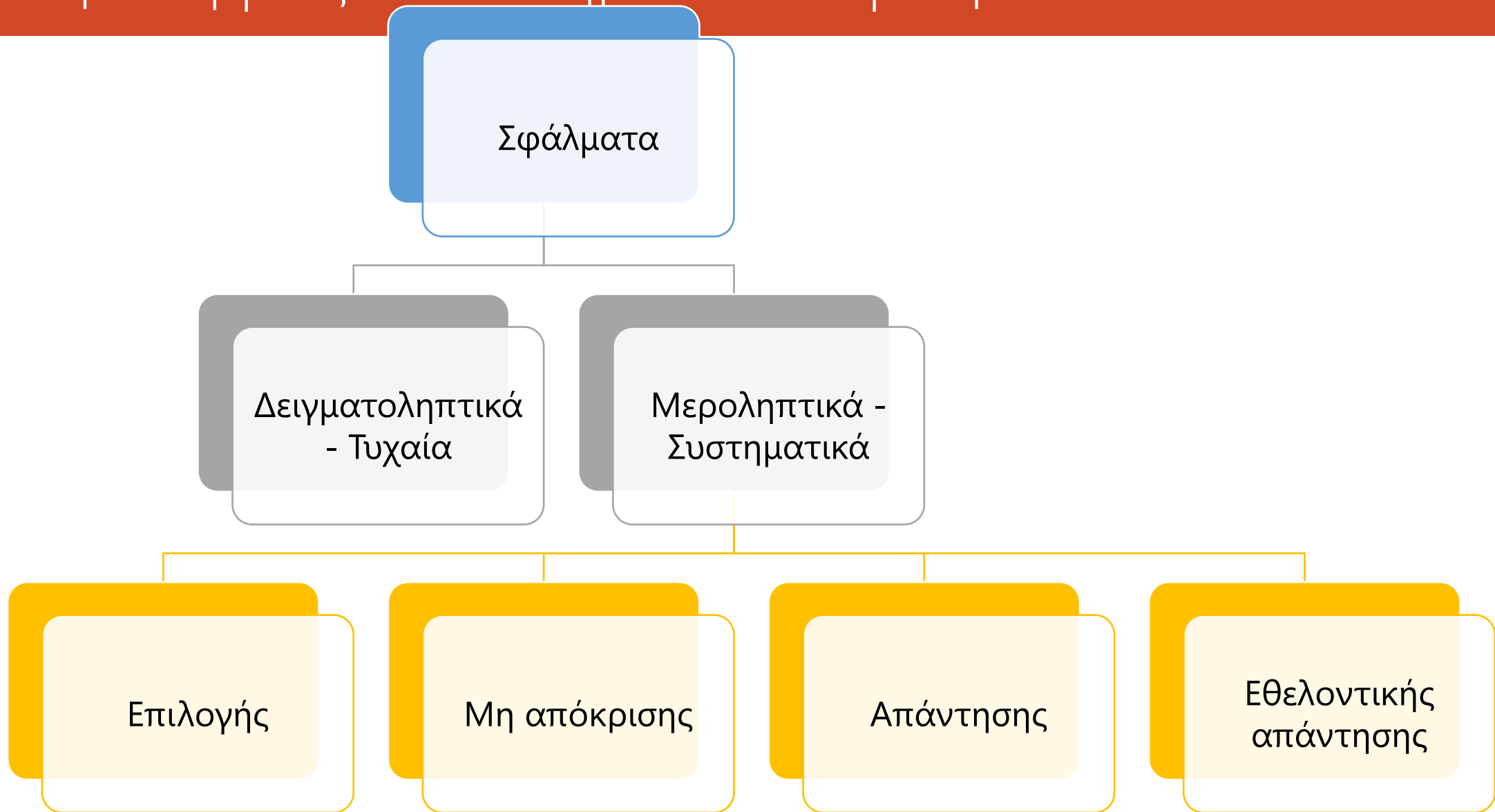


Εισαγωγή στη θεωρία Πιθανοτήτων και την Στατιστική

Επανάληψη

Μεροληψίες – Συστηματικά σφάλματα



Εφαρμογή 1α

Σε μια έρευνα για τις συνήθειες άσκησης, οι συμμετέχοντες καλούνται να απαντήσουν πόσες ώρες άσκησης πραγματοποιούν την εβδομάδα. Μερικοί από αυτούς, για να φανούν πιο υγιείς, αναφέρουν ότι ασκούνται περισσότερες ώρες από όσες πραγματικά κάνουν.

Ποιο είναι το σφάλμα που εμφανίζεται και πώς αυτό μπορεί να παραμορφώσει τα αποτελέσματα της έρευνας;

ΑΠΑΝΤΗΣΗΣ

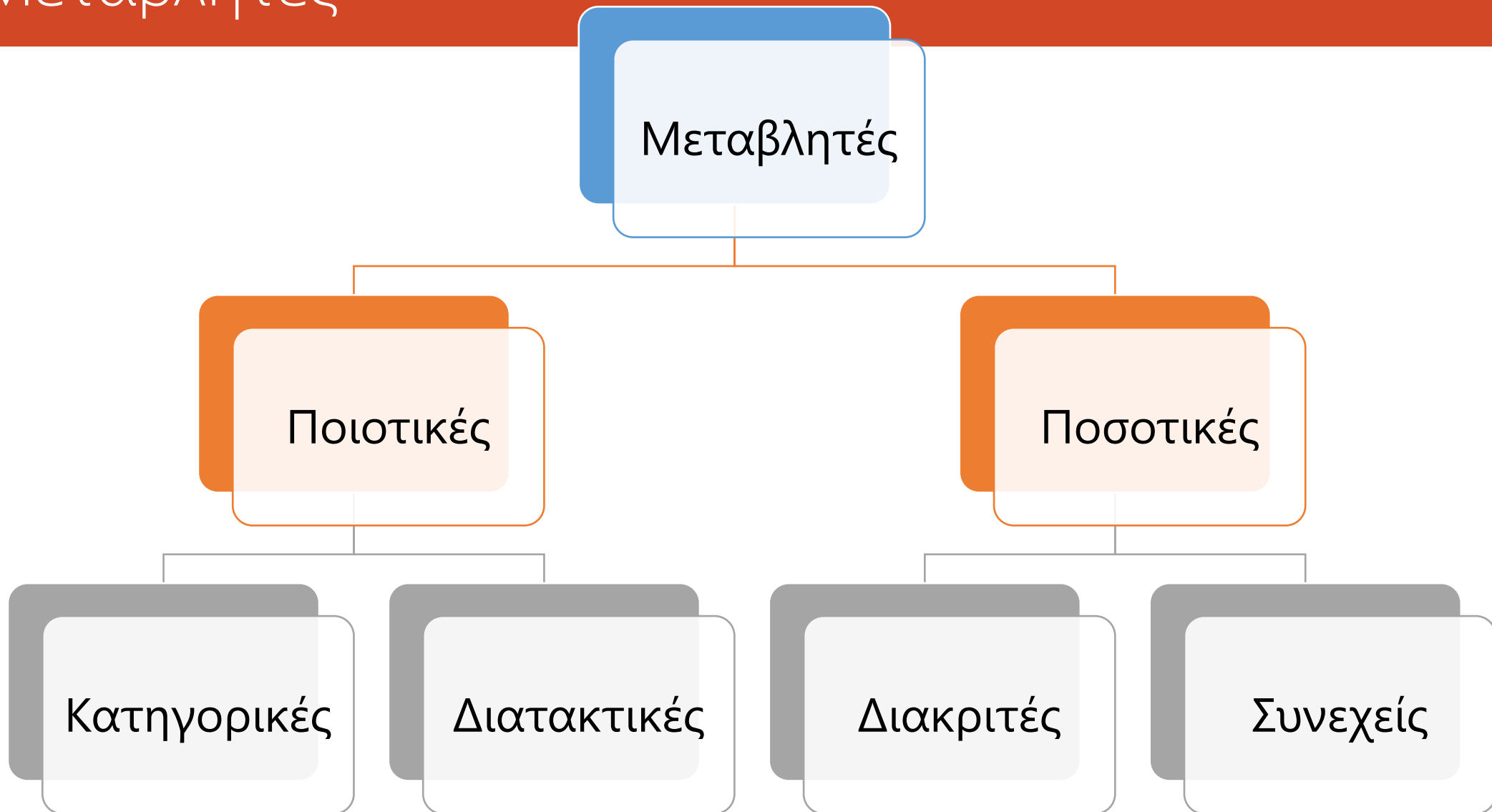
Εφαρμογή 1β

Ένα ειδησεογραφικό site διεξάγει μια δημοσκόπηση για να καθορίσει τη γνώμη των αναγνωστών του σχετικά με τη χρήση ηλεκτρικών αυτοκινήτων. Η δημοσκόπηση είναι ανοιχτή σε όλους τους επισκέπτες του site και οι συμμετέχοντες καλούνται να ψηφίσουν εθελοντικά.

Τι είδους σφάλμα μπορεί να υπάρχει σε αυτή την έρευνα και πώς θα μπορούσε να επηρεάσει τα αποτελέσματα;

**ΕΘΕΛΟΝΤΙΚΗΣ
ΣΥΜΜΕΤΟΧΗΣ**

Μεταβλητές



Εφαρμογή 2

Τι είδους είναι οι παρακάτω μεταβλητές;

1. Ο αριθμός του τηλεφώνου σας.
2. Ο αριθμός των τραγουδιών σε ένα μουσικό CD.
3. Ο αριθμός των ορόφων της οικίας σας.
4. Η μάρκα του αυτοκινήτου του πατέρα σας.
5. Ο αριθμός των φοιτητών που μετεγγράφηκαν στο ΠΤΔΕ πέρυσι.
6. Ο ταχυδρομικός κώδικας της περιοχής όπου μένετε.

Εφαρμογή 2

Τι είδους είναι οι παρακάτω μεταβλητές;

7. Η χρονική διάρκεια των τραγουδιών ενός CD.

8. Η βαθμίδα μέλους ΔΕΠ (λέκτορας, επίκουρος, αναπληρωτής, καθηγητής).

9. Ο αριθμός των απουσιών που είχατε στην Γ Λυκείου.

10. Το ενοίκιο που πληρώνετε στο διαμέρισμα όπου διαμένετε.

11. Ο στρατιωτικός βαθμός

Εφαρμογή 3

Για τις επόμενες δημοτικές εκλογές υπάρχουν 4 υποψήφιοι δήμαρχοι σε μια πόλη. Ένας κάτοικος της πόλης θέλησε να εκτιμήσει τα ποσοστά που θα συγκεντρώσει στις εκλογές ο κάθε υποψήφιος. Για τον λόγο αυτόν τηλεφώνησε σε 50 φίλους του και συγγενείς του και τους ρώτησε ποιον από τους 4 υποψηφίους θα ψηφίσουν.

α) Να βρείτε:

i) ποιος είναι ο πληθυσμός της έρευνας,

ii) ποιο είναι το δείγμα,

iii) ποια είναι η μεταβλητή και ποιο είναι το είδος της μεταβλητής της έρευνας.

β) Να εξετάσετε αν το δείγμα είναι αντιπροσωπευτικό του πληθυσμού.

Εφαρμογή 4

Ο βαθμός του πτυχίου δηλώνεται με δύο τρόπους. Ο ένας είναι ο μέσος όρος των βαθμών όλων των μαθημάτων, στα οποία ένας φοιτητής έχει εξεταστεί επιτυχώς. Για ποιου είδους μεταβλητή πρόκειται;

Ο άλλος τρόπος έγκειται στην απόδοση των χαρακτηρισμών καλώς, λίαν καλώς και άριστα. Τι είδους μεταβλητή έχουμε τώρα;

Κλίμακες Μέτρησης

Αυξανόμενη ακρίβεια μέτρησης

1. Κλίμακες κατηγορίας ή Ονομαστικές
2. Κλίμακες διάταξης ή Τακτικές
3. Κλίμακες διαστήματος ή ισοδιαστημικές
4. Κλίμακες αναλογίας

Ποιοτικές μεταβλητές

Ποσοτικές μεταβλητές

Εφαρμογή 5

Σε τι κλίμακα μπορούν να αποδοθούν οι μεταβλητές:

- Αριθμός μητρώου
- Δαπάνη
- Νομός
- Χρόνος
- Επίπεδο εκπαίδευσης
- Βαθμολογία μαθήματος
- Ηλικία
- Χρώμα ματιών

Εφαρμογή 6

Ρωτήσαμε 50 μαθητές ενός Λυκείου πόσα παιδιά υπάρχουν στην οικογένειά τους και πήραμε τις απαντήσεις του πίνακα.

Να κατασκευάσετε τον πίνακα κατανομής σχετικών συχνοτήτων και σχετικών εκατοστιαίων συχνοτήτων.

Αριθμός παιδιών	Συχνότητα n	Σχετική συχν. f	Σχετική συχν. $f\%$
1	13		
2	30		
3	5		
4	2		
Σύνολο	50		

Εφαρμογή 7α

Τα ύψη (σε cm) 40 μαθητών της Γ' Λυκείου ενός σχολείου είναι τα εξής:

168	160	164	154	169
158	162	176	165	161
163	155	164	166	172
174	164	156	159	165
166	161	167	163	167
165	172	157	175	160
168	160	164	159	167
156	163	177	155	164

Να σχηματίσετε κλάσεις της μορφής $[α, β)$ πλάτους 4 cm

Να κατασκευάσετε πίνακα κατανομής συχνοτήτων

Εφαρμογή 7β

Για τα ύψη των μαθητών της Εφαρμογής 2 να κατασκευάσετε πίνακα κατανομής σχετικών συχνοτήτων και σχετικών εκατοστιαίων συχνοτήτων και ιστόγραμμα συχνοτήτων. Να χαρακτηρίσετε την κατανομή ως προς τη μορφή της

Ύψη	Συχνότητα n	Σχετική συχν. f	Σχετική συχν. $f\%$
[154, 158)	6		
[158, 162)	8		
[162, 166)	12		
[166, 170)	8		
[170, 174)	2		
[174, 178)	4		

Εφαρμογή 8α

Δίνεται ο παρακάτω πίνακας που αναφέρεται σε δεδομένα που έχουν ομαδοποιηθεί σε κλάσεις ίσου πλάτους.

A) Να βρείτε ποιες είναι οι κλάσεις.

Κλάσεις	Συχνότητα n	Σχετική συχν. f
[.....,)	4	
[8,)		0,2
[.....,)		
[....., 20)		
Σύνολο		

Εφαρμογή 8β

Αν γνωρίζουμε ότι:

- 12 παρατηρήσεις είναι μικρότερες του 12 και ότι
- το 70% των παρατηρήσεων είναι μικρότερες του 16.

Να συμπληρώσετε
τον πίνακα

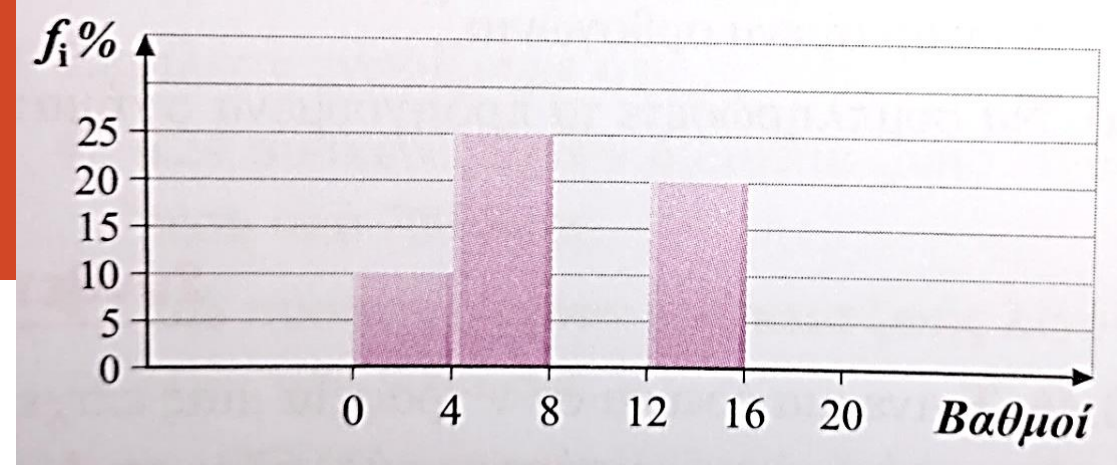
Κλάσεις	Συχνότητα n	Σχετική συχν. f
$[4, 8)$	4	
$[8, 12)$		0,2
$[12, 16)$		
$[16, 20)$		
Σύνολο		

Εφαρμογή 9

- Οι μέγιστες θερμοκρασίες 60 ημερών που καταγράφηκαν σε έναν μετεωρολογικό σταθμό φαίνονται στον πίνακα. Να κατασκευαστεί:
 1. Το ιστόγραμμα και το πολύγωνο συχνοτήτων
 2. Το ιστόγραμμα και το πολύγωνο σχετικών εκατοστιαίων συχνοτήτων

Μέγιστη θερμοκρασία (C)	Αριθμός ημερών
[0 , 2)	6
[2 , 4)	15
[4 , 6)	21
[6 , 8)	12
[8 , 10)	6
Σύνολο	60

Εφαρμογή 10



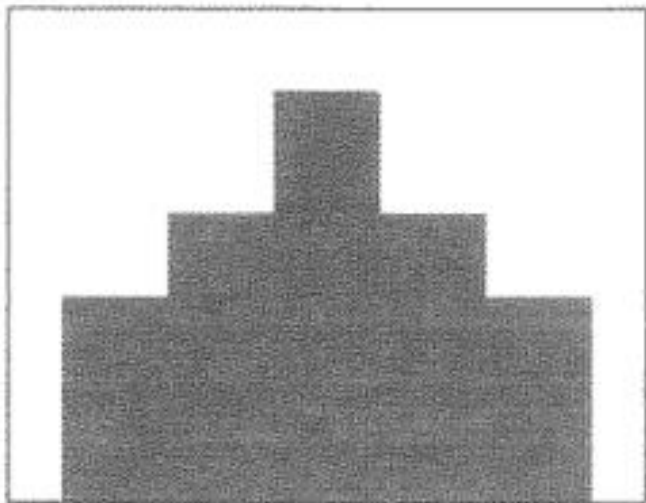
- Οι βαθμοί των μαθητών σε ένα διαγώνισμα φαίνονται στο ιστόγραμμα σχετικών εκατοστιαίων συχνοτήτων. Η σχετική εκατοστιαία συχνότητα της 3ης κλάσης είναι διπλάσια της 5ης κλάσης.
- Να συμπληρωθεί το ιστόγραμμα & να δημιουργηθεί το πολύγωνο σχετικών συχνοτήτων
- Να βρεθεί το ποσοστό των μαθητών που έγραψαν < 8
- Να βρεθεί το ποσοστό των μαθητών που έγραψαν ≥ 16

Μορφολογία κατανομών

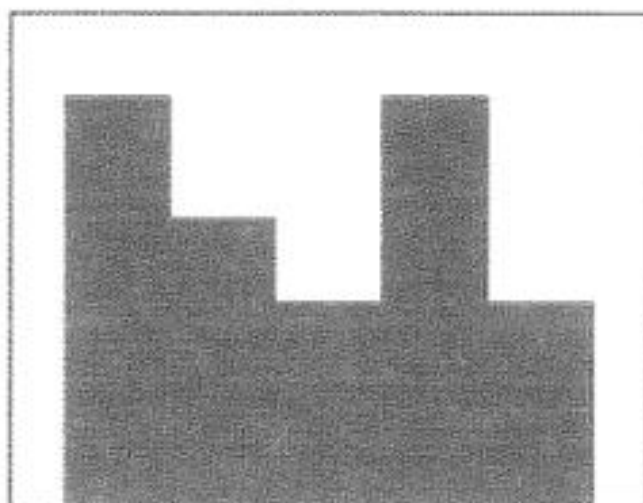
- Το ραβδόγραμμα, το ιστόγραμμα ή το πολύγωνο συχνοτήτων μπορούν να χρησιμοποιηθούν για να δείξουν την **κατανομή της συχνότητας**, δηλαδή πώς οι συχνότητες καταμερίζονται στις τιμές μιας μεταβλητής



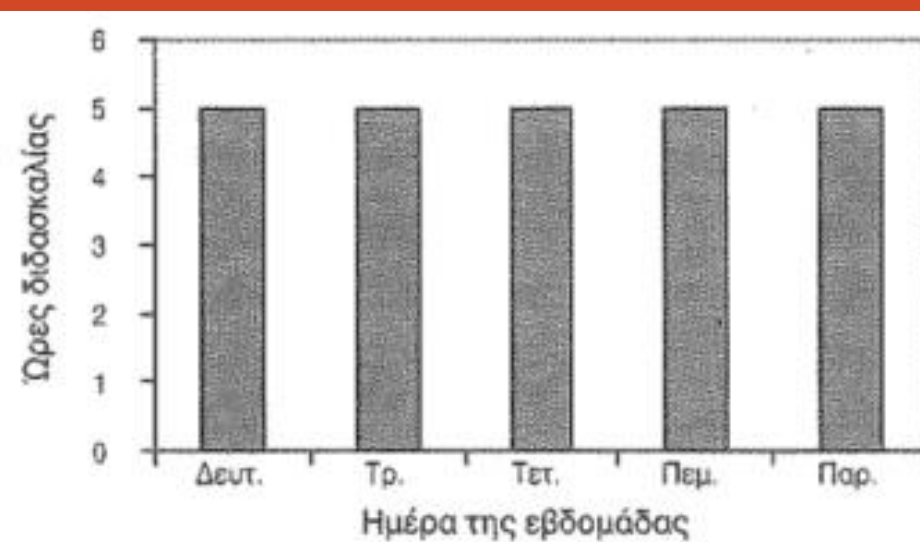
Κορυφές



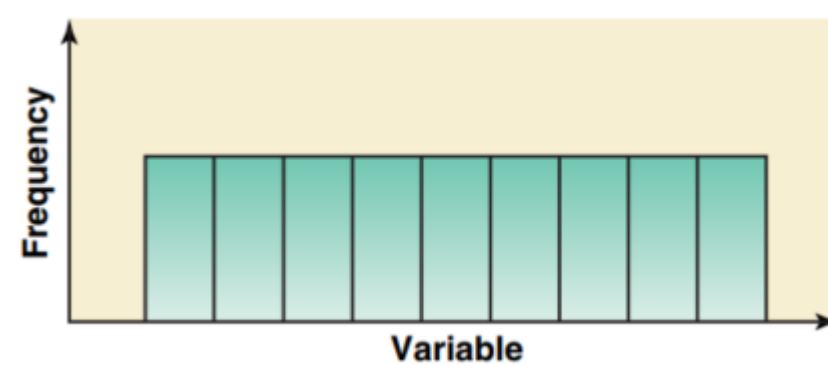
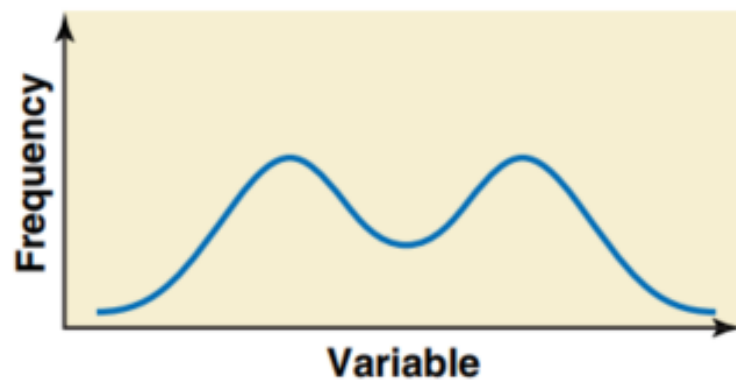
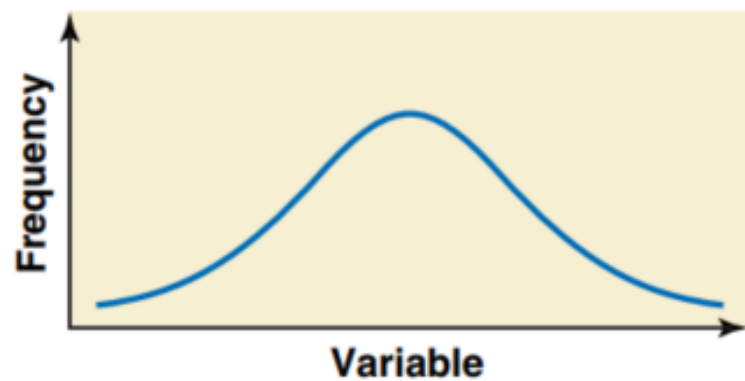
Μονοκόρυφη



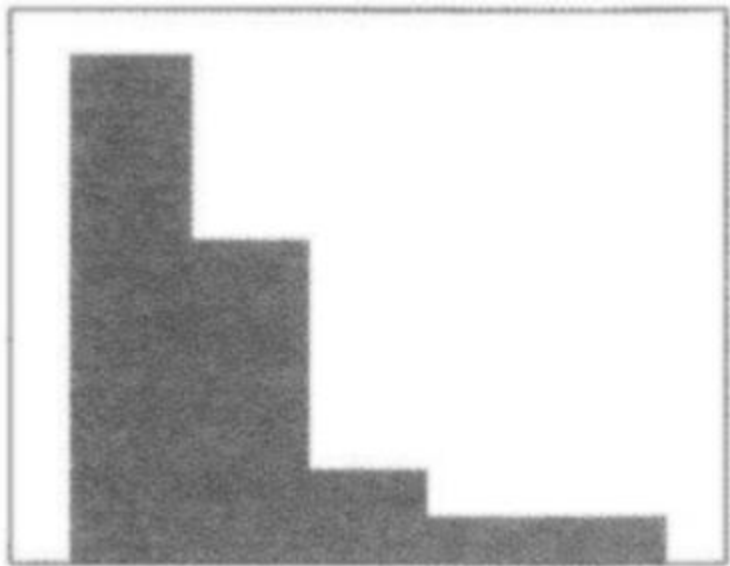
Δικόρυφη



Ισοϋψής

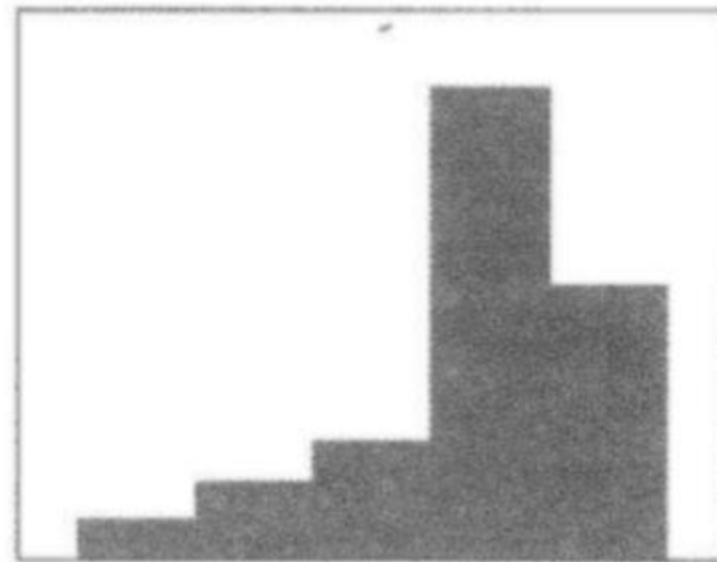


Συμμετρία



Λοξή δεξιά

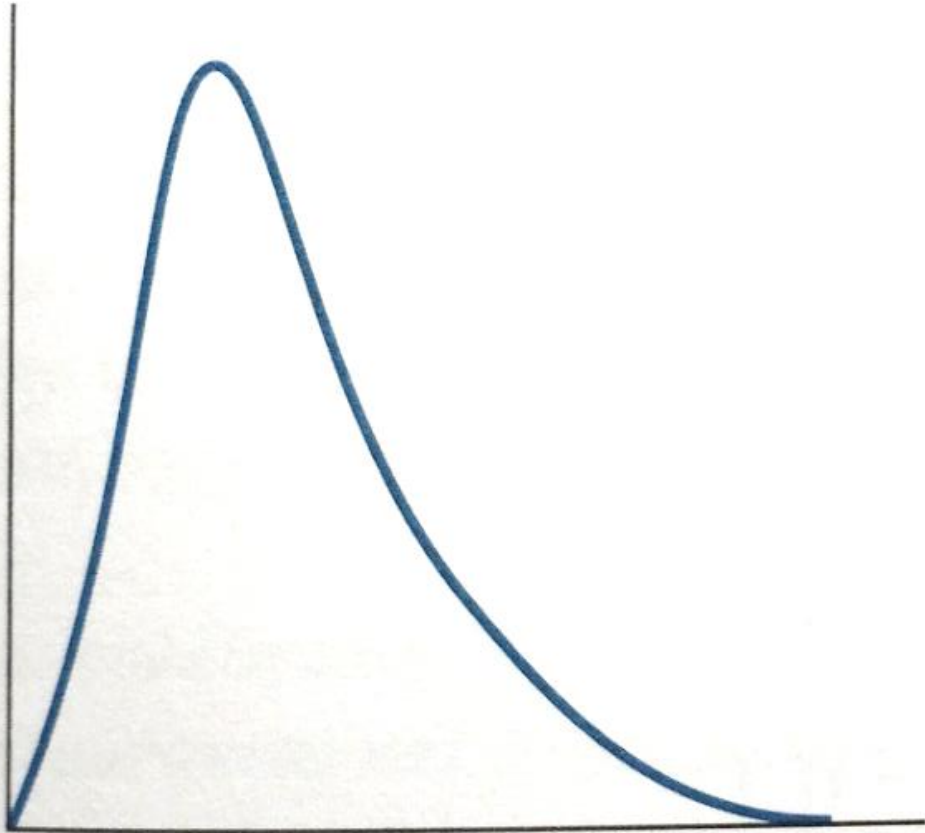
Θετικά λοξή



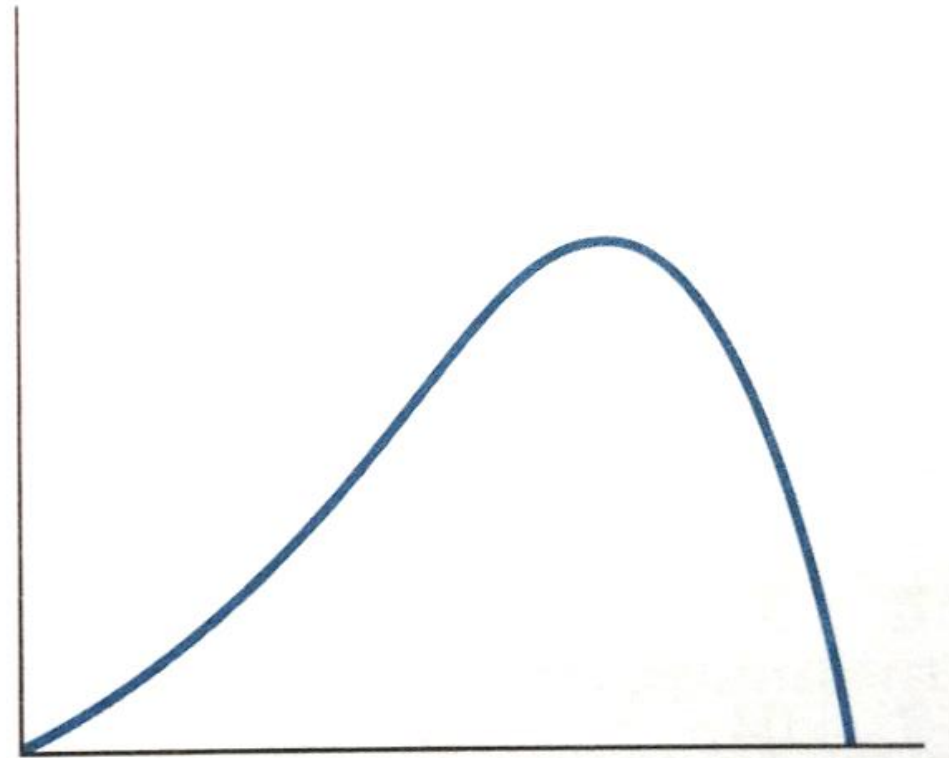
Λοξή αριστερά

Αρνητικά λοξή

Συμμετρία

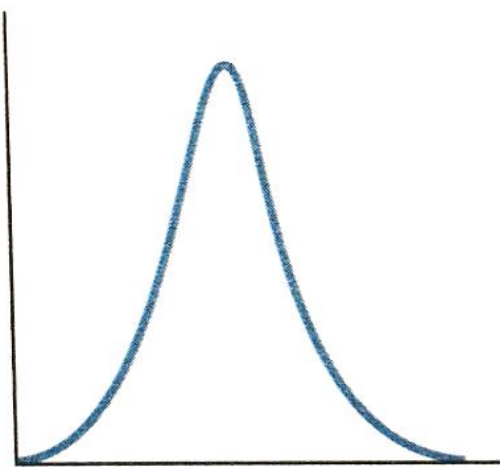
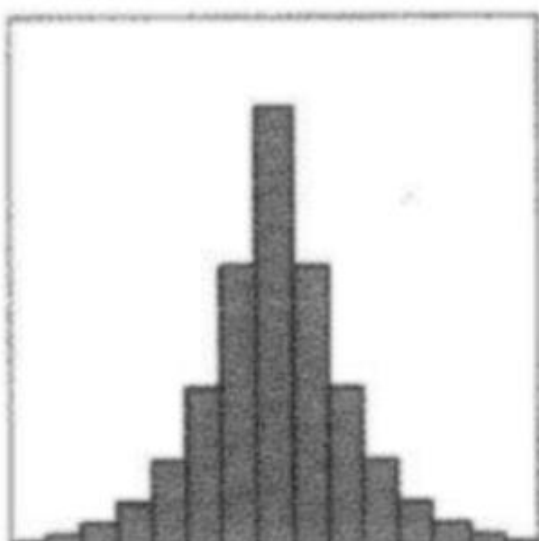


α. Θετική ασυμμετρία.

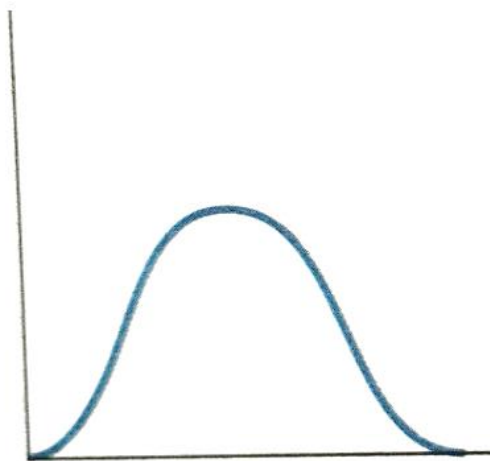
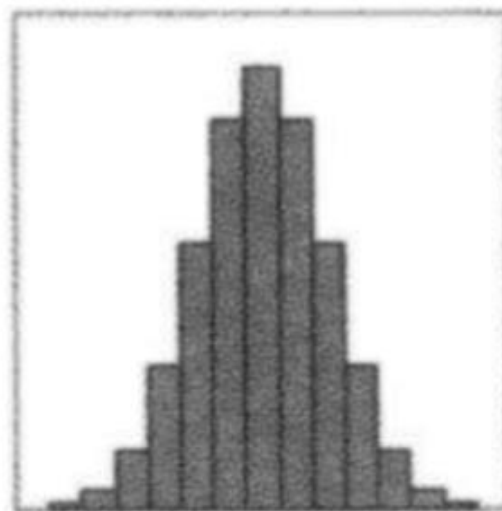


β. Αρνητική ασυμμετρία.

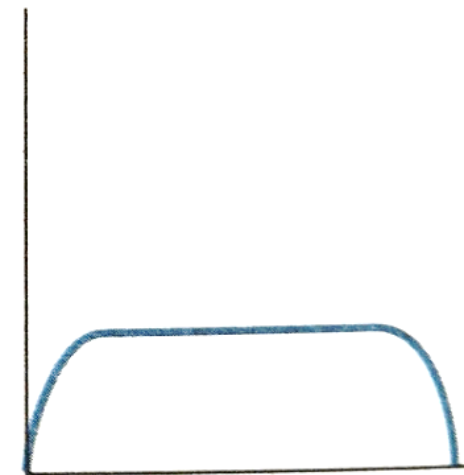
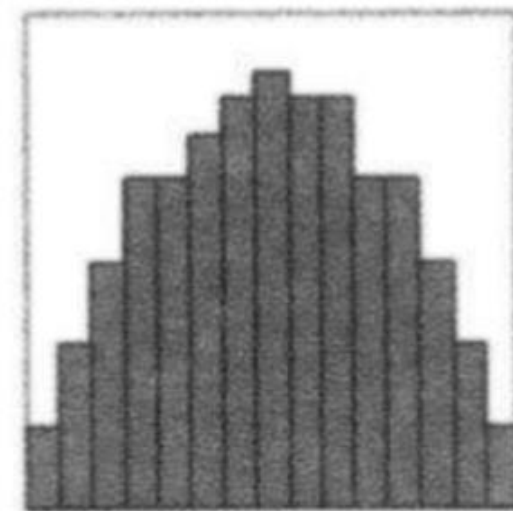
Κυρτότητα



β. Λεπτόκυρτη.



α. Μεσόκυρτη.



γ. Πλατύκυρτη.

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

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

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

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.

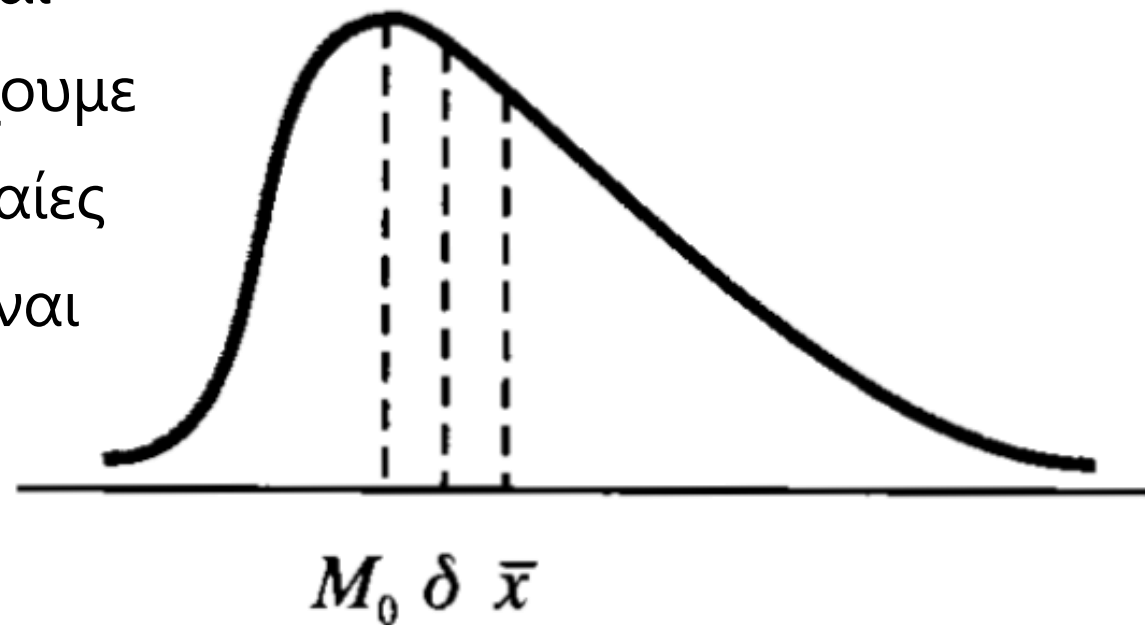
Δεν υπολογίζεται για ποιοτικά δεδομένα.

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

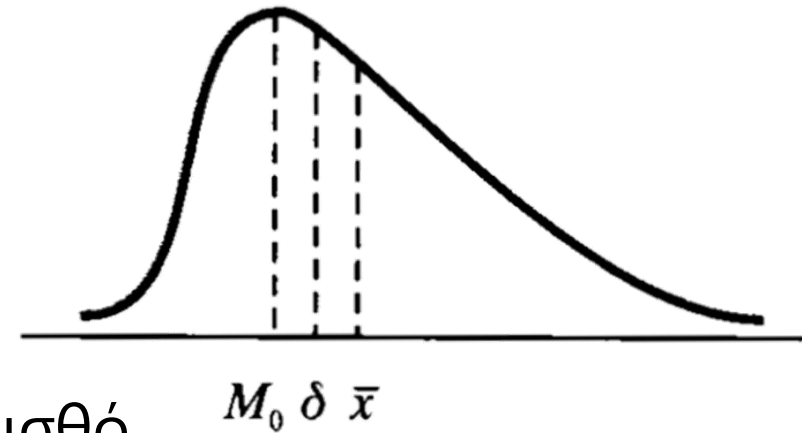
$$\bar{x} > \delta > M_0.$$



Παράδειγμα

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€).

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



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

Εφαρμογή 11

Δίνονται οι αριθμοί:

10, 11, 13, 15, 17, 17, 18, 19, 19, 19, 21, 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 12

Οι βαθμοί των παιδιών μιας τάξης σε ένα διαγώνισμα είναι:

8 6 15 17 12 6 5 20 10 7 20 6

Να υπολογίσετε:

α) το εύρος,

β) την επικρατούσα τιμή,

γ) τον μέσο όρο

δ) τα τεταρτημόρια Q_1 , Q_2 , Q_3

ε) το ενδοτεταρτημοριακό εύρος

στ) να κατασκευάσετε το θηκόγραμμα των βαθμών

Εφαρμογή 13

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.

Αριθμός παιδιών x_i	Οικογένειες n_i
0	3
1	5
2	8
3	3
4	1
Σύνολο	20

Εφαρμογή 14

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

Οι βαθμοί ενός μαθητή σε πέντε διαγωνίσματα είναι:

8, 10, 4, 16, 12

Να υπολογίσετε:

- α) τη μέση βαθμολογία του μαθητή,
- β) την τυπική απόκλιση των βαθμών,
- γ) τον συντελεστή μεταβλητότητας των βαθμών.

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

Εφαρμογή 15

Δίνεται ο διπλανός πίνακας κατανομής συχνοτήτων, στον οποίο φαίνεται η βαθμολογία (1 - 5 αστέρια) που έβαλαν σ' ένα εστιατόριο 20 άτομα. Να υπολογίσετε τη διακύμανση του παραπάνω δείγματος.

Βαθμολογία x_i	v_i
1	4
2	2
3	7
4	4
5	3
Σύνολο	20

Εφαρμογή 15

x_i	v_i	$x_i v_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 v_i$
1	4				
2	2				
3	7				
4	4				
5	3				
Σύνολο	20				

Εφαρμογή 16

Οι διοργανωτές ενός συνεδρίου θέλουν να δώσουν σε κάθε συμμετέχοντα ένα καρτελάκι μ' έναν (μοναδικό) κωδικό. Αποφασίστηκε οι κωδικοί να έχουν δύο σύμβολα: το 1ο να είναι ένα κεφαλαίο γράμμα του ελληνικού αλφαβήτου και το 2ο να είναι ένας μονοψήφιος αριθμός (0- 9). Ετοίμασαν καρτελάκια με όλους τους κωδικούς που ήταν δυνατόν να φτιαχτούν με τον παραπάνω τρόπο.

α) Πόσα καρτελάκια έφτιαξαν οι διοργανωτές;

Αν επιλέξουμε τυχαία ένα καρτελάκι, να υπολογίσετε τις πιθανότητες των ενδεχομένων:

i) A: το γράμμα του κωδικού είναι σύμφωνο,

ii) B: ο αριθμός του κωδικού είναι άρτιος,

iii) Γ: το γράμμα είναι φωνήεν και ο αριθμός είναι μικρότερος του 3.

Διατάξεις (με δυνατότητα επανάληψης)

- Διάταξη των **n στοιχείων** ενός συνόλου A **ανά k** ,
επιτρέποντας επαναλήψεις:
- Το πλήθος των διατάξεων των n ανά k είναι ίσο με n^k
- n : ο αριθμός των στοιχείων του συνόλου
 k : ο αριθμός των «θέσεων»

Εφαρμογή 17

Να υπολογίσετε πόσους:

α) **τετραψήφιους** φυσικούς αριθμούς μπορούμε να σχηματίσουμε χρησιμοποιώντας τα ψηφία 1, 2, 3, 4, 5, 6 όσες φορές θέλουμε το καθένα,

β) **εξαψήφιους** φυσικούς αριθμούς μπορούμε να σχηματίσουμε χρησιμοποιώντας τα ψηφία 7, 8, 9 όσες φορές θέλουμε το καθένα

Εφαρμογή 18

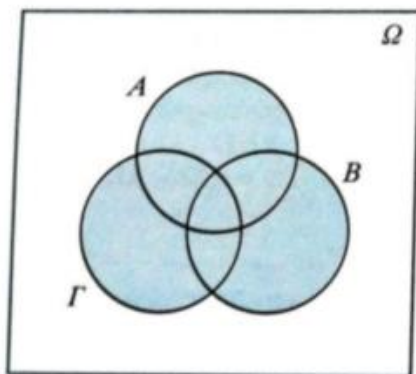
Έχετε μια ελιά (Ε), μια νερατζιά (Ν), μια πορτοκαλιά (Π) και μια λεμονιά (Λ). Πόσες διαφορετικές επιλογές έχετε για να δημιουργήσετε μια δενδροστοιχία των **δύο δένδρων** που θα επιλέξετε **από ένα σύνολο τεσσάρων δένδρων**;

$$4_{(2)} = \frac{4!}{(4-2)!} = \frac{1*2*3*4}{1*2} = 3 * 4 = 12$$

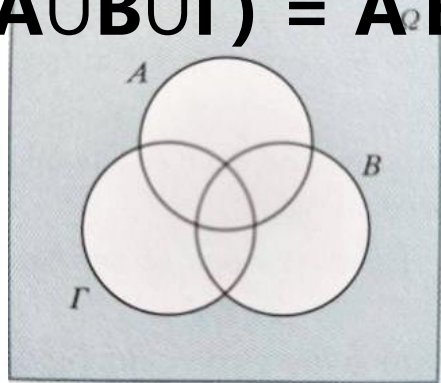
Εφαρμογή 19

Έστω A, B, Γ τρία ενδεχόμενα ενός δειγματικού χώρου Ω . Πώς μπορούν οι σχηματικές αναπαραστάσεις να εκφρασθούν με χρήση πράξεων μεταξύ των A, B, Γ ;

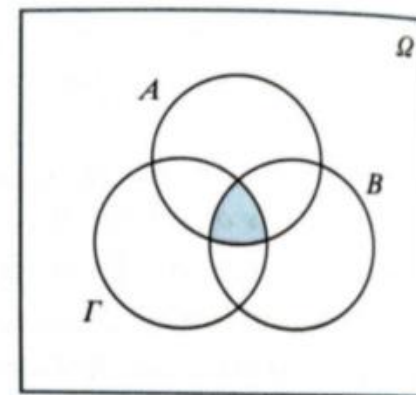
$$A \cup B \cup \Gamma$$



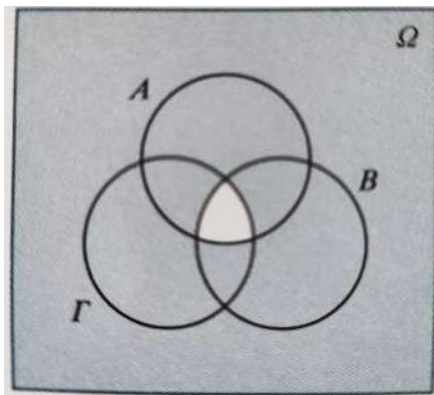
$$(A \cup B \cup \Gamma)' = A' B' \Gamma'$$



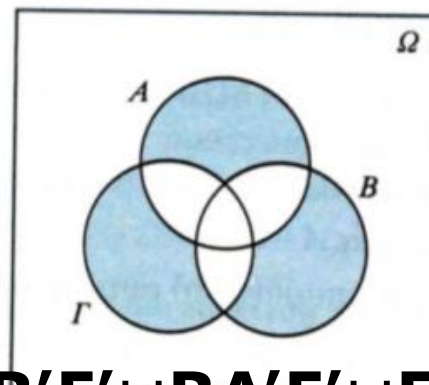
$$A \cap B \cap \Gamma = AB\Gamma$$



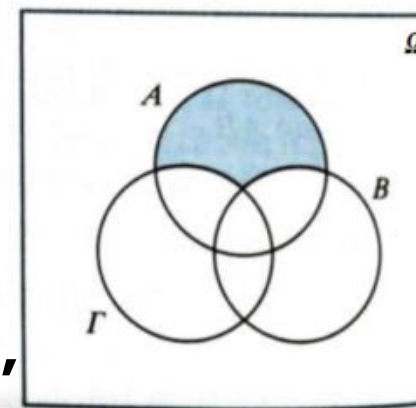
$$(AB\Gamma)'$$



$$AB'\Gamma' \cup BA'\Gamma' \cup \Gamma A'B'$$



$$A(B \cup \Gamma)' = AB'\Gamma'$$



Εφαρμογή 20

Ο δειγματικός χώρος ενός πειράματος τύχης είναι ο $\Omega = \{1, 2, 3, 4, 5\}$. Ισχύει:

$$P(\{1\})=P(\{5\}) = 1/10$$

$$P(\{2\})= 3/10$$

$$P(\{1,3\})= 1/2$$

α) Να υπολογίσετε τις πιθανότητες $P(\{3\})$ και $P(\{4\})$

β) Θεωρούμε τα ενδεχόμενα $A=\{1, 3, 5\}$ και $B= \{2, 4, 5\}$. Να υπολογίσετε τις πιθανότητες των ενδεχομένων A , $B - A$, $A \cup B$, και $A'B'$

Εφαρμογή 21

Έχουμε 10 κάρτες αριθμημένες από το 1 έως και το 10. Επιλέγουμε τυχαία μια κάρτα και γράφουμε τον αριθμό της.

α) Να γράψετε τον δειγματικό χώρο Ω του παραπάνω πειράματος τύχης,

β) Θεωρούμε τα ενδεχόμενα:

- A: ο αριθμός της κάρτας είναι περιττός
- B: ο αριθμός της κάρτας είναι ανάμεσα στο 2 και το 9
- Γ: ο αριθμός της κάρτας είναι τουλάχιστον 6

Να βρείτε τα ενδεχόμενα:

$A \cup B$, $B \cap \Gamma$, Γ' , $\Gamma - B$ $(A \cup B) \cap \Gamma$ $A \cup B'$