

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΕΡΕΥΝΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ  
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ



Β΄ ΓΕΝΙΚΟΥ ΛΥΚΕΙΟΥ

ΦΑΚΕΛΟΣ ΜΑΘΗΤΗ/ΤΡΙΑΣ  
**ΣΤΑΤΙΣΤΙΚΗ**

2019

ΙΝΣΤΙΤΟΥΤΟ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΕΚΔΟΣΕΩΝ «ΔΙΟΦΑΝΤΟΣ»



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΕΡΕΥΝΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ  
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

ΦΑΚΕΛΟΣ ΜΑΘΗΤΗ/ΤΡΙΑΣ  
ΣΤΑΤΙΣΤΙΚΗ

Β ' Γενικού Λυκείου

2019

<b>ΣΥΝΤΟΝΙΣΜΟΣ:</b>	Σκούρας Αθανάσιος, Σύμβουλος Α΄ ΥΠ.Π.Ε.Θ.
<b>ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΟΜΑΔΑ ΜΕΛΩΝ ΔΕΠ:</b>	Θεοδόσιος Ζαχαριάδης, Καθηγητής Ε.Κ.Π.Α. Μιχαήλ Λουλάκης, Αναπληρωτής Καθηγητής Ε.Μ.Π. Δέσποινα Πόταρη, Καθηγήτρια Ε.Κ.Π.Α.
<b>ΑΝΑΜΟΡΦΩΣΗ ΚΑΙ ΕΠΙΜΕΛΕΙΑ ΠΕΡΙΕΧΟΜΕΝΟΥ:</b>	Σκούρας Αθανάσιος, Σύμβουλος Α΄ ΥΠ.Π.Ε.Θ. Στράντζαλος Αθανάσιος, Σύμβουλος Β΄ Ι.Ε.Π.
<b>ΕΚΠΟΝΗΣΗ:</b> <i>Εξωτερικοί εμπειρογνώμονες:</i>	Βρούτσος Νικόλαος, Εκπαιδευτικός Δημοσίου Τομέα (ΠΕΟ3) Διαμαντίδης Δημήτριος, Εκπαιδευτικός Δημοσίου Τομέα (ΠΕΟ3) Καρκάνης Βασίλειος, Εκπαιδευτικός Δημοσίου Τομέα (ΠΕΟ3) Μηλιώνης Χρίστος, Σχολικός Σύμβουλος (ΠΕΟ3) Πατσαλιάς Μιχάλης, Εκπαιδευτικός Δημοσίου Τομέα (ΠΕΟ3) Σκουρκέας Αναστάσιος, Εκπαιδευτικός Δημοσίου Τομέα (ΠΕΟ3) Σπαθάρης Δημήτριος, Σχολικός Σύμβουλος (ΠΕΟ3) Στουραίτης Κωνσταντίνος, Εκπαιδευτικός Δημοσίου Τομέα (ΠΕΟ3)
<i>Μέλη της Μονάδας Ανθρωπιστικών Επιστημών και Φιλολογίας του ΙΕΠ:</i>	Σκούρας Αθανάσιος, Σύμβουλος Α΄ ΥΠ.Π.Ε.Θ. Στράντζαλος Αθανάσιος, Σύμβουλος Β΄ Ι.Ε.Π.
<b>ΠΡΟΕΚΤΥΠΩΤΙΚΕΣ ΕΡΓΑΣΙΕΣ</b>	ΙΤΥΕ “ΔΙΟΦΑΝΤΟΣ”

ΔΡΑΣΗ ΓΙΑ ΤΗΝ ΑΝΑΜΟΡΦΩΣΗ Ή/ΚΑΙ ΕΚΠΟΝΗΣΗ ΠΡΟΓΡΑΜΜΑΤΩΝ ΣΠΟΥΔΩΝ ΚΑΙ ΤΥΧΩΝ ΣΥΜΠΛΗΡΩΜΑΤΙΚΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΥΛΙΚΟΥ ΣΤΑ ΘΕΜΑΤΙΚΑ ΠΕΔΙΑ ΤΩΝ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ, ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ, ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΜΑΘΗΜΑΤΙΚΩΝ ΠΡΩΤΟΒΑΘΜΙΑΣ ΚΑΙ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ (ΠΡΑΞΗ 36/14-09-2017 ΤΟΥ ΔΣ ΤΟΥ ΙΕΠ ΚΑΙ ΣΕ ΣΥΝΕΧΕΙΑ ΤΗΝ ΜΕ ΑΡ. ΠΡΩΤ. 6139/19-09-2017 ΚΑΙ ΑΔΑ: 73Μ10ΞΛΔ-ΗΨΟ ΠΡΟΣΚΛΗΣΗ ΕΚΔΗΛΩΣΗΣ ΕΝΔΙΑΦΕΡΟΝΤΟΣ)

**ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ**

**Γεράσιμος Κουζέλης**

Πρόεδρος του Ινστιτούτου Εκπαιδευτικής Πολιτικής

**Υπεύθυνη Δράσης**

**Γεωργία Φέμελη**

Σύμβουλος Α΄ του Ινστιτούτου Εκπαιδευτικής Πολιτικής

Το παρόν εκπονήθηκε αμισθί, με ευθύνη της Μονάδας Φυσικών Επιστημών, Τεχνολογίας και Μαθηματικών του Ινστιτούτου Εκπαιδευτικής Πολιτικής, στο πλαίσιο της ανωτέρω Δράσης.

## ΠΡΟΛΟΓΟΣ

Το παρόν εκπαιδευτικό υλικό αποτελεί τμήμα από τον Φάκελλο του μαθητή/τριας που εκπονήθηκε στο πλαίσιο του Προγράμματος Σπουδών των Μαθηματικών Κορμού της Β΄ Λυκείου στην Άλγεβρα –Στατιστική, και περιλαμβάνει τις ενότητες της Στατιστικής:

Ενότητα 3.1: Πληθυσμός - Δείγμα – Μεταβλητές

Ενότητα 3.2: Παρουσίαση στατιστικών δεδομένων

Ενότητα 3.3: Μέτρα θέσης και μεταβλητότητας

Η διάρθρωση του υλικού προσομοιάζει με διδακτικό εγχειρίδιο ώστε ο μαθητής να το χρησιμοποιεί ως βασική πηγή για τη οργάνωση της μελέτης του. Συγχρόως, τα διακριτά τμήματα που χαρακτηρίζουν τη δομή του:

- Διερεύνηση,
- Βασικές μαθηματικές έννοιες – ιδέες – διεργασίες,
- Εφαρμογές,
- Ασκήσεις – προβλήματα – δραστηριότητες,
- Πρόσθετο υλικό.

αποτυπώνουν και μια πρόταση (ένα «παράδειγμα») για την πορεία διδασκαλίας - μάθησης που εξελίσσεται σε μια σχολική τάξη:

Με την προσθήκη του ως άνω υλικού στη διδασκαλία της Άλγεβρας της Β΄ Λυκείου, επιδιώκεται η εξοικείωση των μαθητών/τριών με τα στατιστικά δεδομένα, την κριτική τους ανάγνωση και την κατάκτηση των βασικών τεχνικών μελέτης και παρουσίασης των χαρακτηριστικών ενός πληθυσμού. Η προσέγγιση αυτή φιλοδοξεί να λειτουργήσει επωφελώς στην οπτική του αυριανού πολίτη ευρύτερα, ακόμα και στις καθημερινές αποφάσεις του και στον τρόπο που ερμηνεύει το περιβάλλον γύρω του (φυσικό ή κοινωνικό).

## Εισαγωγή

Πολλές φορές, ακούμε από τα μέσα μαζικής ενημέρωσης ή διαβάζουμε στα διάφορα έντυπα για τα αποτελέσματα μιας έρευνας που βασίζεται σε στατιστικά δεδομένα. Αυτές οι πληροφορίες, που αναφέρονται σε συγκεκριμένη έρευνα, συνήθως δίνονται με τη μορφή πινάκων ή διαγραμμάτων. Εκτός από την παρουσίαση αυτών των πληροφοριών, πολλές φορές ακούμε ή διαβάζουμε μελέτες και αναλύσεις των παρατηρήσεων που αναφέρονται στην έρευνα.

Στατιστική είναι η εφαρμοσμένη επιστήμη που συγκεντρώνει και παρουσιάζει πληροφορίες, αλλά ταυτόχρονα μελετά και αναλύει τις παρατηρήσεις, που αναφέρονται στην οικονομία, στη μετεωρολογία, στην υγεία, στην κοινωνιολογία, στον αθλητισμό κ.τ.λ.

Η Στατιστική περιλαμβάνει τόσο τις μεθόδους συλλογής και επεξεργασίας στοιχείων όσο και τις μεθόδους ανάλυσης και μελέτης αυτών. Από τη μελέτη αυτή προκύπτουν οι σχέσεις που υπάρχουν στα διάφορα φαινόμενα και διατυπώνονται συμπεράσματα που είναι χρήσιμα για τη λήψη αποφάσεων.

Ως ορισμό της «Στατιστικής» θα δώσουμε τον συνηθέστερο και πλέον γνωστό ορισμό του R.A. Fisher (1890-1962), πατέρα της σύγχρονης Στατιστικής:

Στατιστική είναι ένα σύνολο αρχών και μεθοδολογιών για:

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

Ο κλάδος της Στατιστικής που ασχολείται με τον πρώτο στόχο λέγεται **σχεδιασμός πειραμάτων** ενώ, με τον δεύτερο ασχολείται η **περιγραφική στατιστική**, που αποτελεί και το αντικείμενο μελέτης μας στη συνέχεια. Τέλος, η **επαγωγική στατιστική ή στατιστική συμπερασματολογία** περιλαμβάνει τις μεθόδους με τις οποίες γίνεται η προσέγγιση των χαρακτηριστικών ενός μεγάλου συνόλου δεδομένων, με τη μελέτη των χαρακτηριστικών ενός μικρού υποσυνόλου των δεδομένων.

## ΕΝΟΤΗΤΑ 3.1: ΠΛΗΘΥΣΜΟΣ - ΔΕΙΓΜΑ - ΜΕΤΑΒΛΗΤΕΣ

### Διερεύνηση

Δυο φοιτητές κοινωνιολογίας έχουν ως θέμα εργασίας να μελετήσουν κατά πόσον οι νέοι ηλικίας μεταξύ 12 και 20 ετών όλης της χώρας συζητούν στα μέσα κοινωνικής δικτύωσης θέματα, που έχουν να κάνουν με την προσωπική τους ζωή και αν το γεγονός αυτό επιφέρει αρνητικές επιπτώσεις στους ίδιους. Για να βοηθηθούν στην έρευνά τους συγκέντρωσαν από το διαδίκτυο τα αποτελέσματα παρόμοιων ερευνών που πραγματοποιήθηκαν σε άλλες χώρες.



