | 16,7% |
| 5 | 6 | 0,2 | 20% |
| 6 | 4 | 0,133 | 13,3% |
| Σύνολο | 30 | 1,00 | 100% |

Ένα ραβδόγραμμα συχνοτήτων φαίνεται παρακάτω



##### Άσκηση 7

Στον πιο κάτω πίνακα δίνεται η συγκέντρωση (mgr/cm3) ενός ρύπου στον αέρα 40 πόλεων της χώρας.

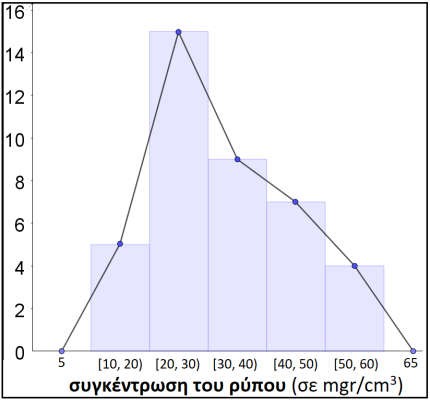
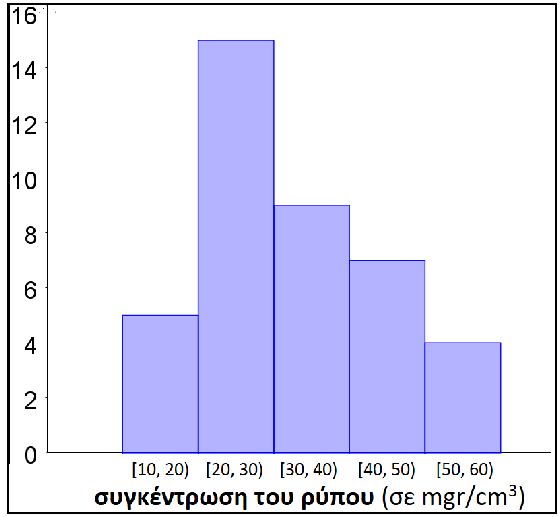
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 16 | 24 | 36 | 47 | 23 | 22 | 43 | 27 | 49 | 48 |
| 12 | 32 | 17 | 38 | 42 | 27 | 31 | 50 | 38 | 21 |
| 36 | 19 | 28 | 31 | 28 | 25 | 45 | 12 | 57 | 51 |
| 22 | 23 | 24 | 25 | 24 | 37 | 43 | 25 | 39 | 51 |

1. Να ομαδοποιήσετε τις παρατηρήσεις στις κλάσεις: [10,20), [20,30), [30,40), [40,50) και [50,60).
2. Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
3. Να κατασκευάσετε το ιστόγραμμα και το πολύγωνο συχνοτήτων. Λύση

Παρακάτω φαίνεται ο πίνακας συχνοτήτων και σχετικών συχνοτήτων, το ιστόγραμμα

συχνοτήτων και το πολύγωνο συχνοτήτων

|  |  |  |  |
| --- | --- | --- | --- |
| **συγκέντρωση του**  **ρύπου** (σε mgr/cm3) | **συχνότητα** | **σχετική συχνότητα** | **σχετική συχνότητα %** |
| [10,20) | 5 | 0,125 | 12,5% |
| [20,30) | 15 | 0,375 | 37,5% |
| [30,40) | 9 | 0,225 | 22,5% |
| [40,50) | 7 | 0,175 | 17,5% |
| [50,60) | 4 | 0,1 | 10% |
| Σύνολο | 40 | 1,00 | 100% |



##### Άσκηση 8

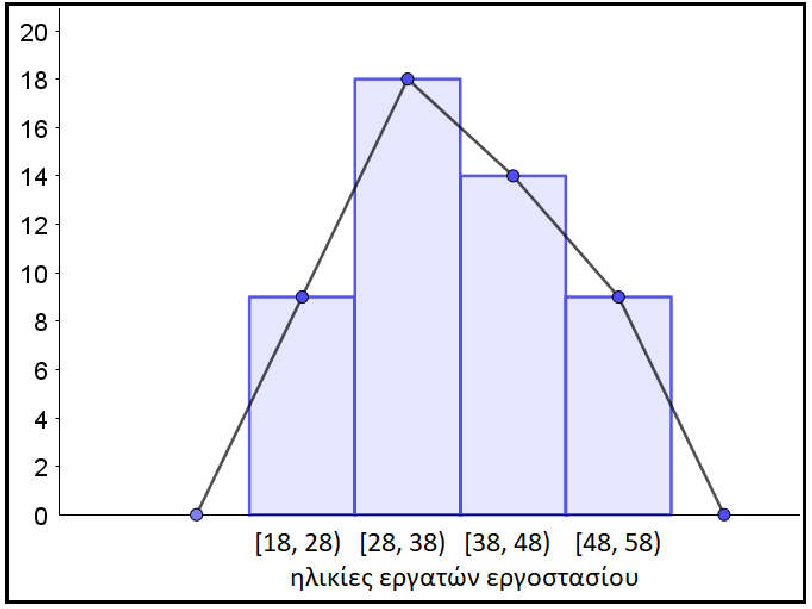
Οι 50 εργάτες ενός εργοστασίου έχουν τις παρακάτω ηλικίες:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 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 | 27 | 27 | 40 | 35 |

1. Να ομαδοποιήσετε τις ηλικίες στις κλάσεις: [18,28), [28,38), [38,48) και [48,58).
2. Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
3. Να κατασκευάσετε το ιστόγραμμα και το πολύγωνο συχνοτήτων. Λύση

Ο πίνακας συχνοτήτων μετά την ομαδοποίηση διαμορφώνεται ως εξής:

|  |  |  |  |
| --- | --- | --- | --- |
| **ηλικίες εργατών** | **συχνότητα** | **σχετική**  **συχνότητα** | **σχετική**  **συχνότητα %** |
| [18,28) | 9 | 0,18 | 18% |
| [28,38) | 18 | 0,36 | 36% |
| [38,48) | 14 | 0,28 | 28% |
| [48,58) | 9 | 0,18 | 18% |
| Σύνολο | 50 | 1,00 | 100% |

Το ιστόγραμμα και το πολύγωνο συχνοτήτων φαίνεται παρακάτω:

# ΕΝΟΤΗΤΑ 2.3: ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ ΘΗΚΟΓΡΑΜΜΑ

**ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ**

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

Οι βαθμοί του Αντρέα σε 4 διαγωνίσματα στα Μαθηματικά ήταν 15, 18, 18, 17. Για τα ίδια διαγωνίσματα, ο Βασίλης είχε πάρει 2 μονάδες περισσότερες, σε κάθε διαγώνισμα, από τον Αντρέα, ενώ ο Γιάννης είχε πάρει 4 μονάδες λιγότερες από τον Αντρέα, σε κάθε διαγώνισμα. Να βρείτε τη μέση τιμή των βαθμών του κάθε παιδιού.

Λύση

Η μέση τιμή (ο μέσος όρος) των βαθμών του Ανδρέα είναι το άθροισμα των βαθμών δια

το πλήθος των μαθημάτων, δηλαδή: 15 + 18 + 18 + 17 = 17. Ομοίως βρίσκουμε ότι για τον

4

Βασίλη είναι 17 + 20 + 20 + 19 = 19 και για τον Γιάννη 11 + 14 + 14 + 13 = 13 . Παρατηρούμε

4 4

ότι η μέση τιμή των βαθμών του Βασίλη είναι κατά 2 μονάδες μεγαλύτερη εκείνης του Ανδρέα. Αυτό οφείλεται στις 2 μονάδες που έχει πάρει παραπάνω ο Βασίλης από τον Ανδρέα σε κάθε μάθημα. Ομοίως, αφού κάθε βαθμός του Γιάννη είναι κατά 4 μικρότερος του αντίστοιχου βαθμού του Ανδρέα, ο μέσος όρος του Γιάννη θα είναι κατά 4 βαθμούς μικρότερος του μέσου όρου του Ανδρέα. (βλέπε και εφαρμογή 2)

##### Άσκηση 2

Αν ο μέσος όρος του μηνιαίου μισθού των υπαλλήλων ενός εργοστασίου πέρυσι ήταν 850€ και φέτος σε κάθε υπάλληλο δοθεί αύξηση 50€, να βρείτε τον νέο μέσο όρο των μισθών.

Λύση

Αν υποθέσουμε ότι το πλήθος των υπαλλήλων είναι ν, τότε το άθροισμα των μισθών όλων των υπαλλήλων την περσινή χρονιά θα ήταν ν  850 € και τη φετινή χρονιά θα είναι

ν850 + ν50

ν  850 + ν  50 €, οπότε ο νέος μέσος όρος των μισθών θα είναι

= 900 €.

## ν

Στο ίδιο συμπέρασμα μπορούμε να καταλήξουμε χωρίς τους παραπάνω υπολογισμούς που έχουν ως αφετηρία την υπόθεση ότι το πλήθος των υπαλλήλων είναι ν. Εφόσον κάθε μισθός αυξήθηκε κατά 50€, και ο μέσος όρος θα έχει αυξηθεί κατά 50€.

Επέκταση: Με δεδομένο ότι ο βασικός μισθός το 2019 είναι 650€, και για έγγαμο εργαζόμενο με προϋπηρεσία τριών χρόνων διαμορφώνεται στα 780€, ποιοι θα μπορούσε να είναι οι μισθοί 100 υπαλλήλων αν η μέση τιμή είναι 900€; Μπορείτε να σκεφτείτε ένα παράδειγμα τέτοιο ώστε η δήλωση "ο μέσος μισθός των υπαλλήλων του εργοστασίου είναι 900€" να είναι παραπλανητική, ακόμα κι αν είναι στατιστικά σωστή;

##### Άσκηση 3

Καθεμία από τις παρακάτω λίστες δεδομένων έχουν μέση τιμή 50.

(Ι) 0, 20, 40, 50, 60, 80, 100

(ΙΙ) 0, 48, 49, 50, 51, 52, 100

(ΙΙΙ) 0, 1, 2, 50, 98, 99, 100

1. Μπορεί να χρησιμοποιηθεί το εύρος για σύγκριση της μεταβλητότητας των δεδομένων αυτών;
2. Χωρίς να γίνουν οι πράξεις, να βρείτε σε ποια λίστα υπάρχει μεγαλύτερη και σε ποια μικρότερη διασπορά παρατηρήσεων.

Λύση

1. Για όλες τις λίστες δεδομένων το εύρος είναι 100  0  100 . Οπότε, δεν μπορούμε να χρησιμοποιήσουμε το εύρος για να διακρίνουμε διαφορές στη μεταβλητότητα μεταξύ των τριών ομάδων δεδομένων.
2. Αναζητώντας ποια ομάδα έχει περισσότερο διάσπαρτα δεδομένα, παρατηρούμε καταρχάς ότι όλες περιλαμβάνουν το 0, το 50 και το 100. Πέραν αυτών των τριών, στη δεύτερη ομάδα (0, 48, 49, 50, 51, 52, 100) τα υπόλοιπα δεδομένα βρίσκονται πολύ κοντά στο κέντρο (που είναι η μέση τιμή, δηλ. το 50). Στην τρίτη ομάδα (0, 1, 2, 50, 98, 99, 100) τα υπόλοιπα δεδομένα βρίσκονται μακριά από το κέντρο (πιο κοντά στα άκρα 0 και 100). Τέλος, στην πρώτη ομάδα (0, 20, 40, 50, 60, 80, 100) οι τιμές 20, 40, 60, 80 είναι πιο ομαλά τοποθετημένες ανάμεσα στα άκρα (0 και 100) και στο κέντρο (50).

Από τα παραπάνω συμπεραίνουμε ότι τη μεγαλύτερη διασπορά έχει η τρίτη λίστα, τη μικρότερη διασπορά έχει η δεύτερη λίστα, ενώ η διασπορά της πρώτης λίστας βρίσκεται κάπου ανάμεσα στις δύο άλλες.

##### Άσκηση 4

Να υπολογίσετε τη μέση τιμή και τη διάμεσο για τα παρακάτω δείγματα δεδομένων και να σχολιάσετε τα αποτελέσματα:

α) 1, 2, 6 β) 2, 4, 12 γ) 11, 12, 16 δ) 12, 14, 22

Λύση

Για τη μέση τιμή των τεσσάρων δειγμάτων έχουμε αντιστοίχως:

∙

## 1 = = 3 , 2 = = 6 , 3 = = 13 , 4 = = 16

## Ενώ για τις διαμέσους έχουμε αντιστοίχως:

δ1 = 2 ,

δ2 = 4 ,

δ3 = 12 ,

δ4 = 14

Παρατηρούμε ότι οι τιμές του δεύτερου δείγματος είναι διπλάσιες των τιμών του πρώτου

και για αυτό ισχύει 2 = 2∙1 και δ2 = 2∙δ1

Οι τιμές του τρίτου δείγματος προέρχονται από τις τιμές του πρώτου αν τις αυξήσουμε

3 = 1 + 10 και δ3 = δ1 + 10 .

κατά 10 και για αυτό ισχύει

Τέλος, οι τιμές του τέταρτου δείγματος προκύπτουν αν διπλασιάσουμε τις τιμές του πρώτου και στο αποτέλεσμα προσθέσουμε 10. Για το λόγο αυτό ισχύει

4 = 2∙1 + 10 και δ4 = 2δ1+10 .

##### Άσκηση 5

Να υπολογίσετε τη διακύμανση και την τυπική απόκλιση για καθεμιά από τις παρακάτω λίστες δεδομένων. Συγκρίνοντας τα δεδομένα και τα αποτελέσματα, τι συμπέρασμα βγάζετε;

α) 1, 3, 4, 5, 7 β) 3, 9, 12, 15, 21 γ) 6, 8, 9, 10, 12 δ) -1, -3, -4, -5, -7

Λύση

Για την (α) ομάδα δεδομένων έχουμε:

α = = 4

= = = 4

και sα = = 2

Για τη (β) ομάδα δεδομένων έχουμε:

β = = 12

= = = 36

και sβ = = 6

Για τη (γ) ομάδα δεδομένων έχουμε:

γ = = 9

= = = 4

και sγ = = 2

Τέλος, για τη (δ) ομάδα δεδομένων έχουμε:

δ = = −4

= = = 4

και sδ = = 2

Παρατηρούμε ότι αν πολλαπλασιάσουμε τα δεδομένα της πρώτης ομάδας επί 3 παίρνουμε τα δεδομένα της δεύτερης. Η διακύμανση της δεύτερης ομάδας είναι 9πλάσια (9=32) της διακύμανσης της πρώτης και η τυπική απόκλιση είναι τριπλάσια. Δηλαδή:

=32∙ και sβ = 3∙sα.

Αν στα δεδομένα της πρώτης ομάδας προσθέσουμε 5, παίρνουμε τα δεδομένα της τρίτης ομάδας. Η διακύμανση και η τυπική απόκλιση της τρίτης ομάδας παραμένουν ίδιες με της

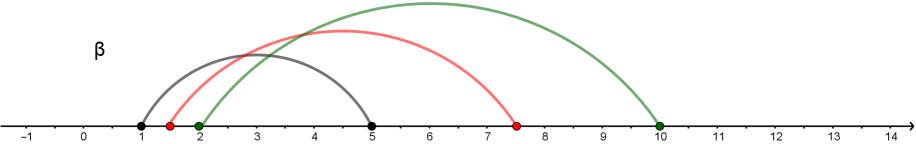
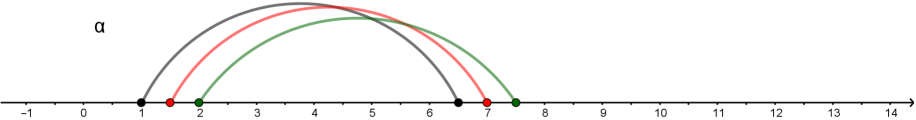
πρώτης. Δηλαδή: = και sγ = sα.

Τέλος, τα δεδομένα της τέταρτης ομάδας είναι τα αντίθετα της πρώτης, ή αλλιώς προκύπτουν

από τα δεδομένα της πρώτης με πολλαπλασιασμό επί (-1). Η διακύμανση και η τυπική απόκλιση της τέταρτης ομάδας παραμένουν ίδιες με της πρώτης. Δηλαδή: = και sδ = sα.

Συνοψίζοντας, θα μπορούσαμε να πούμε ότι πολλαπλασιάζοντας τα δεδομένα με κάποιον αριθμό, η τυπική απόκλιση πολλαπλασιάζεται με την απόλυτη τιμή του αριθμού. Αν όμως προσθέσουμε σε όλα τα δεδομένα τον ίδιο αριθμό, η τυπική απόκλιση δεν μεταβάλλεται.

Σχηματικά, με την πρόσθεση του ίδιου αριθμού σε όλα τα δεδομένα, τα μετατοπίζουμε στον άξονα, χωρίς να αλλάζουμε τις μεταξύ τους αποστάσεις (βλέπε σχήμα α). Αλλά πολλαπλασιάζοντάς τα με τον ίδιο αριθμό τα μετατοπίζουμε και οι μεταξύ τους αποστάσεις πολλαπλασιάζονται με τον ίδιο αριθμό (βλέπε σχήμα β).



