1954

ΙΔΡΥΜΑ ΕΥΓΕΝΙΔΟΥ

ΧΡΥΣΟΥΝ ΜΕΤΑΛΛΙΟΝ ΑΚΑΔΗΜΙΑΣ ΑΘΗΝΩΝ

ΠΡΟΛΟΓΟΣ ΙΔΡΥΜΑΤΟΣ ΕΥΓΕΝΙΔΟΥ

Ὁ Εὐγένιος Εὐγενίδης, ὁ ἱδρυτής καί χορηγός τοῦ «Ἱδρύματος Εὐγενίδου», πολύ νωρίς πρόβλεψε καί σχημάτισε τήν πεποίθηση ὅτι ἡ ἄρτια κατάρτιση τῶν τεχνικῶν μας, σέ συνδυασμό μέ τήν ἐθνική ἀγωγή, θά ἦταν ἀναγκαῖος καί ἀποφασιστικός παράγοντας τῆς προόδου τοῦ Ἔθνους μας.

Τήν πεποίθησή του αὐτή ὁ Εὐγενίδης ἐκδήλωσε μέ τή γενναιόφρονα πράξη εὐεργεσίας, νά κληροδοτήσει σεβαστό ποσό γιά τή σύσταση Ἱδρύματος πού θά εἶχε σκοπό νά συμβάλλει στήν τεχνική ἐκπαίδευση τῶν νέων τῆς Ἑλλάδας.

Ἔτσι τό Φεβρουάριο τοῦ 1956 συστήθηκε τό «Ἵδρυμα Εὐγενίδου», τοῦ ὁποίου τήν διοίκηση ἀνέλαβε ἡ ἀδελφή του κυρία Μαριάνθη Σίμου, σύμφωνα μέ τήν ἐπιθυμία τοῦ διαθέτη.

Ἀπό τό 1956 μέχρι σήμερα ἡ συμβολή τοῦ Ἱδρύματος στήν τεχνική ἐκπαίδευση πραγματοποιεῖται μέ διάφορες δραστηριότητες. Ὅμως ἀπ' αὐτές ἡ σημαντικότερη, πού κρίθηκε ἀπό τήν ἀρχή ὡς πρώτης ἀνάγκης, εἶναι ἡ ἔκδοση βιβλίων γιά τούς μαθητές τῶν τεχνικῶν σχολῶν.

Μέχρι σήμερα ἐκδόθηκαν 150 τόμοι βιβλίων, πού ἔχουν διατεθεῖ σέ πολλά ἑκατομμύρια τεύχη, καί καλύπτουν ἀνάγκες τῶν Κατώτερων καί Μέσων Τεχνικῶν Σχολῶν τοῦ Ὑπ. Παιδείας, τῶν Σχολῶν τοῦ Ὀργανισμοῦ Ἀπασχολήσεως Ἐργατικοῦ Δυναμικοῦ (ΟΑΕΔ) καί τῶν Δημοσίων Σχολῶν Ἐμπορικοῦ Ναυτικοῦ.

Μοναδική φροντίδα τοῦ Ἱδρύματος σ' αὐτή τήν ἐκδοτική του προσπάθεια ἦταν καί εἶναι ἡ ποιότητα τῶν βιβλίων, ἀπό ἄποψη ὄχι μόνον ἐπιστημονική, παιδαγωγική καί γλωσσική, ἀλλά καί ἀπό ἄποψη ἐμφανίσεως, ὥστε τό βιβλίο νά ἀγαπηθεῖ ἀπό τούς νέους.

Γιά τήν ἐπιστημονική καί παιδαγωγική ποιότητα τῶν βιβλίων, τά κείμενα ὑποβάλλονται σέ πολλές ἐπεξεργασίες καί βελτιώνονται πρίν ἀπό κάθε νέα ἔκδοση.

Ἰδιαίτερη σημασία ἀπέδωσε τό Ἵδρυμα ἀπό τήν ἀρχή στήν ποιότητα τῶν βιβλίων ἀπό γλωσσική ἄποψη, γιατί πιστεύει ὅτι καί τά τεχνικά βιβλία, ὅταν εἶναι γραμμένα σέ γλώσσα ἄρτια καί ὁμοιόμορφη ἀλλά καί κατάλληλη γιά τή στάθμη τῶν μαθητῶν, μποροῦν νά συμβάλλουν στήν γλωσσική διαπαιδαγώγηση τῶν μαθητῶν.

Ἔτσι μέ ἀπόφαση πού πάρθηκε ἤδη ἀπό τό 1956 ὅλα τά βιβλία τῆς Βιβλιοθήκης τοῦ Τεχνίτη, δηλαδή τά βιβλία γιά τίς Κατώτερες Τεχνικές Σχολές, ὅπως ἀργότερα καί γιά τίς Σχολές τοῦ ΟΑΕΔ, εἶναι γραμμένα σέ γλώσσα δημοτική μέ βάση τήν γραμματική τοῦ Τριανταφυλλίδη, ἐνῶ ὅλα τά ἄλλα βιβλία εἶναι γραμμένα στήν ἁπλή καθαρεύουσα. Ἡ γλωσσική ἐπεξεργασία τῶν βιβλίων γίνεται ἀπό φιλολόγους τοῦ Ἱδρύματος καί ἔτσι ἐξασφαλίζεται ἡ ἑνιαία σύνταξη καί ὁρολογία κάθε κατηγορίας βιβλίων.

Ἡ ποιότητα τοῦ χαρτιοῦ, τό εἶδος τῶν τυπογραφικῶν στοιχείων, τά σωστά σχήματα καί ἡ καλαίσθητη σελιδοποίηση, τό ἐξώφυλλο καί τό μέγεθος τοῦ βιβλίου περιλαμβάνονται καί αὐτά στίς φροντίδες τοῦ Ἱδρύματος.

Τό Ἵδρυμα θεώρησε ὅτι εἶναι ὑποχρέωσή του, σύμφωνα μέ τό πνεῦμα τοῦ Ἱδρυτή του, νά θέσει στήν διάθεση τοῦ Κράτους ὅλη αὐτή τήν πείρα του τῶν 20 ἐτῶν, ἀναλαμβάνοντας τήν ἔκδοση τῶν βιβλίων καί γιά τίς νέες Τεχνικές καί Ἐπαγγελματικές Σχολές καί τά νέα Τεχνικά καί Ἐπαγγελματικά Λύκεια, σύμφωνα μέ τά Ἀναλυτικά Προγράμματα τοῦ Κ.Ε.Μ.Ε.

Τά χρονικά περιθώρια γι' αὐτή τήν νέα ἐκδοτική προσπάθεια ἦταν πολύ περιορισμένα καί ἴσως γι' αὐτό, ἰδίως τά πρῶτα βιβλία αὐτῆς τῆς σειρᾶς, νά παρουσιάσουν ἀτέλειες στή συγγραφή ἤ στήν ἐκτύπωση, πού θά διορθωθοῦν στή νέα τους ἔκδοση. Γι' αὐτό τό σκοπό ἐπικαλούμαστε τήν βοήθεια ὅλων ὅσων θά χρησιμοποιήσουν τά βιβλία, ὥστε νά μᾶς γνωστοποιήσουν κάθε παρατήρησή τους γιά νά συμβάλλουν καί αὐτοί στή βελτίωση τῶν βιβλίων.

Β' ΤΑΞΗ ΕΠΑΓΓΕΛΜΑΤΙΚΟΥ ΛΥΚΕΙΟΥ

ΣΤΟΙΧΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΔΡΟΣ ΠΕΤΡΟΥ ΑΠΟΣΤ. ΚΙΟΧΟΥ
ΚΑΘΗΓΗΤΟΥ ΣΤΑΤΙΣΤΙΚΗΣ

ΑΘΗΝΑ
1983

ΠΡΟΛΟΓΟΣ

Τό μάθημα τῆς Στατιστικῆς ἔχει προγραμματισθεῖ νά διδαχθεῖ γιά πρώτη φορά στή μέση ἐκπαίδευση καί συγκεκριμένα στούς μαθητές τῆς Β΄ τάξεως Ἐπαγγελματικοῦ Λυκείου. Ἀντικειμενικός σκοπός τοῦ μαθήματος εἶναι νά διδαχθοῦν οἱ μαθητές τήν τεχνική καί τίς στατιστικές μεθόδους πού ἐφαρμόζονται στήν πράξη, κυρίως στό χῶρο τῆς οἰκονομίας καί τῆς διοικήσεως, γιά τήν ἐξαγωγή συμπερασμάτων, τά ὁποῖα εἶναι πολύ χρήσιμα γιά τή λήψη ὀρθῶν ἀποφάσεων, σέ μιά σύγχρονη ἐπιχείρηση ἤ ὀργανισμό.

Ἡ ὕλη πού ἔχει ἐπιλεγεῖ καί περιέχεται στό βιβλίο αὐτό εἶναι κυρίως περιγραφική καί ἐφαρμοσμένη στατιστική, κατάλληλη νά χρησιμοποιηθεῖ στόν ἐπαγγελματικό χῶρο, στόν ὁποῖο θά ἐργασθεῖ ὁ μαθητής τελειώνοντας τό Λύκειο.

Εἶναι ἀλήθεια ὅτι ἡ συγγραφή τοῦ βιβλίου παρουσίασε μεγάλες δυσκολίες. Ἡ προσπάθεια μάλιστα νά παρουσιασθεῖ ὅσο γίνεται πιό ἁπλό καί ὁλοκληρωμένο ἦταν ἰδιαίτερα ἐπίπονη, ἀφοῦ δέν ὑπῆρχε προηγούμενη ἐμπειρία στή χώρα μας ὡς πρός τή διδασκαλία τοῦ μαθήματος τῆς Στατιστικῆς στή μέση ἐκπαίδευση.

Σέ ὅσους βοήθησαν γιά τήν καλύτερη παρουσίαση τοῦ βιβλίου, ἐκφράζω καί ἀπό ἐδῶ τίς εὐχαριστίες μου.

Ὁ συγγραφέας

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

ΣΚΟΠΟΣ ΚΑΙ ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ

1.1 Τί είναι ἡ Στατιστική.

Στήν καθομιλουμένη «Στατιστική» σημαίνει συστηματική ἀπαρίθμηση καί παρουσίαση ἀριθμητικῶν δεδομένων ἤ στοιχείων, τά ὁποῖα προέρχονται ἀπό πολλές παρατηρήσεις ἤ μετρήσεις. Οἱ παρατηρήσεις αὐτές ἤ οἱ μετρήσεις ἀναφέρονται σέ συγκεκριμένο ἀντικείμενο ἤ γεγονός. Ἀνάλογα μέ τό ἀντικείμενο ἤ τό γεγονός στό ὁποῖο ἀναφέρονται τά ἀριθμητικά δεδομένα, ἡ Στατιστική παίρνει καί ἰδιαίτερη ὀνομασία. Ἔτσι π.χ. ὅταν μιλᾶμε γιά «Γεωργική Στατιστική», «Στατιστική Ἐπιχειρήσεων» ἤ «Στατιστική Ἐργατικοῦ δυναμικοῦ», κ.λ.π., ἐννοοῦμε ἀριθμητικά δεδομένα πού ἀναφέρονται ἀντίστοιχα στή γεωργία, στίς ἐπιχειρήσεις ἤ στό ἐργατικό δυναμικό, κ.λ.π. Στήν ἐπιστημονική γλώσσα, ἡ λέξη «Στατιστική» ἔχει εὐρύτερη σημασία· σημαίνει τήν ἐπιστήμη πού ἔχει ὡς ἀντικείμενο ὄχι μόνο τή συγκέντρωση καί παρουσίαση, ἀλλά καί τή μελέτη καί ἀνάλυση τῶν παρατηρήσεως ἤ μετρήσεων, πού ἀναφέρονται σέ ἕνα συγκεκριμένο ἀντικείμενο ἤ γεγονός, ὁποιαδήποτε καί ἄν εἶναι ἡ φύση του. Ἔτσι, ἡ Στατιστική περιλαμβάνει τόσο τίς μεθόδους συλλογῆς καί ἐπεξεργασίας στοιχείων, ὅσο καί τίς μεθόδους ἀναλύσεως καί μελέτης τους, ἀνακαλύπτοντας ἔτσι σχέσεις πού ὑπάρχουν ἀνάμεσα στά διάφορα φαινόμενα καί διατυπώνοντας συμπεράσματα, πού εἶναι χρήσιμα γιά τή λήψη ὀρθῶν ἀποφάσεων. Μποροῦμε λοιπόν νά ποῦμε ὅτι:

Στατιστική εἶναι ἡ ἐπιστήμη πού ἀσχολεῖται μέ τίς ἐπιστημονικές μεθόδους συλλογῆς, ἐπεξεργασίας, παρουσιάσεως καί ἀναλύσεως ἀριθμητικῶν δεδομένων, πού ἀναφέρονται στίς διάφορες μονάδες πολυπληθῶν ὁμάδων καί ἔχει ὡς σκοπό τή διατύπωση συμπερασμάτων, τά ὁποῖα εἶναι χρήσιμα στή λήψη ὀρθῶν ἀποφάσεων.

Ἀναλύοντας τόν ὁρισμό αὐτό τῆς Στατιστικῆς παρατηροῦμε ὅτι τά βασικά στάδια, πού ἀκολουθοῦμε γιά τή μελέτη τῶν ἰδιοτήτων τῶν διαφόρων μονάδων μιᾶς πολυπληθοῦς ὁμάδας, εἶναι τά ἑξῆς:

α) Ἡ συγκέντρωση τῶν ἀπαραιτήτων στατιστικῶν στοιχείων.
β) Ἡ μεθοδική ἐπεξεργασία καί παρουσίαση τῶν στατιστικῶν στοιχείων.
γ) Ἡ ἀνάλυση τῶν στοιχείων αὐτῶν καί ἡ ἐξαγωγή χρησίμων συμπερασμάτων.

1.2 Ἱστορία τῆς Στατιστικῆς.

Ἡ λέξη «στατιστική» προέρχεται ἀπό τή λατινική λέξη status (πού σημαίνει «κράτος») καί δήλωνε ἀρχικά συλλογή στοιχείων γιά τίς κρατικές ἀνάγκες (ἔκταση, παραγωγή, πληθυσμό, κ.λ.π.) Ἔχει ἐξακριβωθεῖ ὅτι ἡ πρώτη ἀπογραφή πληθυσμοῦ ἔγινε στήν Κίνα ἀπό τόν αὐτοκράτορα Υ-άο τό ἔτος 2238 π.Χ., ἐνῶ στούς Ρωμαίους ἡ πρώτη ἀπογραφή πληθυσμοῦ ἔγινε ἐπί Ρωμύλου (753 — 715 π.Χ.) καί ἡ τελευταία ἀπό τόν αὐτοκράτορα Βεσπασιανό τό 73 μ.Χ. Στήν Ἀγγλία ἡ πρώτη καθολική ἀπογραφή τοῦ πληθυσμοῦ καί τοῦ πλούτου γενικά ἔγινε τό 1085 ἀπό τόν Γουλιέλμο τόν κατακτητή.

Τό 1583 γράφεται ἀπό τόν Fr. Sansovino τό πρῶτο βιβλίο στατιστικοῦ περιεχομένου καί λίγο ἀργότερα εἰσάγεται ἀπό τόν Könring (1606 — 1681) ἡ Στατιστική στήν ἀνώτερη παιδεία.

Τήν ἴδια ἐποχή ἐμφανίζεται τό ἐνδιαφέρον γιά τίς ἀσφάλειες ζωῆς καί ὁ περίφημος Ἄγγλος ἀστρονόμος Halley, χρησιμοποιώντας τά ληξιαρχικά βιβλία γεννήσεων καί θανάτων τῆς πόλεως Breslaou, παρουσιάζει τόν πρῶτο πίνακα θνησιμότητας. Τό ρεῦμα αὐτό τῶν δημογραφικῶν μελετῶν ἐπεκτείνεται καί στή Γερμανία, ὅπου ὁ πάστορας Süssmilch (1707 — 1767) συγκεντρώνει στοιχεῖα ἀπό τά ληξιαρχικά βιβλία τῶν ἐφημερίων τῆς Πρωσσίας καί καταλήγει τό 1741 στό συμπέρασμα ὅτι τό ποσοστό γεννήσεως τῶν ἀγοριῶν εἶναι 51% καί τῶν κοριτσιῶν 49%, ἐνῶ τά δύο φύλα ἔχουν ἴσα ποσοστά κατά τήν ἐποχή τοῦ γάμου. Τό φαινόμενο αὐτό γιά τόν συγγραφέα δέν εἶναι τυχαῖο γεγονός ἀλλά νόμος θείας προελεύσεως πού ἀποσκοπεῖ στή διαιώνιση τοῦ εἴδους. Μέχρι τήν ἐποχή αὐτή ἡ Στατιστική ἔχει περιγραφικό χαρακτήρα καί ἀσχολεῖται κυρίως μέ θέματα Δημογραφίας.

Ἡ Στατιστική θά ξεφύγει ἀπό τόν περιγραφικό χαρακτήρα της μέ τήν ἀνάπτυξη ἑνός νέου κλάδου, τοῦ Λογισμοῦ τῶν Πιθανοτήτων, ὁ ὁποῖος προῆρθε ἀπό τή μελέτη τῶν τυχερῶν παιγνιδιῶν (χαρακτηριστική μάλιστα εἶναι ἡ ἀλληλογραφία ἀνάμεσα στούς Γάλλους μαθηματικούς Pascal καί Fermat, μέ ἀφορμή τά ἐρωτήματα πού ἔθεσε στόν Pascal ὁ Ἱππότης De Meré γιά τά παιγνίδια τοῦ κύβου). Ἀπό τούς θεμελιωτές τοῦ Λογισμοῦ τῶν Πιθανοτήτων ἀναφέρομε τόν Bernoulli, ὁ ὁποῖος στό βιβλίο του «Ἡ τέχνη τῶν προβλέψεων» διατυπώνει τόν περίφημο ***νόμο τῶν μεγάλων ἀριθμῶν*** καί τόν Γάλλο μαθηματικό Laplace, στόν ὁποῖο ὀφείλεται ἡ ἐφαρμογή τοῦ Λογισμοῦ τῶν Πιθανοτήτων στή σπουδή τῶν φυσικῶν φαινομένων μέ πολυσύνθετες αἰτίες. Στή νέα αὐτή περίοδο τῆς Στατιστικῆς ὁ Βέλγος ἀστρονόμος Quetelet ἐπεκτείνει τήν ἐφαρμογή τῆς Στατιστικῆς στή σπουδή τῶν φυσικῶν, διανοητικῶν καί ἠθικῶν ἰδιοτήτων τοῦ ἀνθρώπου καί παίρνει τήν πρωτοβουλία γιά τή σύγκληση τοῦ πρώτου Διεθνοῦς Συνεδρίου Στατιστικῆς πού ἔγινε στίς Βρυξέλλες τό 1853, ἐνῶ ἀργότερα ὁ F. Galton ἐφαρμόζει τήν Στατιστική στή Βιολογία καί εἰδικότερα στά προβλήματα τῆς κληρονομικότητας. Ἡ προσπάθεια τοῦ Galton συνεχίσθηκε ἀπό τόν Ἄγγλο μαθηματικό Pearson, στόν ὁποῖο ὀφείλεται κατά πολύ ἡ σημερινή ἀνάπτυξη καί θέση τῆς Στατιστικῆς.

1.3 Χρησιμότητα καί πεδία ἐφαρμογῆς τῆς Στατιστικῆς.

Μιά ἁπλή ἀρίθμηση τῶν ἐφαρμογῶν της δείχνει ὅτι ἡ Στατιστική, ἡ ὁποία εἶναι βασικά ἐφαρμοσμένη ἐπιστήμη, χρησιμοποιεῖται σέ ὅλους σχεδόν τούς τομεῖς τῆς ἀνθρώπινης δραστηριότητας. Ἡ Στατιστική εἶναι ἀπαραίτητη στή ***Διοίκηση*** γενικά, ὅπου ἡ λήψη ὀρθῶν ἀποφάσεων ἔχει μεγάλη σημασία γιά τήν πρόοδο ἑνός κράτους, ἑνός ὀργανισμοῦ, μιᾶς βιομηχανίας ἤ μιᾶς ἐπιχειρήσεως. Γι' αὐτό καί δέν ὑπάρχει σήμερα τομέας σέ καμιά σύγχρονη ἐπιχείρηση, πού νά μή χρησιμοποιεῖ τίς στατιστικές μεθόδους στή λήψη ἐπιχειρηματικῶν ἀποφάσεων.

Μεγάλη σημασία ἔχει ἡ ἐφαρμογή τῆς στατιστικῆς στή ***Δημογραφία,*** ὅπου ἡ μελέτη τῆς γαμηλιότητας, τῆς γεννητικότητας, τῆς θνησιμότητας, τῆς μεταναστεύσεως, κ.λ.π. ἀπαιτεῖ μακροχρόνιες στατιστικές παρατηρήσεις καί ἐπίπονες ἀναλύσεις. Ἐπίσης ἡ Στατιστική ἐφαρμόζεται σήμερα στήν Ἰατρική, Φυσική, Γενετική, Ἀστρονομία, Βιολογία, Μετεωρολογία, Γεωργία, Βιομηχανία, στή μελέτη τοῦ φυσικοῦ περιβάλλοντος, στή μελέτη τῶν ἀνθρωπίνων ἰδεῶν καί προθέσεων, στή Θεωρία τῶν ἀποφάσεων, στόν ἔλεγχο ποιότητας τῶν προϊόντων κ.λ.π. Τέλος ἡ Στατιστική βρίσκει πολύ μεγάλη ἐφαρμογή καί στόν ***Οἰκονομικό τομέα,*** ὅπου ἡ παρακολούθηση τοῦ γενικοῦ ἐπιπέδου τῶν τιμῶν, τοῦ ἐθνικοῦ εἰσοδήματος, τῆς νομισματικῆς ἰσοτιμίας καί τῶν οἰκονομικῶν διακυμάνσεων, τῆς ἀπασχολήσεως, τῆς παραγωγικότητας, τῆς καταρτίσεως δεικτῶν οἰκονομικῆς δραστηριότητας, τῶν ἐθνικῶν πόρων καί τῆς ἐθνικῆς δαπάνης, εἶναι ἀντικείμενα στατιστικῆς ἐπεξεργασίας.

1.4 Τό Κράτος καί ἡ Στατιστική.

Ἕνα καλά ὀργανωμένο Κράτος ὀφείλει νά γνωρίζει κάθε στιγμή τόν πληθυσμό τῆς χώρας, τήν κατανομή τοῦ πληθυσμοῦ κατά φύλο, ἡλικία, ἐπάγγελμα κ.λ.π., καθώς καί τήν κίνηση καί πιθανή ἐξέλιξή του. Πρέπει ἐπίσης νά παρακολουθεῖ τόσο τά οἰκονομικά φαινόμενα τῆς χώρας (παραγωγή, εἰσαγωγές καί ἐξαγωγές, κίνηση καί ἐμπορία τῶν ἀγαθῶν, κ.λ.π.) ὅσο καί τά διοικητικά καί κοινωνικά φαινόμενα τῆς χώρας (διοίκηση, ἐργασία, δημόσια ὑγεία, πρόνοια, κοινωνικές ἀσφαλίσεις, ἐκπαίδευση, δικαιοσύνη, στέγαση, κ.λ.π).

Γιά τό σκοπό αὐτό, κάθε Κράτος ἔχει μία Στατιστική Ὑπηρεσία, ἡ ὁποία συγκεντρώνει τά ἀπαραίτητα στοιχεῖα καί παρακολουθεῖ τήν ἐξέλιξη τῶν παραπάνω φαινομένων. Μία τέτοια ὑπηρεσία πρέπει νά εἶναι καλά ὀργανωμένη καί νά διαθέτει ἕνα πλούσιο κεντρικό ἀρχεῖο στατιστικῶν στοιχείων, ἀπό τό ὁποῖο θά ἀντλεῖ χρήσιμες πληροφορίες κάθε διοικητικός παράγοντας τοῦ Κράτους.

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ

ΣΥΛΛΟΓΗ ΣΤΑΤΙΣΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ

2.1 Στατιστικός πληθυσμός.

Ἡ λέξη «πληθυσμός» χρησιμοποιεῖται στή Στατιστική γιά νά δηλώσει τό σύνολο τῶν ἀτόμων ἤ ἀντικειμένων, στά ὁποῖα ἀναφέρονται οἱ παρατηρήσεις μας.

Τά στοιχεῖα τοῦ συνόλου αὐτοῦ λέγονται ***στατιστικές μονάδες*** ἤ ***ἄτομα*** τοῦ πληθυσμοῦ. Ὡς στατιστική μονάδα λοιπόν μπορεῖ νά θεωρηθεῖ ὁποιοδήποτε πρόσωπο, πράγμα, γεγονός, κ.λ.π τό ὁποῖο ἀνήκει σέ ὁρισμένο πληθυσμό καί ὅλα τά στοιχεῖα τοῦ πληθυσμοῦ αὐτοῦ ἐξετάζονται ὡς πρός μία ἤ περισσότερες χαρακτηριστικές ἰδιότητές τους.

Παράδειγμα.

Ὅταν ἐξετάζομε τή βαθμολογία στά Μαθηματικά τῶν μαθητῶν μιᾶς τάξεως, στατιστικός πληθυσμός εἶναι τό σύνολο τῶν μαθητῶν τῆς τάξεως, ἐνῶ κάθε μαθητής εἶναι μία στατιστική μονάδα. Ἐπειδή ἡ ἰδιότητα ἤ τό χαρακτηριστικό πού μᾶς ἐνδιαφέρει νά μελετήσομε εἶναι «ἡ βαθμολογία τῶν μαθητῶν» τῆς τάξεως στά Μαθηματικά, οἱ διάφοροι βαθμοί τους 10, 12, 11, 13, 17, 18, 19, 12, 10, 13, 12, 17,... ἀποτελοῦν τίς παρατηρήσεις μας (στατιστικά δεδομένα).

Τονίζεται ὅτι ἡ Στατιστική δέν ἀσχολεῖται μέ τή μελέτη τῶν ἴδιων τῶν στατιστικῶν μονάδων, ἀλλά μέ τή μελέτη τῶν ἰδιοτήτων τους.

2.2 Ἔννοια στατιστικῆς μεταβλητῆς – Διακρίσεις αὐτῆς.

Οἱ χαρακτηριστικές ἰδιότητες τῶν στατιστικῶν μονάδων ἑνός πληθυσμοῦ, μέ τή μελέτη τῶν ὁποίων ἀσχολεῖται ἡ Στατιστική, ὀνομάζονται ***μεταβλητές***. Οἱ ἀριθμοί ἤ οἱ ἄλλες συμβολικές ἐκφράσεις, πού ἀντιπροσωπεύουν τίς διάφορες καταστάσεις μιᾶς μεταβλητῆς, ὀνομάζονται ***τιμές τῆς μεταβλητῆς***. Κάθε μεταβλητή συμβολίζεται μέ ἕνα ἀπό τά κεφαλαῖα γράμματα Χ, Υ, Ψ,...

Γιά τήν καλύτερη κατανόηση τῆς ἔννοιας τῆς μεταβλητῆς, δίνονται στόν Πίνακα 2.2.1 διάφορες μεταβλητές καί οἱ ἀντίστοιχες τιμές τους.

ΠΙΝΑΚΑΣ 2.2.1.

ΜΕΤΑΒΛΗΤΕΣ	ΤΙΜΕΣ
Ἀριθμός ἀγοριῶν σέ μιά οἰκογένεια	0, 1, 2, 3, 4, 5, ...
Ἔνδειξη ἑνός ζαριοῦ	1, 2, 3, 4, 5, 6
Ἔνδειξη ἑνός νομίσματος	«πρόσωπο», «γράμματα»
Ἀριθμός δωματίων ἑνός διαμερίσματος	1, 2, 3, 4, ...
Φυλή ἀτόμων	λευκός, μαῦρος, κίτρινος, ἐρυθρόδερμος
Ὕψος ἀτόμων σέ ἑκατοστά	145, 146, 147, ...
Ὑγεία ἀτόμων	ἄριστη, καλή, μέτρια, κακή
Φύλο ἀτόμων	ἀρσενικό, θηλυκό

Οἱ μεταβλητές χωρίζονται σέ δύο κυρίως κατηγορίες:

— Στίς ***ποιοτικές μεταβλητές*** πού δέν ἐπιδέχονται μέτρηση καί οἱ «τιμές» τους ἐκφράζονται μέ λέξεις. Τέτοιες μεταβλητές εἶναι π.χ. «ἡ οἰκογενειακή κατάσταση ἑνός ὑπαλλήλου», «ἡ κατάσταση ὑγείας ἑνός μαθητῆ», «τό ἐπάγγελμα ἑνός ἀτόμου», κ.λ.π.

— Στίς ***ποσοτικές μεταβλητές*** πού ἐπιδέχονται μέτρηση καί οἱ τιμές τους εἶναι ἀριθμοί ἀναφερόμενοι σέ συγκεκριμένες μονάδες, τέτοιες μεταβλητές εἶναι π.χ. τό βάρος ἤ τό ὕψος ἑνός μαθητῆ, ἡ ἡλικία ἤ τό εἰσόδημα ἑνός ἀτόμου, ἡ θερμοκρασία, οἱ ἐξαγωγές, ὁ ἀριθμός τῶν δωματίων ἑνός διαμερίσματος, κ.λ.π. Ἄν μιά ποσοτική μεταβλητή σημειωθεῖ μέ τό γράμμα Χ, οἱ τιμές της θά σημειώνονται μέ x_1, x_2, x_3, ...

Οἱ ποσοτικές μεταβλητές διακρίνονται σέ ***ἀσυνεχεῖς*** καί ***συνεχεῖς***.

Ἀσυνεχεῖς ὀνομάζονται οἱ μεταβλητές ἐκεῖνες πού μποροῦν νά λάβουν πεπερασμένο ἤ ἀριθμήσιμο πλῆθος τιμῶν. Ἔτσι π.χ. ἡ ἔνδειξη ἑνός ζαριοῦ εἶναι μία ἀσυνεχής τυχαία μεταβλητή, γιατί τό σύνολο τιμῶν της {1, 2, 3, 4, 5, 6} εἶναι ***πεπερασμένο.*** Ἐπίσης ὁ ἀριθμός ρίψεων ἑνός νομίσματος μέχρι νά ἐμφανισθεῖ γιά πρώτη φορά ἡ ὄψη «πρόσωπο» εἶναι μία ἀσυνεχής μεταβλητή, γιατί τό σύνολο τιμῶν της {1, 2, 3 . . .ν, . . .} εἶναι ἀριθμήσιμο.

Συνεχεῖς ὀνομάζονται οἱ μεταβλητές ἐκεῖνες πού μποροῦν νά λάβουν ὅλες τίς τιμές ἑνός διαστήματος. Ἔτσι π.χ. τό βάρος ἤ τό ὕψος ἑνός μαθητῆ, τό εἰσόδημα ἤ ἡ ἡλικία ἑνός ἀτόμου, ἡ ταχύτητα, ἡ θερμοκρασία, κ.λ.π εἶναι συνεχεῖς μεταβλητές.

2.3 Πηγές συλλογῆς στατιστικῶν στοιχείων.

Τό πρῶτο στάδιο γιά τή στατιστική μελέτη ἑνός φαινομένου εἶναι ἡ συγκέντρωση τῶν στατιστικῶν στοιχείων του. Τό στάδιο αὐτό χρειάζεται ἰδιαίτερη προσοχή καί φροντίδα, γιατί ἀπό τήν ἀξιοπιστία τῶν στοιχείων πού θά συγκεντρωθοῦν ἐξαρτᾶται καί ἡ ἀξία τῶν στατιστικῶν συμπερασμάτων. Ἄν τά στοιχεῖα εἶναι ἀνακριβή ἤ λανθασμένα εἶναι φανερό ὅτι καί τά στατιστικά συμπεράσματα πού θά προκύψουν θά εἶναι ἐπίσης ἀνακριβή ἤ λανθασμένα.

Ἡ συγκέντρωση στατιστικῶν στοιχείων γίνεται ἀπό πολλούς φορεῖς, π.χ. ἀπό διάφορα ***κέντρα*** καί ***ἰνστιτοῦτα ἐρευνῶν,*** ἀπό ***δημόσιους*** καί ***ἰδιωτικούς ὀργανι-***

σμούς, ἀπό ***βιομηχανικά*** καί ***ἐμπορικά ἐπιμελητήρια,*** ἀπό ***τράπεζες, δημόσιες ὑπηρεσίες, διεθνεῖς ὀργανισμούς*** κ.λ.π. Στή χώρα μας ἡ μεγαλύτερη πηγή παροχῆς στατιστικῶν στοιχείων εἶναι ἡ ***Ἐθνική Στατιστική Ὑπηρεσία*** (Ε.Σ.Υ.Ε), πού ἐκδίδει καί γενικά στατιστικά δημοσιεύματα (ἐτήσια στατιστική ἐπετηρίδα, μηνιαῖο στατιστικό δελτίο) καί εἰδικά στατιστικά δημοσιεύματα πού ἀναφέρονται στόν πληθυσμό, στή γεωργία, στή μεταποίηση, στή συγκοινωνία, στό ἐξωτερικό ἐμπόριο, στήν ἐκπαίδευση, στή δικαιοσύνη κ.λ.π. Ἄλλες βασικές πηγές Ἑλληνικῶν στατιστικῶν στοιχείων εἶναι τό «Μηνιαῖο στατιστικό δελτίο» τῆς Τράπεζας τῆς Ἑλλάδος, «οἱ Ἐθνικοί Λογαριασμοί» ἔντυπο, πού ἐκδίδεται κάθε χρόνο ἀπό τό Ὑπουργεῖο Συντονισμοῦ καί τά δημοσιεύματα διαφόρων κέντρων (Κέντρο Κοινωνικῶν Ἐρευνῶν, Κέντρο Προγραμματισμοῦ, κ.λ.π).

2.4 Μέθοδοι συλλογῆς στατιστικῶν στοιχείων.

Γιά τή συλλογή τῶν στατιστικῶν στοιχείων ἐφαρμόζονται διάφορες μέθοδοι, ἀπό τίς ὁποῖες σπουδαιότερες εἶναι:

α) Ἡ ***ἀπογραφή (ἤ ὁλοκληρωτική μέθοδος),*** ἡ ὁποία συνίσταται στή συγκέντρωση στοιχείων ἀπό ὅλες τίς στατιστικές μονάδες τοῦ πληθυσμοῦ πού ἐξετάζομε.

Ἡ ἀπογραφή ἐφαρμόζεται συνήθως ὅταν ὁ ἀριθμός τῶν στατιστικῶν μονάδων τοῦ πληθυσμοῦ δέν εἶναι πολύ μεγάλος καί, ἀνάλογα μέ τό εἶδος τοῦ πληθυσμοῦ πού ἀπογράφομε, διακρίνομε ἀπογραφές γεωργίας, κτηνοτροφίας, βιομηχανικῶν καί ἐμπορικῶν ἐπιχειρήσεων, ἀπογραφές τῶν μονάδων παραγωγῆς ἑνός ἐργοστασίου ἤ τῶν προϊόντων μιᾶς ἐπιχειρήσεως κ.λ.π. Ἰδιαίτερη σημασία ἔχει ἡ ἀπογραφή τοῦ πληθυσμοῦ, ἡ ὁποία γίνεται κάθε 5 ἤ 10 χρόνια ἀπό τή στατιστική ὑπηρεσία τῆς κάθε χώρας καί ἀποτελεῖ τήν κύρια πηγή πληροφοριῶν ὡς πρός τά δημογραφικά, οἰκονομικά καί κοινωνικά δεδομένα μιᾶς χώρας.

Τά ***βασικά μειονεκτήματα τῶν ἀπογραφῶν εἶναι:***

— ***Ἀπαιτοῦν μεγάλο κόστος,*** γι' αὐτό καί δέν γίνονται συχνά (ἡ ἀπογραφή πληθυσμοῦ στή χώρα μας γίνεται κάθε 10 χρόνια, ἐνῶ ἡ ἀπογραφή τῶν βιομηχανικῶν καί ἐμπορικῶν ἐπιχειρήσεων κάθε 5 χρόνια). Αὐτός εἶναι καί ὁ λόγος, γιά τόν ὁποῖο, κατά τήν ἀπογραφή συγκεντρώνομε ὅσο τό δυνατό περισσότερα στοιχεῖα. Ἔτσι π.χ. στήν ἀπογραφή πληθυσμοῦ συγκεντρώνομε στοιχεῖα ὡς πρός τήν ἡλικία, τό φύλο, τό ἐπάγγελμα, τήν οἰκογενειακή κατάσταση καί ἐκπαίδευση κ.λ.π. κάθε ἀτόμου.

— Ἐπειδή ὁ ἀριθμός τῶν μονάδων πού ἀπογράφονται εἶναι μεγάλος καί οἱ πληροφορίες πάρα πολλές, ἡ ***δημοσίευση τῶν ἀποτελεσμάτων καθυστερεῖ.***

— ***Ἡ ἀπογραφή δέν μπορεῖ νά γίνει μόνο μέ τούς λίγους εἰδικευμένους ὑπαλλήλους τῆς ὑπηρεσίας καί ἀναγκαζόμαστε νά χρησιμοποιήσομε ἔκτακτο προσωπικό, μέ ἀποτέλεσμα οἱ πληροφορίες πού συγκεντρώνονται νά περιέχουν λάθη πού ὀφείλονται στούς ἀπογραφεῖς.***

β) Ἡ ***δειγματοληψία (ἤ δειγματοληπτική μέθοδος)*** συνίσταται στή συγκέντρωση στοιχείων ἀπό ἕνα περιορισμένο ἀριθμό στατιστικῶν μονάδων, οἱ ὁποῖες ἀποτελοῦν ἕνα «δεῖγμα» τοῦ πληθυσμοῦ.

Ἡ δειγματοληψία ἐφαρμόζεται ὅταν ὁ πληθυσμός πού ἐξετάζομε ἀποτελεῖται ἀπό μεγάλο πλῆθος στατιστικῶν μονάδων (ὅπως συνήθως συμβαίνει στήν πράξη) ἤ ὅταν ἡ ἐξέταση τῶν στατιστικῶν μονάδων προϋποθέτει τήν καταστροφή τους, ὁπότε ἡ ἀπογραφή εἶναι παράλογη. Ἔτσι π.χ. ἄν μιά ὑπηρεσία τοῦ Στρατοῦ παρέλαβε μιά ποσότητα ἀπό 100.000 ὀβίδες καί θέλει νά ἐλέγξει τήν ποιότητά τους, θά πραγματοποιήσει μόνο μερικές δοκιμαστικές βολές μέ ὁρισμένες ἀπό τίς ὀβίδες αὐτές καί δέν θά κάνει ἀπογραφή δοκιμάζοντας καί καταστρέφοντας ὅλες τίς ὀβίδες.

Τονίζεται ὅτι ἡ ἐπιλογή τοῦ δείγματος ἀπαιτεῖ εἰδική διαδικασία καί οἱ πληροφορίες, οἱ ἐκτιμήσεις καί τά συμπεράσματα πού προκύπτουν ἀπό τήν ἐξέταση τῶν στοιχείων τοῦ δείγματος ἰσχύουν γιά τό σύνολο τοῦ πληθυσμοῦ μόνον ὅταν τό δεῖγμα εἶναι ἀντιπροσωπευτικό καί ἐφόσον αὐτό ἐπιλέγεται μέ κατάλληλο σχεδιασμό.

Δύο εἶναι τά βασικά ***πλεονεκτήματα*** τῆς δειγματοληψίας σέ σύγκριση μέ τήν ἀπογραφή: ἡ μεγαλύτερη ***ταχύτητα*** μέ τήν ὁποία παίρνομε τίς πληροφορίες μας καί ***τό χαμηλό κόστος.***

Παράλληλα ὅμως μέ τά πλεονεκτήματα τῆς δειγματοληψίας ὑπάρχουν καί ὁρισμένα ***μειονεκτήματα,*** κυριότερα ἀπό τά ὁποῖα εἶναι:

— Ἡ ***δυσκολία πού παρουσιάζεται πολλές φορές στό νά συγκεντρωθεῖ ὁ ἀπαιτούμενος ἀριθμός τῶν στατιστικῶν μονάδων τοῦ δείγματος.*** Ἔτσι π.χ. ἄν θέλομε νά βροῦμε τό ποσοστό τῶν καπνιστῶν μιᾶς μεγάλης πόλεως, πού ἔχουν ἡλικία μεγαλύτερη ἀπό 80 χρόνια, δέν εἶναι εὔκολο νά συγκεντρώσομε τόν ἀπαιτούμενο ἀριθμό ἀτόμων, γιατί τά ἄτομα αὐτῆς τῆς ἡλικίας ἀποτελοῦν μικρό ποσοστό τοῦ ὅλου πληθυσμοῦ τῆς πόλεως.

— Ὁ ***σχεδιασμός καί ἡ ἐκτέλεση τῆς δειγματοληψίας χρειάζονται ἰδιαίτερη προσοχή καί ἔμπειρους ἐπιστήμονες,*** γιατί πρέπει νά ἀκολουθεῖται αὐστηρά μία εἰδική διαδικασία τόσο στήν ἐπιλογή τοῦ δείγματος, ὅσο καί στή στατιστική ἀνάλυση τῶν πληροφοριῶν του.

— Διάφοροι παράγοντες ὅπως εἶναι ἡ ***κακή σχεδίαση καί ἐκτέλεση τῆς δειγματοληψίας,*** ἡ ***μή ἀντιπροσωπευτικότητα*** τοῦ δείγματος, ***ἡ ἐκλογή ἀκατάλληλης μεθόδου,*** κ.λ.π. ὁδηγοῦν σέ παραπλανητικά συμπεράσματα.

Ἡ δειγματοληψία ἔχει σήμερα μεγάλη ἀνάπτυξη καί ὑπάρχουν πολλές μέθοδοι γιά τή διενέργεια δειγματοληπτικῶν ἐρευνῶν, ὅπως εἶναι ***ἡ τυχαία δειγματοληψία, ἡ στρωματοποιημένη δειγματοληψία, ἡ συστηματική δειγματοληψία,*** κ.λ.π. Ἡ ἐκλογή τῆς μεθόδου κάθε φορά ἐξαρτᾶται ἀπό τό ζητούμενο βαθμό ἀκριβείας τῶν ἀποτελεσμάτων, ἀπό τά χρονικά καί χρηματικά περιθώρια τῆς ἔρευνας, ἀπό τή μεταβλητικότητα τῶν μονάδων τοῦ πληθυσμοῦ πού ἐξετάζομε, ἀπό τήν ὕπαρξη καταλόγων γιά τίς μονάδες τοῦ πληθυσμοῦ, ἀπό τή δυνατότητα διαιρέσεως τοῦ πληθυσμοῦ σέ ὁμοιογενεῖς ὑποπληθυσμούς, κ.λ.π. Πάντως γιά κάθε μέθοδο ὑπάρχει εἰδική διαδικασία ὡς πρός τή συλλογή καί ἐπεξεργασία τῶν στατιστικῶν στοιχείων.

γ) Ἡ ***συνεχής ἐγγραφή στατιστικῶν στοιχείων*** συνίσταται στήν ἄμεση καταχώρηση σέ εἰδικά βιβλία ἤ ἔντυπα διαφόρων στοιχείων, μόλις αὐτά ἐμφανισθοῦν.

Ὡς παραδείγματα τῆς μεθόδου αὐτῆς μποροῦμε νά ἀναφέρομε τίς καταχωρήσεις σέ εἰδικά βιβλία τοῦ ληξιαρχείου τῶν γεννήσεων, τῶν θανάτων καί τῶν γάμων, τίς καταχωρήσεις στά εἰδικά βιβλία τῶν νοσοκομείων τῶν ἀσθενῶν πού εἰσέρχονται καί ἐξέρχονται σ' αὐτά, τίς καταχωρήσεις σέ εἰδικό βιβλίο τοῦ Ὑπουργείου Βιομηχανίας τῶν ἀδειῶν πού χορηγοῦνται γιά λειτουργία βιομηχανικῶν ἐργοστασίων, κ.λ.π Ἐπίσης οἱ ἄδειες κυκλοφορίας αὐτοκινήτων, ἡ οἰκονομική δραστηριότητα, ἡ τουριστική κίνηση, οἱ ἐξαγωγές καί εἰσαγωγές, τά μετεωρολογικά στοιχεῖα κ.λ.π., καταχωροῦνται σέ εἰδικά δελτία ἤ βιβλία μόλις παρουσιασθοῦν.

2.5 'Επεξεργασία στατιστικών στοιχείων.

Μετά τή συγκέντρωση τῶν στατιστικῶν στοιχείων ἀπό τά ἐρωτηματολόγια*, ἀκολουθεῖ τό στάδιο τῆς ἐπεξεργασίας τους, τό ὁποῖο περιλαμβάνει:

α) ***Τόν ἔλεγχο ὅλων τῶν ἐρωτηματολογίων*** (γιά νά ἐντοπισθοῦν οἱ ἀσυμπλήρωτες, ἀσαφεῖς, δυσανάγνωστες ἤ ἀσυμβίβαστες ἀπαντήσεις).