Σύμφωνα με τα στοιχεία μιας έρευνας που διεξήχθη πρόσφατα (πηγή ΑΠΕ-ΜΠΕ), τα κοινωνικά δίκτυα δημιουργούν επιπρόσθετο στρες στα νέα παιδιά της Βρετανίας. Για τη διεξαγωγή της έρευνας ερωτήθηκαν 5.000 νέοι/ες, ηλικίας μεταξύ 12 και 20 ετών, από όλη τη χώρα. Εκ των ερωτηθέντων, το 40% δήλωσε ότι αισθάνεται άσχημα όταν κανείς από τους «φίλους» τους δεν κάνει «like» στις «selfies» τους, ενώ το 35% απάντησε ότι το επίπεδο της αυτοπεποίθησής τους

συνδέεται άμεσα από τον αριθμό των «ακολουθών» που έχουν. Περισσότερες λεπτομέρειες για τα αποτελέσματα της συγκεκριμένης έρευνας φαίνονται στην διπλανή εικόνα.

### Ερωτήσεις

- 1) Ποιους/ες εκπροσωπούν οι 5.000 νέοι/ες της έρευνας;
- 2) Ποια μπορεί να ήταν τα ερωτήματα και ποια τα χαρακτηριστικά ως προς τα οποία εξετάστηκαν οι ερωτηθέντες;

Στη συνέχεια οι δυο φοιτητές ετοίμασαν ένα αντίστοιχο ερωτηματολόγιο για την δική τους έρευνα και το έστειλαν στους φίλους τους και τους συμφοιτητές τους για να το απαντήσουν.

### Ερωτήσεις

- 1) Να γράψετε μερικές από τις ερωτήσεις που ενδεχομένως να συμπεριέλαβαν στο ερωτηματολόγιό τους οι δυο φοιτητές, οι οποίες να επιδέχονται άλλες ποσοτική και άλλες ποιοτική απάντηση.
- 2) Να σχολιάσετε τον τρόπο που εργάστηκαν οι φοιτητές για να συλλέξουν τα δεδομένα τους. Ήταν κατάλληλο το δείγμα που επέλεξαν ώστε να είναι αξιόπιστη η έρευνα;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

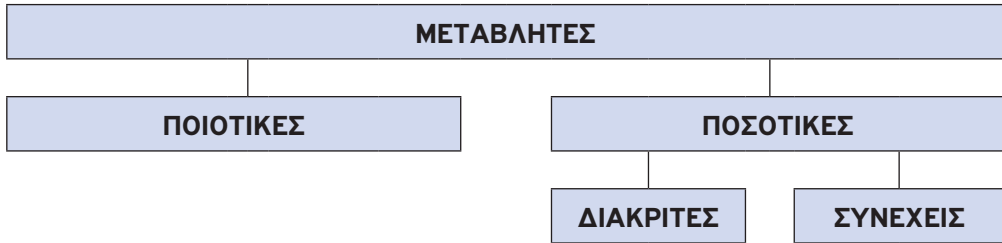
### Ορισμοί

- Το σύνολο, στο οποίο επικεντρωνόμαστε σε μια έρευνα, ονομάζεται **πληθυσμός** της έρευνας. Τα στοιχεία του πληθυσμού συχνά αναφέρονται ως **άτομα** του πληθυσμού.
- Τις περισσότερες φορές είναι πρακτικά πολύ δύσκολο ή και αδύνατο να προσεγγίσουμε καθένα από τα άτομα του πληθυσμού, για τεχνικούς λόγους ή λόγω χρόνου, κόστους κ.τ.λ. Στις περιπτώσεις αυτές επιλέγουμε ένα μέρος του πληθυσμού που είναι αντιπροσωπευτικό του πληθυσμού και το εξετάζουμε. Το μέρος αυτό του πληθυσμού λέγεται **δείγμα**.
- Τα χαρακτηριστικά ως προς τα οποία εξετάζουμε τα άτομα ενός πληθυσμού λέγονται μεταβλητές και συμβολίζονται με ένα κεφαλαίο γράμμα X, Y, Z, ... κ.τ.λ.

### Διάκριση μεταβλητών

Οι μεταβλητές διακρίνονται σε:

- **Ποιοτικές** αν οι τιμές τους δεν είναι αριθμοί, όπως είναι: τα χρώματα των αυτοκινήτων, το φύλο των μαθητών κ.τ.λ.
- **Ποσοτικές** αν οι τιμές τους είναι αριθμητικές και επιδέχονται μέτρηση, όπως είναι: ο ετήσιος αριθμός των τροχαίων ατυχημάτων, το ύψος των μαθητών κ.τ.λ.
  - **Διακριτές** αν παίρνουν μεμονωμένες τιμές, όπως είναι: ο αριθμός των παιδιών των οικογενειών, το νούμερο των γυναικείων παπουτσιών ανά μισό πόντο κ.τ.λ.
  - **Συνεχείς** αν μπορούν να πάρουν οποιαδήποτε τιμή ενός διαστήματος, όπως είναι: το βάρος των μαθητών, ο χρόνος που χρειάζονται οι μαθητές για να απαντήσουν σε ένα διαγώνισμα κ.τ.λ.



## Εφαρμογές

### Εφαρμογή 1η

Μια ομάδα μαθητών θέλει να προβεί σε πρόβλεψη, για το ποιος θα είναι ο νέος πρόεδρος του μαθητικού συμβουλίου του σχολείου. Για το λόγο αυτό τα μέλη της ομάδας κατέγραψαν την πρόθεση ψήφου των μαθητών ενός συγκεκριμένου τμήματος του σχολείου.

- 1) Ποιος είναι ο πληθυσμός της έρευνας;
- 2) Ποιο είναι το δείγμα;
- 3) Ποια είναι η μεταβλητή της έρευνας και ποιο το είδος της;
- 4) Να συζητήσετε πλεονεκτήματα και μειονεκτήματα της μεθόδου που επέλεξαν οι μαθητές, για να συγκεντρώσουν τις πληροφορίες και να κάνουν την πρόβλεψή τους.

### Απάντηση

- 1) Ο πληθυσμός της έρευνας είναι όλοι οι μαθητές του σχολείου.
- 2) Το δείγμα αποτελείται από τους μαθητές του τμήματος των οποίων καταγράφηκε η ψήφος τους.
- 3) Μεταβλητή της έρευνας είναι: «ποιος θα είναι ο νέος πρόεδρος του μαθητικού συμβουλίου του σχολείου». Η μεταβλητή αυτή είναι ποιοτική.
- 4) Το πλεονέκτημα της μεθόδου, που επέλεξαν οι μαθητές, είναι η ευκολία και η συντομία στο να πάρουν τις απαραίτητες απαντήσεις από τα άτομα του δείγματος. Το μειονέκτημα είναι ότι το δείγμα δεν είναι και αντιπροσωπευτικό του πληθυσμού, αφού αποτελείται από τους μαθητές ενός συγκεκριμένου τμήματος. Έτσι τα αποτελέσματα της έρευνας δεν θα είναι αξιόπιστα διότι οι υπόλοιποι μαθητές μπορεί να έχουν διαφορετική γνώμη.



## Εφαρμογή 2η

Η τροχαία θέλει να κάνει μια έρευνα για τα τροχαία ατυχήματα που σημειώθηκαν το 2017 στον αυτοκινητόδρομο Αθηνών - Θεσσαλονίκης. Θέλει να εξετάσει τις πιο κάτω μεταβλητές:

- 1) τον τύπο του οχήματος που έχει εμπλακεί στο δυστύχημα,
- 2) την ταχύτητα του οχήματος,
- 3) τον αριθμό των τραυματιών,

Να χαρακτηρίσετε το είδος καθεμιάς από τις μεταβλητές.

### Απάντηση

Οι μεταβλητές χαρακτηρίζονται ως προς το είδος τους ως ακολούθως:

- 1) ποιοτική,
- 2) ποσοτική συνεχής,
- 3) ποσοτική διακριτή.

## Ασκήσεις - Προβλήματα - Δραστηριότητες

**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) Ποιο πρόβλημα της πόλης θεωρείται το πιο σημαντικό.

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:

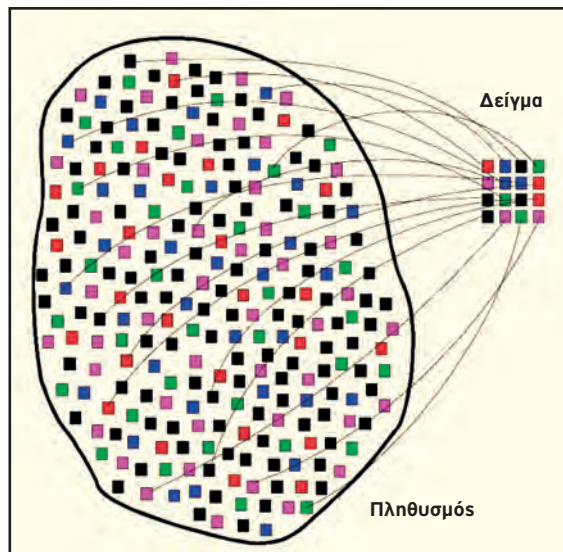
- 1) Εξετάζουμε ένα δείγμα υπαλλήλων μιας εταιρείας.
- 2) Εξετάζουμε ένα δείγμα προϊόντων από μια παραγωγή.
- 3) Εξετάζουμε ένα δείγμα τηλεθεατών.
- 4) Εξετάζουμε τους καλαθοσφαιριστές μιας ομάδας σε έναν αγώνα.

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
Λεωφορείο	Λευκό	34	25
Μοτοσικλέτα	Μαύρο	62	2

Ποιες είναι οι μεταβλητές της έρευνας και ποιό το είδος τους;

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 ολόκληρη την κοινωνία.



## Πρόσθετο υλικό

Η διαδικασία επιλογής δείγματος ονομάζεται **δειγματοληψία**. Για να είναι το δείγμα και αντιπροσωπευτικό, συνήθως χρησιμοποιούνται μέθοδοι κατά τις οποίες κάθε μέλος του πληθυσμού έχει ίσες πιθανότητες επιλογής.

### Ιστορικά στοιχεία

Στις προεδρικές εκλογές στις ΗΠΑ το 1936 υπήρξε πλήρης αποτυχία της πρόγνωσης των εκλογικών αποτελεσμάτων από την εταιρεία «Literary Digest». Η έρευνα, που διενήργησε η εταιρεία, στηρίχτηκε σε αρχικό δείγμα 10.000.000 ατόμων του πληθυσμού από λίστα τηλεφωνικών καταλόγων των διάφορων πόλεων και από τηλεφωνικούς καταλόγους ιδιοκτητών αυτοκινήτων. Στο ερωτηματολόγιο απάντησαν περίπου 2.400.000 άτομα και από την ανάλυση που πραγματοποιήθηκε φάνηκε ότι το εκλογικό σώμα θα έδινε μεγάλη πλειοψηφία στους Ρεπουμπλικάνους με υποψήφιο πρόεδρο τον Landon. Η μέρα των εκλογών όμως επιφύλαξε μια δυσάρεστη έκπληξη στους στατιστικούς αναλυτές της δημοσκόπησης. Ο Δημοκρατικός υποψήφιος Roosevelt εκλέχτηκε με την ιστορική πλειοψηφία 60,7%, σε αντίθεση με το 40,9% που είχε προβλέψει η δημοσκόπηση της «Literary Digest».

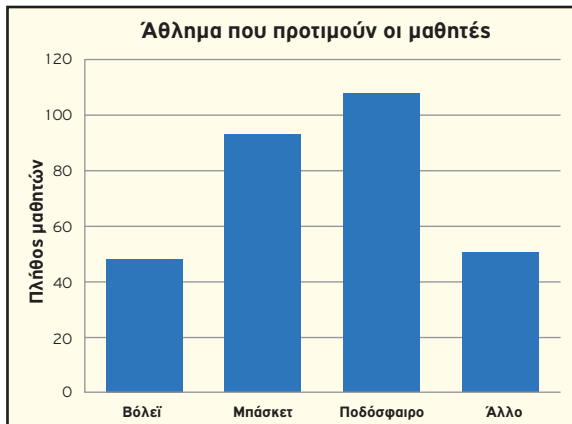
Γιατί νομίζετε η έρευνα απέτυχε και μάλιστα σε τόσο μεγάλο βαθμό;

## ΕΝΟΤΗΤΑ 3.2: ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

### Διερεύνηση

Μια εταιρεία αθλητικών ειδών, πριν επενδύσει σε αθλητικά είδη που προτιμούν οι μαθητές, αποφάσισε να κάνει μια έρευνα. Για το λόγο αυτό επέλεξε, με τυχαίο τρόπο, δείγμα τριακοσίων μαθητών απ' όλη την Ελλάδα. Ο υπεύθυνος που έκανε την έρευνα, μετά την επεξεργασία των στοιχείων που συγκέντρωσε, παρουσίασε στο διευθυντή της εταιρείας τον παρακάτω πίνακα και τα δυο διαγράμματα.

Άθλημα	Πλήθος μαθητών	Ποσοστό
Βόλεϊ	48	16%
Μπάσκετ	93	31%
Ποδόσφαιρο	108	36%
Άλλο	51	17%
<b>Σύνολο</b>	<b>300</b>	<b>100%</b>



- 1) Ποια είναι η μεταβλητή της έρευνας και ποιο το είδος της;
- 2) Πως προκύπτουν τα αντίστοιχα ποσοστά για κάθε άθλημα;
- 3) Ποιο είναι το ύψος της κάθε μπάρας στο ραβδόγραμμα;
- 4) Ποια είναι η γωνία του κάθε κυκλικού τομέα στο κυκλικό διάγραμμα;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

### Ορισμοί

- Έστω  $X$  μια μεταβλητή με τιμές  $x_1, x_2, x_3, \dots, x_k$ . **Συχνότητα**  $v_i$  μιας τιμής  $x_i$  λέγεται ο φυσικός αριθμός που εκφράζει πόσες φορές εμφανίζεται η τιμή αυτή στο σύνολο των παρατηρήσεων του δείγματος. Είναι φανερό ότι το άθροισμα όλων των συχνοτήτων μας δίνει το μέγεθος  $v$  του δείγματος. Δηλαδή:

$$v_1 + v_2 + v_3 + \dots + v_k = v$$

- Η **σχετική συχνότητα**  $f_i$  μιας τιμής  $x_i$  ορίζεται ως ο λόγος της αντίστοιχης συχνότητας  $v_i$  προς το μέγεθος  $v$  του δείγματος. Δηλαδή:

$$f_i = \frac{v_i}{v} \text{ για } i = 1, 2, 3, \dots, k$$

Η σχετική συχνότητα μπορεί να εκφραστεί και ως ποσοστό  $f_i\%$ .

- Σύμφωνα με τους παραπάνω ορισμούς, ο πίνακας του παραδείγματος που αναφέραμε με μεταβλητή «το είδος της μουσικής που προτιμούν οι μαθητές» μετατρέπεται, όπως φαίνεται στη συνέχεια και λέγεται **πίνακας κατανομής συχνοτήτων και σχετικών συχνοτήτων**.

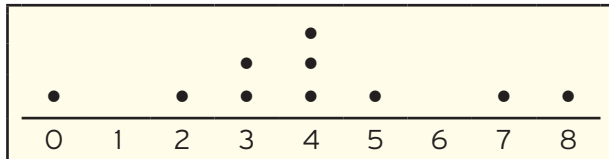
Άθλημα $x_i$	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
Βόλεϊ	48	0,16	16
Μπάσκετ	93	0,31	31
Ποδόσφαιρο	108	0,36	36
Άλλο	51	0,17	17
<b>Σύνολο</b>	<b>300</b>	<b>1,00</b>	<b>100</b>

- Οι πλέον συνηθισμένοι τρόποι γραφικής παρουσίασης ποιοτικών αλλά και ποσοτικών διακριτών δεδομένων είναι το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα.
  - Το **ραβδόγραμμα συχνοτήτων** αποτελείται από ορθογώνιες στήλες, μια για κάθε τιμή της μεταβλητής, όπου το ύψος της κάθε στήλης είναι ίσο με την αντίστοιχη συχνότητα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **ραβδόγραμμα σχετικών συχνοτήτων**.
  - Το **κυκλικό διάγραμμα** χρησιμοποιείται επίσης για τη γραφική παράσταση δεδομένων, κυρίως όταν αυτά παίρνουν λίγες τιμές. Η γωνία  $\alpha_i$  του κάθε κυκλικού

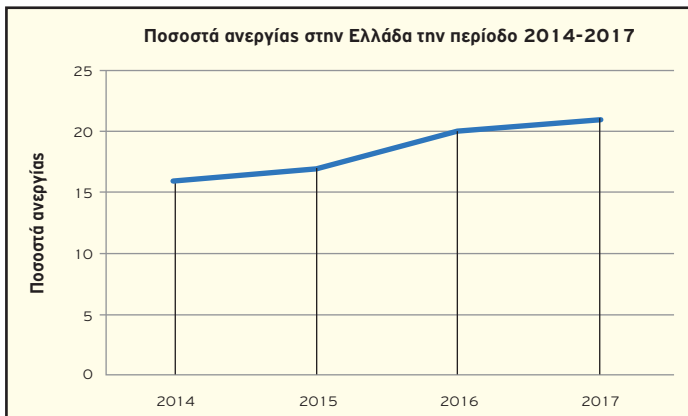
τομέα, είναι ανάλογη της αντίστοιχης σχετικής συχνότητας. Δηλαδή:

$$a_i = 360^\circ \cdot f_i \text{ για } i = 1, 2, 3, \dots, k$$

- Όταν έχουμε λίγες παρατηρήσεις, τότε η κατανομή τους μπορεί να περιγραφεί με το **σημειόγραμμα** στο οποίο οι τιμές παριστάνονται με σημεία υπεράνω ενός άξονα.



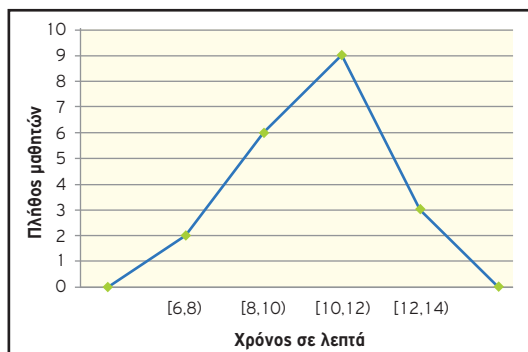
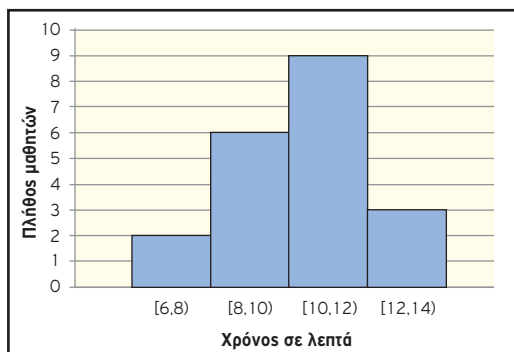
- Το **χρονόγραμμα** χρησιμοποιείται για την γραφική απεικόνιση της εξέλιξης σε σχέση με το χρόνο ενός μεγέθους, συνήθως οικονομικού ή δημογραφικού.



- Στη περίπτωση που έχουμε ποσοτικά συνεχή δεδομένα με πολλές διαφορετικές τιμές, τότε, για την καλύτερη παρουσίασή τους, γίνεται ομαδοποίηση αυτών σε κλάσεις, συνήθως ίσου πλάτους. Σχετικός είναι ο παρακάτω πίνακας με τους χρόνους που χρειάστηκαν οι 20 μαθητές ενός τμήματος για να απαντήσουν σε ένα πρόβλημα.

Χρόνος σε λεπτά των μαθητών	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
[6,8)	2	0,10	10
[8,10)	6	0,30	30
[10,12)	9	0,45	45
[12,14)	3	0,15	15
<b>Σύνολο</b>	<b>20</b>	<b>1,00</b>	<b>100</b>

- Η γραφική παρουσίαση ομαδοποιημένων στατιστικών δεδομένων γίνεται με το **ιστόγραμμα συχνοτήτων**. Στο παρακάτω σχήμα φαίνεται η γραφική παρουσίαση των δεδομένων του προηγούμενου πίνακα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **ιστόγραμμα σχετικών συχνοτήτων**.
- Αν θεωρήσουμε δυο επιπλέον κλάσεις ίσου πλάτους, μια στη αρχή και μια στο τέλος, με συχνότητα 0 και ενώσουμε τα μέσα των άνω βάσεων, τότε προκύπτει το λεγόμενο **πολύγωνο συχνοτήτων**. Σχετικό είναι το παρακάτω σχήμα. Αν, αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **πολύγωνο σχετικών συχνοτήτων**.



## Εφαρμογές

### Εφαρμογή 1η

Σε μια δημοσκόπηση που έγινε για τις δημοτικές εκλογές, 280 άτομα απάντησαν ότι προτιμούν τον υποψήφιο «Α», 320 άτομα τον υποψήφιο «Β» και 200 άτομα τον υποψήφιο «Γ».