Χρησιμοποιώντας συμβολική γλώσσα, τα παραπάνω διατυπώνονται ως εξής:

Αν συμβολίσουμε xα1=1, xα2=3, xα3=4, xα4=5, xα5=7 ,

δηλαδή αν συμβολίσουμε με xαi , i = 1,2,3,4,5 τα δεδομένα της (α) ομάδας

και με xβi , xγi , xδi τα δεδομένα των (β), (γ) και (δ) ομάδων αντιστοίχως, τότε έχουμε:

xβi = 3∙xαi , xγi = xαi + 5 , xδi = −xαi = (−1)∙xαi

Για τις αντίστοιχες διακυμάνσεις και τυπικές αποκλίσεις έχουμε:

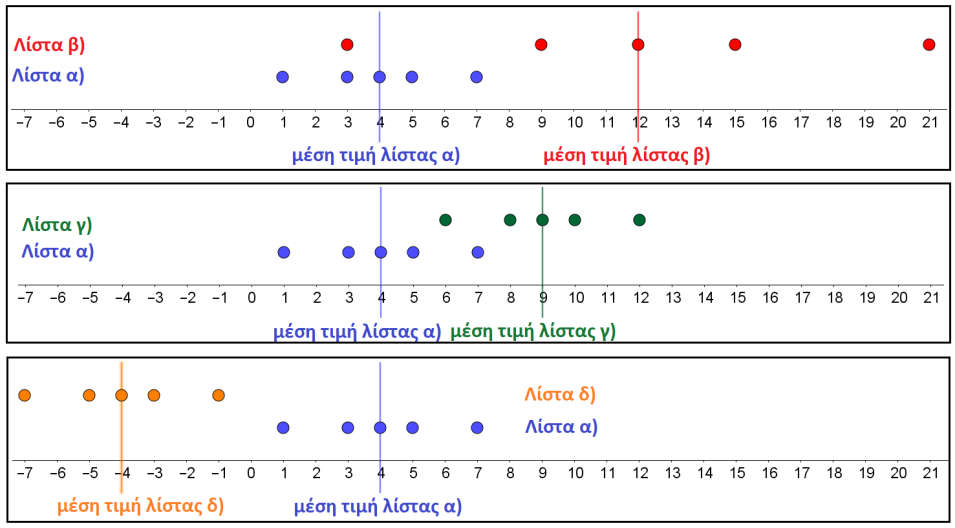
=32∙ και sβ = 3∙sα

= και sγ = sα

=(−1)2∙ και sδ = sα .

Σε κάθε μία από τις τρεις παρακάτω γραφικές παραστάσεις βρίσκονται τοποθετημένες, σε μία νοητή ευθεία, πάνω από μία αριθμογραμμή, οι τιμές της λίστας δεδομένων α). Επίσης στην 1η, τη 2η και την 3η γραφική παράσταση βρίσκονται οι τιμές της λίστας β), της λίστας γ) και της λίστας δ), αντίστοιχα, κατά τον ίδιο τρόπο, ώστε σε κάθε περίπτωση να μπορούν να γίνουν συγκρίσεις.

Επίσης, σε κάθε γραφική παράταση, παριστάνεται η «θέση» της μέσης τιμής κάθε λίστας με μία κατακόρυφη γραμμή, του αντίστοιχου χρώματος. Τα συμπεράσματα των παραπάνω αλγεβρικών συλλογισμών (για τις μεταβολές στη μέση τιμή, αλλά και τη διασπορά) φαίνονται σε αυτές τις γραφικές παραστάσεις.



##### Άσκηση 6

Η βαθμολογία 16 μαθητών σε ένα διαγώνισμα ήταν: 8, 15, 13, 20, 9, 13, 17, 19, 20, 9, 10,

10, 15, 13, 14, 17. Να υπολογίσετε:

* 1. Τα τρία μέτρα θέσης, μέση τιμή, διάμεσο και επικρατούσα τιμή.
  2. Το εύρος, την τυπική απόκλιση και τον συντελεστή μεταβολής. Λύση

1. Για τη μέση τιμή (το μέσο όρο των βαθμών) έχουμε:

= = 13,875

Για να βρούμε τη διάμεσο διατάσσουμε τα δεδομένα (τους βαθμούς των μαθητών) σε αύξουσα σειρά: 8, 9, 9, 10, 10, 13, 13, 13, 14, 15, 15, 17, 17, 19, 20, 20.

Η διάμεσος των 16 παρατηρήσεων είναι ο μέσος όρος της 8ης και της 9ης παρατήρησης,

δηλαδή δ = = 13,5

Η επικρατούσα τιμή είναι το 13, δηλαδή Μ0 = 13.

1. Το εύρος είναι R = 20 - 8 = 12 . Για την τυπική απόκλιση έχουμε:

s =  3, 85

(8 - 13, 875)2 + (9 - 13, 875)2 + ... + (20 - 13, 875)2

16

Για να διευκολύνουμε τους υπολογισμούς, μπορούμε να χρησιμοποιήσουμε ένα λογιστικό φύλλο για να κατασκευάσουμε τον παρακάτω πίνακα:

|  |  |  |  |
| --- | --- | --- | --- |
| παρατήρηση  xi | συχνότητα  νi |  | νi |
| 8 | 1 | 34,51563 | 34,51563 |
| 9 | 2 | 23,76563 | 47,53125 |
| 10 | 2 | 15,01563 | 30,03125 |
| 13 | 3 | 0,765625 | 2,296875 |
| 14 | 1 | 0,015625 | 0,015625 |
| 15 | 2 | 1,265625 | 2,53125 |
| 17 | 2 | 9,765625 | 19,53125 |
| 19 | 1 | 26,26563 | 26,26563 |
| 20 | 2 | 37,51563 | 75,03125 |
| Σύνολο: | | | 237,75 |

Οπότε, s =

 3, 85

Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε έτοιμα εργαλεία του λογιστικού φύλλου για να υπολογίσουμε την τυπική απόκλιση.

237, 75

16

Τέλος, ο συντελεστής μεταβολής είναι: CV = = 28%

##### Άσκηση 7

Η μέση τιμή ηλικίας των υπαλλήλων μιας εταιρείας είναι 32 χρόνια. Ποια θα είναι η μέση τιμή ηλικίας των ίδιων υπαλλήλων ύστερα από τρία χρόνια;

Λύση

Με δεδομένο ότι κανείς υπάλληλος δεν αποχωρεί και κανείς νέος δεν προσλαμβάνεται, εφόσον όλες οι ηλικίες θα έχουν αυξηθεί κατά 3 χρόνια, η μέση τιμή θα έχει αυξηθεί κι αυτή κατά 3 χρόνια και θα είναι 35 χρόνια.

##### Άσκηση 8

Οι βαθμοί στα Μαθηματικά 20 μαθητών της Β΄ τάξης ενός Λυκείου είναι:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 12 | 14 | 15 | 13 | 17 | 15 | 16 | 14 | 18 | 15 | 17 | 13 | 19 | 15 | 16 | 12 | 16 | 18 | 13 | 14 |

* 1. Να βρείτε τη μέση τιμή και την επικρατούσα τιμή.
  2. Να βρείτε τη διάμεσο.
  3. Να βρείτε το πρώτο και το τρίτο τεταρτημόριο.
  4. Να σχεδιάσετε το θηκόγραμμα.

Λύση

Για διευκόλυνση διατάσσουμε τις παρατηρήσεις σε αύξουσα σειρά:

12, 12, 13, 13, 13, 14, 14, 14, 15, 15, 15, 15, 16, 16, 16, 17, 17, 18, 18, 19.

**1)** H μέση τιμή είναι:

= = = 15,1

Η επικρατούσα τιμή είναι Μ0 = 15

1. Η διάμεσος είναι ο μέσος όρος της 10ης και 11ης παρατήρησης, δηλαδή:

## δ = = 15

## Το πρώτο τεταρτημόριο είναι ο μέσος όρος της 5ης και της 6ης παρατήρησης, ενώ το τρίτο τεταρτημόριο είναι ο μέσος όρος της 15ης και 16ης παρατήρησης. Δηλαδή έχουμε:

Q1 = = 13,5 και Q3 = = 16,5

1. Το ενδοτεταρτημοριακό εύρος είναι Q = Q3 - Q1 = 16, 5 - 13, 5 = 3.

Το διάστημα

[Q - 1, 5  Q , Q + 1, 5  Q]

είναι το [9, 21] στο οποίο περιλαμβάνονται

όλες οι τιμές (άρα δεν υπάρχουν ακραίες τιμές). Το θηκόγραμμα φαίνεται παρακάτω:

1 3



Στο θηκόγραμμα βλέπουμε την ελάχιστη τιμή (12), τη μέγιστη τιμή (19), τη διάμεσο (15) και το πρώτο και τρίτο τεταρτημόρια (13,5 και 16,5 αντιστοίχως).

**Πρόσθετο υλικό – Σύνθετες ασκήσεις**

##### Άσκηση 1

Η μέση επίδοση 17 αγοριών και 13 κοριτσιών στο μάθημα των Μαθηματικών μιας τάξης είναι 16,8. Η μέση επίδοση των κοριτσιών είναι 15,6. Να βρείτε τη μέση επίδοση των αγοριών

Λύση

Αν ονομάσουμε α τη μέση επίδοση των αγοριών και κ τη μέση επίδοση των κοριτσιών,

τότε το άθροισμα των βαθμολογιών των αγοριών είναι 17  α και το άθροισμα των βαθμο-

λογιών των κοριτσιών είναι 13∙15,6 (αυτό συμβαίνει επειδή μπορούμε να υποθέσουμε ότι

κάθε αγόρι έχει επίδοση ίση με τη μέση επίδοση των αγοριών και ότι κάθε κορίτσι έχει επίδοση

ίση με τη μέση επίδοση των κοριτσιών).

Τότε θα ισχύει :

= 16,8

Λύνοντας την εξίσωση βρίσκουμε τη μέση επίδοση των αγοριών α  17,7.

##### Άσκηση 2

Σε ένα Λύκειο υπάρχουν 500 μαθητές. Η Α΄ τάξη έχει 200 μαθητές με μέσο όρο ηλικίας 15,7 χρόνια, ενώ η Β΄ τάξη έχει 180 μαθητές με μέσο όρο ηλικίας 16,9 χρόνια. Οι υπόλοιποι μαθητές της Γ΄ τάξης έχουν μέσο όρο ηλικίας 17,7 χρόνια. Να υπολογίσετε τον μέσο όρο ηλικίας όλων των μαθητών του σχολείου.

Λύση

Το πλήθος των μαθητών της Γ΄ τάξης είναι 500 − 200 + 180  = 120 .

Το άθροισμα των ηλικιών των μαθητών της Α΄ τάξης θα είναι 200 15, 7

, το άθροισμα των

ηλικιών των μαθητών της Β΄ τάξης θα είναι 180 ∙16, 9

και το άθροισμα των ηλικιών των

μαθητών της Γ΄ τάξης θα είναι 120  17, 7 . Οπότε η μέση ηλικία των μαθητών του

σχολείου θα είναι

 16,6

##### Άσκηση 3

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

Λύση

Το άθροισμα των τιμών όλων των παρατηρήσεων θα είναι 40 ∙ 20 = 800 και μετά τις

μεταβολές θα είναι 800 − 7  2 + 9  6 = 840 . Οπότε, η νέα μέση τιμή θα είναι

= 21

##### Άσκηση 4

Η τυπική απόκλιση ενός δείγματος με ν παρατηρήσεις είναι ίση με μηδέν. Ποιο συμπέρασμα προκύπτει για τις παρατηρήσεις;

Λύση

Η τυπική απόκλιση είναι η τετραγωνική ρίζα της διακύμανσης (s2), η οποία ορίζεται ως "η μέση τιμή των τετραγώνων των διαφορών της μέσης τιμής των παρατηρήσεων από τις παρατηρήσεις". Για να είναι η τυπική απόκλιση ίση με 0, και η διακύμανση θα είναι ίση με 0, άρα κάθε παρατήρηση θα διαφέρει από τη μέση τιμή κατά 0 (ας θυμηθούμε ότι αν το άθροισμα μη αρνητικών πραγματικών είναι μηδέν τότε όλοι οι προσθετέοι είναι μηδέν). Δηλαδή, όλες οι παρατηρήσεις θα είναι μεταξύ τους ίσες.

##### Άσκηση 5

Η μέση τιμή 6 διαδοχικών ακέραιων αριθμών είναι 7,5. Να υπολογίσετε την τυπική τους απόκλιση.

Λύση

Εύκολα βρίσκουμε ότι οι 6 διαδοχικοί ακέραιοι που έχουν μέση τιμή 7,5 είναι οι: 5, 6, 7, 8, 9, 10.

Η τυπική τους απόκλιση θα είναι:

s =  1, 7

(5 - 7, 5)2 + (6 - 7, 5)2 + (7 - 7, 5)2 + (8 - 7, 5)2 + (9 - 7, 5)2 + (10 - 7, 5)2

6

Σχόλιο: Ένας τρόπος να βρούμε τους 6 διαδοχικούς ακεραίους με μέση τιμή 7,5, είναι να ονομάσουμε με ν τον μικρότερο, οπότε οι επόμενοι είναι ν+1, ν+2, ν+3, ν+4, ν+5 και να

λύσουμε την εξίσωση: ν + (ν + 1) + (ν + 2) + (ν + 3) + (ν + 4) + (ν + 5) = 7, 5

## 6

Ένας άλλος (προφανώς πιο εύκολος) είναι να "μοιράσουμε" έξι διαδοχικούς αριθμούς έτσι ώστε το 7,5 να βρίσκεται στη μέση τους, άρα θα είναι τρεις κάτω από το 7,5 και τρεις πάνω από αυτό.

##### Άσκηση 6

Αν σε μία τάξη ο μέσος όρος της βαθμολογίας ν1 αγοριών είναι και ο μέσος όρος της βαθμολογίας ν2 κοριτσιών είναι , να αποδείξετε ότι ο μέσος όρος της βαθμολογίας όλων των παιδιών της τάξης είναι: =

Λύση

Για τις ανάγκες του συγκεκριμένου προβλήματος, μπορούμε να υποθέσουμε ότι κάθε αγόρι έχει βαθμολογία και κάθε κορίτσι έχει . Άρα το άθροισμα των βαθμών των

αγοριών είναι ν1 , το άθροισμα των βαθμών των κοριτσιών είναι ν2 και το άθροισμα

των βαθμών όλων των μαθητών/τριών είναι ν1 + ν2 .

Αφού όλοι οι μαθητές είναι ν1 + ν2 , ο μέσος όρος της βαθμολογίας θα είναι:

=

##### Άσκηση 7

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

Λύση

Αν οι τιμές των παρατηρήσεων είναι xi  με αντίστοιχες συχνότητες νi και το σύνολο των

παρατηρήσεων είναι ν, τότε:

=

Οπότε: = + + … + = x1 + x2 + … + xκ = f1x1 + f2x2 + … + fκxκ

Δηλαδή η μέση τιμή είναι ίση με το άθροισμα των γινομένων των τιμών της μεταβλητής επί τις αντίστοιχες σχετικές συχνότητες.

##### Άσκηση 8

Σε ένα εργοστάσιο με 100 εργαζόμενους η μέση τιμή των αμοιβών τους είναι 900€. Οι 40 από αυτούς πληρώνονται με μισθό μικρότερο της μέσης τιμής και οι μισθοί τους έχουν μέση τιμή 800€. Αν οι αποδοχές των εργαζομένων με μισθό μικρότερο της μέσης τιμής αυξηθούν και γίνουν όσο η μέση τιμή, τότε ποια θα είναι η νέα μέση τιμή των αμοιβών των 100 εργαζομένων;

Λύση

Το άθροισμα των αμοιβών των 100 εργαζομένων πριν οποιαδήποτε αύξηση είναι

100  900 = 90.000 €. Αφού οι αποδοχές των 40 χαμηλότερα αμειβόμενων εργαζομένων θα γίνουν όσο η μέση τιμή, δηλαδή 900€, αυτοί οι εργαζόμενοι θα έχουν μια μέση αύξηση κατά 100€, άρα αθροιστικά αύξηση 40  100 = 4.000€. Οπότε το άθροισμα των αμοιβών των 100 εργαζομένων διαμορφώνεται στις 94.000 €. Έτσι, η νέα μέση τιμή θα είναι

## = 940 € .

**ΕΝΟΤΗΤΑ 2.4 : ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ ΚΑΙ ΕΦΑΡΜΟΓΕΣ**

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

Το σύνολο των μαθητών/τριών μιας πόλης, ρωτήθηκαν για τον χρόνο που κάνουν να πάνε από το σπίτι στο σχολείο. Το 50% περίπου έδωσε απαντήσεις που ήταν από 12 λεπτά και πάνω, ενώ το 16% περίπου έδωσε απαντήσεις που ήταν από 10 λεπτά και κάτω. Υποθέτουμε ότι η κατανομή του χρόνου της διαδρομής σπίτι-σχολείο των μαθητών είναι κανονική.