β) ***Τή διαλογή τῶν πληροφοριῶν πού ἀναφέρονται στά διάφορα χαρακτηριστικά τῶν στατιστικῶν μονάδων.***

'Η διαλογή τῶν στατιστικῶν πληροφοριῶν ἀνάλογα μέ τόν ἀριθμό τῶν ἐρωτηματολογίων καί τό πλῆθος τῶν χαρακτηριστικῶν πού ἐξετάζομε, μπορεῖ νά γίνει εἴτε μέ τό χέρι, εἴτε μέ μηχανικά μέσα. 'Η διαλογή τῶν στοιχείων μέ τό χέρι γίνεται ὅταν ἔχομε μικρό ἀριθμό ἐρωτηματολογίων καί περιορισμένο πλῆθος χαρακτηριστικῶν, ἐνῶ στήν ἀντίθετη περίπτωση χρησιμοποιοῦμε μηχανογραφικά συστήματα (***ἠλεκτρονικό ὑπολογιστή, ὑπολογιστικές μηχανές,*** κ.λ.π.) καί τότε ἡ ὅλη ἐργασία περιλαμβάνει τίς ἑξῆς φάσεις:

α) Τήν ***κωδικογράφηση,*** κατά τήν ὁποία οἱ ἀπαντήσεις πού περιέχονται στά ἐρωτηματολόγια μετατρέπονται σέ κωδικούς ἀριθμούς πού σημειώνονται στό δεξιό ἄκρο τοῦ δελτίου (ἐρωτηματολογίου).

β) Τή ***διάτρηση*** κατά τήν ὁποία οἱ κωδικοί ἀριθμοί κάθε δελτίου (ἐρωτηματολογίου) μεταφέρονται σέ ειδικές καρτέλλες καί, μέ τή βοήθεια μιᾶς ***διατρητικῆς μηχανῆς,*** μετατρέπονται σέ μία ἤ περισσότερες ὀπές. Στή φάση αὐτή γίνεται καί ***ἐπαλήθευση*** γιά νά διαπιστώσομε (μέ εἰδικές ***ἐπαληθευτικές μηχανές***) ἄν οἱ καρτέλλες ἔχουν διατρηθεῖ σωστά.

γ) Τή ***διαλογή,*** κατά τήν ὁποία τά δελτία πού διατρήθηκαν μεταφέρονται στίς ***διαλογικές μηχανές,*** πού τά χωρίζουν καί τά ταξινομοῦν σέ διάφορες ὁμάδες, ἐνῶ συγχρόνως τά ἀριθμοῦν καί τά τοποθετοῦν χωριστά σέ εἰδικές ὑποδοχές κάθε μηχανῆς διαλογῆς.

δ) Τήν ***πινακογράφηση*** κατά τήν ὁποία μιά εἰδική μηχανή (***πινακογραφική μηχανή***) μπορεῖ νά μετρήσει τά δελτία κάθε κατηγορίας, νά διαβάσει τίς διατρήσεις, νά κάνει ὁρισμένες πράξεις (πρόσθεση, ἀφαίρεση, πολλαπλασιασμό καί διαίρεση) στά δεδομένα νά ἐκτυπώσει τά ἀποτελέσματα τῶν ὑπολογισμοῦ ἤ ὅ,τι ἄλλη πληροφορία ἀναφέρεται στά δελτία καί μᾶς ἐνδιαφέρει (ὄνομα, ἐπώνυμο, ἡμερομηνία, ἔτος γεννήσεως κ.τ.λ.).

Βλέπομε λοιπόν ὅτι ἡ βοήθεια πού προσφέρουν στή Στατιστική τά μηχανογραφικά συγκροτήματα εἶναι μεγάλη καί χωρίς αὐτά πολλές στατιστικές ἐπεξεργασίες θά ἦταν ἀδύνατες. Τά τελευταῖα χρόνια παρατηρήθηκε καί στή χώρα μας μιά τάση χρησιμοποιήσεως τέτοιων μηχανογραφικῶν συγκροτημάτων ἀκόμη καί ἀπό μεγάλες ἐπιχειρήσεις ἤ ὀργανισμούς παρόλο τό μεγάλο κόστους τους.

---

* Τόσο κατά τή δειγματοληψία ὅσο καί κατά τήν ἀπογραφή ἡ συγκέντρωση τῶν στατιστικῶν στοιχείων γίνεται μέ εἰδικά ἔντυπα, στά ὁποῖα εἶναι διατυπωμένες οἱ κατάλληλες ἐρωτήσεις. Τά ἔντυπα αὐτά λέγονται ***ἐρωτηματολόγια.***

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ

ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ

3.1 Γενικά.

Ὕστερα ἀπό τή συλλογή, τήν ἐπεξεργασία καί τήν ταξινόμηση τῶν στατιστικῶν στοιχείων, ἀκολουθεῖ ἡ συνοπτική παρουσίασή τους κατά τρόπο πού νά διευκολύνεται ἡ κατανόησή τους. Ἡ παρουσίαση τῶν στατιστικῶν στοιχείων μπορεῖ νά γίνει μέ τρεῖς τρόπους:

— ***Μέ ἐκθέσεις*** ἤ ***ἀναφορές.***

— ***Μέ πίνακες.***

— ***Μέ γραφικές παραστάσεις.***

Οἱ ***ἐκθέσεις*** ἤ ***ἀναφορές*** εἶναι κείμενα στά ὁποῖα ἀναφέρονται τά κυριότερα σημεῖα τῶν ἀντικειμένων καί σχολιάζεται ἡ σημασία τους. Ὁ ἀναγνώστης μιᾶς τέτοιας ἐκθέσεως, στήν ὁποία ἀναφέρεται συνήθως καί ἡ τεχνική πού ἀκολούθησε ἡ ἔρευνα, πρέπει νά εἶναι πολύ προσεκτικός καί νά διαθέτει μνήμη, ὥστε νά μπορεῖ νά συγκρατήσει ἤ νά συγκρίνει τά στοιχεῖα πού τόν ἐνδιαφέρουν. Καί αὐτό ἀποτελεῖ βασικό μειονέκτημα τοῦ πρώτου τρόπου παρουσιάσεως τῶν στατιστικῶν στοιχείων. Ἐδῶ θά ἀσχοληθοῦμε λεπτομερέστερα μέ τούς δύο ἄλλους τρόπους.

3.2 Στατιστικοί πίνακες.

Ἡ παρουσίαση τῶν στατιστικῶν στοιχείων σέ πίνακες, γίνεται μέ τήν κατάλληλη τοποθέτησή τους σέ στῆλες καί γραμμές κατά τέτοιο τρόπο, ὥστε ὄχι μόνο νά διευκολύνεται ἡ κατανόησή τους, ἀλλά νά γίνεται εὔκολα καί ἡ σύγκριση τῶν στοιχείων τοῦ πληθυσμοῦ πού ἐξετάζομε. Ἔτσι π.χ. ἄν θέλομε νά παρουσιάσομε τόν πληθυσμό πού εἶχε ἡ Γῆ τό 1970, εἶναι προτιμότερο, ἀντί νά ἀπαριθμήσομε τούς πληθυσμούς τῶν διαφόρων περιοχῶν, νά σχηματίσομε τόν Πίνακα 3.2.1.

Ὁ πίνακας αὐτός ὄχι μόνο παρέχει εὔκολη καί γρήγορη ἐνημέρωση, ἀλλά παρέχει καί τή δυνατότητα συγκρίσεως τῶν διαφόρων περιοχῶν τῆς γῆς.

Στή σύνταξη τῶν πινάκων ἀκολουθοῦμε ὁρισμένους κανόνες, οἱ κυριότεροι ἀπό τούς ὁποίους εἶναι:

1) ***Πάνω ἀπό τόν πίνακα γράφομε ἕνα περιληπτικό καί σαφή τίτλο,*** ὁ ὁποῖος δείχνει τό περιεχόμενό του.

2) ***Στό κάτω μέρος τοῦ πίνακα ἀναφέρομε τήν πηγή ἀπό τήν ὁποία προέρχονται τά στοιχεῖα του, ἐκτός ἄν αὐτά ἀποτελοῦν πρωτογενές ὑλικό πού ἐμφανίζεται γιά πρώτη φορά.***

3) ***Στήν κορυφή κάθε στήλης καί στήν ἀρχή κάθε γραμμῆς γράφομε συνοπτικά τή φύση τῶν στατιστικῶν στοιχείων καί τίς μονάδες μετρήσεως.***

ΠΙΝΑΚΑΣ 3.2.1.

Πληθυσμός τῆς Γῆς κατά περιοχές (1970).

Περιοχές	Πληθυσμός (σέ ἑκατομ.)	Ποσοστό (%)
Εὐρώπη	462	12,7
Ἀσία	2066	56,7
Ἀμερική	511	14,0
Ἀφρική	344	9,4
Ὠκεανία	19	0,5
Ρωσία	243	6,7
Ἄθροισμα	3645	100,0

Πηγή: Ο.Η.Ε.

4) ***Ἄν εἶναι δυνατή ἡ ἄθροιση τῶν στοιχείων μιᾶς στήλης, τό ἄθροισμα γράφεται στό κάτω μέρος της, ἐνῶ ἄν εἶναι δυνατή ἡ ἄθροιση τῶν στοιχείων μιᾶς γραμμῆς τό ἄθροισμα γράφεται στό δεξιό ἄκρο της*** (βλ. π.χ. Πίνακα 3.2.6).

Οἱ ἐκτεταμένοι καί λεπτομερεῖς πίνακες πού περιέχουν κάθε διαθέσιμη πληροφορία μιᾶς μεγάλης στατιστικῆς ἔρευνας καί χρησιμοποιοῦνται ὡς πηγές στατιστικῶν πληροφοριῶν χαρακτηρίζονται ὡς ***γενικοί πίνακες.*** Τέτοιους πίνακες κατασκευάζει στή χώρα μας ἡ Στατιστική Ὑπηρεσία κατά τίς ἀπογραφές τοῦ πληθυσμοῦ.

Ἀντίθετα, οἱ πίνακες πού παρουσιάζουν συνοπτικά τά στοιχεῖα τοῦ πληθυσμοῦ καί παρέχουν περιορισμένο πλῆθος πληροφοριῶν χαρακτηρίζονται ὡς ***συνοπτικοί πίνακες.*** Ἐδῶ θά ἀσχοληθοῦμε μόνο μέ τούς συνοπτικούς πίνακες, οἱ ὁποῖοι προέρχονται συνήθως ἀπό τούς γενικούς πίνακες καί χρησιμοποιοῦνται κυρίως στίς στατιστικές ἀναλύσεις διαφόρων φαινομένων.

3.2.1 Τύποι στατιστικῶν πινάκων.

Οἱ συνοπτικοί πίνακες, πού μᾶς ἐνδιαφέρουν, διακρίνονται κυρίως σέ δύο κατηγορίες:

α) ***Πίνακες ἁπλῆς εἰσόδου.*** Οἱ πίνακες αὐτοί μᾶς δίνουν πληροφορίες γιά ἕνα πληθυσμό, τοῦ ὁποίου τά στοιχεῖα ἐξετάζονται ὡς πρός ἕνα ποιοτικό ἤ ποσοτικό χαρακτηριστικό. Πίνακες ἁπλῆς εἰσόδου εἶναι οἱ ἑπόμενοι:

ΠΙΝΑΚΑΣ 3.2.2.

Ἐξέλιξη τοῦ πληθυσμοῦ τῆς Ἑλλάδας.

Ἔτος	Πληθυσμός
1920	5.016.886
1928	6.204.684
1940	7.344.860
1951	7.632.801
1961	8.388.553
1971	8.768.641

Πηγή Ε.Σ.Υ.Ε.

ΠΙΝΑΚΑΣ 3.2.3.

Ἀριθμός βρεφῶν πού γεννιοῦνται κατά μέσο ὅρο ἀπό κάθε γυναίκα, ἡλικίας 15 - 49 χρόνων, στίς διάφορες χῶρες τοῦ κόσμου.

Χῶρες	Βρέφη
Ἀλβανία	7,0
Τουρκία	5,1
Καναδᾶς	3,8
Ἰσραήλ	3,7
Η.Π.Α.	3,5
Γιουγκοσλαβία	3,2
Πορτογαλία	3,0
Ὀλλανδία	3,1
Φινλανδία	3,0
Γαλλία	2,7
Νορβηγία	2,8
Ἰταλία	2,4
Ἑλλάδα	2,2

Πηγή: Ο.Η.Ε. (1958)

ΠΙΝΑΚΑΣ 3.2.4.

Ἐξέταση ἑνός δείγματος 8.530 οἰκογενειῶν ὡς πρός τόν ἀριθμό τῶν τέκνων τους.

Ἀριθμός τέκνων	Ἀριθμός οἰκογενειῶν
0	200
1	1.000
2	2.500
3	2.000
4	1.600
5	700
6	300
7	150
8	80
Ἄθροισμα	8.530

ΠΙΝΑΚΑΣ 3.2.5.

Κατανομή τοῦ πληθυσμοῦ τῆς Ἑλλάδας στά διάφορα γεωγραφικά διαμερίσματα.

Γεωγραφικά διαμερίσματα	Ἀριθμός κατοίκων	Ποσοστό %
1. Περιφέρεια Πρωτ/σας	2.540.241	28,9
2. Ὑπόλοιπη Στ. Ἑλλάδα καί Εὔβοια	992.077	11,3
3. Πελοπόννησος	986.912	11,2
4. Ἰόνιοι Νῆσοι	184.443	2,1
5. Ἤπειρος	310.334	3,5
6. Θεσσαλία	659.913	7,5
7. Μακεδονία	1.890.684	21,5
8. Θράκη	329.582	3,7
9. Νῆσοι Αἰγαίου	417.813	4,7
10. Κρήτη	456.642	5,2
Ἄθροισμα	8.768.641	100

Πηγή: Ε.Σ.Υ.Ε (Ἀπογραφή 1971)

β) ***Πίνακες διπλῆς εἰσόδου.*** Οἱ πίνακες αὐτοί μᾶς παρέχουν πληροφορίες γιά ἕνα πληθυσμό, τά στοιχεῖα τοῦ ὁποίου ἐξετάζονται ὡς πρός δύο ποσοτικά ἤ ποιοτικά χαρακτηριστικά. Τέτοιοι πίνακες διπλῆς εἰσόδου εἶναι οἱ ἑπόμενοι:

ΠΙΝΑΚΑΣ 3.2.6.

Κατανομή τῆς βαθμολογίας 4961 μαθητῶν λυκείου θεωρητικῆς κατευθύνσεως στά μαθηματικά καί στήν ἔκθεση.

Βαθμοί στά μαθηματικά	Βαθμοί στήν ἔκθεση						
	⩽3	4	5	6	7	⩾8	Ἄθροισμα
⩽3	20	51	53	20	2	—	146
4	33	134	246	115	15	—	543
5	24	205	551	489	84	8	1.361
6	3	114	551	746	230	34	1.678
7	1	17	176	453	247	47	941
8⩾	—	1	22	104	124	41	292
Ἄθροισμα	81	522	1599	1927	702	130	4.961

Πηγή: Π. Κιόχου — Test δειγματοληπτικῆς ἔρευνας τοῦ Kolmogorov, Ἀθήνα 1971

ΠΙΝΑΚΑΣ 3.2.7.

Κατανομή τοῦ πληθυσμοῦ τῆς Ἑλλάδας σέ ὁμάδες ἡλικιῶν, κατά τά ἔτη 1907 - 1971.

Ὁμάδες Ἡλικιῶν	Ἔτος					
	1907	1920	1940	1951	1961	1971
0 - 14	39%	35 %	32 %	28%	26 %	25 %
15 - 64	57%	59,3%	62,5%	65%	65,5%	64,8%
65 καί πάνω	4%	5,7%	5,5%	7%	8,5%	11,2%

Πηγή: Π. Κιόχου. Τό δημογραφικό πρόβλημα τῆς Ἑλλάδας, Ἀθήνα 1975.

ΠΙΝΑΚΑΣ 3.2.8.

Κατανομή 600 μαθητών ἑνός σχολείου ὡς πρός τό φύλο καί τό χρῶμα τῶν ματιῶν τους.

Φύλο	Χρῶμα ματιῶν			
	Μαῦρα	Καστανά	Γαλανά	Ἄθροισμα
Ἀγόρια	120	190	40	350
Κορίτσια	86	120	44	250
Ἄθροισμα	206	310	84	600

Πηγή: Ὑποθετικά δεδομένα.

3.3 Πίνακες συχνοτήτων.

Ἡ παρουσίαση τῶν στατιστικῶν στοιχείων πού ἀναφέρονται σέ μία μόνο ἰδιότητα τῶν στοιχείων ἑνός πληθυσμοῦ, μπορεῖ νά γίνει μέ πίνακα ἁπλῆς εἰσόδου, στόν ὁποῖο ἡ πρώτη στήλη παρέχει μιά κατάλληλη κατάταξη ἤ ὁμαδοποίηση τῶν τιμῶν τῆς μεταβλητῆς ἰδιότητας, ἐνῶ ἡ δεύτερη στήλη περιέχει τίς «συχνότητες» τῶν τιμῶν αὐτῶν, δηλαδή τίς παρατηρήσεις πού ἀντιστοιχίζονται στήν κάθε τιμή τῆς μεταβλητῆς. Οἱ πίνακες αὐτοί ὀνομάζονται ***πίνακες συχνοτήτων*** (ἤ ***κατανομές συχνοτήτων***).

Ἄς δοῦμε τώρα πῶς κατασκευάζεται ἕνας πίνακας συχνοτήτων, ὅταν ἡ μεταβλητή εἶναι ἀσυνεχής ἤ συνεχής.

α) ***Ἀσυνεχής μεταβλητή.*** Ἄν ἡ μεταβλητή x εἶναι ἀσυνεχής καί τό πλῆθος τῶν διαφορετικῶν τιμῶν της εἶναι μικρό, τότε ὁ πίνακας συχνοτήτων ἔχει τήν μορφή τοῦ Πίνακα 3.3.1.

ΠΙΝΑΚΑΣ 3.3.1.

Τιμές μεταβλητῆς (x_i)	Ἀριθμός παρατηρήσεων (f_i)
x_1	f_1
x_2	f_2
x_3	f_3
x_4	f_4
.	.
.	.
.	.
x_i	f_i
.	.
.	.
.	.
x_ν	f_ν
Ἄθροισμα	$\sum_i f_i$

ὅπου: x_1, x_2, ..., x_ν εἶναι οἱ τιμές τῆς ἀσυνεχοῦς μεταβλητῆς (διατεταγμένες κατά τήν φυσική τους διάταξη ἀπό τή μικρότερη πρός τή μεγαλύτερη) καί $f_1, f_2, \ldots, f_i, \ldots, f_\nu$ εἶναι οἱ ἀντίστοιχες συχνότητες* τῶν τιμῶν αὐτῶν.

Γιά τήν καλύτερη κατανόηση τοῦ τρόπου κατασκευῆς ἑνός τέτοιου πίνακα συχνοτήτων, ἄς δοῦμε ἕνα παράδειγμα:

Παράδειγμα.

Οἱ παρακάτω παρατηρήσεις δίνουν τίς ἡμέρες ἄδειας 30 ὑπαλλήλων μιᾶς ἐπιχειρήσεως κατά τό θέρος τοῦ 1977.

25, 26, 24, 22, 21, 25, 26, 24, 22, 21, 26, 26, 25, 27, 25,
24, 24, 27, 22, 23, 23, 26, 25, 24, 25, 25, 27, 24, 25, 26

Νά κατασκευασθεῖ ὁ πίνακας συχνοτήτων τῶν πιό πάνω παρατηρήσεων.

Λύση:

Τοποθετοῦμε πρῶτα τίς παρατηρήσεις κατά τή φυσική τους διάταξη:

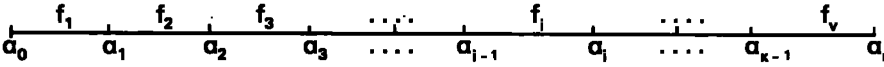
21, 21, 22, 22, 22, 23, 23, 24, 24, 24, 24, 24, 24, 25, 25,
25, 25, 25, 25, 25, 25, 26, 26, 26, 26, 26, 26, 27, 27, 27

καί μετά ὑπολογίζομε τίς συχνότητες τῶν τιμῶν 21, 22, 23,... τῆς μεταβλητῆς X = ἡμέρες ἄδειας. Ἐπειδή ἡ μεταβλητή X παίρνει μόνο 7 τιμές, θά ἔχομε τόν πίνακα συχνοτήτων 3.3.2.

ΠΙΝΑΚΑΣ 3.3.2.

Τιμές μεταβλητῆς (x_i)	Συχνότητες (f_i)
21	2
22	3
23	2
24	6
25	8
26	6
27	3
Ἄθροισμα	30

β) ***Συνεχής μεταβλητή.*** Στήν περίπτωση πού ἡ μεταβλητή X εἶναι συνεχής καί μπορεῖ νά πάρει ὁποιαδήποτε τιμή ἑνός διαστήματος (α, β), καταφεύγομε σέ ***ὁμαδοποίηση τῶν παρατηρήσεων,*** δηλαδή χωρίζομε τό διάστημα:



μεταβολῆς τῆς X σέ Κ διαδοχικά ὑποδιαστήματα (ἴσα ἤ ἄνισα) μέ τούς ἀριθμούς α_1, $\alpha_2 \ldots, \alpha_\kappa$ (ὅπως δείχνει τό παραπάνω σχῆμα) καί βρίσκομε τόν ἀριθμό f_i τῶν

---

* Γενικά, συχνότητα μιᾶς τιμῆς x_i λέγεται ὁ ἀριθμός f_i πού ἐκφράζει πόσες φορές παρατηρεῖται ἡ τιμή x_i στό πλῆθος τῶν παρατηρήσεών μας. Τό ἄθροισμα $f_1 + f_2 + \ldots + f_\nu$ ὅλων τῶν συχνοτήτων σημειώνεται συμβολικά μέ $\sum_{i=1}^{\nu} f_i$ δηλαδή εἶναι $\sum_{i=1}^{\nu} f_i = f_1 + f_2 + \ldots + f_\nu$

παρατηρήσεων πού ὑπάρχει σέ κάθε ὑποδιάστημα. Σέ ὁμαδοποίηση παρατηρήσεων καταφεύγομε καί ὅταν ἔχομε ἀσυνεχή μεταβλητή πού παίρνει πάρα πολλές τιμές καί τήν κάθε μία μέ μικρή συχνότητα.

Κάθε ὑποδιάστημα (a_{i-1}, a_i) λέγεται τώρα ***τάξη*** τῶν τιμῶν τῆς μεταβλητῆς, ἐνῶ τά ἄκρα του a_{i-1} καί a_i λέγονται ἀντίστοιχα ***κατώτερο*** καί ***ἀνώτερο*** ὅριο τῆς τάξεως. Τέλος ὁ ἀριθμός:

$$x_i = \frac{a_{i-1} + a_i}{2}$$

ὁ ὁποῖος παριστάνει τό μέσο τοῦ ὑποδιαστήματος (a_{i-1}, a_i) καί γύρω ἀπό τόν ὁποῖο βρίσκονται ὅλες οἱ παρατηρήσεις πού ἀνήκουν στό ὑποδιάστημα αὐτό, λέγεται ***κεντρική τιμή*** τῆς τάξεως. Ἡ διαφορά $a_i - a_{i-1}$ λέγεται ***πλάτος*** τῆς τάξεως καί σημειώνεται μέ δ.

Ἔτσι, ὁ πίνακας συχνοτήτων τῶν τιμῶν μιᾶς συνεχοῦς μεταβλητῆς ἔχει τή μορφή τοῦ Πίνακα 3.3.3.

ΠΙΝΑΚΑΣ 3.3.3.

Τάξεις	Κεντρικές τιμές (x_i)	Συχνότητες (f_i)
$a_0 - a_1$	x_1	f_1
$a_1 - a_2$	x_2	f_2
$a_2 - a_3$	x_3	f_3
.	.	.
.	.	.
.	.	.
$a_{i-1} - a_i$	x_i	f_i
.	.	.
.	.	.
.	.	.
$a_{\kappa-1} - a_\kappa$	x_κ	f_ν
Ἄθροισμα		$\sum_i f_i$

Μιά μεγάλη δυσκολία πού παρουσιάζεται στήν ὁμαδοποίηση τῶν παρατηρήσεων εἶναι ὁ προσδιορισμός τοῦ ἀριθμοῦ τῶν τάξεων. Δέν ὑπάρχει γενικός κανόνας γιά τόν ἀριθμό Κ τῶν τάξεων κάθε κατανομῆς καί συνήθως χρησιμοποιοῦμε τόν ἐμπειρικό τύπο (κανόνα τοῦ Sturges):

$$K = 1 + 3{,}322\ \lambda o\gamma_{10}\nu$$

πού δίνει τόν ἀριθμό Κ τῶν τάξεων, ὅπου ν εἶναι ὁ ἀριθμός τῶν παρατηρήσεων.

Στήν πράξη, ὁ ἀριθμός τῶν τάξεων προσδιορίζεται κατά τέτοιο τρόπο, ὥστε ἡ κατανομή νά παρουσιάζει ὁμοιογένεια καί ἁπλότητα. Ἡ ὁμοιογένεια ὅμως ἀπαιτεῖ διαίρεση τοῦ ὁλικοῦ διαστήματος μεταβολῆς σέ πολλές τάξεις μικροῦ πλάτους, ἐνῶ ἡ ἁπλότητα ἐπιβάλλει διαίρεση τοῦ ὁλικοῦ εὔρους σέ ὅσο τό δυνατόν λιγότερες τάξεις. Πολλοί συγγραφεῖς δέχονται ὅτι ὁ ἀριθμός τῶν τάξεων στήν ὁμαδοποίηση τῶν παρατηρήσεων δέν πρέπει νά εἶναι μικρότερος ἀπό 5, οὔτε καί μεγαλύτερος ἀπό 20.

Στίς περισσότερες περιπτώσεις (καί κυρίως όταν τό όλικό διάστημα μεταβολῆς εἶναι μικρό), τά πλάτη ὅλων τῶν τάξεων τῆς κατανομῆς εἶναι ἴσα. Σέ όρισμένες ὅμως περιπτώσεις (καί κυρίως ὅταν τό ὁλικό διάστημα μεταβολῆς εἶναι ἀρκετά μεγάλο ἤ ἄπειρο), εἴμαστε ὑποχρεωμένοι νά παίρνομε τάξεις μέ ἄνισα πλάτη. Τέτοιες συνηθισμένες περιπτώσεις στήν πράξη εἶναι οἱ κατανομές ***εἰσοδημάτων*** ἤ ***δαπανῶν.***

Γιά τήν καλύτερη κατανόηση τοῦ τρόπου κατασκευῆς ἑνός πίνακα ὁμαδοποιημένων παρατηρήσεων, ἄς δοῦμε ἕνα παράδειγμα:

Παράδειγμα.

Οἱ παρακάτω παρατηρήσεις παρέχουν τίς ταχύτητες μέ τίς ὁποῖες 40 αὐτοκίνητα πέρασαν ἀπό μία διασταύρωση.

31, 34, 45, 46, 42, 44, 50, 53
54, 55, 56, 57, 60, 61, 62, 63
64, 65, 66, 67, 68, 75, 74, 76
59, 85, 85, 84, 86, 90, 99, 88
78, 87, 92, 93, 94, 95, 96, 89,

Νά γίνει ὁμαδοποίηση τῶν παραπάνω παρατηρήσεων σέ μορφή κατανομῆς συχνοτήτων.

Λύση:

Γιά νά διευκολύνομε τήν ταξινόμηση τῶν παρατηρήσεων, κατατάσσομε αὐτές κατά τή φυσική τους διάταξη.

31, 50, 59, 65, 76, 87, 94
34, 53, 60, 66, 78, 88, 95
42, 54, 61, 67, 84, 89, 96
44, 55, 62, 68, 85, 90, 99
45, 56, 63, 74, 85, 92
46, 57, 64, 75, 86, 93

Ἐπειδή ἡ διαφορά ἀνάμεσα στή μεγαλύτερη τιμή $M = 99$ καί στή μικρότερη τιμή $E = 31$ (ἡ ὁποία ὀνομάζεται ***εὖρος τῆς κατανομῆς***) εἶναι $R = M - E = 99 - 31 = 68$, δηλαδή σχετικά μικρή, χωρίζομε τό ὁλικό διάστημα μεταβολῆς π.χ. σέ 7 τάξεις ἴσου πλάτους καί ἔχομε τόν Πίνακα συχνοτήτων 3.3.4.

ΠΙΝΑΚΑΣ 3.3.4.

Τάξεις	Κεντρικές τιμές (x_i)	Συχνότητες (f_i)
30 - 40	35	2
40 - 50	45	4
50 - 60	55	7
60 - 70	65	9
70 - 80	75	4
80 - 90	85	7
90 - 100	95	7
Ἄθροισμα		40

Σημείωση: Τά ἀνώτερα ὅρια τῶν τάξεων δέν περιέχονται σ' αὐτές.

'Η κατανομή τοῦ Πίνακα 3.3.4 ὀνομάζεται ***κλειστή***, γιατί δέν λείπει οὔτε τό κατώτερο ὅριο τῆς πρώτης τάξεως οὔτε τό ἀνώτερο τῆς τελευταίας. Ἄν σέ μία κατανομή συχνοτήτων δέν ὑπάρχει τό κατώτερο ὅριο τῆς πρώτης ἤ τό ἀνώτερο ὅριο τῆς τελευταίας τάξεως (ἤ καί τά δύο μαζί), ἡ κατανομή ὀνομάζεται ***ἀνοικτή***. Οἱ ἀνοικτές κατανομές παρουσιάζουν ὁρισμένα προβλήματα κατά τόν ὑπολογισμό τῶν στατιστικῶν παραμέτρων, ὅπως θά δοῦμε ἀργότερα.

3.4 Γραφικές παραστάσεις.

Ἕνας συνηθισμένος τρόπος παρουσιάσεως τῶν στατιστικῶν δεδομένων εἶναι οἱ γραφικές παραστάσεις, οἱ ὁποῖες μετατρέπουν τούς ἀφηρημένους ἀριθμούς σέ συγκεκριμένης μορφῆς γεωμετρικά σχήματα καί δίνουν στό φαινόμενο πού ἀναφέρονται μιά μορφή πιό κατανοητή πού διατηρεῖται εὔκολα στή μνήμη μας.

'Αντίθετα ἀπό τούς ἀριθμητικούς πίνακες, οἱ γραφικές παραστάσεις δέν πρέπει νά περιέχουν πολλές λεπτομέρειες, γιατί τότε χάνουν τό κεντρικό τους ἐνδιαφέρον. Ὅπως κάθε στατιστικός πίνακας, ἔτσι καί κάθε γραφική παράσταση πρέπει νά περιλαμβάνει (ἐκτός ἀπό τό σχέδιο) καί ὁρισμένα ἄλλα στοιχεῖα ὅπως εἶναι:

α) 'Ο ***τίτλος***.

β) 'Η ***κλίμακα τῶν τιμῶν*** τῶν μεγεθῶν πού ἀπεικονίζονται.

γ) 'Η ***ἔνδειξη τῶν πηγῶν***.

δ) Τό ***ὑπόμνημα***, τό ὁποῖο τοποθετεῖται συνήθως κάτω ἀπό τό σχῆμα καί ἐξηγεῖ τά διάφορα εἴδη γραμμῶν τῆς γραφικῆς παραστάσεως.

'Υπάρχουν πολλά εἴδη γραφικῶν παραστάσεων, ἐκεῖνα ὅμως πού χρησιμοποιοῦνται περισσότερο στήν πράξη εἶναι:

1) Τά ***ἀκιδωτά διαγράμματα***. Αὐτά χρησιμοποιοῦνται γιά τή γραφική παράσταση ποιοτικῶν μεταβλητῶν ἤ ἀσυνεχῶν ποσοτικῶν μεταβλητῶν. 'Αποτελοῦνται ἀπό ὀρθογώνια παραλληλόγραμμα πού εἶναι τοποθετημένα στόν ὁριζόντιο ἤ κάθετο ἄξονα καί τό μῆκος τους εἶναι ἀνάλογο μέ τούς ἀριθμούς πού ἀντιπροσωπεύουν.

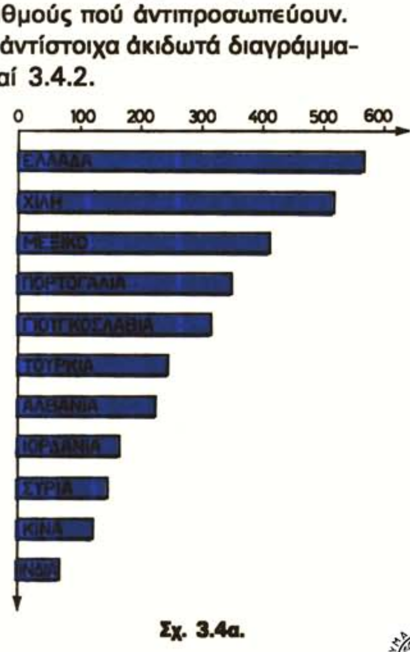
Παραδείγματα δύο στατιστικῶν πινάκων καί τά ἀντίστοιχα ἀκιδωτά διαγράμματα (σχ. 3.4α, 3.4β) παρέχουν οἱ Πίνακες 3.4.1 καί 3.4.2.

ΠΙΝΑΚΑΣ 3.4.1.

Κατά κεφαλή ἐθνικό εἰσόδημα στίς διάφορες χῶρες τοῦ κόσμου κατά τό ἔτος 1965.

Χῶρες	Κατά κεφαλή εἰσόδημα σέ δολλάρια
'Ελλάδα	566
Χιλή	515
Μεξικό	412
Πορτογαλία	351
Γιουγκοσλαβία	319
Τουρκία	244
'Αλβανία	239
'Ιορδανία	179
Συρία	156
Κίνα	147
'Ινδία	86

Πηγή Ο.Η.Ε.



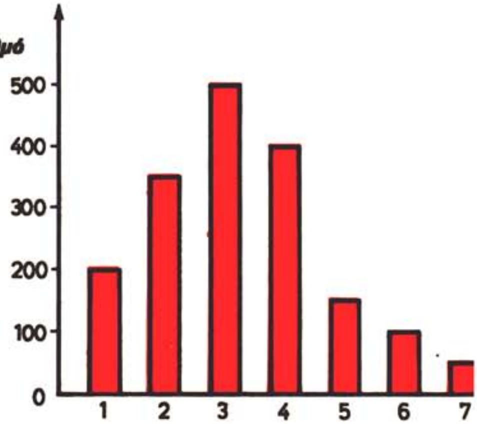
Σχ. 3.4α.

ΠΙΝΑΚΑΣ 3.4.2.

Κατανομή τῶν κατοίκων μιᾶς πόλεως ὡς πρός τόν ἀριθμό τῶν δωματίων τους.

'Αριθμός δωματίων (x_i)	'Αριθμός κατοικιῶν (f_i)
1	200
2	350
3	500
4	400
5	150
6	80
7	40
"Αθροισμα	1.720

Πηγή: 'Υποθετικά δεδομένα



Σχ. 3.4β.

2) Τά ***χρονοδιαγράμματα*** ἤ ***χρονολογικά διαγράμματα.*** Αὐτά χρησιμοποιοῦνται γιά τή γραφική παράσταση ***χρονολογικῶν σειρῶν,*** δηλαδή παρατηρήσεων πού ἀναφέρονται στή διαχρονική ἐξέλιξη ἑνός φαινομένου, ὅπως εἶναι π.χ. ὁ πυρετός τῶν ἀσθενῶν πού παρακολουθεῖται κάθε ὥρα στά νοσοκομεῖα, ἡ καθημερινή τιμή τῶν μετοχῶν, ἡ θερμοκρασία πού σημειώνεται κάθε ὥρα ἀπό τή μετεωρολογική ὑπηρεσία, τά ἀτυχήματα κατά τή διάρκεια μιᾶς σειρᾶς ἐτῶν, οἱ γάμοι, οἱ γεννήσεις καί οἱ θάνατοι κατά τήν τελευταία δεκαετία, τά κέρδη μιᾶς ἐπιχειρήσεως κατά τή διάρκεια μιᾶς χρονικῆς περιόδου κ.λ.π.

Στά χρονοδιαγράμματα τοποθετοῦμε στόν ὁριζόντιο ἄξονα τίς χρονικές στιγμές (ὧρες, μέρες, ἑβδομάδες, μῆνες, χρόνια κ.λ.π.), κατά τίς ὁποῖες ἐλήφθησαν οἱ παρατηρήσεις καί στόν κάθετο ἄξονα τίς ἀντίστοιχες τιμές τῆς μεταβλητῆς πού παρακολουθοῦμε. Οἱ τιμές τῆς μεταβλητῆς παριστάνονται ἤ μέ τίς κορυφές μιᾶς τεθλασμένης γραμμῆς ἤ πάλι μέ ὀρθογώνια παραλληλόγραμμα.

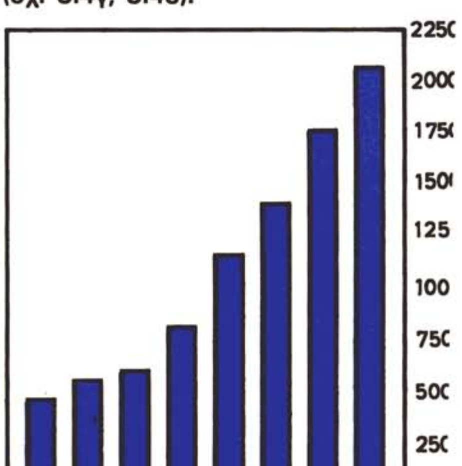
Οἱ Πίνακες 3.4.3 καί 3.4.4 παρέχουν δύο παραδείγματα χρονολογικῶν σειρῶν καί τά ἀντίστοιχα χρονοδιαγράμματά τους (σχ. 3.4γ, 3.4δ).

ΠΙΝΑΚΑΣ 3.4.3.

Κέρδη μιᾶς ἐπιχειρήσεως κατά τά ἔτη 1965 - 1972.

'Ετος	Κέρδη σέ χιλιάδες
1965	490
1966	554
1967	607
1968	782
1969	1130
1970	1375
1971	1750
1972	2050

Πηγή: 'Υποθετικά δεδομένα

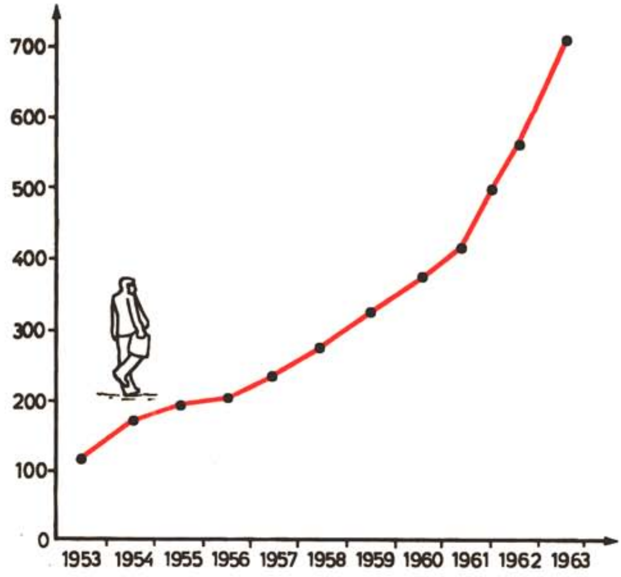


Σχ. 3.4γ.

ΠΙΝΑΚΑΣ 3.4.4.
Ἀριθμός ξένων τουριστῶν ποὺ ἐπισκέφθηκαν τὴν Ἑλλάδα τὰ ἔτη 1953 - 1963.

Ἔτος	Ἀριθμός τουριστῶν
1953	102.032
1954	160.486
1955	195.605
1956	199.707
1957	238.289
1958	257.030
1959	327.155
1960	379.959
1961	471.983
1962	572.503
1963	716.126

Πηγή: Ε.Σ.Υ.Ε



Σχ. 3.46.

3) Τά ***Κυκλικά διαγράμματα.*** Τά κυκλικά διαγράμματα χρησιμοποιοῦνται κυρίως γιά τή γραφική ἀπεικόνιση ποιοτικῶν μεταβλητῶν καί παρέχουν τήν εἰκόνα τοῦ πληθυσμοῦ μέ κυκλικούς τομεῖς τοῦ ἴδιου κυκλικοῦ δίσκου.

Ἄς δοῦμε ὡς παράδειγμα πῶς θά κατασκευάσομε ἕνα κυκλικό διάγραμμα γιά τόν ἀριθμό τῶν αὐτοκινήτων πού κυκλοφοροῦσαν στήν Ἑλλάδα κατά τό 1965, ὁ ὁποῖος δίνεται ἀπό τόν Πίνακα 3.4.5.

ΠΙΝΑΚΑΣ 3.4.5.

Κατηγορία αὐτοκινήτων	Ἀριθμός αὐτοκινήτων (f_i)
Φορτηγά	61.880
Ἐπιβατηγά	96.220
Λεωφορεῖα	8.160
Ταξί	3.740
Ἄθροισμα	170.000

Πηγή: Ε.Σ.Υ.Ε.

Γιά νά ἀπεικονίσομε τά στοιχεῖα τοῦ Πίνακα 3.4.5 μέ κυκλικό διάγραμμα, ἐργαζόμαστε ὡς ἑξῆς:

α) Μετατρέπομε τόν ἀριθμό τῶν αὐτοκινήτων κάθε κατηγορίας σέ ποσοστό ἐπί τοῦ συνόλου.

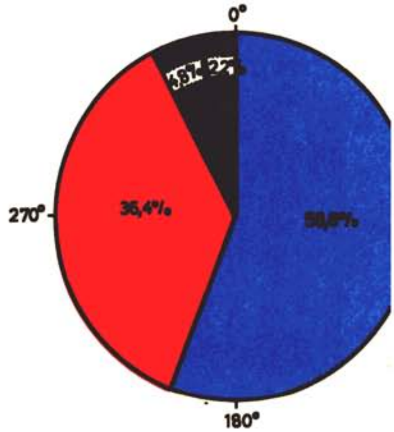
β) Κατασκευάζομε ἕνα ὁποιοδήποτε κυκλικό δίσκο καί ὑποθέτομε ὅτι τό ὑλικό ἐμβαδό του παριστάνει τό συνολικό ἀριθμό τῶν αὐτοκινήτων πού κυκλοφοροῦν. Ὁ ἀριθμός τῶν αὐτοκινήτων κάθε κατηγορίας θά παριστάνεται μέ ἀνάλογο κυκλικό τομέα.

γ) Χωρίζομε τίς 360° μοῖρες τοῦ κύκλου σέ μέρη ἀνάλογα τῶν ποσοστῶν πού βρήκαμε καί σχηματίζομε τούς κυκλικούς τομεῖς, ὁ καθένας ἀπό τούς ὁποίους παριστάνει μία κατηγορία αὐτοκινήτων.

Ἡ ὅλη ἐργασία φαίνεται στόν Πίνακα 3.4.6 μαζί μέ τόν ὁποῖο δίνεται καί τό ἀντίστοιχο κυκλικό διάγραμμα.

ΠΙΝΑΚΑΣ 3.4.6.

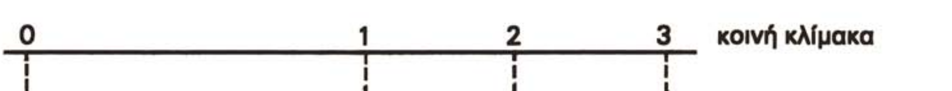
Κατηγορία αὐτοκινήτων	Ἀριθμός αὐτοκινήτων	Ποσοστό %	Μοῖρες
Φορτηγά	61.880	36,4	131°
Ἐπιβατηγά	96.220	56,6	204°
Λεωφορεῖα	8.160	4,8	17°
Ταξί	3.740	2,2	8°
	170.000	100,0	360°



ΥΠΟΜΝΗΜΑ

4) Τά ***ἡμιλογαριθμικά καί λογαριθμικά διαγράμματα.*** Αὐτά χρησιμοποιοῦνται κυρίως ὅταν οἱ τιμές τῆς μεταβλητῆς ἤ οἱ συχνότητες τῶν τιμῶν της παρουσιάζουν μεγάλες διαφορές.

Σέ μία τέτοια περίπτωση, τοποθετοῦμε στόν ἄξονα τῶν τιμῶν τῆς μεταβλητῆς (ἤ στόν ἄξονα συχνοτήτων τῶν τιμῶν της) ***λογαριθμική κλίμακα,*** δηλαδή γράφομε τίς ὑποδιαιρέσεις 1, 2, 3, 4,... τοῦ ἄξονα ὄχι σέ ἴσες ἀποστάσεις μεταξύ τους, ἀλλά σέ ἀποστάσεις ἀπό τήν ἀρχή τοῦ ἄξονα ἀνάλογες πρός τούς ἀριθμούς λογ.1, λογ.2, λογ.3, λογ.4,...