- 1) Ποιο ήταν το μέγεθος του δείγματος;
- 2) Να κάνετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- 3) Να παραστήσετε τα δεδομένα με ραβδόγραμμα αλλά και με κυκλικό διάγραμμα.

### Απάντηση

- 1) Το πλήθος αυτών που απάντησαν είναι:

$$v = 280 + 320 + 200 = 800$$

Επομένως το μέγεθος του δείγματος είναι 800 άτομα.

- 2) Οι συχνότητες είναι:  $v_1 = 280$ ,  $v_2 = 320$  και  $v_3 = 200$ . Οι σχετικές συχνότητες είναι:

$$f_1 = \frac{v_1}{v} = \frac{280}{800} = 0,35, f_2 = \frac{v_2}{v} = \frac{320}{800} = 0,40 \text{ και } f_3 = \frac{v_3}{v} = \frac{200}{800} = 0,25$$

Ο πίνακας κατανομής συχνοτήτων και σχετικών συχνοτήτων είναι:

Υποψήφιος δήμαρχος	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
Υποψήφιος «Α»	280	0,35	35
Υποψήφιος «Β»	320	0,40	40
Υποψήφιος «Γ»	200	0,25	25
<b>Σύνολο</b>	<b>800</b>	<b>1,00</b>	<b>100</b>

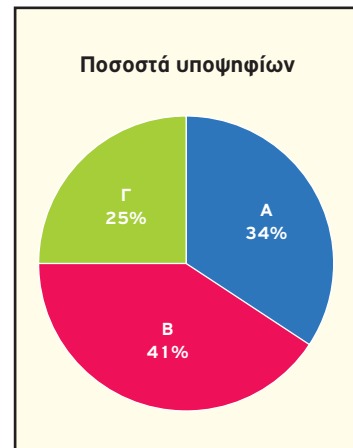
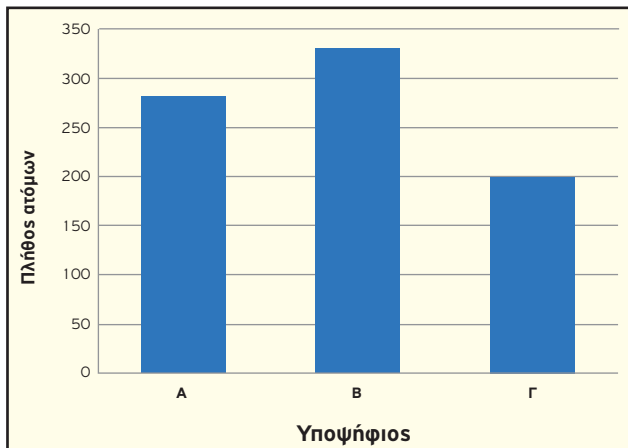
3) Για τις γωνίες των κυκλικών τομέων στο κυκλικό διάγραμμα έχουμε:

$$\alpha_1 = 360^\circ \cdot f_1 = 360^\circ \cdot 0,35 = 126^\circ$$

$$\alpha_2 = 360^\circ \cdot f_2 = 360^\circ \cdot 0,40 = 144^\circ$$

$$\alpha_3 = 360^\circ \cdot f_3 = 360^\circ \cdot 0,25 = 90^\circ$$

Παρακάτω φαίνονται το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα.





**Εφαρμογή 2η**

Παρακάτω δίνονται οι χρόνοι, στρογγυλοποιούμενοι στο δέκατο του δευτερολέπτου, που απαιτήθηκαν για να τρέξουν 50 αθλητές έναν αγώνα δρόμου 1000 m.

52,1	55,3	50	56,4	59,1	54,2	56,7	54,4	57,1	53,7
55,2	55,1	58	59,2	56	55,5	52,5	56,5	58,5	55
55,2	57,3	54,3	53,5	57,9	53	55,4	55,6	52,4	54,5
56,4	59,1	54,2	56,7	55,3	52,4	56,4	54,1	54,3	56,7
54,3	51,5	57	53,2	54,9	55,6	52	55,3	55,1	54,7

- 1) Ποιο είναι το είδος της μεταβλητής, ποιος ο μικρότερος και ποιος ο μεγαλύτερος χρόνος;
- 2) Τι έχετε να παρατηρήσετε για το πλήθος των τιμών των παρατηρήσεων σε σχέση με το πλήθος των παρατηρήσεων;
- 3) Ξεκινώντας από το μικρότερο χρόνο και με βήμα 2 sec, ποιες κλάσεις της μορφής [α,β) δημιουργούνται στις οποίες περιέχονται όλες οι παρατηρήσεις;
- 4) Να παρουσιαστούν οι παρατηρήσεις, ομαδοποιημένες στις παραπάνω κλάσεις ίσου πλάτους, σε έναν πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- 5) Να παραστήσετε τα δεδομένα με ιστόγραμμα συχνοτήτων και με πολύγωνο συχνοτήτων.

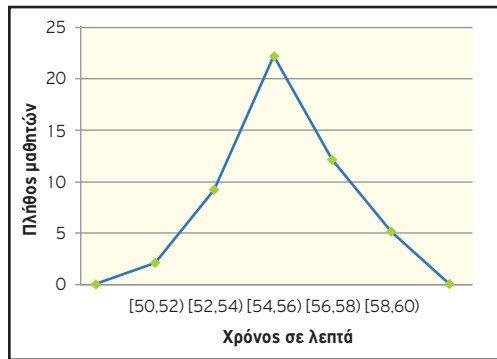
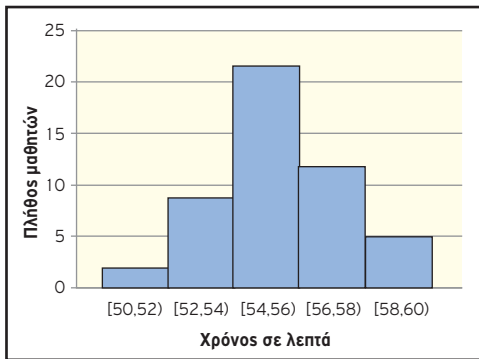
**Απάντηση**

- 1) Η μεταβλητή είναι ποσοτική συνεχής. Ο μικρότερος χρόνος είναι 50 sec και ο μεγαλύτερος 59,2 sec.
- 2) Το πλήθος των τιμών των παρατηρήσεων είναι 37 σε σύνολο 50 παρατηρήσεων. Για να παρουσιαστούν τα δεδομένα, ως έχουν, σε ένα πίνακα συχνοτήτων θα χρειαστούμε 37 γραμμές. Είναι φανερό ότι ένας τέτοιος πίνακας είναι δύσχρηστος και οι πληροφορίες δεν παρουσιάζονται συνοπτικά.
- 3) Οι κλάσεις που δημιουργούνται είναι: [50,52), [52,54), [54,56), [56,58) και [58,60)

4) Μετά τη διαλογή ο πίνακας συχνοτήτων και σχετικών συχνοτήτων είναι ο παρακάτω.

Κλάσεις με χρόνους σε sec	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
[50,52)	2	0,04	4
[52,54)	9	0,18	18
[54,56)	22	0,44	44
[56,58)	12	0,24	24
[58,60)	5	0,10	10
<b>Σύνολο</b>	<b>50</b>	<b>1,00</b>	<b>100</b>

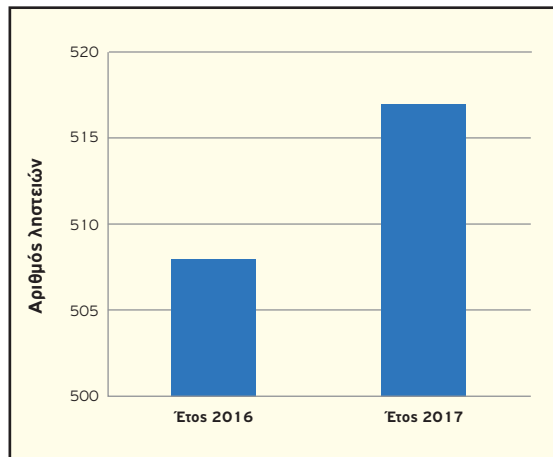
5) Παρακάτω φαίνονται το ιστόγραμμα και το πολύγωνο συχνοτήτων.



### Εφαρμογή 3η

Σε ένα τηλεοπτικό κανάλι, ένας δημοσιογράφος σχολίασε την διπλανή γραφική παράσταση ως εξής: «*Η γραφική παράσταση δείχνει ότι σημειώθηκε τεράστια αύξηση του αριθμού των ληστειών το έτος 2017 σε σχέση με το έτος 2016*».

Νομίζετε ότι ο δημοσιογράφος του καναλιού αυτού ερμήνευσε σωστά την γραφική παράσταση; Να γράψετε ένα επιχειρήμα που να τεκμηριώνει την απάντησή σας.

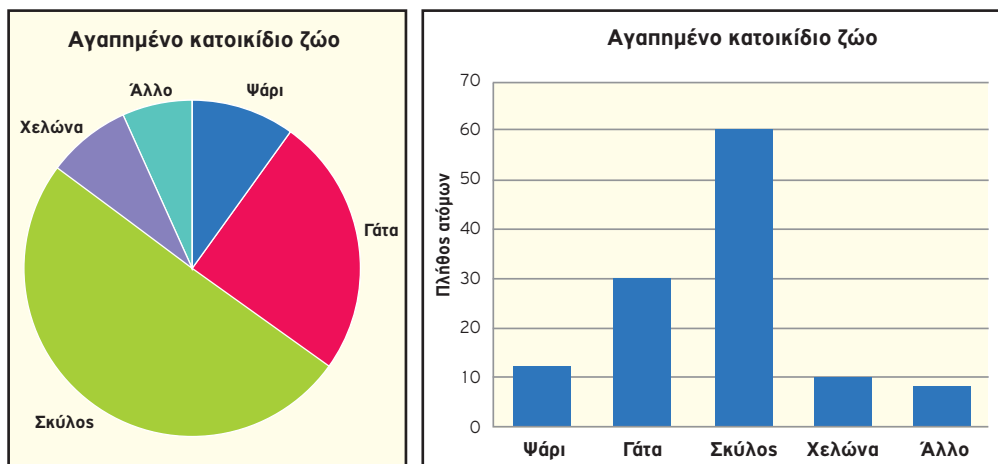


### Απάντηση

Για το λόγο ότι η δεύτερη στήλη φαίνεται περίπου διπλάσια στο ύψος από την πρώτη στήλη, δόθηκε η λανθασμένη ερμηνεία. Όμως αυτό που πραγματικά φαίνεται είναι το άνω τμήμα των στήλων, αφού ο άξονας των τεταγμένων αρχίζει από το 500. Έτσι, το πραγματικό ύψος της πρώτης στήλης είναι 508, ενώ της δεύτερης είναι 517. Επομένως ο αριθμός των ληστειών αυξήθηκε μόνο κατά 9 σε σύνολο 500 και πλέον ληστειών. Το γεγονός αυτό δεν δικαιολογεί την παραπάνω ερμηνεία.

### **Εφαρμογή 4η**

Ρωτήθηκαν 120 άτομα για το αγαπημένο τους κατοικίδιο ζώο και τα αποτελέσματα παρουσιάζονται με ένα κυκλικό διάγραμμα και ένα ραβδόγραμμα.



- 1) Να κάνετε μια ερώτηση που μπορεί να απαντηθεί πιο εύκολα από το ραβδόγραμμα.
- 2) Να κάνετε μια ερώτηση που μπορεί να απαντηθεί πιο εύκολα από το κυκλικό διάγραμμα.

### Απάντηση

- 1) Για το ραβδόγραμμα μια ερώτηση με εύκολη απάντηση μπορεί να είναι: «Πόσες φορές είναι περισσότερα τα άτομα που το αγαπημένο τους κατοικίδιο ζώο είναι ο σκύλος σε σχέση με αυτά που είναι η χελώνα»;
- 2) Για το κυκλικό διάγραμμα μια ερώτηση με εύκολη απάντηση μπορεί να είναι: «Ποιο είναι το ποσοστό των ατόμων που το αγαπημένο τους κατοικίδιο ζώο είναι ο σκύλος»;

## Ασκήσεις - Προβλήματα - Δραστηριότητες

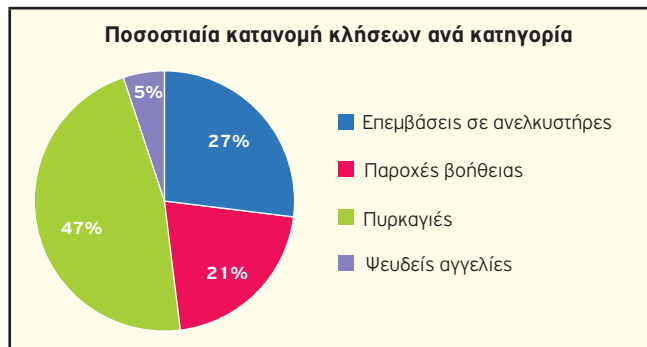
1) Το επάγγελμα του πατέρα 20 μαθητών καταγράφηκε στον διπλανό πίνακα. Να κάνετε πίνακα σχετικών συχνοτήτων και να κατασκευάσετε το αντίστοιχο ραβδόγραμμα συχνοτήτων καθώς και το κυκλικό διάγραμμα.

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

2) 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

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.



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.

Επίπεδο μόρφωσης	Ποσοστό (%)
Πτυχιούχοι Α.Ε.Ι.	20
Πτυχιούχοι Τ.Ε.Ι.	50
Απόφοιτοι Λυκείου	30

2) Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με κυκλικό διάγραμμα.

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

2) Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με σημειόγραμμα.

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.

7) Στον πιο κάτω πίνακα δίνεται η συγκέντρωση ( $\text{mg}/\text{cm}^3$ ) 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.

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.

## Πρόσθετο υλικό

Η εκμάθηση των βασικών εννοιών της στατιστικής γίνεται πιο ενδιαφέρουσα, όταν μπορούμε να εργαστούμε με πληροφορίες που συλλέγονται από εμάς και αποθηκεύονται σε έναν υπολογιστή. Η απλούστερη συλλογή δεδομένων μπορεί να αφορά πληροφορίες για τους συμμαθητές σας. Η έρευνα που ακολουθεί μπορεί να σας δώσει αρκετά δεδομένα, με τα οποία μπορείτε να απαντήσετε σε μερικές ενδιαφέρουσες ερωτήσεις. Αποθηκεύστε τα δεδομένα που θα συλλέξετε.

- 1) Φύλο
- 2) Ηλικία
- 3) Ύψος σε cm
- 4) Ύψος του πατέρα σας σε cm
- 5) Ύψος της μητέρας σας σε cm
- 6) Το πέμπτο ψηφίο του αριθμού της ταυτότητάς σας
- 7) Το τελευταίο ψηφίο του αριθμού ενός αυτοκινήτου της οικογένειάς σας
- 8) Χρώμα μαλλιών
- 9) Χρώμα ματιών

Μπορείτε να προσθέσετε μια ή περισσότερες ερωτήσεις που σας ενδιαφέρουν. Μοιράστε ένα έντυπο ερωτηματολόγιου, όπου κάθε μαθητής θα γράφει τις απαντήσεις του. Επεξεργαστείτε τα δεδομένα που συγκεντρώσατε με ένα λογιστικό φύλλο.

- 1) Προσδιορίστε το είδος κάθε μεταβλητής
- 2) Κατασκευάστε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων για κάθε μεταβλητή ξεχωριστά.
- 3) Κατασκευάστε τα αντίστοιχα διαγράμματα ανάλογα με το είδος της μεταβλητής.

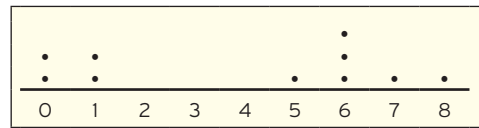
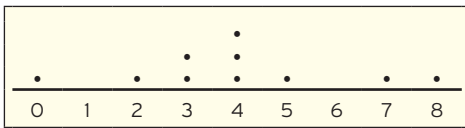
ΥΠΟΔΕΙΓΜΑ

α/α	Φύλο	Ηλικία	Ύψος	Ύψος πατέρα	Ύψος μητέρας	Τρίτο ψηφίο του αριθ. της ταυτότητας	Τελευταίο ψηφίο του αριθ. του αυτοκινήτου	Χρώμα μαλλιών	Χρώμα ματιών
1	θ	17	167	170	160	8	1	Μ	Μ
2	θ	18	165	172	163	4	7	Κ	Μ
3	Α	18	184	185	171	5	5	Κ	Κ
4	θ	17	175	179	168	3	8	Ξ	Μ
5	Α	19	171	175	165	9	4	Κ	Κ
6	Α	18	169	170	167	6	7	Μ	Κ

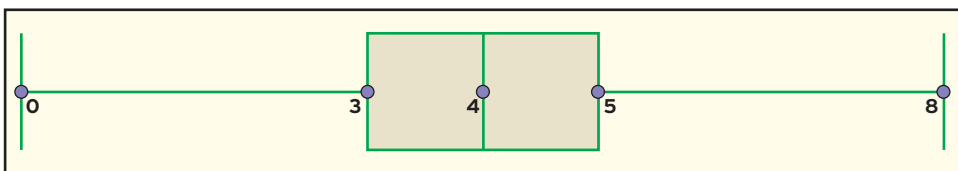
## ΕΝΟΤΗΤΑ 3.3: ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ ΘΗΚΟΓΡΑΜΜΑ ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ

### Διερεύνηση

- 1) Στα επόμενα σημειογράμματα φαίνονται οι πόντοι που σημείωσαν οι 10 καλαθοσφαιριστές κάθε μιας από τις δυο ομάδες μπάσκετ, στο ημίχρονο ενός αγώνα.



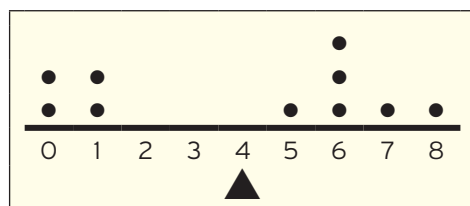
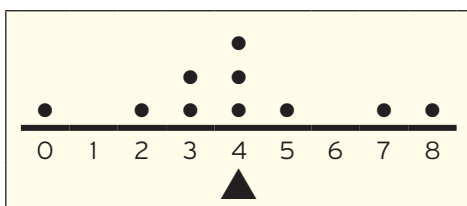
- 1) Να υπολογίσετε τη μέση τιμή των πόντων που σημείωσαν οι 10 καλαθοσφαιριστές κάθε ομάδας στο ημίχρονο.
- 2) Αφού γράψετε τις 10 παρατηρήσεις κάθε ομάδας σε αύξουσα σειρά, να βρείτε για κάθε ομάδα:
  - 1) Τη μικρότερη παρατήρηση.
  - 2) Τη μεσαία παρατήρηση από τις πέντε πρώτες, δηλαδή το πρώτο τεταρτημόριο.
  - 3) Το ημίθροισμα των δυο μεσαίων παρατηρήσεων, δηλαδή το δεύτερο τεταρτημόριο ή αλλιώς τη διάμεσο.
  - 4) Τη μεσαία παρατήρηση από τις πέντε τελευταίες, δηλαδή το τρίτο τεταρτημόριο.
  - 5) Τη μεγαλύτερη παρατήρηση
- 3) Με βάση τα παραπάνω, το θηκόγραμμα για την πρώτη ομάδα είναι αυτό που βλέπετε στο επόμενο σχήμα. Να σχεδιάσετε το θηκόγραμμα για τη δεύτερη ομάδα.



- 4) Ποια είναι η διαφορά της μικρότερης από τη μεγαλύτερη παρατήρηση, δηλαδή το εύρος των παρατηρήσεων, για κάθε ομάδα; Είναι το εύρος ένα αξιόπιστο μέτρο μεταβλητότητας;



- 5) Ποια είναι η διαφορά του πρώτου από το τρίτο τεταρτημόριο, δηλαδή το ενδοτεταρτημοριακό εύρος, για κάθε ομάδα;
  - 6) Ποια τιμή εμφανίζεται τις περισσότερες φορές, δηλαδή ποια είναι η επικρατούσα τιμή για κάθε ομάδα;
  - 7) Πώς θα υπολογίσουμε τη μέση τιμή των τετραγώνων των διαφορών της μέσης τιμής των παρατηρήσεων από τις παρατηρήσεις, δηλαδή τη διακύμανση των παρατηρήσεων; Περιγράψτε τη διαδικασία και τις αριθμητικές πράξεις που απαιτούνται για κάθε ομάδα ξεχωριστά.
  - 8) Να υπολογίσετε την τετραγωνική ρίζα της διακύμανσης, δηλαδή την τυπική απόκλιση για κάθε ομάδα.
- 2) Στα επόμενα δύο σχήματα οι κουκίδες συμβολίζουν δυο ομάδες, από 10 βαρίδια ίδιου βάρους η κάθε μια, τοποθετημένα πάνω σε δυο αβαρείς ράβδους αντίστοιχα, αριθμημένες έτσι ώστε οι αποστάσεις των διαδοχικών αριθμών να είναι ίσες.