1. Να εκτιμήσετε τον μέσο χρόνο διαδρομής σπίτι-σχολείο, των μαθητών και την τυπική απόκλιση του χρόνου διαδρομής τους, σύμφωνα με τις απαντήσεις τους.
2. Αν οι μαθητές/τριες της πόλης είναι 4.000, να εκτιμήσετε πόσοι/ες απάντησαν ότι έχουν χρόνο διαδρομής σπίτι-σχολείο μεταξύ 14 και 16 λεπτών;

Λύση

1. Σε μια κανονική κατανομή το 50% περίπου του πληθυσμού είναι πάνω από μ. Οπότε, αφού το 50% περίπου έδωσε απαντήσεις που ήταν από 12 λεπτά και πάνω, θα είναι μ=12. Επιπλέον, το 100%-68%=32% περίπου του πληθυσμού αναμένεται να έδωσαν απαντήσεις κάτω από μ-σ και πάνω από μ+σ, οπότε το 16% των απαντήσεων αναμένεται να είναι κάτω από μ-σ. Οπότε, θα είναι μ-σ=10, και άρα σ=2. Δηλαδή εκτιμούμε ότι μ=12 λεπτά και σ=2 λεπτά.
2. 14=μ+σ και 16=μ+2σ. Οπότε μεταξύ 14 και 16 αναμένεται ότι απάντησαν περίπου

95% - 68% = 13, 5%

2

των μαθητών

##### Άσκηση 2

Υποθέτουμε ότι το βάρος των μαθητών λυκείου ακολουθεί κανονική κατανομή και παίρνουμε ένα μεγάλο δείγμα μαθητών λυκείου. Το 50% των μαθητών του δείγματος έχουν βάρος το πολύ 65 Kg, ενώ περίπου το 47,5% αυτών έχουν βάρος από 65 Kg έως 75 Kg.

1. Να εκτιμήσετε τη μέση τιμή και την τυπική απόκλιση του βάρους των μαθητών του δείγματος.
2. Να εκτιμήσετε το ποσοστό των μαθητών του δείγματος που έχουν βάρος από 55 Kg έως 70 Kg.

Λύση

1. Αφού το 50% των μαθητών του δείγματος έχουν βάρος το πολύ 65 Kg, θα είναι μ=65kg.

Αφού το = 47, 5% έχουν βάρος από 65 Kg έως 75 Kg θα είναι 75=μ+2σ, και άρα σ=5.

Δηλαδή μπορούμε να εκτιμήσουμε ότι μ=65 kg και σ=5 kg.

1. Επειδή 55=μ−2σ και 70=μ+σ, στο διάστημα (55,70), δηλαδή στο (μ−2σ,μ+σ) αναμένεται

να έχουν βάρος περίπου το 68% + 95% - 68% = 81, 5%

2

των μαθητών.

##### Άσκηση 3

1. Να αντιστοιχίσετε τις γραφικές παραστάσεις των συναρτήσεων που είναι μοντέλα κανονικών κατανομών και περιέχονται στον παρακάτω πίνακα με τα ζεύγη τιμών των παραμέτρων σ και μ που ακολουθούν. Ο κατακόρυφος άξονας των συστημάτων συντεταγμένων ακολουθεί την ίδια κλίμακα σε όλες τις περιπτώσεις.

|  |  |
| --- | --- |
| **Πίνακας** | |
| **(α)** | **(β)** |
| **(γ)** | **(δ)** |

**Α.** μ  1, σ  1 , **Β.** μ  1, σ  0, 3 , **Γ.** μ  0, σ  0,75

1. Να συγκρίνετε την τιμή του σ στο σχήμα (γ) με το 0,3. Λύση
2. α–Β, β–Γ, γ–Δ, δ–Α.

**Δ.** μ  0, σ  1

1. Το εύρος είναι περίπου 4 (μεταξύ του −2 και του 2 φαίνεται να είναι σχεδόν το 100% των παρατηρήσεων). Άρα, 6σ  4 , οπότε σ  0,66 , άρα σ > 0,3 .

# ΕΝΟΤΗΤΑ 2.5 : ΠΙΝΑΚΕΣ ΣΥΝΑΦΕΙΑΣ ΚΑΙ ΡΑΒΔΟΓΡΑΜΜΑΤΑ

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

Για ποια από τα επόμενα ζεύγη μεταβλητών θα κατασκευάζατε πίνακα συνάφειας; Για εκείνα που θα επιλέξετε, να προσδιορίσετε την ανεξάρτητη και την εξαρτημένη μεταβλητή για την έρευνά σας.

α) φύλο και επίπεδο εκπαίδευσης

β) φύλο και χρήση διαδικτύου (σε ώρες/ημέρα)

γ) βαθμός στην Άλγεβρα και ώρες μελέτης (ανά εβδομάδα) δ) ικανοποίηση από το σχολείο και φύλο

ε) ευτυχία στο γάμο και πίστη στα θαύματα

Λύση

Οι πίνακες συνάφειας χρησιμοποιούνται για την αναζήτηση σχέσεων ανάμεσα σε δύο ποιοτικές μεταβλητές. Οπότε, για τις (α) και (ε) μπορούμε να χρησιμοποιήσουμε πίνακες συνάφειας, ενώ για τις (β) και (γ) όχι, εφόσον η χρήση διαδικτύου, ο βαθμός και οι ώρες μελέτης είναι ποσοτικές μεταβλητές. Για την περίπτωση (δ), το φύλο είναι ποιοτική μεταβλητή, αλλά η ικανοποίηση θα μπορούσε να είναι ποιοτική (αν εκφραζόταν με το καθόλου/λίγο/αρκετά/πολύ ή με κάτι παρόμοιο) ή θα μπορούσε να είναι ποσοτική (αν εκφραζόταν πχ σε μια ποσοστιαία κλίμακα). Οπότε αντιστοίχως, θα μπορούσαμε ή όχι να χρησιμοποιήσουμε πίνακες συνάφειας.

##### Άσκηση 2

Σε μια έρευνα του τρόπου ζωής και των καθημερινών συνηθειών στη χώρα μας συλλέχθηκε ένα δείγμα 833 ατόμων απ' όλη την επικράτεια, προκειμένου να εξετάσουμε εάν υπάρχει σχέση μεταξύ του επίπεδου εκπαίδευσης (δευτεροβάθμια, αν έχουν ολοκληρώσει το Λύκειο και τριτοβάθμια, αν έχουν ολοκληρώσει τις σπουδές τους σε ΑΕΙ) και της συστηματικής σωματικής άσκησης τους (ναι: τουλάχιστον 2 φορές την εβδομάδα και όχι: το πολύ μια φορά την εβδομάδα) σε όλη τη διάρκεια ενός έτους. Τα δεδομένα οργανώθηκαν στον παρακάτω πίνακα συνάφειας:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | **Άσκηση** | | |
| **Ναι** | **Όχι** | **Σύνολο** |
| **Εκπαίδευση** | **Δευτεροβάθμια** | **366** | **75** |  |
| **Τριτοβάθμια** | **303** | **89** |  |
|  | **Σύνολο** |  |  |  |

**α)** Να συμπληρώσετε στον παραπάνω πίνακα τα κενά κελιά και να κατασκευάσετε το στοιβαγμένο ραβδόγραμμα των συχνοτήτων του πίνακα συνάφειας.

**β)** Να κατασκευάσετε τον πίνακα συνάφειας σχετικών συχνοτήτων ως προς το επίπεδο εκπαίδευσης και το αντίστοιχο ομαδοποιημένο ραβδόγραμμα του.

**γ)** Μπορείτε να υποστηρίξετε την άποψη ότι το επίπεδο εκπαίδευσης επηρεάζει το αν κάποιο άτομο ασκείται συστηματικά ή όχι;

Λύση

**α)** Ο πίνακας συνάφειας φαίνεται συμπληρωμένος παρακάτω:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Άσκηση | | |
| Ναι | Όχι | Σύνολο |
| Εκπαίδευση | Δευτεροβάθμια | 366 | 75 | 441 |
| Τριτοβάθμια | 303 | 89 | 392 |
|  | Σύνολο | 669 | 164 | 833 |

Το στοιβαγμένο ραβδόγραμμα των συχνοτήτων αυτού του πίνακα είναι το εξής:

Άσκηση Ναι Άσκηση Όχι

Τριτοβάθμια

Δευτεροβάθμια

500

400

300

200

100

0

επίπεδο εκπαίδευσης και

συστηματική άσκηση

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 75 |  | | |
|  |  | 89 |  |
| 366 |
|  |  | 303 |  |
|  |  |  |
|  |  |  |

**β)** Ο πίνακας συνάφειας σχετικών συχνοτήτων ως προς το επίπεδο εκπαίδευσης είναι ο εξής:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Άσκηση | | |
| Ναι | Όχι | Σύνολο |
| Εκπαίδευση | Δευτεροβάθμια | 82,99% | 17,01% | 100,00% |
| Τριτοβάθμια | 77,30% | 22,70% | 100,00% |
|  | Σύνολο |  |  |  |

και το αντίστοιχο ομαδοποιημένο ραβδόγραμμα του:

Άσκηση Ναι Άσκηση Όχι

90,00%

80,00%

70,00%

60,00%

50,00%

40,00%

30,00%

20,00%

10,00%

0,00%

επίπεδο εκπαίδευσης και

συστηματική άσκηση

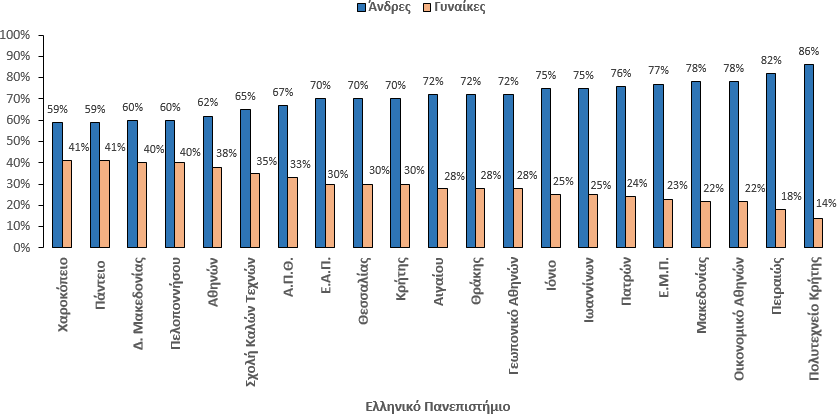
**γ)** Κυρίως από τον πίνακα συνάφειας και το ραβδόγραμμα του (β) ερωτήματος, φαίνεται ότι εκείνοι που έχουν ολοκληρώσει σπουδές σε ΑΕΙ ασκούνται ελαφρώς λιγότερο από εκείνους που έχουν ολοκληρώσει το Λύκειο.

Δευτεροβάθμια

Τριτοβάθμια

Εκπαίδευση

##### Άσκηση 3

Το επόμενο ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων δίνει τα ευρήματα μιας μελέτης που έγινε το 2016 στο Οικονομικό Πανεπιστήμιο Αθηνών για τη θέση των γυναικών στα Ελληνικά Πανεπιστήμια, η οποία δημοσιεύθηκε σε κυριακάτικη εφημερίδα. Πιο συγκεκριμένα, δίνονται τα ποσοστά ανδρών και γυναικών μόνιμων διδασκόντων στα Πανεπιστήμια.

**α)** Σε ποια Πανεπιστήμια παρατηρείται το μεγαλύτερο ποσοστό γυναικών και σε ποια το χαμηλότερο;

**β)** Τη συγκεκριμένη χρονιά δίδαξαν συνολικά 8.483 μόνιμοι διδάσκοντες εκ των οποίων οι 5.894 ήταν άνδρες και οι 2.589 ήταν γυναίκες. Να βρείτε σε ποια πανεπιστήμια το ποσοστό των γυναικών ξεπέρασε το γενικό μέσο ποσοστό τους;

**γ)** Ποιος πιστεύετε ότι μπορεί να ήταν ο τίτλος του άρθρου και για ποιο λόγο μια εφημερίδα το ανέδειξε ως ένα από τα κύρια κοινωνικά άρθρα της;

Λύση

**α)** Το μεγαλύτερο ποσοστό γυναικών μόνιμων διδασκόντων παρατηρείται στο Χαροκόπειο και στο Πάντειο (41%) και το χαμηλότερο στο Πολυτεχνείο Κρήτης (14%)

**β)** Στο σύνολο των Ελληνικών Πανεπιστημίων, το ποσοστό των γυναικών μόνιμων

διδασκόντων είναι  0, 30519  30, 5% . Οπότε, το ποσοστό των γυναικών ξεπέρασε

το γενικό μέσο ποσοστό τους στα εξής Πανεπιστήμια: Χαροκόπειο, Πάντειο, Δυτικής Μακεδονίας, Πελοποννήσου, Αθηνών, Σχολή Καλών Τεχνών και ΑΠΘ.

**γ)** Τίτλοι του άρθρου θα μπορούσαν να είναι: "Η συμμετοχή των γυναικών ως διδάσκουσες στο Πανεπιστήμιο", ή "Το Ελληνικό Πανεπιστήμιο ανδροκρατείται", ή "Μόνο 1 στις 3 διδάσκοντες στο Πανεπιστήμιο είναι γυναίκα". Το σημαντικό κοινωνικό θέμα που αναδεικνύεται από αυτή την έρευνα, είναι το χαμηλό ποσοστό με το οποίο αντιπροσωπεύονται οι γυναίκες μεταξύ των διδασκόντων στα Ελληνικά Πανεπιστήμια. Ένα τέτοιο άρθρο θα μπορούσε να αποτελέσει αφορμή για συζήτηση σχετικά με τις αιτίες και τις παραμέτρους αυτού του προβλήματος. Μερικά ερωτήματα τα οποία θα μπορούσαν να αντιμετωπιστούν σε μια τέτοια συζήτηση είναι τα εξής:

Συσχετίζεται ή όχι το ποσοστό των γυναικών μεταξύ των αποφοίτων των Πανεπιστημίων με το ποσοστό των γυναικών μεταξύ των διδασκόντων των Πανεπιστημίων;

Ποιοι λόγοι μπορεί να εμποδίζουν τις γυναίκες να ακολουθήσουν ακαδημαϊκή καριέρα; Πώς βιώνουν την ακαδημαϊκή καριέρα οι γυναίκες που την έχουν επιλέξει; Τι δυσκολίες αντιμετωπίζουν;

Στη μελέτη κάποιων από αυτά τα ερωτήματα (πχ στο πρώτο και ίσως λιγότερο στο δεύτερο) η στατιστική θα μπορούσε να αξιοποιηθεί. Σε κάποια άλλα (όπως στο τρίτο) ίσως να είναι χρήσιμη μια ποιοτική προσέγγιση, με συνεντεύξεις γυναικών, συνδέσεις με άλλα κείμενα και προσεγγίσεις κλπ.

##### Άσκηση 4

Στη Μυτιλήνη, ο δήμος εγκατέστησε ένα καινοτόμο πρόγραμμα ηλεκτρονικών υπηρεσιών του δήμου από το οποίο εξυπηρετούνται οι κάτοικοι του νησιού, ιδιοκτήτες και ενοικιαστές σπιτιών. Η ικανοποίηση των κατοίκων του νησιού από τις ηλεκτρονικές υπηρεσίες του δήμου φαίνονται από τα παρακάτω ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων (α) και (β).

**α)** Ποιο από τα δύο ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων αντιστοιχεί στο σύνολο των παρατηρήσεων του δείγματος; Τι δείχνει το άλλο ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων;

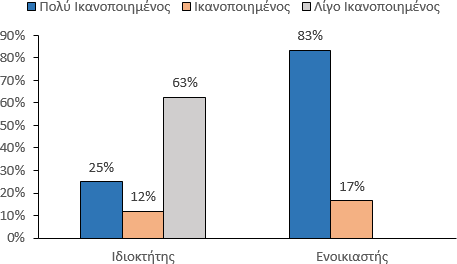
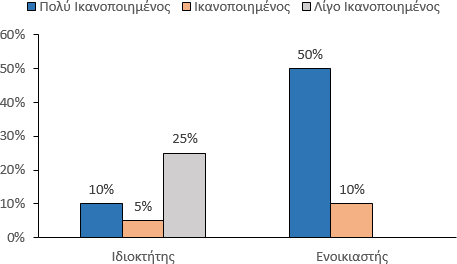
**β)** Ποιο ποσοστό του δείγματος είναι ιδιοκτήτες και ποιο ενοικιαστές σπιτιών;

**γ)** Ποιο ποσοστό του δείγματος είναι πολύ ικανοποιημένοι από τις ηλεκτρονικές υπηρεσίες του δήμου;

**δ)** Τι εκφράζουν τα ποσοστά 63% και 83% στο σχήμα (β);

**ε)** Αν το μέγεθος του δείγματος είναι 200 κάτοικοι, να βρείτε πόσοι είναι οι πολύ ικανοποιημένοι ιδιοκτήτες και πόσοι οι ικανοποιημένοι ενοικιαστές σπιτιών στο νησί.

*Σχήμα (α) Σχήμα (β)*



Λύση

**α)** Στο σύνολο των παρατηρήσεων του δείγματος αντιστοιχεί το ομαδοποιημένο ραβδόγραμμα του σχήματος (α) όπου το άθροισμα των σχετικών συχνοτήτων για τους ιδιοκτήτες και τους ενοικιαστές είναι 100% (10%+5%+25%+50%+10%=100%). Σο σχήμα (β) το ραβδόγραμμα δείχνει το ποσοστό ικανοποίησης σε κάθε μία από τις δύο κατηγορίες (25%+12%+63%=100% για τους ιδιοκτήτες, 83%+17%=100% για τους ενοικιαστές).

**β)** Όπως προκύπτει από το σχήμα (α), ιδιοκτήτες είναι το 40% του δείγματος (10+5+25) ενώ ενοικιαστές είναι το 60% του δείγματος (50+10).

γ) Όπως προκύπτει από το σχήμα (α), πολύ ικανοποιημένοι είναι το 10%+50%=60% του δείγματος

**δ)** Στο (β), το 63% δείχνει το ποσοστό των λίγο ικανοποιημένων ιδιοκτητών στο σύνολο των ιδιοκτητών και το 83% δείχνει το ποσοστό των πολύ ικανοποιημένων ενοικιαστών στο σύνολο των ενοικιαστών.