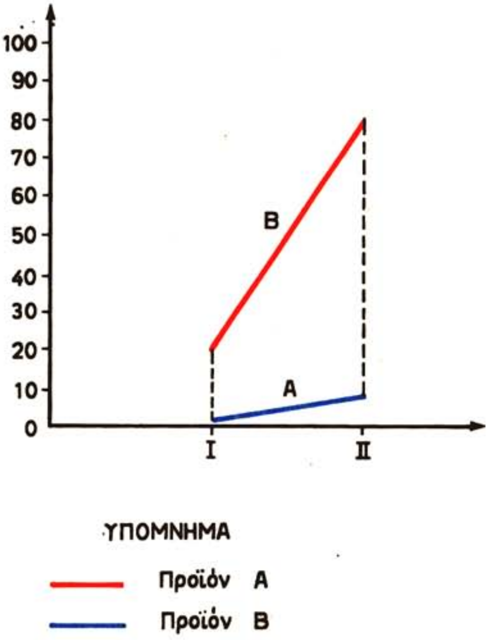
Ἕνα τέτοιο διάγραμμα πού ἔχει μόνο στόν ἕνα ἄξονα λογαριθμική κλίμακα λέγεται ***ἡμιλογαριθμικό.*** Ὅταν ἔχει καί στούς δύο ἄξονες λογαριθμική κλίμακα (δηλαδή ὅταν καί οἱ τιμές τῆς μεταβλητῆς καί οἱ συχνότητές τους παρουσιάζουν μεγάλη διαφορά) λέγεται ***λογαριθμικό.***

Ἄς πάρομε γιά παράδειγμα τίς τιμές τῶν προϊόντων Α καί Β στίς ἐποχές Ι καί ΙΙ πού δίνονται ἀπό τόν Πίνακα 3.4.7.

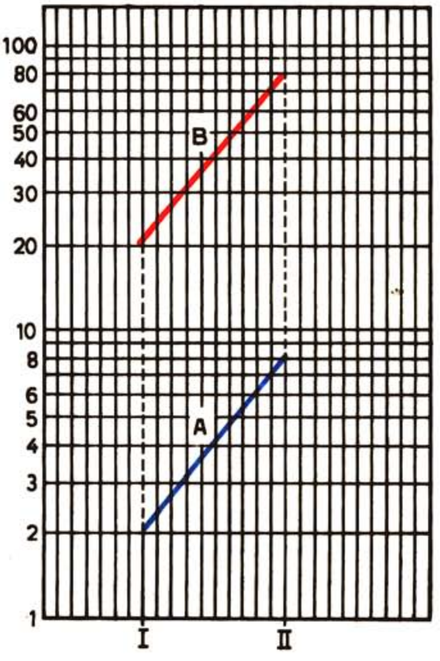
ΠΙΝΑΚΑΣ 3.4.7.

Προϊόντα	Ἐποχή I	Ἐποχή II
A	2	8
B	20	80

Στά σχήματα 3.4ε καί 3.4στ παρουσιάζονται οἱ τιμές αὐτές μέ κοινό διάγραμμα καί μέ ἡμιλογαριθμικό διάγραμμα.



Σχ. 3.4ε.



Σχ. 3.4στ.

Στό παράδειγμά μας τά δύο προϊόντα ἔχουν τό ἴδιο ποσοστό αὐξήσεως (ἀφοῦ ἡ τιμή τους τριπλασιάσθηκε), πράγμα πού δέν ἐμφανίζεται στό ἀριθμητικό διάγραμμα, ἐνῶ φαίνεται στό ἡμιλογαριθμικό διάγραμμα ὅπου τά εὐθύγραμμα τμήματα πού παριστάνουν τά A καί B εἶναι παράλληλα. Καταλαβαίνομε λοιπόν ὅτι θά χρησιμοποιοῦμε ἡμιλογαριθμικά διαγράμματα καί ὅταν ἐνδιαφερόμαστε ὄχι τόσο γιά τήν ἀπόλυτή τιμή μιᾶς μεταβλητῆς ὅσο γιά τίς ποσοστιαῖες μεταβολές τους. Ἔτσι π.χ. ὅταν παρακολουθοῦμε τό ἐθνικό εἰσόδημα, περισσότερο μᾶς ἐνδιαφέρει νά γνωρίζομε τό ποσοστό κατά τό ὁποῖο αὐξάνεται ἤ ἐλαττώνεται κάθε χρόνο παρά τό ἀπόλυτο μέγεθος τῆς αὐξήσεως ἤ τῆς μειώσεώς του.

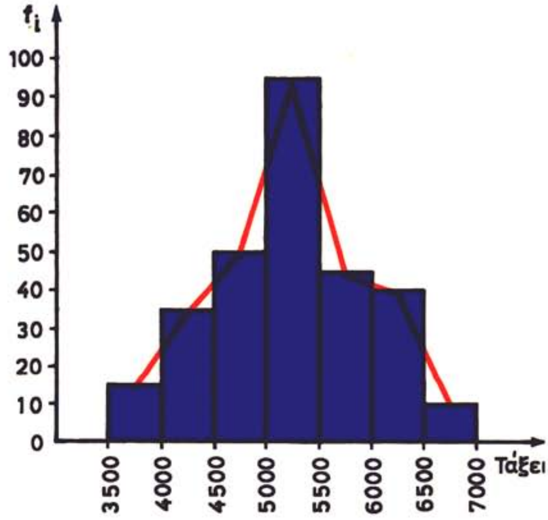
5) Τά *ἱστογράμματα.* Αὐτά χρησιμοποιοῦνται κατά κανόνα γιά τή γραφική ἀπεικόνιση ὁμαδοποιημένων παρατηρήσεων καί ἀποτελοῦνται ἀπό μία σειρά «ἐφαπτομένων» ὀρθογωνίων παραλληλογράμμων, τά ὁποῖα ἔχουν βάσεις τίς τάξεις τῆς μεταβλητῆς τοποθετημένες στόν ὁριζόντιο ἄξονα καί ἐμβαδά ἴσα μέ τίς ἀντίστοιχες συχνότητες τῶν τάξεων.

Στήν περίπτωση πού όλες οἱ τάξεις ἔχουν ἴσα πλάτη, τά ὕψη τοῦ ὀρθογωνίου εἶναι ἀνάλογα τῶν συχνοτήτων.

Στό σχῆμα 3.4ζ βλέπομε τό ἱστόγραμμα τῶν μηνιαίων ἀποδοχῶν τῶν 290 ὑπαλλήλων ἑνός ἐργοστασίου (Πίνακας 3.4.8).

ΠΙΝΑΚΑΣ 3.4.8.

Τάξεις μισθῶν	Συχνότητες (f_i)
3.500 - 4.000	15
4.000 - 4.500	35
4.500 - 5.000	50
5.000 - 5.500	95
5.500 - 6.000	45
6.000 - 6.500	40
6.500 - 7.000	10
Ἄθροισμα	290



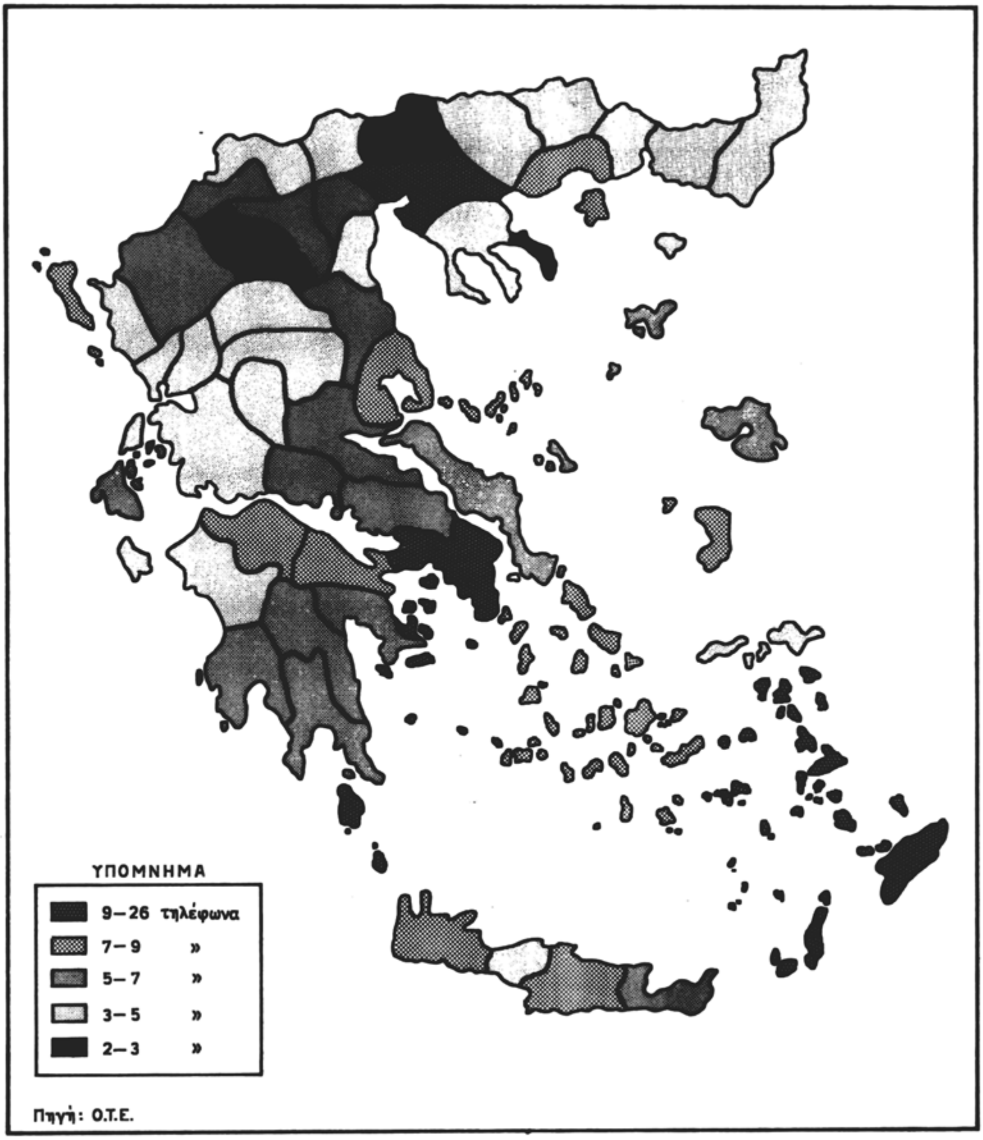
Σχ. 3.4ζ.

Ἄν ἑνώσομε τά μέσα τῶν ἄνω βάσεων ὅλων τῶν ὀρθογωνίων ἑνός ἱστογράμματος, σχηματίζεται μία τεθλασμένη γραμμή πού ὀνομάζεται ***πολύγωνο συχνοτήτων.***

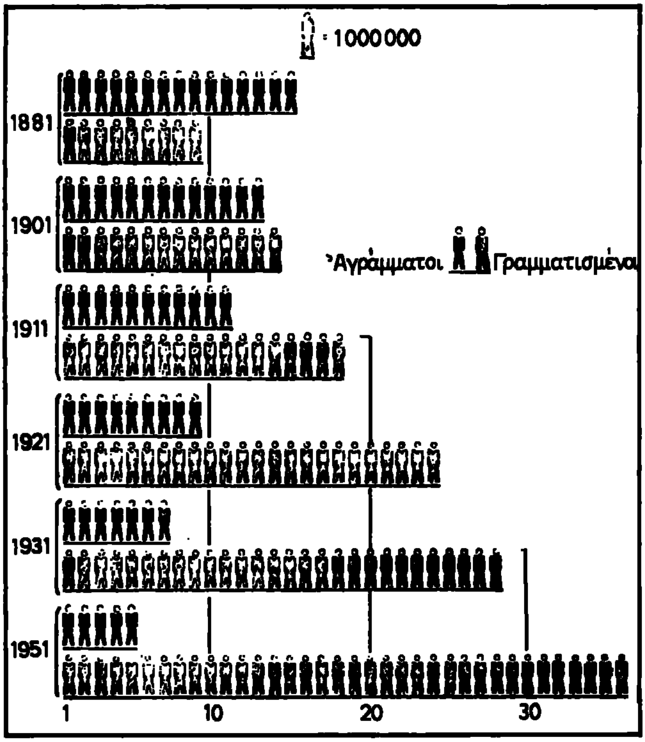
6) Τά ***χαρτογράμματα.*** Εἶναι γραφικές παραστάσεις στατιστικῶν στοιχείων σέ γεωγραφικούς ἤ τοπογραφικούς χάρτες. Οἱ τιμές τῆς μεταβλητῆς στίς διάφορει περιοχές τοῦ χάρτη δείχνονται τώρα μέ διάφορα χρώματα καί ἡ ἀντιστοιχία χρωμάτων καί τιμῶν ἐξηγεῖται στό ὑπόμνημα. Στό σχῆμα 3.4η δίνεται σέ μορφή χαρτογράμματος ἡ πυκνότητα τῶν τηλεφώνων ἀνά 100 κατοίκους στούς διαφόρους νομούς κατά τό 1972.

7) Τά ***εἰδογράμματα.*** Αὐτά εἶναι γραφικές παραστάσεις πού ἔχουν διάφορα σχήματα σέ μορφή προσώπων ἤ πραγμάτων. Χρησιμοποιοῦνται πάρα πολύ, γιατί παρέχουν πιό ἐκφραστικά τήν ἐξέλιξη ἑνός φαινομένου καί μπορεῖ ὁ ἀναγνώστης νι τή συγκρατήσει εὐκολότερα στή μνήμη του. Μέ τά εἰδογράμματα διευκολύνετα ἐπίσης καί ἡ σύγκριση δύο ἤ περισσοτέρων μεγεθῶν.

Στό σχῆμα 3.4θ βλέπομε σέ μορφή εἰδογράμματος τήν κατανομή τοῦ πληθυσμοῦ τῆς Ἰταλίας, ἡλικίας 6 χρόνων καί πάνω, σέ ἀγράμματους καί γραμματισμένους, σύμφωνα μέ τά ἀποτελέσματα τῶν ἀπογραφῶν 1881 - 1951 τῆς στατιστικῆς ὑπηρεσίας τῆς Ἰταλίας.



Σχ. 3.4η.



Σχ. 3.40.

3.5 'Ασκήσεις.

1. Νά άναφέρετε μερικά παραδείγματα ποιοτικών καί ποσοτικών μεταβλητών, καθώς καί τίς άντίστοιχες τιμές τους.

2. 'Ο παρακάτω πίνακας δίνει τήν ποσοστιαία κατανομή τοῦ πληθυσμοῦ τῆς Γῆς τό 1970 κατά περιοχές.

Περιοχές	Ποσοστό (f_i %)
Εὐρώπη	12,7
'Ασία	56,5
'Αμερική	14,0
'Αφρική	9,5
Ρωσία	6,7
'Ωκεανία	0,6
"Αθροισμα	100,0

Νά άπεικονισθοῦν τά παραπάνω δεδομένα σέ μορφή ἀκιδωτοῦ διαγράμματος. Τά σχετικά ὀρθογώνια νά εἶναι παράλληλα πρός τόν ἄξονα τῶν x.

3. 'Ο παρακάτω πίνακας δίνει τήν ποσοστιαία κατανομή τῶν νοικοκυριῶν τοῦ ἀστικοῦ πληθυσμοῦ τῆς χώρας μας, ὡς πρός τόν ἀριθμό τῶν δωματίων τους:

'Αριθμός δωματίων (x_i)	1	2	3	4	5	6	7
f_i%	17	35	26	14	5	2	1

Νά παρασταθοῦν τά παραπάνω δεδομένα σέ μορφή ἀκιδωτοῦ διαγράμματος.

4. Στόν παρακάτω πίνακα δίνεται τό γενικό ποσοστό γεννητικότητας στήν 'Ελλάδα κατά τήν περίοδο 1958 - 1969.

Χρόνια	1958	59	60	61	62	63	64	65	66	67	68	1969
Γεννήσεις	19,0	19,4	18,9	17,9	18	17,5	18	17,7	17,9	18,7	18,2	17,4

Νά παρασταθοῦν τά παραπάνω δεδομένα σέ μορφή τεθλασμένης γραμμῆς.

5. 'Η ποσοστιαία κατανομή τοῦ 'Ελληνικοῦ πληθυσμοῦ κατά περιοχές τό 1971 εἶχε ὡς ἑξῆς:

Περιοχές	Πληθυσμός	Ποσοστό %
'Αστικές	4.667.489	53,2
'Ημιαστικές	1.028.769	11,7
'Αγροτικές	3.072.383	35,1

Νά παρασταθοῦν τά παραπάνω δεδομένα σέ μορφή κυκλικοῦ διαγράμματος.

6. Τό γενικό ποσοστό θνησιμότητας (%) στή χώρα μας κατά τίς χρονικές περιόδους 1911, 1931 καί 1951 εἶχε ὡς ἑξῆς:

	Έτη		
'Ηλικία	1911	1931	1951
0	141,83	108,94	63,26
1	61,85	39,01	10,34
2	27,60	13,21	3,42
3	15,88	7,31	1,99
4	10,86	5,01	1,49

Νά παρασταθοῦν τά παραπάνω δεδομένα σέ ἡμιλογαριθμικό διάγραμμα.

7. Εἶναι δυνατό ἀπό τό ποσοστό τῶν καπνιστῶν τῶν κατοίκων τῆς Καβάλας νά ἐξαχθεῖ συμπέρασμα, ὡς πρός τό συνολικό ποσοστό τῶν καπνιστῶν τῆς χώρας;

8. 'Ο παρακάτω πίνακας δείχνει τήν παραγωγή μερικῶν γεωργικῶν προϊόντων τῆς χώρας μας.

Χρόνια	Παραγωγή (σέ χιλιάδες τόννους)			
	Σιτάρι	Ρύζι	Καπνός	Κριθάρι
1968	1078	21	112	282
1970	985	16	99	341

Νά παρασταθοῦν τά παραπάνω δεδομένα σέ δύο κυκλικά διαγράμματα.

9. Δίνεται παρακάτω τό ὕψος (σέ ἑκατοστά) 30 μαθητῶν ἑνός Λυκείου.

145	157	165
165	164	176
170	162	152
164	160	163
149	169	161
165	167	166
155	163	172
151	170	154
171	175	159
172	176	165

Νά παρασταθοῦν οἱ παραπάνω παρατηρήσεις σέ μορφή κατανομῆς συχνοτήτων. Νά χρησιμοποιηθοῦν 8 τάξεις μέ ἴσα πλάτη.

10. Δίνεται παρακάτω τό μέγεθος τῆς διαμέτρου σέ mm 50 καρφιῶν πού κατασκευάσθηκαν ἀπό τήν ἴδια μηχανή.

13,39	13,35	13,23	13,40	13,41
13,43	13,45	13,45	13,39	13,42
13,50	13,35	13,47	13,38	13,39
13,55	13,40	13,56	13,32	13,35
13,31	13,26	13,42	13,51	13,58
13,62	13,54	13,68	13,45	13,43
13,47	13,47	13,60	13,49	13,41
13,38	13,29	13,31	13,44	13,26
13,37	13,67	13,38	13,44	13,63
13,38	13,50	13,46	13,34	13,13

Νά ταξινομηθοῦν οἱ παραπάνω παρατηρήσεις σέ μορφή κατανομῆς συχνοτήτων. Γιά τόν προσδιορισμό τῶν τάξεων νά χρησιμοποιηθεῖ ὁ κανόνας τοῦ Sturges καί νά κατασκευασθεῖ τό ἱστόγραμμα συχνοτήτων.

11. Ἡ μέση ἡμερήσια θερμοκρασία σέ μία πόλη τῆς χώρας τίς μέρες τοῦ μηνός Ἀπριλίου 1977 εἶχε ὡς ἑξῆς:
15, 16, 15, 18, 20, 18, 17, 19, 16, 18, 17, 16, 15, 20
19, 20, 22, 22, 17, 20, 18, 18, 17, 18, 17, 20, 18, 16
21, 21.
Νά ταξινομηθοῦν οἱ παραπάνω παρατηρήσεις σέ μορφή ἀσυνεχοῦς κατανομῆς συχνοτήτων κα νά κατασκευασθεῖ τό ἀκιδωτό διάγραμμα.

12. Ὁ ἀριθμός τῶν τροχαίων ἀτυχημάτων πού παρατηρήθηκαν τίς μέρες τοῦ περασμένου Μαΐοι στά Γιάννενα ἔχει ὡς ἑξῆς:
1, 2, 0, 0, 0, 1, 2, 3, 4, 0, 1, 0, 5, 2, 1, 3, 0, 1
2, 1, 2, 3, 4, 1, 2, 0, 3, 2, 4, 5, 2.
Νά σχηματισθεῖ ὁ σχετικός πίνακας συχνοτήτων.

13. Δίνεται παρακάτω ἡ κατανομή τῶν ἡμερομισθίων τῶν 100 ὑπαλλήλων μιᾶς βιομηχανίας.

Τάξεις ἡμερομισθίων	f_i
350 - 360	7
360 - 370	10
370 - 380	21
380 - 390	27
390 - 400	22
400 - 410	9
410 - 420	4
Σύνολο	100

Νά κατασκευασθεῖ τό ἱστόγραμμα καί τό πολύγωνο συχνοτήτων.

14. Ἡ βαθμολογία 30 μαθητῶν μιᾶς τάξεως στήν ἔκθεση ἔχει ὡς ἑξῆς:
9, 5, 2, 3, 7, 5, 7, 9, 9, 6, 8, 9, 6, 7, 3, 4, 6, 6, 7, 3, 7, 7, 7, 8, 7, 5, 9, 7, 7, 6
Νά παρασταθοῦν οἱ παραπάνω παρατηρήσεις σέ μορφή κατανομῆς συχνοτήτων καί νά χρησιμοποιηθοῦν 4 τάξεις.

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

ΜΕΤΡΑ ΘΕΣΕΩΣ

4.1 Γενικά.

Στό προηγούμενο κεφάλαιο εἴδαμε ὅτι τό πρῶτο στάδιο τῆς στατιστικῆς ἀναλύσεως ἑνός πληθυσμοῦ, ὕστερα ἀπό τή συλλογή τῶν στατιστικῶν στοιχείων, εἶναι ἡ ταξινόμησή τους καί ἡ παρουσίασή τους σέ πίνακες συχνοτήτων. Οἱ πίνακες αὐτοί περιορίζουν βέβαια τόν ὄγκο τῶν στοιχείων πού συγκεντρώθηκαν, ἀλλά ἐξακολουθοῦν νά παρουσιάζουν μία σύνθετη εἰκόνα τῶν στοιχείων αὐτῶν, μέ τήν ὁποία δέν μποροῦμε νά συγκρίνομε εὔκολα ὁμοειδεῖς ἔρευνες σέ διαφορετικούς πληθυσμούς. Γιά τό λόγο αὐτό, θεωρεῖται πολλές φορές ἀναγκαία μία μεγαλύτερη ἀκόμη συμπύκνωση τῶν στοιχείων καί ἀντικατάστασή τους μέ ὁρισμένους ἀντιπροσωπευτικούς ἀριθμούς, πού ὀνομάζονται ***γενικά χαρακτηριστικά*** (ἤ ***παράμετροι***) τῆς κατανομῆς. Τά χαρακτηριστικά μιᾶς κατανομῆς διακρίνονται:

α) ***Σέ χαρακτηριστικά*** (ἤ ***μέτρα***) ***θέσεως.***

β) ***Σέ χαρακτηριστικά*** (ἤ ***μέτρα***) ***διασπορᾶς.***

Στό κεφάλαιο αὐτό θά ἀναφέρομε τά κυριότερα μέτρα θέσεως. Αὐτά εἶναι ἀριθμοί, γύρω ἀπό τούς ὁποίους βρίσκονται οἱ διάφορες τιμές τῆς μεταβλητῆς πού ἐξετάσαμε. Ἔτσι, κάθε μέτρο θέσεως παριστάνει ἕνα σημεῖο τοῦ ἄξονα Ox, τό ὁποῖο προσδιορίζει τή «θέση» τῶν παρατηρήσεών μας.

4.2 Μέσος ἀριθμητικός.

Ἄν μιά μεταβλητή X παίρνει τίς ν τιμές $x_1, x_2, x_3, \ldots, x_\nu$, ὁ ἀριθμός

$$\frac{1}{\nu}(x_1 + x_2 + x_3 + \ldots + x_\nu)$$

λέγεται ***μέσος ἀριθμητικός*** της (ἤ ***μέση τιμή*** της) καί συμβολίζεται μέ $\overline{x}$, δηλαδή:

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_\nu}{\nu} = \frac{1}{\nu}\sum_{i=1}^{\nu} x_i \qquad (4.1)$$

Ὁ μέσος ἀριθμητικός εἶναι τό πιό βασικό μέτρο θέσεως μιᾶς μεταβλητῆς.

Παράδειγμα 1°.

Οἱ βαθμοί τοῦ Α΄ τριμήνου ἑνός μαθητῆ τῆς Β΄ τάξεως τοῦ Τεχνικοῦ Λυκείου

στά διάφορα μαθήματα εἶναι: 16, 15, 14, 17, 14, 15, 18, 16, 13, 11.
Νά βρεθεῖ ἡ «μέση βαθμολογία» τοῦ μαθητῆ.

Λύση:

$$\bar{x} = \frac{16+15+14+17+14+15+18+16+13+11}{10} = \frac{149}{10} = 14{,}9$$

Παράδειγμα 2ο.

Τά ἡμερομίσθια 8 ἐργατῶν μιᾶς ἐπιχειρήσεως σέ δρχ. εἶναι:
550, 580, 610, 490, 420, 540, 480, 430.
Νά ὑπολογισθεῖ τό «μέσο ἡμερομίσθιο», καί νά ἑρμηνευθεῖ ἡ σημασία του.

Λύση:

Τό μέσο ἡμερομίσθιο εἶναι ὁ ἀριθμητικός μέσος τῶν ἡμερομισθίων. Δηλαδή εἶναι:

$$\bar{x} = \frac{550+580+610+490+420+540+480+430}{8} = \frac{4100}{8} = 512{,}5 \text{ δρχ.}$$

Ἄν ὅλοι οἱ ἐργάτες ἔπαιρναν τό ἴδιο ἡμερομίσθιο καί καθένας τους ἔπαιρνε 512,5 δρχ., τότε ὁ ἐπιχειρηματίας θά πλήρωνε κάθε μέρα γιά τούς 8 ἐργάτες τό ἴδιο ποσό τῶν 4100 δρχ. (8 × 512,5 = 4100 δρχ). Ἀντί ἐπομένως νά συγκρατήσομε στή μνήμη μας τά 8 διαφορετικά ἡμερομίσθια, 550, 580, 610, 490, 420, 540, 480, 430, συγκρατοῦμε μόνο ἕνα ἀριθμό, τόν $\bar{x} = 512{,}5$ ὁ ὁποῖος κατά κάποιο τρόπο ἀντικαθιστᾶ κάθε ἕνα ἀπό τά 8 διαφορετικά ἡμερομίσθια.

4.3 Εὕρεση τοῦ μέσου ἀριθμητικοῦ ἀπό πίνακα συχνοτήτων.

Ὅταν τά στατιστικά δεδομένα δίνονται μέ πίνακα συχνοτήτων, ὁ μέσος ἀριθμητικός ὑπολογίζεται ἄν πολλαπλασιάσομε κάθε τιμή τῆς μεταβλητῆς μέ τήν ἀντίστοιχη συχνότητά της (ἀφοῦ ἡ συχνότητα δείχνει πόσες φορές ἐμφανίζεται ἡ τιμή στό σύνολο τῶν παρατηρήσεων) καί τό ἄθροισμα τῶν γινομένων αὐτῶν τό διαιρέσομε μέ τό ἄθροισμα τῶν συχνοτήτων.

Ἔτσι, ἄν ἡ μεταβλητή μας παίρνει τίς τιμές $x_1, x_2, \ldots, x_\kappa$, μέ συχνότητες $f_1, f_2, \ldots, f_\kappa$, ὁ μέσος ἀριθμητικός της θά δίνεται ἀπό τόν τύπο:

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_\kappa x_\kappa}{f_1 + f_2 + \ldots + f_\kappa}$$

Γιά νά ἐφαρμόζεται εὔκολα ὁ τύπος αὐτός, συμπληρώνομε τόν πίνακα συχνοτήτων μέ μία στήλη, ἡ ὁποία περιέχει τά γινόμενα $f_i\ x_i$ (ὅπως φαίνεται στόν Πίνακα 4.3.1) καί τό ἄθροισμα τῶν γινομένων τῆς στήλης αὐτῆς εἶναι ὁ ἀριθμητής τοῦ $\bar{x}$. Τά ἀθροίσματα $x_1 f_1 + x_2 f_2 + x_3 f_3 + \ldots + f_i x_i + f_\kappa x_\kappa$ καί

$f_1 + f_2 +, \ldots, + f_i +, \ldots, + f_\kappa$ γράφονται σύντομα $\sum_i f_i x_i$ καί $\sum_i f_i$ ὁπότε:

$$\bar{x} = \frac{\sum_i f_i x_i}{\sum_i f_i} \tag{4.2}$$

ΠΙΝΑΚΑΣ 4.3.1.

Τιμές μεταβλητῆς (x_i)	Συχνότητες (f_i)	Γινόμενα $(f_i x_i)$
x_1	f_1	$f_1 x_1$
x_2	f_2	$f_2 x_2$
x_3	f_3	$f_3 x_3$
.	.	.
.	.	.
.	.	.
x_i	f_i	$f_i x_i$
.	.	.
.	.	.
x_κ	f_κ	$f_\kappa x_\kappa$
Ἄθροισμα	Σf_i	$\Sigma f_i x_i$

Παράδειγμα 1ο.

Ὁ Πίνακας 4.3.2 περιλαμβάνει στίς δύο πρῶτες στῆλες τήν ποσοστιαία κατανομή τῶν νοικοκυριῶν τῶν ἀστικῶν περιοχῶν τῆς χώρας ὡς πρός τόν ἀριθμό τῶν δωματίων:

ΠΙΝΑΚΑΣ 4.3.2.

Ἀριθμός δωματίων (x_i)	Ἀναλογία νοικοκυριῶν $(f_i\%)$	Γινόμενα $(f_i x_i)$
1	17	17
2	35	70
3	26	78
4	14	56
5	5	25
6	2	12
7	1	7
Ἄθροισμα	100	265

Ἔτσι ὁ μέσος ἀριθμητικός τῆς μεταβλητῆς x = ἀριθμός δωματίων εἶναι:

$$\bar{x} = \frac{\sum_i f_i x_i}{\sum_i f_i} = \frac{265}{100} = 2{,}65.$$

Καί αὐτό σημαίνει ὅτι κάθε νοικοκυριό μιᾶς ἀστικῆς πόλεως διαθέτει κατά μέσο ὅρο 2,65 δωμάτια.

Στήν περίπτωση πού ἔχομε πίνακα συχνοτήτων μέ ὁμαδοποιημένες παρατηρήσεις, ὡς τιμές τῆς μεταβλητῆς X στόν ὑπολογισμό τοῦ μέσου ἀριθμητικοῦ της παίρνομε τίς κεντρικές τιμές τῶν τάξεων. Στή συνέχεια ἐργαζόμασθε μέ τόν ἴδιο τρόπο, δηλαδή πολλαπλασιάζομε τίς κεντρικές τιμές x_i μέ τίς ἀντίστοιχες συχνότητες f_i κάθε τάξεως, προσθέτομε τά γινόμενα αὐτά καί τό ἄθροισμά τους $\sum_i f_i x_i$ τό διαιροῦμε μέ τό ἄθροισμα $\sum_i f_i$ τῶν συχνοτήτων.

ΠΙΝΑΚΑΣ 4.3.3.

Τάξεις	Συχνότητες (ἀπόλυτες ἤ σχετικές) f_i	Κεντρικές τιμές τῶν τάξεων x_i	Γινόμενα $f_i\,x_i$
$a_0 - a_1$	f_1	x_1	$f_1 x_1$
$a_1 - a_2$	f_2	x_2	$f_2 x_2$
$a_2 - a_3$	f_3	x_3	$f_3 x_3$
.	.	.	.
.	.	.	.
.	.	.	.
$a_{i-1} - a_i$	f_i	x_i	$f_i x_i$
.	.	.	.
.	.	.	.
.	.	.	.
$a_{\kappa-1} - a_\kappa$	f_κ	x_κ	$f_\kappa x_\kappa$
	$\sum_i f_i$		$\sum_i f_i x_i$

Ὅλη αὐτή ἡ ἐργασία φαίνεται στόν Πίνακα 4.3.3 ἀπό τόν ὁποῖο βρίσκομε τελικά:

$$\overline{x} = \frac{\sum_i f_i\,x_i}{\sum_i f_i}$$

Παράδειγμα 2ο.

Οἱ δύο πρῶτες στῆλες τοῦ Πίνακα 4.3.4 δίνουν τήν ποσοστιαία κατανομή τῆς ἡλικίας τοῦ ἑλληνικοῦ πληθυσμοῦ κατά τήν ἀπογραφή τοῦ 1971:

ΠΙΝΑΚΑΣ 4.3.4.

Τάξεις ἡλικῶν	Ἀναλογία πληθυσμοῦ (f_i %)	Κεντρικές τιμές x_i	$f_i\,x_i$
0 – 10	17	5	85
10 – 20	16	15	240
20 – 30	14	25	350
30 – 40	15	35	525
40 – 50	12	45	540
50 – 60	10	55	550
60 – 70	9	65	585
70 – 80	4	75	300
80 – 90	2	85	170
90 – 100	1	95	95
Ἄθροισμα	100		3440

Γιά νά βροῦμε τή μέση ἡλικία τοῦ ἑλληνικοῦ πληθυσμοῦ κατά τό 1971 συμπληρώσαμε τόν παραπάνω πίνακα μέ τή στήλη τῶν κεντρικῶν τιμῶν καί τῶν γινομένων $f_i\ x_i$, ὁπότε βρίσκομε:

$$\overline{x} = \frac{\sum_i f_i\ x_i}{\sum_i f_i} = \frac{3440}{100} = 34,4$$

Αὐτό σημαίνει ὅτι τό 1971 ἡ μέση ἡλικία τῶν Ἑλλήνων ἦταν 34,4 χρόνων.

4.4 Ἰδιότητες τοῦ μέσου ἀριθμητικοῦ.

Ἀπό τόν ὁρισμό τοῦ μέσου ἀριθμητικοῦ προκύπτουν ἀμέσως οἱ παρακάτω χρήσιμες ἰδιότητές του:

1. Ἄν ὅλες οἱ τιμές τῆς μεταβλητῆς X εἶναι ἴσες μέ μία σταθερή ποσότητα α, δηλ. ἄν εἶναι $x_1 = x_2 = x_3 =, \ldots, x_\nu = \alpha$, τότε καί ὁ μέσος ἀριθμητικός εἶναι ἴσος μέ τή σταθερή αὐτή ποσότητα α.

Ἀπόδειξη:

$$\overline{x} = \frac{1}{\nu}(x_1 + x_2 +, \ldots, + x_\nu) = \frac{1}{\nu}(\alpha + \alpha +, \ldots, + \alpha) = \frac{1}{\nu} . \nu\alpha = \alpha$$

2. Ἄν προσθέσομε στίς τιμές τῆς μεταβλητῆς X μία σταθερή ποσότητα x_0, τότε καί ὁ μέσος ἀριθμητικός τους αὐξάνεται κατά τή σταθερή αὐτή ποσότητα.

Ἀπόδειξη:

Ἄν θέσομε $x_i + x_0 = y_i$, ὁ μέσος ἀριθμητικός τῶν τιμῶν y_i εἶναι:

$$\overline{y} = \frac{\sum_i (x_i + x_0)}{\nu} = \frac{1}{\nu}\sum_i x_i + \frac{1}{\nu}\sum_i x_0 = \overline{x} + \frac{1}{\nu}\nu x_0 = \overline{x} + x_0$$

3. Ἄν ἀφαιρέσομε ἀπό τίς τιμές τῆς μεταβλητῆς X μία σταθερή ποσότητα x_0, τότε καί ὁ μέσος ἀριθμητικός τους μειώνεται κατά τή σταθερή αὐτή ποσότητα x_0.

Ἀπόδειξη:

Ἄν θέσομε $x_i - x_0 = y_i$, ὁ μέσος ἀριθμητικός τῶν τιμῶν y_i εἶναι:

$$\overline{y} = \frac{\sum_i (x_i - x_0)}{\nu} = \frac{\sum_i x_i - \sum_i x_0}{\nu} = \frac{\sum_i x_i}{\nu} - \frac{x_0 \nu}{\nu} = \overline{x} - x_0$$

4. Ἄν ἀφαιρέσομε ἀπό ὅλες τίς τιμές x_i τῆς μεταβλητῆς τό μέσο ἀριθμητικό τους, τό ἄθροισμα τῶν διαφορῶν $x_i - \overline{x}$ εἶναι πάντοτε μηδέν, δηλ. εἶναι:

$$\sum_i (x_i - \overline{x}) = 0$$

Ἀπόδειξη:

Ἐπειδή εἶναι $\overline{x} = \frac{1}{\nu} \sum_i x_i$, θά ἔχομε:

$$\sum_i x_i = \nu\overline{x} \text{ καί συνεπῶς}$$

$$\sum_{i=1}^{\nu} (x_i - \overline{x}) = \sum_{i=1}^{\nu} x_i - \nu\overline{x} = \nu\overline{x} - \nu\overline{x} = 0$$

5. Ἄν πολλαπλασιάσομε τίς τιμές τῆς μεταβλητῆς x μέ μία σταθερή ποσότητα λ, τότε καί ὁ μέσος ἀριθμητικός πολλαπλασιάζεται μέ τή σταθερή αὐτή ποσότητα λ.

Ἀπόδειξη:

Ἄν θέσομε $\lambda x_i = y_i$, ὁ μέσος ἀριθμητικός τῶν τιμῶν y_i εἶναι:

$$\overline{y} = \frac{\sum_i \lambda x_i}{\nu} = \lambda \frac{\sum_i x_i}{\nu} = \lambda\overline{x}$$

6. Ἄν χωρίσομε ἕνα πληθυσμό μέ ν ἄτομα σέ Κ «ὑποπληθυσμούς» πού ὁ καθένας τους περιέχει $\nu_1, \nu_2, \ldots \nu_K$ ἄτομα ἀντίστοιχα $(\nu_1 + \nu_2 + \ldots + \nu_K = \nu)$ καί ὀνομάσομε $\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots \overline{x}_K$ τούς μέσους ἀριθμητικούς τῶν τιμῶν μιᾶς μεταβλητῆς Χ στούς ὑποπληθυσμούς αὐτούς, τότε ὁ μέσος ἀριθμητικός τῶν τιμῶν τῆς Χ σέ ὅλο τόν πληθυσμό θά εἶναι:

$$\overline{x} = \frac{\nu_1\overline{x}_1 + \nu_2\overline{x}_2 + \nu_3\overline{x}_3 + \ldots + \overline{x}_K \nu_K}{\nu_1 + \nu_2 + \nu_3 + \ldots + \nu_K} = \frac{1}{\nu} \sum_i \overline{x}_i \nu_i$$

Ἀπόδειξη:

Οἱ ἀριθμοί $\nu_1 \overline{x}_1, \nu_2 \overline{x}_2, \ldots, \nu_K \overline{x}_K$ παριστάνουν τό ἄθροισμα τῶν τιμῶν τῆς μεταβλητῆς Χ στούς ὑποπληθυσμούς καί συνεπῶς ἔχουν ἄθροισμα τόν ἀριθμό $\nu\overline{x}$ πού παριστάνει τό ἄθροισμα τῶν τιμῶν τῆς μεταβλητῆς σέ ὅλο τόν πληθυσμό. Ἔχομε τότε $\nu\overline{x} = \nu_1 \overline{x}_1 + \nu_2 \overline{x}_2 + , \ldots, + \nu_K \overline{x}_K$ καί ἀπό τήν ἰσότητα αὐτή παίρνομε:

$$\overline{x} = \frac{1}{\nu} \sum_i \nu_i \overline{x}_i$$

Ἔτσι, π.χ., ἄν οἱ $\nu_1 = 400$ ἐργάτες ἑνός ἐργοστασίου ἔχουν μέσο ἡμερομίσθιο $\overline{x}_1 = 540$ δρχ. καί οἱ $\nu_2 = 250$ ἐργάτριες τοῦ ἐργοστασίου ἔχουν μέσο ἡμερομίσθιο $\overline{x}_2 = 490$ δρχ., τό μέσο ἡμερομίσθιο ἀνδρῶν καί γυναικῶν τοῦ ἐργοστασίου θά εἶναι:

$$\overline{x} = \frac{v_1 \overline{x}_1 + v_2 \overline{x}_2}{v_1 + v_2} = \frac{400 \times 540 + 250 \times 490}{400 + 250} = 520{,}77 \text{ δρχ.}$$

4.5 Ἔμμεση μέθοδος ὑπολογισμοῦ τοῦ μέσου ἀριθμητικοῦ.

Ὁ ὑπολογισμός τοῦ μέσου ἀριθμητικοῦ παρουσιάζει δυσκολία στίς πράξεις, ὅταν οἱ τιμές x_i τῆς μεταβλητῆς καί οἱ συχνότητές τους (f_i) εἶναι μεγάλοι ἀριθμοί. Στίς περιπτώσεις αὐτές, χρησιμοποιοῦμε μία ἄλλη μέθοδο γιά τόν ὑπολογισμό τοῦ μέσου ἀριθμητικοῦ, ἡ ὁποία χαρακτηρίζεται ὡς ***ἔμμεση μέθοδος*** καί στηρίζεται στίς ἰδιότητες τοῦ μέσου ἀριθμητικοῦ. Στή μέθοδο αὐτή ἀκολουθοῦμε τήν ἑξῆς πορεία:

α) Ἀφαιροῦμε ἀπό ὅλες τίς τιμές x_i ἕνα ἀριθμό x_0 (συνήθως ἀφαιροῦμε τήν τιμή τῆς x μέ τήν πιό μεγάλη συχνότητα) καί βρίσκομε τόν μέσο ἀριθμητικό $\overline{y}$ τῶν τιμῶν $y_i = x_i - x_0$.

Τότε ὅμως εἶναι (ἀπό τήν ἰδιότητα 3 τῆς παραγρ. 4.4) $\overline{y} = \overline{x} - x_0$ καί, λύνοντάς την ὡς πρός $\overline{x}$, βρίσκομε:

$$\overline{x} = x_0 + \overline{y} = x_0 + \frac{\sum_i y_i f_i}{\sum_i f_i}$$

Ὁ ἀριθμός x_0, πού ἀφαιροῦμε ἀπό τίς τιμές x_i, λέγεται ***βοηθητικός μέσος***.

β) Ἄν οἱ διαφορές $y_i = x_0 - x_0$ εἶναι ἐπίσης μεγάλοι ἀριθμοί, τίς διαιροῦμε μέ ἕνα ἀριθμό (συνήθως τόν 10 ἤ τόν 100) καί βρίσκομε τό μέσο ἀριθμητικό τῶν τιμῶν $\xi_i = \frac{y_i}{\lambda} = \frac{x_1 - x_0}{\lambda}$. Τότε ὅμως εἶναι (ἀπό τήν ἰδιότητα 5 τῆς παραγρ. 4.4)

$\overline{\xi} = \frac{1}{\lambda}\,\overline{y} = \frac{1}{\lambda}(\overline{x} - x_0)$ καί, λύνοντας ὡς πρός $\overline{x}$, βρίσκομε:

$$\overline{x} = x_0 + \lambda\overline{\xi} = x_0 + \lambda\,\frac{\sum_i f_i\,\xi_i}{\sum_i f_i\,f_i} \qquad (4.3)$$

Γιά τόν ὑπολογισμό λοιπόν τοῦ $\overline{x}$ ἀρκεῖ νά ὑπολογίσομε τό ἄθροισμα $\sum_i y_i\, f_i$ ἤ τό ἄθροισμα $\sum f_i\, \xi_i$ καί αὐτό γίνεται εὔκολα μέ προσθήκη καταλλήλων στηλῶν στόν πίνακα συχνοτήτων, ὅπως δείχνουν τά παρακάτω παραδείγματα.

Παράδειγμα 1ο.

Νά ὑπολογισθεῖ μέ τήν ἔμμεση μέθοδο ἡ «μέση ἡλικία» τοῦ ἑλληνικοῦ πληθυσμοῦ ἀπό τόν Πίνακα 4.3.4.

Λύση:

Παίρνοντας $x_0 = 45$ καί $\lambda = 10$ σχηματίζομε τόν Πίνακα 4.5.1 ὑπολογισμῶν.

ΠΙΝΑΚΑΣ 4.5.1.

Τάξεις	f_i	x_i	$y_i = x_i - 45$	$\xi_i = \frac{x_i - 45}{10}$	$f_i\ \xi_i$
0 – 10	17	5	– 40	– 4	– 68
10 – 20	16	15	– 30	– 3	– 48
20 – 30	14	25	– 20	– 2	– 28
30 – 40	15	35	– 10	– 1	– 15
40 – 50	12	45	0	0	0
50 – 60	10	55	10	1	10
60 – 70	9	65	20	2	18
70 – 80	4	75	30	3	12
80 – 90	2	85	40	4	8
90 – 100	1	95	50	5	5
Ἄθροισμα	100				– 106

Ἔτσι ἡ μέση ἡλικία θά εἶναι, ἐπειδή βρήκαμε $\sum_i f_i\ \xi_i = -106$:

$$\overline{x} = x_0 + \lambda \frac{\sum_i f_i\ \xi_i}{\sum_i f_i} = 45 + 10\ \frac{(-106)}{100} = 34{,}4 \text{ χρόνια}$$

Παράδειγμα 2ο.

Στόν Πίνακα 4.5.2 δίνεται ἡ κατανομή 100 ἑλληνικῶν ἐπιχειρήσεων ὡς πρός τά καθαρά μηνιαῖα κέρδη τους (σέ χιλιάδες δραχμές) κατά τό ἔτος 1977.

Νά ὑπολογισθεῖ ὁ ἀριθμητικός μέσος τῶν κερδῶν τους.

Λύση:

Ἐφαρμόζοντας τήν «ἄμεση μέθοδο» (δηλαδή χρησιμοποιώντας ἀπ' εὐθείας τόν ὁρισμό τοῦ ἀριθμητικοῦ μέσου) ἔχομε τόν Πίνακα 4.5.2 ὑπολογισμῶν.

ΠΙΝΑΚΑΣ 4.5.2.

Τάξεις	Συχνότητες f_i
20 – 40	10
40 – 100	20
100 – 200	40
200 – 220	30
Ἄθροισμα	100

ΠΙΝΑΚΑΣ 4.5.3.

Τάξεις	f_i	x_i	$f_i\ x_i$
20 – 40	10	30	300
40 – 100	20	70	1.400
100 – 200	40	150	6.000
200 – 220	30	210	6.300
Ἄθροισμα	100		14.000

$$\overline{x} = \frac{\sum_i x_i\ f_i}{\sum_i f_i} = \frac{14.000}{100} = 140 \text{ χιλ. δρχ.}$$

Ἐφαρμόζοντας τήν «ἔμμεση μέθοδο» μέ βοηθητικό μέσο $x_0 = 210$ καί $\lambda = 10$ ἔχομε τόν παρακάτω πίνακα ὑπολογισμῶν.

Τάξεις	f_i	x_i	$x_i - 210$	$\xi_i = \dfrac{x_i - 210}{10}$	$f_i \xi_i$
20 – 40	10	30	−180	−18	−180
40 – 100	20	70	−140	−14	−280
100 – 200	40	150	−60	−6	−240
200 – 220	30	210	0	0	0
Ἄθροισμα	100				−700

Ἔτσι, ὁ μέσος ἀριθμητικός θά ὑπολογισθεῖ τώρα μέ τόν τύπο:

$$\overline{x} = x_0 + \lambda \frac{\sum_i f_i \xi_i}{\sum_i f_i} \quad \text{καί θά εἶναι:}$$

$$\overline{x} = 210 + 10 \frac{(-700)}{100} = 140 \text{ χιλ. δρχ}$$

4.6 Διάμεσος.

Ἄν οἱ τιμές μιᾶς μεταβλητῆς Χ τοποθετηθοῦν κατά τή φυσική τους διάταξη, ἀπό τή μικρότερη πρός τή μεγαλύτερη, τότε ὁ ἀριθμός πού χωρίζει τό πλῆθος τῶν τιμῶν τῆς Χ σέ δύο ἴσες ὁμάδες λέγεται ***διάμεσος τιμή*** τῆς Χ (ἤ ἁπλῶς «διάμεσος» τῆς Χ) καί σημειώνεται μέ M_α.

Ἔτσι, τό 50% τῶν τιμῶν τῆς μεταβλητῆς εἶναι μικρότερες ἀπό τή διάμεσο τιμή M_α καί τό ἄλλο 50% τῶν τιμῶν εἶναι μεγαλύτερες ἀπό τή διάμεσο τιμή M_α ἤ ἴσες μέ αὐτή.

Γιά τόν ὑπολογισμό τῆς τιμῆς τῆς διαμέσου διακρίνομε δύο περιπτώσεις, ὅταν οἱ τιμές τῆς μεταβλητῆς δέν δίνονται σέ πίνακα συχνοτήτων.

α) Τό ***πλῆθος τῶν παρατηρήσεων εἶναι ἀριθμός περιττός.*** Τότε, ἀφοῦ τοποθετήσομε τίς τιμές τῆς μεταβλητῆς κατά τή φυσική τους διάταξη, παίρνομε γιά «διάμεσο» τή μεσαία τιμή, δηλαδή τήν τιμή πού χωρίζει σέ δύο ἴσα μέρη ὅλες τίς τιμές τῆς μεταβλητῆς Χ.

Παράδειγμα.

Πέντε μαθητές ἐξετάσθηκαν προφορικά ἀπό τόν καθηγητή τῶν Μαθηματικῶν καί πῆραν τούς βαθμούς:

13, 13, 14, 11, 15.

Λύση:

Γιά νά βροῦμε τή διάμεσο τῶν βαθμῶν αὐτῶν, τούς τοποθετοῦμε κατά τή φυσική τους διάταξη 11, 13, 13, 14, 15 καί βλέπομε ἀμέσως ὅτι «μεσαῖος ὅρος» εἶναι ὁ 3ος. Συνεπῶς εἶναι:

$$M_\alpha = 13$$

Γενικά, ἄν ἔχομε ν τιμές, ἡ θέση τοῦ «μεσαίου» ὅρου καθορίζεται ἀπό τόν ἀριθμό: $\frac{\nu + 1}{2}$.

β) Τό ***πλῆθος τῶν τιμῶν τῆς μεταβλητῆς εἶναι ἀριθμός ἄρτιος.*** Τότε, ἄν τοποθετήσομε τίς τιμές κατά τή φυσική τους διάταξη, ἔχομε δύο μεσαίους ὅρους καί παίρνομε γιά διάμεσο τό ἡμιάθροισμά τους.

Παράδειγμα.

Νά βρεθεῖ ἡ διάμεσος τῶν παρατηρήσεων: 6, 8, 12, 10, 19, 15.

Λύση:

Τοποθετοῦμε τίς παρατηρήσεις κατά τή φυσική τους διάταξη: 6, 8, 10, 12, 15, 19 καί βλέπομε ἀμέσως ὅτι μεσαῖος ὅρος εἶναι ὁ 3ος καί 4ος, δηλαδή οἱ ἀριθμοί 10 καί 12. Ἔτσι, ἡ διάμεσος εἶναι:

$$M_\alpha = \frac{10 + 12}{2} = 11.$$

Γενικά, ἄν ἔχομε ν τιμές, ὁ ἀριθμός $\frac{\nu + 1}{2}$ καθορίζει πάλι τή θέση τῆς διαμέσου, γιατί εἶναι τώρα κλασματικός ἀριθμός καί οἱ ἀκέραιοι πού τόν περιέχουν παριστάνουν τίς θέσεις τῶν δύο μεσαίων ὅρων.

4.6.1 Εὕρεση διαμέσου ἀπό πίνακα συχνοτήτων.

Ὅταν οἱ τιμές τῆς μεταβλητῆς δίνονται σέ πίνακα συχνοτήτων, γιά νά ὑπολογίσομε τή διάμεσο τιμή, συμπληρώνομε τόν πίνακα μέ μία στήλη πού ἔχει τίς ***ἀθροιστικές συχνότητες*** (ἤ ***ὁλικές συχνότητες***) τῶν τιμῶν. Ὅταν λέμε ἀθροιστική συχνότητα μιᾶς τιμῆς x_i τῆς μεταβλητῆς X, ἐννοοῦμε τά πλήθη τῶν ἀτόμων τοῦ πληθυσμοῦ στά ὁποῖα ἡ X παίρνει τιμή μικρότερη ἤ ἴση μέ x_i. Ἔτσι, ἡ ἀθροιστική συχνότητα μιᾶς τιμῆς x_i, ἡ ὁποία θά σημειώνεται F_i, θά βρίσκεται ἄν προσθέσομε στή συχνότητά της f_i τίς συχνότητες ὅλων τῶν μικροτέρων τιμῶν της $x_1, x_2, \ldots, x_i$, καί συνεπῶς θά εἶναι:

$$F_i = f_1 + f_2 + f_3 + \ldots + f_{i-1} + f_i$$

Οἱ ἑπόμενοι πίνακες εἶναι οἱ 4.3.2 καί 4.5.1 συμπληρωμένοι μέ στήλη ἀθροιστικῶν συχνοτήτων.

Εἶναι φανερό ὅτι ὅσο μεγαλώνουν οἱ τιμές τῆς X, τόσο μεγαλώνουν καί οἱ ἀθροιστικές τους συχνότητες, ἐνῶ ἡ μικρότερη τιμή x_1 ἔχει ἀθροιστική συχνότητα ἴση μέ τή συχνότητά της καί ἡ μεγαλύτερη τιμή x_ν ἔχει ἀθροιστική συχνότητα ἴση μέ τό πλῆθος ὅλων τῶν ἀτόμων τοῦ πληθυσμοῦ.

Ἀριθμός δωματίων	f_i	F_i
1	17	17
2	35	52
3	26	78
4	14	92
5	5	97
6	2	99
7	1	100

τάξεις	f_i	F_i
20 - 40	10	10
40 - 100	20	30
100 - 200	40	70
200 - 220	30	100

Ἄς ἔρθομε τώρα στόν ὑπολογισμό τῆς διαμέσου τῶν τιμῶν μιᾶς μεταβλητῆς X, ὅταν ἔχομε πίνακα συχνοτήτων. Διακρίνομε δύο περιπτώσεις:

α) *Ὅταν ἡ μεταβλητή εἶναι ἀσυνεχής*, ὁπότε:

1. Συμπληρώνομε τόν πίνακα συχνοτήτων μέ τή στήλη τῶν ἀθροιστικῶν συχνοτήτων (f_i).

2. Προσδιορίζομε τήν τιμή ν/2, ὅπου ν τό πλῆθος τῶν παρατηρήσεων.

3. Βρίσκομε τίς δύο διαδοχικές ἀθροιστικές συχνότητες F_{i-1} καί F_i, ἀνάμεσα στίς ὁποῖες περιέχεται ὁ ἀριθμός ν/2 ($F_{i-1} < ν/2 < F_i$).

4. Παίρνομε γιά διάμεσό τους $M_α$ τήν τιμή τῆς μεταβλητῆς πού ἔχει ἀθροιστική συχνότητα F_i, δηλαδή παίρνομε:

$$M_α = x_i$$

Παράδειγμα.

Ἕνα κείμενο ὑπαγορεύθηκε σέ 100 μαθητές καί ὁ Πίνακας 4.6.1 δείχνει τήν κατανομή τῶν ὀρθογραφικῶν σφαλμάτων στά 100 γραπτά.

Νά βρεθεῖ ἡ διάμεσος.

ΠΙΝΑΚΑΣ 4.6.1.

Ἀριθμός σφαλμάτων x_i	Ἀριθμός μαθητῶν (f_i)	Ἀθροιστική σειρά F_i
0	12	12
1	27	39
2	29	68
3	19	87
4	8	95
5	4	99
6	1	100
Ἄθροισμα	100	

$$\frac{ν}{2} = 50$$

Ἐπειδή ἡ κατανομή εἶναι ἀσυνεχής καί ν/2 = 50, βρίσκομε τίς δύο ἀθροιστικές συχνότητες πού περιέχουν τόν ἀριθμό 50. Αὐτές εἶναι οἱ ἀριθμοί 39 καί 68 τῆς τελευταίας στήλης. Ἡ τιμή $x = 2$ τῆς μεταβλητῆς πού ἀντιστοιχεῖ στή μεγαλύτερη ἀθροιστική συχνότητα 68 εἶναι ἡ διάμεσος. Δηλαδή $M_\alpha = 2$.

Καταλαβαίνομε λοιπόν ὅτι τό 50% τῶν μαθητῶν ἔκανε 2 σφάλματα καί κάτω καί τό ἄλλο 50% δυό ἤ περισσότερα σφάλματα.

β) ***Ὅταν ἡ μεταβλητή εἶναι συνεχής,*** ὁπότε ἔχομε ὁμαδοποίηση τῶν παρατηρήσεων καί:

1. ***Συμπληρώνομε τόν πίνακα συχνοτήτων μέ τή στήλη τῶν ἀθροιστικῶν συχνοτήτων*** (F_i).
2. ***Προσδιορίζομε τήν τιμή ν/2, ὅπου ν τό πλῆθος τῶν παρατηρήσεων.***
3. ***Βρίσκομε τίς δύο διαδοχικές ἀθροιστικές συχνότητες F_{i-1} καί F_i πού περιέχουν τόν ἀριθμό ν/2.***
4. ***Ἄν (a_{i-1}, a_i) εἶναι ἡ τάξη πού ἀντιστοιχεῖ στή μεγαλύτερη ἀθροιστική συχνότητα F_i, παίρνομε γιά διάμεσο τόν ἀριθμό:***

$$M_\alpha = \alpha_{i-1} + \frac{\delta}{f_i}\left(\frac{\nu}{2} - F_{i-1}\right) \tag{4.4}$$

ὅπου: α_{i-1} τό κατώτερο ὅριο τῆς τάξεως στήν ὁποία ἐντοπίζεται ἡ διάμεσος, f_i ἡ συχνότητα τῆς τάξεως στήν ὁποία ἐντοπίζεται ἡ διάμεσος, F_{i-1} ἡ ἀθροιστική συχνότητα τῆς τάξεως πού προηγεῖται ἐκείνης στήν ὁποία ἐντοπίζεται ἡ διάμεσος, δ τό πλάτος τοῦ διαστήματος τάξεως στήν ὁποία ἐντοπίζεται ἡ διάμεσος, ν ὁ συνολικός ἀριθμός συχνοτήτων.

Παράδειγμα.

Στόν Πίνακα 4.6.2 δίνεται ἡ ταχύτητα μέ τήν ὁποία πέρασαν ἀπό μία διασταύρωση 100 αὐτοκίνητα.

Νά βρεθεῖ ἡ διάμεσος ταχύτητα.

ΠΙΝΑΚΑΣ 4.6.2.

Τάξεις	f_i	F_1	
30 - 40	5	5	
40 - 50	12	17	
50 - 60	19	36	
			$\left.\right\} \frac{\nu}{2} = 50$
60 - 70	30	66	
70 - 80	17	83	
80 - 90	10	93	
90 - 100	7	100	
Ἄθροισμα	100		

Λύση:

Ἐπειδή ν/2 = 50, καί ὁ ἀριθμός 50 βρίσκεται ἀνάμεσα στίς δύο διαδοχικές ἀθροιστικές συχνότητες 36 καί 66, θά ἔχομε $F_{i-1} = 36$, $F_i = 66$. Ἔτσι, ἡ διάμεσος θά ἀνήκει στήν τάξη 60 - 70, ἡ ὁποία ἀντιστοιχεῖ στή μεγαλύτερη ὁλική συχνότητα

66. Ἔχομε λοιπόν:

$a_{i-1} = 60,\ a_i = 70,\ \delta = a_i - a_{i-1} = 10,\ f_i = 30$, ὁπότε ἡ διάμεσος θά εἶναι:

$$M_a = a_{i-1} + \frac{\delta}{f_i}\left(\frac{\nu}{2} - F_{i-1}\right) = 60 + \frac{10}{30}(50 - 36) = 64{,}66 \text{ km}.$$

Αὐτό σημαίνει ὅτι τά 50% τῶν αὐτοκινήτων πού πέρασαν εἶχαν ταχύτητα 64,66 km καί κάτω καί τά ἄλλα 50% 64,66 καί πάνω.

4.7 Τεταρτημόρια (ἤ τεταρτοτόμοι).

Ἄν οἱ τιμές μιᾶς μεταβλητῆς τοποθετηθοῦν κατά τή φυσική τους διάταξη (ἀπό τή μικρότερη πρός τή μεγαλύτερη), οἱ ἀριθμοί πού χωρίζουν τό πλῆθος τῶν τιμῶν τῆς X σέ 4 ἴσα μέρη λέγονται ***τεταρτημόρια*** (ἤ ***τεταρτοτόμοι***) καί θά σημειώνονται Q_1, Q_2, Q_3.

Ἀπό τόν ὁρισμό μας καταλαβαίνομε ὅτι:

- Ἡ θέση τοῦ πρώτου τεταρτημορίου Q_1 καθορίζεται ἀπό τόν ἀριθμό $\frac{\nu+1}{4}$.
- Ἡ θέση τοῦ δευτέρου τεταρτημορίου Q_2 καθορίζεται ἀπό τόν ἀριθμό $\frac{2(\nu+1)}{4} = \frac{\nu+1}{2}$.
- Ἡ θέση τοῦ τρίτου τεταρτημορίου Q_3 καθορίζεται ἀπό τόν ἀριθμό $\frac{3(\nu+1)}{4}$.

Εἶναι φανερό ὅτι τό δεύτερο τεταρτημόριο εἶναι ἡ διάμεσος, δηλ. $Q_2 = M_a$. Ἐπίσης, τό πρῶτο τεταρτημόριο Q_1 εἶναι ἀριθμός τέτοιος, ὥστε τά 25% τῶν τιμῶν τῆς μεταβλητῆς εἶναι μικρότερες ἤ ἴσες μέ τό Q_1 καί τά ὑπόλοιπα 75% μεγαλύτερες ἤ ἴσες μέ αὐτό. Τό τρίτο τεταρτημόριο Q_3 εἶναι ἀριθμός τέτοιος, ὥστε τά 75% τῶν τιμῶν εἶναι μικρότερες ἤ ἴσες μέ τό Q_3 καί τά 25% μεγαλύτερες ἤ ἴσες μέ αὐτό.

Παράδειγμα.

Νά βρεθοῦν τό πρῶτο καί τρίτο τεταρτημόριο μιᾶς μεταβλητῆς X, πού παίρνει τίς τιμές 15, 6, 4, 10, 12, 5, 20, 10, 4, 13, 7

Λύση:

Τοποθετοῦμε πρῶτα τίς τιμές τῆς μεταβλητῆς κατά τή φυσική τους διάταξη: 4,4,5,6,7,10,10,12,13,15,20 καί στή συνέχεια ὑπολογίζομε τούς ἀριθμούς

$$\frac{\nu+1}{4} = \frac{11+1}{4} = 3 \text{ καί } \frac{3(\nu+1)}{4} = \frac{3(11+1)}{4} = 9.$$

Καταλαβαίνομε λοιπόν ὅτι τό πρῶτο τεταρτημόριο Q_1 εἶναι ἡ τιμή πού ἔχει τήν 3η θέση καί τό τρίτο τεταρτημόριο ἡ τιμή πού ἔχει τήν 9η θέση, δηλαδή:

$$Q_1 = 5, \quad Q_3 = 13$$

Ἄς δοῦμε τώρα πῶς ὑπολογίζονται τά τεταρτημόρια Q_1 καί Q_3 ἀπό πίνακα συχνοτήτων.

Διακρίνομε δύο περιπτώσεις:

α) *Όταν ή μεταβλητή είναι άσυνεχής*, τότε ό ύπολογισμός τοῦ πρώτου καί τρίτου τεταρτημορίου γίνεται μέ τρόπο άνάλογο πρός τόν ύπολογισμό τῆς διαμέσου. Συμπληρώνομε δηλαδή τόν πίνακα μέ τίς άθροιστικές συχνότητες τῶν τιμῶν καί βρίσκομε τίς διαδοχικές άθροιστικές συχνότητες πού περιέχουν τούς άριθμούς $\frac{\nu}{4}$ (γιά τό πρῶτο τεταρτημόριο) καί $\frac{3\nu}{4}$ (γιά τό τρίτο τεταρτημόριο).

Παράδειγμα.

Ό Πίνακας 4.7.1 δίνει τήν κατανομή 100 ήμερῶν ώς πρός τά τροχαῖα δυστυχήματα, πού έγιναν σέ μία όρισμένη πόλη.

Νά βρεθοῦν τό πρῶτο καί τρίτο τεταρτημόριο.

ΠΙΝΑΚΑΣ 4.7.1.

Δυστυχήματα x_i	Μέρες f_i	Άθροιστική σειρά F_i
0	42	42
1	36	78
2	14	92
3	6	98
4	2	100
Άθροισμα	100	

Λύση:

Έπειδή $\nu/4 = 100/4 = 25$ καί ό άριθμός 25 περιέχεται στίς άθροιστικές συχνότητες 0 καί 42, τό Q_1 θά είναι ἴσο μέ τήν τιμή τῆς X πού άντιστοιχεῖ στή μεγαλύτερη άθροιστική συχνότητα 42, δηλαδή:

$$Q_1 = 0.$$

Έπειδή $3\nu/4 = 3 \times 100/4 = 75$ καί ό άριθμός 75 περιέχεται στίς άθροιστικές συχνότητες 42 καί 78, τό Q_3 θά είναι ἴσο μέ τήν τιμή τῆς X πού άντιστοιχεῖ στή μεγαλύτερη άθροιστική συχνότητα 78, δηλαδή:

$$Q_3 = 1.$$

β) *Όταν ή μεταβλητή είναι συνεχής.* Στήν περίπτωση αύτή, κατά τήν όποία έχομε όμαδοποίηση παρατηρήσεων, συμπληρώνομε τόν πίνακα συχνοτήτων μέ τίς άθροιστικές συχνότητες τῶν τιμῶν καί έντοπίζομε πάλι μέ τούς άριθμούς $\nu/4$ καί $3\nu/4$, τίς τάξεις στίς όποῖες βρίσκονται άντιστοίχως οί άριθμοί Q_1 καί Q_3. Άν ύποθέσομε $(a_{i-r} - a_i)$ τήν τάξη πού βρίσκεται ό καθένας καί σημειώσομε μέ f_i τή συχνότητά της καί F_i τήν όλική της συχνότητα, τό πρῶτο καί τρίτο τεταρτημόριο θά ύπολογίζονται άπό τούς τύπους:

$$Q_1 = \alpha_{i-1} + \frac{\delta}{f_i}\left(\frac{\nu}{4} - F_{i-1}\right) \qquad (4.5)$$

$$Q_3 = \alpha_{i-1} + \frac{\delta}{f_i}\left(\frac{3\nu}{4} - F_{i-1}\right) \qquad (4.6)$$

Παράδειγμα.

Νά ὑπολογισθεῖ τό πρῶτο καί τρίτο τεταρτημόριο ἀπό τά δεδομένα τοῦ Πίνακα 4.7.2.

ΠΙΝΑΚΑΣ 4.7.2.

Τάξεις	f_i%	Ἀθροιστική σειρά F_i	
0 - 10	17	17	$\frac{\nu}{4} = 25$
10 - 20	16	33	
20 - 30	14	47	
30 - 40	15	62	
40 - 50	12	74	$\frac{3\nu}{4} = 75$
50 - 60	10	84	
60 - 70	9	93	
70 - 80	4	97	
80 - 90	2	99	
90 - 100	1	100	
Ἄθροισμα	100		

Ἐπειδή $\nu/4 = 100/4 = 25$ καί ὁ ἀριθμός 25 βρίσκεται ἀνάμεσα στίς ἀθροιστικές συχνότητες 17 καί 33, τό πρῶτο τεταρτημόριο θά βρίσκεται στήν τάξη 10 - 20 καί θά εἶναι (ἀφοῦ $\alpha_{i-1} = 10$, $\delta = 10$, $f_i = 16$ καί $F_{i-1} = 17$):

$$Q_1 = 10 + 10\,\frac{25 - 17}{16} = 15 \text{ χρόνια.}$$

Ἡ τιμή $Q_1 = 15$ σημαίνει ὅτι τό 25% τοῦ ἑλληνικοῦ πληθυσμοῦ ἔχει ἡλικία μικρότερη ἤ ἴση τῶν 15 ἐτῶν καί τό 75% μεγαλύτερη ἤ ἴση τῶν 15 ἐτῶν.

Ἐπειδή $3\nu/4 = 3 \times 100/4 = 75$ καί ὁ ἀριθμός 75 βρίσκεται ἀνάμεσα στίς ἀθροιστικές συχνότητες 74 καί 84, τό τρίτο τεταρτημόριο θά βρίσκεται στήν τάξη 50 - 60 καί θά εἶναι (ἀφοῦ τώρα $\alpha_{i-1} = 50$, $f_i = 10$, $\delta = \alpha_i - \alpha_{i-1} = 10$ καί $F_{i-1} = 74$):

$$Q_3 = 50 + \frac{10}{10}(75 - 74) = 51 \text{ χρόνια}$$

Αὐτό σημαίνει ὅτι τό 75% τοῦ ἑλληνικοῦ πληθυσμοῦ ἔχει ἡλικία μικρότερη ἤ ἴση τῶν 51 ἐτῶν καί τό 25% μεγαλύτερη ἤ ἴση τῶν 51 ἐτῶν.

4.8 Δεκατημόρια (ἤ δεκατόμοι).

Ἄν τοποθετήσομε τίς τιμές μιᾶς μεταβλητῆς κατά τή φυσική τους διάταξη (ἀπό τή μικρότερη στή μεγαλύτερη), ***δεκατημόρια*** ὀνομάζονται ἐνιά ἀριθμοί D_1, D_2, D_3, ..., D_9, πού χωρίζουν τό πλῆθος τῶν τιμῶν σέ δέκα ἴσα μέρη.

Ἔτσι, τό πρῶτο δεκατημόριο θά εἶναι ἕνας ἀριθμός τέτοιος, ὥστε τό 10% τῶν παρατηρήσεων νά εἶναι μικρότερες ἤ ἴσες μέ τό D_1, τό δεύτερο δεκατημόριο θά εἶναι ἀριθμός τέτοιος, ὥστε τό 20% τῶν παρατηρήσεων νά εἶναι μικρότερες ἤ ἴσες μέ τό D_2 καί τό 80% τῶν παρατηρήσεων νά εἶναι μεγαλύτερες ἤ ἴσες μέ τό D_2, ... κ.ο.κ.

Παράδειγμα.

Οἱ βαθμοί 14 μαθητῶν στό μάθημα τῆς Φυσικῆς εἶναι κατά τή φυσική τους διάταξη:
9, 9, 10, 11, 11, 11, 12, 12, 15, 16, 16, 17, 18, 18.

Νά ὑπολογισθεῖ τό 6ο δεκατημόριο.

Λύση:

Ἐπειδή ἡ θέση τοῦ 6ου δεκατημορίου καθορίζεται ἀπό τόν ἀριθμό $\frac{6(\nu + 1)}{10} =$

$= \frac{6 \times 15}{10} = 9$, τό ζητούμενο δεκατημόριο θά εἶναι ὁ 9ος ὅρος, δηλαδή $D_6 = 15$.

Ὁ ὑπολογισμός τῶν δεκατημορίων ἀπό πίνακα συχνοτήτων γίνεται κατά τρόπο ἀνάλογο πρός τόν ὑπολογισμό τῶν τεταρτημορίων. Ἔτσι, ἄν ἔχομε ὁμαδοποιημένες παρατηρήσεις, συμπληρώνομε τόν πίνακα μέ τίς ἀθροιστικές συχνότητες καί ἐντοπίζομε τίς δύο διαδοχικές ἀθροιστικές συχνότητες F_{i-1} καί F_i πού περιέχουν τόν ἀριθμό κν/10. Ἄν (a_{i-1}, a_i) εἶναι ἡ τάξη πού ἔχει ἀθροιστική συχνότητα F_i καί f_i εἶναι ἡ συχνότητά της, τό δεκατημόριο D_κ θά ὑπολογίζεται μέ τόν τύπο:

$$D_\kappa = a_{i-1} + \frac{\delta}{f_i}\left(\frac{\kappa\nu}{10} - F_{i-1}\right) \qquad (4.7)$$

Παράδειγμα.

Ὁ Πίνακας 4.8.1 δείχνει τούς βαθμούς πού πῆραν σέ ἕνα τέστ ἐξυπνάδας 40 μαθητές.

Νά βρεθεῖ τό 6ο δεκατημόριο

Λύση:

Ἐπειδή 6ν/10 = 6 × 40/10 = 24 καί ὁ ἀριθμός 24 περιέχεται μεταξύ τῶν ἀθροιστικῶν συχνοτήτων 22 καί 26, τό 6ο δεκατημόριο θά βρίσκεται στήν τάξη 70-80. Ἔχομε λοιπόν $a_{i-1} = 70$, $a_i = 80$, $f_i = 4$, $F_{i-1} = 22$, $\delta = 10$ καί συνεπῶς:

ΠΙΝΑΚΑΣ 4.8.1.

Τάξεις (βαθμοί)	Μαθητές f_i	F_i
30 - 40	2	2
40 - 50	5	7
50 - 60	7	14
60 - 70	8	22
70 - 80	4	26
80 - 90	8	34
90 - 100	6	40
Ἄθροισμα	40	

$$\frac{6\nu}{10} = 24$$

$$D_6 = 60 + \frac{10}{4}(24-22) = 60 + 5 = 65.$$

Αὐτό σημαίνει ὅτι τό 60% τῶν μαθητῶν (δηλαδή 24 μαθητές) πῆραν βαθμό μικρότερο ἤ ἴσο μέ 65 μονάδες καί τό 40% (δηλαδή 16 μαθητές) μεγαλύτερο ἤ ἴσο μέ 65 μονάδες.

4.9 Ἀσκήσεις.

1. Οἱ βαθμοί ἑνός μαθητῆ στά διάφορα μαθήματα τοῦ πρώτου ἑξαμήνου εἶναι:
19, 15, 18, 12, 16, 14, 17, 13, 15, 9, 11.
Νά ὑπολογισθοῦν ἡ μέση βαθμολογία, ἡ διάμεσος βαθμολογία καί τό πρῶτο τεταρτημόριο.

2. Ἡ θερμοκρασία στίς 15 Ἰανουαρίου σέ ὀκτώ πόλεις τῆς Ἑλλάδας ἦταν:
−2,5, 1,5 −0,5 , +0,3, +0,5, +2, +2,8 +3
Νά ὑπολογισθεῖ ἡ μέση θερμοκρασία, ἡ διάμεσος θερμοκρασία καί τό τρίτο τεταρτημόριο.

3. Ὁ μέσος μισθός μιᾶς κατηγορίας ὑπαλλήλων τοῦ δημοσίου εἶναι 12.000 δρχ. Ἄν ὁ μισθός κάθε ὑπαλλήλου αὐξηθεῖ κατά 10%, ποιά μεταβολή ἐπέρχεται στό μέσο μισθό;

4. Ἡ μέση βαθμολογία στά μαθηματικά σέ μία τάξη ἑνός Λυκείου εἶναι 12. Ἄν ὁ καθηγητής ἀποφασίσει νά αὐξήσει τή βαθμολογία κάθε μαθητῆ κατά 2 μονάδες, ποιά μεταβολή ἐπέρχεται στή μέση βαθμολογία;

5. Ἡ κατανομή συχνοτήτων τῶν σπίρτων χωρίς κεφαλή σέ κάθε κουτί τῶν 50 σπίρτων καί σέ σύνολο 100 κουτιῶν ἔχει ὡς ἑξῆς:

Ἀριθμός ἀκεφάλων σπίρτων x_i	Ἀριθμός κουτιῶν f_i
0	12
1	27
2	29
3	19
4	8
5	4
6	1
Ἄθροισμα	100

Ζητεῖται: α) Ὁ μέσος ἀριθμητικός. β) Ἡ διάμεσος, τό πρῶτο καί τρίτο τεταρτημόριο.

6. Στόν παρακάτω πίνακα δίνονται τά μηνιαῖα καθαρά κέρδη 58 ἐπιχειρήσεων

τάξεις	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100 - 110
f_i	2	9	15	13	11	8

Ζητεῖται: α) Τό ἱστόγραμμα καί πολύγωνο συχνοτήτων. β) Ὁ μέσος ἀριθμητικός.

7. Ἡ ποσοστιαία κατανομή τῶν νοικοκυριῶν μιᾶς πόλεως, σχετικά μέ τόν ἀριθμό τῶν ἐργαζομένων μελῶν τους εἶναι:

Ἐργαζόμενα μέλη (x_i)	Ποσοστό (f_i%)
0	2
1	50
2	20
3	15
4 καί ἄνω	13
Ἄθροισμα	100

Νά ὑπολογισθεῖ ὁ ἀριθμητικός μέσος, ἄν γνωρίζομε ὅτι αὐτός εἶναι διπλάσιος ἀπό τή διάμεσο.

8. Δίνεται ἡ παρακάτω ποσοστιαία κατανομή τῶν ἡμερομισθίων 500 ὑπαλλήλων μιᾶς μεγάλης ἐπιχειρήσεως.

Τάξεις ἡμερομίσθ.	350 - 360	360 - 370	370 - 380	380 - 390	390 - 400
ποσοστό f_i%	8	28	44	16	4

Ζητεῖται: α) Ἡ διάμεσος καί τό τρίτο δεκατημόριο. β) Τό ποσοστό τῶν ὑπαλλήλων πού παίρνει:
i) 375 δρχ. καί κάτω.
ii) 365 δρχ. καί ἄνω
iii) μεταξύ 368 καί 395 δρχ.

9. Τά προϊόντα πού παράγονται ἀπό ἕνα ἐργοστάσιο συσκευάζονται σέ κιβώτια τῶν 500 ἀντικειμένων. Σέ ἕνα ἔλεγχο 114 κιβωτίων, σχετικά μέ τόν ἀριθμό τῶν ἐλαττωματικῶν ἀντικειμένων, προέκυψαν τά παρακάτω δεδομένα:

Ἀριθμός Ἐλαττωματικῶν	5	10	15	20	25	30	35
Ἀριθμός δειγμάτων (f_i)	7	13	29	42	16	5	2

Νά ὑπολογισθεῖ ὁ μέσος ἀριθμητικός καί τό 7° δεκατημόριο.

10. Τά νοικοκυριά τῶν δημοσίων ὑπαλλήλων μιᾶς πόλεως ἔχουν κατά μέσο ὅρο $\overline{x} = 2{,}63$ δωμάτια. Ἀπό τόν πίνακα συχνοτήτων ὅμως τῆς μεταβλητῆς Χ = ἀριθμός δωματίων ἔχουν σβησθεῖ οἱ συχνότητες τῶν τιμῶν $x = 2$ καί $x = 3$.

Ἀριθμός δωματίων (x_i)	Ποσοστό (f_i %)
1	16
2	—
3	—
4	26
Ἄθροισμα	100

Νά ὑπολογισθεῖ ἡ διάμεσος καί τό τρίτο τεταρτημόριο.

11. Σέ ἕνα ἐργοστάσιο ἐργάζονται 100 εἰδικευμένοι ἄνδρες μέ μέσο ἡμερομίσθιο $\overline{x}_1 = 520$ δρχ., 70 εἰδικευμένες γυναῖκες μέ μέσο ἡμερομίσθιο $\overline{x}_2 = 380$ δρχ. καί 40 μαθητευόμενα παιδιά μέ μέσο ἡμερομίσθιο $\overline{x}_3 = 220$ δρχ.
Νά βρεθεῖ τό μέσο ἡμερομίσθιο ὅλων τῶν ἐργατῶν μαζύ.

12. Τά ἡμερομίσθια τῶν ἐργατῶν μιᾶς ἐπιχειρήσεως κατανέμονται ὡς ἑξῆς:

Τάξεις	f_i	F_i
100 - 110	30	30
110 - 120	80	110
120 - 130	120	230
130 - 140	150	380
140 - 150	75	455
150 - 160	65	520
Ἄθροισμα	520	

Νά ὑπολογισθεῖ: α) Ὁ μέσος ἀριθμητικός μέ τήν ἔμμεσο μέθοδο. β) Τό πρῶτο καί τρίτο τεταρτημόριο. γ) Τό 8ο δεκατημόριο.

13. Ἀπό τήν ἐξέταση 500 πακέττων σιγαρέττων μιᾶς καπνοβιομηχανίας βρέθηκε ὅτι ὁ μέσος ἀριθμητικός τῶν «σκάρτων» σιγαρέττων σέ κάθε πακέττο ἦταν 3. Ἡ κατανομή τῶν πακέττων αὐτῶν ὡς πρός τόν ἀριθμό τῶν σκάρτων σιγαρέττων δίνεται ἀπό τόν παρακάτω πίνακα, ἀπό τόν ὁποῖο ἔχουν σβησθεῖ οἱ συχνότητες τῶν τιμῶν 1 καί 3.

Σκάρτα σιγαρέττα (x_i)	Ἀριθμός πακέττων (f_i)
0	35
1	—
2	90
3	—
4	85
5	60
6	30
Ἄθροισμα	500

Νά ὑπολογισθεῖ τό πρῶτο καί τρίτο τεταρτημόριο.

14. Ἡ κατανομή τῶν μαθητῶν ἑνός Λυκείου ὡς πρός τίς ὧρες μελέτης τους κάθε ἑβδομάδα ἔχει μέσο ἀριθμητικό $\overline{x} = 25$. Ἄν ὑποθέσομε ὅτι ὁ κάθε μαθητής αὐξάνει τό χρόνο τῆς ἑβδομαδιαίας μελέτης του κατά 3 ὧρες, νά ὑπολογισθεῖ ὁ νέος ἀριθμητικός μέσος τῆς ἑβδομαδιαίας μελέτης.

15. Τά νοικοκυριά μιᾶς πόλεως ἀποτελοῦνται κατά μέσο ὅρο ἀπό τέσσερα ἄτομα (ἀριθμητικός μέσος). Ἀπό τόν παρακάτω πίνακα συχνοτήτων τῆς μεταβλητῆς X = ἀριθμός ἀτόμων ἔχουν σβη-

σθεῖ οἱ συχνότητες τῶν τιμῶν $x = 2$ καί $x = 4$.

Ἀριθμός ἀτόμων (x_i)	Ποσοστό νοικοκυριῶν % (f_i)
1	7
2	—
3	18
4	—
5	17
6	12
7	6
Ἄθροισμα	100

Νά ὑπολογισθεῖ τό τρίτο τεταρτημόριο καί τό 9ο δεκατημόριο.

16. Ἀπό τήν ἔρευνα 200 νοικοκυριῶν βρέθηκε ὅτι ἡ μέση (ἀριθμητικός μέσος) ἡμερήσια δαπάνη αὐτῶν εἶναι $\overline{x} = 141$ δρχ. Ἡ κατανομή τῶν νοικοκυριῶν ὡς πρός τίς δαπάνες κάθε μέρας ἔχει ὡς ἑξῆς:

Δαπάνες (σέ δρχ.)	Ἀριθμός (f_i)
0 - 40	40
40 - 100	70
100 - 200	60
200 - 400	21
400 καί ἄνω	9
Ἄθροισμα	200

Νά ὑπολογισθεῖ τό τρίτο καί τό ἕβδομο δεκατημόριο.

17. Οἱ παρακάτω ἀριθμοί δίνουν τό ὕψος τῶν μαθητῶν μιᾶς τάξεως ἑνός Λυκείου, σέ ἑκατοστά:

170	149	155	156	160	162	163	168	177
190	170	174	163	158	176	164	167	165
183	181	180	160	169	172	165	171	177
180	182	149	152	157	174	168	159	153

Νά ὁμαδοποιήσετε τίς παραπάνω παρατηρήσεις σέ τάξεις μέ ἴσα πλάτη καί νά ὑπολογίσετε τόν μέσο ἀριθμητικό καί τό πρῶτο τεταρτημόριο.

18. Δίνεται παρακάτω ἡ κατανομή τῶν εἰσοδημάτων (σέ χιλιάδες δρχ.) 600 οἰκογενειῶν.

Τάξεις	f_i
50 - 100	110
100 - 200	400
200 - 300	90
Ἄθροισμα	600

Ζητεῖται: α) Ὁ μέσος ἀριθμητικός. β) Ἡ διάμεσος καί τό τρίτο τεταρτημόριο.

19. Ὁ παρακάτω πίνακας δίνει τήν κατανομή τῶν 50 οἰκογενειῶν μιᾶς πολυκατοικίας ὡς πρός τόν ἀριθμό τῶν παιδιῶν τους.

'Αριθμός παιδιῶν (x_i)	Οἰκογένειες (f_i)
0	8
1	9
2	25
3	5
4	2
5	1
Ἄθροισμα	50

Νά ὑπολογισθεῖ ὁ μέσος ἀριθμητικός καί ἡ διάμεσος.

20. Οἱ παρακάτω παρατηρήσεις δίνουν τή βαθμολογία τοῦ Α ἑξαμήνου τῆς Β τάξεως ἑνός Λυκείου στήν Ἔκθεση.

10	13	11	16	18	16	12	14	12
9	10	9	14	10	11	12	13	18
16	15	17	19	13	8	9	10	12
15	16	11	10	14	15	12	11	14

Ζητεῖται: α) Νά ὁμαδοποιήσετε μέ ἴσα πλάτη τίς παραπάνω παρατηρήσεις καί νά κατασκευάσετε τό ἱστόγραμμα συχνοτήτων. β) Νά ὑπολογίσετε τό μέσο ἀριθμητικό καί τό τρίτο τεταρτημόριο.

21. Σέ ἐρώτηση τῶν μαθητῶν μιᾶς τάξεως, γιά τό πόσα δωμάτια ἔχει ἡ κατοικία τους, δόθηκαν οἱ παρακάτω ἀπαντήσεις:

2, 3, 2, 1, 4, 3, 2, 3, 3, 4, 5, 2, 3, 4, 1, 5, 3, 2, 3, 4, 3, 4, 2, 3, 5, 4, 3, 2, 3, 4, 3, 2, 5, 3.

Ζητεῖται: α) Νά κάνετε τόν πίνακα συχνοτήτων τῶν παραπάνω παρατηρήσεων. β) Νά βρεῖτε τό μέσο ἀριθμητικό καί τή διάμεσο.

22. Ὁ παρακάτω πίνακας δείχνει τήν κατανομή τῶν ἡλικιῶν μιᾶς ὁμάδας 264 ἀτόμων.

Ἡλικία σέ ἔτη	Ἀριθμ. ἀτόμων (f_i)
15 - 25	34
25 - 35	38
35 - 45	53
45 - 55	55
55 - 65	46
65 - 75	27
75 - 85	11
Ἄθροισμα	264

Νά ὑπολογισθεῖ ἡ μέση ἡλικία μέ τήν ἔμμεσο μέθοδο.

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

5.1 Ἡ ἔννοια τῆς διασπορᾶς.

Ὁ μέσος ἀριθμητικός, ἡ διάμεσος τιμή καί τά ἄλλα μέτρα θέσεως, τά ὁποῖα ἐξετάσαμε στό προηγούμενο κεφάλαιο, ἐντοπίζουν τή «θέση» ἑνός πληθυσμοῦ, γιατί παριστάνουν σημεῖα γύρω ἀπό τά ὁποῖα βρίσκονται οἱ τιμές τῆς μεταβλητῆς πού ἐρευνοῦμε. Ἡ σημασία κάθε μιᾶς ἀπό τίς πιό πάνω παραμέτρους εἶναι τόσο μεγαλύτερη ὅσο μεγαλύτερη ὁμοιογένεια παρουσιάζει ὁ πληθυσμός. Ἀντίθετα, ἄν ὁ πληθυσμός παρουσιάζει μεγάλη ἀνομοιογένεια, τότε ἡ σημασία τῶν μέτρων θέσεως εἶναι μειωμένη. Ἄς θεωρήσομε π.χ. δύο μεταβλητές Χ καί Υ οἱ ὁποῖες παίρνουν τίς τιμές:

x: 10, 43, 43, 46, 47, 48, 50, 50, 52, 52, 54

y: 7, 14, 15, 23, 38, 48, 50, 50, 75, 85, 90.

Παρατηροῦμε ὅτι ἄν καί οἱ δύο μεταβλητές ἔχουν τόν ἴδιο ἀριθμητικό μέσο 45 καί τήν ἴδια διάμεσο 48, ἡ μία μεταβλητή διαφέρει ἀπό τήν ἄλλη, γιατί οἱ τιμές τῆς Χ κυμαίνονται μεταξύ τῶν ἀριθμῶν 10 καί 54, ἐνῶ οἱ τιμές τῆς Υ μεταξύ τῶν 7 καί 90.

Βλέπομε λοιπόν ὅτι οἱ πληροφορίες πού μᾶς παρέχουν τά μέτρα θέσεως μιᾶς κατανομῆς εἶναι ἀνεπαρκεῖς, γιατί δέν μᾶς ἐξηγοῦν πόσο συγκεντρωμένες ἤ πόσο διασπαρμένες εἶναι οἱ τιμές τῆς μεταβλητῆς γύρω ἀπό τά μέτρα αὐτά. Ἑπομένως, ἐκτός ἀπό τά μέτρα θέσεως, θά πρέπει νά προσδιορίσομε καί ἄλλους ἀριθμούς οἱ ὁποῖοι νά μᾶς δίνουν τό βαθμό συγκεντρώσεως ἤ ***διασπορᾶς*** τῶν τιμῶν τῆς μεταβλητῆς γύρω ἀπό κάποιο μέτρο θέσεως. Οἱ ἀριθμοί ***πού ἐκφράζουν γενικά πόσο διασκορπισμένες ἤ πόσο συγκεντρωμένες εἶναι οἱ παρατηρήσεις μας, ὀνομάζονται παράμετροι διασπορᾶς*** ἤ ***μέτρα διασπορᾶς***.

Τά μέτρα διασπορᾶς πού χρησιμοποιοῦμε συνήθως στή στατιστική εἶναι:

— Τό εὖρος μεταβολῆς.
— Ἡ μέση ἀπόκλιση.
— Ἡ διακύμανση καί ἡ τυπική ἀπόκλιση.
— Ὁ συντελεστής μεταβλητικότητας.

Θά ἐξετάσομε τώρα καθένα ἀπό τά μέτρα αὐτά ξεχωριστά καί θά δοῦμε ποῦ χρησιμοποιεῖται.

5.2 Τό εὖρος μεταβολῆς.

Εὖρος μεταβολῆς μιᾶς μεταβλητῆς ὀνομάζεται ἡ διαφορά ἀνάμεσα στή μεγαλύτερη καί τή μικρότερη τιμή της.

Παράδειγμα.

Ἡ μέση θερμοκρασία σέ 17 πόλεις τῆς χώρας στίς 25 Αὐγούστου τοῦ 1978 ἦταν κατ' αὐξανόμενο μέγεθος:

19, 20, 20, 23, 24, 24, 25, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35.

Λύση:

Τό εὖρος μεταβολῆς τῆς μεταβλητῆς X = μέση θερμοκρασία 25ης Αὐγούστου 1978 εἶναι ἡ διαφορά ἀνάμεσα στή μεγαλύτερη καί μικρότερη θερμοκρασία, δηλαδή εἶναι:

$$R = M - E = 35 - 19 = 16 \text{ βαθμοί.}$$

Τό μέτρο αὐτό τῆς διασπορᾶς ἔχει τό βασικό μειονέκτημα νά μήν ἐξαρτᾶται ἀπό ὅλες τίς τιμές τῆς μεταβλητῆς, ἀλλά μόνο ἀπό τίς ἀκραῖες τιμές της. Ἔτσι, ἄν μία ἀπό δύο ἀκραῖες τιμές ἔχει προκύψει ἀπό λανθασμένη μέτρηση ἤ παρατήρηση, ἔχομε ψεύτικη εἰκόνα τῆς διασπορᾶς τῶν τιμῶν.

Τό εὖρος μεταβολῆς χρησιμοποιεῖται κυρίως στή μετεωρολογία καί στά χρηματιστήρια.

5.3 Μέση ἀπόκλιση.

Ἄν μιά μεταβλητή X παίρνει τίς ν τιμές $x_1, x_2, \ldots, x_\nu$ πού ἔχουν μέσο ἀριθμητικό $\overline{x}$, ***μέση ἀπόκλιση*** αὐτῆς ὀνομάζεται ὁ μέσος ἀριθμητικός τῶν ἀπολύτων τιμῶν ὅλων τῶν διαφορῶν $x_1 - \overline{x}, x_2 - \overline{x}, \ldots, x_\nu - \overline{x}$. Ἄν σημειώσομε λοιπόν τή μέση ἀπόκλιση μέ Μ.Α., θά ἔχομε:

$$\text{Μ.Α.} = \frac{\sum_i |x_i - \overline{x}|}{\nu} \qquad (5.1)$$

Ἔτσι π.χ., ἄν μία μεταβλητή X παίρνει τίς τιμές 3, 5, 7, 9, 16, ὁ μέσος ἀριθμητικός εἶναι:

$\overline{x} = \dfrac{1}{5}(3+5+7+9+16) = \dfrac{40}{5} = 8$ καί συνεπῶς ἡ μέση ἀπόκλισή της εἶναι:

$$\text{Μ.Α.} = \frac{\sum_i |x_i - \overline{x}|}{\nu} = \frac{|3-8| + |5-8| + |7-8| + |9-8| + |16-8|}{5} = 3{,}6$$

Στήν περίπτωση πού οἱ τιμές τῆς X δίνονται σέ πίνακα συχνοτήτων καί $f_1, f_2, \ldots f_\kappa$ εἶναι οἱ συχνότητες τῶν (διαφορετικῶν μεταξύ τους) τιμῶν $x_1, x_2, \ldots x_\kappa$, ὁ τύπο 5.1. γράφεται:

$$\text{Μ.Α} = \frac{\sum_i f_i |x_i - \overline{x}|}{\sum_i f_i} \qquad (5.2$$

Γιά τόν ὑπολογισμό τῆς Μ.Α. ἀπό τόν τύπο αὐτό, ἀφοῦ βροῦμε τόν μέσο ἀριθμητικό $\overline{x}$, συμπληρώνομε τόν πίνακα συχνοτήτων μέ στῆλες πού περιέχουν τίς διαφορές $|x_i - \overline{x}|$ καί τά γινόμενα $f_i\,|x_i - \overline{x}|$.

Παράδειγμα.

Στόν παρακάτω πίνακα δίνεται ἡ κατανομή τῶν 21 μαθητῶν μιᾶς τάξεως ὡς πρός τίς ἀπουσίες τους.

Νά βρεθεῖ ἡ μέση ἀπόκλιση τῶν ἀπουσιῶν.

Λύση:

Σχηματίζομε τόν πιό κάτω πίνακα ἀριθμητικῶν ὑπολογισμῶν.

ΠΙΝΑΚΑΣ 5.3.1.

Ἀπουσίες	Μαθητές f_i	Κεντρικές τιμές x_i	Γινόμενα $f_i\,x_i$	Ἀποκλίσεις $\lvert x_i - \overline{x}\rvert$	Γινόμενα $f_i\lvert x_i - \overline{x}\rvert$
1 – 3	2	2	4	4	8
3 – 5	5	4	20	2	10
5 – 7	7	6	42	0	0
7 – 9	5	8	40	2	10
9 – 11	2	10	20	4	8
Ἄθροισμα	21		126		36

Ἐπειδή $\sum_i f_i\,x_i = 126$, ὁ μέσος ἀριθμητικός εἶναι:

$$\overline{x} = \frac{\sum_i f_i\,x_i}{\sum_i f_i} = \frac{126}{21} = 6.$$

Ἄν λοιπόν συμπληρώσομε τόν πίνακα μέ στήλη τῶν διαφορῶν $|x_i - \overline{x}|$ καί στήλη τῶν γινομένων $f_i\,|x_i - \overline{x}|$ βρίσκομε μία μέση ἀπόκλιση:

$$\text{M.A} = \frac{\sum_i f_i\,|x - \overline{x}|}{\sum_i f_i} = \frac{36}{21} = 1{,}714$$

Ἡ μέση ἀπόκλιση δίνει τό βαθμό συγκεντρώσεως τῶν παρατηρήσεών μας γύρω ἀπό τόν μέσο ἀριθμητικό $\overline{x}$.

Παρατηροῦμε ὅτι τό ἄθροισμα τῶν διαφορῶν $x_1 - \overline{x}$, $x_2 - \overline{x}$, . . ., $x_\nu - \overline{x}$ εἶναι πάντοτε μηδέν καί γι' αὐτό ὁ ἀριθμός $\sum_i (x_i - \overline{x})$ δέν προσφέρεται γιά μέτρο διασπορᾶς. Αὐτός εἶναι ὁ λόγος πού πήραμε γιά μέτρο διασπορᾶς τό ἄθροισμα $\sum_i |x_i - \overline{x}|$ τῶν ἀπολύτων τιμῶν τῶν διαφορῶν.

5.4 Διακύμανση καί τυπική ἀπόκλιση.

Ὁ βαθμός συγκεντρώσεως τῶν τιμῶν $x_1, x_2, \ldots, x_\nu$ μιᾶς μεταβλητῆς X γύρω ἀπό τόν μέσο ἀριθμητικό της $\overline{x}$ μπορεῖ νά δοθεῖ ὄχι μόνο μέ τή βοήθεια τῶν θετικῶν ἀριθμῶν $|x_1 - \overline{x}|, |x_2 - \overline{x}|, \ldots, |x_\nu - \overline{x}|$ ἀλλά καί μέ τή βοήθεια τῶν θετικῶν ἀριθμῶν $(x_1 - \overline{x})^2, (x_2 - \overline{x})^2 \ldots (x_\nu - \overline{x})^2$. Ὁ μέσος ἀριθμητικός τῶν ἀριθμῶν αὐτῶν εἶναι ἐπίσης ἕνα μέτρο διασπορᾶς, τό ὁποῖο λέγεται ***διακύμανση τῆς μεταβλητῆς***. ***Ἑπομένως, διακύμανση μιᾶς μεταβλητῆς ὀνομάζεται ὁ μέσος ἀριθμητικός τῶν τετραγώνων τῶν ἀποκλίσεων τῶν τιμῶν ἀπό τόν ἀριθμητικό μέσο της.***

Ἡ διακύμανση μιᾶς μεταβλητῆς Χ θά σημειώνεται μέ V (X) ἤ μέ s_X^2 ἤ ἁπλῶς μέ s^2, δηλαδή:

$$s^2 = V(x) = \frac{1}{\nu}\left[(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_\nu - \overline{x})^2\right] = \frac{1}{\nu}\sum_i (x_i - \overline{x})^2 \qquad (5.3)$$

Ἔτσι π.χ., ἄν μία μεταβλητή Χ παίρνει τίς τιμές 2, 3, 5, 8, 12, ἡ διακύμανσή της θά εἶναι:

ἀφοῦ ὁ ἀριθμητικός μέσος της εἶναι: $\overline{x} = \dfrac{2+3+5+8+12}{5} = 6$

$$s^2 = V(X) = \frac{1}{5}(2-6)^2 + (3-6)^2 + (5-6)^2 + (8-6)^2 + (12-6)^2 + (12-6)^2) =$$

$$= \frac{1}{5}\,(16 + 9 + 1 + 4 + 36) = \frac{66}{5} = 13{,}2.$$

Οἱ μονάδες, στίς ὁποῖες ἐκφράζεται ἡ διακύμανση, εἶναι τά τετράγωνα τῶν μονάδων στίς ὁποῖες ἐκφράζονται οἱ τιμές τῆς μεταβλητῆς (ἄν π.χ. οἱ τιμές τῆς μεταβλητῆς ἐκφράζονται σέ ἑκατοστά, ἡ διακύμανση ἐκφράζεται σέ ἑκατοστά στό τετράγωνο). Γιά τό λόγο αὐτό, παίρνομε συνήθως ὡς μέτρο διασπορᾶς ὄχι τή διακύμανση, ἀλλά τήν τετραγωνική ρίζα τῆς διακυμάνσεως, ἡ ὁποία ἐκφράζεται στίς ἴδιες μονάδες πού ἐκφράζονται καί οἱ τιμές τῆς μεταβλητῆς μας.

Τό μέτρο αὐτό διασπορᾶς ὀνομάζεται ***τυπική ἀπόκλιση*** καί σημειώνεται* μέ s_X ἤ ἁπλῶς μέ s. Ἔχομε λοιπόν:

$$s = \sqrt{V(X)} = \sqrt{\frac{1}{\nu}\sum_i (x_i - \overline{x})^2} \qquad (5.4$$

Παράδειγμα.

Βρήκαμε ὅτι ἡ διακύμανση μιᾶς μεταβλητῆς Χ πού παίρνει τίς τιμές 2, 3, 5, 8, 12, εἶναι V (X) = 13,2. Συνεπῶς, ἡ τυπική ἀπόκλιση τῆς Χ εἶναι:

$$s = \sqrt{V(X)} = \sqrt{13{,}2} = 3{,}63.$$

Ἄν ἡ τιμή τῆς τυπικῆς ἀποκλίσεως εἶναι μεγάλη, αὐτό σημαίνει ὅτι οἱ παρατηρήσεις μας εἶναι πολύ διασπαρμένες γύρω ἀπό τό μέσο ἀριθμητικό $\overline{x}$, ἐνῶ μικρή τιμή

---

* Ἐπίσης ἡ τυπική ἀπόκλιση σημειώνεται καί μέ σ, ἐνῶ ἡ διακύμανση, πού εἶναι τό τετράγωνο τῆς τυπικῆς ἀποκλίσεως, σημειώνεται καί μέ σ^2.

τῆς τυπικῆς ἀποκλίσεως σημαίνει ὅτι οἱ παρατηρήσεις μας εἶναι συγκεντρωμένες γύρω ἀπό τή μέση τιμή τους.

Εἶναι φανερό ὅτι ἡ τυπική ἀπόκλιση s παίρνει τιμές μεταξύ 0 καί ∞.

Ὑπολογισμός τῆς διακυμάνσεως καί τῆς τυπικῆς ἀποκλίσεως.

Παρατηροῦμε τώρα ὅτι τό ἄθροισμα $\sum\limits_i (x_i - \overline{x})^2$, πού ἐμφανίζεται στούς τύπους (5.3) καί (5.4), μπορεῖ νά γραφεῖ (ἐπειδή, ὅπως γνωρίζομε $\sum\limits_i x_i = v\overline{x}$):

$$\sum_i (x_i - \overline{x})^2 = \sum_i (x_i^2 - 2\,x_i\,\overline{x}^2 + \overline{x}^2) =$$

$$\sum_i x_i^2 - 2\overline{x}\left(\sum_i x_i\right) + v\overline{x}^2 =$$

$$= \sum_i x_i^2 - 2v\overline{x}^2 + vx^2 =$$

$$= \sum_i x_i^2 - v\overline{x}^2$$

Ἔτσι, οἱ δύο τύποι (5.3) καί (5.4) γράφονται ἀκόμη:

$$V(X) = \frac{1}{v}\left(\sum_i x_i^2\right) - \overline{x}^2 \qquad (5.5)$$

$$s_x = \sqrt{\frac{1}{v}\left(\sum_i x_i^2\right) - \overline{x}^2} \qquad (5.6)$$

Βλέπομε λοιπόν ὅτι μέ τόν ὑπολογισμό τῆς διακυμάνσεως (καί τῆς τυπικῆς ἀποκλίσεως) δέν χρειάζεται νά ὑπολογίσομε τά τετράγωνα τῶν διαφορῶν $x_i - \overline{x}$ ἀλλά μόνο τά τετράγωνα τῶν τιμῶν x_i^2 καί συνήθως γιά τόν ὑπολογισμό τῶν V (X) καί s χρησιμοποιοῦμε τούς τύπους (5.3) καί (5.5). Ἔτσι π.χ. ἡ διακύμανση τῆς x_i πού παίρνει τίς τιμές 2, 3, 5, 8, 12, εἶναι (ἀφοῦ $\overline{x} = 6$):

$$s_x^2 = V(X) = \frac{1}{5}(2^2 + 3^2 + 5^2 + 8^2 + 12^2) - 6^2 = \frac{246}{5} - 36 = 49{,}2 - 36 = 13{,}2.$$

Ὑπολογισμός τῆς διακυμάνσεως ἀπό πίνακα συχνοτήτων.

Στήν περίπτωση κατά τήν ὁποία οἱ τιμές τῆς μεταβλητῆς δίνονται μέ πίνακα συχνοτήτων καί $f_1, f_2, \ldots, f_K$ εἶναι οἱ συχνότητες τῶν (διαφορετικῶν μεταξύ τους) τιμῶν $x_1, x_2, \ldots, x_K$ τῆς X, οἱ τύποι (5.3) καί (5.5) τῆς διακυμάνσεως γράφονται ἀντίστοιχα:

$$V(X) = \frac{\sum\limits_i f_i\,(x_i - \overline{x})^2}{\sum\limits_i f_i} \qquad (5.7)$$

καί $$V(X) = \frac{\sum_i f_i x_i^2}{\sum_i f_i} - \overline{x}^2 \qquad (5.8)$$

Ἄν ὁ πίνακας συχνοτήτων ἀναφέρεται σέ ὁμαδοποιημένες παρατηρήσεις γιά τιμές $x_1, x_2, \ldots, x_\kappa$ τῆς μεταβλητῆς, παίρνομε τά κέντρα τῶν κλάσεων.

Παράδειγμα.

Στόν Πίνακα 5.4.1 δίνεται ἡ ταχύτητα μέ τήν ὁποία 10 αὐτοκίνητα πέρασαν ἀπό μία ἐπικίνδυνη διασταύρωση.

Ζητεῖται ἡ διακύμανση καί ἡ τυπική ἀπόκλιση τῶν ταχυτήτων.

ΠΙΝΑΚΑΣ 5.4.1.

Ταχύτητα σέ km/h	Αὐτοκίνητα f_i
0 - 10	4
10 - 20	3
20 - 30	2
30 - 40	1
Ἄθροισμα	10

Α΄ Τρόπος: Γιά νά ὑπολογίσομε τή διακύμανση μέ τόν τύπο:

$$s_x^2 = V(X) = \frac{\sum_i f_i\,(x_i - \overline{x})^2}{\sum_i f_i}$$

βρίσκομε τό μέσο ἀριθμητικό $\overline{x}$ καί σχηματίζομε τόν Πίνακα 5.4.2, πού περιέχει στήλη τῶν διαφορῶν $(x_i - \overline{x})^2$ καί στήλη τῶν γινομένων $f_i\,(x_i - \overline{x})^2$.

ΠΙΝΑΚΑΣ 5.4.2.

Τάξεις	f_i	x_i	$f_i\,x_i$	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$f_i\,(x_i - \overline{x})^2$
0 - 10	4	5	20	− 10	100	400
10 - 20	3	15	45	0	0	0
20 - 30	2	25	50	10	100	200
30 - 40	1	35	35	20	400	400
Ἄθροισμα	10		150			1000

$$\overline{x} = \frac{150}{10} = 15.$$

Ἑπομένως, ἡ διακύμανση καί ἡ τυπική ἀπόκλιση θά εἶναι ἀντίστοιχα:

$$V(X) = \frac{\sum_i f_i\,(x_i - \overline{x})^2}{\sum_i f_i} = \frac{1000}{10} = 100, \quad s_x = \sqrt{100} = 10.$$

Β' τρόπος: Γιά νά ὑπολογίσομε τή διακύμανση μέ τόν τύπο:

$$V(X) = \frac{\sum_i f_i x_i^2}{\sum_i f_i} - \bar{x}^2$$

σχηματίζομε τόν Πίνακα 5.4.3 πού περιέχει στήλη τῶν τετραγώνων x_i^2 καί στήλη τῶν γινομένων $f_i\ x_i^2$.

ΠΙΝΑΚΑΣ 5.4.3.

Τάξεις	f_i	x_i	$f_i x_i$	x_i^2	$f_i x_i^2$
0 - 10	4	5	20	25	100
10 - 20	3	15	45	225	675
20 - 30	2	25	50	625	1250
30 - 40	1	35	35	1225	1225
Ἄθροισμα	10		150		3250

$$\bar{x} = \frac{150}{10} = 15.$$

Ἑπομένως, ἡ διακύμανση καί ἡ τυπική ἀπόκλιση θά εἶναι ἀντίστοιχα:

$$V(X) = \frac{\sum_i f_i x_i^2}{\sum_i f_i} - \bar{x}^2 = \frac{3250}{10} - 15^2 = 100, \quad s_x = \sqrt{100} = 10.$$

Ἰδιότητες τῆς διακυμάνσεως. Ἀπό τόν ὁρισμό τῆς διακυμάνσεως:

$$V(X) = \frac{1}{\nu} \sum_i (x_i - \bar{x})^2,$$

προκύπτουν οἱ ἑξῆς ἰδιότητες:

1) Ἄν οἱ τιμές $x_1, x_2, ..., x_\nu$ εἶναι ἴσες, τότε ἡ διακύμανσή τους εἶναι μηδέν.

Ἀπόδειξη: Θέτοντας $x_1 = x_2 = ... = x_\nu = \alpha$, θά ἔχομε καί $\bar{x} = \alpha$, ὁπότε:

$$V(X) = \frac{1}{\nu}\left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + ... + (x_\nu - \bar{x})^2\right] = \frac{1}{\nu}\left(0 + 0 + ... + 0\right) = 0.$$

2) Ἄν ὅλες οἱ τιμές $x_1, x_2, ..., x_\nu$ αὐξηθοῦν (ἤ μειωθοῦν) κατά μία σταθερά ποσότητα α, ἡ διακύμανση δέν μεταβάλλεται, δηλαδή:

$$V(X + \alpha) = V(X) \quad \text{καί} \quad V(X - \alpha) = V(X).$$

Ἀπόδειξη: Θέτοντας $y_1 = x_i \pm \alpha$ ἔχομε καί $\bar{y} = \bar{x} \pm \alpha$, ὁπότε:

$$V(Y) = V(X \pm \alpha) = \frac{1}{\nu} \sum_i (y_i - \bar{y})^2 = \frac{1}{\nu} \sum_i \left[(x_i \pm \alpha) - (\bar{x} \pm \alpha)\right]^2 =$$

$$\frac{1}{\nu}\sum_i (x_i - \overline{x})^2 = V(X).$$

3) Ἂν ὅλες οἱ τιμές $x_1, x_2, \ldots, x_\nu$ πολλαπλασιασθοῦν μέ ἕνα σταθερό ἀριθμό λ, τότε ἡ διακύμανσή τους πολλαπλασιάζεται ἐπί λ^2, δηλαδή:

$$V(\lambda X) = \lambda^2 V(X).$$

Ἀπόδειξη: Θέτοντας $z_i = \lambda x_i$, ἔχομε καί $\overline{z} = \lambda\overline{x}$, ὁπότε:

$$V(Z) = V(\lambda X) = \frac{1}{\nu}\sum_i (z_i - \overline{z})^2 = \frac{1}{\nu}\sum_i (\lambda x_i - \lambda\overline{x})^2 = \lambda^2 \frac{1}{\nu}\sum_i (x_i - \overline{x})^2 = \lambda^2 V(X).$$

Ἔτσι π.χ., ἄν ὅλες οἱ τιμές μιᾶς μεταβλητῆς X αὐξηθοῦν κατά 8 μονάδες, ἡ διακύμανση τῆς X δέν μεταβάλλεται. Ἂν ὅμως οἱ τιμές τῆς X αὐξηθοῦν κατά 8%, ἡ διακύμανσή της πολλαπλασιάζεται ἐπί $(1{,}08)^2$, γιατί κάθε τιμή x_i γίνεται: $x_i + 0{,}08x_i = (1 + 0{,}008)x_i = (1{,}08)x_i$.

Δηλαδή θά εἶναι:

α) Ἂν οἱ τιμές τῆς μεταβλητῆς X αὐξηθοῦν κατά 8 μονάδες:

$$s_y^2 = \frac{\sum_i (y_i - \overline{y})^2}{\nu} = \frac{\sum_i [(x_i + 8) - (\overline{x} + 8)]^2}{\nu} = \frac{\sum_i (x - \overline{x})^2}{\nu} = s_x^2$$

ἑπομένως καμιά μεταβολή δέν ἐπέρχεται στή διασπορά.

β) Ἂν οἱ τιμές τῆς μεταβλητῆς X αὐξηθοῦν κατά 8%:

$$s_y^2 = \frac{\sum_i (y_i - \overline{y})^2}{\nu} = \frac{\sum_i [(x_i + 0{,}08x_i) - (\overline{x} + 0{,}08\,\overline{x})]^2}{\nu} = 1{,}08^2 \frac{\sum_i (x_i - \overline{x})^2}{\nu} =$$

$$= 1{,}08^2\, s_x^2$$

ἑπομένως ἡ διασπορά ἔχει μεταβληθεῖ.

Ἔμμεση μέθοδος ὑπολογισμοῦ διακυμάνσεως.

Ἀπό τίς ἰδιότητες τῆς διακυμάνσεως συμπεραίνομε ὅτι, γιά τόν ὑπολογισμό τῆς διακυμάνσεως, μποροῦμε νά χρησιμοποιήσομε ἔμμεση μέθοδο ἀνάλογη μέ ἐκείνη πού χρησιμοποιήσαμε στόν ὑπολογισμό τοῦ ἀριθμητικοῦ μέσου, δηλαδή:

— Ἂν οἱ τιμές $x_1, x_2, \ldots, x_\nu$ μιᾶς μεταβλητῆς X εἶναι μεγάλες, μποροῦμε νά ἀφαιρέσομε ἀπό ὅλες τόν ἴδιο ἀριθμό α (συνήθως ἀφαιροῦμε τήν τιμή τῆς x πού ἔχει τή μεγαλύτερη συχνότητα) καί νά βροῦμε τή διακύμανση τῶν τιμῶν $x_1 - \alpha$, $x_2 - \alpha, \ldots, x_\nu - \alpha$ πού εἶναι ἴση μέ τή διακύμανση τῆς X.

— Μποροῦμε ἐπίσης νά πολλαπλασιάσομε (ἤ νά διαιρέσομε) τίς τιμές $x_1 - \alpha$, $x_2 - \alpha, \ldots, x_\nu - \alpha$ μέ ἕνα ἀριθμό λ καί νά βροῦμε τή διακύμανση τῶν νέων τιμῶν, ὁπότε ἡ διακύμανση τῆς X θά βρίσκεται ἄν διαιρέσομε (ἤ πολλαπλασιάσομε) τή διακύμανση τῶν νέων τιμῶν μέ λ^2. Ἂς βροῦμε π.χ. μέ τήν ἔμμεση μέθοδο τή διακύμανση τοῦ παραδείγματος τοῦ Πίνακα 5.4.1. Ἀφαιροῦμε ἀπό ὅλες τίς τιμές x_i

(κέντρο τῶν κλάσεων) τόν ἀριθμό $a = 15$ καί διαιροῦμε τίς διαφορές $x_i - 15$ μέ τόν ἀριθμό $\lambda = 10$. Σχηματίζομε ἔτσι μία νέα μεταβλητή $\Xi = \frac{X-15}{10}$, τῆς ὁποίας ὁ μέσος ἀριθμητικός καί ἡ διακύμανση ὑπολογίζονται ἀπό τόν Πίνακα 5.4.4.

ΠΙΝΑΚΑΣ 5.4.4.

τάξεις	x_i	f_i	$x_i - 15$	$\xi_i = \frac{x_i - 15}{10}$	ξ_i^2	$\xi_i f_i$	$\xi_i^2 f_i$
0-10	5	4	−10	−1	1	−4	4
10-20	15	3	0	0	0	0	0
20-30	25	2	10	1	1	2	2
30-40	35	1	20	2	4	2	4
		10				0	10

Ἐπειδή εἶναι $\bar{\xi} = \frac{0}{10} = 0$ καί $V(\Xi) = \frac{\sum_i f_i \xi_i^2}{\sum_i f_i} - \bar{\xi}^2 = \frac{10}{10} - 0 = 1$, ἡ διακύμανση τῆς X θά εἶναι:

$$V(X) = \lambda^2 V(\Xi) = 10^2 \,.\, V(\Xi) = 10^2 \,.\, 1 = 100.$$

Δηλαδή ἡ διακύμανση μέ τήν ἔμμεση μέθοδο δίνεται ἀπό τόν τύπο:

$$V(X) = \lambda^2 \left[\frac{\sum_i f_i \xi_i^2}{\sum_i f_i} - \left(\frac{\sum_i f_i \xi_i}{\sum_i f_i} \right)^2 \right] \qquad (5.9)$$

5.5 Συντελεστής μεταβλητικότητας.

Ἡ τυπική ἀπόκλιση, ἡ ὁποία εἶναι τό κυρίως χρησιμοποιούμενο μέτρο διασπορᾶς, ἐκφράζεται, ὅπως εἴπαμε, στίς ἴδιες μονάδες μέ τίς ὁποῖες ἐκφράζονται καί οἱ τιμές τῆς μεταβλητῆς καί μᾶς δίνει τό βαθμό συγκεντρώσεως τῶν τιμῶν γύρω ἀπό τό μέσο ἀριθμητικό τους. Ἡ σημασία ὅμως τῆς τυπικῆς ἀποκλίσεως ἐξαρτᾶται καί ἀπό τήν τιμή τοῦ μέσου ἀριθμητικοῦ. Ἔτσι π.χ. μία τυπική ἀπόκλιση 3 cm ἔχει ἄλλη σημασία γιά μετρήσεις πού ἔχουν μέσο ἀριθμητικό 7 cm καί ἄλλη σημασία γιά μετρήσεις πού ἔχουν μέσο ἀριθμητικό 20 cm (οἱ δεύτερες μετρήσεις εἶναι πολύ πιό συγκεντρωμένες τιμές). Καταλαβαίνομε λοιπόν ὅτι δέν μποροῦμε μέ τίς τυπικές ἀποκλίσεις νά συγκρίνομε τή συγκέντρωση τῶν τιμῶν δύο διαφορετικῶν κατανομῶν καί θά ἔπρεπε γιά τό σκοπό αὐτό νά ὁρίσομε ***σχετικά μέτρα διασπορᾶς***, δηλαδή ἀριθμούς πού θά δίνουν τό βαθμό συγκεντρώσεως τῶν παρατηρήσεων σέ σχέση μέ τή θέση τους. Ἕνα τέτοιο μέτρο εἶναι ***ὁ συντελεστής μεταβλητικότητας.***

Συντελεστής μεταβλητικότητας μιᾶς κατανομῆς ὀνομάζεται ὁ λόγος τῆς τυπικῆς ἀποκλίσεώς της πρός τόν ἀριθμητικό μέσο της, ὁ ὁποῖος σημειώνεται μέ CV, δηλαδή:

$$CV = \frac{s_x}{\overline{x}} \quad ἤ \quad CV = \frac{s_x}{\overline{x}}\,100. \qquad (5.10$$

Εἶναι φανερό ὅτι ὁ συντελεστής μεταβλητικότητας εἶναι καθαρός ἀριθμός, ἀνεξάρτητος ἀπό τίς μονάδες μετρήσεως καί ἐκφράζεται συνήθως σέ ἑκατοστιαία ἀναλογία. Ὅσο μικρότερος εἶναι ὁ συντελεστής μεταβλητικότητας, τόσο μεγαλύτερη ὁμοιογένεια ἔχουν οἱ τιμές τῆς μεταβλητῆς.

Παράδειγμα 1ο.

Μία στρατιωτική μονάδα θέλει νά δοκιμάσει δύο πυροβόλα X καί Y διαφορετικοῦ τύπου καί ρίχνει μέ τό καθένα 5 βλήματα δίχως νά ἀλλάξει τή γωνία βολῆς του. Ἄν τά βλήματα ἔπεσαν στίς ἑξῆς ἀποστάσεις:

x: 4400 3700 3600 4600 4200
y: 8800 7900 8600 7500 8200

ποιό ἀπό τά δύο πυροβόλα παρουσιάζει μεγαλύτερη ὁμοιογένεια στίς βολές;

Λύση:

Γιά νά ὑπολογίσομε τό συντελεστή μεταβλητικότητας στήν κάθε περίπτωση, σχηματίζομε τόν Πίνακα 5.5.1.

ΠΙΝΑΚΑΣ 5.5.1.

x	y_i	$x-\overline{x}$	$(x-\overline{x})^2$	$y-\overline{y}$	$(y-\overline{y})^2$
4400	8800	300	90000	600	360000
3700	7900	− 400	160000	− 300	90000
3600	8600	− 500	250000	400	160000
4600	7500	500	250000	− 700	490000
4200	8200	100	10000	0	0
20500	41000		760000		1100000

Ἀπό τόν πίνακα αὐτό, παίρνομε:

Γιά τή X: $\overline{x} = \dfrac{20500}{5} = 4100, \quad V(X) = \dfrac{760000}{5} = 152000,$

$s_x = \sqrt{152000} \simeq 390.$

Γιά τήν Y: $\overline{y} = \dfrac{41000}{5} = 8200, \quad V(Y) = \dfrac{1100000}{5} = 220000,$

$s_y = \sqrt{220000} = 469.$

Ἐπομένως, οἱ συντελεστές μεταβλητικότητας θά εἶναι:

Γιά τή X: $CV = \dfrac{390}{4100} = 0{,}095 = 9{,}5\%$

Γιά τήν Υ: $CV = \frac{469}{8200} = 0{,}057 = 5{,}7\%$

καί κατά συνέπεια τό πυροβόλο Υ ἔχει μεγαλύτερη ὁμοιογένεια στίς βολές.

Παράδειγμα 2ο.

'Εξετάσαμε μία ὁμάδα ἀτόμων ὡς πρός τό ὕψος καί τό βάρος τους καί βρήκαμε:

— Γιά τό βάρος: $\overline{x}_1 = 72$ κιλά, $s_1 = 4$ κιλά

— Γιά τό ὕψος: $\overline{x}_2 = 168$ ἑκατ., $s_2 = 7$ ἑκατ.

Νά ἐξετασθεῖ ποιά ἀπό τίς δυό κατανομές τοῦ βάρους καί τοῦ ὕψους παρουσιάζει μεγαλύτερη διασπορά τιμῶν.

Λύση:

'Επειδή οἱ τιμές τοῦ βάρους καί τοῦ ὕψους ἐκφράζονται σέ διαφορετικές μονάδες, ἡ σύγκριση τῆς διασπορᾶς τῶν τιμῶν γίνεται μόνο μέ τούς συντελεστές μεταβλητικότητας. Αὐτοί ὅμως εἶναι:

— Γιά τό βάρος: $CV = \frac{4}{72} = 0{,}055 = 5{,}5\%$,

— Γιά τό ὕψος: $CV = \frac{7}{168} = 0{,}0417 = 4{,}17\%$.

Συνεπῶς, ἡ κατανομή τοῦ βάρους παρουσιάζει μεγαλύτερη διασπορά τιμῶν.

5.6 'Ασκήσεις.

1. 'Επτά μαθητές ἐξετάσθηκαν προφορικά ἀπό τόν καθηγητή τῆς Στατιστικῆς καί πῆραν τούς πιό κάτω βαθμούς:

   8, 11, 13, 14, 16, 12, 17

   Νά ὑπολογισθεῖ τό εὖρος μεταβολῆς καί ἡ μέση ἀπόκλιση.

2. Νά ὑπολογισθεῖ ἡ μέση ἀπόκλιση τῶν πιό κάτω παρατηρήσεων τῆς μεταβλητῆς χ:
   α) 3, 4, 12.
   β) 2, 4, 6, 5.

3. Νά ὑπολογισθεῖ ἡ μέση ἀπόκλιση τῆς παρακάτω κατανομῆς:

x_i	2	3	6	7	9
f_i	1	3	4	2	2

4. Νά ὑπολογισθεῖ ἡ διακύμανση καί ἡ τυπική ἀπόκλιση τῶν παρακάτω παρατηρήσεων τῆς μεταβλητῆς Χ:
   α) 5, 7, 12, 15, 20.
   β) 2, 4, 6, 5.

5. Νά ὑπολογισθεῖ ἡ διακύμανση καί ἡ τυπική ἀπόκλιση τῆς παρακάτω κατανομῆς:

x_i	10	12	18	23
f_i	2	5	6	4

6. Νά ὑπολογισθεῖ ἡ τυπική ἀπόκλιση καί ὁ συντελεστής μεταβλητικότητας τῶν παρακάτω παρατηρήσεων τῆς μεταβλητῆς Χ:

   4, 6, 10, 24, 32.

7. Ὁ παρακάτω πίνακας παρουσιάζει τόν ἀριθμό δωματίων πού ἔχουν οἱ 30 μαθητές μιᾶς τάξεως:

Ἀριθμός δωματίων (x_i)	Μαθητές (f_i)
2	5
3	14
4	7
5	4
Ἄθροισμα	30

Νά ὑπολογισθεῖ ἡ διακύμανση καί ἡ τυπική ἀπόκλιση.

8. Ὁ παρακάτω πίνακας μᾶς δίνει τίς ἡλικίες τῶν ὑπαλλήλων μιᾶς ἐπιχειρήσεως:

Τάξεις ἡλικιῶν	Ὑπάλληλοι f_i
25 - 35	10
35 - 45	16
45 - 55	8
55 - 65	6
Ἄθροισμα	40

Νά ὑπολογισθεῖ ἡ διακύμανση καί ἡ τυπική ἀπόκλιση.

9. Ὁ παρακάτω πίνακας μᾶς δίνει τή βαθμολογία τῶν μαθητῶν μιᾶς τάξεως ἑνός Λυκείου σέ ἕνα πρόχειρο διαγώνισμα στό μάθημα τῆς Στατιστικῆς:

Βαθμός (x_i)	Μαθητές (f_i)
9	3
10	5
11	8
12	10
14	4
15	5
16	3
17	2
Ἄθροισμα	40

Νά ὑπολογισθεῖ ὁ συντελεστής μεταβλητικότητας.

10. Ὁ παρακάτω πίνακας παρουσιάζει τά ὀρθογραφικά σφάλματα 50 μαθητῶν τῆς τελευταίας τάξεως ἑνός Λυκείου, κατά τήν ὑπαγόρευση ἑνός ἁπλοῦ κειμένου. Τά στοιχεῖα πού λείπουν ἔχουν χαθεῖ:

Λάθη (x_i)	f_i
1	—
2	—
3	5
Ἄθροισμα	50

Ἄν ὁ μέσος ἀριθμητικός τῆς παραπάνω κατανομῆς εἶναι $\overline{x} = 1,4$, νά ὑπολογισθεῖ ἡ διακύμανση.

11. Ὁ παρακάτω πίνακας μᾶς δίνει τήν κατανομή 100 ἡμερῶν ὡς πρός τά αὐτοκινητικά δυστυχήματα πού ἔχουν συμβεῖ σέ μιά πόλη:

Δυστυχήματα x_i	μέρες (f_i)
0	42
1	—
2	14
3	6
4	—
Ἄθροισμα	100

Κατά τήν κατάταξη τῶν παραπάνω δυστυχημάτων χάθηκαν οἱ συχνότητες πού ἀντιστοιχοῦν στήν 2η καί 5η τάξη.
Ἄν ὁ μέσος ἀριθμός δυστυχημάτων εἶναι $\overline{x} = 0,90$, ζητεῖται:
α) Ἡ διάμεσος τιμή καί τό τρίτο τεταρτημόριο.
β) Ἡ διακύμανση καί ἡ τυπική ἀπόκλιση.

12. Δίνεται παρακάτω ἡ κατανομή τῆς ἡλικίας 100 ὑπαλλήλων μιᾶς ἐπιχειρήσεως:

Τάξεις	f_i
25 - 30	2
30 - 35	10
35 - 40	25
40 - 45	35
45 - 50	19
50 - 55	6
55 - 60	3
Ἄθροισμα	100

Ζητεῖται: α) Ἡ μέση ἡλικία. β) Ἡ διάμεσος ἡλικία. γ) Ἡ διακύμανση τῆς ἡλικίας καί ἡ τυπική ἀπόκλιση. δ) Ὁ συντελεστής μεταβλητικότητας.

13. Ὁ παρακάτω πίνακας μᾶς δίνει σέ χιλιάδες δρχ. τά καθαρά ἡμερήσια κέρδη 100 ἐπιχειρήσεων:

Τάξεις	0 - 2	2 - 6	6 - 10	10 - 20	20 - 40	40 - 100	100 καί ἄνω
f_i	10	20	40	20	7	2	1

Ἄν γνωρίζομε ὅτι ὁ ἀριθμητικός μέσος ἰσοῦται μέ τό τρίτο τεταρτημόριο, νά ὑπολογισθεῖ ὁ συντελεστής μεταβλητικότητας.

14. Δίνεται ἡ κατανομή 40 ὑπαλλήλων ὡς πρός τό μηνιαῖο μισθό τους:

Μισθός (σέ χιλιάδες δρχ.)	f_i
2 - 4	6
4 - 6	14
6 - 8	9
8 - 10	8
10 - 12	3
Ἄθροισμα	40

Ζητεῖται: α) Νά ὑπολογισθεῖ ἡ διακύμανση καί ἡ τυπική ἀπόκλιση. β) Νά συγκριθεῖ ἡ διασπορά τιμῶν τῆς παραπάνω κατανομῆς πρός τή διασπορά τιμῶν μιᾶς ἄλλης κατανομῆς ὑπαλλήλων, γιά τήν ὁποία δίνεται ὅτι ἡ διακύμανση εἶναι 2500 καί ἡ μέση ἀριθμητική τιμή εἶναι 8800 δρχ.

15. Δίνεται παρακάτω ἡ κατανομή 28 ἐμπορικῶν καταστημάτων ὡς πρός τίς ἡμερήσιες εἰσπράξεις τους (σέ χιλιάδες δρχ.):

Τάξεις	5 - 15	15 - 25	25 - 35	35 - 45	45 - 55
f_i	1	5	5	10	7

Νά ὑπολογισθεῖ: α) Ὁ συντελεστής μεταβλητικότητας. β) Τό τρίτο τεταρτημόριο.

16. Ἡ κατανομή τῶν ὑπαλλήλων μιᾶς ἐπιχειρήσεως ὡς πρός τίς μηνιαῖες ἀποδοχές τους ἔχει μέσο ἀριθμητικό $\overline{x} = 14.000$ δρχ. καί διακύμανση $V(X) = 160.000$. Ἄν οἱ μηνιαῖες ἀποδοχές κάθε ὑπαλλήλου αὐξηθοῦν κατά 20%, ποιά θά εἶναι ἡ τιμή τῆς νέας διακυμάνσεως;

17. Δίνεται παρακάτω ἡ κατανομή 20 οἰκογενειῶν ὡς πρός τίς μηνιαῖες δαπάνες τους (σέ χιλιάδες δρχ.):

Τάξεις	8 - 10	10 - 12	12 - 14	14 - 16	16 - 18
f_i	2	5	8	3	2

Νά ὑπολογισθεῖ ὁ συντελεστής μεταβλητικότητας. (Νά χρησιμοποιηθεῖ ἡ ἔμμεση μέθοδος ὑπολογισμοῦ τῆς διακυμάνσεως καί τοῦ μέσου ἀριθμητικοῦ).

18. Δίνεται παρακάτω ἡ βαθμολογία ἑνός μαθητῆ στά διάφορα μαθήματα:

5, 6, 7, 8, 6, 4, 9.

Νά ὑπολογισθεῖ ὁ συντελεστής μεταβλητικότητας.

19. Στήν παρακάτω κατανομή νά ὑπολογισθεῖ ἡ διακύμανση καί ἡ τυπική ἀπόκλιση μέ τρεῖς διαφορετικούς τρόπους:

Τάξεις	Συχνότητες f_i
30 - 40	5
40 - 50	4
50 - 60	10
60 - 70	1
Ἄθροισμα	20

20. Ἡ βαθμολογία τῶν μαθητῶν μιᾶς τάξεως στό μάθημα τῆς Ἱστορίας εἶναι 12, 14, 11, 8, 16, 18, 11, 13, 15, 12, ἐνῶ ἡ βαθμολογία τῶν μαθητριῶν τῆς ἴδιας τάξεως στό ἴδιο μάθημα εἶναι 16, 15, 14, 17, 15, 18, 17, 18, 16, 17. Οἱ μαθητές ἤ οἱ μαθήτριες παρουσιάζουν μεγαλύτερη ὁμοιογένεια βαθμολογίας;

21. Στόν παρακάτω πίνακα δίνεται ἡ βαθμολογία στά Μαθηματικά 72 μαθητῶν σέ ἕνα πανελλήνιο διαγώνισμα:

Τάξεις	f_i
2 - 4	8
4 - 6	14
6 - 8	15
8 - 10	14
10 - 12	10
12 - 14	8
14 - 16	3
Ἄθροισμα	72

Ζητεῖται: α) Ὁ μέσος ἀριθμητικός καί τό πρῶτο τεταρτημόριο. β) Ἡ τυπική ἀπόκλιση.

22. Ἡ μέση ἡμερήσια θερμοκρασία σέ 20 ὀρεινά χωριά τῆς αὐτῆς ἐπαρχίας τήν 15η Ἰανουαρίου τοῦ 1978 ἦταν:

1 0 1 2 4 1 0 2 0 3
0 1 2 3 1 1 1 2 2 4

Νά σχηματισθεῖ ὁ πίνακας κατανομῆς συχνοτήτων καί στή συνέχεια νά ὑπολογισθεῖ ὁ συντελεστής μεταβλητικότητας.

23. Δίνεται παρακάτω ἡ κατανομή 150 οἰκογενειῶν ὡς πρός τό καταβαλλόμενο ἐνοίκιο γιά κατοικία:

Μηνιαῖο ἐνοίκιο (σέ χιλιάδες) δρχ.	Οἰκογένειες f_i
2-3	25
3-4	52
4-5	50
5-6	11
6-7	7
7-8	5
Ἄθροισμα	150

Ζητεῖται: α) Νά συμπληρωθεῖ ὁ πίνακας μέ στῆλες σχετικῆς συχνότητας καί ἀθροιστικῆς σχετικῆς συχνότητας. β) Νά ὑπολογισθεῖ ἡ τυπική ἀπόκλιση.

ΚΕΦΑΛΑΙΟ ΕΚΤΟ

ΠΑΛΙΝΔΡΟΜΗΣΗ ΚΑΙ ΣΥΣΧΕΤΙΣΗ ΔΥΟ ΜΕΤΑΒΛΗΤΩΝ

Α. ΑΛΛΗΛΟΕΞΑΡΤΗΣΗ ΔΥΟ ΜΕΤΑΒΛΗΤΩΝ

6.1 Γενικά.

Μέχρι τώρα ἀσχοληθήκαμε μέ τή μελέτη μιᾶς μόνο μεταβλητῆς, δηλαδή ἐξετάζομε τά ἄτομα ἑνός πληθυσμοῦ ὡς πρός μία μόνο μεταβλητή ἰδιότητά τους. Σέ πολλές περιπτώσεις ὅμως, ἀσχολούμαστε συγχρόνως μέ τή μελέτη δύο μεταβλητῶν, μέ σκοπό νά ἐξακριβώσομε ἄν ὑπάρχει ἀλληλοεξάρτηση μεταξύ τους, δηλαδή ἄν οἱ τιμές τῆς μιᾶς ἐπηρεάζονται ἀπό τίς τιμές τῆς ἄλλης καί νά προσδιορίσομε τόν τρόπο ἀλληλοεξαρτήσεώς τους. Τέτοιες περιπτώσεις ἔχομε π.χ. ὅταν θέλομε νά ἐξετάσομε ἄν ὑπάρχει σχέση μεταξύ ἐγκληματικότητας καί ἀνεργίας, ἤ μεταξύ βάρους καί ὕψους μιᾶς ὁμάδας ἀτόμων, ἤ προσφορᾶς καί ζητήσεως τῶν ἀγαθῶν κλπ.

Σέ ὅλα λοιπόν τά ἑπόμενα θά θεωροῦμε ἕνα πληθυσμό μέ ν ἄτομα καί θά ἐξετάζομε καθένα ἀπό τά ἄτομα αὐτά ὡς πρός δύο μεταβλητές ἰδιότητες, τίς ὁποῖες θά σημειώνομε μέ X καί Y. Ἔτσι, οἱ παρατηρήσεις μας θά εἶναι τώρα ν ὁρισμένα ζεύγη τιμῶν:

$$(x_1, y_1), (x_2, y_2), \dots, (x_\nu, y_\nu),$$

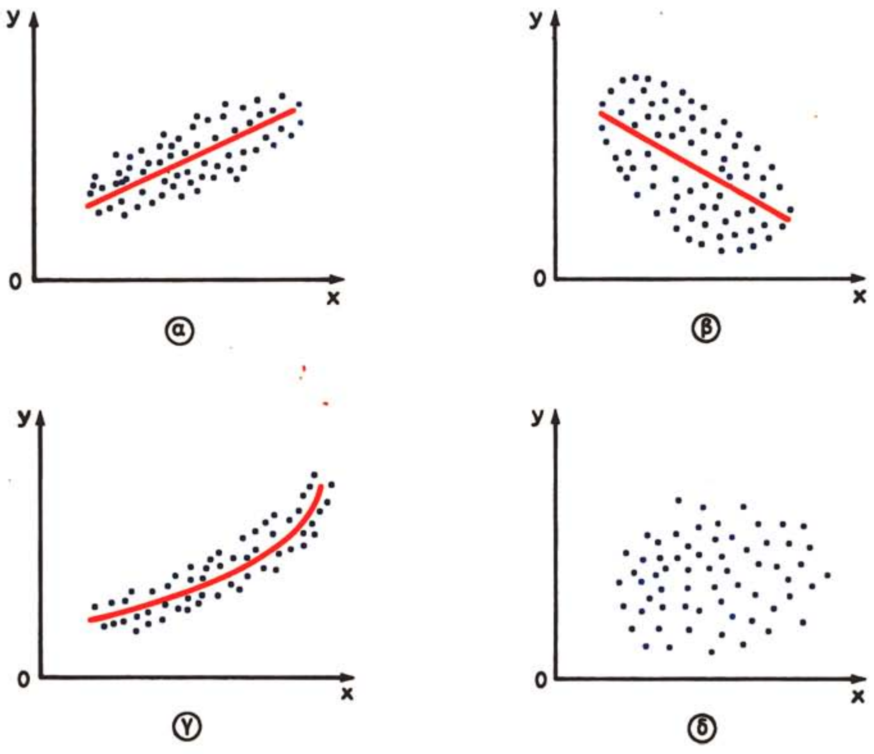
τά ὁποῖα δέν εἶναι ἀπαραιτήτως διαφορετικά μεταξύ τους.

Ἄν πάρομε ἕνα σύστημα ὀρθογωνίων ἀξόνων τοῦ ἐπιπέδου καί σημειώσομε πάνω σ' αὐτό τά σημεῖα $M_1, M_2, \dots, M_\nu$, τά ὁποῖα ἔχουν συντεταγμένες τά ζεύγη πού παριστάνουν τίς παρατηρήσεις μας, σχηματίζεται ἕνα πλῆθος σημείων πού λέγεται ***νέφος σημείων*** ἤ καί ***διάγραμμα διασπορᾶς***.

Μιά πρώτη ἔνδειξη ὅτι ὑπάρχει ἀλληλοεξάρτηση, εἶναι ὅταν τό νέφος τῶν σημείων ἀκολουθεῖ μία νοητή γραμμή τοῦ ἐπιπέδου, ὅπως φαίνεται ὅτι συμβαίνει στό σχῆμα 6.1 (α, β, καί γ). Ἀντίθετα, κάτι τέτοιο δέν συμβαίνει στό (δ) τοῦ ἴδιου σχήματος, αὐτό σημαίνει ὅτι οἱ μεταβλητές δέν ἔχουν ἀλληλοεξάρτηση ἤ ὅτι εἶναι, ὅπως λέμε, ***ἀνεξάρτητες***.

Θά δοῦμε τώρα τούς διάφορους τρόπους ἀλληλοεξαρτήσεως (ἤ ὅπως ἀλλιῶς λέμε, ***συμμεταβολῆς***) δύο μεταβλητῶν. Αὐτοί εἶναι:

α) ἡ ***συναρτησιακή*** καί β) ἡ ***στοχαστική*** ἤ ***στατιστική ἐξάρτηση***.



Σχ. 6.1.

6.2 Συναρτησιακή ἐξάρτηση.

Θά λέμε ὅτι δύο μεταβλητές X καί Y ἔχουν συναρτησιακή ἐξάρτηση, ὅταν σέ κάθε τιμή x_i τῆς μεταβλητῆς X ἀντιστοιχεῖ μία καί μόνη τιμή y_i τῆς μεταβλητῆς Y.

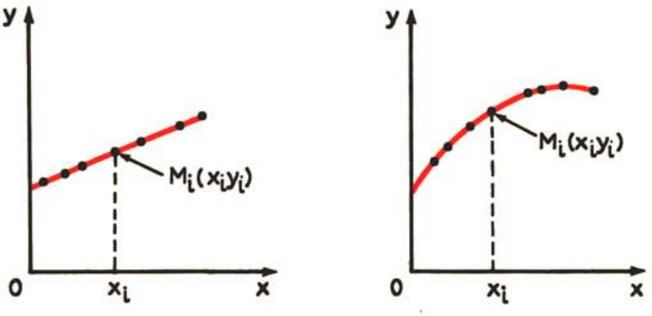
Στήν περίπτωση αὐτή ὑπάρχει, ὅπως γνωρίζομε μιά συνάρτηση μέ τύπο:

$$y = f(x),$$

ὁ ὁποῖος ἐπαληθεύεται ἀπό ὅλα τά ζεύγη τιμῶν (x_i, y_i) πού ἐμφανίζονται. Ἡ σχέση αὐτή ἐπιτρέπει τόν ὑπολογισμό τῶν τιμῶν τῆς Y ἀπό τίς ἀντίστοιχες τιμές τῆς X μέ ἀπόλυτη ἀκρίβεια. Ἑπομένως, ὅλα τά σημεῖα M_i $(x_i\ y_i)$ τοῦ διαγράμματος βρίσκονται ἐπάνω σέ μία καμπύλη πού ἔχει ἐξίσωση $y = f(x)$ (σχ. 6.2).

Σέ μία τέτοια συναρτησιακή ἐξάρτηση, ἡ X χαρακτηρίζεται ὡς ***ἀνεξάρτητη*** μεταβλητή καί ἡ Y ὡς ***ἐξηρτημένη***.

Σχέσεις τῆς μορφῆς αὐτῆς συναντοῦμε στίς θετικές, κυρίως, ἐπιστῆμες. Ἔτσι π.χ. ὑπάρχει συναρτησιακή ἐξάρτηση μεταξύ ἐμβαδοῦ E καί τῆς ἀκτίνας r ἑνός κυκλικοῦ δίσκου, $(E = \pi \cdot r^2)$, μεταξύ κεφαλαίου καί τόκου, ἐφόσον ὁ χρόνος καί τό ἐπιτόκιο παραμένουν σταθερά $(I = K.\ i.\ n)$, μεταξύ ἀριθμοῦ τῶν τηλεφωνημάτων πού πραγματοποιοῦνται καί τοῦ ποσοῦ πού πρέπει νά πληρωθεῖ στόν ΟΤΕ $(y = 75 + 1,20x)$, κ.τ.λ.



Σχ. 6.2.

Ἡ συναρτησιακή ἐξάρτηση εἶναι γενικά ἀντικείμενο τῶν μαθηματικῶν καί γι' αὐτό δέν θά ἀσχοληθοῦμε μέ αὐτήν.

6.3 Στοχαστική ἤ στατιστική ἐξάρτηση.

Θά λέμε ὅτι δύο μεταβλητές X καί Y ἔχουν στοχαστική ἐξάρτηση, ὅταν σέ κάθε τιμή τῆς μεταβλητῆς X δέν ἀντιστοιχίζεται μία ὁρισμένη τιμή τῆς μεταβλητῆς Y, ἀλλά μιά τιμή τῆς Y, ἀπό ἕνα πλῆθος δυνατῶν τιμῶν της, τήν ὁποία δέν μποροῦμε νά προβλέψομε μέ ἀκρίβεια, ὅπως π.χ. ἀπό τό εἰσόδημα μιᾶς οἰκογένειας δέν μποροῦμε νά προβλέψομε μέ ἀκρίβεια τίς δαπάνες διατροφῆς της ἤ ἀπό τόν ἀριθμό τῶν μελῶν της νά προβλέψομε μέ ἀκρίβεια τόν ἀριθμό τῶν δωματίων τῆς κατοικίας της, ἤ ἀπό τό λίπασμα πού χρησιμοποιοῦμε σέ ἕνα ἀγροτεμάχιο δέν μποροῦμε νά προβλέψομε μέ ἀκρίβεια τό ὕψος τῆς παραγωγῆς κ.τ.λ.

Γιά νά μελετήσομε μία στοχαστική ἐξάρτηση, ἀφοῦ συγκεντρώσομε τίς παρατηρήσεις μας:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_\nu, y_\nu)$$

καί σχηματίζομε τό «νέφος» τῶν σημείων τους, ἐργαζόμαστε σέ γενικές γραμμές ὡς ἑξῆς:

α) Παίρνομε ὅλα τά σημεῖα τοῦ νέφους πού ἔχουν τήν ἴδια τετμημένη: π.χ. τήν $x = a_j$, καί βρίσκομε τόν ἀριθμητικό μέσο $\overline{y}_j$ τῶν τεταγμένων τους.

Ὁ ἀριθμός $\overline{y}_j$ λέγεται ***δεσμευμένος ἀριθμητικός μέσος (ἤ δεσμευμένη μέση τιμή)*** τῆς μεταβλητῆς Y γιά $x = a_j$. Εἶναι φανερό ὅτι μέ δύο διαφορετικές τιμές τοῦ a_j (δηλαδή μέ δύο διαφορετικές τιμές τῆς μεταβλητῆς X), θά ἔχομε διαφορετικούς δεσμευμένους ἀριθμητικούς μέσους τῆς Y.

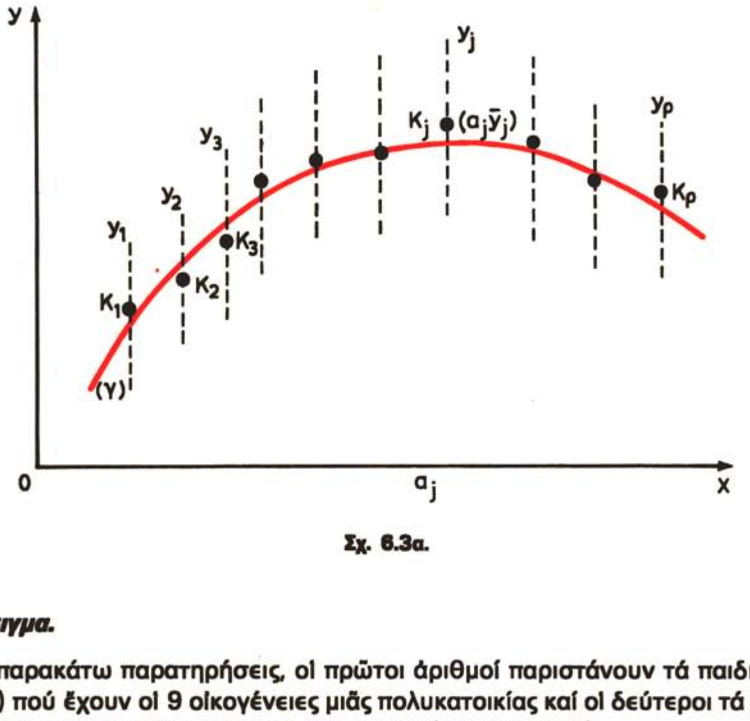
β) Σημειώνομε στό σύστημα τῶν ἀξόνων μας τά σημεῖα K_j $(a_j, \overline{y}_j)$, πού ἔχουν τετμημένες ὅλες τίς διαφορετικές τιμές τῆς X καί τεταγμένες τίς ἀντίστοιχες δεσμευμένες μέσες τιμές τῆς Y.

γ) Φέρνομε μία καμπύλη (γ) πού νά διέρχεται ἀπό τά σημεῖα $K_1, K_2, \ldots, K_j, \ldots, K_\rho$ (ἤ πού νά διέρχεται πολύ κοντά σ' αὐτά), τήν ὁποία ὀνομάζομε ***γραμμή παλινδρομήσεως τῆς y πάνω στή x.***

Ἄν ὑποθέσομε ὅτι ἡ ἐξίσωση τῆς καμπύλης γ εἶναι:

$$y = \phi\,(X),$$

τότε ἡ τιμή τῆς y πού προκύπτει ἀπό αὐτήν, γιά τήν ὁρισμένη τιμή $x = a_j$ παριστάνει ἕνα ἀριθμό (τόν ***δεσμευμένο ἀριθμητικό μέσο τῆς y γιά $x = a_j$***), γύρω ἀπό τόν ὁποῖο θά βρίσκεται ἡ τιμή τῆς Y, ὅταν ἡ μεταβλητή X πάρει τήν τιμή $x = a_j$. Ἔτσι, ἡ γραμμή παλινδρομήσεως μᾶς δίνει μία εἰκόνα τοῦ τρόπου μεταβολῆς τῆς Y, ὅταν μεταβάλλονται οἱ τιμές τῆς X. Δηλαδή μᾶς δίνει μία εἰκόνα τοῦ τρόπου ἀλληλοεξαρτήσεως τῶν δύο μεταβλητῶν (σχ. 6.3α)



Σχ. 6.3α.

Παράδειγμα.

Στίς παρακάτω παρατηρήσεις, οἱ πρῶτοι ἀριθμοί παριστάνουν τά παιδιά (μεταβλητή X) πού ἔχουν οἱ 9 οἰκογένειες μιᾶς πολυκατοικίας καί οἱ δεύτεροι τά χρήματα πού ξοδεύουν τήν ἑβδομάδα γιά θέρμανση (μεταβλητή Y):

(0,3100), (2,1100), (2,1000), (1,1600), (3,720), (0,2700), (1,1800), (1,1400), (2,1050)

Ὑπάρχει ἀλληλοεξάρτηση τῶν X καί Y;

Λύση:

Παρατηροῦμε πρῶτα ὅτι οἱ δεσμευμένοι μέσοι εἶναι:

Γιά $x = 0 \quad \bar{y}_1 = \dfrac{3100 + 2700}{2} = 2900$ δρχ. $= 2{,}90$ χιλ. δρχ.

Γιά $x = 1 \quad \bar{y}_2 = \dfrac{1600 + 1800 + 1400}{3} = 1600$ δρχ. $= 1{,}60$ χιλ. δρχ.

$$\text{Γιά } x = 2 \quad \overline{y}_3 = \frac{1100 + 1000 + 1050}{3} = 1050 \text{ δρχ.} = 1{,}05 \text{ χιλ. δρχ.}$$

Θά πρέπει τώρα νά βροῦμε μία καμπύλη πού διέρχεται πολύ κοντά ἀπό τά 4 σημεῖα:

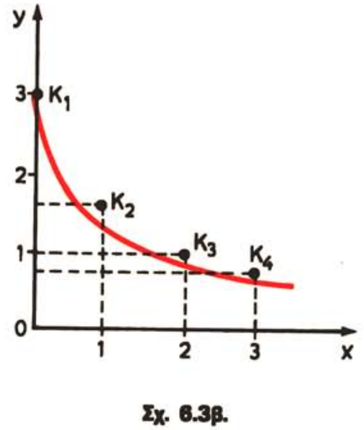
K_1 (0, 2, 90), K_2 (1, 1, 60), K_3 (2, 1,05), K_4 (3, 0, 72).

Μία τέτοια καμπύλη εἶναι π.χ. αὐτή πού ἔχει ἐξίσωση:

$$y = \frac{3}{x+1} \tag{6.1}$$

ἡ ὁποία δίνει: Γιά $x = 0$, $y = 3$, γιά $x = 1$, $y = 1{,}5$ γιά $x = 2$, $y = 1$ καί γιά $x = 3$, $y = 0{,}75$. Μποροῦμε λοιπόν νά ποῦμε ὅτι ὑπάρχει ἀλληλοεξάρτηση τῶν X καί Y καί ἡ γραμμή παλινδρομήσεως τῆς y πάνω στή x ἔχει ἐξίσωση τήν (6.1).

Ἔχει μεγάλη πρακτική ἀξία τό νά μπορέσομε νά περιγράψομε τήν ἀλληλοεξάρτηση τῶν δύο μεταβλητῶν μας μέ μιά καμπύλη $y = \phi(x)$, ἡ ὁποία νά ἔχει ὅσο τόν δυνατό ἁπλούστερη μορφή. Γι' αὐτό, ἀποφεύγομε κατά κανόνα τίς ἐξισώσεις μέ πολύπλοκη μορφή, ἔστω καί ἄν αὐτό εἶναι εἰς βάρος τῆς ἀκρίβειας (σχ. 6.3β).



Σχ. 6.3β.

6.4 Ἡ μέθοδος τῶν ἐλαχίστων τετραγώνων.

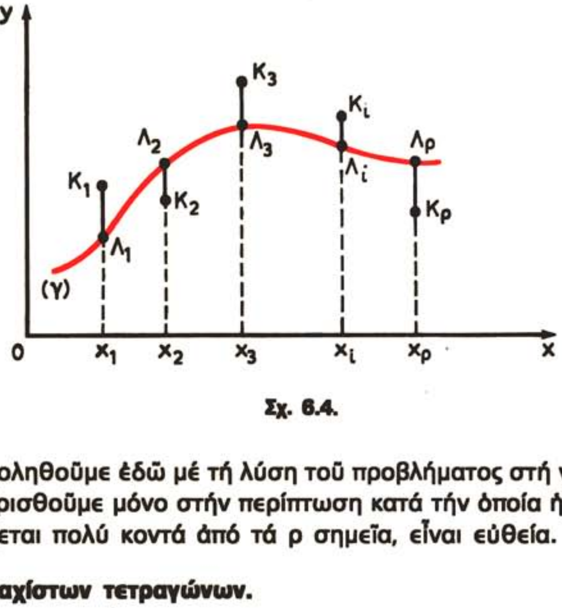
Τό πρόβλημα τῆς στοχαστικῆς ἐξαρτήσεως ἐντοπίζεται, ὅπως εἴδαμε, στήν εὕρεση μιᾶς καμπύλης, ἡ ὁποία νά διέρχεται πολύ κοντά ἀπό ὁρισμένα σημεῖα, δηλαδή ἀνάγεται στό ἑξῆς γενικότερο μαθηματικό πρόβλημα:

Νά βρεθεῖ ἡ ἐξίσωση $y = \phi(x)$ μιᾶς καμπύλης ἡ ὁποία νά διέρχεται «πολύ κοντά» ἀπό ρ ὁρισμένα σημεῖα K_1 (x_1, y_1), K_2 (x_2, y_2), . . ., K_ρ (x_ρ, y_ρ). Γιά νά ἔχει ὅμως νόημα αὐτή ἡ διατύπωση καί κυρίως ἡ φράση «πολύ κοντά», θά πρέπει νά βρίσκομε κάποιο μέτρο τό ὁποῖο νά ἐκφράζει τήν ἀπόσταση τῶν ρ σημείων ἀπό ὁποιαδήποτε καμπύλη τοῦ ἐπιπέδου. Ἄν λοιπόν ὀνομάσομε Λ_1, Λ_2, ..., Λ_ρ τά σημεῖα μιᾶς ὁποιασδήποτε καμπύλης (γ) πού ἔχουν τίς ἴδιες τετμημένες x_1, x_2, . . ., x_ρ ἀντίστοιχα, σάν τέτοιο μέτρο παίρνομε τό ἄθροισμα:

$$A = (K_1\,\Lambda_1)^2 + (K_2\,\Lambda_2)^2 +, \ldots, + (K_\rho\,\Lambda_\rho)^2$$

Ἔτσι, μία καμπύλη θά θεωρεῖται τόσο «πιό κοντά» στά σημεῖα K_1, K_2, ..., K_ρ, ὅσο τό ἄθροισμα αὐτό Α εἶναι μικρότερο, δηλαδή $(K_1\,\Lambda_1)^2 + (K_2\,\Lambda_2)^2 + \ldots + (K_\rho\,\Lambda_\rho)^2 =$ ***ἐλάχιστο.*** Παρατηροῦμε ἀκόμη ὅτι ὑπάρχουν γενικά πολλές καμπύλες μέ διαφορετικά σχήματα πού διέρχονται κοντά ἀπό τά ρ αὐτά σημεῖα (μπορεῖ π.χ. νά διέρχονται κοντά ἀπό τά ρ σημεῖα καί μία εὐθεία καί μία παραβολή). Γι αὐτό ἀκριβῶς προσδιορίζομε ἀπό τήν ἀρχή, ἀνάλογα μέ τή θέση πού ἔχουν τά ρ σημεῖα, τό εἶδος τῆς καμπύλης πού θά τοποθετήσομε ἀνάμεσά τους (π.χ. ἄν θά πάρομε εὐθεία ἤ παραβολή, κ.ο.κ.). Αὐτό σημαίνει ὅτι παίρνομε αὐθαίρετα τή μορφή τῆς ἐξισώσεως $y = \phi\,(X)$ καί μετά προσδιορίζομε τά διάφορα σημεῖα της (συντελεστές, σταθεροί ὅροι, κ.λ.π.) μέ τή βοήθεια τῶν συντεταγμένων τῶν δεδομένων σημείων.

Μέ τίς προϋποθέσεις αὐτές, τό παραπάνω γενικό πρόβλημα λύνεται μέ μία μέθοδο πού λέγεται ***μέθοδος ἐλαχίστων τετραγώνων*** καί ἡ καμπύλη πού βρίσκομε ***καμπύλη ἐλαχίστων τετραγώνων*** (σχ. 6.4).



Σχ. 6.4.

Δέν θά ἀσχοληθοῦμε ἐδῶ μέ τή λύση τοῦ προβλήματος στή γενική του μορφή, ἀλλά θά περιορισθοῦμε μόνο στήν περίπτωση κατά τήν ὁποία ἡ καμπύλη πού θέλομε νά διέρχεται πολύ κοντά ἀπό τά ρ σημεῖα, εἶναι εὐθεία.

6.5 Εὐθεία ἐλαχίστων τετραγώνων.

Ἄς ὑποθέσομε τώρα ὅτι ἡ θέση πού ἔχουν τά ρ σημεῖα $K_1\,(x_1, y_1)$, $K_2\,(x_2, y_2)$,..., $K_\rho\,(x_\rho, y_\rho)$ μᾶς ἐπιτρέπει νά ζητήσομε μία εὐθεία, ἡ ὁποία νά διέρχεται πολύ κοντά ἀπό τά σημεῖα αὐτά.

Στήν περίπτωση αὐτή, θεωροῦμε μία ἐξίσωση πρώτου βαθμοῦ ὡς πρός Χ καί Υ:

$$y = \alpha + \beta x \qquad (6.2)$$

ἡ ὁποία ἀντιπροσωπεύει ὅλες τίς εὐθεῖες τοῦ ἐπιπέδου καί προσπαθοῦμε νά προσδιορίσομε (ἀπό τίς συντεταγμένες τῶν σημείων) τούς συντελεστές της α καί β, ὥστε ἡ (6.2) νά γίνει ἡ ἐξίσωση τῆς εὐθείας πού διέρχεται ὅσο τό δυνατό πιό κοντά ἀπό τά ρ σημεῖα.

Μέ τή μέθοδο τῶν ἐλαχίστων τετραγώνων ἀποδεικνύεται ὅτι, γιά νά συμβαίνει αὐτό, θά πρέπει τά α καί β νά προκύπτουν ἀπό τή λύση τοῦ συστήματος:

$$
\begin{aligned}
\nu\alpha + \left(\sum_i x_i\right)\beta &= \left(\sum_i y_i\right) \\
\left(\sum_i x_i\right)\alpha + \left(\sum_i x_i^2\right)\beta &= \left(\sum_i x_i y_i\right)
\end{aligned}
\qquad (6.3)
$$

Τό σύστημα (6.3) λέγεται ***σύστημα κανονικῶν ἐξισώσεων.***

Γιά τόν προσδιορισμό τῶν παραμέτρων α καί β τοῦ παραπάνω συστήματος, χρησιμοποιοῦμε συνήθως τίς παρακάτω ἰσότητες:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \text{ ὅπου: } \bar{x} = \frac{\sum_i x_i}{\nu}, \ \bar{y} = \frac{\sum_i y_i}{\nu},$$

ὅπου: ν = τό πλῆθος τῶν ζευγῶν τῶν παρατηρήσεών μας.

$$\hat{\beta} = \frac{\nu \cdot \sum_i x_i y_i - \left(\sum_i x_i\right) \cdot \left(\sum_i y_i\right)}{\nu \cdot \sum_i x_i^2 - \left(\sum_i x_i\right)^2},$$

Ἑπομένως, ἡ εὐθεία πού ἔχει ἐξίσωση

$$\hat{y} = \hat{\alpha} + \hat{\beta}x \qquad (6.4)$$

θά εἶναι ἡ ***εὐθεία ἐλαχίστων τετραγώνων*** τῶν ρ σημείων $(x_i\ y_i)$, δηλαδή ἡ εὐθεία πού διέρχεται ὅσο τό δυνατό πιό κοντά ἀπό τά ρ σημεῖα.

Παράδειγμα.

Στόν Πίνακα 6.5.1 δίνεται ἡ βαθμολογία πέντε σπουδαστῶν μιᾶς ἀνώτατης σχολῆς στά μαθήματα «Μαθηματικά» (X) καί «Στατιστική» (Y).

Νά βρεθεῖ ἡ εὐθεία ἐλαχίστων τετραγώνων τῶν πέντε αὐτῶν ζευγῶν.

ΠΙΝΑΚΑΣ 6.5.1.

Μαθη-ματικά x_i	Στατι-στική y_i
2	1
3	2
5	5
6	6
8	7

Λύση:

Γιά νά γράψομε τό σύστημα τῶν κανονικῶν ἐξισώσεων, θά πρέπει νά ὑπολογίσομε τά ἀθροίσματα $\sum_i x_i$, $\sum_i y_i$, $\sum_i x_i^2$, $\sum_i x_i\ y_i$. Συμπληρώνομε λοιπόν τόν πίνακα μέ στῆλες πού περιέχουν τά τετράγωνα τῶν τιμῶν x_i καί τά γινόμενα $x_i\ y_i$ καί βρίσκομε ἀμέσως (Πίνακας 6.5.2):

ΠΙΝΑΚΑΣ 6.5.2

x_i	y_i	x_i^2	$x_i\ y_i$
2	1	4	2
3	2	9	6
5	5	25	25
6	6	36	36
8	7	64	56
24	21	138	125

$\sum\limits_i x_i = 24, \sum\limits_i y_i = 21, \sum\limits_i x_i^2 = 138, \sum\limits_i x_i\ y_i = 125, \nu = 5$. Συνεπῶς, τό σύστημα (6.3) τῶν κανονικῶν ἐξισώσεων εἶναι:

$$5\alpha + 24\beta = 21$$

$$24\alpha + 138\beta = 125$$

καί λύνοντάς το, βρίσκομε:

$$\hat{\beta} = \frac{\nu \cdot \sum\limits_i x_i\ y_i - \left(\sum\limits_i x_i\right)\left(\sum\limits_i y_i\right)}{\nu \cdot \sum\limits_i x_i^2 - \left(\sum\limits_i y_i\right)^2} = \frac{5 \times 125 - 24 \times 21}{5 \times 138 - (24)^2} = 1{,}06$$

$$\overline{x} = \frac{24}{5} = 4{,}8, \qquad \overline{y} = \frac{21}{5} = 4{,}2$$

$$\hat{\alpha} = \overline{y} - \hat{\beta}\overline{x} = 4{,}2 - 1{,}06 \times 4{,}8 = -\ 0{,}888.$$

Ἑπομένως, ἡ εὐθεία ἐλαχίστων τετραγώνων θά ἔχει ἐξίσωση:

$$\hat{y} = -\ 0{,}888 + 1{,}06x.$$

Ἐπειδή οἱ ἀριθμοί $\hat{\alpha}$ καί $\hat{\beta}$ ἐπαληθεύουν τήν πρώτη ἀπό τίς ἐξισώσεις (6.3) (ἀφοῦ ἀποτελοῦν τή λύση τοῦ συστήματος τῶν κανονικῶν ἐξισώσεων), θά ἔχομε:

$$\nu\hat{\alpha} + \left(\sum_i x_i\right)\hat{\beta} = \left(\sum_i y_i\right) \text{ ἤ } \nu\hat{\alpha} + \frac{1}{\nu}\left(\sum_i x_i\right)\hat{\beta} = \frac{1}{\nu}\left(\sum_i y_i\right) \text{ ἤ } \hat{\alpha} + \overline{x}\hat{\beta} = \overline{y}.$$

Ἡ τελευταία ἰσότητα ἐκφράζει ὅτι οἱ ἀριθμοί $\overline{x} = \frac{1}{\nu}\sum\limits_i x_i$ καί $\overline{y} = \frac{1}{\nu}\sum\limits_i y_i$, δηλαδή οἱ ἀριθμητικοί μέσοι τῶν τιμῶν x_i καί τῶν τιμῶν y_i, ἐπαληθεύουν τήν ἐξίσωση (6.4). Συμπεραίνομε λοιπόν ὅτι:

Ἡ εὐθεία ἐλαχίστων τετραγώνων διέρχεται πάντοτε ἀπό τό σημεῖο Κ $(\overline{x}, \overline{y})$, τό ὁποῖο ἔχει συντεταγμένη τούς ἀριθμητικούς μέσους τῶν τετμημένων καί τεταγμένων τῶν ρ σημείων.

Β' ΣΥΣΧΕΤΙΣΜΕΝΕΣ ΜΕΤΑΒΛΗΤΕΣ

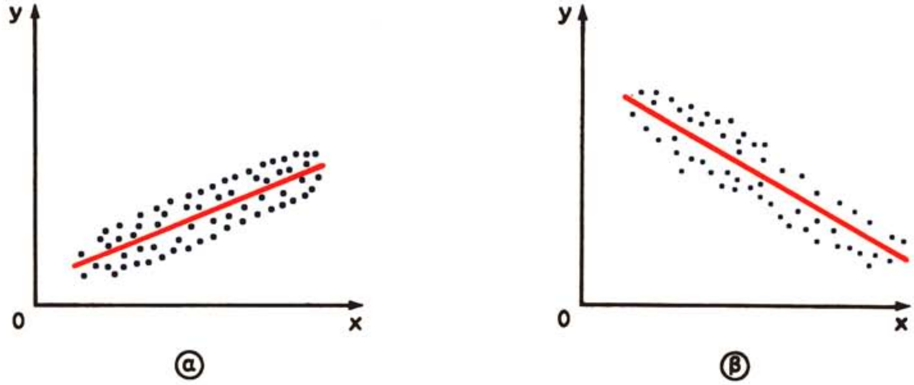
6.6 Ἡ γραμμική συμμεταβολή.

Ἄς ὑποθέσομε πάλι ὅτι ἐξετάζομε τά ν ἄτομα ἑνός πληθυσμοῦ ὡς πρός δύο μεταβλητές ἰδιότητές τους X καί Y. Ἄν σχηματίσομε ἀπό τίς παρατηρήσεις μας:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_\nu, y_\nu)$$

τό διάγραμμα συμμεταβολῆς καί δοῦμε ὅτι τά σημεῖα του βρίσκονται γύρω ἀπό μία εὐθεία (σχ. 6.6), θά λέμε ὅτι οἱ δύο μεταβλητές ἔχουν ***γραμμική συμμεταβολή*** ἤ ὅτι εἶναι ***συσχετισμένες***. Βλέπομε λοιπόν ὅτι στή Στατιστική ἡ ἔννοια «συσχέτιση» εἶναι συνδεδεμένη μέ τήν ἔννοια τῆς γραμμικῆς συμμεταβολῆς της (καί ὄχι μέ ὁποιαδήποτε ἐξάρτηση ἄλλης μορφῆς). Οἱ δύο μεταβλητές θά λέγονται:

— ***Θετικά συσχετισμένες***, ὅταν ἡ αὔξηση τῶν τιμῶν τῆς μιᾶς ἔχει ὡς συνέπεια καί τήν αὔξηση τῶν τιμῶν τῆς ἄλλης [σχ. 6.6 (α)].



Σχ. 6.6.

— ***Ἀρνητικά συσχετισμένες***, ὅταν ἡ αὔξηση τῶν τιμῶν τῆς μιᾶς ἔχει ὡς συνέπεια τήν μείωση τῶν τιμῶν τῆς ἄλλης [σχ. 6.6 (β)].

Σύμφωνα μέ ὅ,τι εἴπαμε στά προηγούμενα, ἡ εἰκόνα τῆς γραμμικῆς συμμεταβολῆς δύο μεταβλητῶν θά δίνεται μέ τήν εὐθεία τῶν ἐλαχίστων τετραγώνων τῶν ν σημείων (x_i, y_i) ἡ ὁποία (ἀφοῦ διέρχεται ὅπως γνωρίζομε ἀπό τό σημεῖο $(\overline{x}, \overline{y})$ μέ:

$\overline{x} = \frac{1}{\nu}\sum_i x_i$ καί $\overline{y} = \frac{1}{\nu}\sum_i y_i$ θά ἔχει ἐξίσωση τῆς μορφῆς:

$$y - \overline{y} = \beta(x - \overline{x}) \tag{6.5}$$

Ἡ εὐθεία αὐτή θά λέγεται τώρα ***εὐθεία παλινδρομήσεως*** τῆς y πάνω στή x καί τό β, τό ὁποῖο παριστάνει τό ***συντελεστή διευθύνσεώς της,*** θά λέγεται ***συντελεστής παλινδρομήσεως*** τῆς y πάνω στήν x καί θά δίνεται ἀπό τήν ἰσότητα:

$$\beta = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sum_i (x_i - \overline{x})^2} \qquad (6.7)$$

Εἶναι φανερό ὅτι οἱ δύο μεταβλητές θά εἶναι θετικά (ἤ ἀρνητικά) συσχετισμένες, ὅταν ὁ β εἶναι θετικός (ἤ ἀρνητικός) ἀριθμός.

6.7 Ἡ συνδιακύμανση δύο μεταβλητῶν.

Συνδιακύμανση τῶν δύο μεταβλητῶν X καί Y ὀνομάζεται ὁ ἀριθμός πού σημειώνεται Cov (X, Y) καί δίνεται ἀπό τήν ἰσότητα:

$$\text{Cov}(X, Y) = \frac{(x_1 - \overline{x})(y_1 - \overline{y}) + (x_2 - \overline{x})(y_2 - \overline{y}) + \ldots + (x_\nu - \overline{x})(y_\nu - \overline{y})}{\nu} =$$

$$= \frac{1}{\nu} \sum_i (x_i - \overline{x})(y_i - \overline{y}) \qquad (6.6)$$

ὅπου: $\overline{x}$ καί $\overline{y}$ εἶναι οἱ ἀριθμητικοί μέσοι τῶν τιμῶν τῆς X καί Y ἀντίστοιχα.

Ὁ τύπος (6.6) πού ὁρίζει τή συνδιακύμανση μπορεῖ νά γραφεῖ ἀκόμη, ἄν κάνομε πράξεις στό δεύτερο μέλος του:

$$\text{Cov}(X, Y) = \frac{1}{\nu} \sum_i (x_i y_i - \overline{x} y_i - \overline{y} x_i + \overline{x} y_i) =$$

$$= \frac{1}{\nu} \sum_i x_i y_i - \overline{x} \left(\frac{1}{\nu} \sum_i y_i\right) - \overline{y} \left(\frac{1}{\nu} \sum_i x_i\right) + \frac{1}{\nu} \nu \overline{x}\,\overline{y} =$$

$$= \frac{1}{\nu} \sum_i x_i y_i - \overline{x}\,\overline{y} - \overline{y}\,\overline{x} + \overline{x}\,\overline{y},$$

ἤ τελικά

$$\text{Cov}(X, Y) = \frac{1}{\nu} \sum_i x_i y_i - \overline{x}\,\overline{y}. \qquad (6.7)$$

Συνήθως ἡ συνδιακύμανση ὑπολογίζεται μέ τόν τύπο (6.7), γιατί, ὅταν τά ζεύγη (x_i, y_i) δίνονται σέ πίνακα, τό ἄθροισμα $\sum x_i y_i$ βρίσκεται ἀμέσως ἄν συμπληρώσομε τόν πίνακα μέ μιά στήλη πού περιέχει τό γινόμενο $x_i y_i$.

Ἄν οἱ δύο μεταβλητές X καί Y ἔχουν γραμμική συμμεταβολή, ὁ συντελεστής παλινδρομήσεως τῆς Y πάνω στή X γράφεται (ἐπειδή $\sum_i (x_i - \overline{x})(y_i - \overline{y}) =$ $\nu \,.\, \text{Cov}(X, Y)$ καί $\sum_i (x_i - \overline{y})^2 = \nu s_x^2$)

$$\beta = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sum_i (x_i - \overline{x})^2} = \frac{\nu \,.\, \text{Cov}(X,Y)}{\nu s_x^2} = \frac{\text{Cov}(X,Y)}{s_x^2} \qquad (6.8)$$

Βλέπομε δηλαδή ὅτι τό πρόσημο τοῦ β συμπίπτει μέ τό πρόσημο τῆς συνδιακυμάνσεως Cov (X, Y) ἀφοῦ τό s_x^2 εἶναι πάντα θετικό) καί συνεπῶς:

— Ὅταν ἡ συνδιακύμανση Cov (X, Y) εἶναι θετικός ἀριθμός, οἱ μεταβλητές X καί Y εἶναι θετικά συσχετισμένες.

— Ὅταν ἡ συνδιακύμανση Cov (X, Y) εἶναι ἀρνητικός ἀριθμός, οἱ μεταβλητές X καί Y εἶναι ἀρνητικά συσχετισμένες.

— Ὅταν ἡ συνδιακύμανση εἶναι μηδέν, τότε ἔχομε καί β = 0 καί αὐτό σημαίνει ὅτι δέν ὑπάρχει γραμμική συμμεταβολή τῶν X καί Y. Στήν περίπτωση αὐτή, λέμε ὅτι οἱ μεταβλητές X καί Y εἶναι ***ἀσυσχέτιστες.***

6.8 Ὁ συντελεστής συσχετίσεως.

Ἐπειδή ἡ συνδιακύμανση Cov (X, Y) ἐξαρτᾶται ἀπό τίς μονάδες μετρήσεως τῶν X καί Y (ἡ μονάδα μετρήσεως τῆς Cov (X, Y) εἶναι ἴση μέ τό γινόμενο τῶν μονάδων μετρήσεως τῶν δύο μεταβλητῶν) δέν μπορεῖ νά ἐκφράσει μέ ἀντικειμενικό τρόπο τό βαθμό τῆς γραμμικῆς συμμεταβολῆς τους οὔτε καί νά χρησιμοποιηθεῖ γιά τή σύγκριση τοῦ βαθμοῦ γραμμικῆς συμμεταβολῆς διαφορετικῶν κατανομῶν. Γι' αὐτό ἀκριβῶς παίρνομε ὡς μέτρο τῆς γραμμικῆς συμμεταβολῆς δύο μεταβλητῶν X καί Y ἕνα καθαρό ἀριθμό, ὁ ὁποῖος σημειώνεται μέ r καί ὁρίζεται ἀπό τήν ἰσότητα:

$$r = \frac{\mathrm{Cov}\,(X, Y)}{s_x\, s_y} = \frac{\mathrm{Cov}\,(X, Y)}{\sqrt{V(X) \,.\, V(Y)}} \qquad (6.9)$$

Ὁ ἀριθμός αὐτός r λέγεται ***συντελεστής συσχετίσεως*** τῶν X καί Y καί ὅπως εἶναι φανερό, ἔχει τό ἴδιο πρόσημο μέ τή συνδιακύμανση Cov (X, Y) (ἀφοῦ οἱ τυπικές ἀποκλίσεις s_x καί s_y εἶναι πάντα θετικοί ἀριθμοί). Γιά τό συντελεστή συσχετίσεως r ἰσχύουν οἱ ἑξῆς ἰδιότητες:

Εἶναι ἀριθμός ἀπόλυτα μικρότερος τῆς μονάδας, δηλαδή $-1 \leqslant r \leqslant 1$.
Ἄν $r = 1$ ἤ $r = -1$, οἱ μεταβλητές X καί Y ἔχουν συναρτησιακή ἐξάρτηση γραμμικῆς μορφῆς καί ἀντιστρόφως.
Ἄν $r = 0$, οἱ μεταβλητές X καί Y εἶναι ἀσυσχέτιστες (δέν ἔχουν γραμμική συμμεταβολή).

Ἀπό τίς ἰδιότητες αὐτές, καταλαβαίνομε ὅτι, γιά νά ὑπάρχει γραμμική συμμεταβολή τῶν X καί Y, θά πρέπει ὁ ἀριθμός $|r|$ νά πλησιάζει πρός τή μονάδα. Ἀντίθετα, ὅταν τό $|r|$ πλησιάζει πρός τό μηδέν, δέν ἔχομε γραμμική συμμεταβολή (δίχως νά ἀποκλείεται ἡ ὕπαρξη καμπυλόγραμμης συμμεταβολῆς).

Ἄν καί οἱ τιμές τοῦ r, γιά τίς ὁποῖες δεχόμαστε τήν ὕπαρξη συσχετίσεως, ἐξαρτῶνται ἀπό τό πλῆθος τῶν παρατηρήσεων, ἐνδεικτικά θεωροῦμε ὅτι:

— Ἄν $|r| \leqslant 0,30$, δέν ἔχομε συσχέτιση.
— Ἄν $0,30 \leqslant |r| \leqslant 0,50$, ἔχομε ἀσθενή συσχέτιση.
— Ἄν $0,50 \leqslant |r| \leqslant 0,70$, ἔχομε μέση συσχέτιση.
— Ἄν $0,70 \leqslant |r| \leqslant 0,80$, ἔχομε ἰσχυρή συσχέτιση.
— Ἄν $|r| \leqslant 0,80$, ἔχομε πολύ ἰσχυρή συσχέτιση.
— Ἄν $|r| = 1$, ἔχομε «τέλεια» συσχέτιση.

Ἄν ἀπό τήν τιμή τοῦ $|r|$ μποροῦμε νά δεχθοῦμε γιά τίς μεταβλητές X καί Y ὕπαρξη συσχετίσεως, τότε γράφομε τήν ἐξίσωση (6.5), ἡ ὁποία δίνει τήν εἰκόνα τῆς συμμεταβολῆς τους. Στήν περίπτωση αὐτή, ὁ συντελεστής παλινδρομήσεως β γράφεται (ἐπειδή εἶναι Cov (X, Y) = $rs_x\, s_y$):

$$\beta = \frac{\mathrm{Cov}\,(X,Y)}{s_x^2} = \frac{rs_x\, s_y}{s_x^2} = r\,\frac{s_y}{s_x} \tag{6.10}$$

καί συνεπῶς ἡ εὐθεία παλινδρομήσεως τῆς Υ πάνω στή Χ θά ἔχει ἐξίσωση:

$$y = \overline{y} + r\frac{s_y}{s_x}(x - \overline{x}) \tag{6.11}$$

6.9 Ὑπολογισμός τοῦ r.

Γιά νά βροῦμε τό συντελεστή συσχετίσεως r ὑπολογίζομε τή συνδιακύμανση Cov(X, Y) καί τίς τυπικές ἀποκλίσεις s_x καί s_y τῶν δύο μεταβλητῶν καί ἐφαρμόζομε τόν τύπο:

$$r = \frac{\mathrm{Cov}\,(X,Y)}{s_x\, s_y} \tag{6.12}$$

Καθένας ἀπό τούς ἀριθμούς Cov (X, Y), s_x, s_y μπορεῖ νά ὑπολογισθεῖ μέ μιά ὁποιαδήποτε ἀπό τίς ἐκφράσεις του:

$$\mathrm{Cov}\,(X,Y) = \frac{1}{\nu}\sum_i (x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{\nu}\sum_i x_i y_i - \overline{x}\,\overline{y} =$$

$$= \frac{1}{\nu}\left[\sum_i x_i y_i - \frac{1}{\nu}\Big(\sum_i x_i\Big)\Big(\sum_i y_i\Big)\right]$$

$$s_x = \sqrt{\frac{1}{\nu}\sum_i (x_i - \overline{x})^2} = \sqrt{\frac{1}{\nu}\sum_i x_i^2 - \overline{x}^2} = \sqrt{\frac{1}{\nu}\left[\sum_i x_i^2 - \frac{1}{\nu}\Big(\sum_i x_i\Big)^2\right]}$$

$$s_y = \sqrt{\frac{1}{\nu}\sum_i (y_i - \overline{y})^2} = \sqrt{\frac{1}{\nu}\sum_i y_i^2 - \overline{y}^2} = \sqrt{\frac{1}{\nu}\left[\sum_i y_i^2 - \frac{1}{\nu}\Big(\sum_i y_i\Big)^2\right]}$$

Ἔτσι, βλέπομε ὅτι ὁ τύπος (6.12) γράφεται ἄν ἀντικαταστήσομε τή Cov (X, Y), s_x, s_y, μέ τίς τελευταῖες ἐκφράσεις τους:

$$r = \frac{\sum_i x_i y_i - \frac{1}{\nu}\Big(\sum_i x_i\Big)\Big(\sum_i y_i\Big)}{\sqrt{\sum_i x_i^2 - \frac{1}{\nu}\Big(\sum_i x_i\Big)^2 \;\; \sum_i y_i^2 - \frac{1}{\nu}\Big(\sum_i y_i\Big)^2}}$$

ἤ τελικά μετά τίς πράξεις:

$$r = \frac{\nu\sum_i x_i y_i - \Big(\sum_i x_i\Big)\Big(\sum_i y_i\Big)}{\sqrt{\nu\sum_i x_i^2 - \Big(\sum_i x_i\Big)^2}\,\sqrt{\nu\sum_i y_i^2 - \Big(\sum_i y_i\Big)^2}} \tag{6.13}$$

Παράδειγμα.

Ὁ Πίνακας 6.9.1 δίνει τήν ἡλικία (x_i) καί τήν πίεση τοῦ αἵματος (y_i) δέκα γυναικῶν.

Νά ἐξετασθεῖ ἄν ὑπάρχει συσχέτιση τῶν δύο μεταβλητῶν καί, ἄν ὑπάρχει, νά βρεθεῖ ἡ γραμμή παλινδρομήσεως τῆς Y πάνω στή X.

ΠΙΝΑΚΑΣ 6.9.1.

x_i	y_i
56	15
42	12
72	16
36	12
63	15
55	15
49	14
38	12
42	14
47	13
500	138

Λύση:

Γιά νά δοῦμε ἄν ὑπάρχει συσχέτιση, θά ὑπολογίσομε τό συντελεστή συσχετίσεως r μέ τόν τύπο (6.12) ἤ (6.13). Σχηματίζομε λοιπόν Πίνακα 6.9.2 πού νά ἔχει ἀπαραίτητα στῆλες γιά τόν ὑπολογισμό τοῦ r καί αὐτός εἶναι:

$$\left(\text{ἐπειδή } \bar{x} = \frac{1}{\nu} \sum_i x_i = \frac{500}{10} = 50 \quad \text{καί} \quad \bar{y} = \frac{1}{\nu} \sum_i y_i = \frac{138}{10} = 13{,}8\right)$$

ΠΙΝΑΚΑΣ 6.9.2.

x_i	y_i	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$x_i y_i$	x_i^2	y_i^2
56	15	6	1,2	7,2	36	1,44	840	3136	225
42	12	−8	−1,8	14,4	64	3,24	504	1764	144
72	16	22	2,2	48,4	484	4,84	1152	5184	256
36	12	14	−1,8	25,2	196	3,84	432	1296	144
63	15	13	1,2	15,6	169	1,44	945	3969	225
55	15	5	1,2	6	25	1,44	825	3025	225
49	14	−1	0,2	−0,2	1	0,04	686	2401	196
38	12	12	−1,8	21,6	144	3,24	456	1444	144
42	14	−8	0,2	−1,6	64	0,04	588	1764	196
47	13	−3	−0,8	2,4	9	0,69	611	2209	169
500	138			139	1192	19,60	7039	26192	1924

Ἀπό τόν πίνακα αὐτό, βρίσκομε ἀμέσως:

$$s_x = \sqrt{\frac{1}{\nu} \sum_i (x_i - \bar{x})^2} = \sqrt{\frac{1192}{10}} = \sqrt{119{,}2} = 10{,}918$$

$$s_y = \sqrt{\frac{1}{v} \sum_i (y_i - \overline{y})^2} = \sqrt{\frac{19{,}60}{10}} = \sqrt{1{,}96} = 1{,}4$$

$$\mathrm{Cov}\,(X, Y) = \frac{1}{v} \sum_i (x_i - \overline{x})(y_i - \overline{y}) = \frac{139}{10} = 13{,}9.$$

Ἔχομε λοιπόν, ἄν χρησιμοποιήσομε τόν τύπο (6.12):

$$r = \frac{\mathrm{Cov}\,(X, Y)}{s_x \, s_y} = \frac{13{,}9}{(10{,}918)\,(1{,}4)} = \frac{13{,}9}{15{,}2852} = 0{,}909$$

Τό r μπορεῖ νά ὑπολογισθεῖ καί μέ τόν τύπο (6.13), στόν ὁποῖο ἀποφεύγομε τόν ὑπολογισμό τῶν Cov (X, Y), s_x s_y καί χρησιμοποιοῦμε μόνο τά ἀθροίσματα $\sum x_i y_i$, $\sum x$, $\sum y_i$, $\sum x_i^2$, $\sum y_i^2$ τοῦ παραπάνω πίνακα. Ἔχομε τότε:

$$r = \frac{v \sum_i x_i y_i - (\sum_i x_i)(\sum_i y_i)}{\sqrt{\left[v \cdot \sum_i x_i^2 - (\sum_i x_i)^2\right] \cdot \left[v \sum_i y_i^2 - (\sum_i y_i)^2\right]}} =$$

$$= \frac{10 \times 7039 - 500 \times 138}{\sqrt{[10 \times 26192 - (500)^2] \times [10 \times 1924 - (138)^2]}} = \frac{1390}{1528{,}5} = 0{,}909$$

Εἶναι φανερό ὅτι, γιά τόν ὑπολογισμό τοῦ r μέ τόν τύπο αὐτό χρειάζονται μόνο οἱ τρεῖς τελευταῖες στῆλες τοῦ παραπάνω πίνακα καί φυσικά οἱ δύο στῆλες πού μᾶς παρέχουν τά δεδομένα. Ἀπό τήν τιμή r = 0,909 καταλαβαίνομε ὅτι ὑπάρχει ἔντονη θετική συσχέτιση (γραμμική συμμεταβολή) μεταξύ τῆς ἡλικίας καί τῆς πιέσεως τῶν 10 γυναικῶν.

Ἡ γραμμική συμμεταβολή τους θά δίνεται ἀπό τήν ἐξίσωση (6.12), ἡ ὁποία γράφεται τώρα (ἀφοῦ $\overline{x} = 50$, $\overline{y} = 13{,}8$, $r = 0{,}909$, $s_x = 10{,}918$, $s_y = 1{,}4$):

$$y = 13{,}8 + 0{,}909 \, \frac{1{,}4}{10{,}918} \, (x_i - 50) = 13{,}8 + 0{,}117\,(x - 50),$$

ἤ τελικά $y = 0{,}117x + 7{,}95$.

6.10 Πίνακες συχνοτήτων δύο μεταβλητῶν.

Ὅταν ἐξετάσομε τά ν ἄτομα ἑνός πληθυσμοῦ ὡς πρός δύο μεταβλητές ἰδιότητές τους, οἱ ν παρατηρήσεις μας:

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_v, y_v)$$

δέν εἶναι ἀπαραίτητα διαφορετικές μεταξύ τους καί μπορεῖ ὁρισμένα ἀπό τά παραπάνω ζεύγη νά εἶναι ἴσα.

Ὁ ἀριθμός πού ἐκφράζει πόσες φορές ἐμφανίζεται στίς ν παρατηρήσεις μας ἕνα ὁρισμένο ζεῦγος (x_i, y_i), ὀνομάζεται ***συχνότητα*** τοῦ ζεύγους αὐτοῦ καί θά σημειώνεται f_{ij}. Οἱ συχνότητες ὅλων τῶν διαφορετικῶν μεταξύ τους ζευγῶν δίνονται μέ τόν Πίνακα 6.10.1 ***διπλῆς εἰσόδου.***

ΠΙΝΑΚΑΣ 6.10.1.

$x_i \backslash y_j$	y_1	y_2	....	y_j	....	y_λ	$f_{i\cdot}$
x_1	f_{11}	f_{12}	....	f_{1j}	....	$f_{1\lambda}$	$f_{1\cdot}$
x_2	f_{21}	f_{22}	....	f_{2j}	....	$f_{2\lambda}$	$f_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	f_{i1}	f_{i2}	....	f_{ij}	....	$f_{i\lambda}$	$f_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_κ	$f_{\kappa 1}$	$f_{\kappa 2}$	....	$f_{\kappa j}$	....	$f_{\kappa\lambda}$	$f_{\kappa\cdot}$
$f_{\cdot j}$	$f_{\cdot 1}$	$f_{\cdot 2}$	....	$f_{\cdot j}$	....	$f_{\cdot\lambda}$	ν

Στόν πίνακα αὐτό, οἱ γραμμές ἀντιστοιχίζονται στίς διαφορετικές μεταξύ τους τιμές $x_1, x_2, ..., x_\kappa$ τῆς μεταβλητῆς X καί οἱ στῆλες στίς διαφορετικές μεταξύ τους τιμές $y_1, y_2, ..., y_\lambda$ τῆς μεταβλητῆς Y, ἐνῶ στό «τετραγωνάκι» πού βρίσκεται στήν i γραμμή καί j στήλη τοῦ πίνακα ἔχομε τή συχνότητα f_{ij} τοῦ ζεύγους (x_i, y_j).

Τό σύνολο τῶν τριάδων (x_i, y_j, f_{ij}) λέγεται ***μικτή κατανομή τῶν μεταβλητῶν X καί Y*** ἤ ***κατανομή τοῦ ζεύγους (X,Y).***

6.11 Περιθωριακές κατανομές.

Ἄν στόν παραπάνω πίνακα διπλῆς εἰσόδου σημειώσομε μέ $f_{i.}$ τό ἄθροισμα συχνοτήτων τῆς i γραμμῆς, δηλαδή ἄν θέσομε:

$$f_{i.} = \sum_{j=1}^{\lambda} f_{ij}$$

οἱ ἀριθμοί $f_1, f_2, ..., f_i, ..., f_\kappa$, παριστάνουν τίς συχνότητες τῶν τιμῶν $x_1, x_2, ..., x_i, ..., x_\kappa$ τῆς μεταβλητῆς X, ἀνεξάρτητα ἀπό τίς τιμές πού παίρνει ἡ μεταβλητή Y. Ἔτσι ἡ μονοδιάστατη κατανομή συχνοτήτων τῆς μεταβλητῆς X, ἡ ὁποία λέγεται ***πρώτη περιθωριακή κατανομή*** τοῦ ζεύγους (X, Y) θά εἶναι αὐτή τοῦ Πίνακα 6.11.1.

ΠΙΝΑΚΑΣ 6.11.1.

x_i	f_i
x_1	$f_{1\cdot}$
x_2	$f_{2\cdot}$
x_3	$f_{3\cdot}$
.	.
.	.
.	.
x_i	$f_{i\cdot}$
.	.
.	.
.	.
x_κ	$f_{\kappa\cdot}$

Ἀπό τόν πίνακα αὐτό, μποροῦμε νά βροῦμε ἀμέσως τίς χαρακτηριστικές τιμές τῆς μεταβλητῆς X, πού θά εἶναι:

$$\bar{x} = \frac{\sum_i f_i . x_i}{\sum_i f_i.} = \frac{1}{\nu} \sum_i f_i . x_i \qquad s_x^2 = \frac{\sum_i f_i . (x_i - \bar{x})^2}{\sum_i f_i} = \frac{1}{\nu} \sum_i f_i . (x_i - \bar{x})^2$$

γιατί, ὅπως εἶναι φανερό, θά ἔχομε $\sum_{i=1}^{\kappa} f_i . = f_1 . + f_2 . + \ldots + f_\kappa . = \nu$

Ἐπίσης, οἱ ἀριθμοί $f_{.1}, f_{.2}, \ldots, f_{.j}, \ldots, f_{.\lambda}$ τῆς τελευταίας γραμμῆς τοῦ πίνακα διπλῆς εἰσόδου, πού παριστάνουν τίς συχνότητες τῶν τιμῶν $y_1, y_2, \ldots y_j, \ldots, y_\lambda$ τῆς μεταβλητῆς Y, μᾶς δίνουν ***τήν περιθωριακή κατανομή τῆς μεταβλητῆς Y,*** ὅπως φαίνεται στόν Πίνακα 6.11.2.

Κατά τόν ἴδιο τρόπο, ἄν στόν πίνακα αὐτόν σημειώσομε μέ $f_{.j}$ τό ἄθροισμα τῶν συχνοτήτων τῆς j στήλης, δηλαδή ἄν θέσομε:

$$f_{.j} = \sum_{i=1}^{\kappa} f_{ij},$$

οἱ ἀριθμοί $f_{.1}, f_{.2}, \ldots, f_{.j}, \ldots, f_{.\lambda}$ παριστάνουν τίς συχνότητες τῶν τιμῶν $y_1, y_2, \ldots y_\lambda$ τῆς μεταβλητῆς Y, ἀνεξάρτητα ἀπό τίς τιμές πού παίρνει ἡ X. Ἔτσι, ἡ μονοδιάστατη κατανομή συχνοτήτων τῆς μεταβλητῆς Y, ἡ ὁποία λέγεται ***δευτέρα περιθωριακή κατανομή*** τοῦ ζεύγους (X, Y) θά εἶναι:

ΠΙΝΑΚΑΣ 6.11.2.

y_j	$f_{.j}$
y_1	$f_{.1}$
y_2	$f_{.2}$
.	.
.	.
.	.
y_j	$f_{.j}$
.	.
.	.
.	.
y_λ	$f_{.\lambda}$

Ἀπό τόν πίνακα αὐτό μποροῦμε νά βροῦμε ἀμέσως τίς χαρακτηριστικές τιμές τῆς μεταβλητῆς Y, πού θά εἶναι:

$$\bar{y} = \frac{\sum_j f_{.j}\, y_j}{\sum_j f_{.j}} = \frac{1}{\nu} \sum_j f_{.j}\, y_j, \quad s_y^2 = \frac{\sum_j f_{.j}\, (y_j - \bar{y})^2}{\sum_j f_{.j}} = \frac{1}{\nu} \sum_j f_{.j}\, (y_j - \bar{y})^2,$$

γιατί, ὅπως εἶναι φανερό, θά ἔχομε $\sum_j f_{.j} = f_{.1} + f_{.2} + \ldots + f_{.\lambda} = \nu$.

6.12 Ὑπολογισμός τῆς συνδιακυμάνσεως ἀπό πίνακα συχνοτήτων.

Γιά νά βροῦμε τή συνδιακύμανση τῶν X καί Y μέ τόν τύπο (6.7), θά πρέπει νά ὑπολογίσομε τό ἄθροισμα ὅλων τῶν γινομένων πού προκύπτουν ἀπό τόν πολλαπλασιασμό τῶν δύο τιμῶν τοῦ ζεύγους κάθε παρατηρήσεως. Τό ἄθροισμα ὅμως αὐτό γράφεται τώρα [ἀφοῦ ἡ κάθε παρατήρηση (x_i, y_j) ἐμφανίζεται f_{ij} φορές]:

$$f_{11}x_1y_1 + f_{12}x_1y_2 + \dots + f_{1\lambda}x_1y_\lambda + f_{21}x_2y_1 + \dots + f_{ij}x_i\,y_j + \dots + f_{\kappa\lambda}\,x_\kappa\,y_\lambda,$$

ἤ πιό σύντομα, $\sum_i\sum_j f_{ij}\,x_i\,y_j$. Ἔτσι λοιπόν, ἡ συνδιακύμανση θά δίνεται ἀπό τήν ἰσότητα:

$$\operatorname{Cov}(X, Y) = \frac{1}{v}\sum_i\sum_j f_{ij}\,x_i\,y_j - \bar{x}\,\bar{y} \tag{6.14}$$

Παράδειγμα.

Στόν Πίνακα 6.12.1 διπλῆς εἰσόδου νά ὑπολογισθεῖ ἡ συνδιακύμανση τῶν X καί Y:

ΠΙΝΑΚΑΣ 6.12.1

$x_i \backslash y_j$	3	2	1	f_i
1	–	2	1	3
2	1	2	–	3
3	1	–	–	1
$f_{\cdot j}$	2	4	1	7

Γιά τόν ὑπολογισμό τῆς διακυμάνσεως σχηματίζομε τόν Πίνακα 6.12.2 ἀριθμητικῶν ὑπολογισμῶν:

ΠΙΝΑΚΑΣ 6.12.2.

$x_i \backslash y_j$	3	2	1	$f_{i\cdot}$	$f_{i\cdot}\cdot x_i$
1	–	2	1	3	3
2	1	2	–	3	6
3	1	–	–	1	3
$f_{\cdot j}$	2	4	1	7	12
$f_{\cdot j}y_j$	6	8	1	15	

καί ἀπό τόν πίνακα αὐτό βρίσκομε:

$$\bar{x} = \frac{1}{v}\sum_i f_{i\cdot}\cdot x_i = \frac{12}{7} = 1{,}714,\ \bar{y} = \frac{1}{v}\sum_j f_{\cdot j}y_j = \frac{15}{7} = 2{,}142,$$

$\sum\sum x_i y_j f_{ij} = 1 \times 0 \times 3 + 1 \times 2 \times 2 + 1 \times 1 \times 1 + 2 \times 1 \times 3 + 2 \times 2 \times 2 + 2 \times 0 \times 1 + 3 \times 1 \times 3 + 3 \times 0 \times 2 + 3 \times 0 \times 1 = 28$

Ἐπομένως, ἡ συνδιακύμανση θά εἶναι:

$$\mathrm{Cov}\,(X, Y) = \frac{1}{v} \sum_i \sum_j f_{ij} x_i y_j - \bar{x}\,\bar{y} =$$

$$= \frac{28}{7} - 1{,}714 \times 2{,}142 = \frac{28}{7} - 3{,}67 = 4 - 3{,}67 = 0{,}33.$$

Εἶναι φανερό ὅτι, ἄν θέλαμε νά ὑπολογίσομε στό παράδειγμά μας καί τίς διακυμάνσεις s_x^2, s_y^2, θά ἔπρεπε νά συμπληρώσομε τόν παραπάνω πίνακα μέ μία ἀκόμη γραμμή πού νά περιέχει τό γινόμενο $y_j^2\ f_{\cdot j}$ καθώς καί μέ μιά ἀκόμη στήλη πού νά περιέχει τό γινόμενο $x_i^2\ f_i$.

6.13 Ὑπολογισμός τοῦ r ἀπό πίνακα συχνοτήτων.

Ἀφοῦ εἴδαμε πῶς ὑπολογίζονται ἀπό πίνακα συχνοτήτων ἡ συνδιακύμανση Cov (X, Y) καί οἱ τυπικές ἀποκλίσεις s_x, s_y, γιά τόν ὑπολογισμό τοῦ r δέν ἔχομε παρά νά ἐφαρμόσομε τή γνωστή ἰσότητα:

$$r = \frac{\mathrm{Cov}\,(X, Y)}{s_x\, s_y} = \frac{\mathrm{Cov}\,(X, Y)}{\sqrt{V(X)\,V(Y)}}$$

Μποροῦμε ἐπίσης νά ἐφαρμόζομε τόν τύπο (6.13), ὁ ὁποῖος, στήν περίπτωση πού τά δεδομένα δίνονται ἀπό πίνακα συχνοτήτων, θά γράφεται:

$$r = \frac{v \sum_i \sum_j f_{ij} x_i y_j - (\sum_i f_{i\cdot} \cdot x_i)(\sum_j f_{\cdot j}\, y_j)}{\sqrt{v \sum_i f_{i\cdot}\, x_i^2 - (\sum_i f_{i\cdot} \cdot x_i)^2}\ \sqrt{v \sum_j f_{\cdot j}\, y_j^2 - (\sum_j f_{\cdot j}\, y_j)^2}} \qquad (6.15)$$

Παράδειγμα.

Νά βρεθεῖ ὁ συντελεστής συσχετίσεως τῶν δύο μεταβλητῶν X καί Y τῆς κατανομῆς τοῦ προηγουμένου παραδείγματος (Πίνακας 6.12.1).

Συμπληρώνομε τόν Πίνακα 6.12.1 μέ μία γραμμή πού περιέχει τό γινόμενο $f_{\cdot j}\, y_j^2$ καί μία στήλη πού περιέχει τό γινόμενο $f_i \cdot x_i^2$ καί καταλήγομε στόν Πίνακα 6.13.1:

ΠΙΝΑΚΑΣ 6.13.1.

x_i \ y_j	3	2	1	$f_{i\cdot}$	$f_{i\cdot} \cdot x_i$	$f_{i\cdot} \cdot x_i^2$
1	–	2	1	3	3	3
2	1	2	–	3	6	12
3	1	–	–	1	3	9
$f_{\cdot j}$	2	4	1	7	12	24
$y_j f_{\cdot j}$	6	8	1	15		
$y_j^2 f_{\cdot j}$	18	16	1	35		

'Εφαρμόζοντας τώρα κατ' εὐθεῖαν τόν τύπο (6.15) ἔχομε:

$$r = \frac{7 \times 28 - 12 \times 15}{\sqrt{7 \times 24 - (12)^2}\sqrt{7 \times 35 - (15)^2}} = 0{,}73$$

καί συνεπῶς βλέπομε ὅτι ὑπάρχει συσχέτιση μεταξύ τῶν X καί Y.

Στήν περίπτωση πού οἱ μεταβλητές X καί Y εἶναι συνεχεῖς, διαιροῦμε πρῶτα τά διαστήματα μεταβολῆς τους σέ κλάσεις καί μετά ὁμαδοποιοῦμε τίς παρατηρήσεις στά διάφορα ζεύγη τῶν κλάσεων, παίρνοντας γιά τιμές x_i καί y_j τά κέντρα τους.

'Η γενική μορφή ἑνός τέτοιου πίνακα παρουσιάζεται στόν Πίνακα 6.13.2.

ΠΙΝΑΚΑΣ 6.13.2.

τάξεις \ τάξεις		$\beta_0 - \beta_1$	$\beta_1 - \beta_2$	...	$\beta_{j-1} - \beta_j$	...	$\beta_{\lambda-1} - \beta_\lambda$	
τάξεις	x_i \ y_j	y_1	y_2	...	y_j	...	y_λ	$f_{i\cdot}$
$\alpha_0 - \alpha_1$	x_1	f_{11}	f_{12}	...	f_{1j}	...	$f_{1\lambda}$	$f_{1\cdot}$
$\alpha_1 - \alpha_2$	x_2	f_{21}	f_{22}	...	f_{2j}	...	$f_{2\lambda}$	$f_{2\cdot}$
.	.	....	....	....	....	....	....	.....
$\alpha_{i-1} - \alpha_i$	x_i	f_{i1}	f_{i2}	...	f_{ij}	...	$f_{i\lambda}$	$f_{i\cdot}$
.	.	....	....	....	....	....	....	.....
$\alpha_{\kappa-1} - \alpha_\kappa$	x_κ	$f_{\kappa 1}$	$f_{\kappa 2}$	...	$f_{\kappa j}$	...	$f_{\kappa\lambda}$	$f_{\kappa\cdot}$
	$f_{\cdot j}$	$f_{\cdot 1}$	$f_{\cdot 2}$	...	$f_{\cdot j}$	...	$f_{\cdot\lambda}$	ν

Στόν παραπάνω πίνακα, οἱ ἀριθμοί $y_1, y_2, \dots, y_j, \dots, y_\lambda$ ἀντιπροσωπεύουν τίς κεντρικές τιμές (y_j) τῶν κλάσεων πού ἀναφέρονται στή μεταβλητή Y, ἐνῶ οἱ ἀριθμοί $x_1, x_2, \dots, x_i, \dots, x_\kappa$ τίς κεντρικές τιμές (x_i) τῶν κλάσεων πού ἀναφέρονται στή μεταβλητή X.

'Ο συντελεστής συσχετίσεως καί στήν περίπτωση αὐτή τῶν συνεχῶν μεταβλητῶν ὑπολογίζεται μέ τόν τύπο (6.15), δηλαδή:

$$r = \frac{\nu \sum_i \sum_j f_{ij} x_i y_j - (\sum_i f_{i.} \cdot x_i)(\sum_j f_{\cdot j} y_j)}{\sqrt{\nu \sum_i f_{i.} x_i^2 - (\sum_i f_{i.} \cdot x_i)^2}\sqrt{\nu \sum_j f_{\cdot j} y_j^2 - (\sum_j f_{\cdot j} y_j)^2}}$$

6.14 'Ανεξάρτητες μεταβλητές.

'Υπολογίζοντας τό συντελεστή συσχετίσεως r δύο μεταβλητῶν X καί Y, διαπιστώνομε ἄν αὐτές εἶναι συσχετισμένες ἤ ὄχι, δηλαδή ἄν ἔχουν ἤ ὄχι γραμμική συμμεταβολή. Στήν περίπτωση τώρα πού οἱ μεταβλητές εἶναι ἀσυσχέτιστες, θά πρέπει νά ἐξετάσομε ἄν ἔχουν ἄλλου εἴδους ἐξάρτηση ἤ ἄν εἶναι μεταξύ τους ἀνεξάρτητες. Τό θέμα τῆς ἀλληλοεξαρτήσεως ἤ ἀνεξαρτησίας δύο μεταβλητῶν εἶναι ἀκόμη πιό ἔντονο στήν περίπτωση πού οἱ μεταβλητές X καί Y εἶναι ποιοτικές, γιατί τότε δέν ὑπάρχει οὔτε ἡ ἔννοια τοῦ συντελεστῆ συσχετίσεως.

Θά δοῦμε τώρα πῶς μποροῦμε νά διακρίνομε ἀπό τόν πίνακα διπλῆς εἰσόδου τῶν συχνοτήτων δύο (ποσοτικῶν ἤ ποιοτικῶν) μεταβλητῶν X καί Y, ἄν οἱ μεταβλητές αὐτές ἔχουν ἀλληλοεξάρτηση ἤ εἶναι μεταξύ τους ἀνεξάρτητες (Πίνακας 6.14.1).

ΠΙΝΑΚΑΣ 6.14.1.

Συχνοτήτων f_{ij}

$x_i \backslash y_j$	y_1	y_2	...	y_j...	y_λ	$f_{i\cdot}$
x_1	f_{11}	f_{12}	...	f_{1j}...	$f_{1\lambda}$	$f_{1\cdot}$
x_2	f_{21}	f_{22}	...	f_{2j}...	$f_{2\lambda}$	$f_{2\cdot}$
. . . x_i	. . . f_{i1}	. . . f_{i2}	. . . ...	. . . f_{ij}...	. . . $f_{i\lambda}$	. . . $f_{i\cdot}$
. . . x_K	...	...	...	...	...	...
$f_{\cdot j}$	$f_{\cdot 1}$	$f_{\cdot 2}$	...	$f_{\cdot j}$	$f_{\cdot \lambda}$	v

Δύο μεταβλητές X καί Y εἶναι ἀνεξάρτητες, ὅταν οἱ τιμές τῆς Y δέν ἐπηρεάζονται ἀπό τίς τιμές τῆς X. Γιά νά συμβαίνει αὐτό, θά πρέπει οἱ σχετικές συχνότητες μιᾶς ὁποιασδήποτε τιμῆς y_j τῆς Y νά εἶναι ἴδιες σέ ὅλα τά ἄτομα τοῦ πληθυσμοῦ, στά ὁποῖα ἡ ἄλλη μεταβλητή X παίρνει μιά ὁποιαδήποτε ἀπό τίς τιμές της $x_1, x_2, ..., x_K$. Παρατηροῦμε ὅμως ὅτι ἡ σχετική συχνότητα τῆς τιμῆς y_j, στά f_i ἄτομα, στά ὁποῖα ἡ X παίρνει τήν τιμή x_i, εἶναι f_{ij}/f_i (ἀφοῦ ἡ μεταβλητή Y παίρνει τήν τιμή y_j μόνο στά f_{ij} ἀπό τά ἄτομα αὐτά). Ἔτσι λοιπόν οἱ X καί Y θά εἶναι ἀνεξάρτητες, ὅταν ἔχομε τίς ἰσότητες:

$$\frac{f_{1j}}{f_{1\cdot}} = \frac{f_{2j}}{f_{2\cdot}} = \ldots = \frac{f_{ij}}{f_{i\cdot}} = \ldots = \frac{f_{Kj}}{f_{K\cdot}} = \frac{\sum_i f_{ij}}{\sum_i f_{i\cdot}} = \frac{f_{\cdot j}}{v}$$

δηλαδή, ὅταν γιά κάθε ζεῦγος (x_i, y_j) ἔχομε τήν ἰσότητα:

$$f_{ij} = \frac{f_{i\cdot} \cdot f_{\cdot j}}{v} \qquad (6.16)$$

Συμπεραίνομε λοιπόν ὅτι ***δύο μεταλητές X καί Y εἶναι ἀνεξάρτητες, ὅταν ἡ συχνότητα f_{ij} κάθε ζεύγους (x_i, y_j) εἶναι ἴση μέ τόν ἀριθμό πού βρίσκομε ἄν πολλαπλασιάσομε τίς συχνότητες $f_{i\cdot}$ καί $f_{\cdot j}$ τῶν τιμῶν x_i καί y_j στίς περιθωριακές κατανομές καί διαιρέσομε μέ τό πλῆθος ὅλων τῶν ἀτόμων τοῦ πληθυσμοῦ μας.***

Ὁ ἀριθμός τοῦ δευτέρου μέλους τῆς ἰσότητας (6.16) λέγεται ***θεωρητική συχνότητα*** τοῦ ζεύγους (x_i, y_j) καί συνεπῶς μποροῦμε νά ποῦμε πιό σύντομα ὅτι ***δύο***

μεταβλητές X καί Y εἶναι ἀνεξάρτητες, ὅταν ἡ συχνότητα κάθε ζεύγους (x_i, y_j) εἶναι ἴση μέ τή θεωρητική συχνότητά του. Αὐτό βέβαια δέν συμβαίνει στήν πράξη καί μᾶς ἀρκεῖ ἁπλῶς οἱ διαφορές $f_{ij} - \frac{f_{i\cdot} f_{\cdot j}}{\nu}$ νά εἶναι ἀρκετά μικροί (θετικοί ἤ ἀρνητικοί) ἀριθμοί. Γιά νά βρίσκομε εὔκολα ὅλες αὐτές τίς διαφορές, σχηματίζομε καί ***πίνακα τῶν θεωρητικῶν συχνοτήτων*** (Πίνακας 6.14.2), δηλαδή σχηματίζομε ἕνα πίνακα ὅμοιο μέ αὐτόν πού μᾶς δόθηκε, ὁ ὁποῖος ἔχει στίς θέσεις τῶν συχνοτήτων f_{ij} τίς ἀντίστοιχες θεωρητικές συχνότητες, τίς ὁποῖες θά σημειώνομε μέ θ_{ij}.

ΠΙΝΑΚΑΣ 6.14.2.

Θεωρητικῶν συχνοτήτων θ_{ij}.

$x_i \backslash y_j$	y_1	y_2	...	y_j	...	y_λ
x_1	θ_{11}	θ_{12}	...	θ_{1j}	...	$\theta_{1\lambda}$
x_2	θ_{21}	θ_{22}	...	θ_{2j}	...	$\theta_{2\lambda}$
.	.	.		.		
.	.	.		.	...	
.	.	.		.		
x_i	θ_{i1}	θ_{i2}	...	θ_{ij}	...	$\theta_{i\lambda}$
.	.	.		.		.
.	.	.		.	...	.
.	.	.		.		.
x_κ	$\theta_{\kappa 1}$	$\theta_{\kappa 2}$	...	$\theta_{\kappa j}$	...	$\theta_{\kappa\lambda}$

$$\theta_{ij} = \frac{f_{i\cdot} \cdot f_{\cdot j}}{\nu}$$

Ὡς μέτρο τοῦ βαθμοῦ ἀνεξαρτησίας τῶν δύο μεταβλητῶν, παίρνομε τόν ἀριθμό:

$$C = \frac{1}{\nu(q-1)} \left[\frac{(f_{11} - \theta_{11})^2}{\theta_{11}} + \frac{(f_{12} - \theta_{12})^2}{\theta_{12}} + \ldots + \frac{(f_{\kappa\lambda} - \theta_{\kappa\lambda})^2}{\theta_{\kappa\lambda}} \right] =$$

$$= \frac{1}{\nu(q-1)} \sum_i \sum_j \frac{(f_{ij} - \theta_{ij})^2}{\theta_{ij}}$$

ὅπου: q εἶναι ὁ μικρότερος ἀπό τούς ἀριθμούς κ καί λ πού παριστάνουν τίς γραμμές καί τίς στῆλες τοῦ πίνακα συχνοτήτων. Τό C λέγεται ***δείκτης μέσης συμπτώσεως*** καί, ἄν εἶναι ἀριθμός πού τείνει πρός τό μηδέν, δεχόμασθε ὅτι οἱ μεταβλητές X καί Y εἶναι ἀνεξάρτητες.

Παράδειγμα.

Ὁ Πίνακας 6.14.3 δείχνει τήν κατανομή 100 ἀνθρώπων ὡς πρός τό χρῶμα τῶν ματιῶν τους (X) καί τό χρῶμα τῶν μαλλιῶν τους (Y).

Νά ἐξετασθεῖ ἄν ὑπάρχει ἀλληλοεξάρτηση τῶν δύο ποιοτικῶν μεταβλητῶν X καί Y.

Λύση:

Ἀπό τόν παραπάνω πίνακα σχηματίζομε τόν Πίνακα 6.14.4 θεωρητικῶν συχνο-

ΠΙΝΑΚΑΣ 6.14.3.

X \ Y	Μαῦρα	Ξανθά	Ἄθροισμα
Μαῦρα	35	17	52
Γαλανά	20	28	48
Ἄθροισμα	55	45	100

τήτων, ὁ ὁποῖος θά ἔχει στοιχεῖα:

$$\theta_{11} = \frac{52 \times 55}{100} = 29, \qquad \theta_{12} = \frac{52 \times 45}{100} = 23{,}4$$

$$\theta_{21} = \frac{48 \times 55}{100} = 26, \qquad \theta_{22} = \frac{48 \times 45}{100} = 22$$

ΠΙΝΑΚΑΣ 6.14.4.

X \ Y	Μαῦρα	Ξανθά
Μαῦρα	29	23
Γαλανά	26	22

Στή συνέχεια, συνθέτομε τόν Πίνακα 6.14.5 γιά τόν ὑπολογισμό τῶν διαφορῶν $f_{ij} - \theta_{ij}$ καί τοῦ ἀθροίσματος τῶν πηλίκων $\frac{(f_{ij} - \theta_{ij})^2}{\theta_{ij}}$.

ΠΙΝΑΚΑΣ 6.14.5.

f_{ij}	θ_{ij}	$f_{ij} - \theta_{ij}$	$\frac{(f_{ij} - \theta_{ij})^2}{\theta_{ij}}$
35	29	6	1,24
17	23	−6	1,56
20	26	−6	1,38
28	22	6	1,64
100	100		5,82

Ἀπό τόν πίνακα αὐτό βρίσκομε ὅτι ὁ δείκτης μέσης συμπτώσεως εἶναι:

$$C = \frac{5{,}82}{100 \times (2-1)} = 0{,}0582$$

καί, ἐπειδή τό C εἶναι πολύ μικρός ἀριθμός, καταλαβαίνομε ὅτι τό χρῶμα τῶν μαλλιῶν δέν ἐξαρτᾶται ἀπό τό χρῶμα τῶν ματιῶν.

6.15 Ἀσκήσεις.

1. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν X καί Y παρατηρήθηκαν τά παρακάτω ζεύγη τιμῶν:

x_i	y_i
1	3
2	5
3	7
4	9
6	10

Νά προσαρμοσθεῖ στά παραπάνω δεδομένα μέ τή μέθοδο τῶν ἐλαχίστων τετραγώνων ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$.

2. Στόν παρακάτω πίνακα νά βρεθεῖ ἡ εὐθεία παλινδρομήσεως (εὐθεία ἐλαχίστων τετραγώνων) $y = \alpha + \beta x$:

x_i	1	2	4	6	7
y_i	12	15	8	4	2

3. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν X καί Y, μετά τούς ἀπαιτούμενους ἀριθμητικούς ὑπολογισμούς, λάβαμε τά παρακάτω δεδομένα:

$\sum_i x_i = 56$, $\sum_i y_i = 40$, $\sum_i x_i y_i = 364$, $\sum_i x_i^2 = 524$, $\sum_i y_i^2 = 256$, $\nu = 8$.

Νά ὑπολογισθεῖ ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$.

4. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν X καί Y λάβαμε τά παρακάτω δεδομένα:

$\sum_i x_i = 50$, $\sum_i y_i = 732$, $\sum_i x_i^2 = 322$, $\sum_i y_i^2 = 67214$, $\sum_i x_i y_i = 4640$, $\nu = 10$.

Νά ὑπολογισθεῖ ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$.

5. Στόν παρακάτω πίνακα δίνεται τό μηνιαῖο εἰσόδημα (σέ χιλιάδες δρχ.) καί ἡ ἀντίστοιχη ζήτηση ἑνός ἀγαθοῦ σέ kg.

x_i	y_i
8	8
10	10
14	11
20	14
25	16

Νά ὑπολογισθεῖ ἡ συνδιακύμανση τῶν μεταβλητῶν X καί Y.

6. Στόν παρακάτω πίνακα δίνεται ἡ βαθμολογία 7 μαθητῶν στά Μαθηματικά (X) καί στή Φυσική (Y).

x_i	y_i
14	12
18	16
15	16
12	14
18	19
10	9
11	10

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως.

7. Στόν παρακάτω πίνακα δίνεται ἡ ἡλικία (X) καί τό βάρος (Y), 9 μαθητῶν.

ἡλικία x_i	βάρος y_i
4	16
5	18
5	20
6	21
7	24
8	23
8	25
9	27
10	30

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως.

8. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν X καί Y παρατηρήθηκαν τά παρακάτω ζεύγη τιμῶν:

x_i	14	16	18	20	22	26
y_i	9	14	17	21	25	30

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως.

9. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν X καί Y προέκυψαν τά παρακάτω ζεύγη τιμῶν:

x_i	y_i
1	1
3	2
4	4
6	4
8	5
9	7
11	8
14	9

Νά ὑπολογισθεῖ: α) ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$ καί β) ὁ συντελεστής συσχετίσεως.

10. Ὁ παρακάτω πίνακας δείχνει τή βαθμολογία 10 σπουδαστῶν σύμφωνα μέ τήν ἀπόδοσή τους στά ἐργαστήρια καί σέ γραπτή φροντιστηριακή ἄσκηση:

Βαθμός ἐργαστηρίου (x_i)	8	3	9	2	7	10	4	6	1	5
Βαθμός φροντ. ἀσκήσεως (y_i)	9	5	10	1	8	7	3	4	2	6

Νά ὑπολογισθεῖ ἡ εὐθεία ἐλαχίστων τετραγώνων $y = \alpha + \beta x$ καί ἡ συνδιακύμανση.

11. Δίδονται τά παρακάτω δεδομένα:

x_i	y_i
8	7
3	0
5	10

Ζητεῖται: α) Ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$ καί ὁ συντελεστής συσχετίσεως. β) Ποιά ἡ τιμή τῆς y ὅταν $x = 2{,}4$. γ) Νά ὑπολογισθεῖ ἡ συνδιακύμανση.

12. Ὁ παρακάτω πίνακας μᾶς δίνει τό κεφάλαιο (X), πού ἔχει χρησιμοποιήσει μία ἑταιρεία σέ 9 χρόνια, καί τό ἀντίστοιχο κέρδος (Y) (καί τά δύο σέ ἑκατομμύρια δρχ.).

Κεφάλαιο (x_i)	10	20	30	40	60	70	80	90	100
Κέρδη (y_i)	2	4	8	10	15	14	20	22	30

1) Νά ὑπολογισθεῖ: α) Ὁ συντελεστής συσχετίσεως. β) Ἡ εὐθεία ἐλαχίστων τετραγώνων $y = \alpha + \beta x$. 2) Νά γίνει πρόβλεψη τῆς τιμῆς y ὅταν $x = 25$.

13. Ὀκτώ μαθητές ἔλαβαν σέ δύο tests τούς παρακάτω βαθμούς:

Test (x)	10	9	8	8	7	6	6	5
Test (y_i)	8	10	7	9	6	7	4	5

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως καί ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$.

14. Ὁ παρακάτω πίνακας δίνει τόν ἀριθμό τῶν καταδικαστικῶν ἀποφάσεων γιά ἐγκληματικές πράξεις (σέ χιλιάδες) καί τόν ἀριθμό τῶν ἀνέργων (σέ ἑκατομμύρια) γιά τά ἔτη 1970 - 1976 στή Γαλλία.

Ἀποφάσεις (x_i)	7,8	8,2	7,8	7,2	7,4	8,3	8,8	9,5
Ἄνεργοι (y_i)	1,2	1,3	1,4	1,2	1,3	2,5	2,6	2,2

Νά ἐξετασθεῖ ἄν ὑπάρχει ἀλληλοεξάρτηση μεταξύ ἀνεργίας καί ἐγκληματικότητας.

15. Ὁ συντελεστής συσχετίσεως μεταξύ δύο μεταβλητῶν X καί Y εἶναι 0,60. Ἄν $s_x = 1{,}50$, $s_y = 2$, $\bar{x} = 10$ καί $\bar{y} = 20$, νά ὑπολογισθεῖ ἡ εὐθεία ἐλαχίστων τετραγώνων $y = \alpha + \beta x$.

16. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν X καί Y, μετά τούς ἀπαιτούμενους ἀριθμητικούς ὑπολογισμούς λάβαμε τά παρακάτω χρήσιμα δεδομένα:

$\sum_i x_i = 50$, $\sum_i y_i = 70$, $\sum_i x_i y_i = 460$, $\sum_i x_i^2 = 360$, $\sum_i y_i^2 = 612$, $\nu = 10$.

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως.

17. Νά κατασκευασθεῖ τό διάγραμμα διασπορᾶς στά δεδομένα τοῦ παρακάτω πίνακα καί νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως.

x_i	1	2	3	5	6	7	8	9	10
y_i	2,2	2,8	4,2	5	5,4	5,6	5	6,3	6

18. Δίνονται τά παρακάτω ζεύγη τιμῶν:

x_i	0	1	2	3
y_i	4	7	8	10

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως καί ἡ συνδιακύμανση.

19. Στούς παρακάτω πίνακες διπλῆς εἰσόδου νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως:

α)

$x_i \backslash y_j$	1	2	3	$f_{i\cdot}$
0	–	1	2	3
2	2	3	4	9
$f_{\cdot j}$	2	4	6	12

β)

$x_i \backslash y_j$	1	2	$f_{i\cdot}$
1	2	4	6
2	1	3	4
$f_{\cdot j}$	3	7	10

γ)

$x_i \backslash y_j$	1	2	3
12	10	0	0
14	0	20	0
16	0	0	10

δ)

$x_i \backslash y_j$	4	5	6
3	1	1	1
2	1	0	1
1	1	1	1

ε)

$x_i \backslash y_j$	4	5	6
3	0	1	10
2	1	5	1
1	10	1	0

ζ)

$x_i \backslash y_j$	0	1	$f_{i\cdot}$
2	1	3	4
1	4	2	6
$f_{\cdot j}$	5	5	10

η)

$x_i \backslash y_j$	0	1	$f_{i\cdot}$
0	5	3	8
1	10	2	12
$f_{\cdot j}$	15	5	20

20. Στόν παρακάτω πίνακα δίνεται τό ὕψος (y_j) καί τό βάρος (x_i) 42 ἀτόμων:

Βάρος x_i \ Ὕψος y_j		Κεντρικές τιμές τάξεων σέ cm					$f_{i\cdot}$
		160	165	170	175	180	
Κεντρικές	65	3	2	1	–	–	6
τιμές τάξεων	70	1	3	5	1	–	10
σέ kg*	75	–	2	5	7	1	15
	80	–	–	4	5	2	11
	$f_{\cdot j}$	4	7	15	13	3	42

Νά ὑπολογισθεῖ ἡ συνδιακύμανση.

21. Ὁ παρακάτω πίνακας διπλῆς εἰσόδου ἀναφέρεται στή βαθμολογία 40 μαθητῶν στή Στατιστική καί στήν Ἔκθεση:

X \ Y	1 – 4	4 – 7	7 – 10	$f_{i\cdot}$
1 – 4	3	4	0	7
4 – 7	5	6	2	13
7 – 10	4	8	8	20
$f_{\cdot j}$	12	18	10	40

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως.

22. Δίνεται ὁ παρακάτω πίνακας διπλῆς εἰσόδου:

x_i \ y_j	1	2	3	4	$f_{i\cdot}$
0	2	1	3	4	10
1	2	3	3	1	9
2	1	2	3	1	7
3	1	2	1	–	4
$f_{\cdot j}$	6	8	10	6	30

Ζητεῖται: 1) Νά ὑπολογισθοῦν οἱ περιθωριακοί μέσοι τῶν μεταβλητῶν X καί Y καί ἡ διασπορά V (X)
ἤ s_x^2 τῆς μεταβλητῆς X.

23. Δίνεται ὁ παρακάτω πίνακας διπλῆς εἰσόδου:

x_i \ y_j	0	1	2	Ἄθροισμα
0	.	1	.	4
1	1	.	3	.
Ἄθροισμα	2	.	.	10

Νά βρεθοῦν πρῶτα οἱ συχνότητες πού ἔχουν σβησθεῖ καί μετά νά ὑπολογισθεῖ ὁ συντελεστής μεταβλητικότητας τῆς μεταβλητῆς X.

24. Κατά τή μελέτη τῆς ἀλληλοεξαρτήσεως τῶν μεταβλητῶν x καί y προέκυψε ὁ παρακάτω πίνακας διπλῆς εἰσόδου:

x_i \ y_j	1	3	5	7	9	$f_{i\cdot}$
10	60	.	30	.	15	150
20	.	10	20	.	.	100
30	20	5	10	10	.	50
40	.	20	40	40	.	200
$f_{\cdot j}$	200	50	100	100	50	500

Ἄν ὑποθέσομε ὅτι οἱ μεταβλητές X καί Y εἶναι ἀνεξάρτητες μεταξύ τους, νά βρεθοῦν πρῶτα οἱ συχνότητες πού ἔχουν σβησθεῖ καί μετά νά ὑπολογισθεῖ ὁ περιθωριακός μέσος τῆς μεταβλητῆς X.

25. Δεῖγμα 95 ἀτόμων κατανέμεται ὡς πρός τό χρῶμα τῶν ματιῶν καί τῶν μαλλιῶν τους ὡς ἑξῆς:

Χρῶμα μαλλιῶν (y) / Χρῶμα ματιῶν (x)	Ξανθά	Καστανά	Ἄθροισμα
Γαλανά	32	12	44
Καστανά	14	22	36
Ἄλλο χρῶμα	6	9	15
Ἄθροισμα	52	43	95

Νά ἐξετασθεῖ ἄν ὑπάρχει ἀλληλεξάρτηση τῶν δύο ποιοτικῶν μεταβλητῶν Χ καί Υ.

26. Δίνεται παρακάτω ἡ κατανομή συχνοτήτων 100 ἀνδρῶν ὡς πρός τό ὕψος Χ (σέ ἑκατοστά) καί τό βάρος Υ (σέ κιλά):

Τάξεις (Y)		50 - 60	60 - 70	70 - 80	80 - 90	
Τάξεις (X)	y_i / x_i	55	65	75	85	f_i
150 - 160	155	3	10	1	0	14
160 - 170	165	5	13	12	4	34
170 - 180	175	2	8	16	5	31
180 - 190	185	0	4	7	10	21
	$f_{.j}$	10	35	36	19	100

Νά ὑπολογισθεῖ ὁ συντελεστής συσχετίσεως καί ὁ ἀριθμητικός μέσος τῆς μεταβλητῆς Χ.

27. Στόν παρακάτω πίνακα διπλῆς εἰσόδου δίνεται ὁ ἀπαιτούμενος χρόνος (Υ) σέ δευτερόλεπτα γιά τή λύση ἑνός προβλήματος ἀπό ἕνα μαθητή καί ὁ συντελεστής ἐξυπνάδας (Χ) τοῦ μαθητῆ:

Y / X	50 - 60	40 - 50	30 - 40	20 - 30
60 - 80	1	2	—	—
80 - 100	2	4	5	3
100 - 120	—	2	3	3
120 - 140	—	—	1	2

Ζητεῖται: ὁ συντελεστής συσχετίσεως καί ἡ συνδιακύμανση Cov (X, Y).

28. Ἀπό τά δεδομένα τοῦ παραπάνω πίνακα διπλῆς εἰσόδου:

Y / X	1	2	3	4	$f_{i.}$
0	2	1	3	4	10
1	2	3	3	1	9
2	1	2	3	1	7
3	1	2	1	0	4
$f_{.j}$					30

νά ὑπολογίσετε: α) Τούς μέσους ἀριθμητικούς τῶν περιθωριακῶν κατανομῶν, δηλαδή $\overline{x}$ καί $\overline{y}$. β) Τή συνδιακύμανση COV (x, y) τῶν μεταβλητῶν X καί Y.

29. Σέ ἕνα διμεταβλητό στατιστικό πληθυσμό (x_i y_i), $i = 1, 2, ..., \nu$ ἔχει προσαρμοσθεῖ (μέ τή μέθοδο τῶν ἐλαχίστων τετραγώνων) ἡ εὐθεία παλινδρομήσεως $y = \alpha + \beta x$.
Ποιές εἶναι οἱ τιμές τῶν συντελεστῶν α καί β, ὅταν: $\overline{x} = 160$, $\overline{y} = 70$, $r = 0{,}80$, συντελεστής μεταβλητικότητας τῆς X, C V(X) = 0,10 καί συντελεστής μεταβλητικότητας τῆς Y, CY(y) = 0,20.

30. Ἄν ὑποθέσομε ὅτι οἱ μεταβλητές X καί Y εἶναι ἀνεξάρτητες μεταξύ τους, νά ὑπολογισθοῦν οἱ συχνότητες πού ἔχουν σβησθεῖ στόν παρακάτω πίνακα καί μετά νά ὑπολογισθεῖ ὁ ἀριθμητικός μέσος καί ἡ διακύμανση τῆς μεταβλητῆς X.

x_i \ y_j	1	2	3	4	f_i
1	.	12	.	.	40
2	.	.	20	.	.
3	.	.	.	24	60
$f_{.j}$	.	.	.	.	200

ΚΕΦΑΛΑΙΟ ΕΒΔΟΜΟ

ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ

7.1 Γενικά.

Μέ τόν ὅρο ***χρονολογική σειρά*** ἐννοοῦμε ἕνα πλῆθος παρατηρήσεων, οἱ ὁποῖες λαμβάνονται σέ ὁρισμένες χρονικές στιγμές πού ἀπέχουν μεταξύ τους κατά ἴσα διαστήματα.

Ἄν σημειώσομε μέ y_i τήν τιμή τῆς παρατηρήσεως πού λαμβάνεται κατά τή χρονική στιγμή x_i, τότε μία χρονολογική σειρά μέ ν παρατηρήσεις θά ἀποτελεῖται ἀπό ν ζεύγη τῆς μορφῆς:

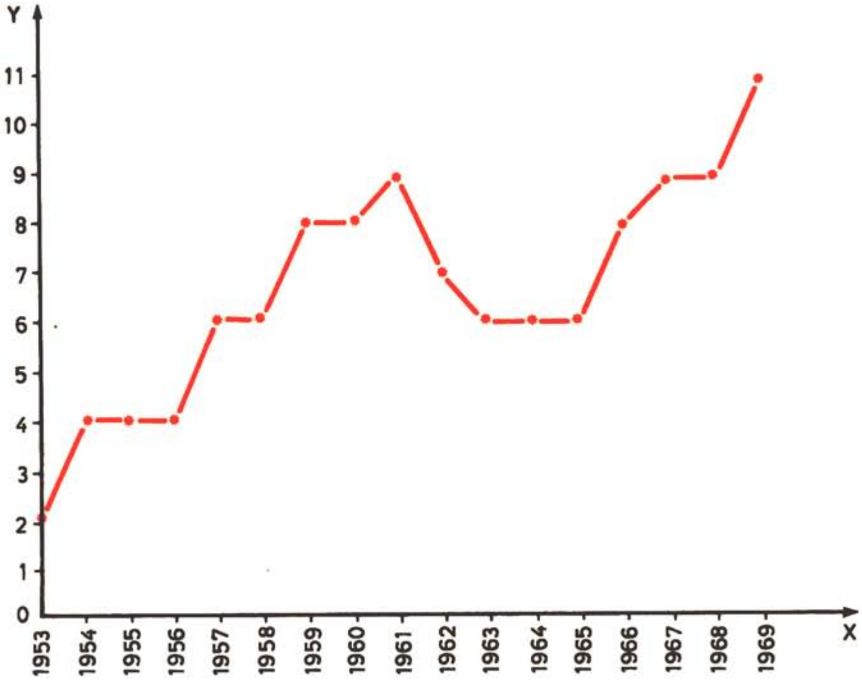
$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_i, y_i), \ldots, (x_\nu, y_\nu)$$

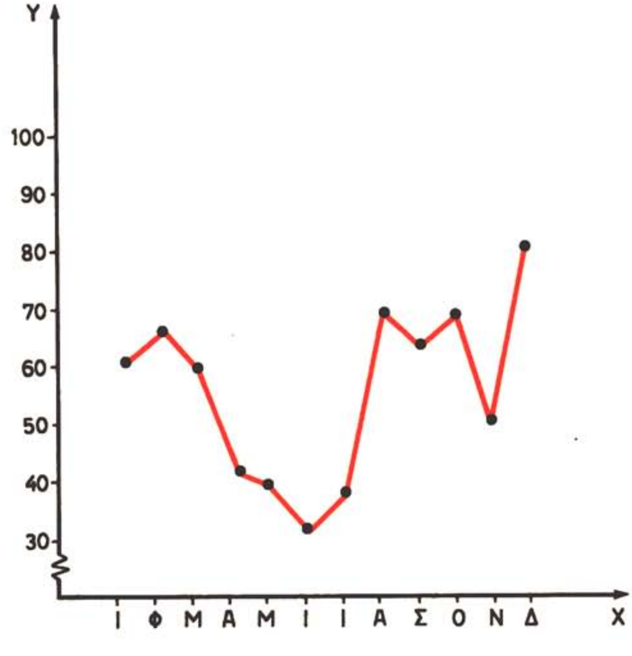
Βλέπομε λοιπόν πάλι ὅτι ἔχομε οὐσιαστικά ν παρατηρήσεις ἑνός ζεύγους μεταβλητῶν (X, Y), ἀπό τίς ὁποῖες ἡ πρώτη X παριστάνει πάντα χρόνο καί οἱ τιμές της καθορίζονται μέ ἀκρίβεια, ἐνῶ ἡ δεύτερη Y παριστάνει ἕνα ὁποιοδήποτε μέγεθος πού οἱ τιμές του προκύπτουν ἀπό μέτρηση ἤ καταγραφή.

Ἡ τεθλασμένη γραμμή, ἡ ὁποία συνδέει τά διαδοχικά σημεῖα (x_1, y_1), (x_2, y_2), ..., (x_ν, y_ν) παριστάνει τό διάγραμμα συμμεταβολῆς τοῦ ζεύγους (X, Y), τό ὁποῖο θά λέγεται τώρα ***χρονοδιάγραμμα*** τῆς Y.

Στό σχῆμα 7.1 παριστάνονται δύο τέτοια χρονοδιαγράμματα.

Οἱ χρονολογικές σειρές ἔχουν μεγάλη σημασία στή μελέτη οἰκονομικῶν φαινομένων, γιατί μέ τήν ἀνάλυσή τους μποροῦμε νά διατυπώσομε πολλές φορές προβλέψεις γιά μελλοντικές ἐξελίξεις ἤ καταστάσεις καί νά προχωρήσομε ἔτσι στήν κατάστρωση διαφόρων προγραμμάτων.


Y
11
10
9
8
7
6
5
4
3
2
1
0
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
X


Y
100
90
80
70
60
50
40
30
Ι
Φ
Μ
Α
Μ
Ι
Ι
Α
Σ
Ο
Ν
Δ
Χ

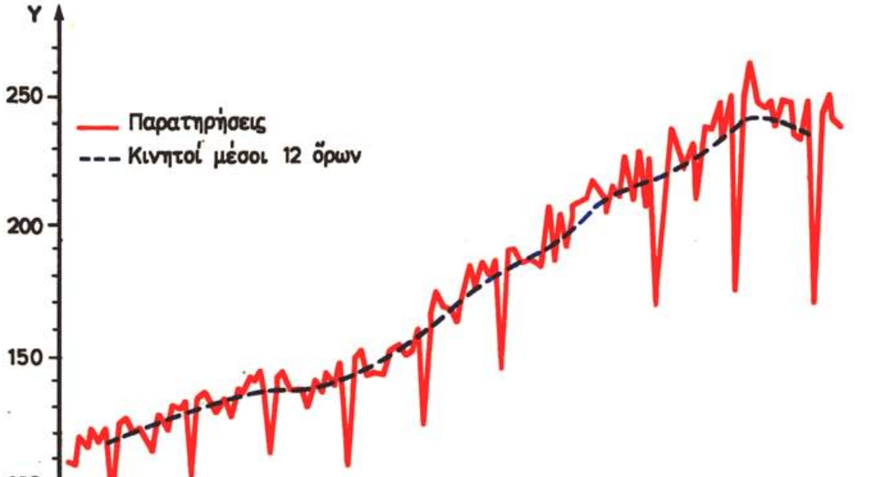
Σχ. 7.1.

7.2 Οἱ μεταβολές μιᾶς χρονολογικῆς σειρᾶς.

Σέ μία χρονολογική σειρά μᾶς ἐνδιαφέρουν συνήθως ὄχι μόνο οἱ τιμές $y_1, y_2, \ldots, y_\nu$ τῆς μεταβλητῆς y, ἀλλά καί οἱ διαφορές τῶν τιμῶν αὐτῶν, δηλαδή οἱ ἀριθμοί $y_2 - y_1, y_3 - y_2, \ldots, y_\nu - y_{\nu-1}$, πού προκύπτουν, ἄν ἀπό κάθε τιμή ἀφαιρέσομε τήν προηγούμενή της. Εἶναι φανερό ὅτι, ἐάν μία διαφορά $y_i - y_{i-1}$ εἶναι θετικός (ἤ ἀρνητικός) ἀριθμός, τή χρονική στιγμή x_i, θά ἔχομε αὔξηση (ἤ μείωση) τῆς τιμῆς τῆς Υ.

Γενικά τώρα, κάθε ἀπότομη αὔξηση, μείωση ἤ κάθε οὐσιαστική ἐναλλαγή ἀπό αὔξηση σέ μείωση καί ἀντιστρόφως, χαρακτηρίζεται ὡς μεταβολή τῆς χρονολογικῆς σειρᾶς. Οἱ μεταβολές μιᾶς χρονολογικῆς σειρᾶς διακρίνονται κυρίως σέ:

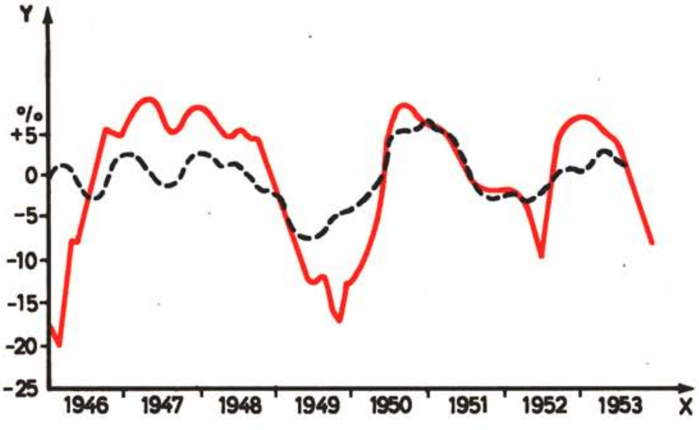
α) ***Περιοδικές***, δηλαδή μεταβολές πού ἐπαναλαμβάνονται σέ ὁρισμένα χρονικά διαστήματα (συνήθως μικρῆς διάρκειας). Τέτοιες περιοδικές μεταβολές μέ περίοδο ἑνός ἔτους εἶναι οἱ ***ἐποχιακές μεταβολές*** πού ἐπαναλαμβάνονται τήν ἴδια ἐποχή κάθε ἔτους (π.χ. μείωση τῶν τιμῶν τῶν λαχανικῶν τούς θερινούς μῆνες, αὔξηση τῆς ζητήσεως τῶν ἀγαθῶν τίς ἐποχές πού χορηγοῦνται δῶρα στούς ὑπαλλήλους, κλπ). Τό χρονοδιάγραμμα τοῦ σχήματος 7.2α δείχνει ἀκριβῶς μιά χρονολογική σειρά μέ ἐποχιακές μεταβολές.



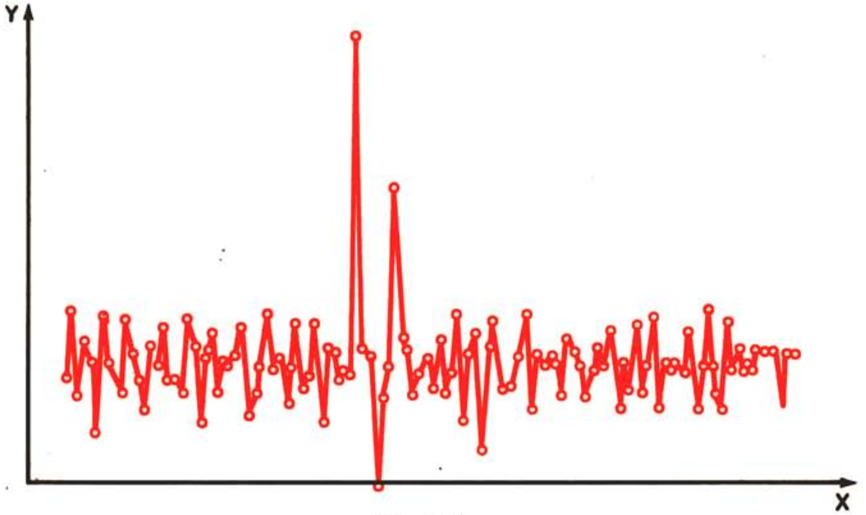
Σχ. 7.2α.

β) ***Κυκλικές***, δηλαδή μεταβολές πού προκύπτουν, ὅταν μετά ἀπό κάθε χρονική περίοδο αὐξήσεων ἀκολουθεῖ μιά χρονική περίοδος μειώσεων ἤ ἀντιστρόφως. Οἱ κυκλικές μεταβολές, ὅπως φαίνεται καί στό σχῆμα 7.2β παρουσιάζουν μιά περιοδικότητα (μέ περίοδο μεγαλύτερη τοῦ ἔτους) ἡ ὁποία ὅμως συνήθως δέν εἶναι κανονική.

γ) ***Ἀκανόνιστες***, δηλαδή μεταβολές πού ὀφείλονται σέ συμπτωματικά ἤ ἀπρό-



Σχ. 7.2β.



Σχ. 7.2γ.

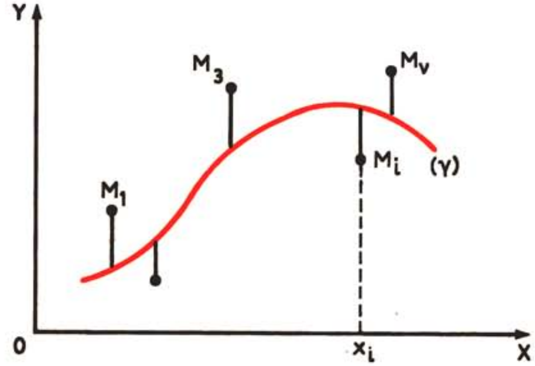
βλεπτα γεγονότα (σεισμούς, ἀπεργίες, ἐπιδημίες, πολέμους, κ.λ.π) καί γενικά σέ τυχαῖα γεγονότα, τά ὁποῖα εἶναι ἀδύνατο νά προβλέφθοῦν (σχ. 7.2γ).

7.3 Ἡ εὐθεία τάσεως μιᾶς χρονολογικῆς σειρᾶς.

Τό χρονοδιάγραμμα μιᾶς χρονολογικῆς σειρᾶς μᾶς δίνει ὄχι μόνο τίς τιμές μιᾶς μεταβλητῆς Y στίς διάφορες χρονικές στιγμές, ἀλλά καί τόν τρόπο μεταβολῆς τῆς Y στό θεωρούμενο χρονικό διάστημα (σχ. 7.3α). Πραγματικά, ἀφοῦ ἡ Y ἔχει γενικά στοχαστική ἐξάρτηση ἀπό τό χρόνο X καί τό χρονοδιάγραμμα εἶναι τό διάγραμμα συμμεταβολῆς τους, ὁ τρόπος μεταβολῆς τῆς Y θά δίνεται, ὅπως γνωρίζομε, ἀπό μία καμπύλη γ πού διέρχεται ὅσο τό δυνατό πλησιέστερα ἀπό τά σημεῖα:

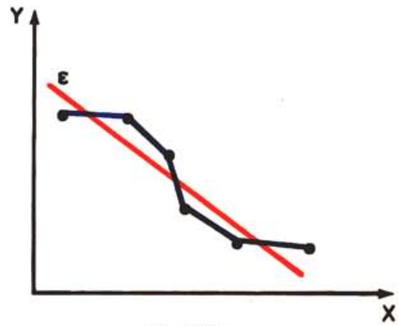
$$M_1\ (x_1, y_1),\ M_2\ (x_2, y_2),\ \ldots,\ M_\nu\ (x_\nu\ y_\nu),$$

τά όποῖα εἶναι κορυφές τοῦ χρονοδιάγραμματος. Ἡ καμπύλη αὐτή λέγεται τώρα ***καμπύλη τάσεως*** τῆς χρονολογικῆς σειρᾶς.



Σχ. 7.3α.

Σχ. 7.3β.



Σχ. 7.3γ.

Μιά μερική περίπτωση, εἶναι ἐκείνη κατά τήν ὁποία μποροῦμε νά πάρομε ὡς καμπύλη τάσεως μία εὐθεία ε καί τότε θά μιλᾶμε γιά ***εὐθεία τάσεως*** τῆς χρονολογικῆς σειρᾶς. Ἡ περίπτωση αὐτή, στήν ὁποία καί θά περιορισθοῦμε, παρουσιάζεται ὅταν οἱ διαδοχικές τιμές τῆς Υ ἔχουν μιά μόνιμη τάση αὐξήσεως (σχ. 7.3β) ἤ μειώσεως (σχ. 7.3γ).

Ἕνας ἄλλος πρακτικός τρόπος κατασκευῆς τῆς εὐθείας τάσεως μιᾶς χρονολογικῆς σειρᾶς εἶναι ὁ ἑξῆς: Χωρίζομε τίς παρατηρήσεις μας σέ δύο ἴσες ὁμάδες καί ὑπολογίζομε τούς ἀριθμητικούς μέσους τῶν τιμῶν τῆς Υ σέ κάθε ὁμάδα. Μετά, βρίσκομε τά δύο σημεῖα, πού ἔχουν τετμημένες τά μέσα τῶν χρονικῶν διαστημάτων τῶν δύο ὁμάδων καί τεταγμένες τούς δύο μέσους ἀριθμητικούς πού βρήκαμε, καί χαράσσομε τήν εὐθεία πού διέρχεται ἀπό τά σημεῖα αὐτά. Ὁ τρόπος αὐτός λέγεται ***μέθοδος τῶν δύο μέσων σημείων.***

Παράδειγμα.

Στόν Πίνακα 7.3.1 δίνεται ἡ παραγωγή ἑνός ἐργοστασίου, σέ χιλιάδες τόννους κατά τή χρονική περίοδο 1965 — 1970.

Νά κατασκευασθεῖ τό χρονοδιάγραμμα καί νά βρεθεῖ ἡ εὐθεία τάσεως.

ΠΙΝΑΚΑΣ 7.3.1.

Έτος	Παραγωγή (y_i)
1965	5
1966	8
1967	12
1968	15
1969	20
1970	23

Χωρίζομε τή χρονολογική σειρά σέ δύο ὁμάδες, ἀπό τίς ὁποῖες ἡ πρώτη περιέχει τίς τρεῖς πρῶτες παρατηρήσεις καί ἡ δεύτερη τίς τρεῖς τελευταῖες, καί προσδιορίζομε τούς μέσους ἀριθμητικούς τῶν τιμῶν τῆς y, ὅπως φαίνεται στόν πίνακα τοι σχήματος 7.3δ. Μετά, παίρνομε στό χρονοδιάγραμμα δύο σημεῖα Α καί Β πού προ βάλλονται στά μέσα τῶν χρονικῶν διαστημάτων τῶν δύο ὁμάδων (δηλαδή στά ἔτ 1966 καί 1969) καί ἔχουν τεταγμένες τούς μέσους 8,3 καί 19,3 πού βρήκαμε. Ἡ εὐθεία πού περνᾶ ἀπό τά σημεῖα αὐτά Α καί Β θά εἶναι ἡ ***εὐθεία τάσεως τῆς χρο νολογικῆς σειρᾶς*** (σχ. 7.3δ).

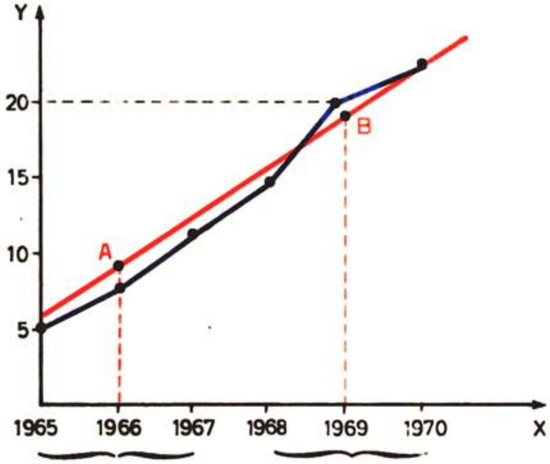
Έτος	x_i		y_i	
1965	0		5	
1966	1	3:3=1	8	$\dfrac{25}{3} = 8{,}3$
1967	2		12	
1968	3		15	
1969	4	12:3=4	20	$\dfrac{58}{3} = 19{,}3$
1970	5		23	

Θά μπορούσαμε ἀκόμη νά βροῦμε καί τήν ἐξίσωση τῆς ΑΒ, ἄν θεωρήσομε μί ἀντιστοιχία τῶν ἐτῶν τῆς χρονολογικῆς σειρᾶς μέ τίς ὑποδιαιρέσεις τοῦ ἄξονα C (πχ. τήν 1965 → 0, 1966 → 1, ..., 1970 → 5). Τότε ἡ ΑΒ, ἡ ὁποία διέρχεται ἀπό τ δύο σημεῖα Α ($x_1 = 1$, $y_1 = 8{,}3$) καί Β ($x_2 = 4$, $y_2 = 19{,}3$), θά ἔχει ἐξίσωση τῆς μορ φῆς:

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1},$$

δηλαδή:

$$\frac{x-1}{4-1}=\frac{y-8{,}3}{19{,}3-8{,}3} \Rightarrow y=4{,}6+3{,}7x$$



Σχ. 7.3δ.

Συνήθως, γιά τήν κατασκευή τῆς εὐθείας τάσεως δέν ἀκολουθοῦμε τούς πιό πάνω πρακτικούς τρόπους, ἀλλά ἐφαρμόζομε τή ***μέθοδο τῶν ἐλαχίστων τετραγώνων*** πού μάθαμε στό προηγούμενο κεφάλαιο. Ἔτσι, παίρνομε γιά εὐθεία τάσεως τήν εὐθεία ἐλαχίστων τετραγώνων τῶν ν σημείων:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_\nu, y_\nu),$$

ὅπου: τώρα τά $x_1, x_2, \ldots, x_\nu$ εἶναι οἱ ὑποδιαιρέσεις τοῦ ἄξονα Ox, στίς ὁποῖες ἀντιστοιχίσαμε τίς χρονικές στιγμές τῆς χρονολογικῆς σειρᾶς. Ἔτσι, ἡ εὐθεία τάσεως θά ἔχει ἐξίσωση $\hat{y}=\hat{\alpha}+\hat{\beta}x$, στήν ὁποία τά $\hat{\alpha}$ καί $\hat{\beta}$ εἶναι (παράγρ. 6.5) ἡ λύση τοῦ κανονικοῦ συστήματος:

$$\nu\alpha + (\sum_i x_i)\,\beta = (\sum_i y_i)$$
$$(\sum_i x_i)\,\alpha + (\sum_i x_i)\,\beta = (\sum_i x_i\, y_i)$$

Στήν περίπτωση τῶν χρονολογικῶν σειρῶν, εἶναι πάντα δυνατό νά κάνομε τέτοια ἀντιστοιχία τῶν χρονικῶν στιγμῶν της μέ τίς ὑποδιαιρέσεις τοῦ ἄξονα Ox, ὥστε νά ἔχομε $\sum x_i = 0$. Τότε, οἱ ἐξισώσεις τοῦ παραπάνω συστήματος τῶν κανονικῶν ἐξισώσεων γράφονται $\nu\alpha = (\sum_i y_i)$ καί $(\sum_i x_i^2)\,\beta = (\sum_i x_i\, y_i)$ καί συνεπῶς ἔχομε ἀμέσως:

$$\hat{\alpha}=\frac{\sum_i y_i}{\nu}\quad,\ \hat{\beta}=\frac{\sum_i x_i\, y_i}{\sum_i x_i^2}$$

Παράδειγμα 1ο.

Στόν Πίνακα 7.3.2 δίνεται σέ χιλιάδες ἀτόμων ὁ πληθυσμός τῆς περιφέρεια Ἀθηνῶν γιά τή χρονική περίοδο 1961 — 1965.

Νά βρεθεῖ ἡ ἐξίσωση τῆς εὐθείας τάσεως μέ τή μέθοδο τῶν ἐλαχίστων τετραγώνων.

ΠΙΝΑΚΑΣ 7.3.2.

Ἔτος	Πληθυσμός
1961	1885
1962	1938
1963	1981
1964	2050
1965	2088

Λύση:

Ἐπειδή ἔχομε περιττό ἀριθμό ἐτῶν (τιμῶν τῆς X), ἀντιστοιχίζομε τό μεσαῖο ἔτος στήν ὑποδιαίρεση 0 τοῦ ἄξονα Ox, δηλαδή ὁρίζομε τήν ἀντιστοιχία.

−2	−1	0	1	2
1961	1962	1963	1964	1965

ὁπότε ἔχομε $\sum_i x_i = 0$. Ἔτσι, ἄν κατασκευάσομε τόν πίνακα τοῦ σχήματος 7.3ε πού περιέχει τίς στῆλες x_i, x_i^2, $x_i\ y_i$, βρίσκομε:

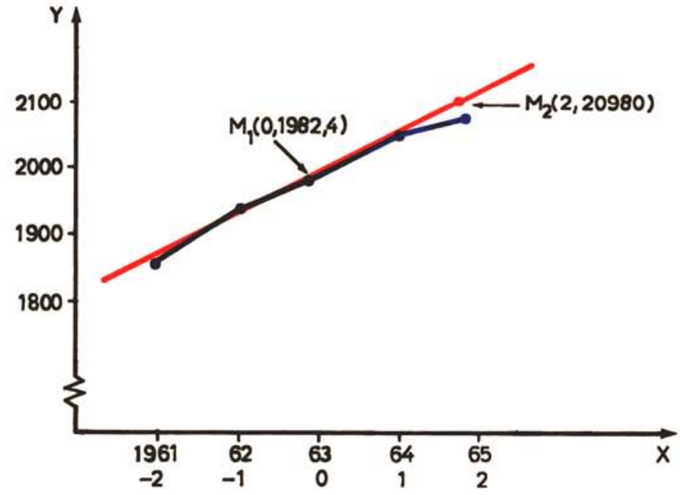
Ἔτος	y_i	x_i	$x_i\ y_i$	x_i^2
1961	1855	−2	−3710	4
1962	1938	−1	−1938	1
1963	1981	0	0	0
1964	2050	1	2050	1
1965	2088	2	4176	4
Ἄθροισμα	9912	$\sum_i x_i = 0$	578	10

$$\hat{\alpha} = \frac{\sum_i y_i}{\nu} = \frac{9912}{5} = 1982{,}4$$

$$\hat{\beta} = \frac{\sum_i x_i\ y_i}{\sum_i x_i^2} = \frac{578}{10} = 57{,}8$$

Ἐπομένως, ἡ εὐθεία τάσεως θά ἔχει ἐξίσωση:

$$\hat{y} = 1982{,}4 + (57{,}8)\,x.$$



Σχ. 7.3ε.

Παράδειγμα 2ο.

Ὁ Πίνακας 7.3.3 δίνει (σέ χιλιάδες λίτρα) τή βενζίνη πού πούλησε ἕνα πρατήριο κατά τά ἔτη 1970 — 1975:

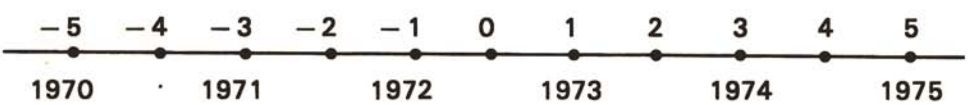
ΠΙΝΑΚΑΣ 7.3.3.

Ἔτος	1970	1971	1972	1973	1974	1975
Βενζίνη (Y)	289	305	319	339	368	390

Νά βρεθεῖ ἡ εὐθεία τάσεως καί νά ἐκτιμηθεῖ ἡ βενζίνη πού πούλησε τό πρατήριο κατά τό 1977.

Λύση:

Ἐπειδή ὁ ἀριθμός τῶν ἐτῶν (τιμῶν τῆς X) εἶναι ἄρτιος, ἀντιστοιχίζομε τά δύο μεσαῖα ἔτη στίς ὑποδιαιρέσεις — 1 καί 1 τοῦ ἄξονα Ox, δηλαδή ὁρίζομε τήν ἀντιστοιχία:



ὁπότε ἔχομε $\sum_i x_i = 0$. Στή συνέχεια, κατασκευάζομε τόν πίνακα ὑπολογισμῶν 7.3.4:

ΠΙΝΑΚΑΣ 7.3.4.

Έτος	y_i	x_i	$x_i\ y_i$	x_i^2
1970	289	−5	−1445	25
1971	305	−3	−915	9
1972	319	−1	−319	1
1973	339	1	339	1
1974	368	3	1104	9
1975	390	5	1950	25
Άθροισμα	2010	$\sum_i x_i = 0$	714	70

ἀπό τόν ὁποῖο βρίσκομε ἀμέσως:

$$\hat{\alpha} = \frac{\sum_i y_i}{\nu} = \frac{2010}{6} = 335,$$

$$\hat{\beta} = \frac{\sum_i x_i\ y_i}{\sum_i x_i} = \frac{714}{70} = 10{,}2.$$

Ἐπομένως, ἡ εὐθεία τάσεως θά ἔχει ἐξίσωση:

$$\hat{y} = 335 + (10{,}2)x.$$

Ἐπειδή τώρα τό ἔτος 1977 ἀντιστοιχίζεται στήν ὑποδιαίρεση 9 τοῦ ἄξονα Ox, ἡ πώληση βενζίνης κατά τό ἔτος αὐτό θά δίνεται (κατά προσέγγιση) ἀπό τήν τεταγμένη τοῦ σημείου τῆς εὐθείας τάσεως, τό ὁποῖο ἔχει τετμημένη $x = 9$ καί συνεπῶς θά εἶναι:

$$\hat{y} = 335 + (10{,}2) \times 9 = 426{,}8 \text{ χιλιάδες λίτρα.}$$

7.4 Ἐποχιακές μεταβολές. Δεῖκτες ἐποχικότητας.

Ἄς περιορισθοῦμε τώρα σέ χρονολογικές σειρές πού παρουσιάζουν ἐποχιακές μεταβολές, δηλαδή μεταβολές πού ἐπαναλαμβάνονται τήν ἴδια ἐποχή κάθε ἔτους καί ὀφείλονται, ὅπως εἴπαμε, σέ ἐποχιακούς λόγους. Ειδικότερα, ἄν σέ μία χρονολογική σειρά οἱ τιμές τῆς μεταβλητῆς Y ἀναφέρονται σέ διάφορες ἐποχές (π.χ. στούς μῆνες) κάθε ἔτους σέ μιά σειρά ἐτῶν, θά ἐξετάσομε πῶς διαπιστώνεται ἡ ὕπαρξη ἐποχιακῶν μεταβολῶν καί πῶς μποροῦμε νά «μετρήσομε» τίς μεταβολές αὐτές.

Γιά νά διαπιστώσομε τήν ὕπαρξη ἐποχιακῶν μεταβολῶν σέ μία χρονολογική σειρά οἱ τιμές τῆς ὁποίας ἀναφέρονται σέ μῆνες, ἐκφράζομε πρῶτα κάθε τιμή τῆς

Υ σέ ποσοστό τῆς μέσης μηνιαίας τιμῆς τοῦ ἴδιου ἔτους. Ἔτσι, ἄν οἱ τιμές τῆς μεταβλητῆς Υ γιά ἕνα ὁρισμένο ἔτος εἶναι π.χ.:

ΙΑΝ.	ΦΕΒ.	ΜΑΡΤ.	ΑΠΡ.	ΜΑΪ.	ΙΟΥΝ.
109	108	123	116	129	119
ΙΟΥΛ.	ΑΥΓ.	ΣΕΠΤ.	ΟΚΤ.	ΝΟΕΜ.	ΔΕΚ.
125	102	126	127	125	124

βρίσκομε τό μέσο ὅρο $\overline{y}$ τῶν 12 αὐτῶν τιμῶν καί ἀντικαθιστοῦμε κάθε τιμή y_i ἀπό αὐτές μέ τό ποσοστό τοῦ $\overline{y}$ πού ἀντιπροσωπεύει, δηλαδή μέ τόν ἀριθμό $y_i/\overline{y}$. Ἔτσι, ἐπειδή ὁ μέσος ὅρος τῶν 12 τιμῶν εἶναι 118,83, ἀντικαθιστοῦμε κάθε μία ἀπό τίς παραπάνω τιμές τῆς Υ μέ αὐτές πού προκύπτουν ἄν τήν πολλαπλασιάσομε μέ 100/118,83 καί βρίσκομε τελικά:

ΙΑΝ.	ΦΕΒΡ.	ΜΑΡ.	ΑΠΡ.	ΜΑΪ.	ΙΟΥΝ.
91,7	90,9	103,5	97,6	164,3	100,1
ΙΟΥΛ.	ΑΥΓ.	ΣΕΠΤ.	ΟΚΤ.	ΝΟΕΜ.	ΔΕΚ.
105,2	85,8	106	106,5	103,5	104,3

Στόν πίνακα αὐτό, ὁ ἀριθμός π.χ. $91,7 = 100\ \dfrac{100}{118,33}$ σημαίνει ὅτι ἡ τιμή τῆς Υ τόν Ἰανουάριο εἶναι τά 91,7% τῆς μέσης τιμῆς ὅλου τοῦ ἔτους. Εἶναι φανερό ὅτι τό ἄθροισμα ὅλων αὐτῶν τῶν ποσοστῶν εἶναι 1200. Ἀπό τόν νέο πίνακα τῶν τιμῶν τοῦ ἔτους, βλέπομε π.χ. ὅτι τόν Αὔγουστο παρουσιάζεται μία μεγάλη πτώση τῆς τιμῆς τῆς Υ. Μέ τίς παρατηρήσεις ὅμως μόνο ἑνός ἔτους, δέν μποροῦμε νά δοῦμε ἄν ἡ πτώση αὐτή εἶναι ἐποχιακή μεταβολή ἤ συμπτωματική μεταβολή. Γιά νά καταλήξομε στό συμπέρασμα ὅτι πρόκειται γιά ἐποχιακή μεταβολή, θά πρέπει νά ξέρομε γιά τήν ἴδια χρονολογική σειρά τίς παρατηρήσεις καί ἄλλων ἐτῶν καί νά διαπιστώσομε ὅτι κάθε Αὔγουστο παρουσιάζεται αὐτή ἡ πτώση τῆς τιμῆς τῆς Υ.

Γενικά λοιπόν θά ὑπάρχει ἐποχιακή μεταβολή σέ ἕνα ὁρισμένο μήνα, ὅταν τά ποσοστά τοῦ μήνα αὐτοῦ στά διάφορα ἔτη εἶναι αἰσθητά μειωμένα ἤ αὐξημένα. Ὁ μέσος ὅρος τῶν ποσοστῶν ἑνός μήνα στά διάφορα ἔτη, λέγεται ***δείκτης ἐποχικότητας*** τοῦ μήνα αὐτοῦ.

Ἀπό τά παραπάνω, καταλαβαίνομε ὅτι, γιά νά βρίσκομε εὔκολα τούς δεῖκτες ἐποχικότητας, γράφομε τίς παρατηρήσεις μας κάθε ἔτους (ὑπό μορφή ποσοστῶν) σέ μία γραμμή καί, ἀθροίζοντας τίς στῆλες τῶν μηνῶν βρίσκομε τούς μέσους ὅρους τῶν ποσοστῶν κάθε μήνα, τούς ὁποίους τοποθετοῦμε σέ μία τελευταία γραμμή τοῦ πίνακα. Ὁ μηχανισμός αὐτός φαίνεται στό ἑπόμενο παράδειγμα.

Παράδειγμα.

Στόν Πίνακα 7.4.1 δίνονται οἱ πωλήσεις (σέ χιλιάδες kg) ἑνός προϊόντος πού ἔγιναν κάθε μήνα ἀπό ἕνα μεγάλο ἐμπορικό κατάστημα κατά τά ἔτη 1975 — 1977.

ΠΙΝΑΚΑΣ 7.4.1.
Μῆνες

Ἔτος	Ι	Φ	Μ	Α	Μ	Ι	Ι	Α	Σ	Ο	Ν	Δ
1975	109	108	123	116	124	119	125	102	126	127	123	124
1976	120	115	130	123	132	131	135	108	135	143	135	129
1977	134	129	140	138	144	136	148	116	144	146	104	139

Νά βρεθοῦν οἱ δεῖκτες ἐποχικότητας.

Λύση:

Ὑπολογίζομε πρῶτα σέ μιά πρόσθετη στήλη τά ἀθροίσματα τῶν τιμῶν κάθε ἔτους καί σέ μία ἄλλη διπλανή στήλη γράφομε τούς μηνιαίους μέσους κάθε ἔτους (πού βρίσκονται ἄν διαιρέσομε τά ἐτήσια ἀθροίσματα διά 12). Ἔχομε ἔτσι τόν Πίνακα 7.4.2.

ΠΙΝΑΚΑΣ 7.4.2.
Μῆνες

Ἔτος	Ι	Φ	Μ	Α	Μ	Ι	Ι	Α	Σ	Ο	Ν	Δ	ΕΤΗΣΙΑ ΑΘΡΟΙΣΜΑΤΑ	ΜΕΣΟΙ ΜΗΝΙΑΙΟΙ
1975	109	108	123	116	124	119	125	102	126	127	123	124	1426	118,83
1976	120	115	130	123	132	131	135	108	135	143	135	129	1536	128,00
1977	134	129	140	138	144	136	148	116	144	146	104	139	1612	134,33

Στή συνέχεια, ἐκφράζομε κάθε τιμή τοῦ παραπάνω πίνακα ὡς ποσοστό τῆς μέσης μηνιαίας τιμῆς τοῦ ἴδιου ἔτους καί ἀθροίζομε κατά στῆλες. Τέλος, γράφομε κάτω ἀπό κάθε ἄθροισμα τό μέσο ὅρο τῶν ποσοστῶν (δηλαδή τό πηλίκο τοῦ ἀθροίσματος διά τοῦ πλήθους τῶν ἐτῶν) καί ἔχομε ἔτσι τούς δεῖκτες ