

**ΜΑΘΗΜΑΤΙΚΑ
ΚΑΙ
ΣΤΟΙΧΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ**

Γ' ΓΕΝΙΚΟΥ ΛΥΚΕΙΟΥ

Τόμος 2ος

ΣΤΟΙΧΕΙΑ ΑΡΧΙΚΗΣ ΕΚΔΟΣΗΣ

ΣΥΓΓΡΑΦΕΙΣ:

Αδαμόπουλος Λεωνίδας
Επ. Σύμβουλος Παιδαγωγικού Ιν-
στιτούτου

Δαμιανού Χαράλαμπος
Αναπλ. Καθηγητής Παν/μίου Αθη-
νών

Σβέρκος Ανδρέας
Σχολικός Σύμβουλος

ΚΡΙΤΕΣ:

Κουνιάς Στρατής
Καθηγητής Παν/μίου Αθηνών

Μακρής Κωνσταντίνος
Σχολικός Σύμβουλος

Τσικαλουδάκης Γεώργιος
Καθηγητής Β/θμιας Εκπαίδευσης

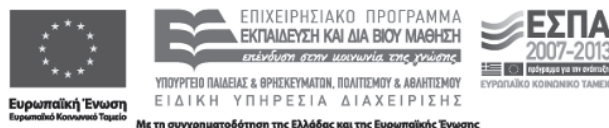
Γλωσσική Επιμέλεια:
Μπουσούνη Λία
Καθηγήτρια Β/θμιας Εκπαίδευσης

Δακτυλογράφηση:
Μπολιώτη Πόπη

Σχήματα:
Μπούτσικας Μιχάλης

ΣΤΟΙΧΕΙΑ ΕΠΑΝΕΚΔΟΣΗΣ

Η επανέκδοση του παρόντος βιβλίου πραγματοποιήθηκε από το Ινστιτούτο Τεχνολογίας Υπολογιστών & Εκδόσεων «Διόφαντος» μέσω ψηφιακής μακέτας, η οποία δημιουργήθηκε με χρηματοδότηση από το ΕΣΠΑ / ΕΠ «Εκπαίδευση & Διά Βίου Μάθηση» / Πράξη «ΣΤΗΡΙΖΩ».



Οι διορθώσεις πραγματοποιήθηκαν κατόπιν έγκρισης του Δ.Σ. του Ινστιτούτου Εκπαιδευτικής Πολιτικής

Η αξιολόγηση, η κρίση των προσαρμογών και η επιστημονική επιμέλεια του προσαρμοσμένου βιβλίου πραγματοποιείται από τη Μονάδα Ειδικής Αγωγής του Ινστιτούτου Εκπαιδευτικής Πολιτικής.

Η προσαρμογή του βιβλίου για μαθητές με μειωμένη όραση από το ΙΤΥΕ – ΔΙΟΦΑΝΤΟΣ πραγματοποιείται με βάση τις προδιαγραφές που έχουν αναπτυχθεί από ειδικούς εμπειρογνώμονες για το ΙΕΠ.

**ΠΡΟΣΑΡΜΟΓΗ ΤΟΥ ΒΙΒΛΙΟΥ
ΓΙΑ ΜΑΘΗΤΕΣ
ΜΕ ΜΕΙΩΜΕΝΗ ΟΡΑΣΗ**

ΙΤΥΕ - ΔΙΟΦΑΝΤΟΣ

**ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΕΡΕΥΝΑΣ
ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ**

**ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ
ΠΟΛΙΤΙΚΗΣ**

**ΑΔΑΜΟΠΟΥΛΟΣ ΛΕΩΝΙΔΑΣ
ΔΑΜΙΑΝΟΥ ΧΑΡΑΛΑΜΠΟΣ
ΣΒΕΡΚΟΣ ΑΝΔΡΕΑΣ**

**Η συγγραφή και η επιστημονική
επιμέλεια του βιβλίου
πραγματοποιήθηκε υπό την αιγίδα
του Παιδαγωγικού Ινστιτούτου**

**ΜΑΘΗΜΑΤΙΚΑ
ΚΑΙ
ΣΤΟΙΧΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ**

Γ' ΓΕΝΙΚΟΥ ΛΥΚΕΙΟΥ

Τόμος 2ος

Ι.Τ.Υ.Ε. «ΔΙΟΦΑΝΤΟΣ»

2 ΣΤΑΤΙΣΤΙΚΗ

Εισαγωγή

Ο όρος “Στατιστική” ενδεχομένως να προέρχεται από τη λατινική λέξη “status” (πολιτεία, κράτος) η οποία, χρησιμοποιήθηκε αρχικά για το χαρακτηρισμό αριθμητικών δεδομένων που αναφέρονται κυρίως στον πληθυσμό μιας χώρας. Μπορεί όμως να προέρχεται από την αρχαία ελληνική λέξη στατίζω (τοποθετώ, ταξινομώ, συμπεραίνω). Με την εμφάνιση της Στατιστικής και στα πρώτα στάδια της ανάπτυξής της οι άνθρωποι την ταύτισαν

με την παράθεση τεράστιων πινάκων με δεδομένα σχετικά με τους θανάτους, τις γεννήσεις, τους φόρους, τα προϊόντα, τους άνδρες σε στρατεύσιμη ηλικία κτλ., προσπαθώντας έτσι να περιγράψουν διάφορα δημογραφικά, οικονομικά και πολιτικά φαινόμενα. Η αρχαιότερη ίσως συλλογή στατιστικών στοιχείων θεωρείται η απογραφή πληθυσμού που έγινε το 2238 π.Χ. στην Κίνα από τον αυτοκράτορα Υαο. Επίσης, στοιχειώδεις απογραφές φαίνεται να έχουν πραγματοποιηθεί από τους Σίνες, τους Αιγυπτίους και τους Πέρσες. Ο όρος Στατιστική αναφέρεται επίσης και από το Σωκράτη (Ξενοφώντας “Απομνημονεύματα”) και από τον Αριστοτέλη (“Πολιτεία”). Όπως γνωρίζουμε

απογραφή πληθυσμού είχε επίσης διαταχθεί και από τον καίσαρα Αύγουστο στην περίοδο της γέννησης του Χριστού.

Στην αρχαιότητα, η συγκέντρωση στατιστικών στοιχείων είχε στόχο τον εντοπισμό των πολιτών που είχαν υποχρέωση να υπηρετήσουν ως πολεμιστές ή να πληρώσουν φόρο. Συστηματική συλλογή δεδομένων για τον πληθυσμό και την οικονομία άρχισε κατά τη διάρκεια της Αναγέννησης στις πόλεις Βενετία και Φλωρεντία στην Ιταλία, και γρήγορα επεκτάθηκε και σε άλλες χώρες της Δυτικής Ευρώπης. Ο μεγάλος ρυθμός θνησιμότητας στην Ευρώπη οφειλόταν στις επιδημικές ασθένειες, στους πολέμους και στις

λιμοκτονίες. Στις αρχικές καταγραφές των θανάτων από την πανώλη, τη φοβερή ασθένεια που εμφανίστηκε το 1348 και κράτησε πάνω από 400 χρόνια, προστέθηκαν στη συνέχεια και οι θάνατοι από άλλες αιτίες. Στα 1620 ο Άγγλος εμπορευόμενος Graunt από δειγματοληπτική έρευνα που έκανε σε οικογένειες του Λονδίνου βρήκε ότι σε κάθε 88 άτομα υπήρχαν 3 θάνατοι. Χρησιμοποιώντας τους καταλόγους του Λονδίνου, που έδιναν 13.200 θανάτους το 1620, εκτίμησε τον πληθυσμό του Λονδίνου το έτος αυτό στα 387.200 άτομα.

Μια πραγματικά σπουδαία στατιστική απογραφή στην εποχή του Γουλιέλμου του Κατακτητή, στο τέλος του 11ου αιώνα, αναφέρεται

σε διάφορες μονάδες παραγωγής της Αγγλίας όπως μεταλλεία, ιχθυοτροφεία κ.ά. Από το 16ο έως το 19ο αιώνα, η ραγδαία ανάπτυξη του εμπορίου ώθησε τις πολιτειακές αρχές στη μελέτη οικονομικών δεδομένων, όπως είναι το εξαγωγικό εμπόριο, το πλήθος και η δυναμικότητα των βιομηχανιών κτλ.

Ενώ παλαιότερα η Στατιστική ασχολείτο μόνο με την παράθεση τεράστιων πινάκων με δεδομένα και αναρίθμητων διαγραμμάτων, σήμερα μπορούμε να διακρίνουμε σε μια στατιστική έρευνα τρία στάδια: Τη συλλογή του στατιστικού υλικού, την επεξεργασία και παρουσίασή του και τέλος την ανάλυση αυτού του υλικού και την εξαγωγή χρήσιμων συμπερασμάτων. Τα τρία

αυτά στάδια επιτυγχάνονται με την εφαρμογή καταλλήλων για κάθε περίπτωση στατιστικών μεθόδων, όπως και με τη βοήθεια των Υπολογιστών, οι οποίοι σημείωσαν τεράστια ανάπτυξη στις μέρες μας. Συμπερασματικά λοιπόν μπορούμε να δώσουμε ως ορισμό της “Στατιστικής” το 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 (experimental design) ενώ, με τον δεύτερο ασχολείται η περιγραφική στατιστική (descriptive statistics), που αποτελεί και το αντικείμενο μελέτης μας στη συνέχεια. Τέλος, η επαγωγική στατιστική ή στατιστική συμπερασματολογία (inferential statistics) 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. Έτσι αν, για παράδειγμα, ο Διευθυντής ενός σχολείου εξετάζοντας ένα δείγμα 100 απουσιών των μαθητών από το σύνολο των απουσιών

ενός τριμήνου αναφέρει στο σύλλογο των καθηγητών ότι 20 από τις 100 απουσίες είναι αδικαιολόγητες, τότε απλώς περιγράφει αυτό που παρατήρησε. Αν όμως αναφέρει ότι το 20% των απουσιών είναι αδικαιολόγητες, τότε συμπεραίνει ότι το ποσοστό των απουσιών όλων των μαθητών του σχολείου είναι (περίπου) το ίδιο με αυτό του δείγματος. Προβαίνει δηλαδή σε μια επαγωγή από το δείγμα στον πληθυσμό.

Η Στατιστική σήμερα χρησιμοποιείται ευρύτατα σε όλους σχεδόν τους τομείς της ανθρώπινης δραστηριότητας. Βασικές έννοιες της Στατιστικής έχουν εισχωρήσει και ενσωματωθεί σε όλες τις επιστήμες. Από τις Ανθρωπιστικές, Νομικές και Κοινωνικές Επιστήμες (Αρχαιολογία,

Λαογραφία, Κοινωνιολογία, Δημογραφία, ...), τις Φυσικές Επιστήμες (Φυσική, Χημεία, Αστρονομία, ...), τις Επιστήμες Υγείας (Ιατρική, Φαρμακευτική, Βιολογία, ...), τις Τεχνολογικές Επιστήμες (Μηχανολογία, Τοπογραφία, Ναυπηγική, ...) μέχρι τις Επιστήμες Οικονομίας και Διοίκησης (Οικονομικά, Χρηματιστηριακά, Διαφήμιση, Marketing, ...), βλέπουμε να υπεισέρχεται η Στατιστική είτε με την αρχική περιγραφική μορφή της είτε με τις προηγμένες αναλυτικές τεχνικές της.

Η ανάλυση στατιστικών ερευνών είναι το κυριότερο εργαλείο έρευνας σε ένα μεγάλο φάσμα εφαρμογών των παραπάνω επιστημών. Οι έρευνες των ανθρώπινων πληθυσμών (συχνά αναφερόμενες και ως

δημοσκοπήσεις) αποτελούν σπουδαίες πηγές βασικής γνώσης των κοινωνικών επιστημών. Οικονομολόγοι, ψυχολόγοι, κοινωνιολόγοι και πολιτικοί επιστήμονες μελετούν ποικίλα θέματα όπως πρότυπα εσόδων-εξόδων των οικογενειών και των επιχειρήσεων, την επίδραση της επαγγελματικής απασχόλησης των γυναικών στην οικογενειακή ζωή, τις συγκοινωνιακές και ταξιδιωτικές συνήθειες των κατοίκων μιας πόλης, τις προτιμήσεις των ψηφοφόρων για τους υποψηφίους και τις θέσεις τους. Πολλά προβλήματα που αντιμετωπίζουν σήμερα οι επιχειρήσεις αφορούν τη διατήρηση, αντικατάσταση ή το κρίσιμο σημείο αντοχής συσκευών ή προσωπικού. Ο διευθυντής μιας βιομηχανίας πρέπει να είναι σε θέση να

κατανοεί στατιστικές έρευνες που αφορούν την ποιότητα του προϊόντος και την αποδοτικότητα της παραγωγικής διαδικασίας. Πρέπει επίσης να αντιλαμβάνεται την αποτελεσματικότητα της διαφήμισης και τις προτιμήσεις του καταναλωτή σε μια έρευνα αγοράς. Συμβουλευόμενος και τον στατιστικό μπορεί να πάρει σωστές αποφάσεις αναφορικά με την επέκταση ή μη της επιχείρησης. Σήμερα κάθε γιατρός πρέπει να έχει βασικές γνώσεις Στατιστικής που θα τον βοηθήσουν τόσο στην έρευνα όσο και στην καθημερινή άσκηση του κάθε μορφής και είδους ιατρικού ή βιοϊατρικού, γενικότερα, επαγγέλματος. Η Εθνική Στατιστική Υπηρεσία κάθε χώρας διενεργεί σε τακτά χρονικά διαστήματα δειγματοληπτικές έρευνες, για

να πάρει πληροφορίες για τον πληθωρισμό, την απασχόληση και την ανεργία στη χώρα. Ανάλογα με τα αποτελέσματα διαμορφώνεται και η κυβερνητική πολιτική στα θέματα αυτά.

Πέρα από όλα αυτά, διαπιστώνουμε ολοένα και περισσότερο να γίνεται χρήση μεθόδων της Στατιστικής για την υποστήριξη διάφορων θέσεων. Ακόμα και σε τηλεοπτικές αντιπαραθέσεις (κυρίως σε προεκλογικές περιόδους) βλέπουμε τους συνομιλητές να κάνουν χρήση αριθμών, στατιστικών στοιχείων, γραφημάτων και διαγραμμάτων, για να δώσουν εγκυρότητα στις απόψεις τους και να πείσουν για τα λεγόμενά τους.

Παραπάνω έχουν αναφερθεί ελάχιστα από τα πεδία εφαρμογών της Στατιστικής. Προφανώς μια λεπτομερής περιγραφή όλων των εφαρμογών δεν είναι δυνατή. Η μελέτη όμως και η γνώση της Στατιστικής βοηθά όχι μόνο στη σωστή χρήση των γνωστών μεθόδων αλλά και στην ανάπτυξη νέων τεχνικών για την αποτελεσματικότερη εξαγωγή χρήσιμων συμπερασμάτων.

2.1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Πληθυσμός - Μεταβλητές

Όπως αναφέρθηκε και προηγουμένως, αυτό που μας ενδιαφέρει είναι να εξετάσουμε τα στοιχεία ενός συνόλου ως προς ένα ή περισσότερα χαρακτηριστικά τους. Αυτό συμβαίνει, για παράδειγμα, όταν ενδιαφερόμαστε για:

- α) τις προτιμήσεις των ψηφοφόρων εν όψει των προσεχών εκλογών
- β) τον αριθμό των υπαλλήλων μιας επιχείρησης
- γ) το ύψος, το βάρος, την ομάδα αίματος και το φύλο των μαθητών της Γ΄ τάξης Λυκείου

**δ) ΤΙΣ ΣΥΝΕΠΕΙΕΣ ΤΟΥ ΚΑΠΝΙΣΜΑΤΟΣ
ΣΤΗΝ ΥΓΕΙΑ ΤΩΝ ΚΑΠΝΙΣΤΩΝ ΚΤΛ.**

Σε καθένα από τα παραδείγματα αυτά έχουμε ένα σύνολο και θέλουμε να εξετάσουμε τα στοιχεία του ως προς ένα ή περισσότερα χαρακτηριστικά τους. Ένα τέτοιο σύνολο λέγεται πληθυσμός (population). Τα στοιχεία του πληθυσμού συχνά αναφέρονται και ως μονάδες ή άτομα του πληθυσμού. Στο πρώτο παράδειγμα έχουμε το σύνολο των ψηφοφόρων και μας ενδιαφέρει η προτίμησή τους, ποιο “κόμμα” π.χ. υποστηρίζουν. Στο τρίτο παράδειγμα έχουμε το σύνολο των μαθητών της Γ΄ Λυκείου και μας ενδιαφέρουν τα τέσσερα χαρακτηριστικά τους: ύψος, βάρος, ομάδα αίματος και φύλο.

Τα χαρακτηριστικά ως προς τα οποία εξετάζουμε έναν πληθυσμό λέγονται μεταβλητές (variables) και τις συμβολίζουμε συνήθως με τα κεφαλαία γράμματα X, Y, Z, B, \dots . Οι δυνατές τιμές που μπορεί να πάρει μια μεταβλητή λέγονται τιμές της μεταβλητής. Από τη διαδοχική εξέταση των ατόμων του πληθυσμού ως προς ένα χαρακτηριστικό τους προκύπτει μια σειρά από δεδομένα, που λέγονται στατιστικά δεδομένα ή παρατηρήσεις. Τα στατιστικά δεδομένα δεν είναι κατ' ανάγκη διαφορετικά. Για παράδειγμα, αν εξετάζουμε την ομάδα αίματος δέκα ατόμων, τα στατιστικά δεδομένα ή παρατηρήσεις που θα προκύψουν μπορεί να είναι: A, B, A, AB, O, AB, AB, AB, O, B. Οι δυνατές όμως τιμές

που μπορεί να πάρει η μεταβλητή “ομάδα αίματος” είναι οι εξής τέσσερις: A, B, AB και O.

Τις μεταβλητές τις διακρίνουμε:

1. Σε ποιοτικές ή κατηγορικές μεταβλητές, των οποίων οι τιμές τους δεν είναι αριθμοί. Τέτοιες είναι, για παράδειγμα, η ομάδα αίματος (με τιμές A, B, AB, O), το φύλο (με τιμές αγόρι, κορίτσι), οι συνέπειες του καπνίσματος (με τιμές καρδιακά νοσήματα, καρκίνος κτλ), όπως επίσης και η οικονομική κατάσταση και η υγεία των ανθρώπων (που μπορεί να χαρακτηριστεί ως κακή, μέτρια, καλή ή πολύ καλή), καθώς και το ενδιαφέρον των μαθητών για τη Στατιστική, που μπορεί να χαρακτηριστεί ως υψηλό, μέτριο, χαμηλό ή μηδαμινό.

2. Σε ποσοτικές μεταβλητές, των οποίων οι τιμές είναι αριθμοί και διακρίνονται:

- i) Σε διακριτές μεταβλητές, που παίρνουν μόνο “μεμονωμένες” τιμές. Τέτοιες μεταβλητές είναι, για παράδειγμα, ο αριθμός των υπαλλήλων μιας επιχείρησης (με τιμές 1, 2, ...), το αποτέλεσμα της ρίψης ενός ζαριού (με τιμές 1, 2, ..., 6) κτλ.**
- ii) Σε συνεχείς μεταβλητές, που μπορούν να πάρουν οποιαδήποτε τιμή ενός διαστήματος πραγματικών αριθμών (α , β). Τέτοιες μεταβλητές είναι το ύψος και το βάρος των μαθητών της Γ΄ Λυκείου, ο χρόνος που χρειάζονται οι μαθητές να απαντήσουν στα θέματα μιας**

εξέτασης, η διάρκεια μιας τηλεφωνικής συνδιάλεξης κτλ.

Συλλογή Στατιστικών Δεδομένων

Ένας τρόπος για να πάρουμε τις απαραίτητες πληροφορίες που χρειαζόμαστε για κάποιο πληθυσμό είναι να εξετάσουμε όλα τα άτομα (στοιχεία) του πληθυσμού ως προς το χαρακτηριστικό που μας ενδιαφέρει. Η μέθοδος αυτή συλλογής των δεδομένων καλείται απογραφή (census). Για παράδειγμα, η Στατιστική Υπηρεσία της χώρας μας (ΕΣΥΕ) κάνει κάθε 10 χρόνια απογραφή του πληθυσμού, η οποία αποτελεί κύρια πηγή δεδομένων

δημογραφικού, οικονομικού, εμπορικού και βιομηχανικού χαρακτήρα. Η τελευταία απογραφή έγινε το 1991.

