

# Εισαγωγή στη θεωρία Πιθανοτήτων και την Στατιστική

Μέτρα θέσης

# Μέτρα θέσης (ή Δείκτες κεντρικής τάσης)

- Προσδιορισμός της θέσης του «κέντρου» κατανομής των παρατηρήσεων, δηλ. τη θέση γύρω από την οποία είναι συγκεντρωμένες οι περισσότερες τιμές
- **Επικρατούσα τιμή**
- **Μέσος όρος**
- **Διάμεσος**
- **Ποσοστιαία σημεία (Τεταρτημόρια)**

# Επικρατούσα τιμή

- Η τιμή με τη μεγαλύτερη συχνότητα στο δείγμα λέγεται **επικρατούσα** ή **δεσπόζουσα τιμή** και συμβολίζεται με  $M_0$
- Η τιμή αυτή είναι η πιθανότερη μέσα στο δείγμα
- Σε αυτή την τιμή αντιστοιχεί η κορυφή του ιστογράμματος
- Είναι ο μόνος δείκτης κεντρικής τάσης που έχει νόημα για ποιοτικές, ονομαστικές μεταβλητές.

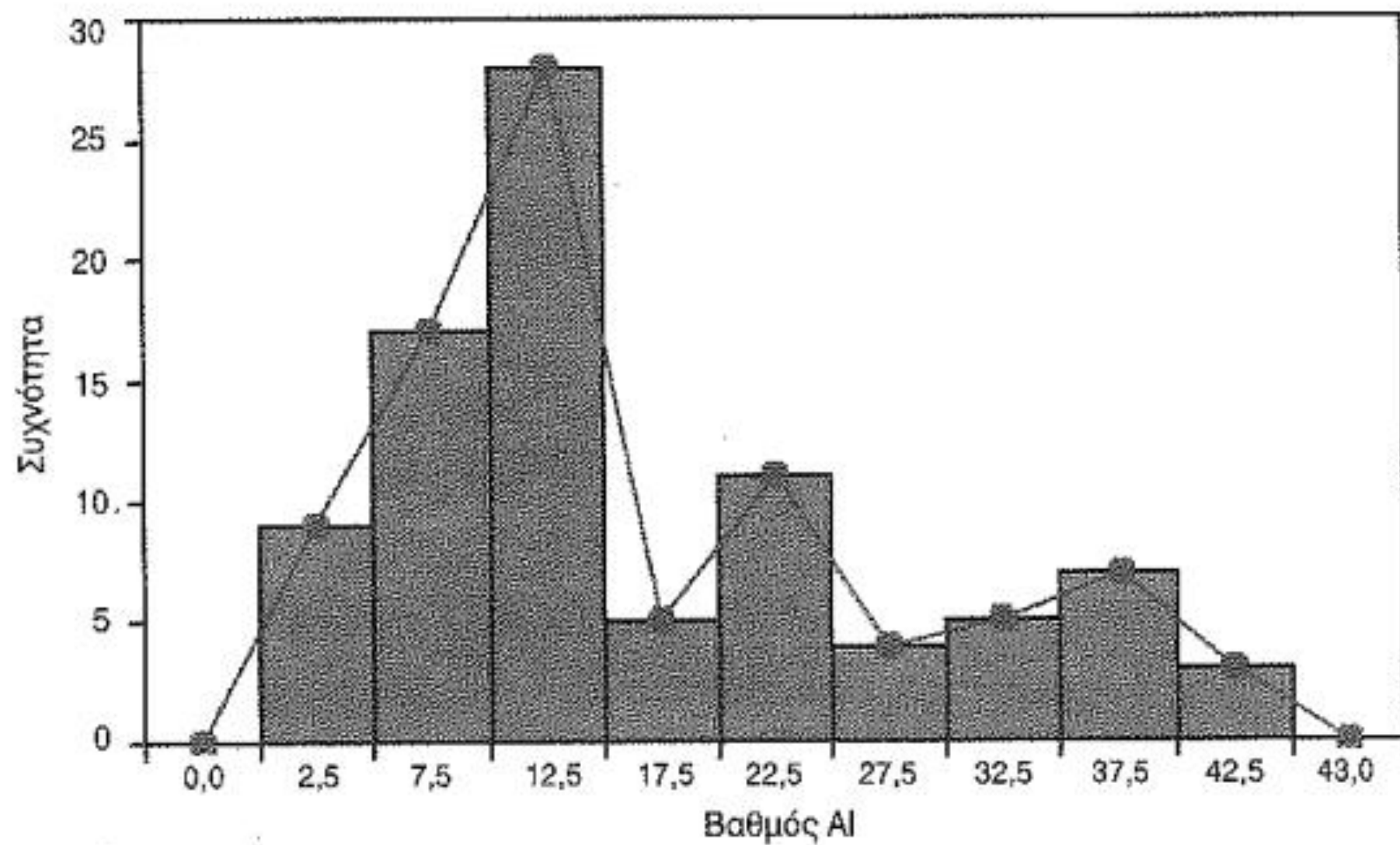
# Επικρατούσα τιμή

Πίνακας 2.2γ. Πίνακας κατανομής του βαθμού AI (βήμα 3).

| Βαθμός στο τεστ AI | Συχνότητα (f) |
|--------------------|---------------|
| 0                  | 1             |
| 1                  | 2             |
| 2                  | 1             |
| 3                  | 3             |
| 4                  | 2             |
| 5                  | 5             |
| 6                  | 4             |
| 7                  | 3             |
| 8                  | 2             |
| 9                  | 3             |
| 10                 | 6             |
| 11                 | 9             |
| 12                 | 3             |
| 13                 | 6             |
| 14                 | 4             |
| 15                 | 2             |
| 16                 | 0             |
| 17                 | 1             |
| 18                 | 1             |
| 19                 | 1             |
| 20                 | 3             |

| Βαθμός στο τεστ AI | Συχνότητα (f) |
|--------------------|---------------|
| 21                 | 3             |
| 22                 | 3             |
| 23                 | 2             |
| 24                 | 0             |
| 25                 | 0             |
| 26                 | 1             |
| 27                 | 1             |
| 28                 | 2             |
| 29                 | 0             |
| 30                 | 2             |
| 31                 | 2             |
| 32                 | 0             |
| 33                 | 1             |
| 34                 | 0             |
| 35                 | 0             |
| 36                 | 1             |
| 37                 | 2             |
| 38                 | 3             |
| 39                 | 1             |
| 40                 | 0             |
| 41                 | 2             |
| 42                 | 1             |

# Επικρατούσα τιμή

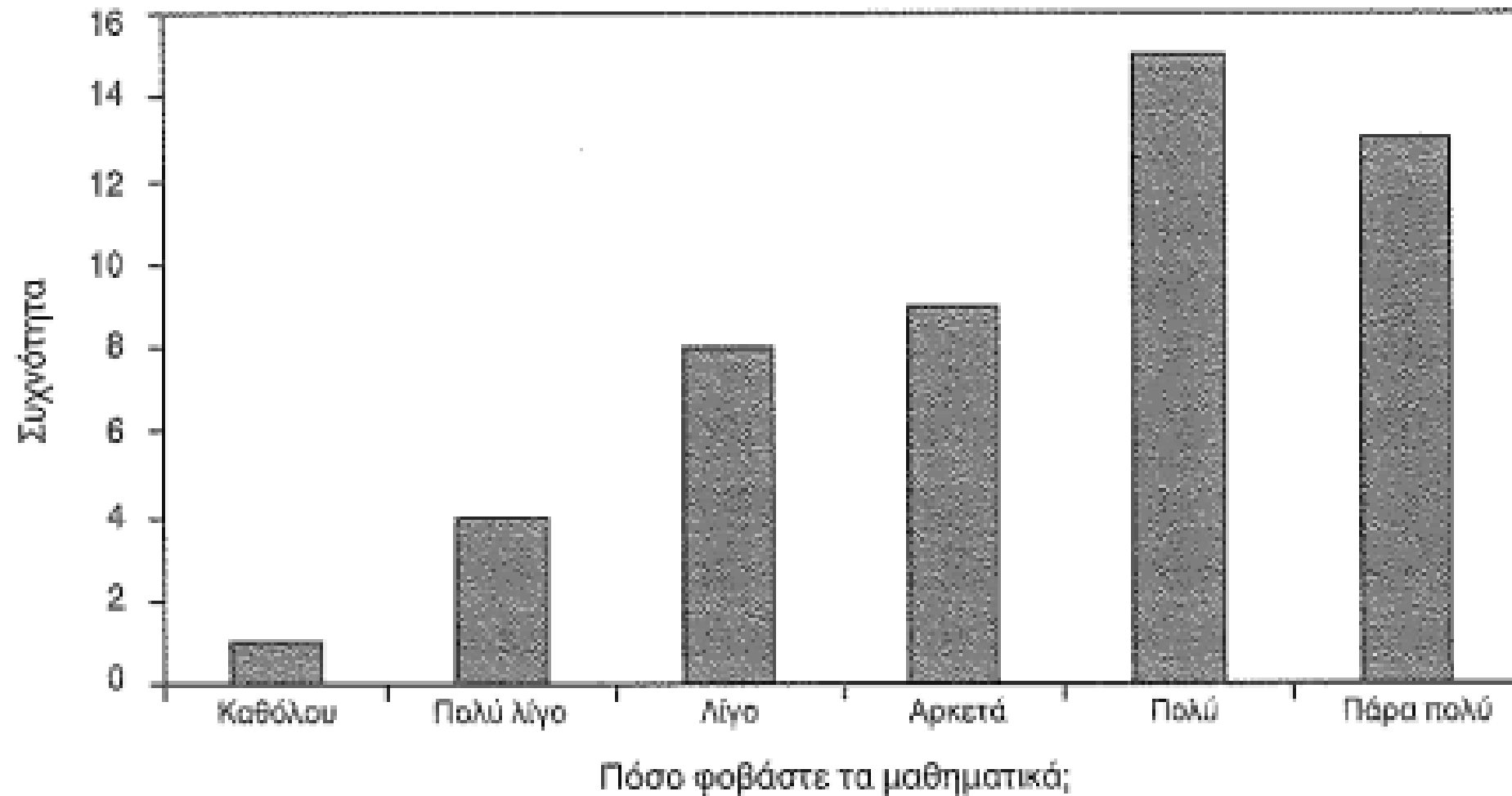


# Επικρατούσα τιμή

Πίνακας 2.1. Κατανομή συχνοτήτων τιμών της κλίμακας του φόβου.

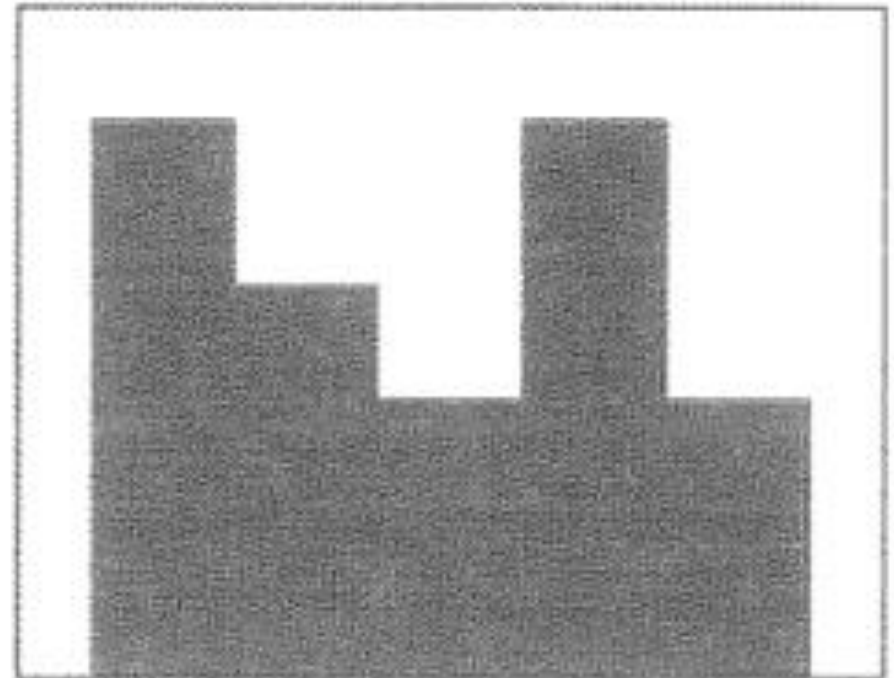
| Πόσο φοβόσασταν τα μαθηματικά στο λύκειο; | Αριθμός φοιτητών |
|-------------------------------------------|------------------|
| 0 (καθόλου)                               | 1                |
| 1 (πολύ λίγο)                             | 4                |
| 2 (λίγο)                                  | 8                |
| 3 (αρκετά)                                | 9                |
| 4 (πολύ)                                  | 15               |
| 5 (πάρα πολύ)                             | 13               |
|                                           | Σύνολο: 50       |

# Επικρατούσα τιμή



# Επικρατούσα τιμή

- Η επικρατούσα τιμή μπορεί να μην είναι μοναδική
- Κατανομές με δύο δεσπόζουσες τιμές λέγονται δικόρυφες



# Επικρατούσα τιμή

- Όταν όλες οι παρατηρήσεις είναι διαφορετικές, τότες λέμε ότι δεν υπάρχει επικρατούσα τιμή



# Εφαρμογή 1

Ο πίνακας δείχνει τον αριθμό των λαθών που έκαναν 60 μαθητές στο τεστ Ιστορίας. Αν ισχύει ότι  $k \leq \lambda$  και υπάρχουν 2 επικρατούσες τιμές, να βρείτε τα  $k$  και  $\lambda$ .

| Αριθμός λαθών | $n$       |
|---------------|-----------|
| 1             | $k$       |
| 2             | 12        |
| 3             | 14        |
| 4             | $\lambda$ |
| 5             | 8         |
| <b>Σύνολο</b> | <b>60</b> |

# Επικρατούσα τιμή

## **Πλεονεκτήματα**

*Υπολογίζεται εύκολα.*

*Είναι εύκολα κατανοητή.*

*Υπολογίζεται και από ελλιπή δεδομένα.*

*Δεν επηρεάζεται από ακραίες τιμές.*

*Υπολογίζεται και για ποιοτικά δεδομένα.*

## **Μειονεκτήματα**

*Δε χρησιμοποιούνται όλες οι τιμές για τον υπολογισμό της.*

*Στη στατιστική συμπερασματολογία έχει περιορισμένη σημασία.*

# Διάμεσος

- Η διάμεσος σε ένα δείγμα παρατηρήσεων, **διατεταγμένο από τη μικρότερη στη μεγαλύτερη τιμή**, είναι η τιμή εκείνη που χωρίζει το δείγμα σε δύο ίσα μέρη.
- Το ένα έχει τιμές μικρότερες ή ίσες με τη διάμεσο, και το άλλο μεγαλύτερες ή ίσες με αυτήν.
- Αν το πλήθος του δείγματος είναι περιττό τότε η διάμεσος είναι μια από τις παρατηρήσεις, ενώ αν είναι άρτιο η διάμεσος μπορεί να μην είναι μια από τις παρατηρήσεις.

# Διάμεσος

- Αν το **πλήθος των παρατηρήσεων είναι περιττός** αριθμός, τότε η διάμεσος είναι **η μεσαία παρατήρηση**
- Για να την βρούμε...

$$\delta = t_{\frac{(n+1)}{2}}$$

# Διάμεσος

- Αν το **πλήθος των παρατηρήσεων είναι άρτιος** αριθμός, τότε η διάμεσος είναι **το ημιάθροισμα των δύο μεσαίων παρατηρήσεων**
- Για να την βρούμε...

$$\delta = \frac{t_{\frac{n}{2}} + t_{(\frac{n}{2} + 1)}}{2}$$

## Εφαρμογή 2

Να βρείτε τη διάμεσο των θερμοκρασιών:

3   6   8   5   2   -1   10

# Εφαρμογή 3

Να βρείτε τη διάμεσο των παρατηρήσεων:

2   6   3   1   5   8   10   4   14   3

# Εφαρμογή 4α

Να βρείτε τη διάμεσο των δειγμάτων:

| $x$ | $v$ |
|-----|-----|
| 1   | 4   |
| 2   | 7   |
| 3   | 9   |
| 4   | 5   |

## Εφαρμογή 4β

Να βρείτε τη διάμεσο των δειγμάτων:

| $\chi$ | $\nu$ |
|--------|-------|
| 1      | 9     |
| 2      | 7     |
| 3      | 4     |
| 4      | 12    |
| 5      | 8     |

# Διάμεσος

## **Πλεονεκτήματα**

*Είναι εύκολα κατανοητή.*

*Δεν επηρεάζεται από ακραίες τιμές.*

*Ο υπολογισμός της είναι απλός.*

*Είναι μοναδική.*

## **Μειονεκτήματα**

*Δε χρησιμοποιούνται όλες οι τιμές για τον υπολογισμό της.*

*Δεν υπολογίζεται για ποιοτικά δεδομένα.*

# Μέσος όρος

- Ο μέσος όρος ενός συνόλου ορίζεται ως το άθροισμα των παρατηρήσεων δια το πλήθος  $n$  των παρατηρήσεων.

$$\bar{x} = \frac{t_1 + t_2 + \dots + t_n}{n} = \frac{\sum_{i=1}^n t_i}{n} :$$

- Ο μέσος όρος στηρίζεται για τον υπολογισμό της σε όλες τις μετρήσεις του δείγματος δίνοντας ίδια σημασία σε καθεμία από αυτές.
- Αντίθετα, η διάμεσος υπολογίζεται από μία ή το πολύ δύο μετρήσεις του δείγματος

# Εφαρμογή 5

Η βαθμολογία 8 μαθητών μιας τάξης στο τεστ Ιστορίας είναι:

6   8   10   13   9   15   14   5

A) Να υπολογίσετε τον μέσο όρο των βαθμών

B) Να βρείτε ποιος πρέπει να είναι ο βαθμός του καινούριου μαθητή που ήρθε στην τάξη για να γίνει ο Μ.Ο. ίσος με 11.

# Μέσος όρος

- Όταν μας δίνεται ο πίνακας κατανομής συχνοτήτων υπολογίζουμε τον μέσο όρο ως εξής:

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

$$\bar{X} = \frac{X_1 V_1 + X_2 V_2 + \dots + X_K V_K}{V}$$

| χ                | ν               |
|------------------|-----------------|
| x <sub>1</sub>   | v <sub>1</sub>  |
| x <sub>2</sub>   | v <sub>2</sub>  |
| x <sub>3</sub>   | v <sub>3</sub>  |
| x <sub>...</sub> | v <sub>..</sub> |
| x <sub>k</sub>   | v <sub>k</sub>  |

## Εφαρμογή 6

Ο αριθμός των επιβατών που υπήρχαν σε καθένα από τα 40 αυτοκίνητα που έλεγξε η τροχαία σε κάποιο σημείο του εθνικού οδικού δικτύου φαίνεται στον διπλανό πίνακα.

Να υπολογίσετε τον μέσο όρο του πλήθους των επιβατών που υπήρχαν στα παραπάνω αυτοκίνητα.

| Αριθμός επιβατών $x_i$ | Αριθμός αυτοκινήτων $n_i$ |
|------------------------|---------------------------|
| 1                      | 8                         |
| 2                      | 14                        |
| 3                      | 10                        |
| 4                      | 6                         |
| 5                      | 2                         |
| <b>Σύνολο</b>          | <b>40</b>                 |

## Εφαρμογή 6

Στον πίνακα κατανομής συχνοτήτων θα προσθέσουμε μια στήλη με τα γινόμενα  $x_i v_i$ .

| Αριθμός επιβατών $x_i$ | Αριθμός αυτοκινήτων $v_i$ | $x_i v_i$ |
|------------------------|---------------------------|-----------|
| 1                      | 8                         |           |
| 2                      | 14                        |           |
| 3                      | 10                        |           |
| 4                      | 6                         |           |
| 5                      | 2                         |           |
| <b>Σύνολο</b>          | <b>40</b>                 |           |

# Μέσος όρος

- Από αυτή την σχέση προκύπτει ότι μπορούμε να υπολογίσουμε τον μέσο όρο όταν γνωρίζουμε τις σχετικές συχνότητες

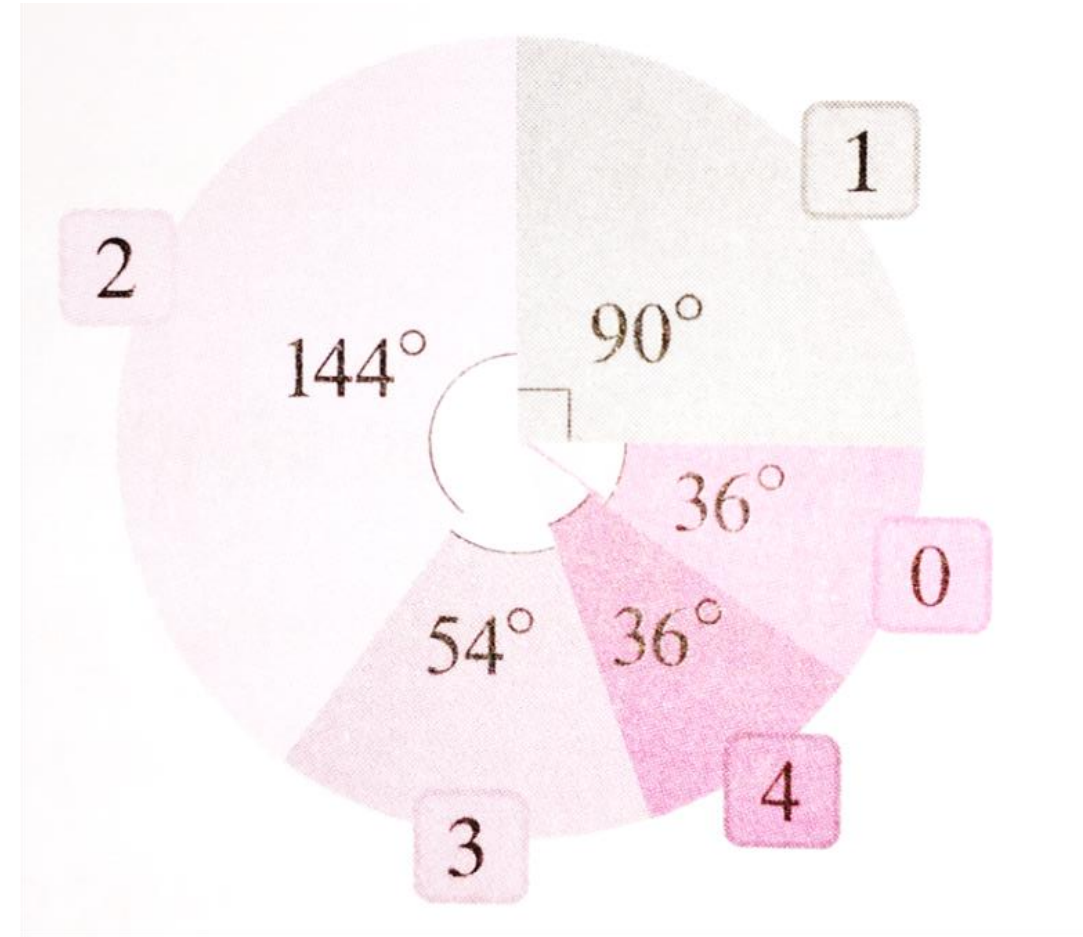
$$\bar{X} = \frac{X_1 V_1 + X_2 V_2 + \dots + X_K V_K}{V}$$

$$\bar{X} = \frac{X_1 V_1}{V} + \frac{X_2 V_2}{V} + \dots + \frac{X_K V_K}{V}$$

$$\bar{X} = x_1 f_1 + x_2 f_2 + \dots + x_K f_K$$

# Εφαρμογή 7

Το κυκλικό διάγραμμα δείχνει τον αριθμό των βιβλίων που δανείστηκαν από τη σχολική βιβλιοθήκη τον περασμένο μήνα τα παιδιά ενός σχολείου. Να φτιάξετε τον πίνακα κατανομής σχετικών συχνοτήτων και να υπολογίσετε τη μέσο όρο του πλήθους των βιβλίων που δανείστηκαν οι μαθητές.



# Μέσος όρος

## **Πλεονεκτήματα**

*Για τον υπολογισμό του χρησιμοποιούνται όλες οι τιμές.*

*Είναι μοναδικός.*

*Ο υπολογισμός του είναι απλός.*

*Χρησιμοποιείται στη στατιστική συμπερασματολογία.*

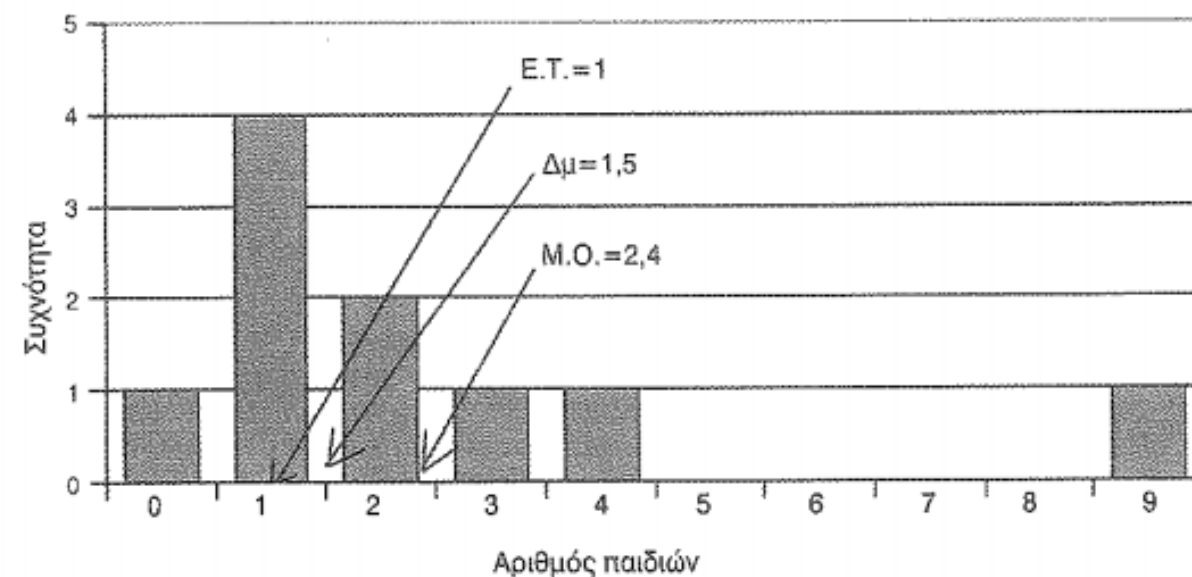
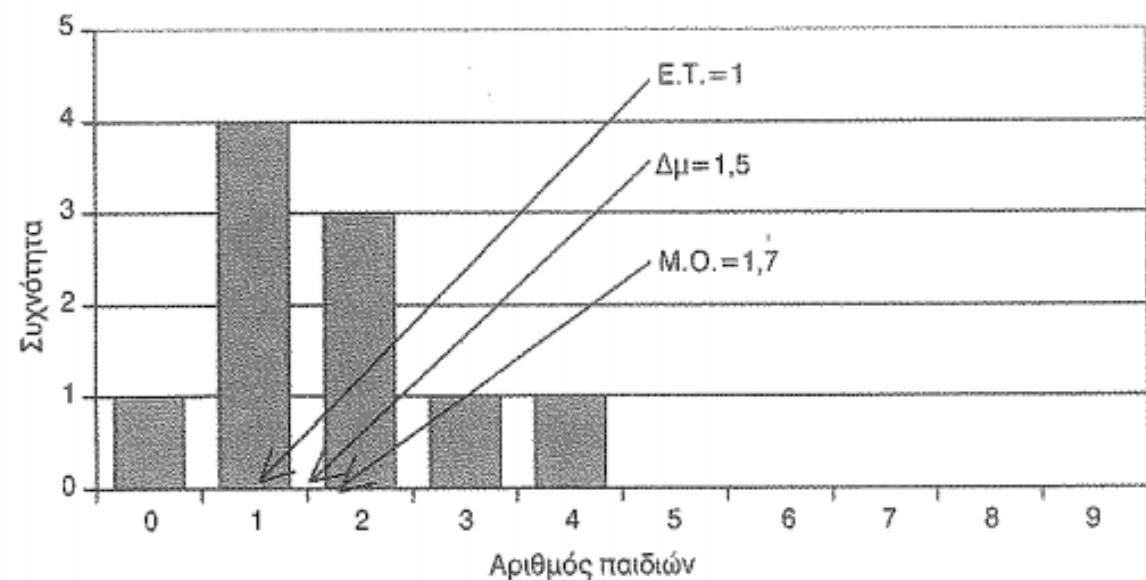
## **Μειονεκτήματα**

*Επηρεάζεται από ακραίες τιμές.*

*Μπορεί να μην αντιστοιχεί σε δυνατή τιμή της μεταβλητής.*

*Δεν υπολογίζεται για ποιοτικά δεδομένα.*

# Μέσος όρος



Σχήμα 3.1. Αλλαγή μίας και μόνο μέτρησης επηρέασε σημαντικά τον μέσο όρο, αλλά καθόλου την επικρατούσα τιμή και τη διάμεσο.

# Ποσοστιαία σημεία

Τα ποσοστιαία σημεία της κατανομής του δείγματος αποτελούν γενίκευση της έννοιας της διαμέσου και βοηθούν στην πληρέστερη περιγραφή της θέσης της κατανομής του δείγματος.

Το  $p$  -ποσοστιαίο σημείο είναι ένας αριθμός για τον οποίο ισχύει ότι το πολύ  $p\%$  των τιμών του δείγματος είναι μικρότερες από αυτόν και το πολύ  $(100 - p)\%$  των τιμών του δείγματος είναι μεγαλύτερες από αυτόν.

# Ποσοστιαία σημεία: Τεταρτημόρια

Υπάρχουν τρία τεταρτημόρια, τα οποία συμβολίζονται  $Q_1$ ,  $Q_2$  και  $Q_3$  και χωρίζουν σε 4 ίσα μέρη τις παρατηρήσεις. Συγκεκριμένα:

- Για το  $Q_1$  ισχύει ότι αριστερά του έχουμε το πολύ 25% των παρατηρήσεων και δεξιά του έχουμε το πολύ 75% των παρατηρήσεων.
- Για το  $Q_2$  ισχύει ότι αριστερά του έχουμε το πολύ 50% των παρατηρήσεων και δεξιά του έχουμε το πολύ 50% των παρατηρήσεων, άρα το  $Q_2$  συμπίπτει με τη διάμεσο.

Είναι δηλαδή  **$Q_2 = \delta$**

- Για το  $Q_3$  ισχύει ότι αριστερά του έχουμε το πολύ 75% των παρατηρήσεων και δεξιά του έχουμε το πολύ 25% των παρατηρήσεων.

# Ποσοστιαία σημεία: Τεταρτημόρια

## Τρόπος υπολογισμού

Όταν το **πλήθος** των διατεταγμένων παρατηρήσεων είναι **άρτιος αριθμός**, τότε θεωρούμε δύο σύνολα: το πρώτο μισό (με τις μικρότερες παρατηρήσεις) και το δεύτερο μισό (με τις μεγαλύτερες παρατηρήσεις). Το  $Q_1$  είναι η διάμεσος του πρώτου μισού και το  $Q_3$  η διάμεσος του δεύτερου μισού.

Όταν το **πλήθος** των διατεταγμένων παρατηρήσεων είναι **περιττός αριθμός**, τότε αφαιρούμε από το δείγμα τη διάμεσο (δηλαδή τη μεσαία παρατήρηση) και χωρίζουμε τις υπόλοιπες παρατηρήσεις σε δύο σύνολα: το πρώτο μισό (με τις μικρότερες παρατηρήσεις) και το δεύτερο μισό (με τις μεγαλύτερες παρατηρήσεις). Το  $Q_1$  είναι η διάμεσος του πρώτου μισού και το  $Q_3$  η διάμεσος του δεύτερου μισού.

## Εφαρμογή 8

**Να βρείτε τα τεταρτημόρια των παρακάτω παρατηρήσεων:**

**6, 13, 20, 1, 2, 11, 14, 4, 18**

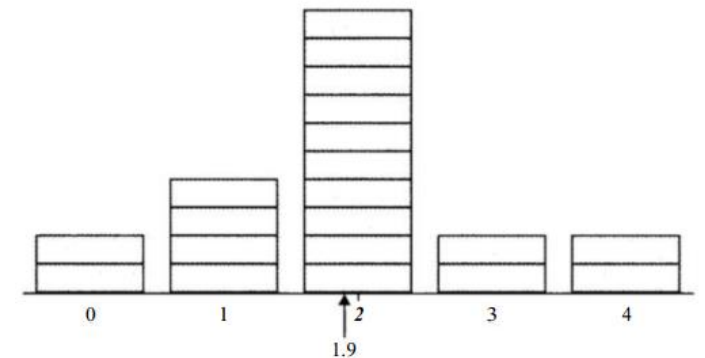
## Εφαρμογή 9

**Να βρείτε τα τεταρτημόρια των παρακάτω παρατηρήσεων:**

**12, 7, 10, 3, 1, 6, 8, 15, 9, 17**

# Σύγκριση μέτρων θέσης

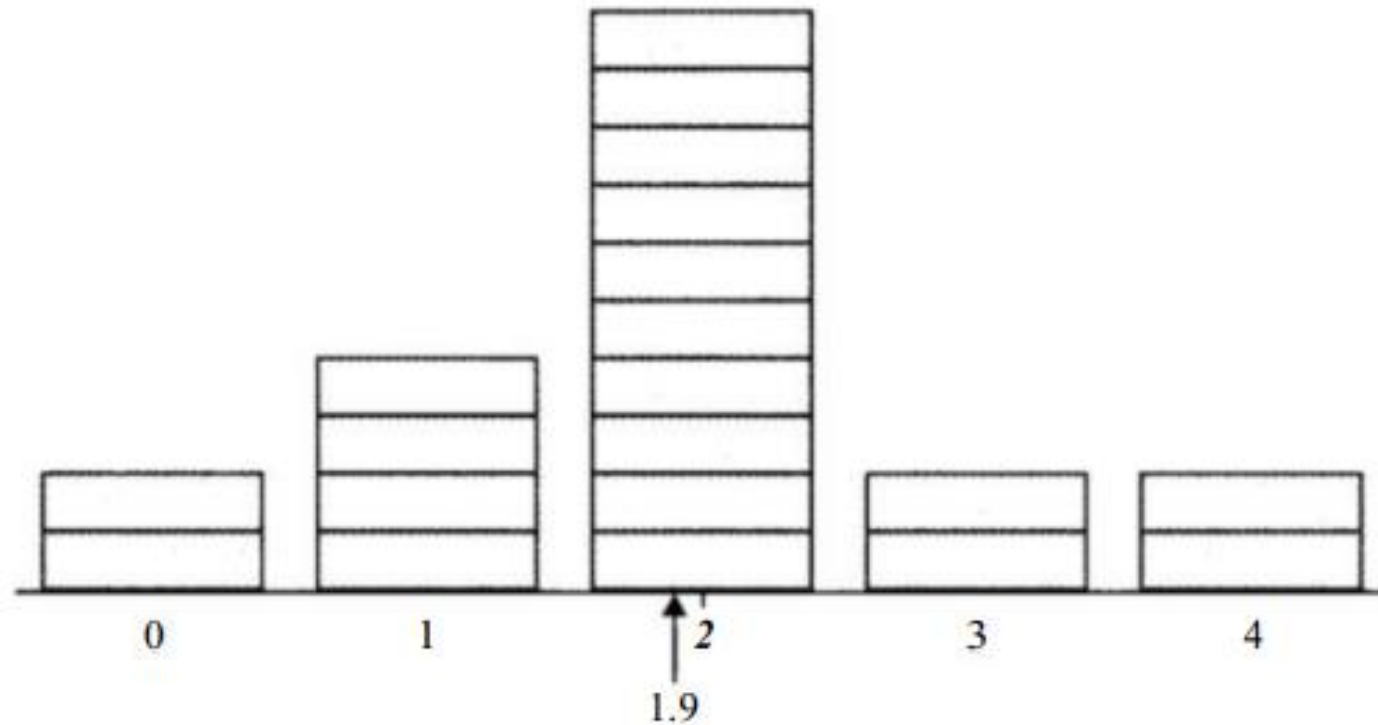
- Η **επικρατούσα τιμή** αντιστοιχεί στην **κορυφή** του ιστογράμματος
- Η **διάμεσος** αντιστοιχεί στη «**διχοτόμο**» του ιστογράμματος, αφού το 50% των μετρήσεων βρίσκονται αριστερά της και το υπόλοιπο 50% δεξιά της
- Ο **μέσος όρος** μπορεί να ερμηνευθεί και ως το **σημείο ισορροπίας** της κατανομής του δείγματος



Σχήμα 9.1.12 .  
Ερμηνεία του δειγματικού μέσου ως το σημείο ισορροπίας της κατανομής του δείγματος.

# Σύγκριση μέτρων θέσης

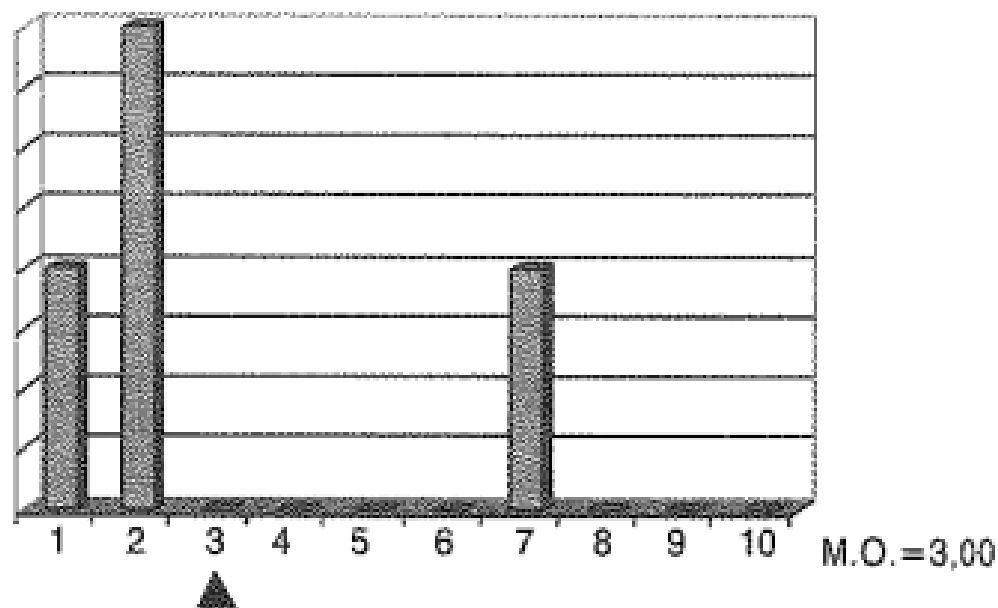
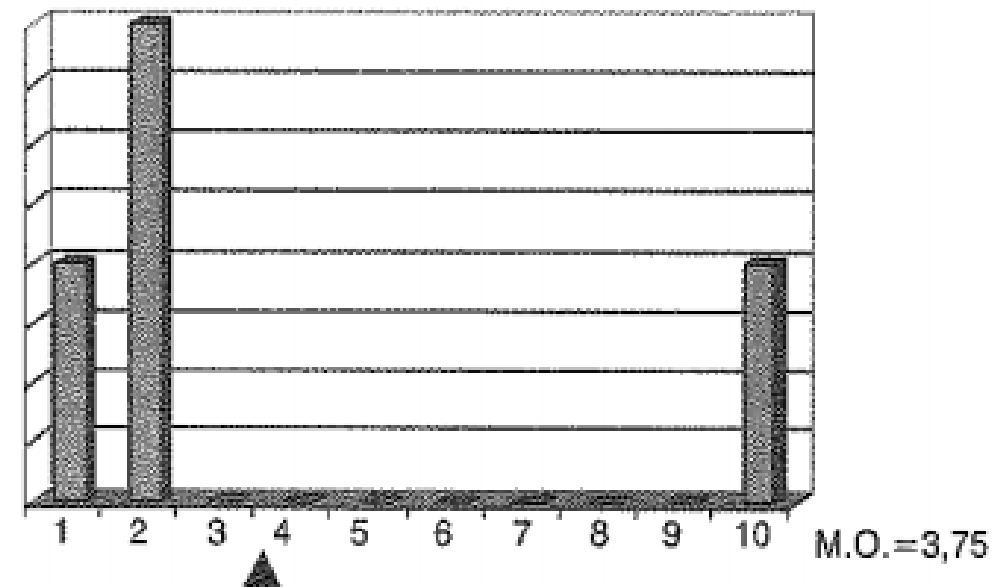
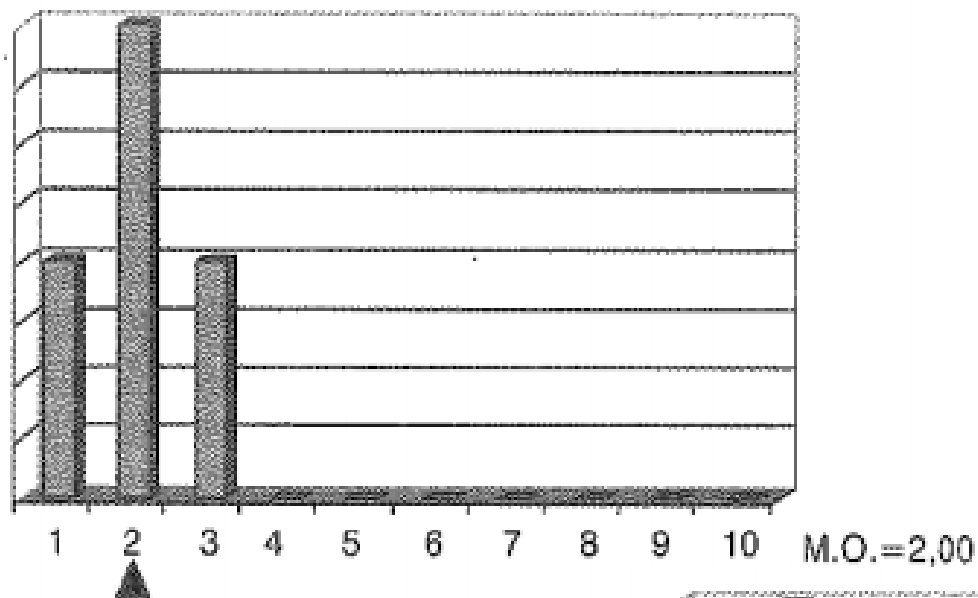
- 



Σχήμα 9.1.12 .

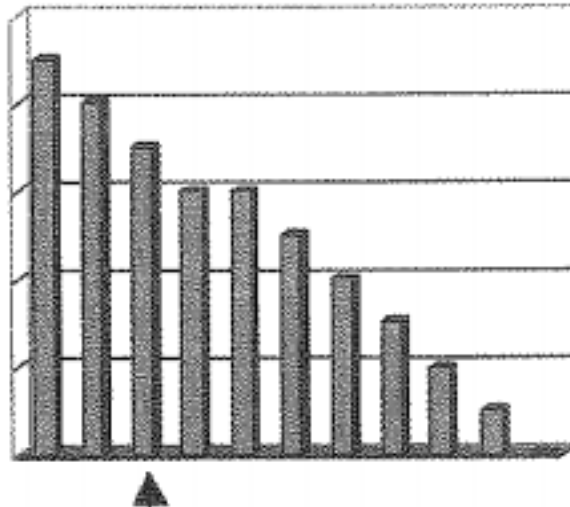
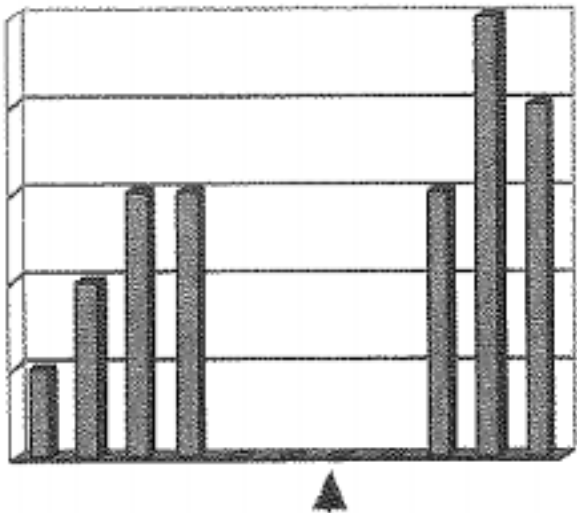
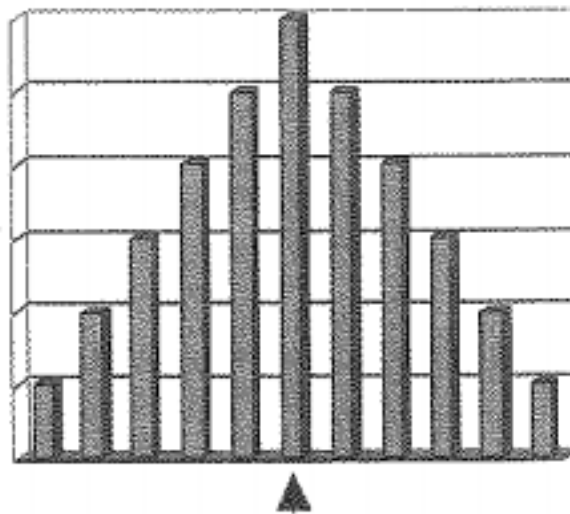
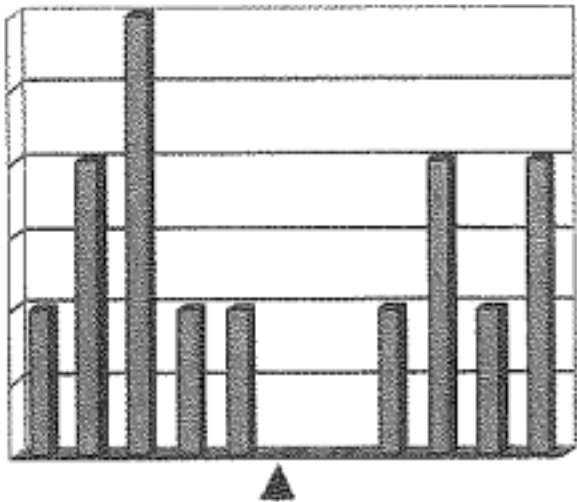
*Ερμηνεία του δειγματικού μέσον ως το σημείο ισορροπίας της κατανομής του δείγματος.*

# Σύγκριση μέτρων θέσης



Αύξηση της μεγαλύτερης παρατήρησης «συμπαράσύρει» τον μέσο όρο ενώ δεν επηρεάζει τη διάμεσο ούτε την επικρατούσα τιμή

# Σύγκριση μέτρων θέσης



Ο μέσος όρος είναι το σημείο στο οποίο η κατανομή ισορροπεί.  
Το σημείο ισορροπίας δεν εντοπίζεται κατ' ανάγκην σε κάποια τιμή του δείγματος ή στη μέση της κλίμακας

# Σύγκριση μέτρων θέσης

- Οι δείκτες που χρησιμοποιούνται περισσότερο είναι ο μέσος όρος και η διάμεσος.
- Όταν η κατανομή είναι συμμετρική, τότε ο μέσος όρος, η διάμεσος και η επικρατούσα τιμή συμπίπτουν. Αυτό συμβαίνει γιατί το σημείο ισορροπίας σε ένα συμμετρικό σχήμα οφείλει να είναι ακριβώς στη μέση.
- Αν η κατανομή είναι και μονοκόρυφη, τότε οι τρεις δείκτες κεντρικής τάσης συμπίπτουν, όπως συμβαίνει, για παράδειγμα, στην κανονική κατανομή

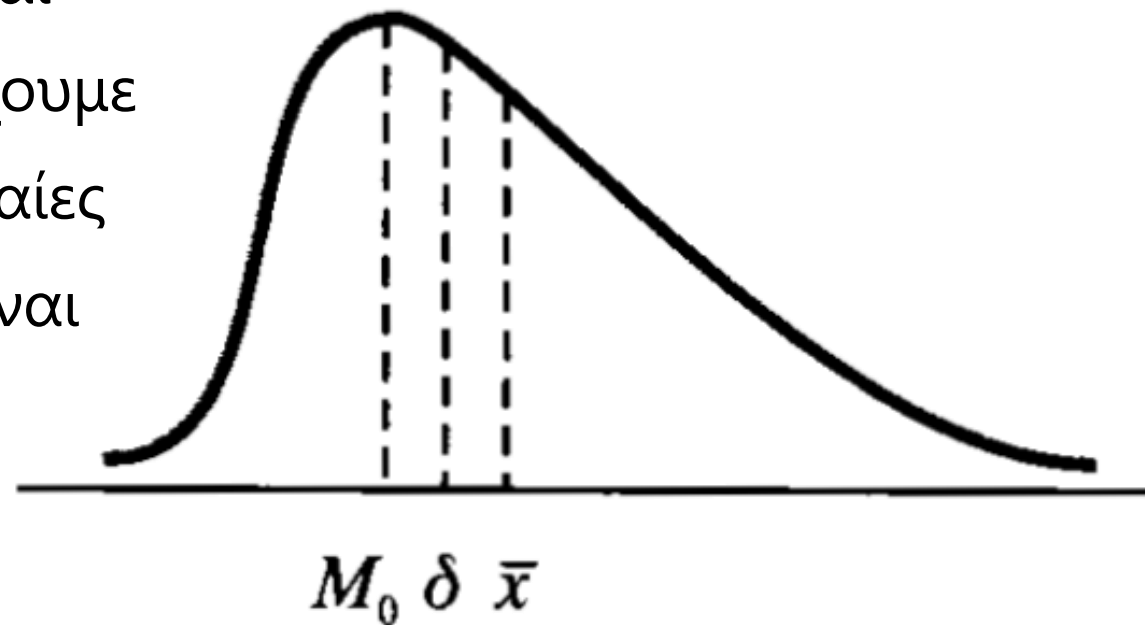


# Σύγκριση μέτρων θέσης

Αν η κατανομή είναι λοξή, ο μέσος όρος δεν ισούται με τη διάμεσο.

Σε μια δεξιό λοξή κατανομή ο μέσος όρος είναι μεγαλύτερος από τη διάμεσο, αφού, όπως έχουμε δει, ο μέσος όρος επηρεάζεται από λίγες ακραίες μετρήσεις, οι οποίες στην περίπτωση αυτή είναι προς τα δεξιά (στις μεγαλύτερες τιμές) της κλίμακας μέτρησης

$$\bar{x} > \delta > M_0.$$



# Παράδειγμα

Αν σε μια γειτονιά με 10 οικογένειες οι 9 έχουν ετήσια εισοδήματα μεταξύ 10.000€ και 50.000€ , και η δέκατη είναι οικογένεια εφοπλιστή με ετήσιο εισόδημα 30.000.000€, η διάμεσος του εισοδήματος θα είναι κάπου μεταξύ 10.000€ και 50.000€ , αλλά το μέσο εισόδημα θα είναι της τάξης των μερικών εκατομμυρίων ευρώ , το οποίο, στην περίπτωση αυτή, δεν αντανακλά προφανώς το «κέντρο»

# Παράδειγμα

Ο ιδιοκτήτης μιας μικρής επιχείρησης που απασχολεί πέντε εργαζομένους ισχυρίσθηκε σε δημοσιογράφο τοπικής εφημερίδας ότι οι εργαζόμενοι στην επιχείρησή του είναι πολύ καλά αμειβόμενοι αφού ο μέσος μισθός τους είναι 2000€.

Ο δημοσιογράφος ερεύνησε λεπτομερέστερα το θέμα και βρήκε ότι οι μισθοί των εργαζομένων ήταν 400€, 400€, 500€, 700€ και 8000€ αντίστοιχα! Ο μισθός των 8000€ ήταν του manager και συνιδιοκτήτη!

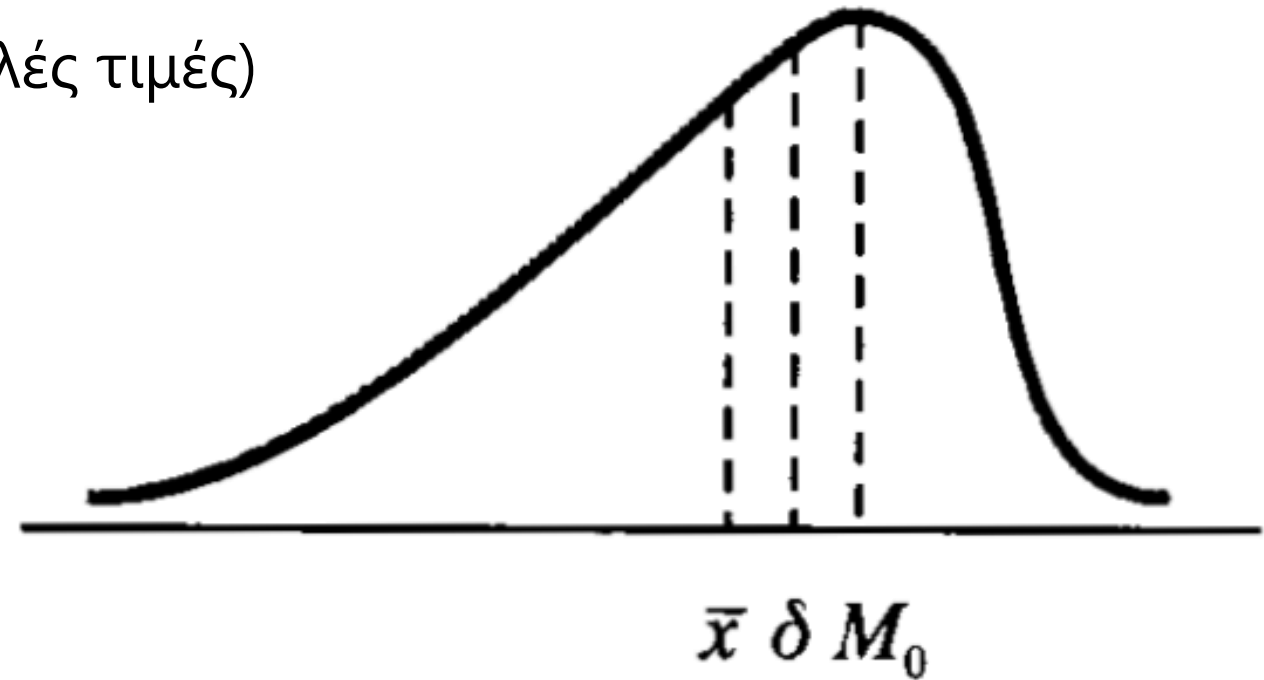
# Παράδειγμα

Έστω ότι η κατανομή των ( μηνιαίων) μισθών των εργαζομένων μιας επιχείρησης παρουσιάζει θετική ασυμμετρία με μέσο 2000€ και διάμεσο 1500€.  
Σε μια συνάντηση των εκπροσώπων των εργαζομένων με τον εργοδότη, ο εργοδότης αναφέρεται στον υψηλό μέσο μισθό (2000€).

Τι αντεπιχειρήματα, που να προκύπτουν από το είδος της ασυμμετρίας της κατανομής, έχουν οι εργαζόμενοι;

# Σύγκριση μέτρων θέσης

Σε μια αριστερά λοξή κατανομή, ο μέσος όρος είναι μικρότερος από τη διάμεσο, αφού στην περίπτωση αυτή οι λίγες ακραίες μετρήσεις βρίσκονται προς τα αριστερά (στις χαμηλές τιμές) της κλίμακας μέτρησης.



# Παράδειγμα

Αν μια φοιτήτρια συγκεντρώνει μέχρι τώρα βαθμούς από 8 έως 10 στα μαθήματα που έχει περάσει, ένα ενδεχόμενο 5 στο μάθημα της στατιστικής θα «ρίξει» τον μέσο όρο της βαθμολογίας της ενώ η διάμεσος θα παραμείνει υψηλά.

# Σύγκριση μέτρων θέσης

- Σε μια πολύ λοξή κατανομή, η διάμεσος πρέπει να προτιμάται από τον μέσο όρο, γιατί δίνει αντικειμενικότερη εικόνα για το «κέντρο» της κατανομής.
- Όσο λοξότερη είναι η κατανομή, τόσο περισσότερο διαφέρει ο μέσος όρος από τη διάμεσο, με αποτέλεσμα πολλές φορές οι δύο αυτοί δείκτες να οδηγούν σε διαφορετικές ερμηνείες.

# Εφαρμογή 10

Σε μια εταιρεία εργάζονται 9 άτομα. Η μέση τιμή των ηλικιών τους είναι 32 έτη.

α) Αν φύγει ένα άτομο 36 ετών, πόση θα είναι η μέση τιμή των ηλικιών των υπόλοιπων ατόμων;

β) Στα 9 άτομα που υπήρχαν αρχικά προστέθηκε ένα ακόμη άτομο και η μέση τιμή των ηλικιών όλων των ατόμων έγινε 33 έτη. Να βρείτε την ηλικία του ατόμου που προστέθηκε.

γ) Από τα 9 άτομα που υπήρχαν αρχικά, οι 4 είναι άντρες και οι 5 είναι γυναίκες. Αν η μέση τιμή των ηλικιών των ανδρών είναι 34,5 έτη, να υπολογίσετε τη μέση τιμή των ηλικιών των γυναικών.

# Εφαρμογή 11

Εξετάσαμε 20 οικογένειες ως προς τον αριθμό των παιδιών που έχουν και πήραμε τα αποτελέσματα του πίνακα:

- α) Να βρείτε τη διάμεσο.
- β) Να υπολογίσετε τη μέση τιμή.
- γ) Να βρείτε τα τεταρτημόρια  $Q_1$  ,  $Q_2$  και  $Q_3$

| Αριθμός παιδιών $x_i$ | Οικογένειες $v_i$ |
|-----------------------|-------------------|
| 0                     | 3                 |
| 1                     | 5                 |
| 2                     | 8                 |
| 3                     | 3                 |
| 4                     | 1                 |
| <b>Σύνολο</b>         | <b>20</b>         |

# Εφαρμογή 12

Δίνονται οι αριθμοί:

10, 11, 11, 13, 15, 17, 17, 17, 17, 18, 19, 19, 19, 19, 22, 25, 27, 27, 27, 30

α) Να βρείτε την επικρατούσα τιμή.

β) Να υπολογίσετε το μέσο όρο.

γ) Να βρείτε τα τεταρτημόρια  $Q_1$ ,  $Q_2$  και  $Q_3$

## Εφαρμογή 13

Σε μια εταιρεία εργάζονται 11 υπάλληλοι με μέσο μηνιαίο μισθό 1.300€ .

Από τους παραπάνω υπαλλήλους συνταξιοδοτήθηκαν 3 που είχαν μισθούς 1.600€, 1.500€ και 1.400€ , ενώ προσλήφθηκαν 2 υπάλληλοι με μισθούς 1.100€ και 1.200€.

Να υπολογίσετε τον μέσο μηνιαίο μισθό των υπαλλήλων που υπάρχουν τελικά στην εταιρεία.