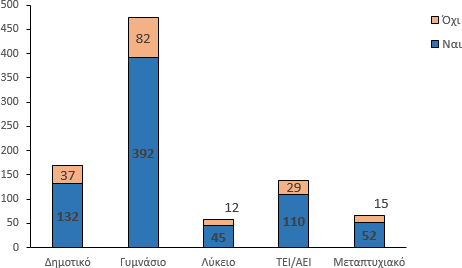
**ε)** Αν το δείγμα είναι 200 κάτοικοι, οι πολύ ικανοποιημένοι ιδιοκτήτες είναι 10% 200 = 20

άτομα και οι ικανοποιημένοι ενοικιαστές είναι 10% 200 = 20

άτομα.

##### Άσκηση 5

Πρόσφατα πραγματοποιήθηκε μια έρευνα προκειμένου να ερευνηθεί εάν υπάρχει σχέση μεταξύ της πίστης του ανθρώπου στα θαύματα και του επιπέδου εκπαίδευσης που έχει αποκτήσει. Τα αποτελέσματα της έρευνας συνοψίζονται στο επόμενο στοιβαγμένο ραβδόγραμμα συχνοτήτων.



**α)** Τι ποσοστό αποφοίτων Γυμνασίου, Λυκείου και ΑΕΙ/ΤΕΙ έχουμε στο δείγμα μας;

**β)** Τι ποσοστό αποφοίτων Λυκείου φαίνεται να πιστεύουν στα θαύματα και τι ποσοστό αποφοίτων ΤΕΙ/ΑΕΙ φαίνεται ότι δεν πιστεύουν στα θαύματα;

**γ)** Να μετατρέψετε το παραπάνω στοιβαγμένο ραβδόγραμμα συχνοτήτων σε ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς την πίστη στα θαύματα.

**δ)** Με βάση τα αποτελέσματα της συγκεκριμένης έρευνας φαίνεται να υπάρχει κάποια σχέση μεταξύ του επιπέδου εκπαίδευσης και της πίστης στα θαύματα;

Λύση

Για την οργάνωση των δεδομένων μας θα μπορούσαμε να κατασκευάσουμε τον πίνακα συνάφειας:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Πίστη στα θαύματα | |  |
| Ναι | Όχι | Σύνολο |
| Εκπαίδευση | Δημοτικό | 132 | 37 | 169 |
| Γυμνάσιο | 392 | 82 | 474 |
| Λύκειο | 45 | 12 | 57 |
| ΤΕΙ/ΑΕΙ | 110 | 29 | 139 |
| Μεταπτυχιακό | 52 | 15 | 67 |
|  | Σύνολο | 731 | 175 | 906 |

**α)** Οι απόφοιτοι Γυμνασίου είναι

474

906

 0, 523  52,3%

του δείγματος, οι απόφοιτοι

57

Λυκείου είναι

906

 0, 063  6,3%

του δείγματος και οι απόφοιτοι ΑΕΙ/ΤΕΙ είναι

139  0,153  15, 3% του δείγματος.

906

**β)** Το ποσοστό αποφοίτων Λυκείου που φαίνεται να πιστεύουν στα θαύματα είναι

45  0, 789  78, 9% . Το ποσοστό αποφοίτων ΤΕΙ/ΑΕΙ που φαίνεται ότι δεν πιστεύουν

57

στα θαύματα είναι

29

139

 0, 209  20, 9% .

**γ)** Ο πίνακας συνάφειας σχετικών συχνοτήτων ως προς την πίστη στα θαύματα είναι ο εξής:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Πίστη στα θαύματα | |  |
| Ναι | Όχι | Σύνολο |
| Εκπαίδευση | Δημοτικό | 18,1% | 21,1% |  |
| Γυμνάσιο | 53,6% | 46,9% |  |
| Λύκειο | 6,2% | 6,9% |  |
| ΤΕΙ/ΑΕΙ | 15,0% | 16,6% |  |
| Μεταπτυχιακό | 7,1% | 8,6% |  |
|  | Σύνολο | 100,0% | 100,0% |  |

Το ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς την πίστη στα θαύματα φαίνεται παρακάτω:

Δημοτικό

Γυμνάσιο

Λύκειο

ΤΕΙ/ΑΕΙ Μεταπτυχιακό

Εκπαίδευση

Πίστη στα θαύματα Όχι

Πίστη στα θαύματα Ναι

60,0%

50,0%

40,0%

30,0%

20,0%

10,0%

0,0%

επίπεδο εκπαίδευσης και πίστη στα θαύματα

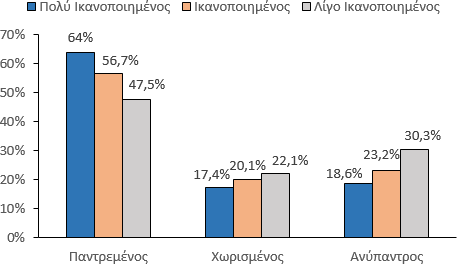
**δ)** Από το ομαδοποιημένο ραβδόγραμμα του (ε) ερωτήματος φαίνεται μια μεγαλύτερη συμμετοχή στο "όχι" από ότι στο "ναι" για τα τρία υψηλότερα επίπεδα εκπαίδευσης, αλλά και για το Δημοτικό. Αντίστροφη είναι η εικόνα για τους απόφοιτους Γυμνασίου. Αν κατασκευάσουμε τον πίνακα συνάφειας ως προς το επίπεδο εκπαίδευσης, αυτό φαίνεται από τα μεγαλύτερα ποσοστά του "όχι" που συγκεντρώνουν οι τρεις υψηλότερες βαθμίδες (21,1%, 20,9% και 22,4%) σε σύγκριση με το ποσοστό του "όχι" που αφορά το σύνολο των συμμετεχόντων (19,3%).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Πίστη στα θαύματα | |  |
| Ναι | Όχι | Σύνολο |
| Εκπαίδευση | Δημοτικό | 78,1% | 21,9% | 100,0% |
| Γυμνάσιο | 82,7% | 17,3% | 100,0% |
| Λύκειο | 78,9% | 21,1% | 100,0% |
| ΤΕΙ/ΑΕΙ | 79,1% | 20,9% | 100,0% |
| Μεταπτυχιακό | 77,6% | 22,4% | 100,0% |
|  | Σύνολο | 80,7% | 19,3% | 100,0% |

Αντίστοιχη παρατήρηση ισχύει και για το Δημοτικό (21,9%), ενώ η αντίστροφη παρατήρηση ισχύει για το Γυμνάσιο (17,3%).

Ωστόσο, αν κανείς παραμέριζε τις λεπτομέρειες, θα μπορούσε να πει ότι περίπου 4 στους 5 συμμετέχοντες πιστεύουν στα θαύματα ανεξάρτητα από το επίπεδο εκπαίδευσης. Οι μικροδιαφορές που υπάρχουν ανά επίπεδο εκπαίδευσης δεν αλλάζουν την γενική εικόνα.

##### Άσκηση 6

Μια κυριακάτικη εφημερίδα περιέχει ένα άρθρο σχετικά με την ικανοποίηση των ανθρώπων από την εργασία τους και την οικογενειακή τους κατάσταση μαζί με το παρακάτω γράφημα. Οι συντάκτες του άρθρου καταλήγουν στο συμπέρασμα ότι ο γάμος κάνει τους ανθρώπους περισσότερο ικανοποιημένους με την εργασία τους, δεδομένου ότι το 64% των πολύ ικανοποιημένων ανθρώπων είναι παντρεμένοι, ενώ μόνο το 18,6% δεν έχουν παντρευτεί ποτέ.

**α)** Συμφωνείτε με το συμπέρασμα του συντάκτη του άρθρου;

**β)** Συμφωνείτε με την αιτιολόγηση του συντάκτη του άρθρου;

Λύση

**α)** Από τα δεδομένα που παρουσιάζονται εδώ φαίνεται κάποια συσχέτιση μεταξύ οικογενειακής κατάστασης και ικανοποίησης από την εργασία. Ωστόσο, το συμπέρασμα ότι ο γάμος κάνει τους ανθρώπους περισσότερο ικανοποιημένους με την εργασία τους εμπεριέχει ή υπονοεί μια αιτιώδη σχέση: ο γάμος είναι η αιτία της ικανοποίησης από την εργασία ή κατά κάποιο τρόπο την προκαλεί. Προφανώς αυτό δεν ισχύει, διότι μια σχέση ανάμεσα στα δύο δεν είναι κατ' ανάγκην αιτιώδης. Θα μπορούσε να συμβαίνει το αντίστροφο: η ικανοποίηση από την εργασία να προκαλεί τον γάμο. Ή τίποτα από τα δύο δεν προκαλεί το άλλο: θα μπορούσαν και τα δύο να συνδέονται αιτιωδώς με κάτι άλλο, πχ. το πόσο καλά αμειβόμενη είναι η εργασία.

**β)** Στην αιτιολόγησή του ο συντάκτης συγκρίνει το 64% των πολύ ικανοποιημένων που είναι οι παντρεμένοι με το 18,6% των πολύ ικανοποιημένων που είναι οι ανύπαντροι. Αυτή η σύγκριση όμως δεν είναι ασφαλής, εφόσον φαίνεται ότι το πλήθος των παντρεμένων στο δείγμα είναι σαφώς μεγαλύτερο από το πλήθος των άλλων κατηγοριών.

Από τα δεδομένα φαίνεται να υπάρχει μια σχέση ανάμεσα στην ικανοποίηση από την εργασία και την οικογενειακή κατάσταση των συμμετεχόντων στην έρευνα. Πιο συγκεκριμένα, μεταξύ των παντρεμένων φαίνεται να αυξάνεται το ποσοστό όταν από τους λίγο ικανοποιημένους από την εργασία περνάμε στους πολύ ικανοποιημένους (από το 47,5% στο 56,7% και μετά στο 64%). Το αντίθετο συμβαίνει μεταξύ των ανύπαντρων (από το 30,3% στο 23,2% και μετά στο 18,6%). Ας σημειώσουμε ότι, αν η ικανοποίηση από την εργασία δεν συσχετιζόταν με την οικογενειακή κατάσταση, θα περιμέναμε τα ποσοστά των τριών βαθμίδων ικανοποίησης να είναι περίπου ίδια στους παντρεμένους (δηλαδή, οι ράβδοι στην ομάδα παντρεμένοι να έχουν περίπου ίδιο ύψος). Και αντίστοιχα για τις δύο άλλες ομάδες (αν και τα ύψη των ράβδων στους ανύπαντρους αναμένεται να είναι χαμηλότερα από των παντρεμένων, εφόσον στο σύνολο του δείγματος οι ανύπαντροι είναι λιγότεροι από τους παντρεμένους)

**Πρόσθετο Υλικό – θέματα για διερεύνηση**

**1)** Το φθινόπωρο του 1973 παρατηρήθηκε το ακόλουθο παράδοξο κατά την εισαγωγή μεταπτυχιακών φοιτητών στο Πανεπιστήμιο του Berkeley στην Καλιφόρνια. Τα στοιχεία της έρευνας έδειξαν ότι οι άνδρες που υπέβαλλαν αίτηση ήταν πιο πιθανό από ό, τι οι γυναίκες να γίνουν δεκτοί. Στον πίνακα (α) δίνονται τα στοιχεία αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές τη χρονιά του 1973. Στον πίνακα (β) δίνονται τα στοιχεία αιτούντων και εισακτέων ανδρών και γυναικών σε έξι διακεκριμένα τμήματα του Πανεπιστημίου του Berkeley για μεταπτυχιακές σπουδές τη χρονιά του 1973.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό**  **εισακτέων** |
| **Άνδρες** | 8442 | 3714 | …… |
| **Γυναίκες** | 4321 | 1512 | …... |

*Πίνακας (α): Πίνακας αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές στο Πανεπιστήμιο του Berkeley*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Άνδρες** | | | **Γυναίκες** | | |
| **Αιτούντες** | **Εισακτέοι** | **Ποσοστό**  **εισακτέων** | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό**  **εισακτέων** |
| **Τμήμα Α** | 825 | 512 | …….. | 108 | 89 | …….. |
| **Τμήμα Β** | 560 | 353 | …….. | 25 | 17 | …….. |
| **Τμήμα Γ** | 325 | 120 | …….. | 593 | 202 | …….. |
| **Τμήμα Δ** | 417 | 138 | …….. | 375 | 131 | …….. |
| **Τμήμα Ε** | 191 | 54 | …….. | 393 | 94 | …….. |
| **Τμήμα Ζ** | 373 | 23 | …….. | 341 | 24 | …….. |

*Πίνακας (β): Πίνακας αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές σε έξι διακεκριμένα τμήματα Πανεπιστήμιο του Berkeley*

**α)** Συμπληρώστε τα ποσοστά των εισακτέων του πίνακα (α). Τι διαπιστώνετε, συγκρίνοντας τα ποσοστά των ανδρών και των γυναικών εισακτέων;

**β)** Συμπληρώστε τα ποσοστά των εισακτέων του πίνακα (β) και συγκρίνετε:

**(ι)** τα ποσοστά των ανδρών εισακτέων στα τμήματα Γ και Ε με τα αντίστοιχα ποσοστά των γυναικών,

**(ιι**) τα ποσοστά των εισακτέων γυναικών στα τμήματα Α, Β, Δ και Ζ με τα αντίστοιχα ποσοστά των ανδρών.

**γ)** Τι διαπιστώνετε, συγκρίνοντας τα παραπάνω ποσοστά σε κάθε τμήμα του Πανεπιστήμιου με τα ποσοστά του ερωτήματος α; Υπάρχει κάποια σύγχυση;

**δ)** Σε τι είδους τμήματα ως προς την ανταγωνιστικότητα και το ποσοστό των εισακτέων φαίνεται να απευθύνονται οι γυναίκες και σε τι είδους τμήματα οι άνδρες;

Λύση

**α)** Ο πίνακας (α) συμπληρωμένος φαίνεται παρακάτω:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό**  **εισακτέων** |
| **Άνδρες** | 8442 | 3714 | 44% |
| **Γυναίκες** | 4321 | 1512 | 35% |

Από τον πίνακα αυτό φαίνεται ότι πράγματι, οι άνδρες που υπέβαλλαν αίτηση ήταν πιο πιθανό από ό, τι οι γυναίκες να γίνουν δεκτοί.

**β)** Ο πίνακας (β) συμπληρωμένος φαίνεται παρακάτω:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Άνδρες** | | | **Γυναίκες** | | |
| **Αιτούντες** | **Εισακτέοι** | **Ποσοστό**  **εισακτέων** | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό**  **εισακτέων** |
| **Τμήμα Α** | 825 | 512 | 62% | 108 | 89 | 82% |
| **Τμήμα Β** | 560 | 353 | 63% | 25 | 17 | 68% |
| **Τμήμα Γ** | 325 | 120 | 37% | 593 | 202 | 34% |
| **Τμήμα Δ** | 417 | 138 | 33% | 375 | 131 | 35% |
| **Τμήμα Ε** | 191 | 54 | 28% | 393 | 94 | 24% |
| **Τμήμα Ζ** | 373 | 23 | 6% | 341 | 24 | 7% |

**(ι)** τα ποσοστά των ανδρών εισακτέων στα τμήματα Γ και Ε είναι υψηλότερα από τα αντίστοιχα ποσοστά των γυναικών, κατά 3 και 4 ποσοστιαίες μονάδες αντιστοίχως.

**(ιι**) τα ποσοστά των εισακτέων γυναικών στα τμήματα Α, Β, Δ και Ζ είναι υψηλότερα από τα αντίστοιχα ποσοστά των ανδρών κατά 20, 5, 2 και 1 ποσοστιαίες μονάδες αντιστοίχως. **γ)** Στο σύνολο των 6 τμημάτων το ποσοστό των εισακτέων γυναικών είναι μικρότερο από εκείνο των ανδρών, παρά το ότι το ποσοστό των εισακτέων γυναικών είναι μεγαλύτερο από εκείνο των ανδρών σε 4 από τα 6 τμήματα. Αυτό οφείλεται στο ότι οι αριθμοί των γυναικών αιτουσών και εισακτέων για τα τμήματα Γ και Ε είναι από τους μεγαλύτερους αριθμούς που διαμορφώνουν το σύνολο των αιτουσών και των εισακτέων γυναικών.

Δηλαδή τα τμήματα Γ και Ε "βαραίνουν" στον υπολογισμό του ποσοστού περισσότερο από τα άλλα.

**δ)** Οι άνδρες φαίνεται να απευθύνονται σε όλα τα τμήματα, αλλά έχουν σχεδόν "αποκλειστικότητα" στο τμήμα Β. Ο γυναίκες απευθύνονται κυρίως στα Γ και Ε (όπου είναι περισσότερες από τους άνδρες υποψηφίους), λιγότερο στα Δ και Ζ και πολύ λίγο στα Α και Β.

Όσον αφορά το ποσοστό εισακτέων, τόσο οι άνδρες, όσο και οι γυναίκες έχουν υψηλά ποσοστά εισακτέων στα τμήματα Α και Β (αν και είναι μικρός ο αριθμός γυναικών εισακτέων στο Β) και πολύ χαμηλά στο Ζ. Γενικά, στα ποσοστά εισακτέων ανδρών και γυναικών ανά τμήμα δεν φαίνεται διαφορά τέτοια, ώστε να μιλήσει κανείς για σημαντικά διαφορετικές ευκαιρίες μεταξύ ανδρών και γυναικών.