- 1) Να εξηγήσετε το λόγο για τον οποίο, αν τοποθετήσουμε μια τριγωνική σφήνα στη θέση 4 σε κάθε ράβδο, τότε αυτές ισορροπούν. Ποια είναι η φυσική ερμηνεία της μέσης τιμής των βαρών σε σχέση με το σημείο ισορροπίας αυτών;
- 2) Είναι το εύρος, στην παραπάνω περίπτωση, ένα αξιόπιστο μέτρο μεταβλητότητας των βαρών των δυο ομάδων;
- 3) Ποια είναι η μέση τιμή των τετραγώνων των διαφορών των παρατηρήσεων από τη μέση τιμή τους, δηλαδή η διακύμανση των βαρών κάθε ομάδας;
- 4) Εκφράζονται η διακύμανση και οι παρατηρήσεις με την ίδια μονάδα μέτρησης;
- 5) Ποια είναι η τετραγωνική ρίζα της διακύμανσης, δηλαδή η τυπική απόκλιση των βαρών κάθε ομάδας;
- 6) Εκφράζονται η τυπική απόκλιση και οι παρατηρήσεις με την ίδια μονάδα μέτρησης;
- 7) Να εκφράσετε το πηλίκο  $\frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}}$  ως ποσοστό για κάθε ομάδα βαρών.  
Ποια ομάδα έχει μικρότερο πηλίκο οπότε είναι περισσότερο ομοιογενής;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

### Ορισμοί για τα μέτρα θέσης ποσοτικών δεδομένων

Η **μέση τιμή**  $\bar{x}$  ενός συνόλου  $n$  παρατηρήσεων ορίζεται, ως το άθροισμα των παρατηρήσεων διά του πλήθους τους.

$$\bar{x} = \frac{t_1 + t_2 + t_3 + \dots + t_n}{n}$$

Αν για παράδειγμα οι βαθμοί ενός μαθητή σε δέκα μαθήματα είναι: 18, 17, 19, 17, 18, 16, 18, 17, 16, 18, τότε:

$$\bar{x} = \frac{18+17+19+17+18+16+18+17+16+18}{10} = \frac{174}{10} = 17,4$$

Αν οι τιμές των παρατηρήσεων είναι  $x_i$  με αντίστοιχες συχνότητες  $v_i$ , τότε:

$$\bar{x} = \frac{x_1 \cdot v_1 + x_2 \cdot v_2 + \dots + x_k \cdot v_k}{n}$$

Για το παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα έχουμε:

$$\bar{x} = \frac{16 \cdot 2 + 17 \cdot 3 + 18 \cdot 4 + 19 \cdot 1}{10} = \frac{174}{10} = 17,4$$

$x_i$	$v_i$	$x_i \cdot v_i$
16	2	32
17	3	51
18	4	72
19	1	19
Σύνολο	10	174

• **Διάμεσος  $\delta$**  ενός δείγματος  $n$  παρατηρήσεων, οι οποίες έχουν διαταχθεί σε αύξουσα σειρά, ορίζεται ως ακολούθως:

- η μεσαία παρατήρηση, όταν το  $n$  είναι περιττός αριθμός,
- το ημίαθροισμα των δύο μεσαίων παρατηρήσεων, όταν το  $n$  είναι άρτιος αριθμός.

Για το παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα, αν τους διατάξουμε σε αύξουσα σειρά, έχουμε: 16, 16, 17, 17, 17, 18, 18, 18, 18, 19. Επομένως:

$$\delta = \frac{17+18}{2} = 17,5$$

Αν έχουμε τις παρατηρήσεις 4, 4, 5, 5, 6, 7, 39 σε αύξουσα σειρά, τότε η διάμεσος είναι  $\delta = 5$  ενώ η μέση τιμή είναι:

$$\bar{x} = \frac{4+4+5+5+6+7+39}{7} = \frac{70}{7} = 10$$

Παρατηρούμε ότι, στην παραπάνω περίπτωση, η διάμεσος είναι ένα μέτρο θέσης που «αντιπροσωπεύει» καλύτερα τις παρατηρήσεις.

Είναι φανερό ότι το πολύ 50% των παρατηρήσεων είναι μικρότερες και το πολύ 50% είναι μεγαλύτερες από τη διάμεσο  $\delta$ .

- Τα **τεταρτημόρια** συμβολίζονται με  $Q_1$ ,  $Q_2$  και  $Q_3$ .

- Για το  $Q_1$  έχουμε αριστερά το πολύ 25% και δεξιά το πολύ 75% των παρατηρήσεων.
- Για το  $Q_2$  έχουμε  $Q_2 = \delta$ , δηλαδή συμπίπτει με τη διάμεσο.
- Για το  $Q_3$  έχουμε αριστερά το πολύ 75% και δεξιά το πολύ 25% των παρατηρήσεων.

Για το ευκολότερο υπολογισμό των  $Q_1$  και  $Q_3$  έχουμε:

- Όταν το πλήθος των διατεταγμένων παρατηρήσεων είναι άρτιος αριθμός, διακρίνουμε το πρώτο μισό, του οποίου η διάμεσος είναι το  $Q_1$  και το δεύτερο μισό, του οποίου η διάμεσος είναι το  $Q_3$ .
- Όταν το πλήθος των διατεταγμένων παρατηρήσεων είναι περιττός αριθμός, αφαιρούμε από το δείγμα τη διάμεσο και διακρίνουμε το πρώτο μισό, του οποίου η διάμεσος είναι το  $Q_1$  και το δεύτερο μισό, του οποίου η διάμεσος είναι το  $Q_3$ .

Για το παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα είναι:

$$Q_1 = 17, Q_2 = \delta = 17,5, Q_3 = 18$$

- **Επικρατούσα τιμή  $M_0$**  είναι η παρατήρηση με τη μεγαλύτερη συχνότητα. Στο παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα, επικρατούσα τιμή είναι  $M_0 = 18$ . Η επικρατούσα τιμή μπορεί να μην είναι μοναδική. Όταν όλες οι παρατηρήσεις είναι διαφορετικές, τότε λέμε ότι δεν υπάρχει επικρατούσα τιμή.

### Ορισμοί για τα μέτρα διασποράς ποσοτικών δεδομένων

- Το **εύρος  $R$**  είναι το απλούστερο από τα μέτρα διασποράς και ορίζεται ως η διαφορά της ελάχιστης  $x_{\min}$  παρατήρησης από τη μέγιστη  $x_{\max}$  παρατήρηση. Δηλαδή:

$$R = \text{μέγιστη παρατήρηση} - \text{ελάχιστη παρατήρηση} = x_{\max} - x_{\min}$$

- Το **ενδοτεταρτημοριακό εύρος  $Q$**  είναι η διαφορά του πρώτου τεταρτημορίου  $Q_1$  από το τρίτο τεταρτημόριο  $Q_3$ . Δηλαδή:

$$Q = Q_3 - Q_1$$

Όσες παρατηρήσεις βρίσκονται έξω από το διάστημα  $[Q_1 - 1,5 \cdot Q, Q_3 + 1,5 \cdot Q]$  ονομάζονται **ακραίες τιμές**. Είναι φανερό ότι η διάμεσος είναι ένα μέτρο θέσης το οποίο δεν επηρεάζεται από τις ακραίες τιμές.

Το εύρος και το ενδοτεταρτημοριακό εύρος είναι πολύ απλά μέτρα διασποράς και δεν εκφράζουν την απόκλιση που έχει κάθε παρατήρηση από το «κέντρο» των παρα-

τηρήσεων. Για το λόγο αυτό χρησιμοποιούμε τα επόμενα μέτρα διασποράς.

- Η **διακύμανση**  $s^2$  ορίζεται ως η μέση τιμή των τετραγώνων των διαφορών της μέσης τιμής των παρατηρήσεων από τις παρατηρήσεις. Δηλαδή:

$$s^2 = \frac{(t_1 - \bar{x})^2 + (t_2 - \bar{x})^2 + (t_3 - \bar{x})^2 + \dots + (t_v - \bar{x})^2}{v}$$

Για παράδειγμα, αν οι βαθμοί έξι μαθητών στα Μαθηματικά είναι: 18, 15, 18, 18, 15, 18, τότε:

$$\bar{x} = \frac{18+15+18+18+15+18}{6} = \frac{102}{6} = 17 \text{ και}$$

$$s^2 = \frac{(18-17)^2 + (15-17)^2 + (18-17)^2 + (18-17)^2 + (15-17)^2 + (18-17)^2}{6} = \frac{12}{6} = 2$$

Αν οι τιμές των παρατηρήσεων είναι  $x_i$  με αντίστοιχες συχνότητες  $v_i$ , τότε:

$$s^2 = \frac{(x_1 - \bar{x})^2 \cdot v_1 + (x_2 - \bar{x})^2 \cdot v_2 + \dots + (x_k - \bar{x})^2 \cdot v_k}{v}$$

Για το παράδειγμα των βαθμών των έξι μαθητών στα Μαθηματικά, σχετικός είναι ο επόμενος πίνακας:

$x_i$	$v_i$	$x_i \cdot v_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot v_i$
15	2	30	-2	4	8
18	4	72	1	1	4
Σύνολο	6	102			12

$\bar{x} = \frac{102}{6} = 17$   
 και  
 $s^2 = \frac{12}{6} = 2$

Η διακύμανση δεν εκφράζεται με τις μονάδες που εκφράζονται οι παρατηρήσεις.

- Η **τυπική απόκλιση**  $s$  είναι:  $s = \sqrt{s^2}$ . Στο παραπάνω παράδειγμα είναι:  $s = \sqrt{2} \approx 1,41$ . Η τυπική απόκλιση έχει το πλεονέκτημα ότι εκφράζεται με τις μονάδες, που εκφράζονται οι παρατηρήσεις.

### Θηκόγραμμα

- Έστω ότι έχουμε τον αριθμό των εργαζομένων σε 20 βιοτεχνίες.

10	14	25	7	31	8	12	19	10	24	9	13	5	28	24	19	26	51	68	14
----	----	----	---	----	---	----	----	----	----	---	----	---	----	----	----	----	----	----	----

Κατατάσσουμε τις παρατηρήσεις σε αύξουσα σειρά:

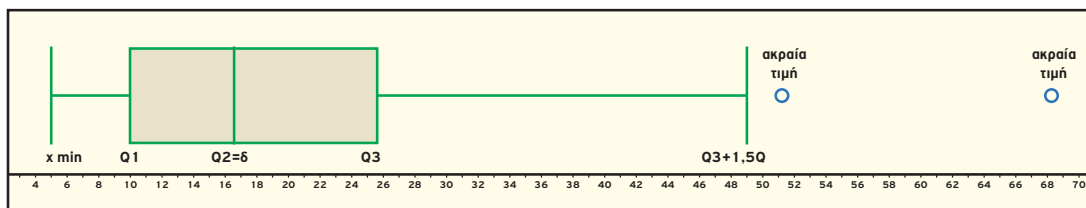
5, 7, 8, 9, 10, 10, 12, 13, 14, 14, 19, 19, 24, 24, 25, 26, 28, 31, 51, 68

Είναι:  $Q_1 = \frac{10+10}{2} = 10$ ,  $Q_2 = \delta = \frac{14+19}{2} = 16,5$ ,  $Q_3 = \frac{25+26}{2} = 25,5$

Το ενδοτεταρτημοριακό εύρος είναι:  $Q = 25,5 - 10 = 15,5$  και  $1,5 \cdot Q = 23,25$

Επομένως  $[Q_1 - 1,5 \cdot Q, Q_3 + 1,5 \cdot Q] = [-13,25, 48,75]$ . Οπότε ακραίες παρατηρήσεις είναι το 51 και το 68.

Τα μέτρα αυτά απεικονίζονται στο επόμενο διάγραμμα που λέγεται **θηκόγραμμα**.



Από τα μέσα των πλευρών, που παριστάνουν το πρώτο και το τρίτο τεταρτημόριο, φέρουμε ευθύγραμμα τμήματα με μήκος το πολύ  $1,5 \cdot Q$ .

Η ελάχιστη τιμή, δηλαδή το 5, είναι σημείο του ευθύγραμμου τμήματος που φέρουμε από το  $Q_1$ , οπότε το άκρο είναι το 5.

Η μέγιστη τιμή, δηλαδή το 68 δεν είναι σημείο του ευθύγραμμου τμήματος που φέρουμε από το  $Q_3$  με μήκος  $1,5 \cdot Q$ , οπότε το άκρο είναι το  $25,5 + 23,25 = 48,75$

Τέλος σημειώνουμε τις ακραίες τιμές, δηλαδή το 51 και το 68.

### Συντελεστής μεταβλητότητας

- Δυο όμιλοι επιχειρήσεων, που ο καθένας αποτελείται από 5 επιχειρήσεις, είχαν ετήσιες δαπάνες για το οικονομικό έτος 2017 τα ποσά, σε χιλιάδες ευρώ, που αναγράφονται στον παρακάτω πίνακα.

<b>Όμιλος Α</b>	200	250	300	300	350
<b>Όμιλος Β</b>	5.000	5.050	5.100	5.100	5.150

Οι μέσες τιμές των δαπανών για κάθε όμιλο είναι:

$$\bar{x}_A = \frac{200 + 250 + 300 + 300 + 350}{5} = 280 \text{ χιλιάδες ευρώ.}$$

$$\bar{x}_B = \frac{5.000 + 5.050 + 5.100 + 5.100 + 5.150}{5} = 5.080 \text{ χιλιάδες ευρώ.}$$

Οι διακυμάνσεις των δαπανών για κάθε όμιλο είναι:

$$s_A^2 = \frac{(200-280)^2 + (250-280)^2 + (300-280)^2 + (300-280)^2 + (350-280)^2}{5} = 2.600$$

$$s_B^2 = \frac{(5.000-5.080)^2 + (5.050-5.080)^2 + (5.100-5.080)^2 + (5.100-5.080)^2 + (5.150-5.080)^2}{5} \\ = 2.600$$

Οι τυπικές αποκλίσεις των δαπανών για κάθε όμιλο είναι ίσες με:

$$s_A = s_B = \sqrt{2.600} \approx 51 \text{ χιλιάδες ευρώ.}$$

Μπορούμε άραγε να ισχυριστούμε ότι οι 51.000 ευρώ των ίσων τυπικών αποκλίσεων έχουν την ίδια βαρύτητα για τους ομίλους Α και Β, που έχουν διαφορετικές μέσες τιμές  $\bar{x}_A = 280.000$  ευρώ και  $\bar{x}_B = 5.080.000$  ευρώ; Ένας τέτοιος ισχυρισμός θα ήταν λανθασμένος, γιατί η σχέση τιμών της μεταβλητής «ετήσιες δαπάνες» με την τιμή της τυπικής απόκλισης, είναι διαφορετικές για τις επιχειρήσεις των δυο ομίλων.

Ένα μέτρο, που μας βοηθάει να αντιμετωπίσουμε τέτοια προβλήματα και μας δίνει τη δυνατότητα συγκρίσεων, είναι ο **συντελεστής μεταβλητότητας CV**, ο οποίος είναι καθαρός αριθμός και εκφράζεται συνήθως ως ποσοστό. Ορίζεται ως ακολούθως:

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

Στο παράδειγμά μας ο συντελεστής μεταβλητότητας για τις δυο επιχειρήσεις είναι:

$$CV_A = \frac{51}{280} \approx 0,1821 = 18,21\% \text{ και } CV_B = \frac{51}{5.080} \approx 0,01004 \approx 1,00\%$$

Όσο μικρότερο είναι το ποσοστό αυτό, τόσο περισσότερη ομοιογένεια υπάρχει στις τιμές της μεταβλητής που μας ενδιαφέρει. Στο παράδειγμά μας είναι  $CV_A > CV_B$ , επομένως η διασπορά των τιμών στον όμιλο Α σε σχέση με τη μέση τιμή είναι μεγαλύτερη από την αντίστοιχη διασπορά στον όμιλο Β.

Γενικά δεχόμαστε ότι ένα δείγμα τιμών μιας μεταβλητής θα είναι **ομοιογνές**, εάν ο συντελεστής μεταβλητότητας δεν ξεπερνά το 10%.

## Εφαρμογές

### Εφαρμογή 1η

Μία ομάδα δέκα μαθητών μέτρησε το μήκος ενός θρανίου με ακρίβεια δέκατου του εκατοστού του μέτρου (χιλιοστού του μέτρου). Οι μαθητές χρησιμοποίησαν την ίδια μετροταινία και κατέγραψαν τις τιμές των δέκα μετρήσεων, όπως φαίνεται στον παρακάτω πίνακα.

Μήκος σε cm	120,2	120,1	120,1	119,8	119,7	120,3	120,2	119,9	120,1	119,8
-------------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

- 1) Να αναφέρετε δύο λόγους, που κατά τη γνώμη σας, δικαιολογούν τις διαφορές στις μετρήσεις του θρανίου.
- 2) Πώς υπολογίζουμε τη μέση τιμή των μετρήσεων;
- 3) Ποια είναι η τιμή του μήκους του θρανίου που θα χρησιμοποιήσουμε;
- 4) Για ποιο λόγο είναι χρήσιμος ο υπολογισμός της μέσης τιμής πολλών μετρήσεων;
- 5) Δικαιολογήστε με παράδειγμα, που θα δημιουργήσετε από τις τιμές του πίνακα, γιατί ζητήθηκε από τους μαθητές να κάνουν δέκα μετρήσεις και όχι λιγότερες (π.χ. τρεις μετρήσεις).

### Απάντηση

- 1) Δύο λόγοι που δικαιολογούν τις διαφορές μπορεί να είναι:
  - 1) Η αρχή της μετροταινίας δεν συμπίπτει με την αρχή της μετρούμενης απόστασης
  - 2) Η μετροταινία δεν ακολουθεί ευθεία και παράλληλη γραμμή προς τη μετρούμενη απόσταση.
- 2) Για τη μέση τιμή των μετρήσεων έχουμε:

$$\begin{aligned}\bar{x} &= \frac{120,2 + 120,1 + 120,1 + 119,8 + 119,7 + 120,3 + 120,2 + 119,9 + 120,1 + 119,8}{10} \\ &= \frac{1200,2}{10} = 120,02\end{aligned}$$

- 3) Για την άγνωστη τιμή του μήκους του θρανίου, μπορούμε να χρησιμοποιήσουμε τη μέση τιμή  $\bar{x} = 120,02$  cm του δείγματος.
- 4) Θέλουμε τη μέση τιμή πολλών μετρήσεων, δηλαδή θέλουμε μεγάλο μέγεθος δείγματος, ώστε το αποτέλεσμα να έχει μεγαλύτερη πιθανότητα να προσεγγίσει ικανοποιητικά το πραγματικό μήκος του θρανίου.

5) Αν για παράδειγμα πάρουμε τη μέση τιμή των τριών πρώτων μετρήσεων έχουμε:

$$\bar{x}' = \frac{120,2 + 120,1 + 120,1}{3} = \frac{360,4}{3} = 120,133... \text{ cm}$$

Αυτή είναι η μέση τιμή ενός άλλου δείγματος. Το μήκος όμως του θρανίου είναι μοναδικό. Το μεγαλύτερο μέγεθος του δείγματος μας δίνει τη δυνατότητα ακριβέστερης εκτίμησης του πραγματικού μήκους.

### Γενικότερα

Σε ένα δείγμα παρατηρήσεων το μέσο  $\bar{x}$  τον λέμε δειγματικό μέσο και την τυπική απόκλιση  $s$  τη λέμε δειγματική τυπική απόκλιση και μπορούμε να τα υπολογίσουμε. Εξαρτώνται βέβαια κάθε φορά από την επιλογή του δείγματος. Όμως, μέσω του δείγματος, τον μέσο  $\mu$  του πληθυσμού και την τυπική του απόκλιση  $\sigma$  μπορούμε μόνο να τα εκτιμήσουμε και όχι να τα υπολογίσουμε.

### **Εφαρμογή 2η**

Η μέση τιμή των μηνιαίων μισθών των υπαλλήλων μιας εταιρείας είναι 1.200 ευρώ και η τυπική απόκλιση 100 ευρώ.

- 1) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 50 ευρώ, τότε ποια μεταβολή επέρχεται στη μέση τιμή, στην τυπική απόκλιση και στο συντελεστή μεταβλητότητας των μισθών;
- 2) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 5%, τότε ποια μεταβολή επέρχεται στη μέση τιμή, στην τυπική απόκλιση και στο συντελεστή μεταβλητότητας των μισθών;

### **Απάντηση**

Έστω ότι  $t_1, t_2, t_3, \dots, t_v$  είναι οι μηνιαίοι μισθοί των υπαλλήλων της εταιρείας. Τότε:

$$\checkmark \bar{x} = \frac{t_1 + t_2 + t_3 + \dots + t_v}{v} = 1.200$$

$$\checkmark s = \sqrt{\frac{(t_1 - 1.200)^2 + (t_2 - 1.200)^2 + \dots + (t_v - 1.200)^2}{v}} = 100$$

$$\checkmark CV = \frac{s}{|\bar{x}|} = \frac{100}{1.200} \approx \mathbf{8,33\%}$$



1) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 50 ευρώ, τότε:

$$\bar{y} = \frac{(t_1 - 50) + (t_2 - 50) + \dots + (t_v - 50)}{v} = \frac{t_1 + t_2 + \dots + t_v}{v} - \frac{50 \cdot v}{v} = 1\,200 - 50 = 1\,150$$

$$s_y = \sqrt{\frac{(t_1 - 50 - 1.150)^2 + \dots + (t_v - 50 - 1.150)^2}{v}} = \sqrt{\frac{(t_1 - 1.200)^2 + \dots + (t_v - 1.200)^2}{v}}$$

$$= s = 100$$

$$CV_y = \frac{s_y}{|\bar{y}|} = \frac{100}{1.150} \approx \mathbf{8,7\%}$$

Επομένως:

✓ Η μέση τιμή  $\bar{x} = 1.200\text{€}$  μειώνεται κατά 50€ και γίνεται  $\bar{y} = 1.150\text{€}$

✓ Η τυπική απόκλιση δεν μεταβάλλεται.

✓ Ο συντελεστής μεταβλητότητας  $CV = 8,33\%$  αυξάνεται και γίνεται  $CV_y = 8,7\%$

2) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 5%, τότε όλοι οι μηνιαίοι μισθοί πολλαπλασιάζονται με 0,95. Έτσι έχουμε:

$$\bar{z} = \frac{0,95 \cdot t_1 + 0,95 \cdot t_2 + \dots + 0,95 \cdot t_v}{v} = 0,95 \cdot \frac{t_1 + t_2 + \dots + t_v}{v} = 0,95 \cdot 1\,200 = 1\,140$$

$$s_z = \sqrt{\frac{(0,95 \cdot t_1 - 0,95 \cdot \bar{x})^2 + (0,95 \cdot t_2 - 0,95 \cdot \bar{x})^2 + \dots + (0,95 \cdot t_v - 0,95 \cdot \bar{x})^2}{v}}$$

$$= \sqrt{0,95^2 \cdot s^2} = |0,95| \cdot s = 0,95 \cdot 100 = 95$$

$$CV_z = \frac{s_z}{|\bar{z}|} = \frac{0,95 \cdot s}{0,95 \cdot \bar{x}} = CV = 8,33\%$$

Επομένως:

✓ Η μέση τιμή  $\bar{x} = 1.200\text{€}$  μειώνεται κατά 5% και γίνεται  $\bar{z} = 1.140\text{€}$

✓ Η τυπική απόκλιση  $s = 100\text{€}$  μειώνεται κατά 5% και γίνεται  $s_z = 95\text{€}$

✓ Ο συντελεστής μεταβλητότητας δεν μεταβάλλεται.

**Γενικότερα**

Έστω ότι έχουμε τις τιμές  $x_i$ ,  $i = 1, 2, \dots, n$  μιας ποσοτικής μεταβλητής  $X$  με μέση τιμή  $\bar{x}$  και τυπική απόκλιση  $s_x$ . Τότε για τη μέση τιμή  $\bar{y}$  και την τυπική απόκλιση  $s_y$  των τιμών της μεταβλητής  $Y$  για τις οποίες ισχύει  $y_i = ax_i + \beta$ ,  $i = 1, 2, \dots, n$  και  $a, \beta \in \mathbb{R}$  είναι:

$$\bar{y} = a\bar{x} + \beta \text{ και } s_y = |a| \cdot s_x$$

**Εφαρμογή 3η**

Στην εθνική οδό Αθηνών - Λαμίας, τα αυτοκίνητα τρέχουν με σταθερή ταχύτητα. Εμείς, με το δικό μας αυτοκίνητο, τρέχουμε με 100 Km/h και προσπερνάμε σε πλήθος όσα ακριβώς μας προσπερνούν. Να βρείτε τη διάμεσο των ταχυτήτων του δικού μας αυτοκινήτου και όλων όσων μας προσπέρασαν ή προσπεράσαμε στο ταξίδι μας αυτό.

Απάντηση

Η διάμεσος είναι  $\delta = 100$  Km/h διότι το πολύ 50% των ταχυτήτων είναι μικρότερες και το πολύ 50% είναι μεγαλύτερες από 100 Km/h.

**Εφαρμογή 4η**

Σε μια κάλπη υπάρχουν άσπρες, μαύρες, κόκκινες και πράσινες μπάλες σε αναλογία 10%, 20%, 30% και 40% αντίστοιχα. Μια άσπρη μπάλα έχει βάρος 10 gr, μια μαύρη 11 gr, μια κόκκινη 12 gr και μια πράσινη 13 gr. Να βρείτε τη μέση τιμή του βάρους τους, αν γνωρίζουμε ότι στην κάλπη υπάρχουν:

- 1) 10 μπάλες.
- 2) 20 μπάλες.
- 3) Δε γνωρίζουμε πόσες μπάλες υπάρχουν στην κάλπη.

Απάντηση

1) Με βάση την αναλογία της εκφώνησης, στις 10 μπάλες θα έχουμε 1 άσπρη, 2 μαύρες, 3 κόκκινες και 4 πράσινες. Αν οι τιμές των βαρών είναι οι παρατηρήσεις  $x_i$  με αντίστοιχες συχνότητες  $v_i$  και σχετικές συχνότητες  $f_i$ , τότε:

$$\bar{x} = \frac{1 \cdot 10 + 2 \cdot 11 + 3 \cdot 12 + 4 \cdot 13}{10}$$

$$= 0,10 \cdot 10 + 0,20 \cdot 11 + 0,30 \cdot 12 + 0,40 \cdot 13 = 12 \text{ gr}$$

$x_i$	$v_i$	$f_i$
10	1	0,10
11	2	0,20
12	3	0,30
13	4	0,40
Σύνολο	10	1,00

2) Με βάση την αναλογία της εκφώνησης, στις 20 μπάλες θα έχουμε 2 άσπρες, 4 μαύρες, 6 κόκκινες και 8 πράσινες. Αν οι τιμές των βαρών είναι οι παρατηρήσεις  $x_i$  με αντίστοιχες συχνότητες  $v_i$  και σχετικές συχνότητες  $f_i$ , τότε:

$$\bar{x} = \frac{2 \cdot 10 + 4 \cdot 11 + 6 \cdot 12 + 8 \cdot 13}{20}$$

$$= 0,10 \cdot 10 + 0,20 \cdot 11 + 0,30 \cdot 12 + 0,40 \cdot 13 = 12 \text{ gr}$$

$x_i$	$v_i$	$f_i$
10	2	0,10
11	4	0,20
12	6	0,30
13	8	0,40
Σύνολο	20	1,00

3) Παρατηρώντας προσεκτικά την παραπάνω διαδικασία καταλήγουμε, ότι ανεξαρτήτως του πλήθους από τις μπάλες έχουμε:

$$\bar{x} = f_1 \cdot x_1 + f_2 \cdot x_2 + f_3 \cdot x_3 + f_4 \cdot x_4 = 0,10 \cdot 10 + 0,20 \cdot 11 + 0,30 \cdot 12 + 0,40 \cdot 13 = 12 \text{ gr}$$

**Γενικότερα**

Έστω ότι έχουμε τις τιμές  $x_i$ ,  $i = 1, 2, \dots, v$  με σχετικές συχνότητες  $f_i$ ,  $i = 1, 2, \dots, v$ . Τότε για τη μέση τιμή  $\bar{x}$  είναι:

$$\bar{x} = f_1 \cdot x_1 + f_2 \cdot x_2 + f_3 \cdot x_3 + \dots + f_v \cdot x_v$$

## Ασκήσεις - Προβλήματα - Δραστηριότητες

- 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.
- 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.
- 3)** Καθεμία από τις παρακάτω λίστες δεδομένων έχουν μέση τιμή 50.
- (I) 0, 20, 40, 50, 60, 80, 100
- (II) 0, 48, 49, 50, 51, 52, 100
- (III) 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.
- 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
- 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
- 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.
- 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;

**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.

## Πρόσθετο υλικό

### Σύνθετες ασκήσεις

- 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.
- 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.
- 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.
- 4) Η τυπική απόκλιση ενός δείγματος με  $n$  παρατηρήσεις είναι ίση με μηδέν. Ποιο συμπέρασμα προκύπτει για τις παρατηρήσεις;
- 5) Η μέση τιμή 6 διαδοχικών ακέραιων αριθμών είναι 7,5. Να υπολογίσετε την τυπική τους απόκλιση.
- 6) Αν σε μία τάξη ο μέσος όρος της βαθμολογίας  $v_1$  αγοριών είναι  $\bar{x}$  και ο μέσος όρος της βαθμολογίας  $v_2$  κοριτσιών είναι  $\bar{y}$ , να αποδείξετε ότι ο μέσος όρος της βαθμολογίας όλων των παιδιών της τάξης είναι 
$$\bar{z} = \frac{v_1\bar{x} + v_2\bar{y}}{v_1 + v_2}.$$
- 7) Να αποδείξετε ότι η μέση τιμή είναι το άθροισμα των γινομένων των τιμών της μεταβλητής επί τις αντίστοιχες σχετικές συχνότητες.
- 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;







Βάσει του ν. 3966/2011 τα διδακτικά βιβλία του Δημοτικού, του Γυμνασίου, του Λυκείου, των ΕΠΑ.Λ. και των ΕΠΑ.Σ. τυπώνονται από το ΙΤΥΕ - ΔΙΟΦΑΝΤΟΣ και διανέμονται δωρεάν στα Δημόσια Σχολεία. Τα βιβλία μπορεί να διατίθενται προς πώληση, όταν φέρουν στη δεξιά κάτω γωνία του εμπροσθόφυλλου ένδειξη «ΔΙΑΤΙΘΕΤΑΙ ΜΕ ΤΙΜΗ ΠΩΛΗΣΗΣ». Κάθε αντίτυπο που διατίθεται προς πώληση και δεν φέρει την παραπάνω ένδειξη θεωρείται κλεψίτυπο και ο παραβάτης διώκεται σύμφωνα με τις διατάξεις του άρθρου 7 του νόμου 1129 της 15/21 Μαρτίου 1946 (ΦΕΚ 1946,108, Α').

*Απαγορεύεται η αναπαραγωγή οποιουδήποτε τμήματος αυτού του βιβλίου, που καλύπτεται από δικαιώματα (copyright), ή η χρήση του σε οποιαδήποτε μορφή, χωρίς τη γραπτή άδεια του Υπουργείου Παιδείας, Έρευνας και Θρησκευμάτων / ΙΤΥΕ - ΔΙΟΦΑΝΤΟΣ.*



Ινστιτούτο  
Τεχνολογίας  
Υπολογιστών & Εκδόσεων



(01) 000000 0 22 0258 7

Κωδικός Βιβλίου: 0-22-0258  
ISBN 978-960-06-5994-8