Σε πολλές όμως περιπτώσεις η εξέταση όλων των μονάδων του πληθυσμού είναι δύσκολη ή ακόμα και αδύνατη. Ένας υποψήφιος βουλευτής, για παράδειγμα, πριν από τις εκλογές είναι δύσκολο να εξετάσει όλους τους ψηφοφόρους, για να προσδιορίσει τι αντίληψη έχουν για τις θέσεις του. Επίσης ο κόπος, ο χρόνος και τα έξοδα που χρειάζονται για τη διεξαγωγή μιας απογραφής είναι πολλές φορές αρκετά μεγάλα, ιδίως όταν ο πληθυσμός που εξετάζεται είναι αρκετά μεγάλος. Εξάλλου ένας κατασκευαστής εκρηκτικών μηχανισμών ή ηλεκτρικών

λυχνιών είναι αδύνατο να δοκιμάζει όλους τους παραγόμενους μηχανισμούς, για να ελέγχει την αποτελεσματικότητά τους, ή όλες τις παραγόμενες λυχνίες για να ελέγχει το χρόνο ζωής τους. Ομοίως ο γιατρός για να υπολογίσει την αποτελεσματικότητα ενός νέου φαρμάκου στην καταπολέμηση μιας ασθένειας είναι αδύνατο να περιμένει να δοκιμαστεί το φάρμακο σε όλα τα άτομα που πάσχουν από τη συγκεκριμένη ασθένεια. Όπου λοιπόν η απογραφή είναι δύσκολη, αδύνατη ή οικονομικά και χρονικά ασύμφορη, ο ερευνητής μαζεύει πληροφορίες από κάποια μικρή ομάδα ή υποσύνολο του πληθυσμού, το οποίο καλείται δείγμα. Κάνει τις παρατηρήσεις του στο δείγμα αυτό και

μετά γενικεύει τα συμπεράσματά του για ολόκληρο τον πληθυσμό. Τα συμπεράσματα όμως που θα προκύψουν από τη μελέτη του δείγματος θα είναι αξιόπιστα, θα ισχύουν δηλαδή με ικανοποιητική ακρίβεια για ολόκληρο τον πληθυσμό, αν η επιλογή του δείγματος γίνει με σωστό τρόπο, ώστε το δείγμα να είναι, όπως λέμε, αντιπροσωπευτικό του πληθυσμού. Στην πράξη, ένα δείγμα θεωρείται αντιπροσωπευτικό ενός πληθυσμού, εάν έχει επιλεγεί κατά τέτοιο τρόπο, ώστε κάθε μονάδα του πληθυσμού να έχει την ίδια δυνατότητα να επιλεγεί. Η επιλογή του αντιπροσωπευτικού δείγματος είναι “εκ των ων ουκ άνευ”. Αποτελεί πολύ σοβαρή και

δύσκολη διαδικασία. Ο κακός σχεδιασμός και η εκτέλεση της στατιστικής έρευνας, η μη αντιπροσωπευτικότητα του δείγματος, ο μη σωστός καθορισμός του μεγέθους του δείγματος αποτελούν μερικά βασικά μειονεκτήματα στη διαδικασία επιλογής ενός δείγματος. Από την άλλη πλευρά, στις απογραφές απαιτείται συνήθως μεγάλος αριθμός απογραφέντων. Παρουσιάζεται έτσι η ανάγκη πρόσληψης και εκπαίδευσης μεγάλου αριθμού υπαλλήλων. Λόγω του μεγάλου χρόνου και κυρίως των σημαντικών εξόδων που απαιτούνται, πολλές φορές χρησιμοποιούνται ανεπαρκώς εκπαιδευμένοι απογραφείς με κίνδυνο να σημειώνονται λάθη οφειλόμενα σ' αυτούς.

Αξίζει να σημειωθεί ότι μία “προσεκτική” επιλογή μικρότερου δείγματος είναι δυνατόν να δώσει καλύτερα αποτελέσματα από ένα μεγαλύτερο δείγμα που δεν έχει εκλεγεί κατάλληλα. Ενδεικτικό είναι το παράδειγμα των προεδρικών εκλογών των ΗΠΑ το 1936. Το περιοδικό Literary Digest χρησιμοποιώντας δείγμα 2.400.000 ατόμων πρόβλεψε νίκη του Landon με ποσοστό 57%. Αντίθετα, το δημοσκοπικό γραφείο του G. Gallup χρησιμοποιώντας δείγμα 50.000 ατόμων πρόβλεψε το σωστό αποτέλεσμα που ήταν νίκη του Roosevelt με ποσοστό 62%! Η παταγώδης αποτυχία της δημοσκόπησης του περιοδικού οφειλόταν στο γεγονός ότι το δείγμα που επελέγη δεν ήταν αντιπροσωπευτικό του πληθυσμού.

Οι αρχές και οι μέθοδοι για τη συλλογή και ανάλυση δεδομένων από πεπερασμένους πληθυσμούς είναι το αντικείμενο της Δειγματοληψίας (Sampling), που αποτελεί τη βάση της Στατιστικής. Γενικά, μπορούμε να πούμε ότι η οργάνωση της συλλογής και επεξεργασίας των σχετικών δεδομένων και πληροφοριών γίνεται κατά τρόπο που για δεδομένη ακρίβεια να επιτυγχάνεται το χαμηλότερο δυνατό κόστος ή, αντιστρόφως, να εξασφαλίζεται η μέγιστη δυνατή ακρίβεια την οποία επιτρέπουν τα μέσα που διαθέτουμε.

ΑΣΚΗΣΕΙΣ

- 1. Ποιες από τις παρακάτω μεταβλητές είναι ποιοτικές και ποιες ποσοτικές;
Από τις ποσοτικές ποιες είναι διακριτές και ποιες συνεχείς;**
- α) Βάρος**
 - β) Αριθμός τροχαίων ατυχημάτων**
 - γ) Φύλο**
 - δ) Οικογενειακή κατάσταση**
 - ε) Στάθμη της λίμνης του Μαραθώνα**
 - στ) Τόπος καταγωγής**
 - ζ) Επάγγελμα**
 - η) Αριθμός παιδιών στην οικογένεια**

- θ) Βαθμολογία στο σκάκι**
- ι) Νούμερο γυναικείων παπουτσιών.**

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:

- α) Εξετάζουμε ένα δείγμα υπαλλήλων μιας εταιρείας.**
- β) Εξετάζουμε ένα δείγμα προϊόντων από μια παραγωγή.**
- γ) Εξετάζουμε ένα δείγμα τηλεθεατών.**

δ) Εξετάζουμε τους καλαθοσφαιριστές μιας ομάδας σε έναν αγώνα.

3. Για να βρούμε ποιες εκπομπές στην τηλεόραση έχουν τη μεγαλύτερη ακροαματικότητα αποφασίσαμε να πάρουμε δείγμα 500 τηλεθεατών. Ποιος είναι, κατά τη γνώμη σας, ο καλύτερος από τους παρακάτω τρόπους, για να πάρουμε το δείγμα; Είναι καλύτερο να πάρουμε:

α) μόνο άνδρες, β) μόνο γυναίκες, γ) άτομα από τις μεγάλες πόλεις, δ) άτομα μόνο από την επαρχία, ε) άτομα από διάφορες περιοχές.

4. Τι έχετε να παρατηρήσετε για τα παρακάτω επιλεγόμενα δείγματα;

α) Για να βρούμε τα ποσοστά των ανδρών και των γυναικών στην Ελλάδα, πηγαίνουμε σε μια μεγάλη στρατιωτική μονάδα και ρωτάμε όλους τους στρατιώτες, πόσοι άνδρες και πόσες γυναίκες υπάρχουν στην οικογένειά τους.

β) Κάποιος θέλει να σχηματίσει μια ιδέα για το αποτέλεσμα των επερχόμενων βουλευτικών εκλογών. Τηλεφωνεί λοιπόν σε συγγενείς και φίλους του και τους ρωτάει σχετικά.

- γ) Για να εκτιμήσουμε το κατά κεφαλή εισόδημα των Ελλήνων παίρνουμε ένα δείγμα από το Κολωνάκι των Αθηνών.**
- δ) Για να δούμε πώς διασκεδάζουν οι νέοι της χώρας μας επιλέγουμε κάποιους μαθητές από διάφορα Λύκεια της Αττικής.**
- ε) Ο διευθυντής ενός Λυκείου αποφάσισε να καταγράψει τους λόγους της απουσίας των μαθητών από το Λύκειο κατά τη διάρκεια της ακαδημαϊκής χρονιάς. Γι' αυτό τον λόγο πήρε ως δείγμα όσους απουσίασαν το Νοέμβριο.**

2.2 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 σωστών συμπερασμάτων. Η παρουσίαση των στατιστικών δεδομένων σε πίνακες γίνεται με την κατάλληλη τοποθέτηση των πληροφοριών σε γραμμές και στήλες, με τρόπο που να διευκολύνεται η σύγκριση των στοιχείων και η καλύτερη ενημέρωση του αναγνώστη σχετικά με τη δομή του πληθυσμού που ερευνάμε.

Οι πίνακες διακρίνονται στους:

α) γενικούς πίνακες, οι οποίοι περιέχουν όλες τις πληροφορίες που προκύπτουν από μία στατιστική έρευνα (συνήθως με αρκετά λεπτομερειακά στοιχεία) και αποτελούν πηγές στατιστικών πληροφοριών στη διάθεση των επιστημόνων-ερευνητών για παραπέρα ανάλυση και εξαγωγή συμπερασμάτων,

β) ειδικούς πίνακες, οι οποίοι είναι συνοπτικοί και σαφείς. Τα στοιχεία τους συνήθως έχουν ληφθεί από τους γενικούς πίνακες.

Κάθε πίνακας που έχει κατασκευαστεί σωστά πρέπει να περιέχει:

α) τον τίτλο, που γράφεται στο επάνω μέρος του πίνακα και δηλώνει με σαφήνεια και συνοπτικά το περιεχόμενο του πίνακα,

β) τις επικεφαλίδες των γραμμών και στηλών, που δείχνουν συνοπτικά τη φύση και τις μονάδες μέτρησης των δεδομένων,
γ) το κύριο σώμα (κορμό), που περιέχει διαχωρισμένα μέσα στις γραμμές και στις στήλες τα στατιστικά δεδομένα,
δ) την πηγή, που γράφεται στο κάτω μέρος του πίνακα και δείχνει την προέλευση των στατιστικών στοιχείων, έτσι ώστε ο αναγνώστης να ανατρέχει σ' αυτήν, όταν επιθυμεί, για επαλήθευση στοιχείων ή για λήψη περισσότερων πληροφοριών.

Παρακάτω δίνονται μερικοί στατιστικοί πίνακες, που διευκρινίζουν την εφαρμογή των προηγούμενων εννοιών.

Πίνακας 1

Πληθυσμός της Ελλάδος (σε εκατομμύρια) κατά μεγάλες ομάδες ηλικιών

Ηλικία (σε έτη)	Απογραφή 1971	Απογραφή 1981	
0 - 14	2,22	2,31	
15 - 64	5,58	6,19	
≥ 65	0,96	1,24	

Πηγή: ΕΣΥΕ, 1996

**Απογραφή
1991**

**Εκτίμηση
1993**

**Εκτίμηση
1994**

1,97

1,85

1,81

6,88

6,99

7,04

1,40

1,54

1,58

Πίνακας 2

Επιφάνεια και πληθυσμός των κατοικημένων νησιών της Ελλάδας με πληθυσμό, κατά την απογραφή του 1991, άνω των 10.000 κατοίκων.

Κατοικημένες νήσοι	Επιφάνεια σε τ.χμ.
Κρήτη	8.261,183
Εύβοια	3.661,637
Λέσβος	1.635,998
Ρόδος	1.401,459
Χίος	842,796
Κεφαλληνία	734,014
Κέρκυρα	585,312
Σάμος	477,942
Λήμνος	476,288
Ζάκυνθος	406,612
Νάξος	389,434

Πληθυσμός κατά τις απογραφές

1971

1981

1991

456.471

502.082

539.938

162.986

185.626

205.502

97.008

88.601

87.151

66.606

87.831

98.181

52.487

48.700

51.060

31.787

27.649

29.392

89.578

96.533

104.781

32.664

31.629

33.032

17.367

15.721

17.645

30.180

30.011

32.556

14.201

14.037

14.838

Κατοικημένες νήσοι	Επιφάνεια σε τ.χμ.	
Θάσος	383,672	
Λευκάδα	301,106	
Κως	287,611	
Κάλυμνος	110,581	
Σαλαμίνα	91,503	
Σύρος	84,069	
Αίγινα	77,014	

Πηγή: ΕΣΥΕ, Απογραφή 1991

Πληθυσμός κατά τις απογραφές

1971

1981

1991

13.316

13.111

13.527

22.917

19.947

19.350

16.650

20.350

26.379

13.097

14.295

15.706

23.065

28.574

34.272

18.642

19.668

19.870

9.553

11.127

11.639

Πίνακας 3

Εργατικά ατυχήματα κατά ομάδες ηλικιών

Έτη 1990-94

Ηλικία	1990	1991	
Κάτω των 15	18	9	
15 - 19	731	564	
20 - 24	3323	2785	
25 - 29	4277	3921	
30 - 34	3952	3700	
35 - 39	3589	3146	
40 - 44	3237	2803	
45 - 49	2839	2593	
50 - 54	2727	2564	
55 - 59	2304	2230	
60 - 64	728	720	
65 - 69	121	150	
Σύνολο	27846	25185	

Πηγή: ΙΚΑ, Ελληνικό Ινστιτούτο

Υγιεινής και Ασφάλειας της Εργασίας

	1992	1993	1994
	10	16	5
	437	735	442
	2755	2981	2696
	4246	3881	3717
	3388	3348	3282
	3233	3230	3000
	2911	2880	2903
	2784	2608	2403
	2286	2095	1877
	2185	1699	1664
	688	420	523
	140	66	96
	25063	23959	22608

Πίνακας 4

Χαρακτηριστικά 40 μαθητών Γ΄ τάξης
ενός Λυκείου.

α.α	Φύλο	Ασχολία *	Αριθμός αδελφών	
1	Κ	4	1	
2	Α	1	0	
3	Κ	4	2	
4	Κ	5	1	
5	Κ	5	0	
6	Κ	4	3	
7	Κ	4	2	
8	Α	4	1	
9	Α	2	3	
10	Κ	3	1	
11	Κ	4	1	
12	Α	2	1	

	Βαθμός μαθη- ματι- κών Β' λυκείου	Ύψος (cm)	Βάρος (kg)	Ύψος πατέ- ρα (cm)	Ύψος μητέ- ρας (cm)
	15	170	60	172	168
	17	180	68	185	165
	12	178	62	181	160
	18	165	47	180	162
	15	170	54	180	168
	16	168	56	185	168
	15	175	58	193	162
	15	175	72	174	174
	13	173	67	182	160
	15	162	50	176	170
	16	160	51	176	164
	11	170	58	182	165

α.α	Φύλο	Ασχολία *	Αριθμός αδελφών	
13	Κ	7	3	
14	Α	1	1	
15	Α	1	0	
16	Κ	2	2	
17	Α	2	0	
18	Α	7	2	
19	Κ	4	1	
20	Α	5	1	
21	Κ	5	1	
22	Κ	5	1	
23	Α	5	2	
24	Κ	6	1	
25	Κ	6	2	
26	Α	5	0	
27	Κ	6	1	
28	Α	3	1	

	Βαθμός μαθη- ματι- κών Β' λυκείου	Ύψος (cm)	Βάρος (kg)	Ύψος πατέ- ρα (cm)	Ύψος μητέ- ρας (cm)
	20	167	50	174	170
	18	177	81	177	169
	17	180	70	170	165
	19	170	63	165	174
	14	182	71	176	173
	17	178	73	182	170
	14	165	58	180	161
	16	178	74	173	168
	12	156	44	170	158
	13	175	53	170	165
	18	172	60	178	165
	16	173	64	182	162
	14	167	57	172	157
	14	187	85	185	170
	17	170	62	180	165
	12	180	80	180	167

α.α	Φύλο	Ασχολία *	Αριθμός αδελφών	
29	A	3	0	
30	A	2	1	
31	A	2	0	
32	K	4	1	
33	K	4	2	
34	K	4	1	
35	A	3	1	
36	A	3	1	
37	A	5	1	
38	K	5	1	
39	K	3	0	
40	K	4	1	

***1: Υπολογιστές, 2:Αθλητισμός,
3:Διασκέδαση - Ντίσκο, 4:Μουσι-
κή, 5:Τηλεόραση -Κινηματογράφος,
6:Διάβασμα εξωσχολικών βιβλίων,**

Βαθμός μαθη- ματι- κών Β' λυκείου	Ύψος (cm)	Βάρος (kg)	Ύψος πατέ- ρα (cm)	Ύψος μητέ- ρας (cm)
15	178	73	173	170
10	191	86	180	170
16	176	65	180	172
12	169	57	170	167
14	167	61	179	158
19	166	62	178	165
19	179	76	178	160
16	178	68	180	160
19	180	85	170	163
19	164	64	184	170
15	170	63	165	167
15	173	63	186	162

7:Άλλο

Πηγή: Δειγματοληπτική έρευνα με-
ταξύ μαθητών 1ου Λυκείου Αμαρου-
σίου (Σεπτ. '98).

Πίνακες Κατανομής Συχνοτήτων

Ας υποθέσουμε ότι x_1, x_2, \dots, x_k είναι οι τιμές μιας μεταβλητής X , που αφορά τα άτομα ενός δείγματος μεγέθους n , $k \leq n$. Στην τιμή x_i αντιστοιχίζεται η (απόλυτη) συχνότητα (frequency) v_i , δηλαδή ο φυσικός αριθμός που δείχνει πόσες φορές εμφανίζεται η τιμή x_i της εξεταζόμενης μεταβλητής X στο σύνολο των παρατηρήσεων. Είναι φανερό ότι το άθροισμα όλων των συχνοτήτων είναι ίσο με το μέγεθος n του δείγματος, δηλαδή:

$$v_1 + v_2 + \dots + v_k = n \quad (1)$$

Για παράδειγμα, για τη μεταβλητή X : “αριθμός αδελφών” του πίνακα 4 οι συχνότητες για τις τιμές $x_1 = 0$,

$x_2 = 0$, $x_3 = 2$, $x_4 = 3$ είναι, αντίστοιχα, $v_1 = 8$, $v_2 = 22$, $v_3 = 7$, $v_4 = 3$ με $v_1 + v_2 + v_3 + v_4 = 40$. Ο υπολογισμός των συχνοτήτων γίνεται με τη διαλογή των παρατηρήσεων, όπως φαίνεται στον παρακάτω πίνακα 5. Διατρέχοντας με τη σειρά τη λίστα των δεδομένων καταγράφουμε κάθε παρατήρηση με συμβολικό τρόπο σαν μια γραμμή “ | ” στην αντίστοιχη τιμή της μεταβλητής.

Πίνακας 5

Κατανομή συχνοτήτων της μεταβλητής X : “αριθμός αδελφών” των μαθητών του πίνακα 4.

Αριθμός αδελφών x_i	Διαλογή
0	### III
1	### ### ### ### II
2	### II
3	III
Σύνολο:	

	Συχνότητα v_i	Σχετική Συχνότητα f_i	Σχετική Συχνότητα $f_i \%$
	8	0,200	20,0
	22	0,550	55,0
	7	0,175	17,5
	3	0,075	7,5
	40	1,000	100,0

Αν διαιρέσουμε τη συχνότητα v_i με το μέγεθος v του δείγματος, προκύπτει η **σχετική συχνότητα** (relative frequency) f_i της τιμής x_i , δηλαδή

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

Για τη σχετική συχνότητα ισχύουν οι ιδιότητες:

i) $0 \leq f_i \leq 1$ για $i = 1, 2, \dots, k$ αφού
 $0 \leq v_i \leq v$.

ii) $f_1 + f_2 + \dots + f_k = 1$, αφού

$$\begin{aligned} f_1 + f_2 + \dots + f_k &= \frac{v_1}{v} + \frac{v_2}{v} + \dots + \frac{v_k}{v} = \\ &= \frac{v_1 + v_2 + \dots + v_k}{v} = \frac{v}{v} = 1. \end{aligned}$$

Συνήθως, τις σχετικές συχνότητες f_i

τις εκφράζουμε επί τοις εκατό, οπότε συμβολίζονται με $f_i\%$, δηλαδή $f_i\% = 100f_i$. Για παράδειγμα, οι σχετικές συχνότητες για τις τιμές $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, $x_4 = 3$ της μεταβλητής X : “αριθμός αδελφών” είναι αντιστοίχως:

$$f_1 = \frac{8}{40} = 0,20, \quad f_2 = \frac{22}{40} = 0,55,$$

$$f_3 = \frac{7}{40} = 0,175 \quad \text{και} \quad f_4 = \frac{3}{40} = 0,075$$

$$\begin{aligned} \text{με} \quad f_1 + f_2 + f_3 + f_4 &= \\ &= 0,20 + 0,55 + 0,175 + 0,075 = 1. \end{aligned}$$

$$\begin{aligned} \text{Συνεπώς} \quad f_1\% &= 20, \quad f_2\% = 55, \\ f_3\% &= 17,5 \quad \text{και} \quad f_4\% = 7,5 \quad \text{με} \\ f_1\% + f_2\% + f_3\% + f_4\% &= 100. \end{aligned}$$

Οι ποσότητες x_i , v_i , f_i για ένα δείγμα συγκεντρώνονται σε ένα συνοπτικό πίνακα, που ονομάζεται πίνακας κατανομής συχνοτήτων ή απλά πίνακας συχνοτήτων.

Για μια μεταβλητή, το σύνολο των ζευγών (x_i, v_i) λέμε ότι αποτελεί την κατανομή συχνοτήτων και το σύνολο των ζευγών (x_i, f_i) , ή των ζευγών $(x_i, f_i\%)$, την κατανομή των σχετικών συχνοτήτων. Στον πίνακα 5 παρουσιάζονται οι κατανομές συχνοτήτων και σχετικών συχνοτήτων της μεταβλητής X : “αριθμός αδελφών” των μαθητών του πίνακα 4.

Αθροιστικές Συχνότητες

Στην περίπτωση των ποσοτικών μεταβλητών εκτός από τις συχνότητες v_i και f_i χρησιμοποιούνται

συνήθως και οι λεγόμενες αθροιστικές συχνότητες (cumulative frequencies) N_i και οι αθροιστικές σχετικές συχνότητες (cumulative relative frequencies) F_i , οι οποίες εκφράζουν το πλήθος και το ποσοστό αντίστοιχα των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i . Συχνά οι F_i πολλαπλασιάζονται επί 100 εκφραζόμενες έτσι επί τοις εκατό, δηλαδή $F_i\% = 100F_i$, βλέπε πίνακα 6. Αν οι τιμές x_1, x_2, \dots, x_k μιας ποσοτικής μεταβλητής X είναι σε αύξουσα διάταξη, τότε η αθροιστική συχνότητα της τιμής x_i είναι $N_i = v_1 + v_2 + \dots + v_i$. Όμοια, η αθροιστική σχετική συχνότητα είναι $F_i = f_1 + f_2 + \dots + f_i$, για $i = 1, 2, \dots, k$. Για παράδειγμα, για τη μεταβλητή X : “αριθμός αδελφών” του πίνακα

**4 είναι $N_1 = v_1 = 8$, $N_2 = v_1 + v_2 = 30$,
 $N_3 = v_1 + v_2 + v_3 = 37$ και
 $N_4 = v_1 + v_2 + v_3 + v_4 = v = 40$, οπότε
 $F_1 = f_1 = 0,20$, $F_2 = f_1 + f_2 = 0,75$,
 $F_3 = f_1 + f_2 + f_3 = 0,925$ και
 $F_4 = f_1 + f_2 + f_3 + f_4 = 1$, οπότε
 $F_1\% = 20$, $F_2\% = 75$, $F_3\% = 92,5$ και
 $F_4\% = 100$. Είναι φανερό ότι ισχύ-
 ουν οι σχέσεις:**

$$v_1 = N_1, \quad v_2 = N_2 - N_1, \quad \dots,$$

$$v_k = N_k - N_{k-1}$$

και

$$f_1 = F_1, \quad f_2 = F_2 - F_1, \quad \dots,$$

$$f_k = F_k - F_{k-1}.$$

Πίνακας 6

Κατανομή συχνοτήτων και αθροιστικών συχνοτήτων της μεταβλητής “αριθμός αδελφών” των μαθητών του πίνακα 4.

Αριθμός αδελφών x_i	Συχνότητα v_i	Σχετ. Συχν. f_i	
0	8	0,200	
1	22	0,550	
2	7	0,175	
3	3	0,075	
Σύνολο:	40	1,000	

	ΣΧΕΤ. ΣΥΧΝ. $f_i\%$	Αθροισ. ΣΥΧΝ. N_i	Αθροισ. ΣΧΕΤ. ΣΥΧΝ. F_i	Αθροισ. ΣΧΕΤ. ΣΥΧΝ. $F_i\%$
	20,0	8	0,200	20,0
	55,0	30	0,750	75,0
	17,5	37	0,925	92,5
	7,5	40	1,000	100,0
	100,0	–	–	–

Γραφική Παράσταση Κατανομής Συχνοτήτων

Τα στατιστικά δεδομένα παρουσιάζονται πολλές φορές και υπό μορφή γραφικών παραστάσεων ή διαγραμμάτων. Οι γραφικές παραστάσεις παρέχουν πιο σαφή εικόνα του χαρακτηριστικού σε σχέση με τους πίνακες, είναι πολύ πιο ενδιαφέρουσες και ελκυστικές, χωρίς βέβαια να προσφέρουν περισσότερη πληροφορία από εκείνη που περιέχεται στους αντίστοιχους πίνακες συχνοτήτων. Επί πλέον με τα διαγράμματα διευκολύνεται η σύγκριση μεταξύ ομοειδών στοιχείων για το ίδιο ή για διαφορετικά χαρακτηριστικά.

Υπάρχουν διάφοροι τρόποι γραφικής παρουσίασης, ανάλογα με το είδος των δεδομένων που έχουμε. Όπως όμως οι στατιστικοί πίνακες έτσι και τα στατιστικά διαγράμματα πρέπει να συνοδεύονται από α) τον τίτλο, β) την κλίμακα με τις τιμές των μεγεθών που απεικονίζονται, γ) το υπόμνημα που επεξηγεί συνήθως τις τιμές της μεταβλητής και δ) την πηγή των δεδομένων.

α) Ραβδόγραμμα

Το ραβδόγραμμα (barchart) χρησιμοποιείται για τη γραφική παράσταση των τιμών μιας ποιοτικής μεταβλητής. Το ραβδόγραμμα αποτελείται από ορθογώνιες στήλες που οι βάσεις τους βρίσκονται

πάνω στον οριζόντιο ή τον κατακόρυφο άξονα. Σε κάθε τιμή της μεταβλητής X αντιστοιχεί μια ορθογώνια στήλη της οποίας το ύψος είναι ίσο με την αντίστοιχη συχνότητα ή σχετική συχνότητα. Έτσι έχουμε αντίστοιχα το ραβδόγραμμα συχνοτήτων και το ραβδόγραμμα σχετικών συχνοτήτων. Τόσο η απόσταση μεταξύ των στηλών όσο και το μήκος των βάσεων τους καθορίζονται αυθαίρετα. Στον πίνακα 7 έχουμε την κατανομή συχνοτήτων της μεταβλητής X : “απασχόληση στον ελεύθερο χρόνο” και στα σχήματα 1(α), (β) τα αντίστοιχα ραβδογράμματα συχνοτήτων και σχετικών συχνοτήτων.

Πίνακας 7

Κατανομή συχνοτήτων για την απασχόληση στον ελεύθερο χρόνο τους των μαθητών του πίνακα 4.

i	Απασχόληση x_i	Συχνότητα v_i	
1	Υπολογιστές	3	
2	Αθλητισμός	6	
3	Διασκέδαση- ντίσκο	6	
4	Μουσική	11	
5	Τηλεόραση-Κι- νηματογράφος.	9	
6	Διάβασμα εξωσχ. Βιβλίων	3	
7	Άλλο	2	
Σύνολο:		40	

**Σχετική
συχνότητα
 f_i**

**Σχετική
συχνότητα
 $f_i\%$**

0,075

7,5

0,150

15,0

0,150

15,0

0,275

27,5

0,225

22,5

0,075

7,5

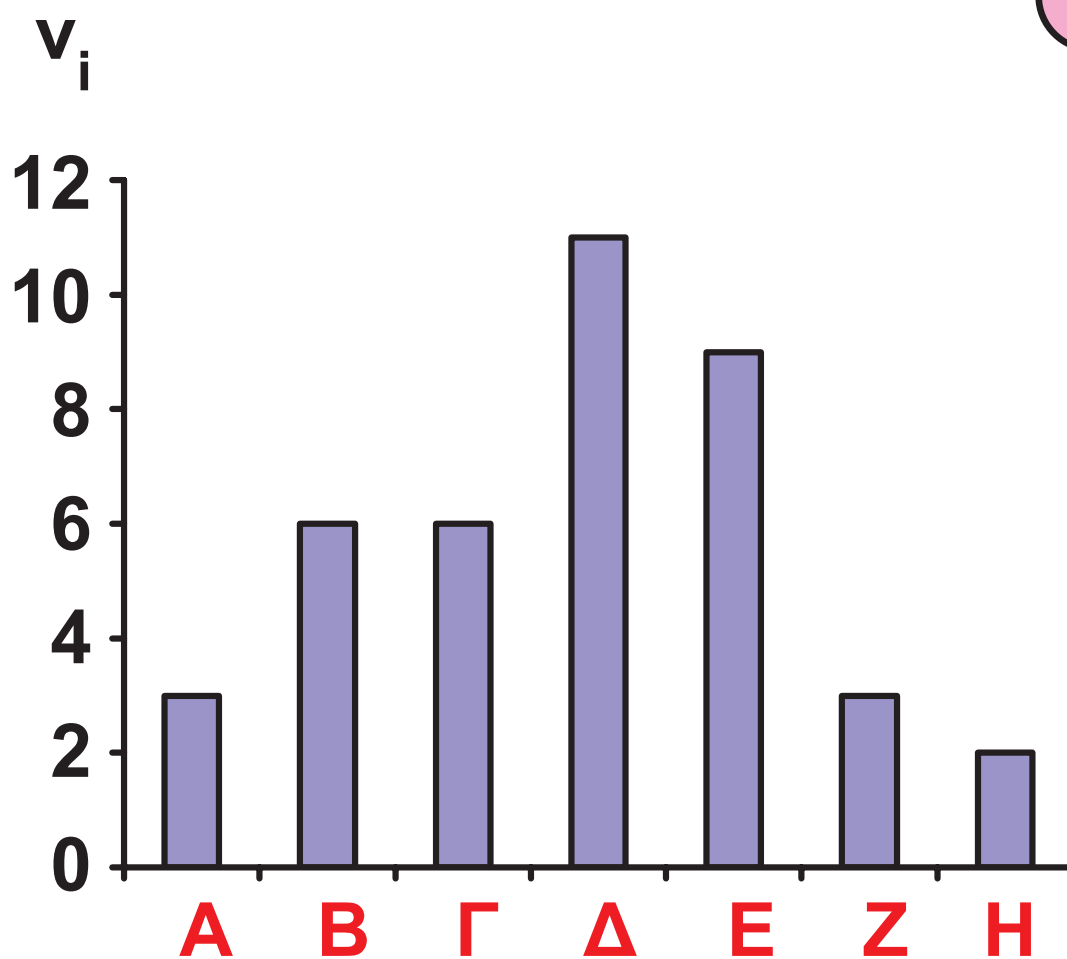
0,050

5,0

1,000

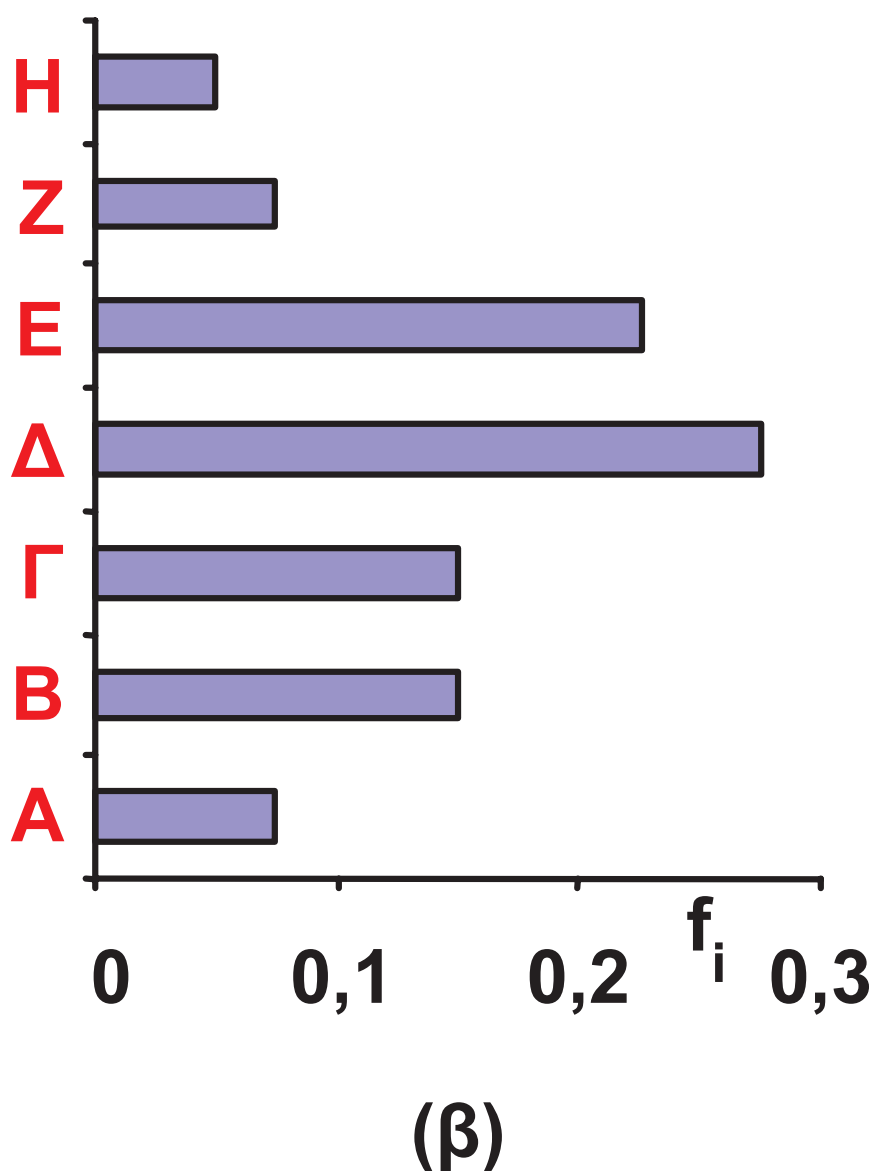
100,0

Μερικές φορές σε ένα ραβδόγραμμα συχνοτήτων ο ρόλος των δύο αξόνων είναι δυνατόν να αντιστραφεί, όπως φαίνεται στο σχήμα 1(β), που παριστάνεται το ραβδόγραμμα σχετικών συχνοτήτων της ίδιας μεταβλητής. Αν θέλουμε να συγκρίνουμε τον τρόπο που περνούν τον ελεύθερο χρόνο τους τα αγόρια και τα κορίτσια, τότε κατασκευάζουμε το ραβδόγραμμα σχετικών συχνοτήτων του σχήματος 1(γ), όπως προκύπτει από τον πίνακα 4.



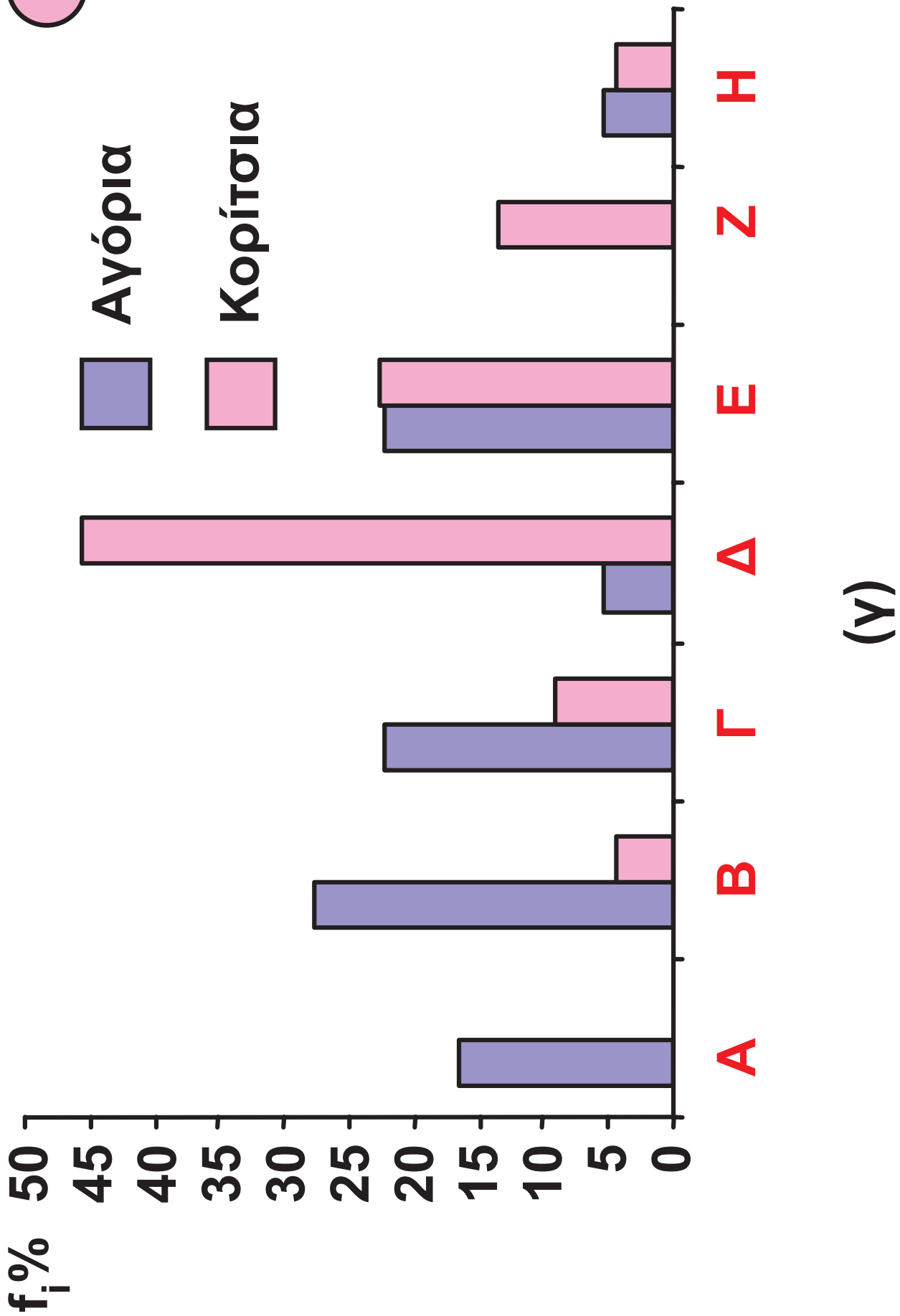
(α)

Ραβδόγραμμα συχνοτήτων (α) και σχετικών συχνοτήτων (β) για την απασχόληση των μαθητών του πίνακα 7.



A. Η/Υ **B.** Αθλητισμός **Γ.** Διασκέδαση-Ντίσκο **Δ.** Μουσική **E.** Τηλεόραση-Κινηματογράφος. **Z.** Διάβασμα εξωσχ. Βιβλίων **H.** Άλλο

1



Ραβδόγραμμα σχετικών συχνοτήτων για την απασχόληση των μαθητών του πίνακα 4 ανάλογα με το φύλο.

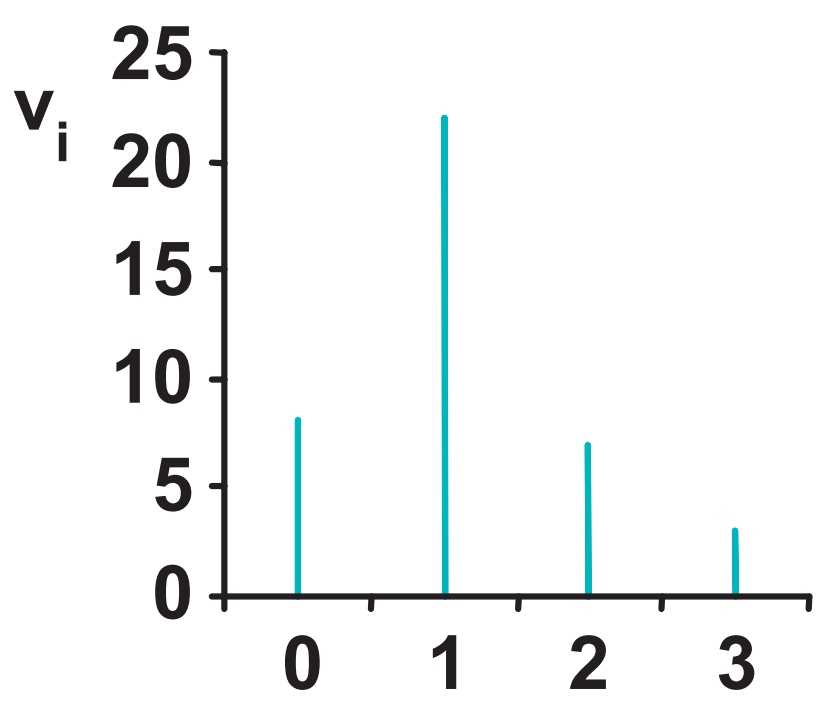
- A.** Η/Υ **B.** Αθλητισμός **Γ.** Διασκέδαση-Ντίσκο
- Δ.** Μουσική **Ε.** Τηλεόραση-Κινηματογράφος.
- Ζ.** Διάβασμα εξωσχ. Βιβλίων **Η.** Άλλο

β) Διάγραμμα Συχνοτήτων

Στην περίπτωση που έχουμε μια ποσοτική μεταβλητή αντί του ραβδογράμματος χρησιμοποιείται το **διάγραμμα συχνοτήτων** (line diagram). Αυτό μοιάζει με το ραβδόγραμμα με μόνη διαφορά ότι αντί να χρησιμοποιούμε συμπαγή ορθογώνια υψώνουμε σε κάθε x_i (υποθέτοντας ότι $x_1 < x_2 < \dots < x_k$) μία κάθετη γραμμή με μήκος ίσο προς την αντίστοιχη συχνότητα, όπως φαίνεται στο σχήμα 2(α). Μπορούμε επίσης αντί των συχνοτήτων v_i στον κάθετο άξονα να βάλουμε τις σχετικές συχνότητες f_i , οπότε έχουμε το **διάγραμμα σχετικών συχνοτήτων**. Ενώνοντας τα σημεία (x_i, v_i) ή (x_i, f_i) έχουμε το λεγόμενο **πολύγωνο συχνοτήτων** ή **πολύγωνο σχετικών**

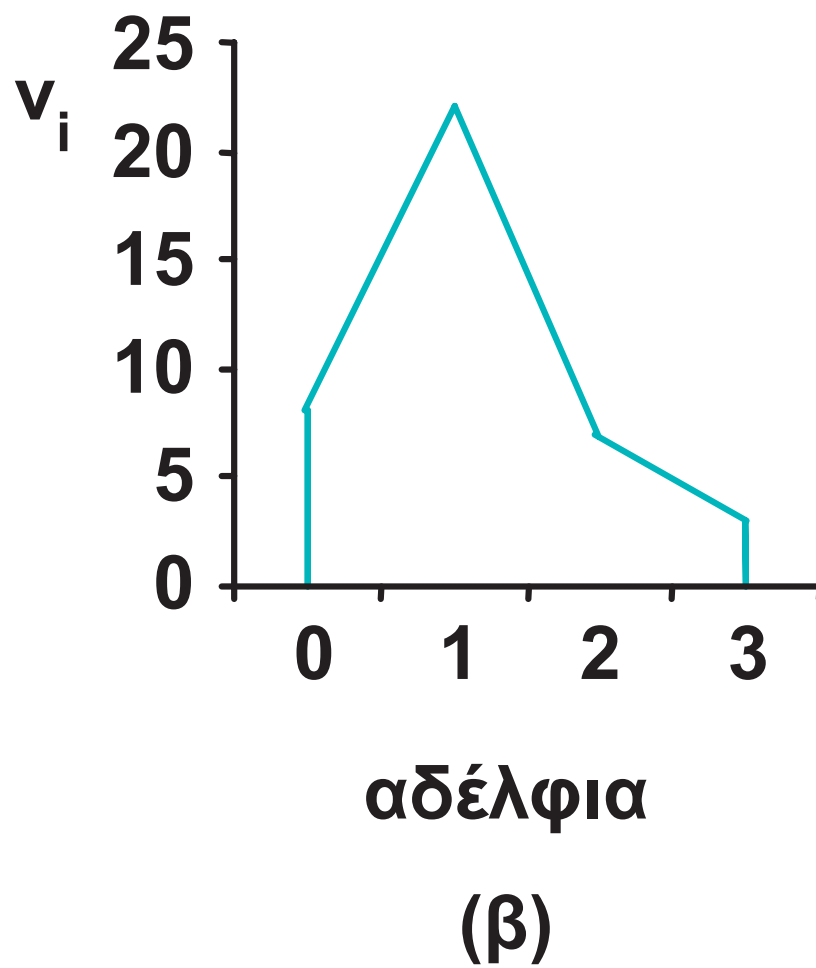
συχνοτήτων, αντίστοιχα, που μας δίνουν μια γενική ιδέα για τη μεταβολή της συχνότητας ή της σχετικής συχνότητας όσο μεγαλώνει η τιμή της μεταβλητής που εξετάζουμε, βλέπε σχήμα 2(β).

2α



αδέλφια
(α)

2β



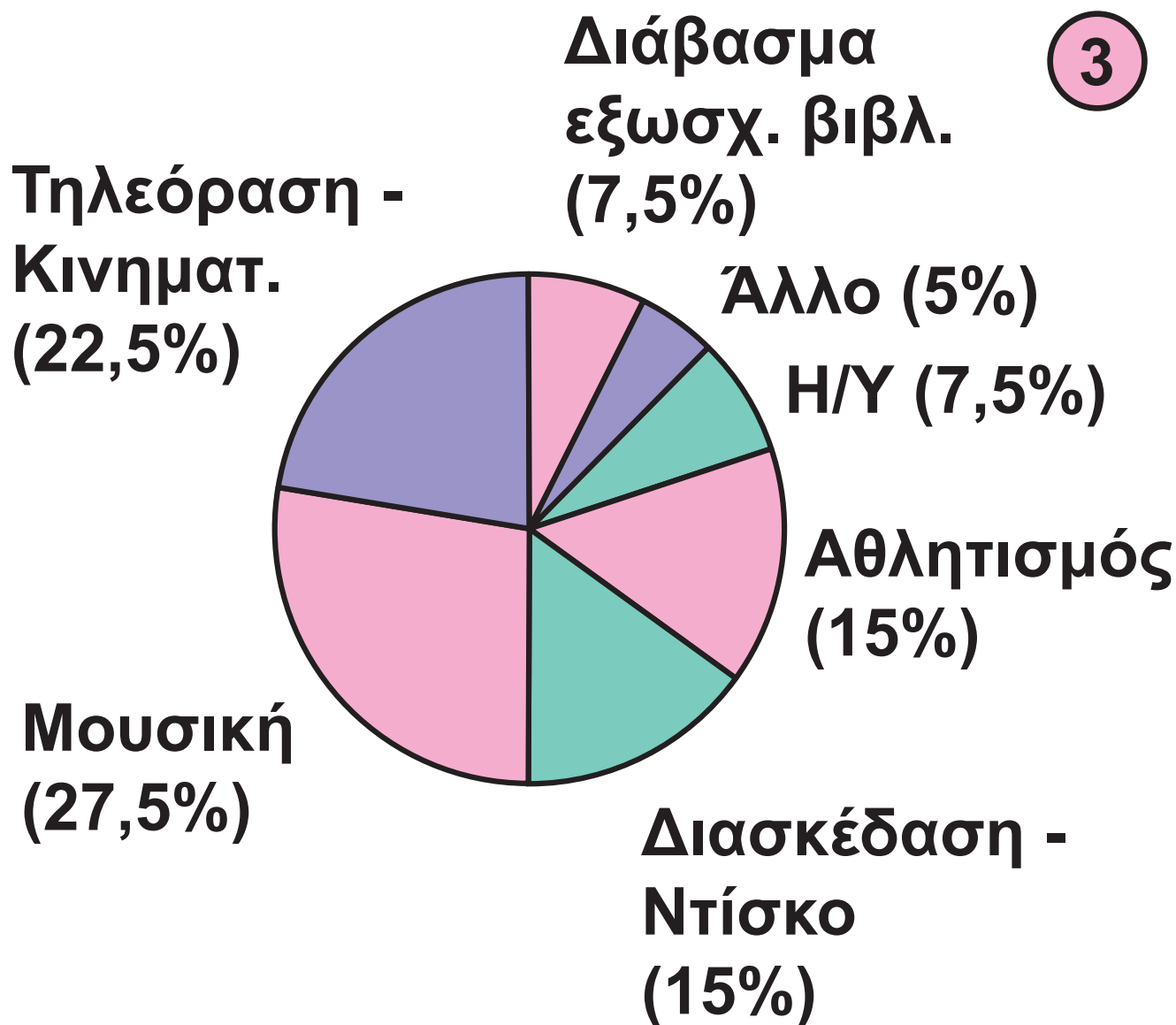
Διάγραμμα συχνοτήτων (α) και πολύγωνο συχνοτήτων (β) για τη μεταβλητή “αριθμός αδελφών” του πίνακα 4.

γ) Κυκλικό Διάγραμμα

Το κυκλικό διάγραμμα (pie chart) χρησιμοποιείται για τη γραφική παράσταση τόσο των ποιοτικών όσο και των ποσοτικών δεδομένων, όταν οι διαφορετικές τιμές της μεταβλητής είναι σχετικά λίγες. Το κυκλικό διάγραμμα είναι ένας κυκλικός δίσκος χωρισμένος σε κυκλικούς τομείς, τα εμβαδά ή, ισοδύναμα, τα τόξα των οποίων είναι ανάλογα προς τις αντίστοιχες συχνότητες v_i ή τις σχετικές συχνότητες f_i των τιμών x_i της μεταβλητής. Αν συμβολίσουμε με α_i το αντίστοιχο τόξο ενός κυκλικού τμήματος στο κυκλικό διάγραμμα συχνοτήτων, τότε

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

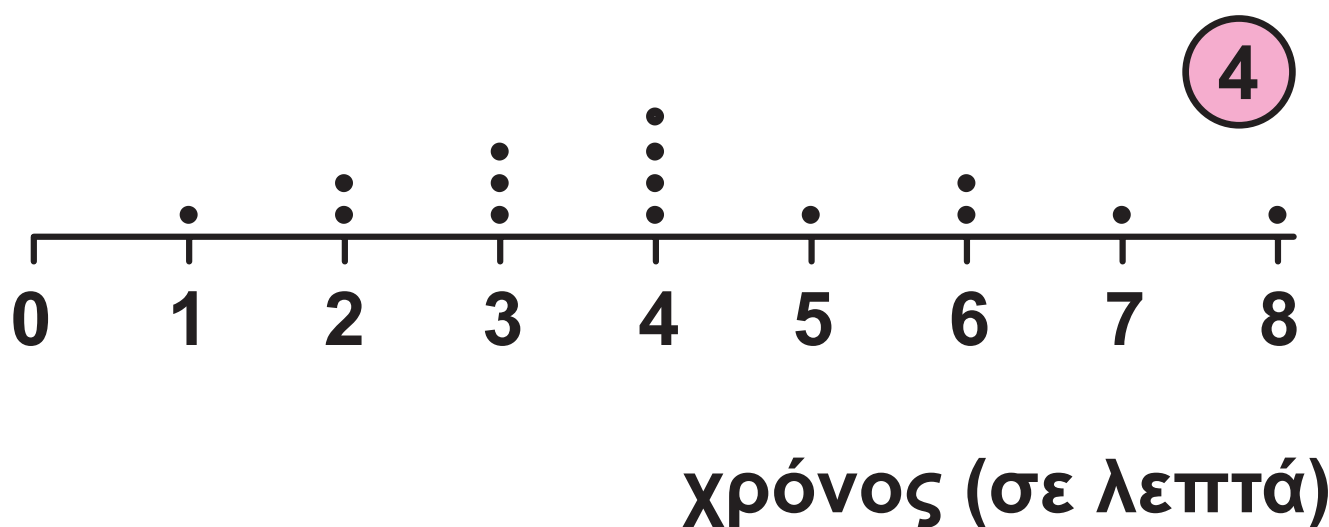
Στο σχήμα 3 παριστάνεται το αντίστοιχο κυκλικό διάγραμμα σχετικών συχνοτήτων της “απασχόλησης των μαθητών” για τα δεδομένα του πίνακα 4.



Κυκλικό διάγραμμα σχετικών συχνοτήτων της απασχόλησης των μαθητών για τα δεδομένα του πίνακα 4.

δ) Σημειόγραμμα

Όταν έχουμε λίγες παρατηρήσεις, η κατανομή τους μπορεί να περιγραφεί με το σημειόγραμμα (dot diagram), στο οποίο οι τιμές παριστάνονται γραφικά σαν σημεία υπεράνω ενός οριζόντιου άξονα. Στο σχήμα 4 έχουμε το σημειόγραμμα των χρόνων (σε λεπτά) 4, 2, 3, 1, 5, 6, 4, 2, 3, 4, 7, 4, 8, 6, 3 που χρειάστηκαν δεκαπέντε μαθητές, για να λύσουν ένα πρόβλημα.



ε) Χρονόγραμμα

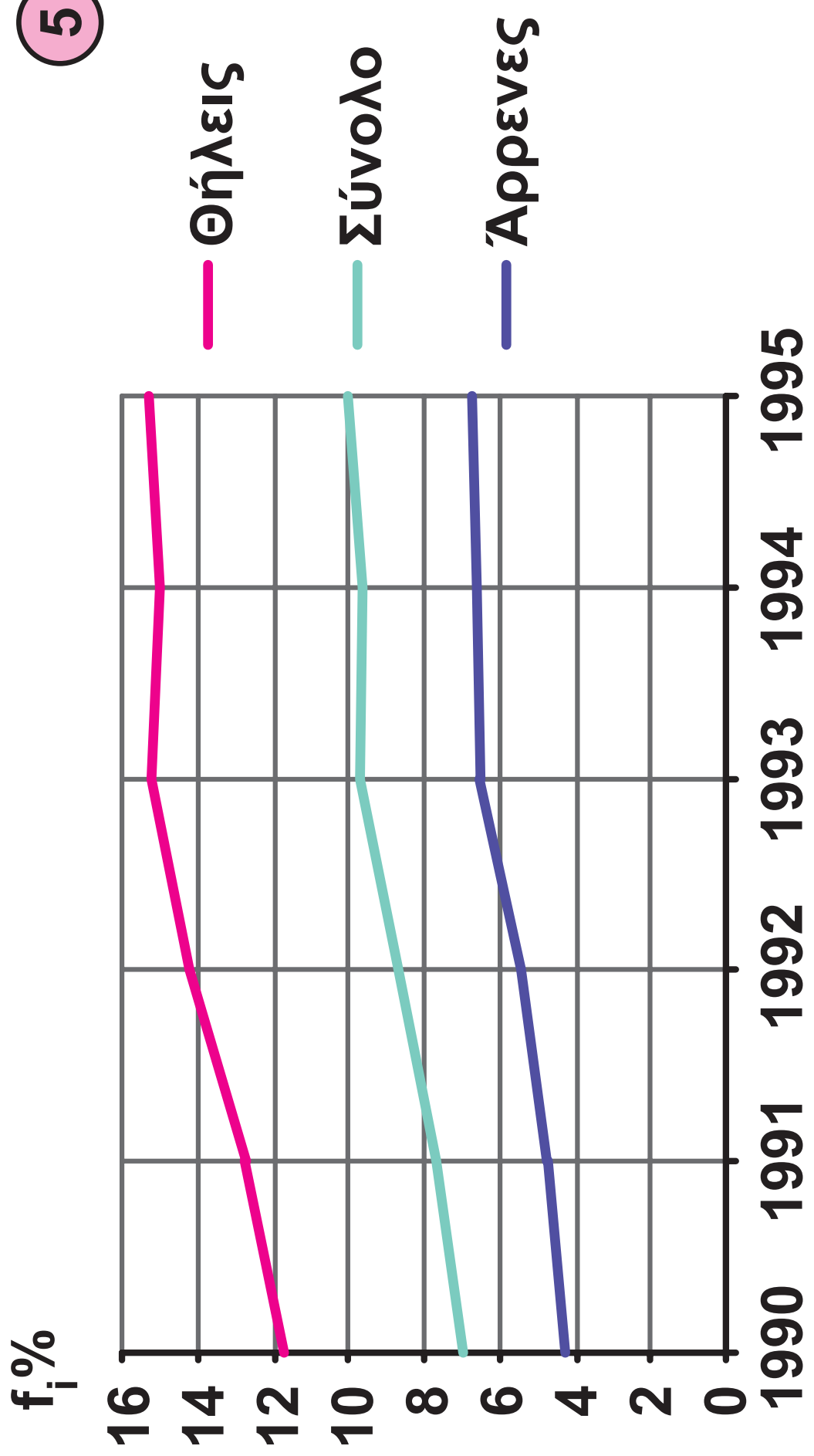
Το χρονόγραμμα ή χρονολογικό διάγραμμα χρησιμοποιείται για τη γραφική απεικόνιση της διαχρονικής εξέλιξης ενός οικονομικού, δημογραφικού ή άλλου μεγέθους. Ο οριζόντιος άξονας χρησιμοποιείται συνήθως ως άξονας μέτρησης του χρόνου και ο κάθετος ως άξονας μέτρησης της εξεταζόμενης μεταβλητής.

Στο σχήμα 5 έχουμε το χρονόγραμμα του ποσοστού ανεργίας στη χώρα μας από το 1990 έως το 1995. (Πηγή ΕΣΥΕ).

Παρατηρούμε ότι στο γυναικείο πληθυσμό υπάρχει συστηματικά μεγαλύτερο ποσοστό ανεργίας, γύρω στις 8 εκατοστιαίες μονάδες. Στο διάστημα 1993-95 το ποσοστό

**ανεργίας έχει σταθεροποιηθεί γύρω
στο 6,5% για τους άνδρες και γύρω
στο 15% για τις γυναίκες.**

5



Ποσοστά ανεργίας στην Ελλάδα

Ομαδοποίηση των Παρατηρήσεων

Οι πίνακες συχνοτήτων και κατ' αναλογία τα αντίστοιχα διαγράμματα είναι δύσκολο να κατασκευαστούν, όταν το πλήθος των τιμών μιας μεταβλητής είναι αρκετά μεγάλο. Αυτό μπορεί να συμβεί είτε στην περίπτωση μιας διακριτής μεταβλητής είτε, πολύ περισσότερο, στην περίπτωση μιας συνεχούς μεταβλητής, όπου αυτή μπορεί να πάρει οποιαδήποτε τιμή στο διάστημα ορισμού της. Σ' αυτές τις περιπτώσεις είναι απαραίτητο να ταξινομηθούν (ομαδοποιηθούν) τα δεδομένα σε μικρό πλήθος ομάδων, που ονομάζονται και κλάσεις (class intervals), έτσι ώστε κάθε τιμή να ανήκει μόνο σε μία κλάση. Τα άκρα των κλάσεων καλούνται όρια των

κλάσεων (class boundaries). Συνήθως υιοθετούμε την περίπτωση που μια κλάση περιέχει το κάτω άκρο της (κλειστή αριστερά) αλλά όχι το άνω άκρο της (ανοικτή δεξιά), δηλαδή που οι κλάσεις είναι της μορφής $[,)$. Οι παρατηρήσεις κάθε κλάσης θεωρούνται όμοιες, οπότε μπορούν να “αντιπροσωπευθούν” από τις κεντρικές τιμές, τα κέντρα δηλαδή κάθε κλάσης.

- Το πρώτο βήμα στην ομαδοποίηση των δεδομένων είναι η εκλογή του αριθμού k των ομάδων ή κλάσεων. Ο αριθμός αυτός συνήθως ορίζεται αυθαίρετα από τον ερευνητή σύμφωνα με την πείρα του. Γενικά όμως μπορεί να χρησιμοποιηθεί ως οδηγός ο παρακάτω πίνακας:

Μέγεθος δείγματος ν	Αριθμός κλάσεων κ
< 20	5
20 - 50	6
50 - 100	7
100 - 200	8
200 - 400	9
400 - 700	10
700 - 1000	11
≥ 1000	12

- Το δεύτερο βήμα είναι ο προσδιορισμός του πλάτους των κλάσεων. Πλάτος μιας κλάσης ονομάζεται η διαφορά του κατωτέρου από το ανώτερο όριο της κλάσης. Στην πλειονότητα των πρακτικών

εφαρμογών οι κλάσεις έχουν το ίδιο πλάτος. Φυσικά υπάρχουν και περιπτώσεις όπου επιβάλλεται οι κλάσεις να έχουν άνισο πλάτος, όπως, για παράδειγμα, στις κατανομές εισοδήματος, ημερών απεργίας κτλ. Για να κατασκευάσουμε ισοπλατείς κλάσεις, χρησιμοποιούμε το εύρος (range) R του δείγματος, δηλαδή τη διαφορά της μικρότερης παρατήρησης από τη μεγαλύτερη παρατήρηση του συνολικού δείγματος. Τότε υπολογίζουμε το πλάτος c των κλάσεων διαιρώντας το εύρος R διά του αριθμού των κλάσεων k , στρογγυλεύοντας, αν χρειαστεί για λόγους διευκόλυνσης, πάντα προς τα πάνω.

- Το επόμενο βήμα είναι η κατασκευή των κλάσεων. Ξεκινώντας

από την μικρότερη παρατήρηση, ή για πρακτικούς λόγους λίγο πιο κάτω από την μικρότερη παρατήρηση, και προσθέτοντας κάθε φορά το πλάτος c δημιουργούμε τις k κλάσεις. Αυτονόητο είναι ότι η μεγαλύτερη τιμή του δείγματος θα (πρέπει να) ανήκει οπωσδήποτε στην τελευταία κλάση.

- Τέλος, γίνεται η διαλογή των παρατηρήσεων. Το πλήθος των παρατηρήσεων n_i που προκύπτουν από τη διαλογή για την κλάση i καλείται συχνότητα της κλάσης αυτής ή συχνότητα της κεντρικής τιμής x_i , $i = 1, 2, \dots, k$.

Έστω, για παράδειγμα, ότι από τα δεδομένα του πίνακα 4 εξετάζουμε το ύψος των μαθητών. Το ύψος των

μαθητών, όπως έχει καταγραφεί με τη σειρά, δίνεται στον παρακάτω πίνακα 8.

Πίνακας 8

Το ύψος (σε cm) των μαθητών της Γ΄ Λυκείου, όπως έχει καταγραφεί στον πίνακα 4. Σε αγκύλες έχουμε τη μικρότερη και τη μεγαλύτερη τιμή.

170	180	178	165	170
168	175	175	173	162
160	170	167	177	180
170	182	178	165	178
[156]	175	172	173	167
187	170	180	178	[191]
176	169	167	166	179
178	180	164	170	173

Παρατηρούμε ότι το εύρος του δείγματος είναι $R = 191 - 156 = 35$. Επειδή έχουμε $n = 40$ παρατηρήσεις, χρησιμοποιούμε $k = 6$ κλάσεις.

Το πλάτος των κλάσεων είναι

$$c = \frac{R}{k} = \frac{35}{6} = 5,83 \approx 6. \text{ Αν θεωρήσου-}$$

με ως αρχή της πρώτης κλάσης το 156, θα έχουμε τον επόμενο πίνακα 9.

Πρέπει να προσεχτεί ότι:

- Καμία παρατήρηση δεν μπορεί να μείνει έξω από κάποια κλάση.**
- Οι κεντρικές τιμές διαφέρουν μεταξύ τους όσο και το πλάτος των κλάσεων, που εδώ είναι ίσο με 6.**
- Μία παρατήρηση που συμπίπτει με το άνω άκρο μιας κλάσης θα**

τοποθετηθεί κατά τη διαλογή στην αμέσως επόμενη κλάση. Για παράδειγμα, ο μαθητής με ύψος 180 θα τοποθετηθεί στην πέμπτη κλάση [180, 186).

Πίνακας 9

Κατανομές συχνοτήτων (απόλυτων, σχετικών, αθροιστικών) για τα δεδομένα του πίνακα 8.

Κλάσεις [-)	Κεντρικές τιμές x_i	Διαλογή	
156-162	159		
162-168	165		
168-174	171		
174-180	177		
180-186	183		
186-192	189		
	Σύνολο	—	

	Συχν. v_i	Σχετική Συχνό- τητα $f_i\%$	Αθρ. συχν. N_i	Αθρ. Σχετ. Συχν. $F_i\%$
	2	5,0	2	5,0
	8	20,0	10	25,0
	12	30,0	22	55,0
	11	27,5	33	82,5
	5	12,5	38	95,0
	2	5,0	40	100,0
	40	100	—	—

Ιστογράμμα Συχνοτήτων

Η αντίστοιχη γραφική παράσταση ενός πίνακα συχνοτήτων με ομαδοποιημένα δεδομένα γίνεται με το λεγόμενο **ιστόγραμμα** (histogram) συχνοτήτων. Στον οριζόντιο άξονα ενός συστήματος ορθογωνίων αξόνων σημειώνουμε, με κατάλληλη κλίμακα, τα όρια των κλάσεων. Στη συνέχεια, κατασκευάζουμε διαδοχικά ορθογώνια (ιστούς), από καθένα από τα οποία έχει βάση ίση με το πλάτος της κλάσης και ύψος τέτοιο, ώστε το εμβαδόν του ορθογωνίου να ισούται με τη συχνότητα της κλάσης αυτής.

α) Κλάσεις Ίσου Πλάτους

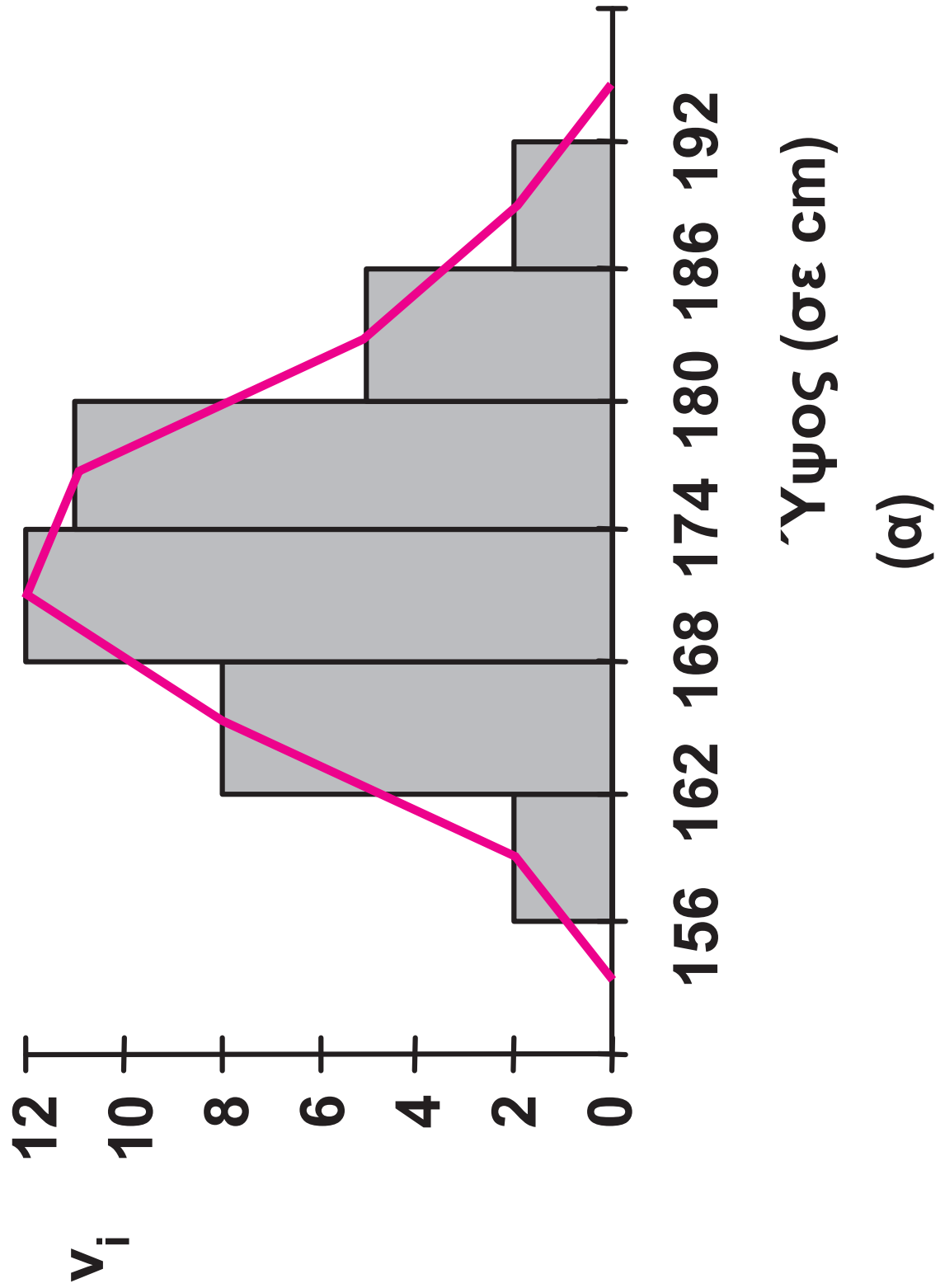
Θεωρώντας το πλάτος c ως μονάδα μέτρησης του χαρακτηριστικού

στον οριζόντιο άξονα, το ύψος κάθε ορθογωνίου είναι ίσο προς τη συχνότητα της αντίστοιχης κλάσης, έτσι ώστε να ισχύει πάλι ότι το εμβαδόν των ορθογωνίων είναι ίσο με τις αντίστοιχες συχνότητες. Επομένως, στον κατακόρυφο άξονα σε ένα ιστόγραμμα συχνοτήτων βάζουμε τις συχνότητες. Με ανάλογο τρόπο κατασκευάζεται και το **ιστόγραμμα σχετικών συχνοτήτων**, οπότε στον κάθετο άξονα βάζουμε τις σχετικές συχνότητες.

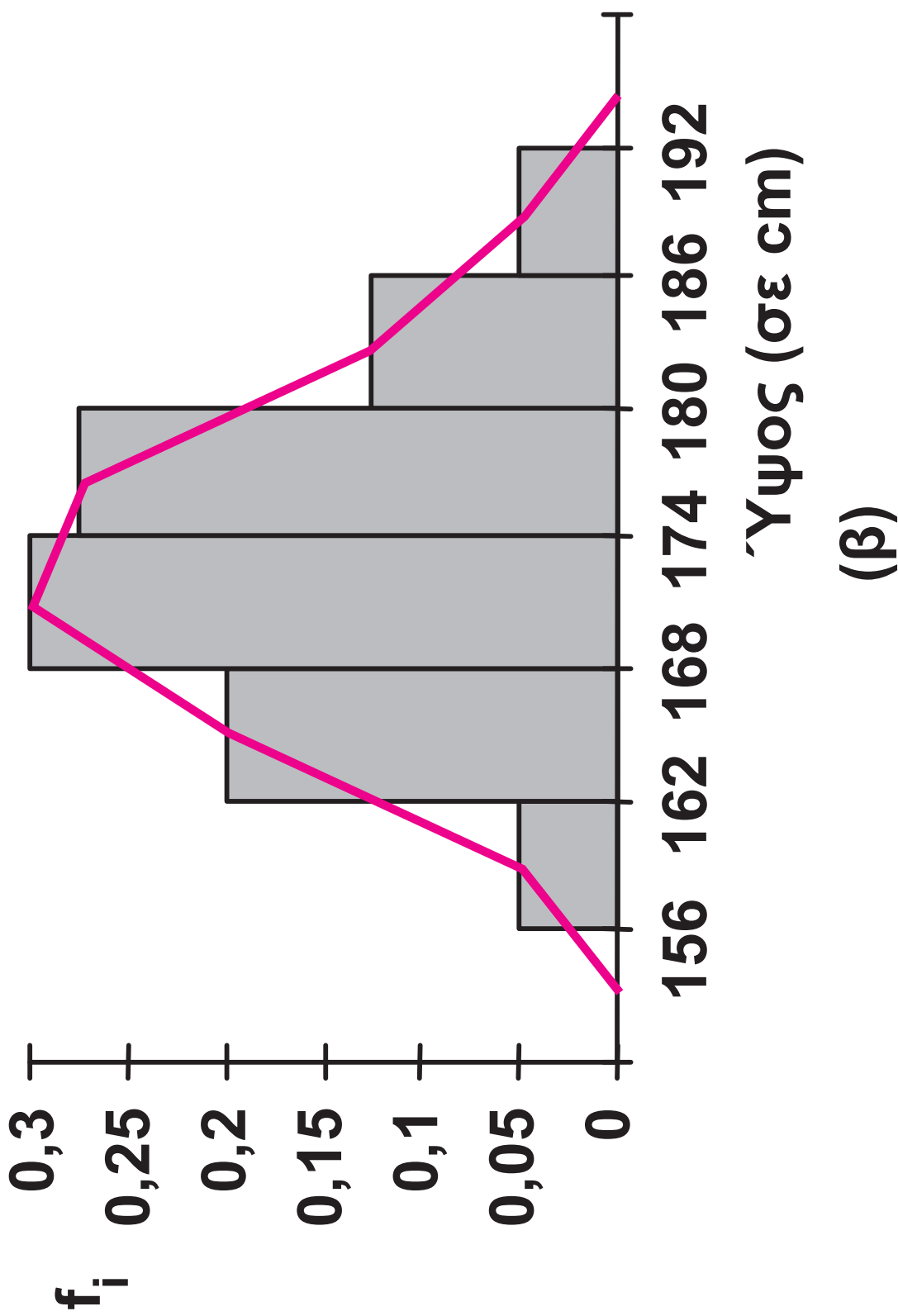
Αν στα ιστογράμματα συχνοτήτων θεωρήσουμε δύο ακόμη υποθετικές κλάσεις, στην αρχή και στο τέλος, με συχνότητα μηδέν και στη συνέχεια ενώσουμε τα μέσα των άνω βάσεων των ορθογωνίων με ευθύγραμμα τμήματα, σχηματίζεται το

Λεγόμενο πολύγωνο συχνοτήτων (frequency polygon). Το εμβαδόν του χωρίου που ορίζεται από το πολύγωνο συχνοτήτων και τον οριζόντιο άξονα είναι ίσο με το άθροισμα των συχνοτήτων, δηλαδή με το μέγεθος του δείγματος n . Όμοια κατασκευάζεται από το ιστόγραμμα σχετικών συχνοτήτων και το πολύγωνο σχετικών συχνοτήτων με εμβαδόν ίσο με 1, (βλέπε σχήμα 6).

6α



102 / 74

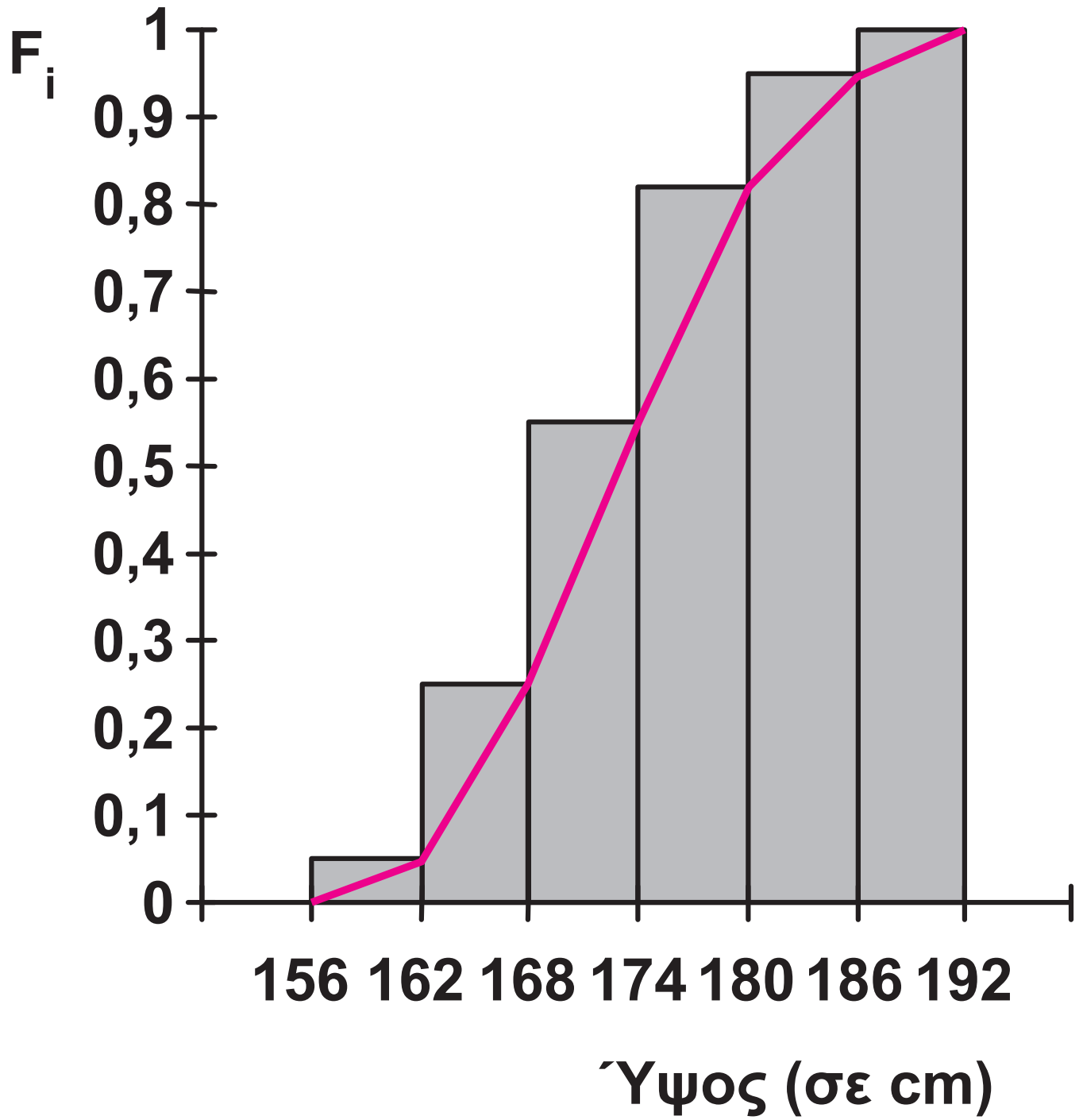


6β

Ιστόγραμμα και πολύγωνο (α) συχνοτήτων και (β) σχετικών συχνοτήτων για τα δεδομένα του πίνακα 9.

Με τον ίδιο τρόπο κατασκευάζονται και τα ιστογράμματα αθροιστικών συχνοτήτων και αθροιστικών σχετικών συχνοτήτων. Αν ενώσουμε σε ένα ιστόγραμμα αθροιστικών συχνοτήτων τα δεξιά άκρα (όχι μέσα) των άνω βάσεων των ορθογωνίων με ευθύγραμμα τμήματα βρίσκουμε το πολύγωνο αθροιστικών συχνοτήτων (ogive) της κατανομής. Στο σχήμα 7 παριστάνεται το ιστόγραμμα και το πολύγωνο αθροιστικών σχετικών συχνοτήτων για το ύψος των μαθητών του πίνακα 9.

7



105 / 74

β) Κλάσεις Άνισου Πλάτους

Όπως προαναφέραμε, συνήθως επιλέγουμε κλάσεις ίσου πλάτους. Υπάρχουν όμως και περιπτώσεις που είναι απαραίτητο να έχουμε κλάσεις διαφορετικού πλάτους όπως, για παράδειγμα, στην κατανάλωση νερού και ηλεκτρικού ρεύματος ή ακόμα και περιπτώσεις όπου οι συχνότητες σε κάποιες κλάσεις να είναι πολύ μικρές οπότε γίνεται συγχώνευση κλάσεων. Έστω, για παράδειγμα, η διάρκεια (σε sec) $n = 80$ τηλεφωνημάτων που έγιναν τυχαία από ένα κινητό τηλέφωνο, η οποία δίνεται στον παρακάτω πίνακα συχνοτήτων.

Διάρκεια τηλεφ. σε sec	Συχνότητα v_i
0-20	20
20-25	20
25-30	24
30-40	16
Σύνολο	$v = 80$

Το αντίστοιχο ιστόγραμμα συχνοτήτων κατασκευάζεται πάλι, έτσι ώστε το εμβαδόν κάθε ορθογωνίου να ισούται με τη συχνότητα της αντίστοιχης κλάσης. Άρα, αν c_i είναι το πλάτος της κλάσης i με συχνότητα v_i , το ύψος του ορθογωνίου θα είναι

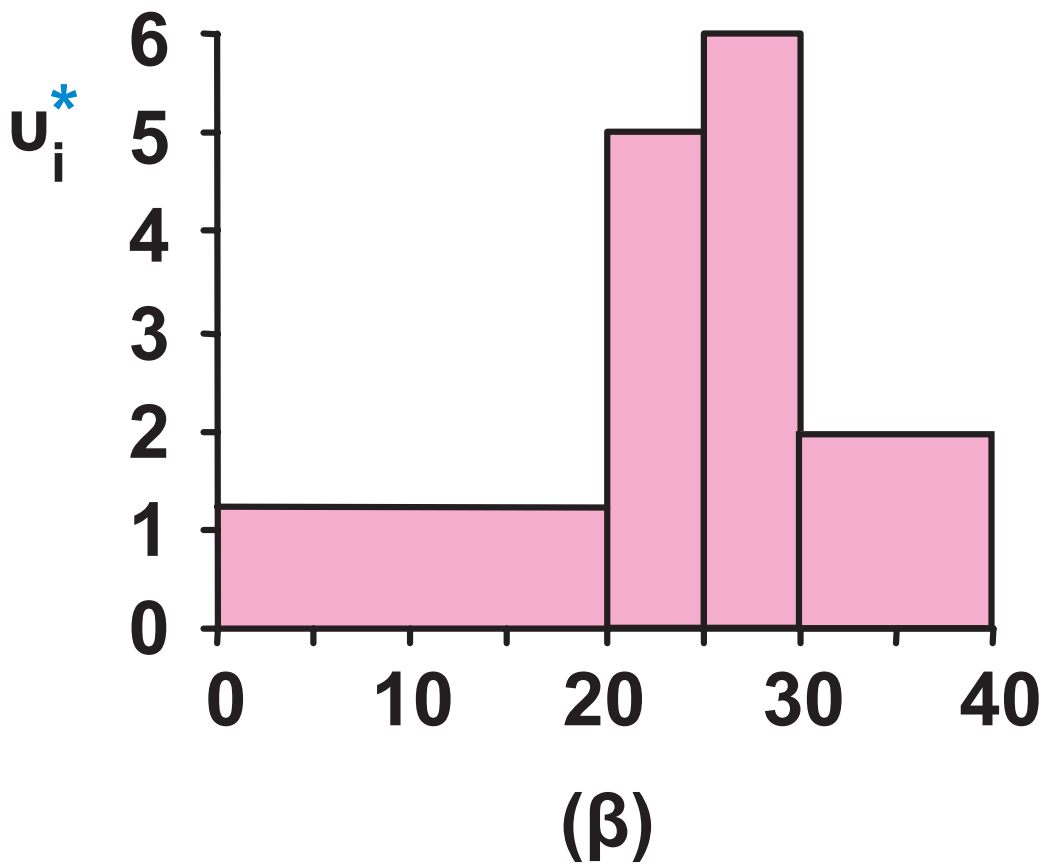
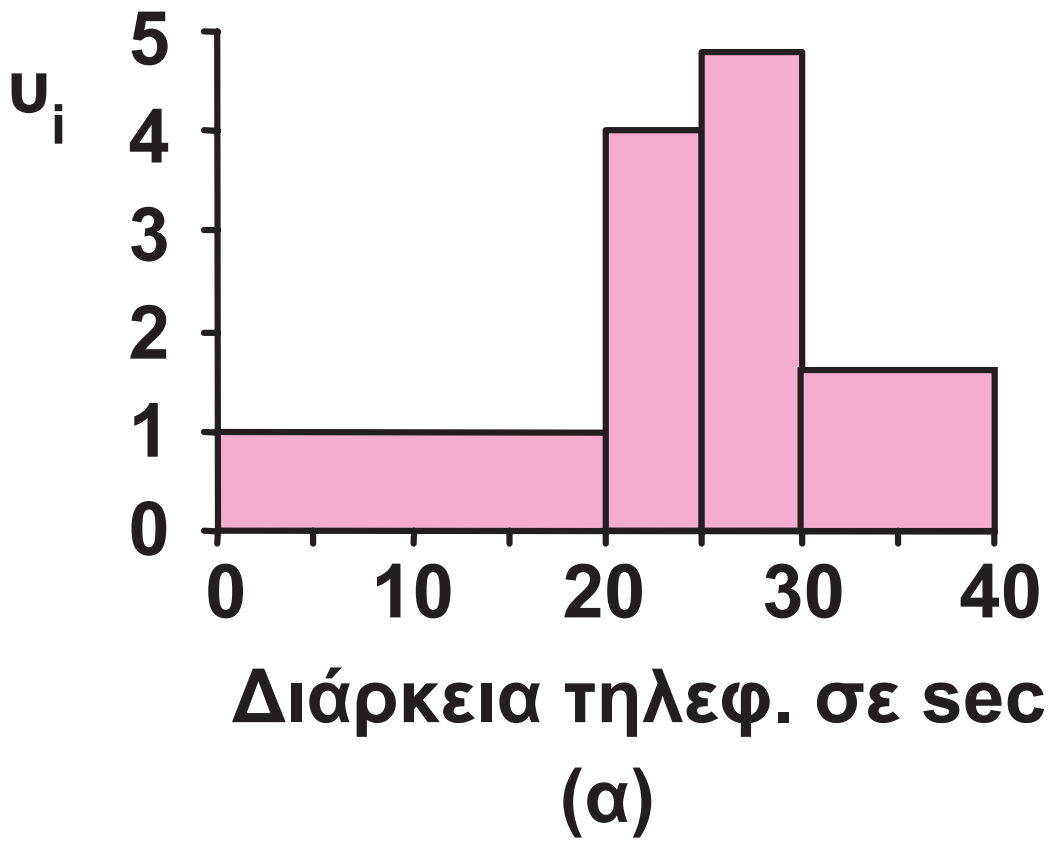
$u_i = \frac{v_i}{c_i}$, $i = 1, 2, \dots, k$. Επομένως, για την κατασκευή του ιστογράμματος

συχνοτήτων χρειαζόμαστε τα πλάτη των κλάσεων και τα ύψη των ορθογωνίων. Αυτά δίνονται στον επόμενο πίνακα.

Διάρκεια τηλεφ. σε sec	Πλάτος κλάσης c_i	Συχνότητα v_i	Ύψος $u_i = \frac{v_i}{c_i}$	Ύψος $u_i^* = \frac{f_i\%}{c_i}$
0-20	20	20	1,0	1,25
20-25	5	20	4,0	5,00
25-30	5	24	4,8	6,00
30-40	10	16	1,6	2,00

Τότε το ιστόγραμμα συχνοτήτων δίνεται στο σχήμα 8(α). Παρατηρούμε ότι το άθροισμα των εμβαδών όλων

των ορθογωνίων είναι ίσο με το συνολικό μέγεθος δείγματος n , όπως δηλαδή συμβαίνει και στο ιστόγραμμα με κλάσεις ίσου πλάτους.



Ιστόγραμμα συχνοτήτων (α) και

110 / 75

σχετικών συχνοτήτων (β) της διάρκειας τηλεφωνημάτων.

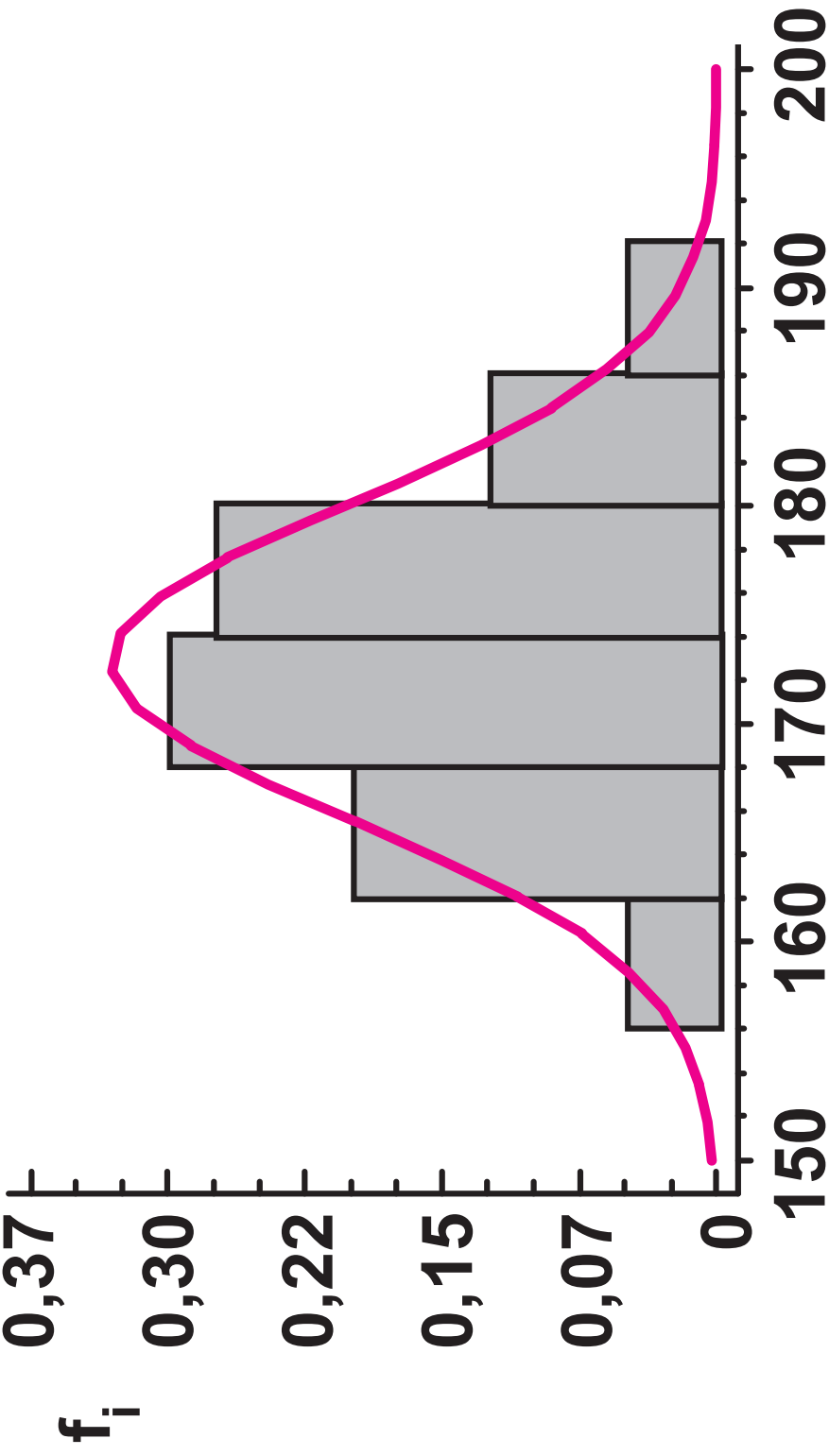
Με ανάλογο τρόπο κατασκευάζεται και το ιστόγραμμα σχετικών συχνοτήτων, (σχήμα 8(β)) αρκεί να χρησιμοποιήσουμε ως ύψος των ορθογωνίων το λόγο των σχετικών συχνοτήτων προς το πλάτος των

κλάσεων, δηλαδή $u_i^* = \frac{f_i \%}{c_i}$.

Καμπύλες Συχνοτήτων

Εάν υποθέσουμε ότι ο αριθμός των κλάσεων για μια συνεχή μεταβλητή είναι αρκετά μεγάλος (τείνει στο άπειρο) και ότι το πλάτος των κλάσεων είναι αρκετά μικρό (τείνει στο μηδέν), τότε η πολυγωνική γραμμή

συχνοτήτων τείνει να πάρει τη μορφή μιας ομαλής καμπύλης, η οποία ονομάζεται καμπύλη συχνοτήτων (frequency curve), όπως δείχνει το σχήμα 9. Οι καμπύλες συχνοτήτων έχουν μεγάλη εφαρμογή στη Στατιστική, όπου οι ιδιότητες τους μπορούν να χρησιμοποιηθούν για την εξαγωγή χρήσιμων συμπερασμάτων.

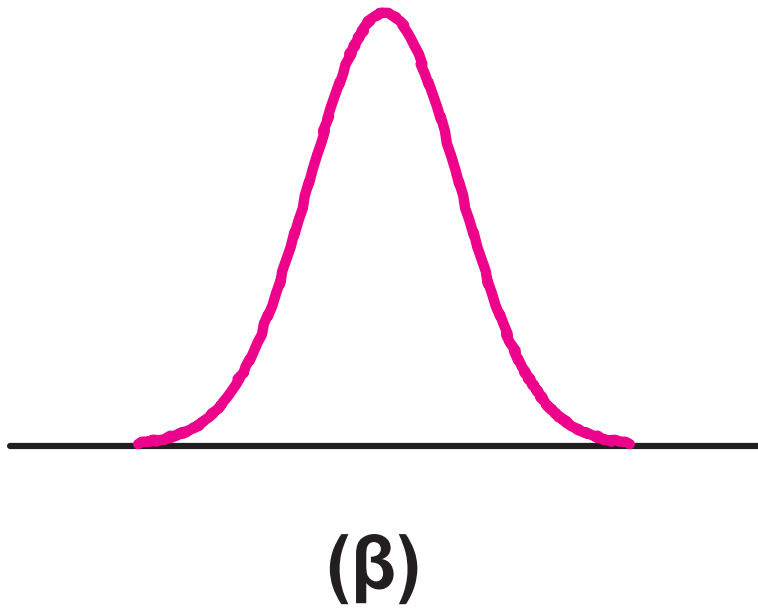
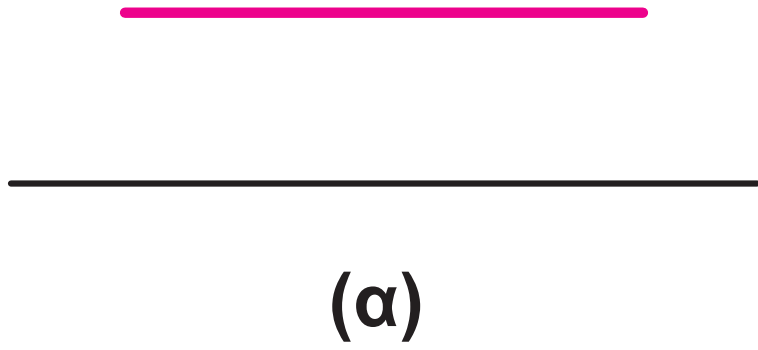


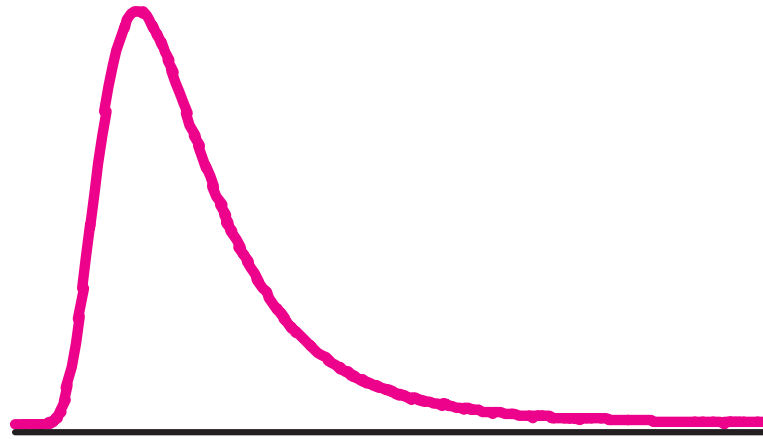
Υψος (σε cm)

Καμπύλη συχνοτήτων για το ύψος των μαθητών του
πίνακα 4

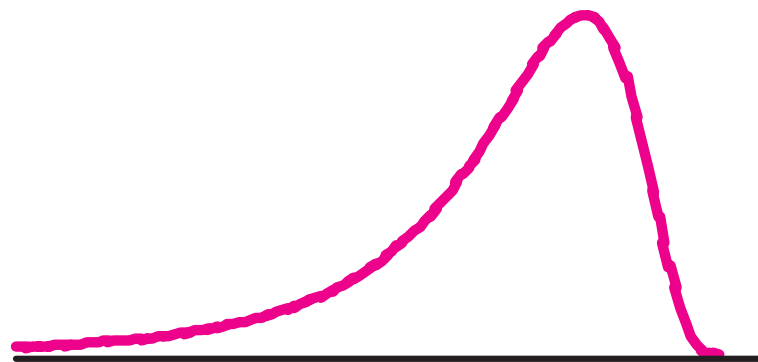
9

Η μορφή μιας κατανομής συχνοτήτων εξαρτάται από το πώς είναι κατανεμημένες οι παρατηρήσεις σε όλη την έκταση του εύρους τους. Μερικές χαρακτηριστικές καμπύλες συχνοτήτων που συναντάμε συχνά στις εφαρμογές δίνονται στο σχήμα 10. Η κατανομή (β), με “κωδωνοειδή” μορφή λέγεται κανονική κατανομή (normal distribution) και παίζει σπουδαίο ρόλο στη Στατιστική. Όταν οι παρατηρήσεις “κατανέμονται” ομοιόμορφα σε ένα διάστημα $[\alpha, \beta]$, όπως στην κατανομή (α), η κατανομή λέγεται ομοιόμορφη. Όταν οι παρατηρήσεις δεν είναι συμμετρικά κατανεμημένες, η κατανομή λέγεται ασύμμετρη με θετική ασυμμετρία όπως στην κατανομή (γ) ή αρνητική ασυμμετρία όπως στην κατανομή (δ).





(γ)



(δ)

Μερικές χαρακτηριστικές κατανομές συχνοτήτων

ΕΦΑΡΜΟΓΕΣ

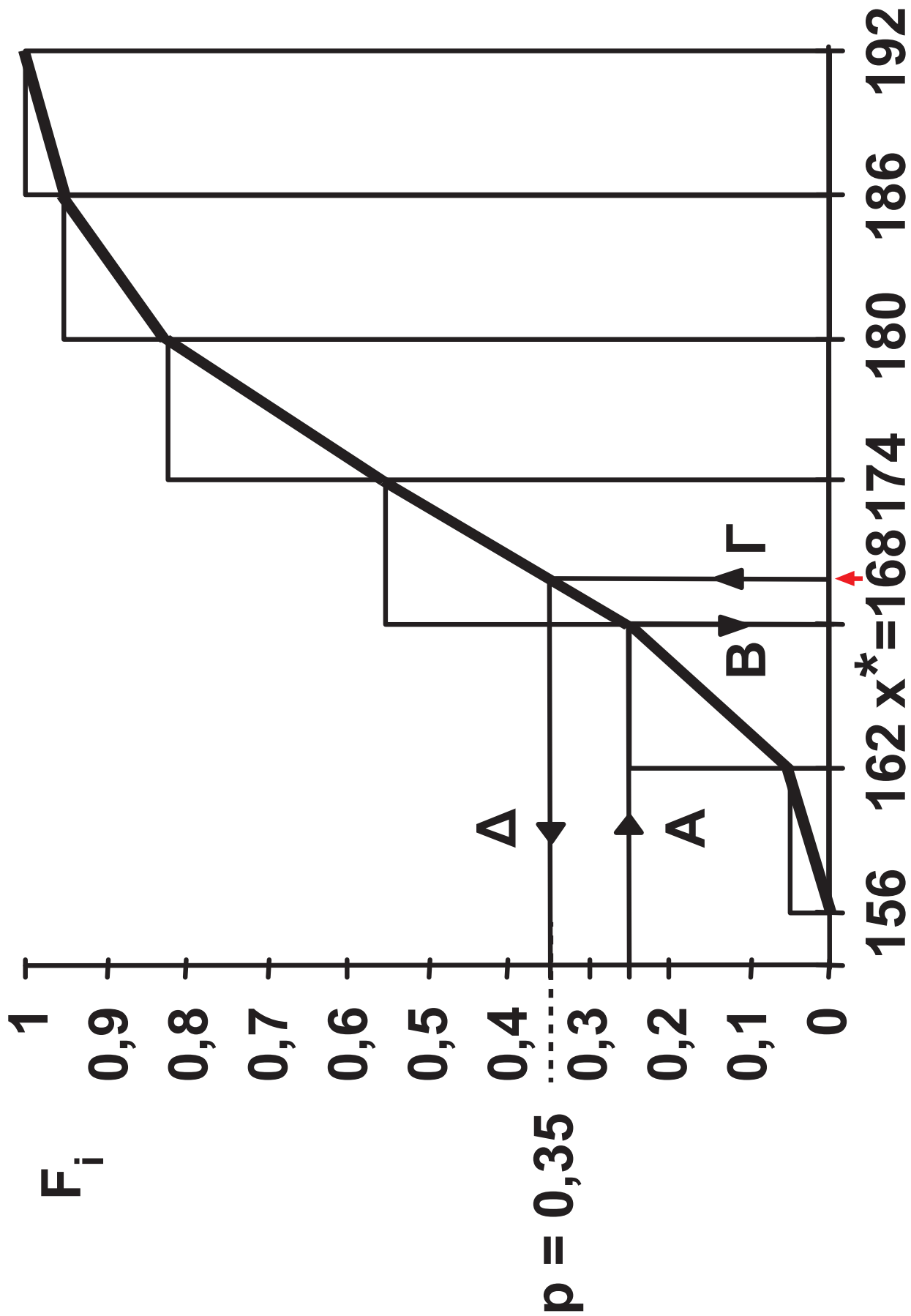
1. Από το πολύγωνο σχετικών αθροιστικών συχνοτήτων του παρακάτω διαγράμματος να βρεθεί

α) το ύψος x^* , κάτω από το οποίο ανήκει το 25% των μαθητών

β) το ποσοστό p των μαθητών που έχουν ύψος μέχρι και 170 cm.

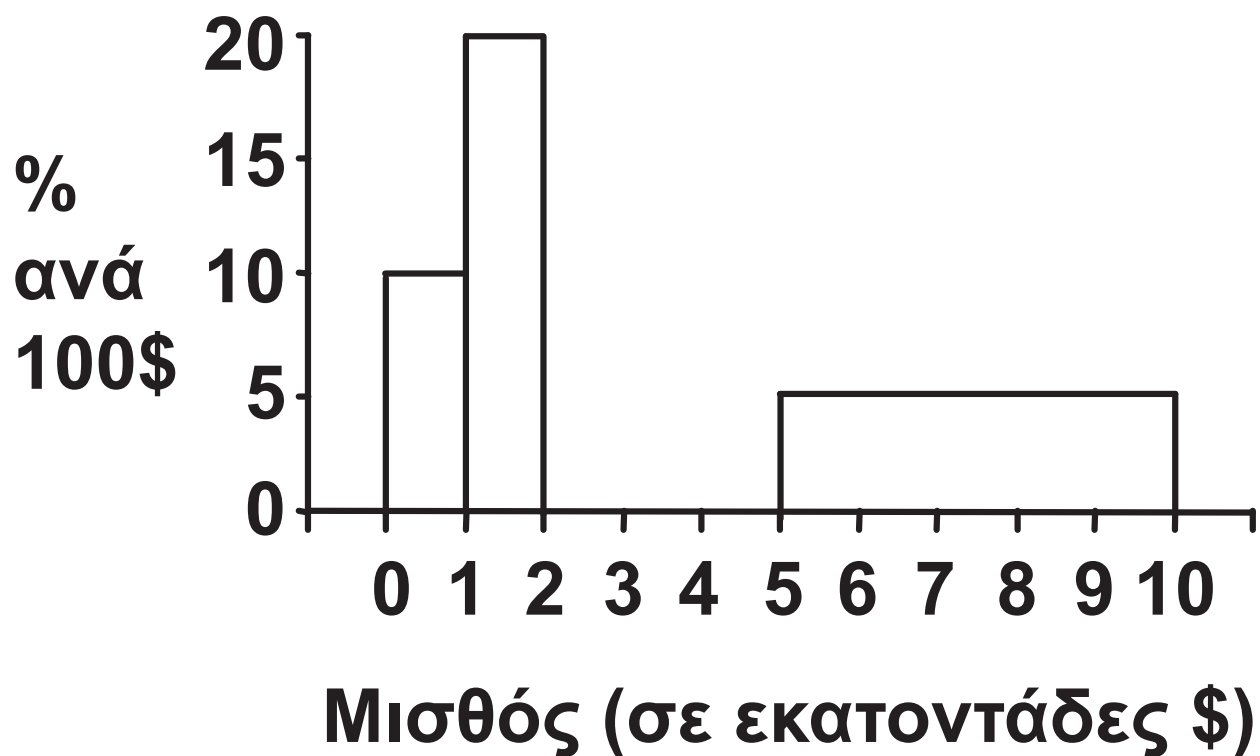
ΛΥΣΗ

α) Ακολουθούμε τη διαδρομή AB, όπως φαίνεται στο διάγραμμα, και ξεκινώντας από το σημείο (0, 0,25) πηγαίνουμε παράλληλα προς τον άξονα $0x$ μέχρι το αθροιστικό διάγραμμα και μετά κάθετα στον άξονα $0x$ μέχρι το σημείο (x^* , 0). Το $x^* = 168$ είναι το ζητούμενο ύψος.



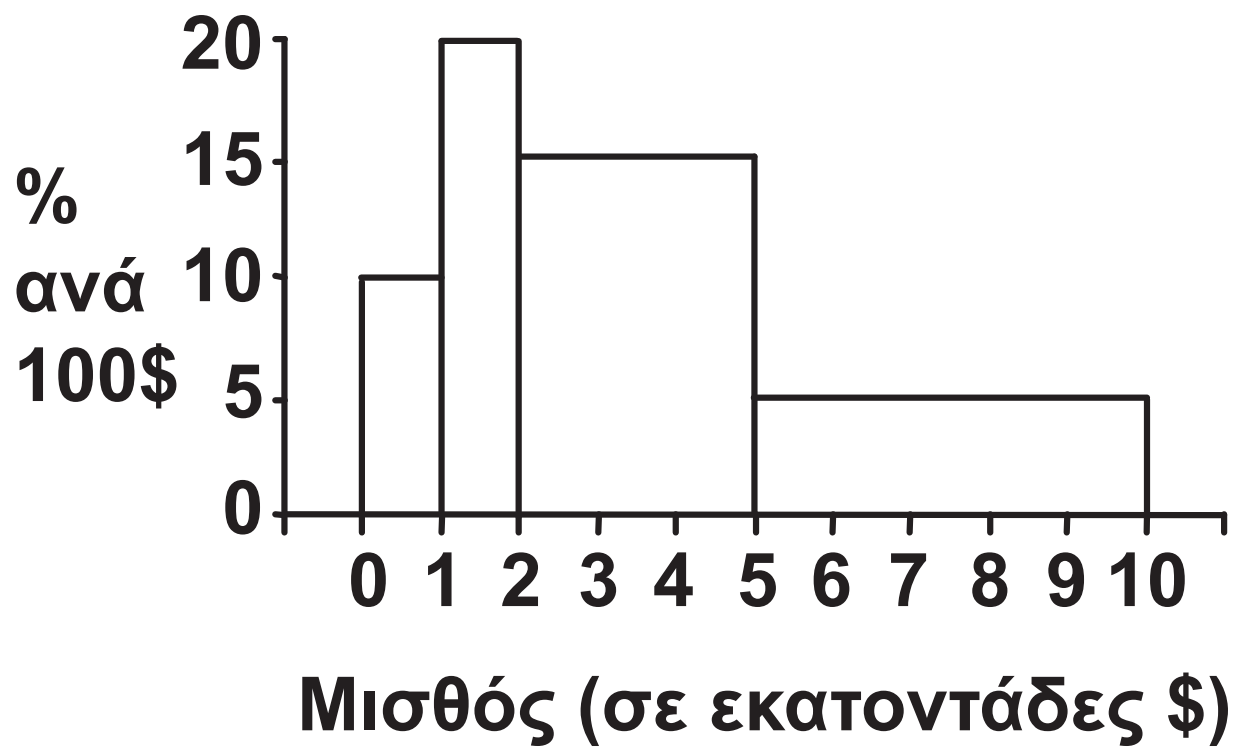
β) Όμοια, ακολουθώντας τη διαδρομή $\Gamma\Delta$ από το σημείο $(170, 0)$ καταλήγουμε, όπως φαίνεται στο σχήμα, στο σημείο $(0, p)$. Το $p = 0,35 = 35\%$ είναι το ζητούμενο ποσοστό.

2. Στο παρακάτω ιστόγραμμα σχετικών συχνοτήτων σβήστηκε κατά λάθος το ορθογώνιο της κλάσης [2-5). Εάν είναι γνωστό ότι δεν υπάρχει μισθός άνω των \$1000, να κατασκευάσετε το ορθογώνιο αυτό.



ΛΥΣΗ

Επειδή έχουμε ένα ιστόγραμμα σχετικών συχνοτήτων ($f_i\%$), το άθροισμα των εμβαδών όλων των ορθογωνίων θα πρέπει να ισούται με 100. Το εμβαδόν του πρώτου ορθογωνίου είναι $E_1 = (1 - 0) \cdot 10 = 10$, του δεύτερου ορθογωνίου $E_2 = (2 - 1) \cdot 20 = 20$, και του τέταρτου $E_4 = (10 - 5) \cdot 5 = 25$. Άρα, το εμβαδόν του τρίτου ορθογωνίου θα είναι $E_3 = 100 - (10 + 20 + 25) = 45$. Επειδή το πλάτος του ορθογωνίου είναι $5 - 2 = 3$, το ύψος του θα είναι $\frac{45}{3} = 15$ όπως φαίνεται στο παρακάτω σχήμα.



ΑΣΚΗΣΕΙΣ

Α' ΟΜΑΔΑΣ

1. Η βαθμολογία 50 φοιτητών στις εξετάσεις ενός μαθήματος είναι:

3	4	5	8	9	7	6	8	7	10
8	7	6	5	9	3	8	5	6	6
6	3	5	6	4	2	9	8	7	7
1	6	3	1	5	8	1	2	3	4
5	6	7	9	10	9	8	7	6	5

α) Να κατασκευάσετε τον πίνακα κατανομής συχνοτήτων και σχετικών συχνοτήτων (απολύτων και αθροιστικών).

β) Από τον πίνακα αυτό να εκτιμήσετε το ποσοστό των φοιτητών που πήραν βαθμό

i) κάτω από τη βάση (μικρότερο του 5)

ii) άριστα (9 ή 10)

iii) τουλάχιστον 7 αλλά το πολύ 9.

2. Οι παραπάνω φοιτητές ήταν αντίστοιχα αγόρια (Α) ή κορίτσια (Κ):

A	A	K	A	K	A	A	A	K	K
K	K	A	A	A	K	A	K	A	A
A	A	A	A	K	K	A	K	A	K
K	K	K	A	K	K	A	A	A	A
A	A	K	A	K	K	A	A	A	K

Να συμπληρώσετε τον επόμενο πίνακα χρησιμοποιώντας απόλυτες συχνότητες.

Φύλο	Βαθμολογία		Σύνολο	
	≤ 5	> 5		
A				
K				
Σύνολο				

3. Να μετατρέψετε τον προηγούμενο πίνακα συχνοτήτων της άσκησης 2 σε πίνακα σχετικών συχνοτήτων επί τοις εκατό:

α) ως προς το σύνολο των φοιτητών

β) ως προς το φύλο (γραμμές)

γ) ως προς τη βαθμολογία (στήλες) και να ερμηνεύσετε τα αποτελέσματα.

4. Χρησιμοποιώντας τον παρακάτω πίνακα συχνοτήτων, που δίνει την κατανομή του αριθμού των ημερών απουσίας από την εργασία τους λόγω ασθένειας 50 εργατών, να βρεθεί ο αριθμός και το ποσοστό των εργατών που απουσίασαν:

- α) τουλάχιστον 1 ημέρα**
- β) πάνω από 5 ημέρες**
- γ) από 3 έως 5 ημέρες**
- δ) το πολύ 5 ημέρες**
- ε) ακριβώς 5 ημέρες.**

Αριθ- μός ημε- ρών	Συχνό- τητα	Αριθ- μός ημε- ρών	Συχνό- τητα
0	12	5	8
1	8	6	0
2	5	7	5
3	4	8	2
4	5	9	1

5. Να συμπληρώσετε τον παρακάτω πίνακα.

x_i	v_i	f_i	N_i	F_i	$f_i\%$	$F_i\%$
1						10
2	4	0,20	6			
3				0,60		
4					25	
5	2					
6						
Σύνολο						

6. Να κατασκευάσετε το διάγραμμα συχνοτήτων του βαθμού Μαθηματικών για τα αγόρια και κορίτσια (χωριστά) του πίνακα 4.

7. Τα δημοφιλέστερα ξένα μουσικά συγκροτήματα των 18 αγοριών του πίνακα 4 ήσαν: Metallica, Iron Maiden, Άλλο, Scorpions, Oasis, Άλλο, Άλλο, Rolling Stones, Metallica, Metallica, Rolling Stones, Metallica, Iron Maiden, Iron Maiden, Scorpions, Scorpions, Scorpions, Metallica.

Να κατασκευάσετε α) το ραβδόγραμμα και β) το κυκλικό διάγραμμα σχετικών συχνοτήτων.

8. Σε ένα κυκλικό διάγραμμα παριστάνεται η βαθμολογία των 450 μαθητών ενός Γυμνασίου σε τέσσερις κατηγορίες “Άριστα”, “Λίαν Καλώς”, “Καλώς” και “Σχεδόν Καλώς”. Το 30% των μαθητών έχουν επίδοση “Λίαν Καλώς”. Η γωνία του κυκλικού τομέα για την επίδοση “Καλώς” είναι 144° . Οι μαθητές με βαθμό “Σχεδόν Καλώς” είναι διπλάσιοι των μαθητών με “Άριστα”. Να μετατρέψετε το κυκλικό διάγραμμα σε ραβδόγραμμα σχετικών συχνοτήτων. Πόσοι μαθητές έχουν επίδοση τουλάχιστον λίαν καλώς;

9. Από το 1960 έως το 1998 (Πρωταθλήματα Α΄ Εθνικής) ο Παναθηναϊκός έχει κατακτήσει 15 τίτλους, ο Ολυμπιακός 12, η ΑΕΚ 9, ο ΠΑΟΚ 2 και η Λάρισα 1. Να κατασκευάσετε το ραβδόγραμμα και το κυκλικό διάγραμμα σχετικών συχνοτήτων.

10. Παρακάτω δίνονται τα μετάλλια που πήραν μερικές χώρες στο 17ο Ευρωπαϊκό Πρωτάθλημα Στίβου, το 1998. Να παρασταθούν τα δεδομένα αυτά σε ένα ραβδόγραμμα.

Χώρα	Χρυ- σά	Αση- μένια	Χάλ- κινα
Μ. Βρετανία	9	4	3
Γερμανία	8	7	8
Ρωσία	6	9	7
Πολωνία	3	4	1
Ρουμανία	3	2	2
Ουκρανία	3	2	1
Ιταλία	2	4	3
Πορτογαλία	2	3	1
Ισπανία	2	1	4
Γαλλία	2	1	1
Ελλάδα	1	0	2

11. Τα κρούσματα δύο λοιμωδών νόσων από το 1987 έως το 1997 δίνονται στον παρακάτω πίνακα. (Πηγή: ΕΚΕ-ΠΑΠ.)

Να κατασκευάσετε τα αντίστοιχα χρονογράμματα και να τα σχολιάσετε.

Έτος	Έρπης ζωστήρ	Ηπατίτιδα Α
1987	85	351
1988	58	254
1989	123	273
1990	178	172
1991	134	213
1992	201	127
1993	241	123
1994	252	259
1995	338	295
1996	296	107
1997	256	131

12. Τα παρακάτω δεδομένα αντιπροσωπεύουν την επίδοση 50 υποψηφίων για την πρόσληψή τους σε μια ιδιωτική σχολή (κλίμακα 0-10).

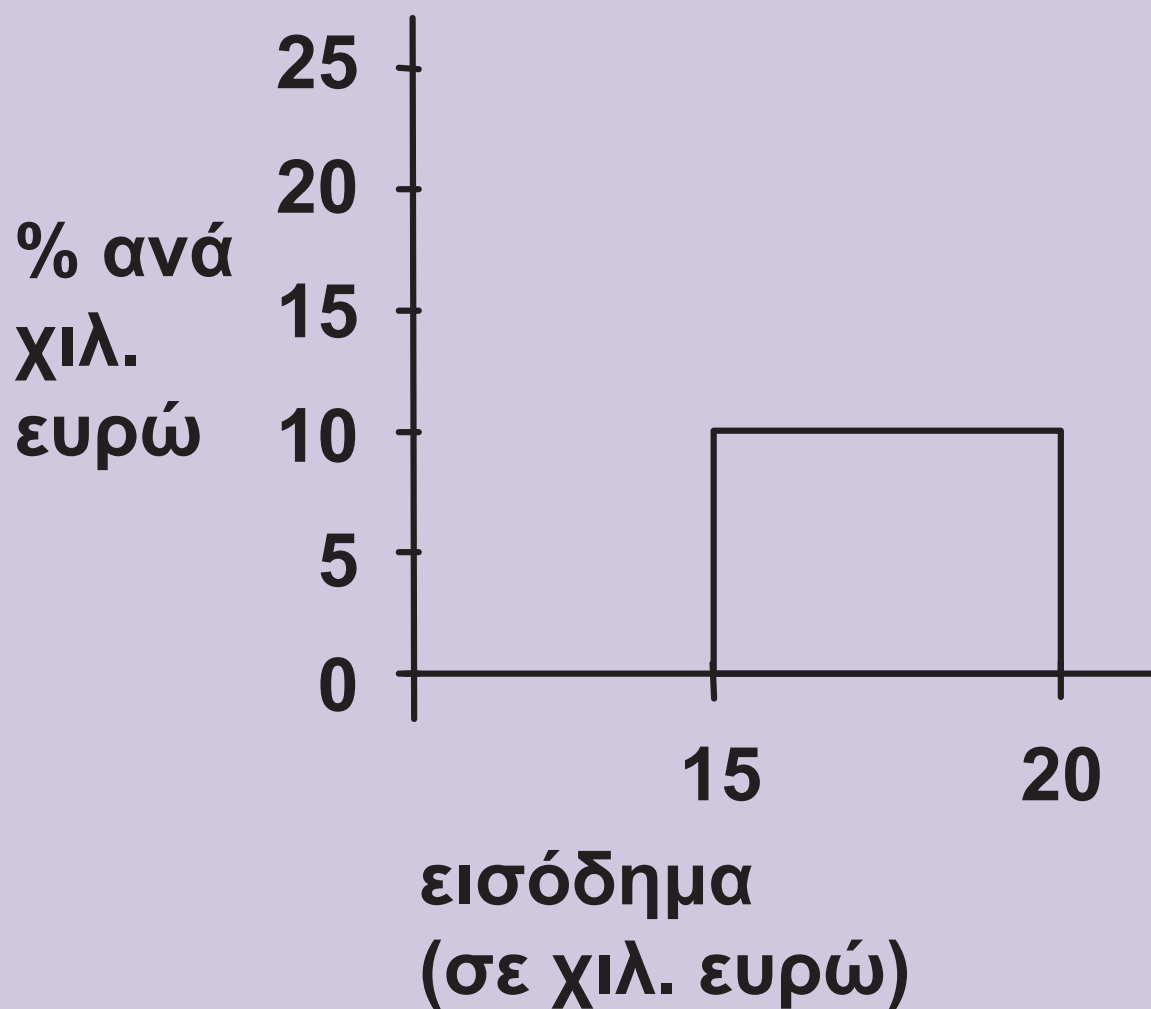
6 7 8 5 1 4 7 3 9 9 2 5 3 8 6
7 7 6 8 1 3 0 1 4 9 0 9 7 8 6
1 2 3 5 4 6 6 4 3 2 8 8 7 7 6
5 5 9 2 4

- α) Να παραστήσετε τα δεδομένα σε έναν πίνακα συχνοτήτων.
- β) Να κατασκευάσετε το διάγραμμα σχετικών και αθροιστικών σχετικών συχνοτήτων.
- γ) Αν η σχολή θελήσει να πάρει όσους είχαν επίδοση

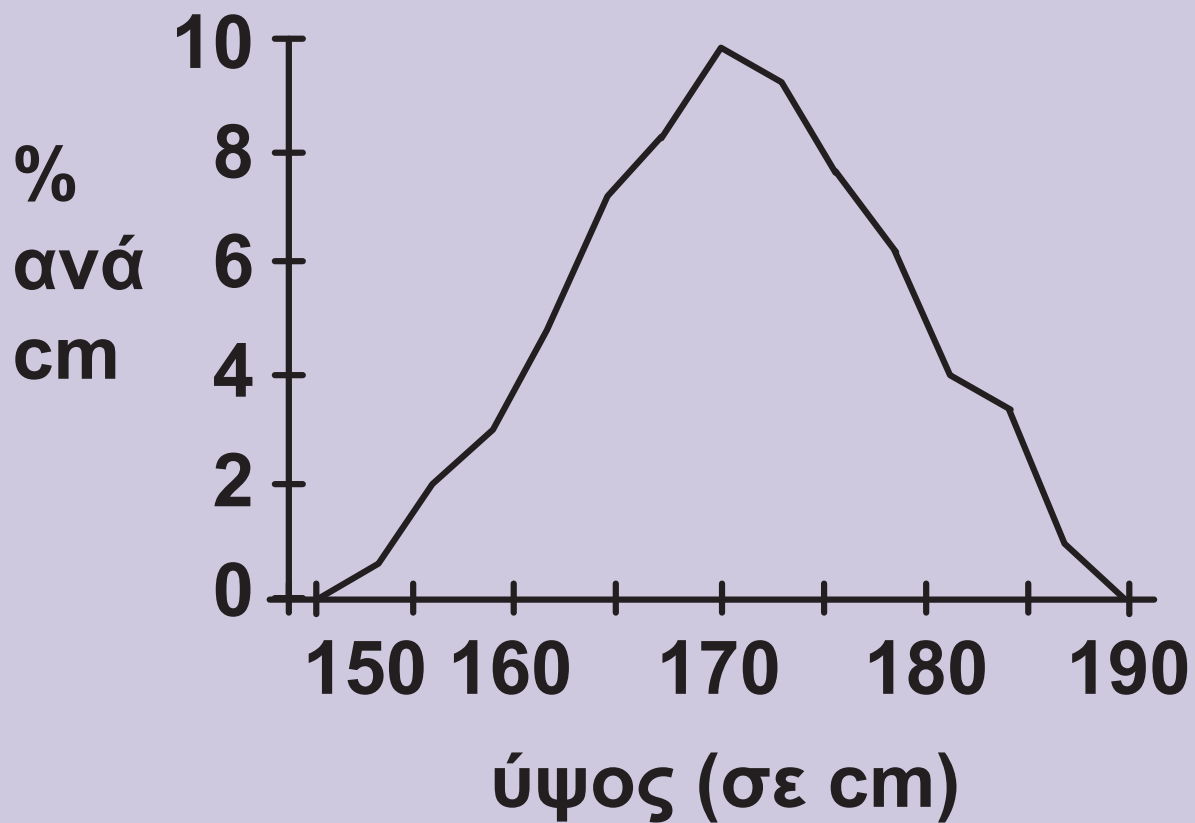
μεγαλύτερη ή ίση του 8, πόσους θα πάρει;

δ) Αν η σχολή πάρει μόνο το 36% των υποψηφίων, τι επίδοση πρέπει να έχει κάποιος για να επιλεγεί;

13. Παρακάτω δίνεται μόνο ένα ορθογώνιο από το ιστόγραμμα του ετήσιου εισοδήματος των οικογενειών μιας περιοχής. Τι ποσοστό οικογενειών της περιοχής είχαν εισόδημα 15.000 ευρώ έως 20.000 ευρώ;



14. Ένας μαθητής έκανε το παρακάτω πολύγωνο σχετικών συχνοτήτων για το ύψος των αγοριών της τάξης του και ο καθηγητής το διέγραψε σαν λάθος. Είχε δίκιο ο καθηγητής;



Β' ΟΜΑΔΑΣ

1. Να κατασκευάσετε τα αντίστοιχα χρονογράμματα για τον πληθυσμό των νησιών α) Λέσβου, β) Θάσου, γ) Σαλαμίνας με βάση τα δεδομένα του πίνακα 2. Τι συμπέρασμα συνάγετε;

2. Οι βεβαιωθέντες θάνατοι από χρήση ναρκωτικών κατά τα έτη 1988-1998 (για το 1998 έως 8 Απριλίου) σύμφωνα με τον Οργανισμό κατά των Ναρκωτικών (ΟΚΑΝΑ) ήταν 62, 72, 66, 79, 79, 78, 146, 176, 222, 222 και 65 αντίστοιχα. Από αυτούς είχαμε 7, 4, 2, 2, 1, 4, 8, 7, 14, 22 και 6 μέχρι και 20 ετών, 43, 51, 34, 44, 47, 49, 71, 90, 98, 99 και 33 από 21-30 ετών και οι υπόλοιποι ήσαν άνω των 30 ετών.

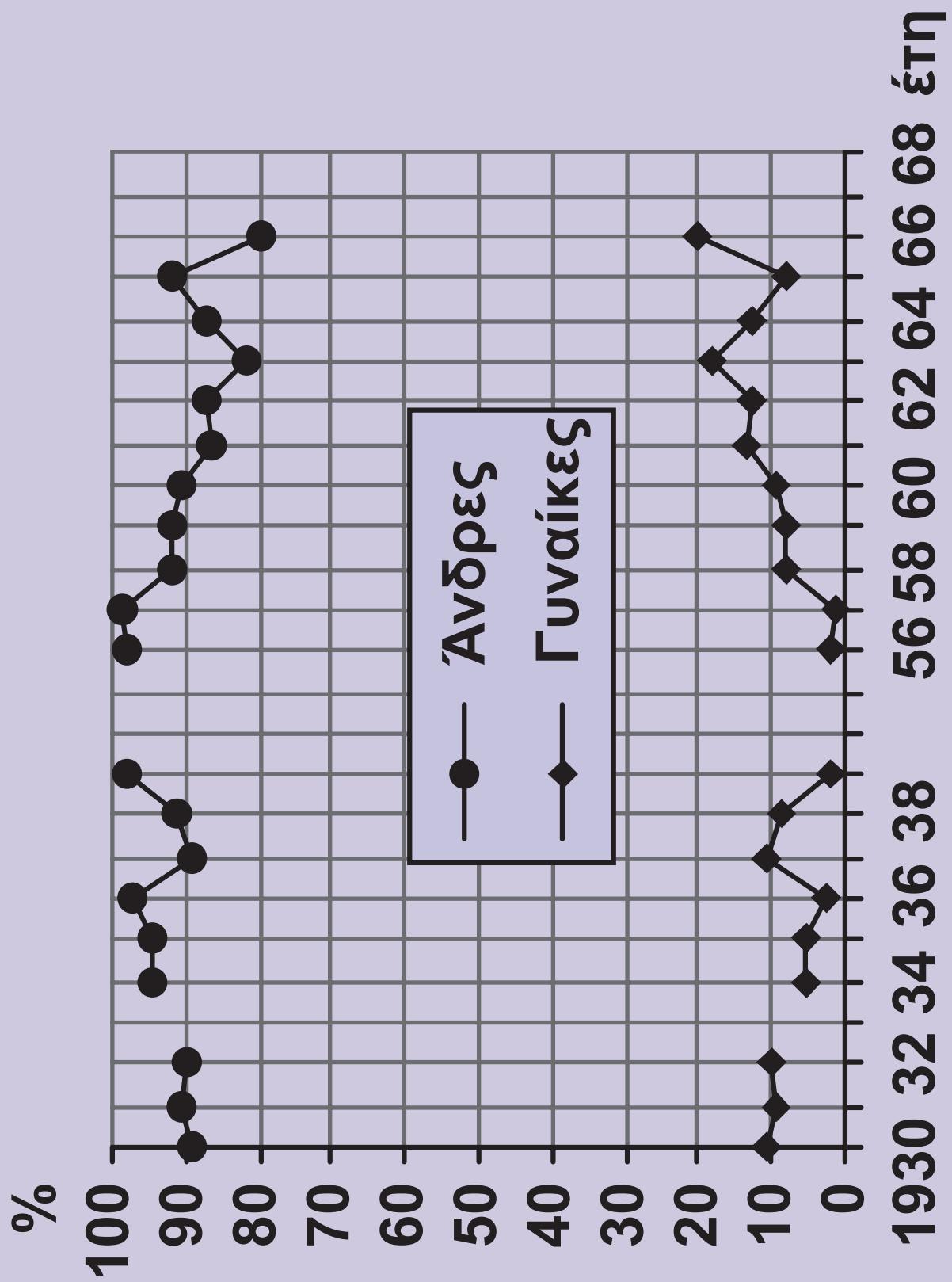
Να παρασταθούν τα δεδομένα αυτά σε μορφή πίνακα.

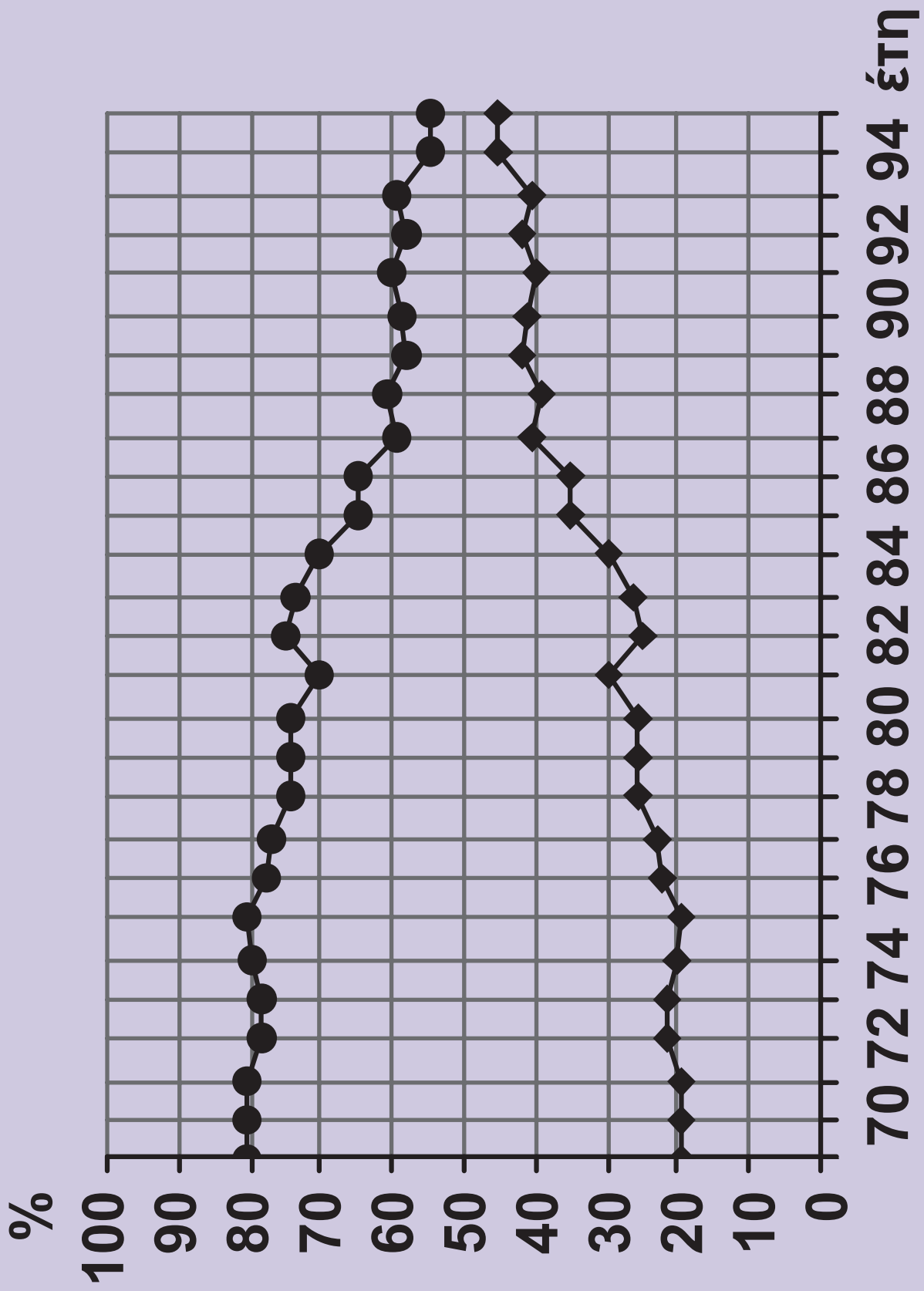
3. Να παρασταθούν τα παραπάνω δεδομένα της άσκησης 2 σε μορφή πίνακα αναφορικά με το έτος και το φύλο των ατόμων, αν γνωρίζουμε ότι από τους βεβαιωθέντες θανάτους από χρήση ναρκωτικών κατά τα έτη 1988-1998 οι 8, 10, 7, 5, 9, 8, 11, 14, 20, 20 και 9 αντίστοιχα ήταν γυναίκες.

4. Το παρακάτω χρονόγραμμα δίνει τη σχετική συχνότητα των νέων πτυχιούχων Μαθηματικών σε όλη την Ελλάδα από το 1930 έως το 1995 ανάλογα με το φύλο. α) Μελετώντας προσεκτικά το χρονόγραμμα αυτό ποια συμπεράσματα εξάγονται; β) Ο συνολικός αριθμός νέων πτυχιούχων Μαθηματικών το έτος 1995 ήταν 789. Πόσες ήσαν οι γυναίκες και πόσοι οι άνδρες; γ) Ο αριθμός των γυναικών που έγιναν πτυχιούχοι Μαθηματικών το έτος 1974 ήσαν 173. Πόσοι ήσαν οι άνδρες που έγιναν πτυχιούχοι Μαθηματικοί το ίδιο έτος; δ) Πόσοι

**άνδρες και πόσες γυναίκες
πήραν πτυχίο Μαθηματικών
στην Ελλάδα το 1985;**

143 / 82





5. Να δοθεί και να ερμηνευτεί το χρονόγραμμα των δεδομένων του πίνακα 1 για κάθε ομάδα ηλικιών.

6. Στον παρακάτω πίνακα δίνεται η κατανομή συχνοτήτων της συστολικής πίεσης 150 γυναικών ηλικίας 17-24 ετών που χρησιμοποιούν το φάρμακο A για κάποια πάθηση και 200 γυναικών, ανάλογης ηλικίας, που χρησιμοποιούν το φάρμακο B.

α) Να συγκρίνετε τα ποσοστά γυναικών που παίρνουν τα φάρμακα A και B και έχουν συστολική πίεση μεγαλύτερη ή ίση των 130 mm Hg

β) Να κατασκευάσετε τα πολύγωνα αθροιστικών σχετικών συχνοτήτων, χρησιμοποιώντας τους ίδιους άξονες συντεταγμένων.

Συστολική πίεση (σε mm Hg)	Φάρμακο A	Φάρμακο B
	v_i	v_i
95-99	6	4
100-104	15	14
105-109	16	18
110-114	22	24
115-119	30	32
120-124	20	28
125-129	15	28
130-134	12	26
135-139	6	12
140-144	5	8
145-149	3	6
Σύνολο	150	200

Πηγή: Υποθετικά δεδομένα

7. Οι χρόνοι (σε λεπτά) που χρειάστηκαν 55 μαθητές να λύσουν ένα πρόβλημα δίνονται παρακάτω:

3,4	4,4	3,3	1,4	3,1
13,2	3,6	9,7	3,5	8,7
6,7	1,4	2,0	2,8	2,8
1,4	2,4	4,4	10,4	3,8
1,3	3,6	3,3	11,9	13,0
3,8	3,1	8,7	12,3	3,0
3,9	7,5	3,9	2,9	6,4
2,9	6,9	11,6	2,8	3,2
13,8	7,8	5,6	1,5	5,9
3,9	12,7	9,0	4,1	7,0
2,7	3,9	3,4	5,9	8,2

α) Να ομαδοποιήσετε τα δεδομένα σε κατάλληλο αριθμό κλάσεων.

- β) Να κατασκευάσετε τον πίνακα με τις συχνότητες n_i , $f_i\%$, N_i , $F_i\%$.**
- γ) Να κατασκευάσετε το πολύγωνο σχετικών συχνοτήτων και αθροιστικών σχετικών συχνοτήτων.**

ΥΠΟΔΕΙΞΕΙΣ – ΑΠΑΝΤΗΣΕΙΣ ΑΣΚΗΣΕΩΝ

2 ΣΤΑΤΙΣΤΙΚΗ

§ 2.1

1. Ποιοτικές: $\gamma, \delta, \sigma\tau, \zeta$
Ποσοτικές-διακριτές: $\beta, \eta, \theta, \iota$
Ποσοτικές-συνεχείς: α, ϵ
2. Είναι δυνατόν να έχουμε διάφορες μεταβλητές για κάθε περίπτωση. Για παράδειγμα:
α) μισθός (ποσοτική-συνεχής),
ηλικία (ποσοτική-διακριτή), φύλο (ποιοτική) κτλ.
3. (ϵ).

4. Είναι δυνατό να έχουμε διάφορους λόγους ακαταλληλότητας των επιλεγομένων δειγμάτων. Για παράδειγμα:
- α) Θα έχουμε υπερεκτίμηση των ανδρών κτλ.

§ 2.2 Α' Ομάδας

1. β) i) 26% ii) 14% iii) 38%.
2. Έχουμε 11 αγόρια με βαθμό < 5 , κτλ.
3. α) Το 22% των φοιτητών είναι αγόρια με βαθμό < 5 , κτλ.
4. α) 76% β) 16% γ) 34% δ) 84%
ε) 16%.

5. Από τη σχέση $f_2 = \frac{v_2}{v}$ βρίσκουμε πρώτα το μέγεθος $v = 20$. Να βρεθεί μετά το N_1 κτλ.
6. Να εργαστείτε όπως στο σχήμα 1(γ).
7. Να κατασκευάσετε πρώτα τον πίνακα συχνοτήτων και μετά να εργαστείτε όπως στα σχήματα 1(β) και 3.
8. Για τους $v = 450$ μαθητές έχουμε 30% με τιμή $x_2 = \text{“Λίαν καλώς”}$.
Δηλαδή
 $f_2\% = 30 \Leftrightarrow v_2 = 135 \Leftrightarrow \alpha_2 = 108^\circ$,
κτλ.
9. Να εργαστείτε όπως στα σχήματα 1(β) και 3.

10. Να εργαστείτε όπως στο σχήμα 1(γ).
11. Να εργαστείτε όπως στο σχήμα 5.
12. γ) 11 άτομα
δ) επίδοση ≥ 7 .
13. 50%.
14. ΝΑΙ. Το εμβαδόν πρέπει να είναι 100.

§ 2.2 Β' Ομάδας

1. Για τη Λέσβο υπάρχει πτωτική τάση ενώ για τη Σαλαμίνα υπάρχει ανοδική τάση.

Για τη Θάσο υπάρχει περίπου σταθερή κατάσταση.

2. Να συμπληρώσετε τον πίνακα:

Έτος	Ηλικία			Σύνολο
	≤ 20	21 - 30	≥ 31	

3. Ανάλογα με την άσκηση 2.

4. β) 355 γυναίκες, 434 άνδρες

γ) 692

δ) Δεν μπορούμε να ξέρουμε με τα στοιχεία που μας δίνονται.

5. Να εργαστείτε όπως στο σχήμα 5.

6. α) Φάρμακο A 17,3% και Φάρμακο B 26%.

7. Επειδή $n = 55$ να χρησιμοποιήσετε $k = 7$ κλάσεις με πλάτος $c = 1,8$.

Ευρετήριο Όρων

Στο Ευρετήριο όρων τα γράμματα Α, Β, Γ, Δ και Ε δηλώνουν τον 1ο, 2ο, 3ο, 4ο και 5ο τόμο αντίστοιχα, ενώ οι αριθμοί αναφέρονται στην πρώτη από τις δύο ενδείξεις που αναγράφονται σε κάθε σελίδα.

αδύνατο ενδεχόμενο	Δ' 103
αθροιστικές συχνότητες	Β' 58
αθροιστικές σχετικές συχνότητες	Β' 59
ανεξάρτητα ενδεχόμενα	Ε' 84
ανεξάρτητη μεταβλητή	Α' 14, Γ' 95

αξιοματικός ορισμός πιθανότητας	E' 15
απλός προσθετικός νόμος	E' 20
απογραφή	B' 23
ασυμβίβαστα ενδεχόμενα	Δ' 111, E' 19
βασική αρχή απαρίθμησης	E' 46
βέβαιο ενδεχόμενο	Δ' 103
γραμμική συσχέτιση	Δ' 5, Δ' 18
γραφική παράσταση συνάρτησης	A' 18
δείγμα	B' 25
δειγματικός χώρος	Δ' 100
δειγματοληψία	B' 29

δειγματοληψία με επανατοποθέτηση	Δ' 120
δεντροδιάγραμμα	Ε' 44
δεσμευμένη πιθανότητα	Ε' 78
δεύτερη παράγωγος	Α' 79
δημοσκόπηση	Β' 14
διάγραμμα διασποράς	Γ' 100
διάγραμμα συχνοτήτων	Β' 76
διακριτή μεταβλητή	Β' 22
διακύμανση	Γ' 44
διαλογή	Β' 53, Β' 91
διάμεσος	Γ' 23

διάμεσος ομαδοποιημένης κατανομής	Γ' 26
διασπορά	Γ' 44
διατάξεις	Ε' 49
εκατοστημόριο	Γ' 28
εκθετική συνάρτηση	Α' 23
εκτιμήτριες ελαχίστων τετραγώνων	Γ' 112
ενδεχόμενο	Δ' 101
ενδοτεταρτημοριακό εύρος	Γ' 42
εξαρτημένα ενδεχόμενα	Ε' 85
εξαρτημένη μεταβλητή	Α' 14, Γ' 103
εξίσωση γραφικής παράστασης	Α' 19

επαγωγή	B' 12
επικρατούσα τιμή	Γ' 32
ευθεία παλινδρόμησης	Γ' 103
ευνοϊκές περιπτώσεις	Δ' 102, Ε' 14
εύρος	B' 90, Γ' 40
εφαπτομένη καμπύλης	A' 48
ισοπίθανα απλά ενδεχόμενα	Ε' 13
ιστόγραμμα	B' 98
ιστόγραμμα αθροιστικών συχνοτήτων	B' 104
καμπύλη συχνοτήτων	B' 112
κανόνες παραγωγίσης	A' 89
κανονική κατανομή	B' 114

καμπύλη συνάρτησης	A' 18
κατανομή συχνοτήτων	B' 58
κατηγορική μεταβλητή	B' 21
κεντρική τιμή κλάσης	B' 88
κλάσεις	B' 87
κλάσεις ανίσου πλάτους	B' 106
κλάσεις ίσου πλάτους	B' 98
κλασικός ορισμός πιθανότητας	E' 18
κορυφή	Γ' 32
κριτήριο δεύτερης παραγώγου	A' 129
κριτήριο πρώτης παραγώγου	A' 117

κυκλικό διάγραμμα	B' 80
κύμανση	Γ' 40
λογαριθμική συνάρτηση	A' 24
μέθοδος ελαχίστων τετραγώνων	Γ' 106
μέση τιμή	Γ' 12
μεταβλητή	B' 18
μεταθέσεις	E' 50
μέτρα ασυμμετρίας	Γ' 9
μέτρα διασποράς	Γ' 38
μέτρα θέσης	Γ' 11
μονοτονία	A' 26
ολικό ελάχιστο	A' 29

ολικό μέγιστο	A' 29
ομαδοποίηση παρατηρήσεων	B' 87
ομοιογένεια	Γ' 60
ομοιόμορφη κατανομή	B' 114
όρια κλάσης	B' 87
όριο συνάρτησης	A' 32
παλινδρόμηση	Γ' 94
παραβολή	A' 21
παραγοντικό	E' 51
παράγωγος συνάρτησης	A' 78
παράγωγος σύνθετης συνάρτησης	A' 94

παράγωγος της f στο x_0	A' 61
πείραμα τύχης	Δ' 96
περιγραφική στατιστική	B' 11
πίνακας συχνοτήτων	B' 58
πλάτος κλάσης	B' 89
πληθυσμός	B' 19
ποιοτική μεταβλητή	B' 21
πολλαπλασιαστικός νόμος	E' 82
πολύγωνο συχνοτήτων	B' 76, B' 100
ποσοτική μεταβλητή	B' 22
πράξεις με ενδεχόμενα	Δ' 104
πράξεις με συναρτήσεις	A' 16

προσθετικός νόμος	Ε' 23
ραβδόγραμμα	Β' 65
ρυθμός μεταβολής	Α' 64
σημειόγραμμα	Β' 83
σταθμικός μέσος	Γ' 20
στατιστική ομαλότητα	Ε' 11
στατιστικοί πίνακες	Β' 35
στιγμιαία ταχύτητα	Α' 52
συμπληρωματικά ενδεχόμενα	Ε' 21
συνάρτηση αύξουσα	Α' 28
συνάρτηση γνησίως μονότονη	Α' 28

συνάρτηση ημίτονο	A' 25
συνάρτηση πραγματική	A' 13
συνάρτηση συνεχής	A' 39
συνάρτηση συνημίτονο	A' 25
συνάρτηση φθίνουσα	A' 28
συνδυασμοί	E' 55
συνεχής μεταβλητή	B' 22
συντελεστής γραμμικής συσχέτισης	Δ' 10
συντελεστής μεταβολής	Γ' 57
συχνότητα	B' 52
συχνότητα κλάσης	B' 91
σχεδιασμός πειραμάτων	B' 11

σχετική συχνότητα	B' 56
τεταρτημόριο	Γ' 29
τοπικό ελάχιστο	A' 31
τοπικό μέγιστο	A' 31
τυπική απόκλιση	Γ' 51
υπερβολή	A' 22
χρονόγραμμα	B' 84

ΠΕΡΙΕΧΟΜΕΝΑ 2ου ΤΟΜΟΥ

Σελίδα

ΚΕΦΑΛΑΙΟ 2ο: Στατιστική

2.1	Βασικές Έννοιες	18
2.2	Παρουσίαση Στατιστικών Δεδομένων	35

ΥΠΟΔΕΙΞΕΙΣ - ΑΠΑΝΤΗΣΕΙΣ ΑΣΚΗΣΕΩΝ	151
-------------------------------------	-----

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.