# ΕΝΟΤΗΤΑ 2.6 : ΣΥΓΚΡΙΣΗ ΠΟΣΟΤΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΣΤΙΣ ΚΑΤΗΓΟΡΙΕΣ ΕΝΟΣ ΠΟΙΟΤΙΚΟΥ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟΥ

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

1) Στους επόμενους πίνακες δίνονται οι βαθμολογίες των μαθητών/τριών δύο τμημάτων της Β΄ τάξης ενός γενικού λυκείου σε μια γραπτή αξιολόγηση της Άλγεβρας:

|  |  |  |  |
| --- | --- | --- | --- |
| **Τμήμα Β1** | | | |
| **20** | **17** | **14** | **10** |
| **20** | **17** | **13** | **9** |
| **19** | **16** | **12** | **9** |
| **19** | **16** | **11** | **8** |
| **17** | **15** | **10** | **8** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Τμήμα Β2** | | | |
| **20** | **19** | **14** | **11** |
| **20** | **19** | **14** | **10** |
| **20** | **18** | **14** | **9** |
| **20** | **15** | **13** | **9** |
| **19** | **15** | **12** | **8** |

**α)** Να υπολογίσετε τη μέση τιμή και την τυπική απόκλιση της βαθμολογίας των μαθητών/τριών σε κάθε τμήμα. Ποια πρώτη εικόνα σας δίνουν τα αποτελέσματα των παραπάνω στατιστικών μέτρων για την επίδοση κάθε τμήματος;

**β)** Να βρείτε τους συντελεστές μεταβλητότητας (CV) και να συγκρίνετε τα δύο τμήματα ως προς την ομοιογένειά τους.

**γ)** Ο καθηγητής θέλει να δώσει βραβείο στους μαθητές/μαθήτριες κάθε τμήματος που πήραν βαθμό μεγαλύτερο ή ίσο από το 75% των μαθητών/τριών του τμήματος και να δώσει επιπλέον εργασία για το σπίτι στους μαθητές/μαθήτριες που πήραν βαθμό μικρότερο ή ίσο από το 25% των μαθητών/τριών του τμήματος. Να βρείτε τους βαθμούς των μαθητών/τριών που θα βραβευτούν και τους βαθμούς των μαθητών/τριών που θα πάρουν επιπλέον εργασία.

**δ)** Να κατασκευάσετε τα θηκογράμματα για κάθε τμήμα.

Λύση

**α)** Για το Β1 είναι:

1 = = = 14

= = 16,3 ή αλλιώς

= = 16,3

και τελικά s1=  4, 04

Ομοίως βρίσκουμε για το Β2: 2 = 14,95 και s2=  4,09

Από τα παραπάνω, το Β2 φαίνεται να έχει ελαφρώς υψηλότερη μέση επίδοση, χωρίς όμως τα δύο τμήματα να έχουν σημαντική διαφορά στη διασπορά των βαθμολογιών τους.

**β)** Για το Β1 βρίσκουμε ότι CV1 =  28,8%

και για το Β2 βρίσκουμε ότι CV2 =  27,4%.

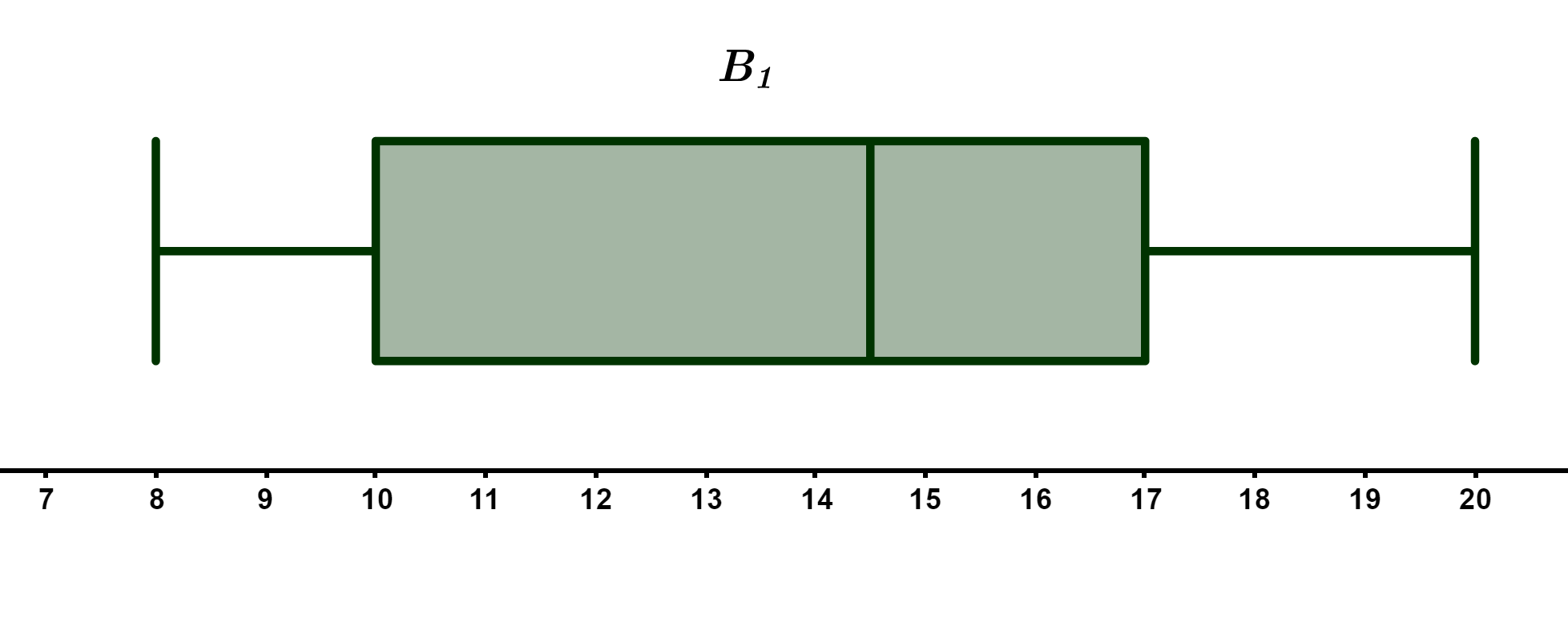
Φαίνεται ότι το Β2 είναι λίγο περισσότερο ομοιογενές, κάτι που οφείλεται στη μεγαλύτερη μέση τιμή του.

**γ)** Βρίσκουμε τα τεταρτημόρια για κάθε τμήμα:

για το Β1 έχουμε: Q1=10, Q2=δ=14,5, Q3=17, ενώ για το Β2 έχουμε: Q1=11,5, Q2=δ=14,5, Q3=19

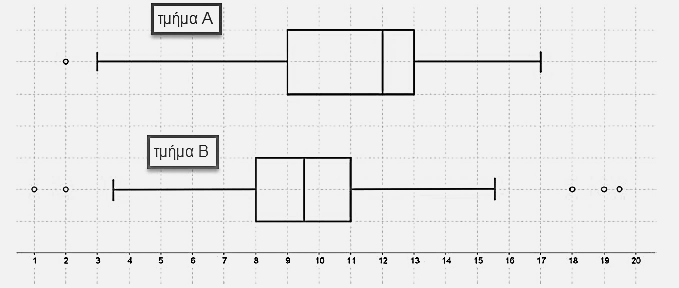
Από αυτά προκύπτει ότι οι βαθμοί των μαθητών/τριών που θα βραβευτούν θα είναι για το Β1 από 17 και πάνω, ενώ για το Β2 από 19 και πάνω. Οι βαθμοί των μαθητών/τριών που θα πάρουν επιπλέον εργασία θα είναι για το Β1 από 10 και κάτω, ενώ για το Β2 από 11 και κάτω.

**δ)** Τα θηκογράμματα φαίνονται παρακάτω:



##### Άσκηση 2

Τα παρακάτω θηκογράμματα παρουσιάζουν τους βαθμούς των μαθητών/τριών δύο τμημάτων Α και Β σε ένα μαθηματικό διαγωνισμό.



**α)** Να βρείτε ποιο από τα δύο τμήματα έχει το μεγαλύτερο εύρος βαθμών.

**β)** Να βρείτε ποιο από τα δύο τμήματα έχει το μεγαλύτερο ενδοτεταρτημοριακό εύρος βαθμών;

**γ)** Σε ποιο από τα δύο τμήματα φαίνεται να είναι πιο συμμετρική γύρω από τη διάμεσο η κατανομή των βαθμών;

**δ)** Να γράψετε μια μικρή αναφορά για το ποιο τμήμα θα μπορούσε να είναι το καλύτερο.

**ε)** Να βρείτε το τμήμα και τη βαθμολογία των δύο μαθητών/τριών με τον καλύτερο βαθμό.

Λύση

**α)** Το εύρος για το τμήμα Α είναι RΑ =17-2=15 και για το Β είναι RΒ =19,5-1=18,5. Οπότε το τμήμα Β έχει μεγαλύτερο εύρος βαθμών.

**β)** Για το Α έχουμε: Q=Q3-Q1=13-9=4 και για το Β: Q=Q3-Q1=11-8=3. Άρα μεγαλύτερο

ενδοτεταρτημοριακό εύρος βαθμών έχει το τμήμα Α.

**γ)** Στο τμήμα Β η κατανομή των βαθμών φαίνεται να είναι πιο συμμετρική γύρω από τη διάμεσο.

**δ)** Το "καλύτερο" θα μπορούσε να αναζητηθεί με διαφορετικά κριτήρια.

Με κριτήριο την μέση επίδοση καλύτερο φαίνεται να είναι το Α. Με κριτήριο τη συμμετρική κατανομή των βαθμών μάλλον είναι το Β. Με κριτήριο τη μικρότερη διασπορά των βαθμών καλύτερο φαίνεται μάλλον το Α. Με κριτήριο τους υψηλότερους βαθμούς (πάνω από 17) καλύτερο είναι το Β.

Ποιο όμως από όλα αυτά τα κριτήρια θα μπορούσε να είναι επικρατέστερο; Μιλώντας για τη συμμετοχή σε μαθηματικό διαγωνισμό, το κριτήριο της υψηλότερης βαθμολογίας ίσως να είναι αυτό που θα κυριαρχήσει. Αλλά σχετικά με τη διδασκαλία και την πρόοδο των μαθητών ίσως να είναι προτιμότερη μια μεγαλύτερη ομοιογένεια του τμήματος.

Βέβαια, σε αυτή τη συζήτηση δεν λαμβάνουμε καθόλου υπόψη άλλα χαρακτηριστικά του τμήματος (πχ πολιτισμικό υπόβαθρο, φύλο των μαθητών, ιδιαίτερα ενδιαφέροντα, κλίσεις, δυσκολίες και ικανότητες των μαθητών κ.α.) που μπορεί να μετατοπίζουν τη συζήτηση από το "ποιο τμήμα είναι καλύτερο" στο "ποια είναι τα χαρακτηριστικά του κάθε τμήματος". Επιπλέον, ούτε οι επιδόσεις, ούτε τα χαρακτηριστικά είναι αναλλοίωτα στο χρόνο και στις διαφορετικές συνθήκες.

**ε)** Οι δύο υψηλότερες βαθμολογίες είναι 19 και 19,5 και τις πέτυχαν μαθητές/τριες από το τμήμα Β.

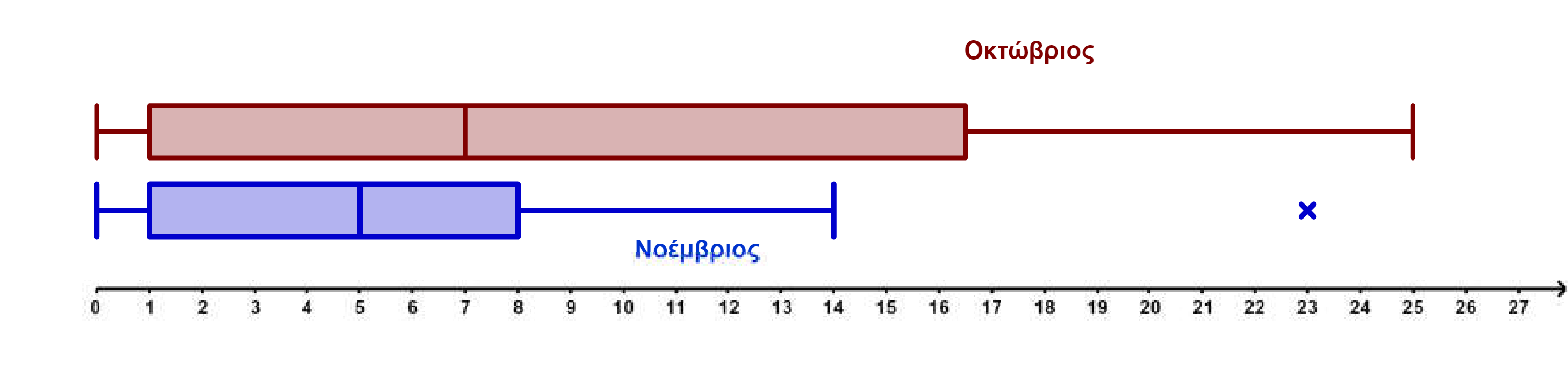
##### Άσκηση 3

Ο επόμενος πίνακας παρουσιάζει τον αριθμό των απουσιών που έκαναν οι μαθητές ενός τμήματος τους μήνες Οκτώβριο και Νοέμβριο:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Απουσίες μαθητών ενός τμήματος** | | | | | | | | | | | | | | | | |
| Οκτώβριος | 14 | 7 | 0 | 19 | 21 | 7 | 3 | 0 | 0 | 25 | 2 | 9 | 8 | 7 | 14 | 20 | 0 |
| Νοέμβριος | 7 | 8 | 2 | 2 | 1 | 23 | 14 | 7 | 0 | 14 | 4 | 7 | 5 | 0 | 1 | 8 | 0 |

Να κατασκευάσετε το θηκόγραμμα των απουσιών των μαθητών του τμήματος για κάθε μήνα χωριστά και να συγκρίνετε τις απουσίες των μαθητών.

Λύση

Τα θηκογράμματα των απουσιών φαίνονται παρακάτω:

Από τα θηκογράμματα φαίνεται ότι τον Οκτώβριο έγιναν περισσότερες απουσίες.

##### Άσκηση 4

Για δύο τύπους μπαταριών Α και Β επιλέχθηκαν δύο δείγματα μεγέθους 5 το καθένα. Οι χρόνοι ζωής των μπαταριών για το κάθε δείγμα σε χιλιάδες ώρες δίνονται στον επόμενο πίνακα:

|  |  |
| --- | --- |
| **Μπαταρία τύπου Α** | **Μπαταρία τύπου Β** |
| **20** | **26** |
| **26** | **32** |
| **24** | **19** |
| **22** | **20** |
| **18** | **23** |

**α)** Να βρείτε τη δειγματική μέση τιμή της διάρκειας ζωής μιας μπαταρίας τύπου Α και μιας μπαταρίας τύπου Β.

**β)** Με βάση το παραπάνω δείγμα και το γεγονός ότι μια μπαταρία τύπου Α στοιχίζει 38 ευρώ, ποιου τύπου μπαταρία θα προτιμήσετε αν μια μπαταρία τύπου Β στοιχίζει:

**ι)** 40 ευρώ **ιι)** 42 ευρώ

Να δικαιολογήσετε την απάντησή σας σε κάθε μία από τις περιπτώσεις ι) και ιι).

**γ)** Να βρείτε τις δειγματικές τυπικές αποκλίσεις sA

τύπων μπαταριών.

και sΒ της διάρκειας ζωής των δύο

**δ)** Να βρείτε ποιο από τα δύο παραπάνω δείγματα μπαταριών τύπου Α και Β παρουσιάζει τη μεγαλύτερη ομοιογένεια ως προς τη διάρκεια ζωής.

Λύση

**α)** Οι μέσες τιμές για τη διάρκεια ζωής στα δύο δείγματα είναι Α = 22 χιλιάδες ώρες και

Β = 24 χιλιάδες ώρες.

**β)** Για την μπαταρία Α το κόστος ανά χίλιες ώρες λειτουργίας είναι 38 : 22  1, 72 ευρώ. Για την μπαταρία Β το κόστος ανά χίλιες ώρες λειτουργίας είναι για την περίπτωση (ι) 40 : 24  1, 66 ευρώ και για την περίπτωση (ιι) είναι 42 : 24  1, 75 . Οπότε στην περίπτωση (ι) προτιμότερη είναι η μπαταρία Β, ενώ στην περίπτωση (ιι) προτιμότερη είναι η Α.

Τα παραπάνω βασίζονται στα δεδομένα των δύο δειγμάτων και στην υπόθεση ότι η

μπαταρία που θα αγοράσουμε θα έχει τον ίδιο χρόνο ζωής που έχει το αντίστοιχο δείγμα. Αυτή η υπόθεση δεν είναι βέβαιο ότι θα επαληθευτεί.

**γ)** Είτε χρησιμοποιώντας τον τύπο για τη διακύμανση και βρίσκοντας την τετραγωνική ρίζα, είτε με αξιοποίηση κάποιου λογιστικού φύλλου, βρίσκουμε τις τυπικές αποκλίσεις των δύο δειγμάτων: sA 2,8 και sΒ 4,7

**δ)** Έχουμε αντίστοιχα:

CVΑ =  12,7% και CVΒ =  19,6%

Οπότε μεγαλύτερη ομοιογένεια παρουσιάζει το δείγμα των μπαταριών τύπου Α.

##### Άσκηση 5

Στις 12μ.μ. η θερμοκρασία (σε βαθμούς Κελσίου) στη Λαμία και στη Θεσσαλονίκη το τελευταίο δεκαήμερο του Μαρτίου ήταν:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Θερμοκρασία σε βαθμούς Κελσίου** | | | | | | | | | |
| **Λαμία (Λ)** | **20** | **18** | **20** | **17** | **18** | **17** | **16** | **17** | **16** | **10** |
| **Θεσσαλονίκη (Θ)** | **18** | **16** | **17** | **15** | **16** | **12** | **16** | **17** | **20** | **22** |

**α)** Να βρείτε τη μέση, τη διάμεσο και την επικρατούσα θερμοκρασία των δειγμάτων της Λαμίας και της Θεσσαλονίκης.

**β)** Αν η δειγματική τυπική απόκλιση (σε βαθμούς Κελσίου) για τη Λαμία και τη

Θεσσαλονίκη είναι sΛ  2, 66 και s  2, 59 αντίστοιχα, να δικαιολογήσετε σε ποια από τα

Θ

δύο δείγματα οι τιμές της θερμοκρασίας έχουν μεγαλύτερη διασπορά.

**γ)** Εκ των υστέρων διαπιστώθηκε ότι το θερμόμετρο που χρησιμοποιήθηκε για τη μέτρηση της θερμοκρασίας στη Λαμία παρουσίαζε, λόγω κατασκευαστικού λάθους, αυξημένη θερμοκρασία κατά 5 βαθμούς. Αφού υπολογίσετε τις σωστές θερμοκρασίες της Λαμίας,

να βρείτε σε ποια πόλη από τις δύο το συγκεκριμένο δεκαήμερο οι τιμές της θερμοκρασίας έχουν μεγαλύτερη ομοιογένεια. Να δικαιολογήσετε την απάντησή σας.

Λύση

**α)** Οι δύο πόλεις έχουν την ίδια μέση θερμοκρασία:

**α)** Οι δύο πόλεις έχουν την ίδια μέση θερμοκρασία: Λ = 16,9 και Θ = 16,9 .

Για τις διαμέσους έχουμε δΛ = 17 και δΘ = 16,5.

Επικρατούσα θερμοκρασία για τη Λαμία είναι *M*Λ 17 και για τη Θεσσαλονίκη είναι *M*Θ 16.

**β)** Εφόσον Λαμίας.

sΛ > sΘ , μεγαλύτερη διασπορά έχουν οι θερμοκρασίες του δείγματος της

**γ)** Οι πραγματικές θερμοκρασίες στη Λαμία θα είναι 5 βαθμούς μικρότερες από εκείνες

του πίνακα, άρα θα είναι ως εξής:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Θερμοκρασία σε βαθμούς Κελσίου** | | | | | | | | | |
| **Λαμία (Λ)** | **15** | **13** | **15** | **12** | **13** | **12** | **11** | **12** | **11** | **5** |
| **Θεσσαλονίκη (Θ)** | **18** | **16** | **17** | **15** | **16** | **12** | **16** | **17** | **20** | **22** |

Η μέση τιμή θα είναι κατά 5 βαθμούς μικρότερη εκείνης που υπολογίστηκε, άρα Λ = 11,9 ,

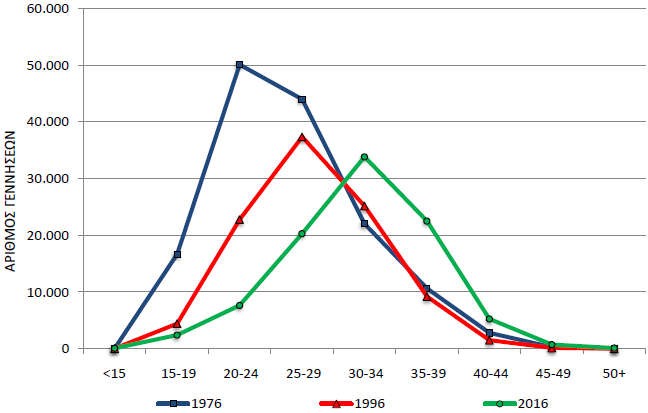
ενώ η τυπική απόκλιση παραμένει ίδια. Έτσι έχουμε:

CVΛ = =  22% και CVΘ = =  15% .

Οπότε μεγαλύτερη ομοιογένεια έχουν οι θερμοκρασίες στη Θεσσαλονίκη.

##### Άσκηση 6

Στο επόμενο διάγραμμα φαίνονται τα πολύγωνα συχνοτήτων των γεννήσεων ζώντων κατά ομάδες ηλικιών της μητέρας για τα έτη 1976, 1996 και 2016 (Πηγή: ΕΛ.ΣΤΑΤ: Διεύθυνση Στατιστικών Πληθυσμού & Αγοράς Εργασίας, Τμήμα Φυσικής & Μεταναστευτικής Κίνησης Πληθυσμού).



**α)** Σε ποια ηλικία των μητέρων έχουμε τις περισσότερες γεννήσεις το 1976, το 1996 και το 2016; Μπορείτε να το δικαιολογήσετε;

**β)** Να σχολιάσετε τον αριθμό γεννήσεων κατά ηλικιακή ομάδα των μητέρων για το 2016. **γ)** Σε ποιες ηλικιακές ομάδες των μητέρων φαίνεται ο αριθμός των γεννήσεων το 2016 να ξεπερνάει τους αντίστοιχους αριθμούς για τα έτη 1996 και 1976; Γιατί πιστεύετε ότι συνέβη αυτό;

Λύση

**α)** Τις περισσότερες γεννήσεις το 1976 έχουμε στις ηλικίες 20-24 των μητέρων, το 1996 στις ηλικίες 25-29 και το 2016 στις ηλικίες 30-34. Κάποιοι λόγοι για τους οποίους μπορεί να συμβαίνει αυτό συνδέονται με την αύξηση τη πρόσβασης των γυναικών στις σπουδές και την εργασία που μεταθέτει για αργότερα τις γεννήσεις των παιδιών. Θα μπορούσε να γίνει μια συζήτηση στην τάξη αφενός για άλλους λόγους που μπορεί να συμβαίνει αυτό και αφετέρου για την αξιολόγηση (ως θετικά ή αρνητικά, για ποιον και με τι κριτήριο) των γεγονότων αυτών (τόσο της αύξησης της ηλικίας των μητέρων κατά τη γέννηση του παιδιού τους, όσο και των λόγων που θεωρείται ότι αυτό συμβαίνει)

**β)** Η κορύφωση των γεννήσεων για το 2016 συμβαίνει στην ηλικία των 30-34. Το πολύγωνο συχνοτήτων φαίνεται κάπως συμμετρικό πριν και μετά την κορυφή του.

**γ)** Ο αριθμός των γεννήσεων για το 2016 ξεπερνάει τους αντίστοιχους αριθμούς για τα έτη 1996 και 1976 στις ηλικιακές ομάδες 30-34, 35-39, 40-44 και 45-49. Αυτό συνδέεται με το πρώτο ερώτημα, δηλαδή με τη "μετατόπιση" των γεννήσεων σε μεγαλύτερες ηλικίες των μητέρων. Έτσι, βλέπουμε οι λιγότερες γεννήσεις στις μικρότερες ηλικίες (για το 2016) να εξισορροπούνται με περισσότερες γεννήσεις στις μεγαλύτερες.

# ΕΝΟΤΗΤΑ 2.7 : ΓΡΑΜΜΙΚΗ ΣΥΣΧΕΤΙΣΗ ΠΟΣΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΚΑΙ ΔΙΑΓΡΑΜΜΑΤΑ ΔΙΑΣΠΟΡΑΣ

**Λύσεις των Ασκήσεων**

##### Άσκηση 1

Να διατάξετε τις παρακάτω τιμές του r σε αύξουσα τάξη του βαθμού γραμμικής

συσχέτισης δύο ποσοτικών μεταβλητών Χ και Υ: 0,6 ,

0,9 , 0,7 , 0, 2 ,

0 , 1 .

Λύση

Με κριτήριο το πόσο ισχυρή είναι η γραμμική συσχέτιση και ανεξάρτητα από το αν αυτή είναι θετική ή αρνητική, η διάταξη των τιμών του r σε αύξουσα σειρά είναι:

0, 0, 2, - 0, 6, - 0, 7, 0, 9, −1

##### Άσκηση 2

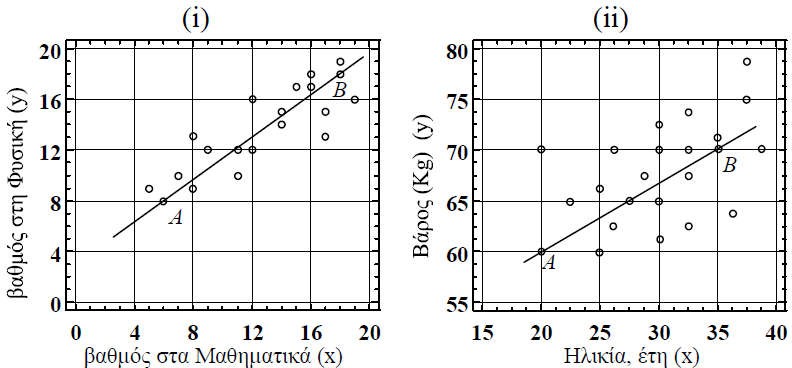
Ο συντελεστής γραμμικής συσχέτισης δύο μεταβλητών Χ και Υ είναι 0,96 , ενώ ο συντελεστής γραμμικής συσχέτισης δύο μεταβλητών Ζ και Φ είναι 0,96 . Ποια είναι η διαφορά τους;

Λύση

Τα δύο ζευγάρια μεταβλητών έχουν το ίδιο ισχυρή γραμμική συσχέτιση, αλλά για τις μεταβλητές Χ και Υ η συσχέτιση είναι θετική, ενώ για τις Ζ και Φ είναι αρνητική. Αυτό σημαίνει ότι όταν οι τιμές της μεταβλητής Χ αυξάνονται, οι τιμές της Υ τείνουν να αυξάνονται, ενώ όταν οι τιμές της μεταβλητής Ζ αυξάνονται, οι τιμές της Φ τείνουν να μειώνονται.

##### Άσκηση 3

Δίνονται δυο διαγράμματα διασποράς με χαραγμένες «με το μάτι» δύο ευθείες από έναν μαθητή.



**α)** Χρησιμοποιώντας τα σημεία Α και Β να βρείτε τις εξισώσεις των δύο ευθειών.

**β)** Πώς θα μπορούσατε να χρησιμοποιήσετε τις ευθείες του ερωτήματος α);

Λύση

**α)** Για την περίπτωση (i) έχουμε Α(6,8), Β(18,18) και αναζητούμε την ευθεία y = α + βx η οποία διέρχεται από τα Α και Β. Οπότε έχουμε το σύστημα

 8 = α + β 6



18 = α + β18

από τη λύση του οποίου προκύπτει ότι α = 3 και β = 0, 83 . Οπότε η ευθεία είναι η

y = 3 + 0, 83x.

Ομοίως για την περίπτωση (ii) βρίσκουμε την ευθεία y = 46, 66 + 0, 66x .

**β)** Οι ευθείες του ερωτήματος (α) μπορούν να χρησιμοποιηθούν για την εκτίμηση του βαθμού στη Φυσική αν ξέρουμε το βαθμό στα Μαθηματικά (για την περ i) και για την εκτίμηση του βάρους αν ξέρουμε το ύψος (για την περ. ii). Αυτές οι εκτιμήσεις μπορούν να γίνουν με κάποιους περιορισμούς. Για παράδειγμα, δεν μπορούμε να εκτιμήσουμε το βάρος για ηλικία 15 ετών ή για ηλικία 60 ετών, γιατί οι τιμές αυτές βρίσκονται έξω από το πεδίο των δεδομένων που έχουμε.

##### Άσκηση 4

Μια εταιρεία διαφημίσεων παρουσίασε τον επόμενο πίνακα:

|  |  |
| --- | --- |
| **Αριθμός**  **διαφημίσεων** | **Έσοδα**  **από πωλήσεις** |
| **10** | **20000** |
| **18** | **28000** |
| **24** | **35000** |
| **32** | **44000** |
| **35** | **48000** |
| **37** | **50000** |
| **42** | **55000** |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς και να εκτιμήσετε από αυτό τον συντελεστή γραμμικής συσχέτισης μεταξύ του αριθμού διαφημίσεων της εταιρείας και των εσόδων της από τις πωλήσεις.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ του αριθμού διαφημίσεων της εταιρείας και των εσόδων της από τις πωλήσεις.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Αν το πλήθος των διαφημίσεων που αγόραζε ένα κανάλι ήταν 30, πόσο εκτιμάτε ότι θα ήταν τα έσοδα της εταιρείας διαφημίσεων;

**ε)** Αν το πλήθος των διαφημίσεων που αγόραζε ένα κανάλι ήταν 60, θα ήταν ασφαλές να εκτιμήσετε τα έσοδα της εταιρείας διαφημίσεων, όπως στο δ);

Λύση

**α)** Το διάγραμμα διασποράς φαίνεται παρακάτω:



Αριθμός διαφημίσεων

45

40

35

30

25

20

15

10

5

0

60000

50000

40000

30000

20000

10000

0

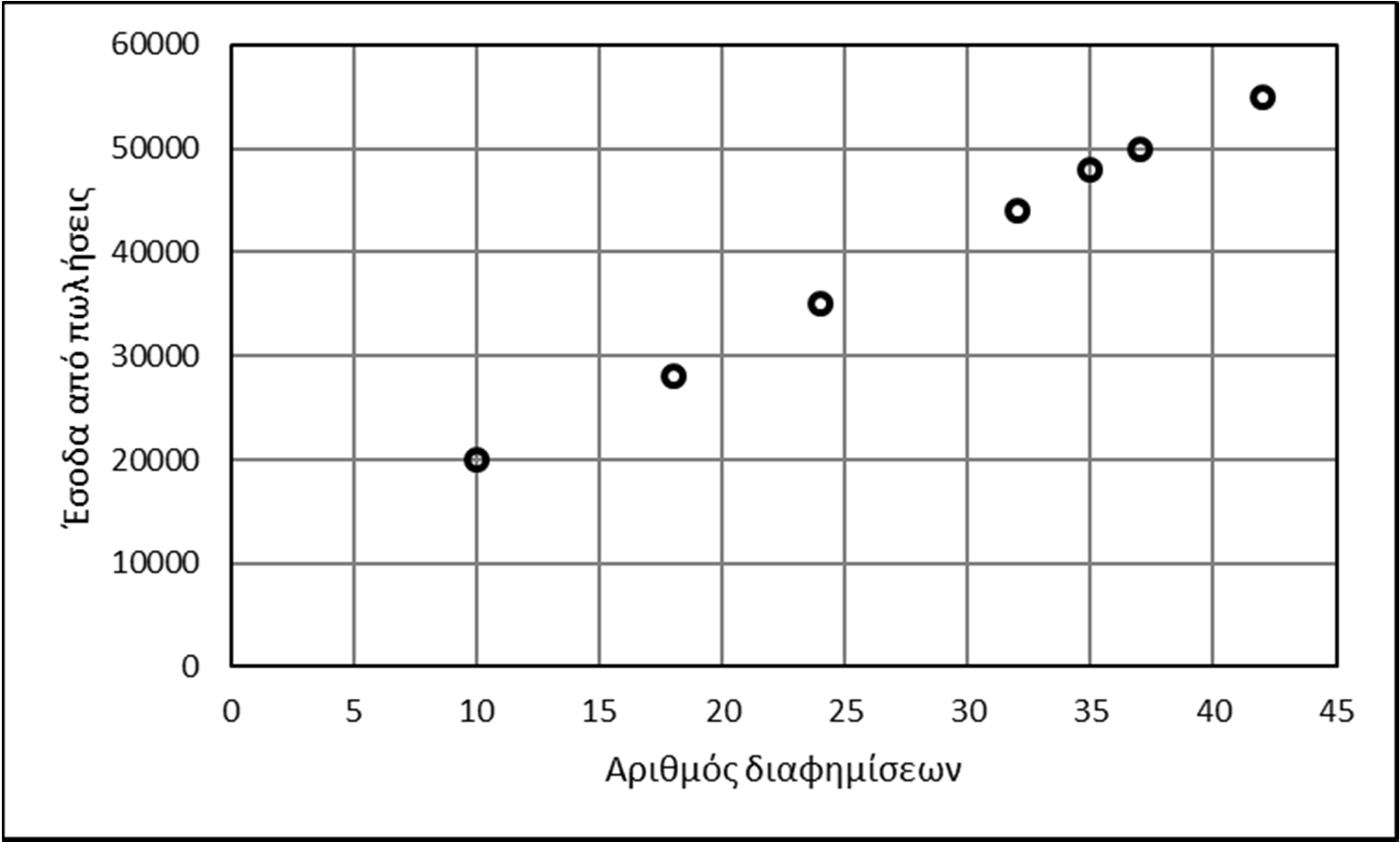
Έσοδα από πωλήσεις

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

Από το διάγραμμα διασποράς φαίνεται να υπάρχει ισχυρή θετική γραμμική συσχέτιση μεταξύ του αριθμού των διαφημίσεων και των εσόδων από πωλήσεις.

**β)** Χρησιμοποιώντας λογιστικό φύλλο ή κάποια εφαρμογή στατιστικής επεξεργασίας ή ακόμη και τον τύπο για τον συντελεστή Pearson, βρίσκουμε r = 0,9995 . Επειδή η τιμή του συντελεστή είναι πολύ κοντά στο 1, ο αριθμός των διαφημίσεων και τα έσοδα από πωλήσεις έχουν σχεδόν τέλεια θετική γραμμική συσχέτιση.

**γ)** Επιλέγουμε την ευθεία που διέρχεται από τα σημεία (10,20000) και (42,55000).



**δ)** Αν το πλήθος των διαφημίσεων που αγόραζε ένα κανάλι ήταν 30, τα έσοδα της εταιρείας διαφημίσεων εκτιμάμε ότι θα ήταν περίπου 42000.

**ε)** Για πλήθος διαφημίσεων 60, δεν μπορούμε να εκτιμήσουμε τα έσοδα, εφόσον το 60 βρίσκεται έξω από το πεδίο των δεδομένων που έχουμε στη διάθεσή μας.

##### Άσκηση 5

Τα δεδομένα του επόμενου πίνακα παριστάνουν τους βαθμούς (στην κλίμακα του 100) 10 μαθητών/τριών της Β΄ τάξης του Γενικού Λυκείου στα μαθήματα της Φυσικής (Χ) και των Μαθηματικών (Υ) κορμού σε μια γραπτή αξιολόγηση.

|  |  |  |  |
| --- | --- | --- | --- |
| **Βαθμός- Χ** | **Βαθμός- Υ** | **Βαθμός- Χ** | **Βαθμός- Υ** |
| **67** | **63** | **81** | **85** |
| **74** | **67** | **93** | **89** |
| **67** | **70** | **81** | **89** |
| **78** | **74** | **96** | **96** |
| **89** | **81** | **89** | **100** |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς και να εκτιμήσετε από αυτό τον συντελεστή γραμμικής συσχέτισης μεταξύ των βαθμών της Φυσικής (Χ) και των Μαθηματικών (Υ) των 10 μαθητών/τριών της Β τάξης του Γενικού Λυκείου.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των βαθμών της Φυσικής (Χ) και των Μαθηματικών (Υ) των 10 μαθητών/τριών της Β΄ τάξης του Γενικού Λυκείου.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Πώς θα μπορούσατε να εκτιμήσετε τον βαθμό των Μαθηματικών ενός μαθητή της Β΄ Λυκείου, εάν γνωρίζατε ότι στη Φυσική έγραψε 70;

Λύση

**α)** Το διάγραμμα διασποράς φαίνεται παρακάτω:



Βαθμός Φυσικής (Χ)

100

90

80

70

60

50

100

95

90

85

80

75

70

65

60

55

50

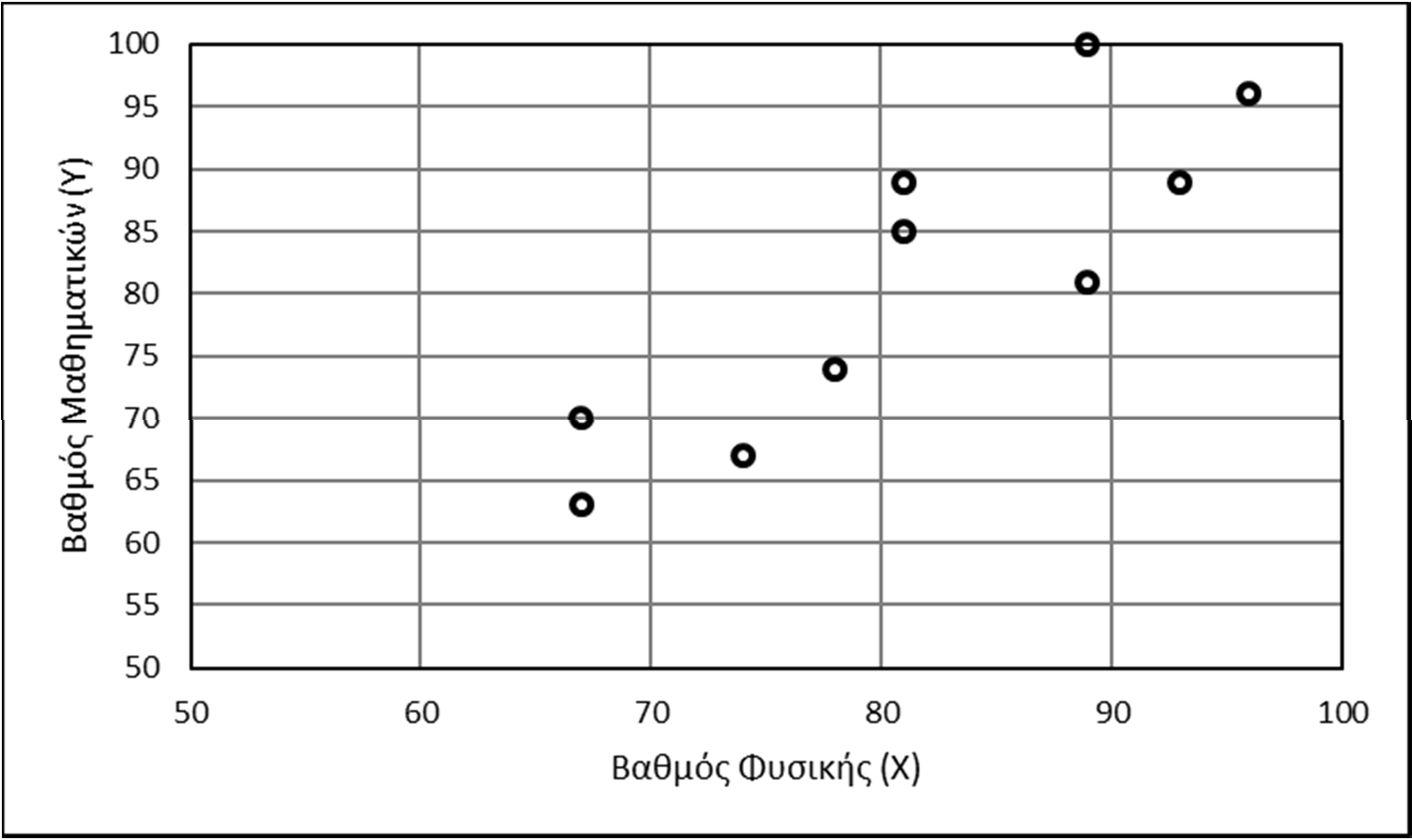
Βαθμός Μαθηματικών (Υ)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Από το διάγραμμα διασποράς φαίνεται ισχυρή θετική γραμμική συσχέτιση μεταξύ των βαθμών των δύο μαθημάτων.

**β)** Υπολογίζουμε (πχ με χρήση λογιστικού φύλλου) ότι r = 0,86 . Η τιμή αυτή του συντελεστή γραμμικής συσχέτισης Pearson δείχνει ισχυρή θετική γραμμική συσχέτιση μεταξύ των βαθμών των δύο μαθημάτων.

**γ)** Επιλέγουμε την ευθεία που φαίνεται στο παρακάτω σχήμα



**δ)** Για έναν μαθητή που έγραψε στη Φυσική 70, ένας αναμενόμενος βαθμός για το μάθημα των Μαθηματικών είναι περίπου 68.

##### Άσκηση 6

Στον παρακάτω πίνακα δίνονται οι ηλικίες και οι (συστολικές) πιέσεις αίματος 10 γυναικών.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Ηλικία (x)** | **56** | **42** | **72** | **36** | **63** | **47** | **55** | **49** | **38** | **60** |
| **Πίεση αίματος (y)** | **17** | **12** | **14** | **10** | **13** | **9** | **11** | **8** | **11** | **15** |

**α)** Να σημειώσετε σε ορθοκανονικό σύστημα αξόνων τα σημεία (x, y) σύμφωνα με τον παραπάνω πίνακα, όπου x είναι η ηλικία των γυναικών σε έτη και y είναι η πίεση αίματος των γυναικών σε cm Hg.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των μεταβλητών της ηλικίας των γυναικών σε έτη (x) και της πίεσης τους σε cm Hg (y).

Λύση

**α)** Το διάγραμμα διασποράς φαίνεται παρακάτω:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**β)** Βρίσκουμε ότι r = 0,57 . Αυτή η τιμή δηλώνει ότι υπάρχει θετική γραμμική συσχέτιση, η οποία όμως δεν είναι ισχυρή.



Ηλικίες (Χ)

80

70

60

50

40

30

0

5

10

15

20

25

Συστολική πίεση (Υ)

##### Άσκηση 7

Στον ακόλουθο πίνακα παρουσιάζονται στην πρώτη γραμμή οι τιμές (σε €) για διαφορετικά κράνη ποδηλασίας και στη δεύτερη γραμμή η βαθμολογία ποιότητάς τους που έγινε από ειδικούς (σε μια κλίμακα από 0 έως 100, όπου όσο μεγαλύτερη είναι η τιμή τόσο πιο ποιοτικό είναι το κράνος).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Τιμή (€)** | **35** | **22** | **33** | **42** | **50** | **23** | **29** | **18** | **39** | **28** | **20** | **25** |
| **Βαθμολογία ποιότητας** | **64** | **60** | **58** | **55** | **54** | **45** | **47** | **43** | **42** | **41** | **40** | **32** |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς.

**β)** Υπάρχει γραμμική συσχέτιση ανάμεσα στην τιμή και τη βαθμολογία ποιότητας;

**γ)** Θα μπορούσαμε να πούμε με βεβαιότητα ότι αν αγοράσουμε πιο φθηνό κράνος θα έχει πιο χαμηλή ποιότητα;

**δ)** Να σχεδιάσετε «με το μάτι» στο διάγραμμα διασποράς μια ευθεία που θα μπορούσε να περιγράφει τη σχέση του αναμενόμενου βαθμού ποιότητας ενός ποδηλατικού κράνους με την τιμή του.

Λύση

**α)** Το διάγραμμα διασποράς φαίνεται παρακάτω:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**β)** Από το διάγραμμα φαίνεται να υπάρχει κάποια ασθενής γραμμική συσχέτιση ανάμεσα στην τιμή και την βαθμολογία ποιότητας. Υπολογίζοντας τον συντελεστή Pearson βρίσκουμε r = 0,40 , κάτι που επιβεβαιώνει την εκτίμηση για ασθενή γραμμική συσχέτιση.



Τιμή (Χ)

60

50

40

30

20

10

20

30

40

50

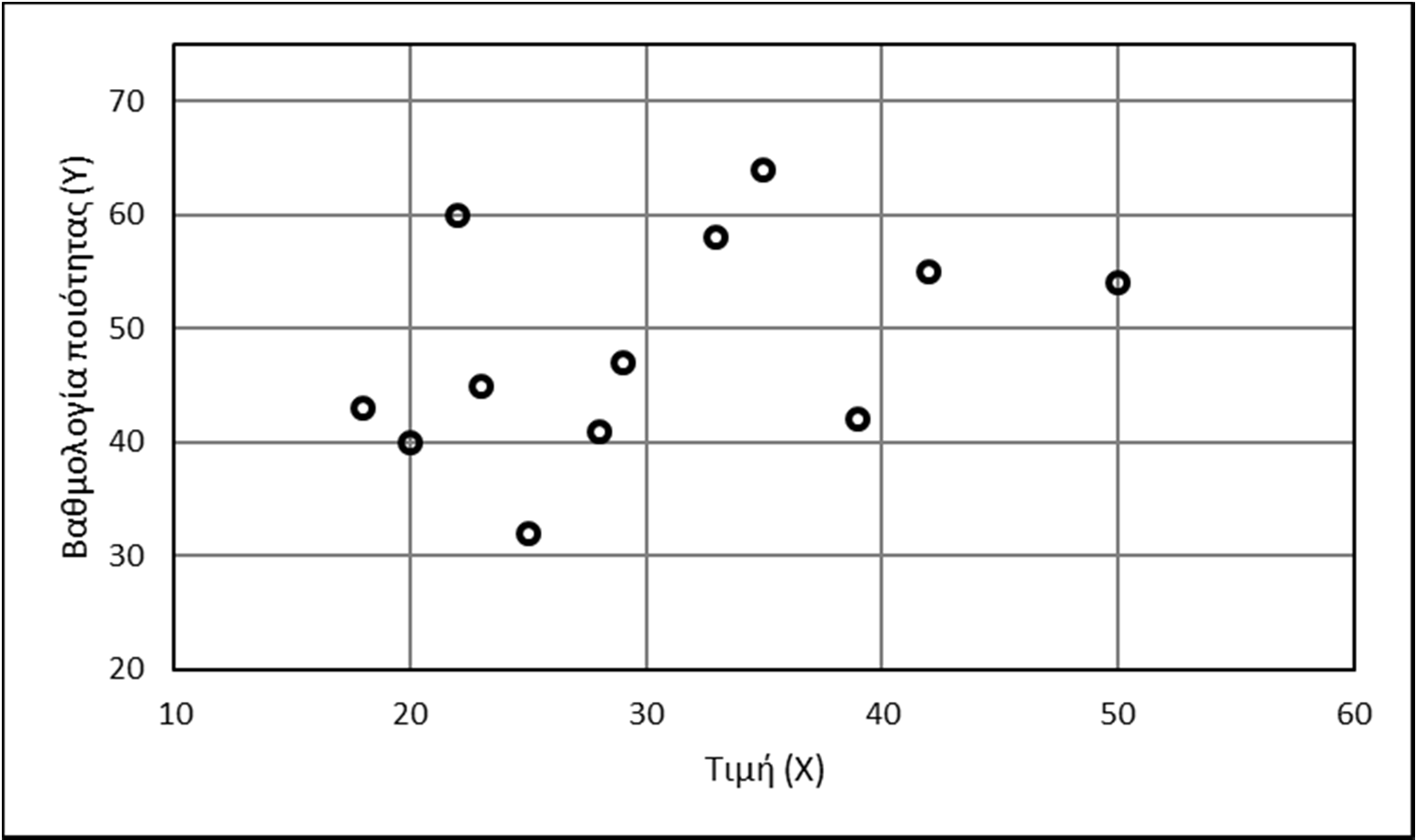
60

70

Βαθμολογία ποιότητας (Υ)

**γ)** Επειδή η συσχέτιση είναι ασθενής, δεν μπορούμε να πούμε με βεβαιότητα ότι χαμηλότερη τιμή σημαίνει χαμηλότερη ποιότητα. Για παράδειγμα υπάρχει κράνος με τιμή 39 € και βαθμολογία 42, και άλλο με τιμή 22 € και βαθμολογία 60.

**δ)** Μια ευθεία που θα μπορούσε να περιγράφει τη σχέση του αναμενόμενου βαθμού ποιότητας ενός ποδηλατικού κράνους με την τιμή του φαίνεται στο παρακάτω σχήμα.



##### Άσκηση 8

Στον ακόλουθο πίνακα παρουσιάζονται οι συντελεστές γραμμικής συσχέτισης των γραπτών βαθμολογιών στις εξετάσεις Ιουνίου σε 5 μαθήματα ενός τμήματος Β΄ τάξης γενικού λυκείου.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Άλγεβρα** | **Βιολογία** | **Γλώσσα** | **Φυσική** | **Χημεία** |
| **Άλγεβρα** | **1,00** |  |  |  |  |
| **Βιολογία** | **0,54** | **1,00** |  |  |  |
| **Γλώσσα** | **0,76** | **0,81** | **1,00** |  |  |
| **Φυσική** | **0,70** | **0,73** | **0,71** | **1,00** |  |
| **Χημεία** | **0,41** | **0,80** | **0,67** | **0,66** | **1,00** |

Να εξετάσετε κατά πόσο υπάρχει, ισχυρή ή όχι, γραμμική συσχέτιση ανάμεσα στις βαθμολογίες στα 5 εξεταζόμενα μαθήματα των μαθητών/τριών αυτών.

Λύση

Καταρχάς φαίνεται μια θετική γραμμική συσχέτιση ανάμεσα σε όλα τα μαθήματα, αν τα πάρουμε ανά δύο. Ισχυρή συσχέτιση υπάρχει μεταξύ Γλώσσας και Βιολογίας (0,81), Χημείας και Βιολογίας (0,80), Γλώσσας και Άλγεβρας (0,76).

##### Άσκηση 9

Το πλήθος x των οχημάτων σε εκατομμύρια και ο αριθμός y των ατυχημάτων σε εκατοντάδες, σε 15 διαφορετικές χώρες, δίνονται από τον επόμενο πίνακα:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Χώρα** | Α | Β | Γ | Δ | Ε | Ζ | Η | Θ | Ι | Κ | Λ | Μ | Ν | Ξ | Ο |
| **x** | 8,6 | 13,4 | 12,8 | 9,3 | 1,3 | 9,4 | 13,1 | 4,9 | 13,5 | 9,6 | 7,5 | 9,8 | 23,3 | 21 | 19,4 |
| **y** | 33 | 51 | 30 | 48 | 12 | 23 | 46 | 18 | 36 | 50 | 34 | 35 | 95 | 99 | 69 |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς μεταξύ των μεταβλητών του πλήθους των οχημάτων και του αριθμού ατυχημάτων για τις 15 χώρες.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των μεταβλητών του πλήθους των οχημάτων και του αριθμού ατυχημάτων για τις 15 χώρες.

Λύση

**α)** Το διάγραμμα διασποράς φαίνεται παρακάτω:



100

80

60

40

20

0

5

10

15

Πλήθος οχημάτων (Χ)

20

25

αριθμός ατυχημάτων (Υ)

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**β)** Υπολογίζουμε ότι r = 0,90 . Η τιμή αυτή δείχνει ότι υπάρχει ισχυρή γραμμική συσχέτιση μεταξύ του πλήθους των οχημάτων και του αριθμού ατυχημάτων για τις 15 χώρες. Είναι ίσως αναμενόμενο μεγαλύτερος αριθμός αυτοκινήτων να συνδέεται με μεγαλύτερο αριθμό ατυχημάτων. Ωστόσο, έχει ενδιαφέρον να διερευνηθούν οι παράγοντες που κάνουν τις δύο μεταβλητές να μην είναι πλήρως γραμμικά συσχετισμένες. Αυτό θα οδηγούσε σε επέκταση της έρευνας ώστε να συμπεριλάβει και άλλες διαστάσεις του προβλήματος.

##### Άσκηση 10

Από 8 γάμους που έγιναν σε μια εκκλησία ενός χωριού κατά τη διάρκεια ενός μηνός , οι ηλικίες των ανδρόγυνων ήταν:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Ηλικία γαμπρού** | **20** | **22** | **24** | **25** | **28** | **30** | **33** | **38** |
| **Ηλικία νύφης** | **20** | **20** | **22** | **27** | **24** | **25** | **28** | **34** |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς μεταξύ των ηλικιών της νύφης (Υ) και του γαμπρού (Χ) και να περιγράψετε το είδος της σχέσης που φαίνεται να έχουν οι δύο μεταβλητές.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των ηλικιών νύφης και γαμπρού.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Να βρείτε την αναμενόμενη ηλικία της νύφης για έναν υποψήφιο γαμπρό ετών 34.

Λύση

**α)** Το διάγραμμα διασποράς φαίνεται παρακάτω:



45

40

35

30

Ηλικία γαμπρού (Χ)

25

20

15

15

20

25

30

35

40

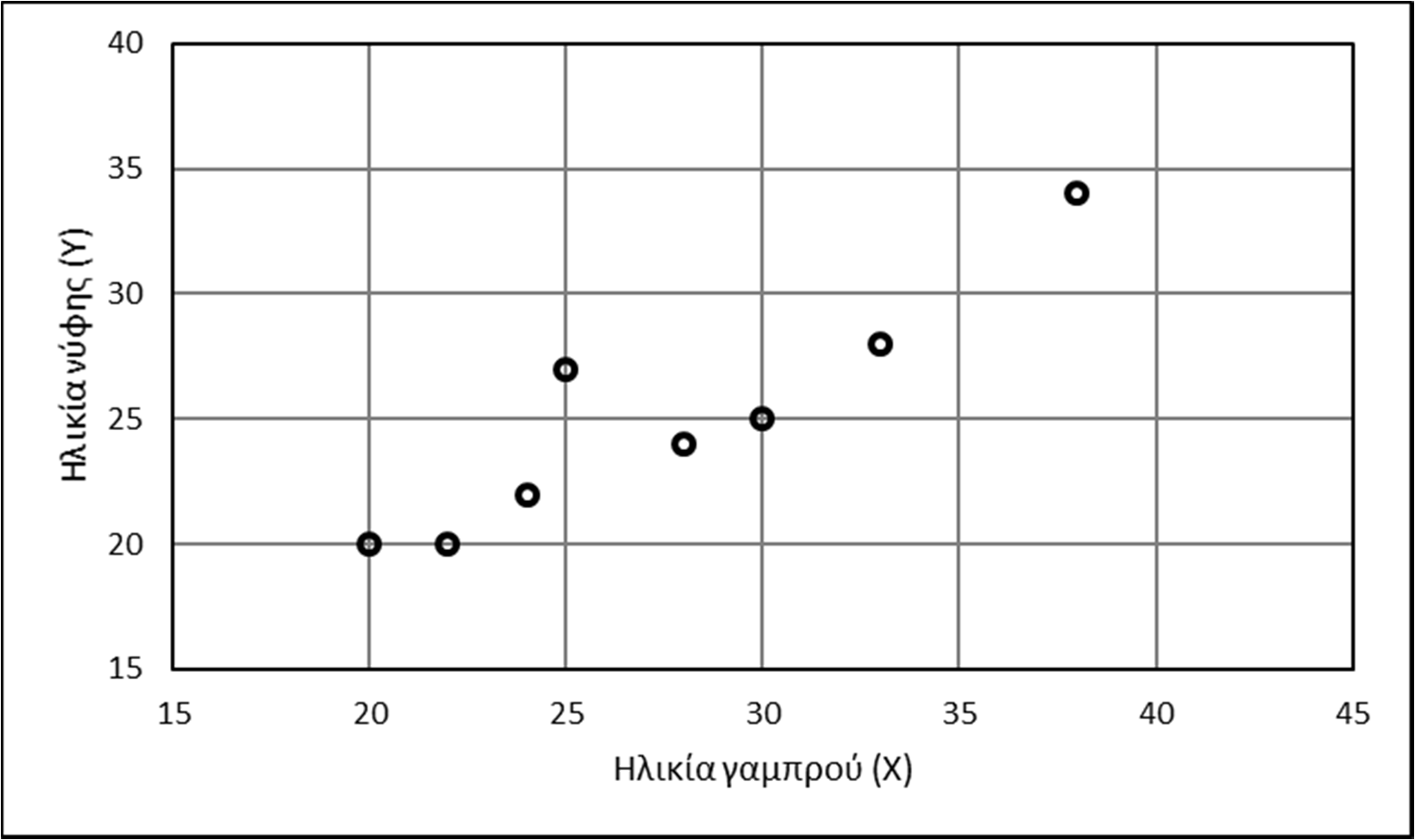
Ηλικία νύφης (Υ)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Από το διάγραμμα φαίνεται να υπάρχει ισχυρή γραμμική συσχέτιση μεταξύ των δύο ηλικιών.

**β)** Υπολογίζουμε ότι r = 0,92 , τιμή που επιβεβαιώνει την ισχυρή γραμμική συσχέτιση της ηλικίας του γαμπρού με εκείνη της νύφης.

**γ)** Η ευθεία φαίνεται στο παρακάτω σχήμα



**δ)** Αξιοποιώντας την παραπάνω ευθεία μπορούμε να εκτιμήσουμε ότι αν ο γαμπρός είναι 34 ετών, η αναμενόμενη ηλικία της νύφης είναι περίπου 31 ετών.

##### Άσκηση 11

Δίνεται δείγμα ν ζευγών παρατηρήσεων (xi , yi), i  1, 2,..., ν δύο μεταβλητών Χ και Υ και έστω r ο συντελεστής γραμμικής συσχέτισης των μεταβλητών Χ και Υ. Να αποδείξετε ότι αν όλα τα παραπάνω σημεία βρίσκονται πάνω σε μια ευθεία με εξίσωση:

yi = α + βxi , i  1,2,..., ν ,

τότε: r = 1 αν β > 0 και r = −1 αν β < 0 .

Λύση

α) Εφόσον ισχύει ότι: yi = α+βxi , i  1,2,..., ν , σύμφωνα με την εφαρμογή 3 της §2.3 θα

= α + β και sy = sx

Οπότε έχουμε:

r = = =

Επειδή: να = να = έχουμε:

r = = (1)

Αποδεικνύεται ότι:

= ,

οπότε είναι:

= -

Επιπλέον:

– ν = – ν = – ν = -

Με βάση αυτά η (1) γίνεται:

r = =

Οπότε, αν β > 0 θα είναι: r = = 1

ενώ αν β < 0 θα είναι: r = = -1 .

**Πρόσθετο Υλικό – θέματα για διερεύνηση**

1. Σε μια έρευνα που έγινε με σκοπό να διερευνηθεί η αποτελεσματικότητα μιας δίαιτας, μετρήθηκε το βάρος 15 ατόμων πριν και μετά τη δίαιτα.

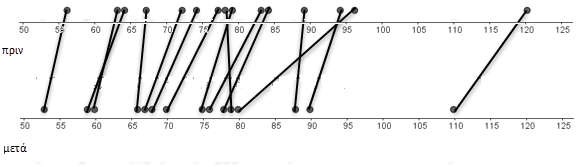
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
| **Πριν** | **89** | **83** | **78** | **56** | **96** | **120** | **67** | **63** | **72** | **74** | **79** | **94** | **84** | **64** | **77** |
| **Μετά** | **88** | **76** | **79** | **53** | **80** | **110** | **66** | **60** | **67** | **67** | **75** | **90** | **78** | **59** | **70** |

**α)** Να υπολογίσετε τη μέση τιμή και την τυπική απόκλιση της μεταβλητής του βάρους των 15 ατόμων πριν (Χπριν) και μετά (Χμετά) τη δίαιτα. Να σχολιάσετε τα αποτελέσματα.

**β)** Να κατασκευάσετε τα θηκογράμματα για τη μεταβλητή του βάρους των 15 ατόμων πριν και μετά τη δίαιτα. Να τα συγκρίνετε και να διατυπώσετε την άποψή σας για το αν υπάρχει διαφορά πριν και μετά και τη δίαιτα.

**γ)** Δίνονται τα επόμενα σημειογράμματα για τη μεταβλητή του βάρους των 15 ατόμων πριν και μετά τη δίαιτα, στα οποία έχουν ενωθεί τα βάρη των 15 ατόμων πριν και μετά τη

δίαιτα. Τι παρατηρείτε από το γράφημα; Επιβεβαιώνεται η παρατήρησή σας, συγκρίνοντας με τα αποτελέσματα των προηγούμενων ερωτημάτων;



**δ)** Να δημιουργήσετε τη μεταβλητή Ζ = Χπριν - Χμετά για κάθε άτομο και να κατασκευάσετε το θηκόγραμμά της. Πώς μπορείτε να αναδείξετε τη διαφορά των τιμών του βάρους των 15 ατόμων πριν και μετά τη δίαιτα από αυτό το θηκόγραμμα;

**ε)** Να κατασκευάσετε τα διαγράμματα διασποράς ανάμεσα στις μεταβλητές του βάρους (Χπριν) και (Χμετά) των 15 ατόμων.

**στ)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των

μεταβλητών του βάρους (Χπριν) και (Χμετά) των 15 ατόμων.

**ζ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**η)** Θα μπορούσατε να εκτιμήσετε το βάρος ενός ατόμου που πρόκειται να ακολουθήσει αυτή τη δίαιτα, εάν το αρχικό του βάρος ήταν 91 κιλά;

Λύση

**α)** Με χρήση λογιστικού φύλλου βρίσκουμε

Χπριν = 79,7,

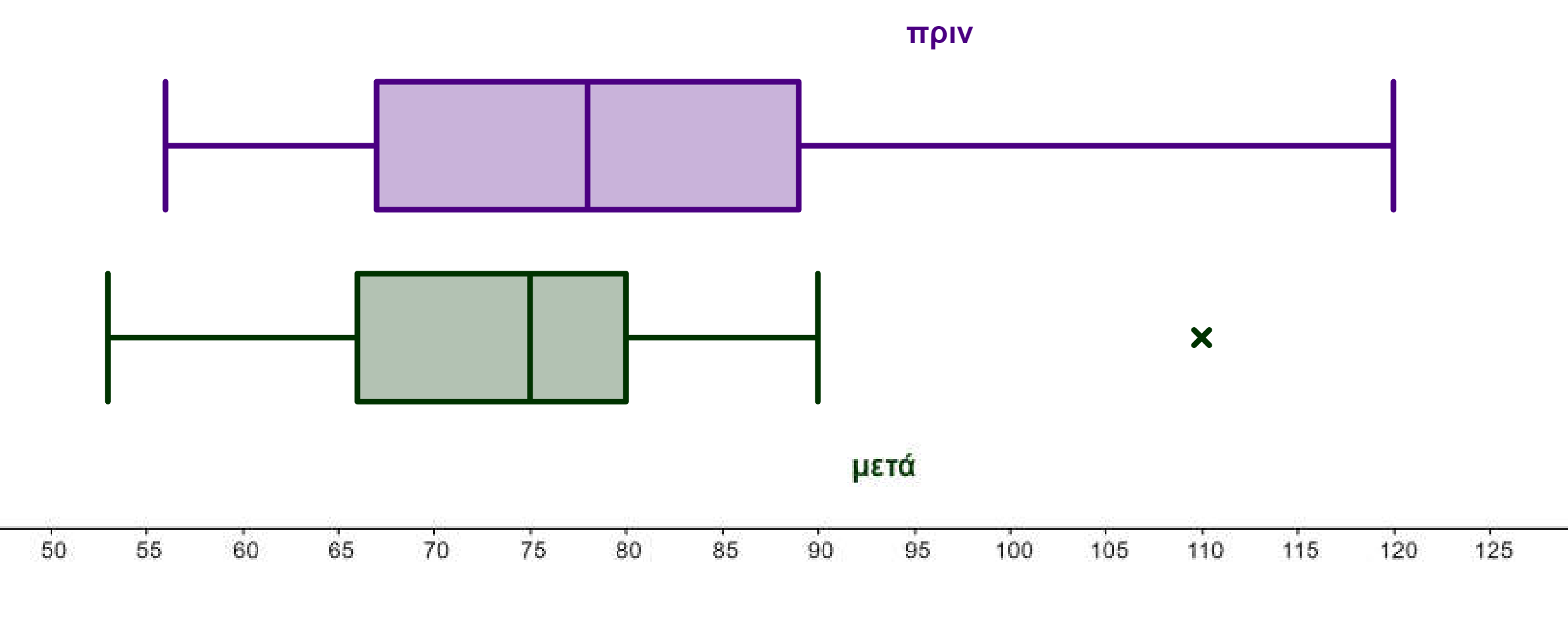
Χμετά = 74,5,

sπριν =15,4 και

sμετά =13,8

Κάποιος σχολιασμός θα μπορούσε να συνδέεται με τη μείωση του βάρους, αλλά πιθανόν και με τη μείωση της διασποράς.

**β)** Τα θηκογράμματα φαίνονται παρακάτω:

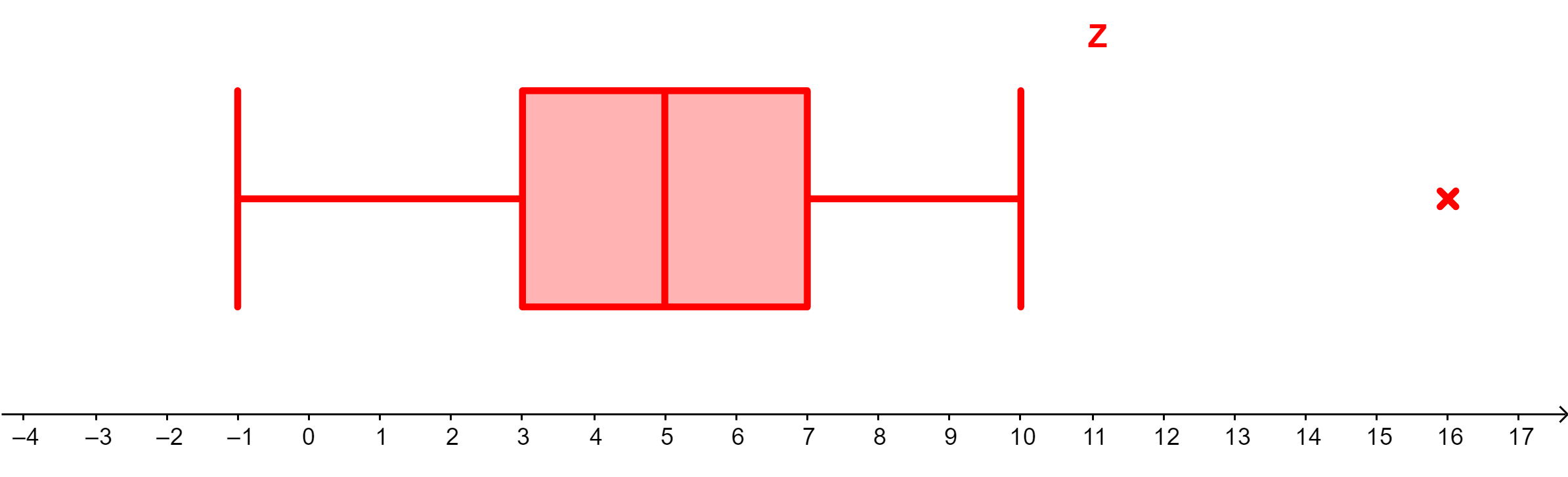


Φαίνεται να υπάρχει διαφορά μετά τη δίαιτα, ο αναγνώστης καλείται να την σχολιάσει.

**γ)** Ο σχολιασμός αφήνεται στον αναγνώστη.

**δ)** Η μεταβλητή Ζ φαίνεται στον παρακάτω πίνακα:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
| **Ζ = Χπριν - Χμετά** | **1** | **7** | **-1** | **3** | **16** | **10** | **1** | **3** | **5** | **7** | **4** | **4** | **6** | **5** | **7** |

και το αντίστοιχο θηκόγραμμα:

Ο σχολιασμός αφήνεται στον αναγνώστη

**ε), στ), ζ)** και **η)** Δίνεται το διάγραμμα διασποράς και ότι ο συντελεστής Pearson είναι

r = 0, 97 . H γραμμή, ο σχολιασμός και η εκτίμηση αφήνονται στον αναγνώστη.



125

115

105

95

85

πριν

75

65

55

45

115

105

95

85

75

65

55

45

μετά

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |