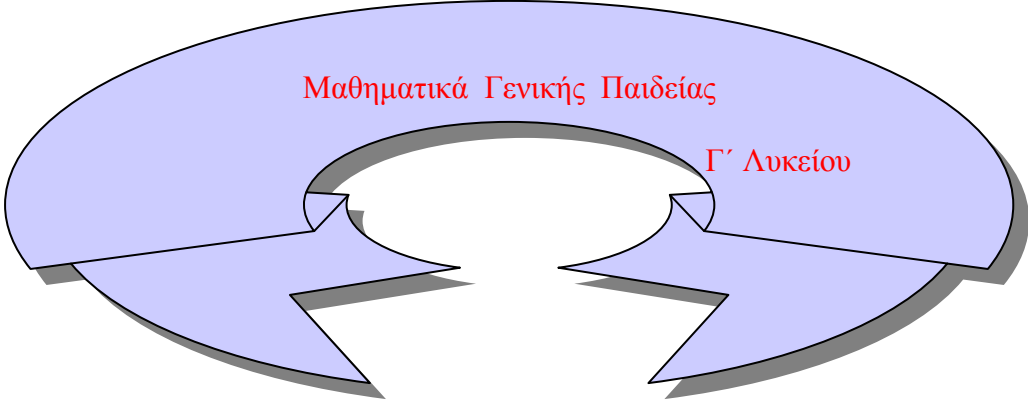
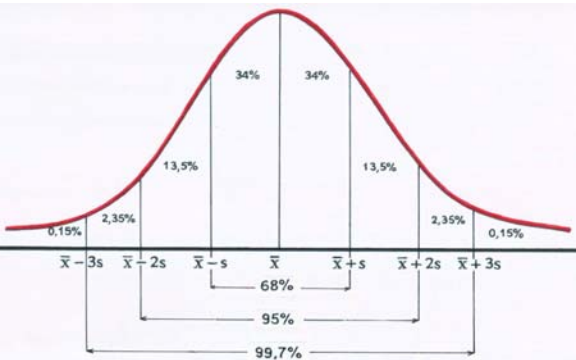
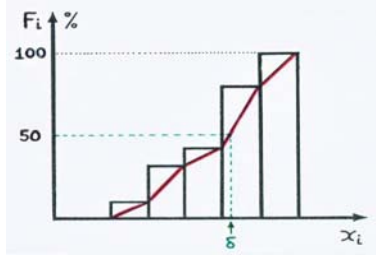


ΜΑΘΗΜΑΤΙΚΑ ΓΕΝΙΚΗΣ ΠΑΙΔΕΙΑΣ

ΣΤΑΤΙΣΤΙΚΗ



ΣΤΑΤΙΣΤΙΚΗ

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Στατιστική είναι ο κλάδος των εφαρμοσμένων επιστημών, η οποία βασίζεται σ' ένα σύνολο αρχών και μεθοδολογιών που έχουν σκοπό:

- Το σχεδιασμό της διαδικασίας **συλλογής** δεδομένων.
- Τη συνοπτική και αποτελεσματική **παρουσίασή** τους.
- Την **ανάλυση** και την **εξαγωγή χρήσιμων συμπερασμάτων**.

Στατιστικός πληθυσμός ή απλά **πληθυσμός** ονομάζεται κάθε σύνολο, τα στοιχεία του οποίου εξετάζουμε ως προς ένα ή περισσότερα χαρακτηριστικά τους. Τα στοιχεία του πληθυσμού ονομάζονται **μονάδες** ή **άτομα**.

Μεταβλητές ονομάζονται τα χαρακτηριστικά εκείνα, ως προς τα οποία εξετάζουμε έναν πληθυσμό. Οι δυνατές τιμές που μπορεί να πάρει μια μεταβλητή λέγονται **τιμές της μεταβλητής**.

Οι μεταβλητές χωρίζονται σε δύο κατηγορίες, στις **ποιοτικές** και στις **ποσοτικές**.

Ποιοτικές μεταβλητές είναι εκείνες που δεν επιδέχονται μέτρηση και οι τιμές τους δεν είναι αριθμοί.

Ποσοτικές μεταβλητές είναι εκείνες που επιδέχονται μέτρηση και οι τιμές τους είναι αριθμοί.

Οι ποσοτικές μεταβλητές διακρίνονται σε **συνεχείς** και **διακριτές (ασυνεχείς)**.

Συνεχείς είναι οι ποσοτικές μεταβλητές που μπορούν να πάρουν **οποιαδήποτε τιμή ενός διαστήματος** (α , β).

Διακριτές είναι οι ποσοτικές μεταβλητές που παίρνουν **μόνο μεμονωμένες τιμές**.

ΣΥΛΛΟΓΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Στατιστικά δεδομένα ή **παρατηρήσεις** είναι μια σειρά δεδομένων, που προκύπτουν από τη διαδοχική εξέταση των ατόμων ενός πληθυσμού ως προς ένα χαρακτηριστικό τους.

Οι κυριότερες μέθοδοι συλλογής στατιστικών δεδομένων είναι η απογραφή και η δειγματοληψία.

Απογραφή είναι μια μέθοδος συλλογής στατιστικών δεδομένων, που ακολουθούμε για να πάρουμε όλες τις απαραίτητες πληροφορίες, για έναν πληθυσμό εξετάζοντας όλα τα άτομα του πληθυσμού ως προς τα χαρακτηριστικά που μας ενδιαφέρουν.

Η απογραφή όμως είναι δύσκολη, οικονομικά και χρονικά ασύμφορη και πολλές φορές αδύνατη. Γι' αυτό το λόγο επιλέγουμε μια μικρή ομάδα, δηλαδή ένα υποσύνολο του πληθυσμού το οποίο ονομάζεται **δείγμα**.

Συλλέγουμε τις παρατηρήσεις από το δείγμα και στη συνέχεια γενικεύουμε τα συμπεράσματα για ολόκληρο τον πληθυσμό. Τα συμπεράσματα όμως, που θα προκύψουν από τη μελέτη του δείγματος θα είναι αξιόπιστα, δηλαδή θα ισχύουν με ικανοποιητική προσέγγιση για ολόκληρο τον πληθυσμό, μόνο όταν η επιλογή του δείγματος έχει γίνει με τέτοιο τρόπο, ώστε το δείγμα να είναι αντιπροσωπευτικό.

Ένα δείγμα θεωρείται **αντιπροσωπευτικό**, όταν κάθε άτομο του πληθυσμού έχει την ίδια δυνατότητα να επιλεγεί.

Δειγματοληψία είναι αυτή η μέθοδος συλλογής στατιστικών δεδομένων, κατά την οποία συγκεντρώνονται πληροφορίες **μόνο για ένα υποσύνολο του πληθυσμού**.

ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Στατιστικοί πίνακες

Μετά τη συλλογή των στατιστικών δεδομένων είναι αναγκαία η κατασκευή συνοπτικών πινάκων ή γραφικών παραστάσεων, ώστε να είναι εύκολη η κατανόησή τους και η εξαγωγή ορθών συμπερασμάτων.

Η παρουσίαση των στατιστικών δεδομένων σε πίνακες γίνεται με την κατάλληλη τοποθέτηση των πληροφοριών σε γραμμές και στήλες, ώστε να διευκολύνεται η σύγκριση και να έχουμε καλύτερη ενημέρωση.

Οι πίνακες διακρίνονται σε

- **Γενικούς πίνακες**, οι οποίοι περιέχουν όλες τις πληροφορίες μιας στατιστικής έρευνας και αποτελούν πηγές στατιστικών πληροφοριών για παραπέρα ανάλυση και εξαγωγή συμπερασμάτων.
- **Ειδικούς πίνακες**, οι οποίοι είναι συνοπτικοί και σαφείς. Τα στοιχεία τους έχουν ληφθεί από τους γενικούς πίνακες.

Κάθε πίνακας πρέπει να περιλαμβάνει

- **Τον τίτλο**, που γράφεται στο επάνω μέρος του πίνακα και δηλώνει συνοπτικά και με σαφήνεια το περιεχόμενό του.
- **Τις επικεφαλίδες** των γραμμών και των στηλών, που δείχνουν συνοπτικά τη φύση και τις μονάδες μέτρησης των δεδομένων.
- **Το κύριο σώμα**, που περιέχει διαχωρισμένα σε γραμμές και στήλες τα στατιστικά δεδομένα.
- **Την πηγή**, που δείχνει την προέλευση των στατιστικών δεδομένων, ώστε ο αναγνώστης να ανατρέξει σ' αυτή όταν το επιθυμεί, για επαλήθευση στοιχείων ή για άντληση περισσότερων πληροφοριών.

Συμβολισμοί :

◇ Μεταβλητή :	X
◇ Μέγεθος δείγματος :	N
◇ Παρατηρήσεις :	t₁ , t₂ , t₃ , ... , t_v
◇ Τιμές μεταβλητής :	x₁ , x₂ , x₃ , ... , x_κ
◇ Συχνότητες τιμών :	v₁ , v₂ , v₃ , ... , v_κ (κ ≤ v)
◇ Σχετικές συχνότητες τιμών :	f₁ , f₂ , f₃ , ... , f_κ
◇ Αθροιστικές συχνότητες τιμών :	N₁ , N₂ , N₃ , ... , N_κ
◇ Αθροιστικές σχετικές συχνότητες τιμών : (◇ Μόνο για ποσοτικές μεταβλητές)	F₁ , F₂ , F₃ , ... , F_κ

Ορισμοί :

α) Συχνότητα (v_i)

Αν $x_1, x_2, x_3, \dots, x_k$ οι τιμές μιας μεταβλητής X ενός δείγματος μεγέθους v , τότε **συχνότητα** της τιμής x_i , $i = 1, 2, 3, \dots, k$ λέγεται ο φυσικός αριθμός v_i που δείχνει πόσες φορές η μεταβλητή X παίρνει την τιμή x_i .

β) Σχετική συχνότητα (f_i)

Αν $x_1, x_2, x_3, \dots, x_k$ οι τιμές μιας μεταβλητής X ενός δείγματος μεγέθους v και $v_1, v_2, v_3, \dots, v_k$ οι αντίστοιχες συχνότητές τους, τότε **σχετική συχνότητα** της τιμής x_i , $i = 1, 2, 3, \dots, k$ λέγεται ο αριθμός

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

γ) Αθροιστική συχνότητα (N_i)

Αν οι τιμές $x_1, x_2, x_3, \dots, x_k$ μιας ποσοτικής μεταβλητής X , ενός δείγματος μεγέθους v , είναι σε αύξουσα διάταξη και $v_1, v_2, v_3, \dots, v_k$ οι αντίστοιχες συχνότητές τους, τότε **αθροιστική συχνότητα** της τιμής x_i , $i = 1, 2, 3, \dots, k$ λέγεται ο φυσικός αριθμός $N_i = v_1 + v_2 + v_3 + \dots + v_i$ που δείχνει πόσες παρατηρήσεις είναι μικρότερες ή ίσες της τιμής x_i .

δ) Αθροιστική σχετική συχνότητα (F_i)

Αν οι τιμές $x_1, x_2, x_3, \dots, x_k$ μιας ποσοτικής μεταβλητής X , ενός δείγματος μεγέθους v , είναι σε αύξουσα διάταξη και $f_1, f_2, f_3, \dots, f_k$ οι αντίστοιχες συχνότητές τους, τότε **αθροιστική σχετική συχνότητα** της τιμής x_i , $i = 1, 2, 3, \dots, k$ λέγεται ο αριθμός $F_i = f_1 + f_2 + f_3 + \dots + f_i$ που δείχνει το ποσοστό των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i .

ε) Κατανομή συχνοτήτων (x_i, v_i)

Αν $x_1, x_2, x_3, \dots, x_k$ οι τιμές μιας μεταβλητής X ενός δείγματος μεγέθους v και $v_1, v_2, v_3, \dots, v_k$ οι αντίστοιχες συχνότητές τους, τότε η **κατανομή συχνοτήτων** αποτελείται από το σύνολο των ζευγών της μορφής (x_i, v_i) , $i = 1, 2, 3, \dots, k$.

στ) Κατανομή σχετικών συχνοτήτων (x_i, f_i)

Αν $x_1, x_2, x_3, \dots, x_k$ οι τιμές μιας μεταβλητής X ενός δείγματος μεγέθους v και $f_1, f_2, f_3, \dots, f_k$ οι αντίστοιχες σχετικές συχνότητές τους, τότε η **κατανομή σχετικών συχνοτήτων** αποτελείται από το σύνολο των ζευγών της μορφής (x_i, f_i) , $i = 1, 2, 3, \dots, k$.

Πρόταση

Αν $x_1, x_2, x_3, \dots, x_k$ είναι οι τιμές μιας μεταβλητής X ενός δείγματος μεγέθους v και $v_1, v_2, v_3, \dots, v_k$ οι αντίστοιχες συχνότητές τους, όπου k, v είναι μη μηδενικοί φυσικοί αριθμοί με $k \leq v$, τότε ισχύουν:

α) $v_1 + v_2 + v_3 + \dots + v_k = v$.

β) $0 \leq f_i \leq 1$ για $i = 1, 2, 3, \dots, k$.

γ) $f_1 + f_2 + f_3 + \dots + f_k = 1$.

Απόδειξη

α) Είναι φανερό ότι το άθροισμα όλων των συχνοτήτων είναι ίσο με το μέγεθος n του δείγματος, δηλαδή $v_1 + v_2 + v_3 + \dots + v_k = n$.

β) Είναι $0 \leq v_i \leq n \Leftrightarrow \frac{0}{n} \leq \frac{v_i}{n} \leq \frac{n}{n} \Leftrightarrow 0 \leq f_i \leq 1$ για $i = 1, 2, 3, \dots, k$.

γ) $f_1 + f_2 + f_3 + \dots + f_k = \frac{v_1}{n} + \frac{v_2}{n} + \frac{v_3}{n} + \dots + \frac{v_k}{n} = \frac{v_1 + v_2 + v_3 + \dots + v_k}{n} = \frac{n}{n} = 1$.

Το σύμβολο " Σ "

Το άθροισμα $v_1 + v_2 + v_3 + \dots + v_k$ συμβολίζεται $\sum_{i=1}^k v_i$. Άρα ισχύουν:

$$\alpha) \sum_{i=1}^k v_i = n \quad \beta) \sum_{i=1}^k f_i = 1,$$

$$\gamma) \sum_{i=1}^{\lambda} v_i = N_{\lambda} \quad \delta) \sum_{i=1}^{\lambda} f_i = F_{\lambda} \quad \text{όπου } \lambda = 1, 2, 3, \dots, k.$$

Ιδιότητες " Σ "

$$\alpha) \sum_{i=1}^k (a_i + b_i) = \sum_{i=1}^k a_i + \sum_{i=1}^k b_i \quad \beta) \sum_{i=1}^k (\lambda a_i) = \lambda \sum_{i=1}^k a_i \quad \gamma) \sum_{i=1}^k \lambda = k\lambda$$

Π.χ.

$$\sum_{i=1}^k (3v_i - 5f_i + 4) = \sum_{i=1}^k 3v_i - \sum_{i=1}^k 5f_i + \sum_{i=1}^k 4 = 3 \sum_{i=1}^k v_i - 5 \sum_{i=1}^k f_i + 4k = 3n - 5 \cdot 1 + 4k = 3n - 5 + 4k.$$

Ομαδοποίηση παρατηρήσεων

Η ομαδοποίηση των παρατηρήσεων γίνεται όταν το πλήθος των τιμών μιας ποσοτικής μεταβλητής είναι μεγάλο.

Στην περίπτωση αυτή οι παρατηρήσεις κατανέμονται σε σύνολα (ομάδες) που λέγονται **κλάσεις** και οι οποίες είναι δυνατόν να έχουν ίσο ή άνισο πλάτος .

Για την ομαδοποίηση n παρατηρήσεων σε **κλάσεις ίσου πλάτους** εργαζόμαστε ως εξής :

- Προσδιορίζουμε τον αριθμό k των κλάσεων. Ο αριθμός αυτός εξαρτάται από το μέγεθος n του δείγματος και δίνεται από τον πίνακα της σελίδας 72 του σχολικού βιβλίου ή ορίζεται από τον ερευνητή σύμφωνα με την εμπειρία του.
- Προσδιορίζουμε το εύρος R των παρατηρήσεων. Είναι $R = t_{\max} - t_{\min}$
- Προσδιορίζουμε το πλάτος c κάθε κλάσης . Είναι $c = \frac{R}{k} = \frac{t_{\max} - t_{\min}}{k}$
- Καθορίζουμε τις κλάσεις $[a_0, a_1), [a_1, a_2), [a_2, a_3), \dots, [a_{k-1}, a_k]$.
- Κάνουμε διαλογή των παρατηρήσεων. Κατασκευάζουμε πίνακα .

Κεντρική τιμή της κλάσης $[a_{i-1}, a_i)$ είναι ο αριθμός $x_i = \frac{a_{i-1} + a_i}{2}$, $i = 1, 2, 3, \dots, k$.

Παρατήρηση

Κατά την ομαδοποίηση των παρατηρήσεων πρέπει:

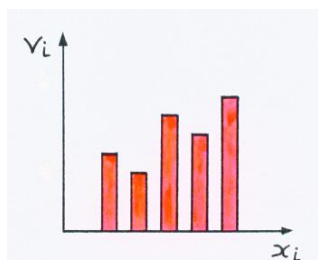
- ❖ Κάθε παρατήρηση να ανήκει σε μια κλάση.
- ❖ Η απόσταση δύο διαδοχικών κεντρικών τιμών να είναι ίση με το πλάτος c κάθε κλάσης

Γραφικές παραστάσεις

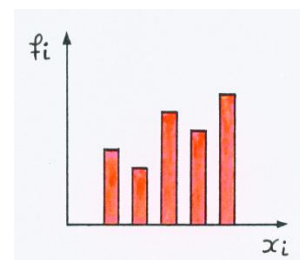
Τα στατιστικά δεδομένα παρουσιάζονται και με τη μορφή γραφικών παραστάσεων ή διαγραμμάτων. Με τα διαγράμματα διευκολύνεται η σύγκριση μεταξύ ομοειδών στοιχείων για το ίδιο χαρακτηριστικό ή και για διαφορετικά χαρακτηριστικά. Όπως και οι στατιστικοί πίνακες έτσι και τα στατιστικά διαγράμματα πρέπει να συνοδεύονται από τον τίτλο, την κλίμακα με τις τιμές των μεγεθών που απεικονίζονται, το υπόμνημα που επεξηγεί τις τιμές της μεταβλητής και την πηγή των δεδομένων.

A. Ποιοτική μεταβλητή

**Ραβδόγραμμα
συχνοτήτων
(x_i, v_i)**

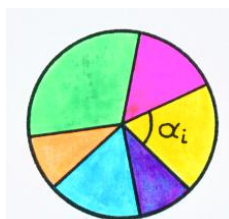


**Ραβδόγραμμα
σχετικών
συχνοτήτων
(x_i, f_i)**



Το **ραβδόγραμμα** χρησιμοποιείται για τη γραφική παράσταση των τιμών μιας **ποιοτικής** μεταβλητής. Αποτελείται από ορθογώνιες στήλες, που οι βάσεις τους βρίσκονται πάνω στον οριζόντιο άξονα ή και στον κατακόρυφο άξονα. Σε κάθε τιμή της μεταβλητής X αντιστοιχεί μία ορθογώνια στήλη της οποίας το ύψος είναι ίσο με την αντίστοιχη συχνότητα ή τη σχετική συχνότητα, έτσι έχουμε το **ραβδόγραμμα συχνοτήτων** ή το **ραβδόγραμμα σχετικών συχνοτήτων**. Η απόσταση μεταξύ των στηλών και το μήκος των βάσεων τους καθορίζονται αυθαίρετα.

**Κυκλικό διάγραμμα
συχνοτήτων
ή
σχετικών συχνοτήτων**



**Σημειόγραμμα
συχνοτήτων
ή
σχετικών συχνοτήτων**



Το **κυκλικό διάγραμμα** χρησιμοποιείται για τη γραφική παράσταση τόσο των **ποιοτικών** όσο και των **ποσοτικών** μεταβλητών, όταν οι τιμές της μεταβλητής είναι σχετικά λίγες.

Το **κυκλικό διάγραμμα** είναι ένας κυκλικός δίσκος χωρισμένος σε κυκλικούς τομείς των οποίων τα τόξα ή ισοδύναμα τα εμβαδά είναι ανάλογα προς τις αντίστοιχες συχνότητες v_i ή τις σχετικές συχνότητες f_i ή $f_i \%$ των τιμών x_i της μεταβλητής.

Αν συμβολίσουμε με α_i το αντίστοιχο τόξο ενός κυκλικού τομέα στο κυκλικό διάγραμμα συχνοτήτων ή σχετικών συχνοτήτων, τότε

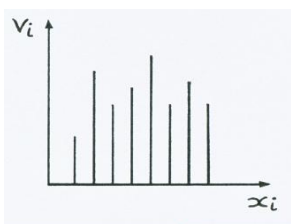
$$\alpha_i = 360^\circ \frac{v_i}{v} = 360^\circ f_i, \quad i = 1, 2, 3, \dots, \kappa.$$

Το **σημειόγραμμα** χρησιμοποιείται για τη γραφική παράσταση τόσο των **ποιοτικών** όσο και των **ποσοτικών** μεταβλητών, όταν τα στατιστικά δεδομένα είναι λίγα.

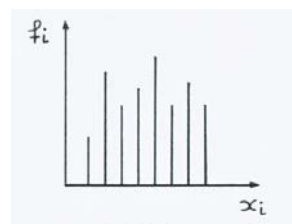
Το **σημειόγραμμα** αποτελείται από έναν οριζόντιο άξονα πάνω στον οποίο τοποθετούμε τις τιμές της μεταβλητής. Σε κάθε μία τιμή σημειώνουμε κατακόρυφα προς τα πάνω τόσα σημεία (τελείες) όσο και η συχνότητα της τιμής x_i ή σχετική συχνότητα f_i ή $f_i \%$.

Β. Ποσοτική μεταβλητή (Μη ομαδοποιημένες παρατηρήσεις)

**Διάγραμμα
συχνοτήτων
(x_i, v_i)**



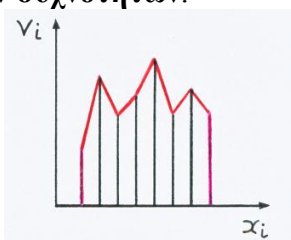
**Διάγραμμα
σχετικών συχνοτήτων
(x_i, f_i)**



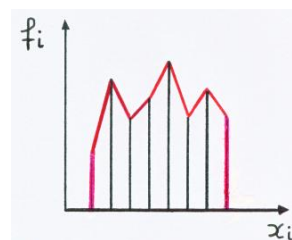
Το **διάγραμμα συχνοτήτων** χρησιμοποιείται για τη γραφική παράσταση των τιμών μιας **ποσοτικής** μεταβλητής (μη ομαδοποιημένης). Στο διάγραμμα συχνοτήτων, αντί να χρησιμοποιούμε συμπαγή ορθογώνια, υψώνουμε σε κάθε τιμή x_i της μεταβλητής, (υποθέτουμε ότι οι τιμές x_i είναι ταξινομημένες κατά αύξουσα σειρά), ένα κάθετο ευθύγραμμο τμήμα που έχει μήκος ίσο προς την αντίστοιχη συχνότητα v_i .

Αν αντί των συχνοτήτων v_i χρησιμοποιήσουμε τις σχετικές συχνότητες f_i ή $f_i \%$ παίρνουμε το **διάγραμμα σχετικών συχνοτήτων**.

**Πολύγωνο
συχνοτήτων
(x_i, v_i)**



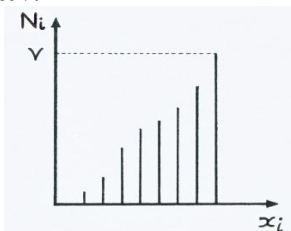
**Πολύγωνο
σχετικών συχνοτήτων
(x_i, f_i)**



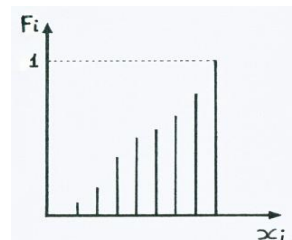
Αν σ' ένα διάγραμμα συχνοτήτων ενώσουμε τα σημεία (x_i, v_i) δημιουργείται μια τεθλασμένη γραμμή που ονομάζεται **πολύγωνο συχνοτήτων**.

Αν σ' ένα διάγραμμα σχετικών συχνοτήτων ενώσουμε τα σημεία (x_i, f_i) δημιουργείται μια τεθλασμένη γραμμή που ονομάζεται **πολύγωνο σχετικών συχνοτήτων**.

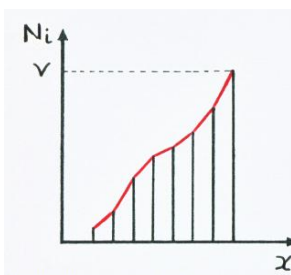
**Διάγραμμα
αθροιστικών
συχνοτήτων
(x_i, N_i)**



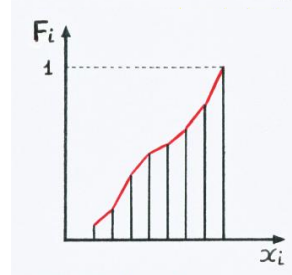
**Διάγραμμα
αθροιστικών
σχετικών
συχνοτήτων
(x_i, F_i)**



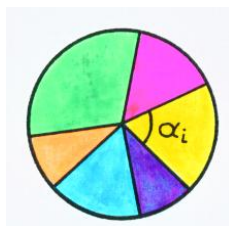
**Πολύγωνο
αθροιστικών
συχνοτήτων
(x_i, N_i)**



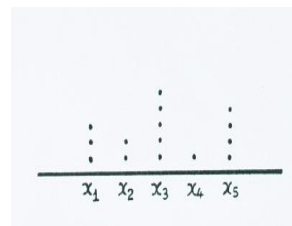
**Πολύγωνο
αθροιστικών
σχετικών
συχνοτήτων
(x_i, F_i)**



**Κυκλικό διάγραμμα
συχνοτήτων
ή
σχετικών συχνοτήτων**



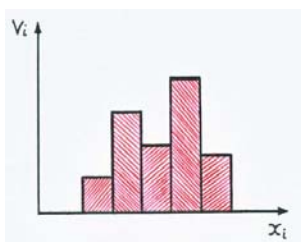
**Σημειόγραμμα
συχνοτήτων
ή
σχετικών συχνοτήτων**



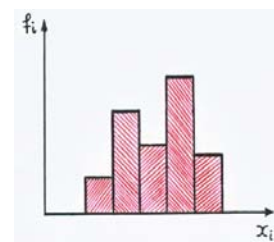
$$\alpha_i = 360^\circ \frac{v_i}{v} = 360^\circ f_i, \quad i = 1, 2, 3, \dots, \kappa.$$

Γ. Ποσοτική μεταβλητή (Ομαδοποιημένες παρατηρήσεις σε κλάσεις ίσου πλάτους)

**Ιστόγραμμα
συχνοτήτων**
(x_i, v_i)



**Ιστόγραμμα
σχετικών συχνοτήτων**
(x_i, f_i)



Το **ιστόγραμμα συχνοτήτων** χρησιμοποιείται για τη γραφική παράσταση ενός πίνακα συχνοτήτων με ομαδοποιημένα δεδομένα και είναι ένα σύστημα συντεταγμένων, όπου στον οριζόντιο άξονα σημειώνουμε με κατάλληλη κλίμακα τα όρια των κλάσεων.

Στη συνέχεια κατασκευάζουμε διαδοχικά ορθογώνια (ιστούς), καθένα από τα οποία έχει βάση ίση με το πλάτος της κλάσης και ύψος τέτοιο, ώστε το εμβαδόν του ορθογωνίου να είναι ίσο με τη συχνότητα της αντίστοιχης κλάσης.

Αν θεωρήσουμε το πλάτος c ως μονάδα μέτρησης του χαρακτηριστικού στον οριζόντιο άξονα, τότε το ύψος κάθε ορθογωνίου θα είναι ίσο με τη συχνότητα της αντίστοιχης κλάσης, αφού το εμβαδόν κάθε ορθογωνίου πρέπει να είναι ίσο με την αντίστοιχη συχνότητα.

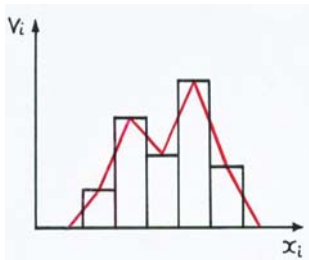
Συνεπώς σε κάθε ιστόγραμμα συχνοτήτων στον κατακόρυφο άξονα σημειώνουμε τις συχνότητες. Αν στον κατακόρυφο άξονα σημειώσουμε τις σχετικές συχνότητες, τότε με ανάλογο τρόπο κατασκευάζουμε το **ιστόγραμμα σχετικών συχνοτήτων**.

Αν συμβολίσουμε με E το άθροισμα των εμβαδών των ορθογωνίων σ' ένα :

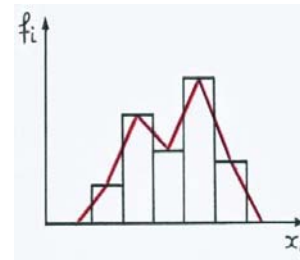
- Ιστόγραμμα συχνοτήτων, τότε $E = v_1 + v_2 + \dots + v_k \Leftrightarrow E = v$
- Ιστόγραμμα σχετικών συχνοτήτων, τότε:

$$E = f_1 + f_2 + \dots + f_k \Leftrightarrow E = 1 \quad \text{ή} \quad E = f_1 \% + f_2 \% + \dots + f_k \% \Leftrightarrow E = 100$$

**Πολύγωνο
συχνοτήτων**
(x_i, v_i)



**Πολύγωνο
σχετικών συχνοτήτων**
(x_i, f_i)



Αν σ' ένα **ιστόγραμμα συχνοτήτων** θεωρήσουμε δύο ακόμη υποθετικές κλάσεις, μία στην αρχή και μία στο τέλος με συχνότητα μηδέν και στη συνέχεια ενώσουμε τα μέσα των άνω βάσεων των ορθογωνίων, δηλαδή τα σημεία (x_i, v_i) , όπου x_i η κεντρική τιμή της κλάσης και v_i η αντίστοιχη συχνότητά της, τότε σχηματίζεται το **πολύγωνο συχνοτήτων**.

Το **εμβαδόν του χωρίου** που ορίζεται από το πολύγωνο συχνοτήτων και τον οριζόντιο άξονα είναι ίσο με το άθροισμα των εμβαδών των ορθογωνίων, δηλαδή

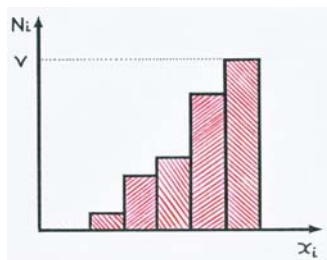
$$E = v$$

Αν σ' ένα **ιστόγραμμα σχετικών συχνοτήτων** θεωρήσουμε δύο ακόμη υποθετικές κλάσεις, μία στην αρχή και μία στο τέλος με σχετική συχνότητα μηδέν και στη συνέχεια ενώσουμε τα μέσα των άνω βάσεων των ορθογωνίων, δηλαδή τα σημεία (x_i, f_i) , όπου x_i η κεντρική τιμή της κλάσης και f_i η αντίστοιχη σχετική συχνότητά της, τότε σχηματίζεται το **πολύγωνο σχετικών συχνοτήτων**.

Το **εμβαδόν του χωρίου** που ορίζεται από το πολύγωνο σχετικών συχνοτήτων και τον οριζόντιο άξονα είναι ίσο με το άθροισμα των εμβαδών των ορθογωνίων, δηλαδή

$$E = 1 \quad \text{ή} \quad E = 100$$

**Ιστόγραμμα
αθροιστικών
συχνοτήτων
(x_i, N_i)**



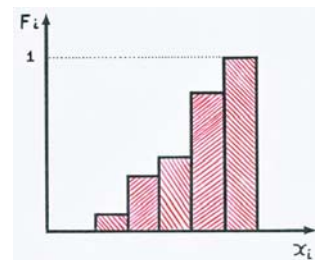
Με όμοιο τρόπο κατασκευάζονται :

Το ιστόγραμμα αθροιστικών συχνοτήτων

Το εμβαδόν του τελευταίου ιστού στο ιστόγραμμα αθροιστικών συχνοτήτων είναι

$$E_k = v$$

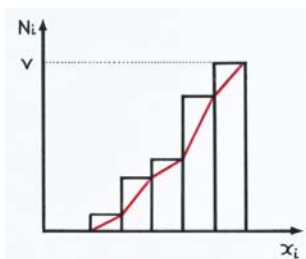
**Ιστόγραμμα
αθροιστικών
σχετικών
συχνοτήτων
(x_i, F_i)**



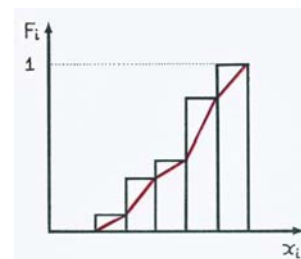
Το ιστογράμματα αθροιστικών σχετικών συχνοτήτων.

Το εμβαδόν του τελευταίου ιστού στο ιστόγραμμα αθροιστικών σχετικών συχνοτήτων είναι $E_k = 1$ ή $E_k = 100$

**Πολύγωνο
αθροιστικών
συχνοτήτων
(x_i, N_i)**



**Πολύγωνο
αθροιστικών
σχετικών
συχνοτήτων
(x_i, F_i)**

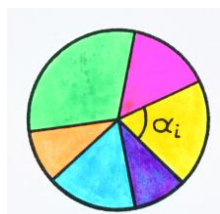


Αν σ' ένα ιστόγραμμα αθροιστικών συχνοτήτων ενώσουμε τα δεξιά άκρα των άνω βάσεων με ευθύγραμμα τμήματα, τότε προκύπτει το **πολύγωνο αθροιστικών συχνοτήτων**. Με όμοιο τρόπο κατασκευάζεται το **πολύγωνο αθροιστικών σχετικών συχνοτήτων**.

Με το **πολύγωνο αθροιστικών συχνοτήτων** βρίσκουμε **πλήθος** που δεν προκύπτει άμεσα από πίνακα.

Με το **πολύγωνο αθροιστικών σχετικών συχνοτήτων** βρίσκουμε **ποσοστό** που δεν προκύπτει άμεσα από πίνακα.

**Κυκλικό διάγραμμα
συχνοτήτων
ή
σχετικών συχνοτήτων**

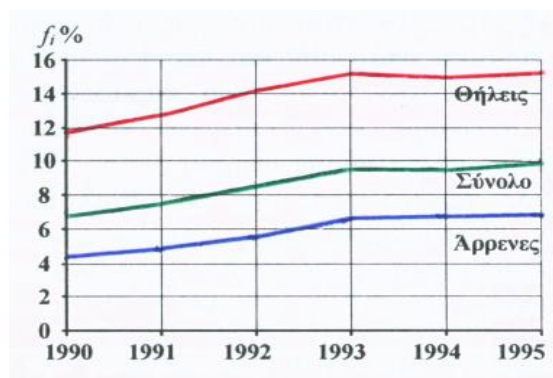


$$\alpha_i = 360^\circ \frac{v_i}{v} = 360^\circ f_i, \quad i=1,2,3,\dots,k.$$

Δ. Χρονόγραμμα

Το **χρονόγραμμα** χρησιμοποιείται για τη γραφική απεικόνιση της **διαχρονικής εξέλιξης** ενός οικονομικού, δημογραφικού ή άλλου μεγέθους, όπως η αξία μιας μετοχής, τα ποσοστά ανεργίας, το ύψος του πληθωρισμού, το πλήθος των γεννήσεων κ.τ.λ.

Αποτελείται από ένα ορθογώνιο σύστημα αξόνων, όπου ο οριζόντιος άξονας θεωρείται ως άξονας μέτρησης του χρόνου και ο κατακόρυφος άξονας θεωρείται ως άξονας μέτρησης της συχνότητας v_i ή της σχετικής συχνότητας f_i ή $f_i \%$ των τιμών της εξεταζόμενης μεταβλητής.



Ποσοστά ανεργίας στην Ελλάδα κατά τα έτη 1990-1995

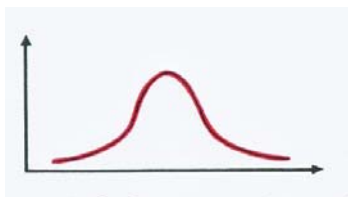
Ε. Καμπύλη συχνοτήτων

Καμπύλη συχνοτήτων είναι η ομαλή καμπύλη προς την οποία τείνει η πολυγωνική γραμμή συχνοτήτων, όταν το πλήθος των κλάσεων είναι αρκετά μεγάλο (τείνει στο άπειρο) και το πλάτος κάθε κλάσης είναι μικρό (τείνει στο μηδέν).

Η μορφή της καμπύλης συχνοτήτων εξαρτάται από το πώς είναι κατανομημένες οι παρατηρήσεις σ' όλη την έκταση του εύρους τους.

ΣΤ. Μορφές καμπύλης συχνοτήτων

Κανονική κατανομή



Αν οι παρατηρήσεις κατανέμονται συμμετρικά ως προς μια κατακόρυφη ευθεία, τότε λέμε ότι ακολουθούν την **κανονική κατανομή**.

Η καμπύλη που αντιστοιχεί στην κανονική κατανομή έχει “**κωδωνοειδή**” μορφή.

Η συχνότητα που αντιστοιχεί στον άξονα συμμετρίας της καμπύλης είναι η μεγαλύτερη από όλες τις άλλες συχνότητες και σημαίνει ότι η αντίστοιχη τιμή της μεταβλητής εμφανίζεται πιο συχνά από κάθε άλλη τιμή.

Ομοιόμορφη κατανομή

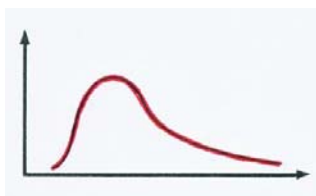


Αν οι παρατηρήσεις κατανέμονται ομοιόμορφα σ' ένα διάστημα $[α, β]$, τότε λέμε ότι ακολουθούν την **ομοιόμορφη κατανομή**.

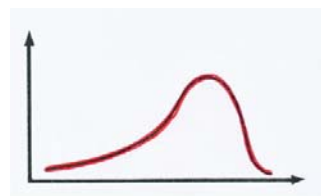
Στην ομοιόμορφη κατανομή κάθε παρατήρηση ή κλάση έχει την ίδια συχνότητα

Αν οι παρατηρήσεις δεν είναι συμμετρικά κατανομημένες, τότε η κατανομή λέγεται **ασύμμετρη** με:

Θετική ασυμμετρία



Αρνητική ασυμμετρία



ΜΕΤΡΑ ΘΕΣΗΣ

α) Μέση Τιμή μεταβλητής X (\bar{x})

Ορισμός

Αν $t_1, t_2, t_3, \dots, t_v$ είναι οι παρατηρήσεις μιας ποσοτικής μεταβλητής X ενός δείγματος μεγέθους v , $x_1, x_2, x_3, \dots, x_k$ είναι οι τιμές της μεταβλητής X με αντίστοιχες συχνότητες $v_1, v_2, v_3, \dots, v_k$ και σχετικές συχνότητες $f_1, f_2, f_3, \dots, f_k$, τότε **μέση τιμή** της μεταβλητής της X είναι ο αριθμός \bar{x} , όπου

$$\bar{x} = \frac{1}{v} \sum_{i=1}^v t_i = \frac{1}{v} \sum_{i=1}^k x_i v_i = \sum_{i=1}^k x_i f_i$$

Εύρεση Μέσης Τιμής (\bar{x})

1^η Περίπτωση:

Αν δίνονται οι παρατηρήσεις $t_1, t_2, t_3, \dots, t_v$, τότε χρησιμοποιούμε τον τύπο

$$\bar{x} = \frac{1}{v} \sum_{i=1}^v t_i = \frac{t_1 + t_2 + t_3 + \dots + t_v}{v}$$

2^η Περίπτωση:

Αν δίνονται οι τιμές $x_1, x_2, x_3, \dots, x_k$ και οι αντίστοιχες συχνότητες τους $v_1, v_2, v_3, \dots, v_k$, τότε :

α) Κάνουμε πίνακα συχνοτήτων

x_i	v_i	$x_i v_i$
ΣΥΝΟΛΟ	v	$\sum x_i v_i$

β) Χρησιμοποιούμε τον τύπο

$$\bar{x} = \frac{1}{v} \sum_{i=1}^k x_i v_i$$

3^η Περίπτωση:

Αν δίνονται οι τιμές $x_1, x_2, x_3, \dots, x_k$ και οι αντίστοιχες σχετικές συχνότητες τους $f_1, f_2, f_3, \dots, f_k$, τότε :

α) Κάνουμε πίνακα σχετικών συχνοτήτων

x_i	f_i	$x_i f_i$
ΣΥΝΟΛΟ	1	$\sum x_i f_i$

β) Χρησιμοποιούμε τον τύπο

$$\bar{x} = \sum_{i=1}^k x_i f_i$$

Σημείωση

Όταν οι παρατηρήσεις είναι ομαδοποιημένες εργαζόμαστε όπως στη 2^η ή 3^η περίπτωση και ως x_i θεωρούμε τις κεντρικές τιμές των κλάσεων.

β) Σταθμικός μέσος μεταβλητής X (\bar{x})

Ορισμός

Αν $x_1, x_2, x_3, \dots, x_v$ είναι οι παρατηρήσεις μιας ποσοτικής μεταβλητής X ενός δείγματος μεγέθους v και $w_1, w_2, w_3, \dots, w_v$ οι αντίστοιχοι συντελεστές στάθμισης, τότε **σταθμικός μέσος** της μεταβλητής X ονομάζεται ο αριθμός \bar{x} , όπου

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 + \dots + x_v w_v}{w_1 + w_2 + w_3 + \dots + w_v} = \frac{\sum_{i=1}^v x_i w_i}{\sum_{i=1}^v w_i}$$

γ) Διάμεσος παρατηρήσεων (δ)

Ορισμός

Αν $t_1, t_2, t_3, \dots, t_v$ είναι οι παρατηρήσεις μιας ποσοτικής μεταβλητής X , οι οποίες είναι διατεταγμένες κατά αύξουσα σειρά, τότε η **διάμεσος** των παρατηρήσεων είναι

$$\delta = \begin{cases} \frac{t_k + t_{k+1}}{2} & \text{αν } v = 2k \\ t_{k+1} & \text{αν } v = 2k + 1 \end{cases}$$

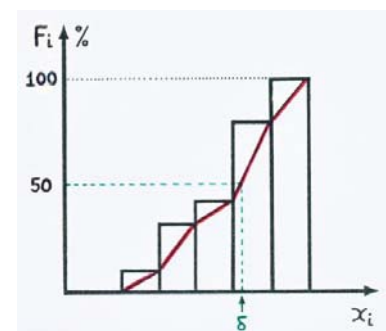
Δηλαδή, η διάμεσος είναι ίση με

- το ημίαθροισμα των δύο μεσαίων παρατηρήσεων, αν v άρτιος.
- τη μεσαία παρατήρηση, αν v περιττός.

Η διάμεσος σε ομαδοποιημένα δεδομένα

Για να προσδιορίσουμε τη διάμεσο v παρατηρήσεων, όταν αυτές είναι ομαδοποιημένες, εργαζόμαστε ως εξής:

- Κατασκευάζουμε πίνακα κατανομής σχετικών συχνοτήτων.
- Σχεδιάζουμε ιστόγραμμα και πολύγωνο αθροιστικών σχετικών συχνοτήτων (F_i %).
- Προσδιορίζουμε την τιμή που αντιστοιχεί σε ποσοστό 50 %.



Σημείωση

Η διάμεσος είναι η τιμή που χωρίζει ένα σύνολο παρατηρήσεων σε δύο ίσα μέρη, όταν οι παρατηρήσεις είναι διατεταγμένες κατά αύξουσα σειρά, δηλαδή η διάμεσος είναι η τιμή για την οποία το πολύ το 50 % των παρατηρήσεων είναι μικρότερες από αυτήν και το πολύ το 50 % των παρατηρήσεων είναι μεγαλύτερες από την τιμή αυτήν.

Σύγκριση μέσης τιμής και διαμέσου

	Πλεονεκτήματα	Μειονεκτήματα
Μέση Τιμή	<ul style="list-style-type: none">• Για τον υπολογισμό της χρησιμοποιούνται όλες οι τιμές.• Είναι μοναδική για κάθε σύνολο δεδομένων.• Είναι εύκολα κατανοητή.• Ο υπολογισμός της είναι σχετικά εύκολος.• Έχει μεγάλη εφαρμογή για περαιτέρω στατιστική ανάλυση.	<ul style="list-style-type: none">• Επηρεάζεται πολύ από ακραίες τιμές.• Συνήθως δεν αντιστοιχεί σε τιμή της μεταβλητής.• Δεν υπολογίζεται για ποιοτικά δεδομένα
Διάμεσος	<ul style="list-style-type: none">• Είναι εύκολα κατανοητή.• Δεν επηρεάζεται από ακραίες τιμές.• Ο υπολογισμός της είναι απλός.• Είναι μοναδική για κάθε σύνολο δεδομένων.	<ul style="list-style-type: none">• Δεν χρησιμοποιούνται όλες οι τιμές για τον υπολογισμό της.• Είναι δύσκολη η εφαρμογή της για περαιτέρω στατιστική ανάλυση.• Δεν υπολογίζεται για ποιοτικά δεδομένα

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

A) Εύρος παρατηρήσεων (R)

Ορισμός

Αν $t_1, t_2, t_3, \dots, t_n$ είναι οι παρατηρήσεις μιας ποσοτικής μεταβλητής X , τότε **εύρος** των παρατηρήσεων είναι ο αριθμός **R**, που είναι ίσος με τη διαφορά της ελάχιστης παρατήρησης από τη μέγιστη παρατήρηση. Δηλαδή

$$R = t_{\max} - t_{\min}$$

B) Διακύμανση (S^2)

Ορισμός

Αν $t_1, t_2, t_3, \dots, t_n$ είναι οι παρατηρήσεις μιας ποσοτικής μεταβλητής X ενός δείγματος μεγέθους v , $x_1, x_2, x_3, \dots, x_k$ είναι οι τιμές της μεταβλητής X με αντίστοιχες συχνότητες $v_1, v_2, v_3, \dots, v_k$, τότε **διακύμανση** S^2 της μεταβλητής X είναι

$$S^2 = \frac{1}{v} \sum_{i=1}^v (t_i - \bar{x})^2 = \frac{1}{v} \sum_{i=1}^k v_i (x_i - \bar{x})^2$$

όπου \bar{x} είναι η μέση τιμή της μεταβλητής X .

Σημείωση

Οι προηγούμενοι τύποι γράφονται και ως εξής :

$$\bullet \quad S^2 = \frac{1}{v} \left[\sum_{i=1}^v t_i^2 - \frac{\left(\sum_{i=1}^v t_i \right)^2}{v} \right] = \frac{1}{v} \sum_{i=1}^v t_i^2 - \left(\frac{\sum_{i=1}^v t_i}{v} \right)^2 = \overline{t^2} - (\bar{t})^2$$

$$\bullet \quad S^2 = \frac{1}{v} \left[\sum_{i=1}^k x_i^2 v_i - \frac{\left(\sum_{i=1}^k x_i v_i \right)^2}{v} \right] = \frac{1}{v} \sum_{i=1}^k x_i^2 v_i - \left(\frac{\sum_{i=1}^k x_i v_i}{v} \right)^2 = \overline{x^2} - (\bar{x})^2$$

Εύρεση Διακύμανσης (S^2)

1^η Περίπτωση:

Αν δίνονται οι παρατηρήσεις $t_1, t_2, t_3, \dots, t_v$, τότε

α) Βρίσκουμε τη μέση τιμή από τον τύπο $\bar{x} = \frac{1}{v} \sum_{i=1}^v t_i = \frac{t_1 + t_2 + t_3 + \dots + t_v}{v}$

β) Βρίσκουμε τη διακύμανση από τον τύπο

$$S^2 = \frac{1}{v} \sum_{i=1}^v (t_i - \bar{x})^2 = \frac{(t_1 - \bar{x})^2 + (t_2 - \bar{x})^2 + \dots + (t_v - \bar{x})^2}{v} \quad (\text{Συνήθως αν } \bar{x} \text{ ακέραιος})$$

ή από τον τύπο

$$S^2 = \frac{1}{v} \sum_{i=1}^v t_i^2 - (\bar{x})^2 = \overline{x^2} - (\bar{x})^2 \quad (\text{Συνήθως αν } \bar{x} \text{ όχι ακέραιος})$$

2^η Περίπτωση:

Αν δίνονται οι τιμές $x_1, x_2, x_3, \dots, x_k$ και οι αντίστοιχες συχνότητες τους $v_1, v_2, v_3, \dots, v_k$, τότε:

α) Κάνουμε πίνακα συχνοτήτων

x_i	v_i	$x_i v_i$	$x_i^2 v_i$
ΣΥΝΟΛΟ	v	$\sum x_i v_i$	$\sum x_i^2 v_i$

β) Βρίσκουμε τη μέση τιμή από τον τύπο

$$\bar{x} = \frac{1}{v} \sum_{i=1}^k x_i v_i$$

γ) Βρίσκουμε τη διακύμανση από τον τύπο

$$S^2 = \frac{1}{v} \sum_{i=1}^k x_i^2 v_i - (\bar{x})^2$$

3^η Περίπτωση:

Αν δίνονται οι τιμές $x_1, x_2, x_3, \dots, x_k$ και οι αντίστοιχες σχετικές συχνότητες τους $f_1, f_2, f_3, \dots, f_k$, τότε :

α) Κάνουμε πίνακα σχετικών συχνοτήτων

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
ΣΥΝΟΛΟ	1	$\sum x_i f_i$	$\sum x_i^2 f_i$

β) Βρίσκουμε τη μέση τιμή από τον τύπο

$$\bar{x} = \sum_{i=1}^k x_i f_i$$

γ) Βρίσκουμε τη διακύμανση από τον τύπο

$$S^2 = \sum_{i=1}^k x_i^2 f_i - (\bar{x})^2$$

Σημείωση

Όταν οι παρατηρήσεις είναι ομαδοποιημένες εργαζόμαστε όπως στη 2^η ή 3^η περίπτωση και ως x_i θεωρούμε τις κεντρικές τιμές των κλάσεων.

γ) Τυπική Απόκλιση (s)

Ορισμός

Τυπική απόκλιση ονομάζεται η θετική τετραγωνική ρίζα της διακύμανσης, δηλαδή

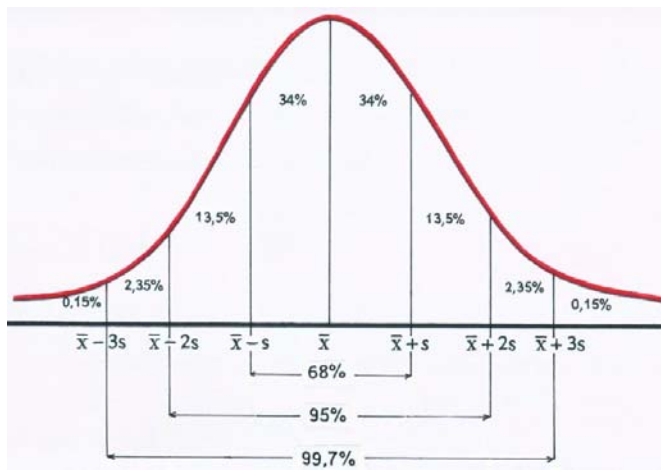
$$s = \sqrt{s^2}$$

Ιδιότητες

Αν η καμπύλη συχνοτήτων για το χαρακτηριστικό που εξετάζουμε είναι κανονική ή περίπου κανονική, τότε η τυπική απόκλιση s έχει τις παρακάτω ιδιότητες:

i) Στο διάστημα $(\bar{x}-s, \bar{x}+s)$ ανήκει το 68% περίπου των παρατηρήσεων, οπότε σε καθένα από τα διαστήματα $(\bar{x}-s, \bar{x})$ και $(\bar{x}, \bar{x}+s)$ ανήκει το 34% περίπου των παρατηρήσεων.

ii) Στο διάστημα $(\bar{x}-2s, \bar{x}+2s)$ ανήκει το 95% περίπου των παρατηρήσεων, οπότε σε καθένα από τα διαστήματα $(\bar{x}-2s, \bar{x}-s)$ και $(\bar{x}+s, \bar{x}+2s)$ ανήκει το $\frac{95-68}{2} = 13,5\%$ περίπου των παρατηρήσεων.



iii) Στο διάστημα $(\bar{x}-3s, \bar{x}+3s)$ ανήκει περίπου το 99,7% των παρατηρήσεων, οπότε σε καθένα από τα διαστήματα $(\bar{x}-3s, \bar{x}-2s)$ και $(\bar{x}+2s, \bar{x}+3s)$ ανήκει περίπου το $\frac{99,7-95}{2} = 2,35\%$ των παρατηρήσεων.

iv) Το εύρος είναι περίπου ίσο με έξι τυπικές αποκλίσεις, δηλαδή $R \approx 6s$.

δ) Συντελεστής Μεταβολής (CV)

Ορισμός

Συντελεστής μεταβολής ονομάζεται ο λόγος της τυπικής απόκλισης προς τη μέση τιμή, δηλαδή

$$CV = \frac{s}{\bar{x}} \quad (\bar{x} \neq 0)$$

Αν $\bar{x} < 0$, τότε αντί της \bar{x} χρησιμοποιούμε την $|\bar{x}|$.

Παρατηρήσεις

- ❖ Ο συντελεστής μεταβολής είναι ένα μέτρο ομοιογένειας, το οποίο χρησιμοποιείται για τη σύγκριση ομάδων τιμών που εκφράζονται
 - σε διαφορετικές μονάδες μέτρησης ή
 - στην ίδια μονάδα μέτρησης αλλά έχουν σημαντικά διαφορετικές μέσες τιμές.
- ❖ Ο συντελεστής μεταβολής
 - είναι ανεξάρτητος από τις μονάδες μέτρησης.
 - εκφράζεται επί τοις εκατό.
 - είναι ένα μέτρο σχετικής διασποράς των τιμών και όχι της απόλυτης διασποράς.

- ❖ Αν για δύο δείγματα τιμών A και B ισχύει $CV_A < CV_B$, τότε λέμε ότι το δείγμα A παρουσιάζει μεγαλύτερη ομοιογένεια από το δείγμα B.
- ❖ Ένα δείγμα τιμών A θεωρείται ομοιογενές, όταν $CV_A \leq 10\%$.

Σύγκριση μέτρων διασποράς

	Πλεονεκτήματα	Μειονεκτήματα
Εύρος	<ul style="list-style-type: none"> • Ο υπολογισμός του είναι σχετικά εύκολος. • Χρησιμοποιείται συχνά στον έλεγχο ποιότητας. • Είναι δυνατόν να χρησιμοποιηθεί για την εκτίμηση της τυπικής απόκλισης. 	<ul style="list-style-type: none"> • Δε θεωρείται αξιόπιστο μέτρο διασποράς, επειδή βασίζεται μόνο στις δύο ακραίες παρατηρήσεις. • Δε χρησιμοποιείται για περαιτέρω στατιστική ανάλυση.
Διακύμανση – Τυπική Απόκλιση	<ul style="list-style-type: none"> • Λαμβάνονται υπόψη για τον υπολογισμό τους όλες οι παρατηρήσεις. • Έχουν μεγάλη εφαρμογή στη στατιστική συμπερασματολογία. • Σε κανονικούς στατιστικούς πληθυσμούς το 68 %, το 95 %, και το 99,7 % των παρατηρήσεων ανήκουν στα διαστήματα $(\bar{x} - s, \bar{x} + s)$, $(\bar{x} - 2s, \bar{x} + 2s)$ και $(\bar{x} - 3s, \bar{x} + 3s)$ αντιστοίχως. 	<ul style="list-style-type: none"> • Απαιτούνται περισσότερες αλγεβρικές πράξεις για τον υπολογισμό τους παρά στα άλλα μέτρα. • Το κυριότερο μειονέκτημα της διακύμανσης είναι ότι δεν εκφράζεται στις ίδιες μονάδες με το χαρακτηριστικό. Το μειονέκτημα αυτό παύει να υπάρχει με τη χρησιμοποίηση της τυπικής απόκλισης
Συντελεστής Μεταβολής	<ul style="list-style-type: none"> • Είναι καθαρός αριθμός. • Χρησιμοποιείται ως μέτρο σύγκρισης της μεταβλητότητας, όταν έχουμε ίδιες ή και διαφορετικές μονάδες μέτρησης. • Χρησιμοποιείται ως μέτρο ομοιογένειας ενός στατιστικού πληθυσμού. 	<ul style="list-style-type: none"> • Δεν ενδείκνυται στην περίπτωση που η μέση τιμή είναι κοντά στο μηδέν.

Πρόταση

Έστω $x_1, x_2, x_3, \dots, x_n$ οι παρατηρήσεις μιας ποσοτικής μεταβλητής X με μέση τιμή \bar{x} και τυπική απόκλιση s_x και $y_1, y_2, y_3, \dots, y_n$ οι παρατηρήσεις μιας ποσοτικής μεταβλητής Y με μέση τιμή \bar{y} και τυπική απόκλιση s_y . Αν $y_i = \alpha x_i + \beta$ για κάθε $i = 1, 2, 3, \dots, n$ και $\alpha, \beta \in \mathbf{R}$, τότε:

$$\text{i) } \bar{y} = \alpha \bar{x} + \beta \qquad \text{ii) } s_y = |\alpha| \cdot s_x$$

Απόδειξη

i) Είναι

$$\begin{aligned} \bar{y} &= \frac{y_1 + y_2 + y_3 + \dots + y_n}{n} = \frac{(\alpha x_1 + \beta) + (\alpha x_2 + \beta) + (\alpha x_3 + \beta) + \dots + (\alpha x_n + \beta)}{n} = \\ &= \alpha \cdot \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} + \frac{n\beta}{n} = \alpha \bar{x} + \beta. \end{aligned}$$

ii) Είναι

$$\begin{aligned} s_y^2 &= \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + (y_3 - \bar{y})^2 + \dots + (y_n - \bar{y})^2}{n} = \\ &= \frac{(\alpha x_1 + \beta - \alpha \bar{x} - \beta)^2 + (\alpha x_2 + \beta - \alpha \bar{x} - \beta)^2 + (\alpha x_3 + \beta - \alpha \bar{x} - \beta)^2 + \dots + (\alpha x_n + \beta - \alpha \bar{x} - \beta)^2}{n} = \\ &= \frac{(\alpha x_1 - \alpha \bar{x})^2 + (\alpha x_2 - \alpha \bar{x})^2 + (\alpha x_3 - \alpha \bar{x})^2 + \dots + (\alpha x_n - \alpha \bar{x})^2}{n} = \\ &= \alpha^2 \cdot \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \alpha^2 \cdot s_x^2. \quad \text{Άρα } s_y = |\alpha| \cdot s_x. \end{aligned}$$

Έχουμε λοιπόν τον παρακάτω πίνακα

n	$y_i = x_i + \beta$	$i = 1, 2, 3, \dots, n$ και $\alpha, \beta \in \mathbf{R}$	τότε	$\bar{y} = \bar{x} + \beta$	$s_y = s_x$
	$y_i = \alpha x_i$			$\bar{y} = \alpha \bar{x}$	$s_y = \alpha \cdot s_x$
	$y_i = \alpha x_i + \beta$			$\bar{y} = \alpha \bar{x} + \beta$	$s_y = \alpha \cdot s_x$

Ασκήσεις για λύση

1) Να συμπληρώσετε τον παρακάτω πίνακα συχνοτήτων – σχετικών συχνοτήτων.

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1		12				
2						
3		75		40		
4						87,5
5						
Σύνολο	120					

2) Να συμπληρώσετε τον παρακάτω πίνακα συχνοτήτων – σχετικών συχνοτήτων.

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1			0,12			
2			0,35			
3			0,28			
4			0,20			
5						
Σύνολο	200					

3) Να συμπληρώσετε τον παρακάτω πίνακα συχνοτήτων – σχετικών συχνοτήτων.

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	2					
2			0,25			
3					0,75	
4					0,95	
5						
Σύνολο	20					

4) Να συμπληρώσετε τον παρακάτω πίνακα συχνοτήτων – σχετικών συχνοτήτων.

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1						
2	4	6	0,20			
3					0,60	
4				25		
5	2					
6						
Σύνολο						

5) Να συμπληρώσετε τον παρακάτω πίνακα συχνοτήτων – σχετικών συχνοτήτων.

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1						
2				37,5		
3			0,20			80
4						
5	2	40				
Σύνολο						

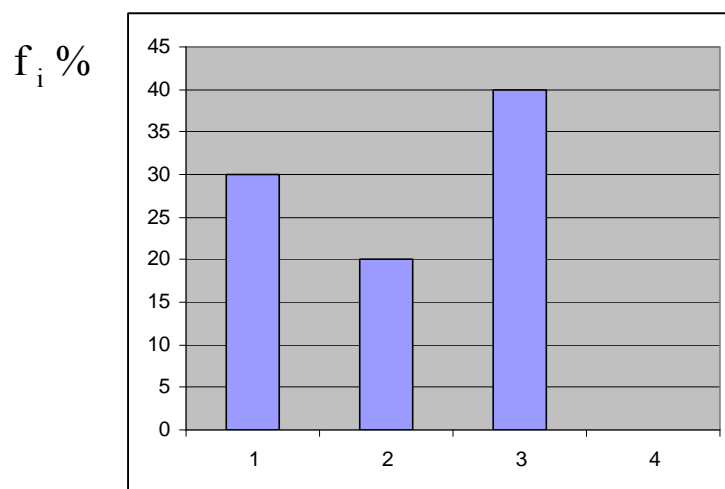
6) Μια μεταβλητή X παίρνει τις τιμές x_1, x_2, x_3, x_4 . Αν η συχνότητα v_1 της x_1 είναι διπλάσια της συχνότητας v_3 της τιμής x_3 , η σχετική συχνότητα f_2 της τιμής x_2 είναι διπλάσια της σχετικής συχνότητας f_1 της τιμής x_1 και η γωνία κυκλικού διαγράμματος της τιμής x_4 είναι 108° , τότε:

- i) Να υπολογίσετε τις f_1, f_2, f_3, f_4 .
- ii) Να κατασκευάσετε το κυκλικό διάγραμμα.

Παρατήρηση : Το κυκλικό διάγραμμα χρησιμοποιείται για την παράσταση τόσο των ποιοτικών όσο και των ποσοτικών μεταβλητών.

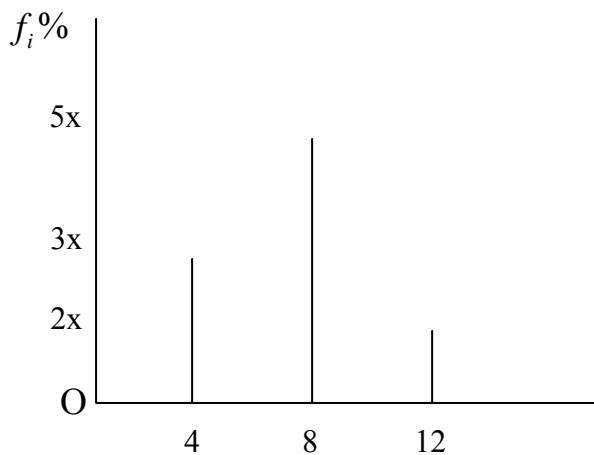
7) Το παρακάτω ραβδόγραμμα δίνει το ποσοστό τηλεθέασης 200 ατόμων, οι οποίοι παρακολουθούν τα κανάλια 1, 2, 3, 4.

- i) Να βρείτε το πλήθος των τηλεθεατών που παρακολουθεί το κανάλι 4.
- ii) Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- iii) Να μετατρέψετε το ραβδόγραμμα σε κυκλικό διάγραμμα σχετικών συχνοτήτων.



Παρατήρηση : Το ραβδόγραμμα χρησιμοποιείται για την παράσταση ΜΟΝΟΝ των ποιοτικών μεταβλητών.

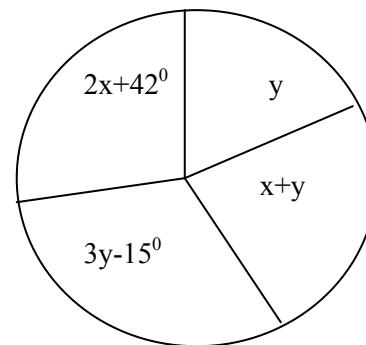
- 8) Η βαθμολογία μιας ομάδας φοιτητών σ' ένα διαγώνισμα είναι 4, 5, 6, 7 και 8. Το 80 % έχει βαθμό τουλάχιστον 5. Οι φοιτητές που έχουν βαθμό 4 είναι διπλάσιοι αυτών που έχουν βαθμό 8. Είκοσι ένας φοιτητές έχουν βαθμό κάτω από 6. Είκοσι τέσσερις φοιτητές έχουν βαθμό πάνω από 6 και το 55 % έχει βαθμό 6 ή 7. Να κατασκευάσετε τον πίνακα κατανομής συχνοτήτων : $n_i, N_i, f_i\%, F_i\%$.
- 9) Έστω x_1, x_2, x_3, x_4 οι τιμές μιας μεταβλητής X με αντίστοιχες σχετικές συχνότητες f_1, f_2, f_3, f_4 , οι οποίες δίνονται από τον τύπο $f_i = \frac{i^2}{5\kappa}$, $i = 1, 2, 3, 4$.
- Να αποδείξετε ότι $\kappa = 6$.
 - Αν η απόλυτη συχνότητα n_4 είναι 64, να βρείτε το πλήθος n .
 - Αν $n = 120$ να χαράξετε το κυκλικό διάγραμμα της παραπάνω κατανομής.
- 10) Το παρακάτω διάγραμμα σχετικών συχνοτήτων δίνει τα ποσοστά επιτυχίας ενός καλαθοσφαιριστή στις ελεύθερες βολές (πέτυχε 4), στα δίποντα (πέτυχε 8) και τρίποντα (πέτυχε 12) αντιστοίχως σε έναν αγώνα.



- Να βρείτε τα ποσοστά αυτά.
- Να κατασκευάσετε το κυκλικό διάγραμμα.

Παρατήρηση : Το διάγραμμα χρησιμοποιείται για την παράσταση ΜΟΝΟΝ των διακριτών ποσοτικών μεταβλητών.

- 11) Το διπλανό κυκλικό διάγραμμα αντιστοιχεί σε πλήθος 720 μαθητών. Αν είναι $x+y=81$
- i) Να αποδείξετε ότι $x=36^0$ και $y=45^0$.
- ii) Να βρείτε τις συχνότητες (απόλυτες) και να κατασκευάσετε το ραβδόγραμμα συχνοτήτων.



- 12) Μια μεταβλητή X παίρνει τις τιμές x_1, x_2, x_3 , με αντίστοιχες σχετικές συχνότητες f_1, f_2, f_3 , για τις οποίες ισχύει $f_1 = 3f_2$, ενώ η γωνία α_3 του κυκλικού διαγράμματος είναι 80^0 . Να κατασκευάσετε το κυκλικό διάγραμμα και να υπολογίσετε τις σχετικές συχνότητες των τιμών.
- 13) Μια μεταβλητή X παίρνει μόνο τις τιμές x_1, x_2, x_3 που σε κυκλικό διάγραμμα τα αντίστοιχα τόξα είναι $\alpha^0, \frac{3\alpha^0}{8}, \frac{7\alpha^0}{8}$. Να βρείτε τις αντίστοιχες σχετικές συχνότητες.

- 14) Στα σχολεία ενός Δήμου υπηρετούν 100 εκπαιδευτικοί. Ο συνολικός χρόνος υπηρεσίας των εκπαιδευτικών δίνεται από τον παρακάτω πίνακα:

A) Πόσοι εκπαιδευτικοί έχουν τουλάχιστον 15 χρόνια υπηρεσίας;

B) Με την προϋπόθεση ότι κάθε εκπαιδευτικός θα συνταξιοδοτηθεί, όταν συμπληρώσει 35 χρόνια:

α) Πόσοι εκπαιδευτικοί θα συνταξιοδοτηθούν μέσα στα επόμενα 12,5 χρόνια;

Να δικαιολογήσετε την απάντησή σας.

β) Πόσοι συνολικά εκπαιδευτικοί πρέπει να προσληφθούν μέσα στα επόμενα πέντε χρόνια, ώστε ο αριθμός των εκπαιδευτικών που υπηρετούν στα σχολεία του Δήμου να παραμείνει ο ίδιος;

Να δικαιολογήσετε την απάντησή σας.

Χρόνια υπηρεσίας [,)	Σχετική Συχνότητα f_i %
0 - 5	10
5 - 10	15
10 - 15	12
15 - 20	15
20 - 25	18
25 - 30	18
30 - 35	12

(Θέμα Πανελληνίων Εξετάσεων 2000)

- 15) Μια μεταβλητή X παίρνει τιμές $a, a+2, 2a+3, 4a-1$. Να βρείτε την τιμή του a , αν η μέση τιμή τους είναι 11.
- 16) Η μέση βαθμολογία ενός αθλητή ενόργανης γυμναστικής σε 6 αγωνίσματα είναι 9,2. Η μέση βαθμολογία των πέντε καλύτερων επιδόσεων είναι 9,5. Να βρείτε τη μικρότερη επίδοση του αθλητή.

- 17) Να βρείτε τους αριθμούς $x+y$, x , $x+8y$ και $x+3y$ με $x, y > 0$, αν αυτοί έχουν μέση τιμή 27 και διάμεσο 26.
- 18) Ο μέσος μηνιαίος μισθός v υπαλλήλων μιας εταιρείας είναι 1200 ευρώ. Αν σε κάθε υπάλληλο χορηγηθεί αύξηση 5% και επίδομα 50 ευρώ να βρείτε πόσος γίνεται ο μέσος μηνιαίος μισθός των υπαλλήλων;
- 19) Η μέση τιμή των ετήσιων αποδοχών των εργαζομένων σε μια επιχείρηση είναι 11800 €. Πως θα διαμορφωθεί η μέση τιμή στις παρακάτω περιπτώσεις;
- Σε όλους τους εργαζόμενους δίνεται ως δώρο το ποσό των 1000 €.
 - Σε όλους τους εργαζόμενους δίνεται αύξηση 5%
 - Σε όλους τους εργαζόμενους δίνεται αύξηση 5% και το ποσό των 1000 €.
- 20) Η κατανομή των 150 μαθητών ενός Λυκείου ως προς τις ώρες μελέτης τους ανά εβδομάδα έχει μέση τιμή $\bar{x} = 25$. Αν οι 40 μαθητές της Α' τάξης μελετούν κατά μέσο όρο 18 ώρες την εβδομάδα, ενώ οι 50 μαθητές της Β' κατά 22 ώρες, να βρείτε το μέσο χρόνο μελέτης των μαθητών της Γ' τάξης του Λυκείου.
- 21) Ο μέσος όρος στα Μαθηματικά των μαθητών μιας τάξης ενός Λυκείου είναι 14. Στην τάξη αυτή ήρθαν από άλλο σχολείο δύο μαθητές με βαθμούς, ο ένας 19 και ο άλλος 13. Ο νέος μέσος όρος είναι ίσος με 14,2. Να βρεθεί ο αρχικός αριθμός των μαθητών της τάξης.
- 22) Ο παρακάτω πίνακας δίνει τις εισπράξεις (σε δεκάδες χιλιάδες €) που έγιναν από τους αντιπροσώπους μιας εταιρείας κατά τη διάρκεια ενός έτους.

Κλάσεις [,)	0 – 2	2 – 4	4 – 6	6 – 8	8 – 10
Αριθμός Αντιπροσώπων	10	16	12	7	5

i) Να συμπληρώσετε τον πίνακα:

Κλάσεις	Κέντρο Κλάσης	v_i	f_i	$f_i \%$	F_i	$F_i \%$	$v_i x_i$

ii) Να υπολογίσετε τη μέση τιμή \bar{x} και τη διάμεσο.

- 23)** Ρωτήθηκαν 50 απόφοιτοι Πανεπιστημίου για το βαθμό του πτυχίου τους. Οκτώ από αυτούς δήλωσαν βαθμό κατά 2 μονάδες μεγαλύτερο του πραγματικού, δεκατέσσερις από αυτούς δήλωσαν βαθμό κατά μία μονάδα μεγαλύτερο του πραγματικού, ενώ οι υπόλοιποι δήλωσαν τον πραγματικό βαθμό. Αν με τους βαθμούς που δηλώθηκαν προκύπτει μέση τιμή $\bar{x} = 7,6$, τότε να βρείτε την πραγματική μέση τιμή.
- 24)** Εργολάβος οικοδομών απασχολεί 12 χτίστες, 5 εργάτες για τη μεταφορά των υλικών και 3 εργάτες ως χειριστές μηχανημάτων. Πληρώνει το ίδιο ημερομίσθιο για τους χτίστες και τους χειριστές μηχανημάτων, ενώ για τη μεταφορά των υλικών πληρώνει ημερομίσθιο κατά 10 ευρώ μικρότερο των προηγούμενων. Αν το μέσο ημερομίσθιο είναι 37,5 ευρώ, τότε να βρείτε:
- Ποιο είναι το ημερομίσθιο κάθε κατηγορίας εργαζομένων.
 - Πόσα € πληρώνει συνολικά ο εργολάβος κάθε ημέρα.
- 25)** Τέσσερις γυναίκες έχουν ηλικίες με διάμεσο και μέση τιμή 35 έτη. Αν η μεγαλύτερη απ' αυτές κρύψει το $\frac{1}{10}$ της ηλικίας της, τότε η μέση τιμή των ηλικιών τους γίνεται 34 έτη, ενώ αν επιπλέον και η τρίτη σε ηλικία κρύψει το $\frac{1}{9}$ της ηλικίας της, η μέση τιμή των ηλικιών τους γίνεται 33 έτη. Πόσο ετών είναι κάθε γυναίκα;
- 26)** Να υπολογίσετε το εύρος, τη διακύμανση και την τυπική απόκλιση των παρατηρήσεων
3, 4, 5, 5, 6, 7, 7, 8, 8, 8.
μιας ποσοτικής μεταβλητής X.
- 27)** Σε ένα δείγμα n υποψηφίων που διαγωνίστηκαν στο μάθημα της Έκθεσης βρέθηκαν οι βαθμολογίες, με άριστα το 20, που δίνονται στον παρακάτω πίνακα βαθμολογιών και αθροιστικών σχετικών συχνοτήτων F_i %.
- Να βρείτε τη μέση τιμή και τη διακύμανση της βαθμολογίας.
 - Να εξετάσετε αν το δείγμα είναι ομοιογενές.
 - Να εξετάσετε αν το δείγμα ακολουθεί την κανονική κατανομή.
 - Να βρείτε το πλήθος του δείγματος, αν γνωρίζουμε ότι 90 υποψήφιοι πήραν βαθμολογία 16-20, σύμφωνα με το διπλανό πίνακα. (Δίνεται $\sqrt{18,4} \approx 4,28$).

Βαθμολογίες [,)	F_i %
0 – 4	5
4 – 8	20
8 – 12	40
12 – 16	85
16 – 20	100

(Απάντηση : α) $\bar{x} = 12$, β) $s = 4,28$, γ) $CV = \frac{1}{3}$, δ) $n = 600$)

28) Το 50% των προϊόντων ενός θερμοκηπίου ζυγίζουν το πολύ 2.800 γραμμάρια, ενώ το 34% από 2.400 έως 2.800 γραμμάρια. Υποθέτουμε ότι η κατανομή του βάρους είναι από προσέγγιση κανονική.

- α) Να βρείτε το μέσο βάρος και την τυπική απόκλιση του βάρους των προϊόντων.
β) Να εξετάσετε αν το δείγμα είναι ομοιογενές. Αν όχι, πόσο τουλάχιστον πρέπει να μεταβληθεί το βάρος κάθε προϊόντος για να γίνει ομοιογενές.
γ) Αν τα προϊόντα που ζυγίζουν πάνω από 3,2 κιλά είναι 36, να βρείτε τον αριθμό των προϊόντων που παράγονται στο θερμοκήπιο.

(Απάντηση: α) $\bar{x} = 2800$, $S = 400$, β) πρέπει να αυξηθεί τουλάχιστον 1200 γρ., γ) 225 προϊόντα)

29) Ο παρακάτω πίνακας δείχνει την κατανομή της ημερήσιας αμοιβής των 80 εργατών μιας επιχείρησης.

- α) Να κατασκευάσετε τον πίνακα συχνοτήτων και αθροιστικών συχνοτήτων.
β) Να βρείτε το ποσοστό των εργατών που η αμοιβή τους είναι τουλάχιστον 8 χιλ. δρχ.
γ) Να υπολογίσετε τη διάμεσο και τη μέση τιμή.
δ) Η εταιρεία θέλει να προσλάβει ειδικευμένους εργάτες με ημερήσια αμοιβή μεγαλύτερη ή ίση των 12 χιλ. και μικρότερη των 14 χιλ. δρχ. Πόσους τέτοιους εργάτες πρέπει να προσλάβει ώστε το μέσο ημερομίσθιο να διαμορφωθεί στις 8 χιλ. δρχ.;

Κλάσεις [,)	v_i
0 – 2	5
2 – 4	10
4 – 6	5
6 – 8	20
8 – 10	25
10 – 12	15

(Απάντηση : α) F1% , β) 50% , δ = 8 , γ) $\bar{x} = 7375$ δρχ , 10 εργάτες)

6,25
8,75
25
50
81,25
100

30) Ένα προϊόν πωλείται σε 10 διαφορετικά καταστήματα στις παρακάτω τιμές σε €:

8 , 10 , 13 , 13 , 15 , 16 , 18 , 14 , 14 , 9.

- α) Να υπολογίσετε τη μέση τιμή και τη διάμεσο.
β) Να υπολογίσετε το εύρος, την τυπική απόκλιση και το συντελεστή μεταβολής.
γ) Αν οι τιμές του προϊόντος σε όλα τα καταστήματα υποστούν έκπτωση 10%, να εξετάσετε αν θα μεταβληθεί ο συντελεστής μεταβολής.

(Πανελλαδικές Εξετάσεις 2001-2002)

(Απάντηση : α) $\bar{x} = 13$ €, δ = 13,5 €, β) $R = 10$ €, $S^2 = \frac{1}{10} \sum_{i=1}^{10} (t_i - \bar{x})^2$)

- 31)** Σε μια έρευνα που έγινε στους μαθητές μιας πόλης για τον χρόνο που κάνουν να πάνε από το σπίτι στο σχολείο, διαπιστώθηκε ότι το 50% περίπου των μαθητών χρειάζεται περισσότερο από 12 min, ενώ το 16% περίπου των μαθητών χρειάζεται λιγότερο από 10 min. Υποθέτουμε επίσης ότι η κατανομή του χρόνου της διαδρομής είναι κατά προσέγγιση κανονική.
- α)** Να βρείτε το μέσο χρόνο διαδρομής των μαθητών και την τυπική απόκλιση του χρόνου διαδρομής τους.
- β)** Να εξετάσετε αν το δείγμα είναι ομοιογενές.
- γ)** Αν οι μαθητές αυτής της πόλης είναι 4.000, να βρείτε πόσοι από αυτούς θα κάνουν χρόνο διαδρομής από 14 έως 16 λεπτά.
- δ)** Μια μέρα λόγω έργων στον κεντρικό δρόμο της πόλης κάθε μαθητής καθυστέρησε για 5 λεπτά. Να βρείτε πόσο μεταβάλλεται ο συντελεστής μεταβολής CV.

(Απάντηση : α) $\bar{x} = \delta = 12$ λεπτά , $S = 2$, β) $CV = 16,67\%$ άρα δεν είναι ομοιογενής ,
γ) $13,5\%$, $n=540$, δ) $cv' = 11,8$ άρα ο συντελεστής μεταβολής μειώνεται κατά 4,9).

- 32)** Σε μια επιχείρηση οι μισθοί των 100 εργαζομένων κυμαίνονται από 200 έως 1200 €. Ειδικότερα γνωρίζουμε ότι:
- Κάτω από 400 € παίρνει το 10% των εργαζομένων.
 - Κάτω από 600 € παίρνουν οι μισοί εργαζόμενοι.
 - Από 400 € έως 800 € παίρνουν 70 εργαζόμενοι, ενώ
 - Κάτω από 1000 € παίρνουν 95 εργαζόμενοι.
- α)** Να βρείτε το μέσο μισθό των εργαζομένων.
- β)** Να βρείτε την τυπική απόκλιση των μισθών.
- γ)** Είναι οι μισθοί των εργαζομένων ομοιογενής μεταβλητή;

(Απάντηση : α) $\bar{x} = 630$ €, β) $S = \sqrt{4170} = 202,7$ €, γ) $CV = 0,32 = 32\%$ άρα δεν είναι ομοιογενής μεταβολή).

- 33)** Η μέση ημερήσια θερμοκρασία σε μια πόλη κατά το μήνα Σεπτέμβριο είναι 30° με τυπική απόκλιση 2° . Υποθέτοντας ότι έχουμε περίπου κανονική κατανομή, να βρείτε κατά προσέγγιση το ποσοστό των ημερών που έχουν θερμοκρασία:
- α)** Πάνω από 32° .
- β)** Κάτω από 26° .
- γ)** Τουλάχιστον 30° .
- δ)** Μεταξύ 26° και 32° .

(Απάντηση : α) 16% , β) 2,5% , γ) 50% , δ) 81,5%)

34) Εξετάσαμε ένα δείγμα μαθητών ως προς το βαθμό που πήραν στα Μαθηματικά Γενικής Παιδείας στις Πανελλήνιες Εξετάσεις και διαπιστώσαμε ότι:

Κάτω από 20 πήραν 5 μαθητές, κάτω από 40 πήραν 13 μαθητές, από 40 και πάνω πήρε το 48% των μαθητών, κάτω από 60 πήρε το 76% των μαθητών, ενώ από 80 και πάνω πήρε το 8% των μαθητών.

α) Να συμπληρώσετε το διπλανό πίνακα.

β) Να βρείτε το ποσοστό των μαθητών που πήρε βαθμό από 50 μέχρι και 70.

γ) Να βρείτε τη μέση τιμή της βαθμολογίας των μαθητών και την τυπική απόκλιση.

Είναι ομοιογενές το δείγμα;

Βαθμός [,)	v_i	$f_i\%$	N_i	$F_i\%$
0 – 20				
20 – 40				
60 – 80				
80–100				
Σύνολο				

35) Ο διπλανός πίνακας παρουσιάζει τα ορθογραφικά σφάλματα, που έκαναν οι μαθητές ενός τμήματος Θεωρητικής Κατεύθυνσης της Γ' Λυκείου κατά την υπαγόρευση ενός κειμένου.

Ο μέσος αριθμητικός είναι $\bar{x} = 2,63$, όπου X η μεταβλητή των λαθών.

Από τον πίνακα αυτόν έχουν σβηστεί οι σχετικές συχνότητες των τιμών x_2 και x_3 .

α) Να βρείτε τη διακύμανση των παρατηρήσεων.

β) Αν ο αριθμός των μαθητών του τμήματος είναι $n=25$, να βρείτε τη διάμεσο.

Αριθμός λαθών x_i	Σχετική Συχνότη. $f_i\%$	Αθροιστική Σχετ. Συχν. $F_i\%$
1	16	16
2		48
3		76
4	24	100
Σύνολο	100	

36. Στο διπλανό πίνακα παρουσιάζεται η χρηματική παροχή από τους γονείς, σε €, δείγματος έξι μαθητών της πρώτης τάξης (Ομάδα Α') και έξι μαθητών της δεύτερης τάξης (Ομάδα Β') ενός Γυμνασίου.

α) Να βρείτε τη μέση τιμή και τη διάμεσο των παρατηρήσεων κάθε ομάδας.

β) Να συγκρίνετε μεταξύ τους, ως προς την ομοιογένεια τις δύο ομάδες.

γ) Αν σε κάθε παρατήρηση της ομάδας Α' γίνει αύξηση 20 % και οι παρατηρήσεις της ομάδας Β' αυξηθούν κατά 5 € η κάθε μία, πώς διαμορφώνονται οι νέες μέσες τιμές των δύο ομάδων;

δ) Να συγκρίνετε μεταξύ τους, ως προς την ομοιογένεια τις δύο ομάδες με τα νέα δεδομένα.

Ομάδα Α'	Ομάδα Β'
1	7
8	14
9	6
5	4
3	12
4	5

(Θέμα Πανελληνίων Εξετάσεων 2003)

37) Η τυπική απόκλιση και ο συντελεστής μεταβολής από την εξέταση ενός τμήματος 20 φοιτητών στο μάθημα της Στατιστικής είναι αντίστοιχα $S_1 = 1$ και $CV_1 = 0,20$.

Η τυπική απόκλιση και ο συντελεστής μεταβολής από την εξέταση ενός δεύτερου τμήματος 12 φοιτητών στο ίδιο μάθημα είναι αντίστοιχα $S_2 = 3$ και $CV_2 = 0,60$.

Να βρείτε:

α) Τις μέσες τιμές \bar{x}_1 του πρώτου τμήματος και \bar{x}_2 του δεύτερου τμήματος.

β) Τη μέση τιμή της επίδοσης όλων των φοιτητών.

γ) Το συντελεστή μεταβολής CV όλων των φοιτητών.

δ) Πόσο πρέπει να αυξηθεί η βαθμολογία κάθε φοιτητή, ώστε το δείγμα να γίνει ομοιογενές.

Απάντηση : **α)** $\bar{x}_1 = 5$, $\bar{x}_2 = 5$ **β)** $\bar{x} = \frac{20 \cdot 5 + 12 \cdot 5}{32} = 5$.

$$\gamma) s_1^2 = \frac{1}{20} \sum_{i=1}^{20} (t_i - \bar{x}_1)^2 \Rightarrow \dots \sum_{i=1}^{20} (t_i - \bar{x}_1)^2 = 20 \quad , \quad s_2^2 = \dots \sum_{i=1}^{12} (t_i - \bar{x}_2)^2 = 108 \quad ,$$

$$s^2 = \frac{1}{32} \sum_{i=1}^{32} (t_i - \bar{x})^2 = \frac{20 + 108}{32} = 4 \Rightarrow s = 2 \quad , \quad CV = 0,40 = 40\%$$

δ) $x \geq 15$ απορρίπτεται γιατί $0 \leq x \leq 10$.

38) Δίνεται ο παρακάτω πίνακας σχετικών συχνοτήτων.

x_i	0	1	2	3	4	5
f_i	0,08	0,30	0,30	f_4	f_5	0,02

Αν είναι $\bar{x} = 2$, τότε:

α) Να υπολογίσετε τις f_4 , f_5 .

β) Αν η συνάρτηση $f(x) = 5f_4 x^3 + 40f_5 x^2 + \kappa x$, $\kappa \in R$, έχει ακριβώς ένα μέγιστο και ένα ελάχιστο, να βρείτε τη μέγιστη ακέραια τιμή του κ .

Απάντηση : **α)** $f_4 = 0,20$, $f_5 = 0,10$ **β)** $\kappa = 5$

Απαντήσεις – Υποδείξεις των ασκήσεων

1)

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	12	12	0,10	10	0,10	10
2	15	27	0,125	12,5	0,225	22,5
3	48	75	0,40	40	0,625	62,5
4	30	105	0,25	25	0,875	87,5
5	15	120	0,125	12,5	1	100
Σύνολο	120		1	100		

2)

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	24	24	0,12	12	0,12	12
2	70	94	0,35	35	0,47	47
3	56	150	0,28	28	0,75	75
4	40	190	0,20	20	0,95	95
5	10	200	0,05	5	1	100
Σύνολο	200		1	100		

3)

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	2	2	0,10	10	0,10	10
2	5	7	0,25	25	0,35	35
3	8	15	0,40	40	0,75	75
4	4	19	0,20	20	0,95	95
5	1	20	0,05	5	1	100
Σύνολο	20		1	100		

4)

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	2	2	0,10	10	0,10	10
2	4	6	0,20	20	0,30	30
3	6	12	0,30	30	0,60	60
4	5	17	0,25	25	0,85	85
5	2	19	0,10	10	0,95	95
6	1	20	0,05	5	1	100
Σύνολο	20		1	100		

5)

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	9	9	0,225	22,5	0,225	22,5
2	15	24	0,375	37,5	0,60	60
3	8	32	0,20	20	0,80	80
4	6	38	0,15	15	0,95	95
5	2	40	0,05	5	0,05	100
Σύνολο	40		1	100		

6) i) $f_1 = 0,2$, $f_2 = 0,4$, $f_3 = 0,1$, $f_4 = 0,3$

ii) $a_1 = 72^0$, $a_2 = 144^0$, $a_3 = 36^0$.

7) i) 10% , δηλαδή 20 άτομα. ii)

iii) $a_1 = 108^0$, $a_2 = 72^0$,
 $a_3 = 144^0$, $a_4 = 36^0$.

x_i	v_i	N_i	f_i	$f_i\%$	F_i	$F_i\%$
1	60	60	0,30	30	0,30	30
2	40	100	0,20	20	0,50	50
3	80	180	0,40	40	0,90	90
4	20	200	0,10	10	1	100
Σύνολο	200		1	100		

8)

x_i	v_i	N_i	$f_i\%$	$F_i\%$
4	12	12	20	20
5	9	21	15	35
6	15	36	25	60
7	18	54	30	90
8	6	60	10	100
Σύνολο	60		100	

9)

f_i	α_i
$\frac{1}{30}$	12^0
$\frac{4}{30}$	48^0
$\frac{9}{30}$	108^0
$\frac{16}{30}$	192^0

10) i) $2x + 3x + 5x = 100$, άρα $x = 10$. Επομένως $f_1\% = 30$, $f_2\% = 50$, $f_3\% = 20$

ii) $\phi_1 = \frac{30}{100}360 = 108^\circ$, $\phi_2 = 180^\circ$, $\phi_3 = 72^\circ$.

11) ii) Χρησιμοποιούμε τον τύπο $v_i = \frac{a_i}{360}v$. Βρίσκουμε $v_1 = 162$, $v_2 = 228$, $v_3 = 240$, $v_4 = 90$.

12) $\alpha_1 = 210^\circ$, $\alpha_2 = 70^\circ$, $\alpha_3 = 80^\circ$. $f_1 \cong 58,33\%$, $f_2 \cong 19,45\%$, $f_3 \cong 22,22\%$

13) $f_1 = \frac{4}{9}$, $f_2 = \frac{1}{6}$, $f_3 = \frac{7}{18}$.

14) A) $15+18+18+12=63$, B) α) $\frac{18}{2}+18+12=39$, β) 12

15) $\alpha=5$

16) 7,7

17) $x = 24$, $y = 1$

18) 1310 €

19) α) 12800 € β) 12390 € γ) 13390 €

20) $25 = \frac{40\bar{x}_A + 50\bar{x}_B + 60\bar{x}_\Gamma}{150} \Rightarrow \bar{x}_\Gamma = \frac{3750 - 1820}{60} \dots$

21) $v = 18$.

22) i)

Κλάσεις	Κέντρο Κλάσης	v_i	f_i	$f_i\%$	F_i	$F_i\%$	$v_i x_i$
[0, 2)	1	10	0,20	20	0,20	20	10
[2, 4)	3	16	0,32	32	0,52	52	48
[4, 6)	5	12	0,24	24	0,76	76	60
[6, 8)	7	7	0,14	14	0,90	90	49
[8,10)	9	5	0,10	10	1,00	100	45
Σύνολο		50	1	100			212

ii) $\bar{x} = \frac{1}{v} \sum_{i=1}^5 v_i x_i = \frac{1}{50} 212 = 4,24$, $\delta = 3,8750$

23) $\bar{y} = 7$.

24) i) 40 € και 30 € αντίστοιχα ii) 750 €

25) $x_1 = 30$, $x_2 = 34$, $x_3 = 36$, $x_4 = 40$

26) $R = 5$, $s^2 = 2,89$, $s = 1,7$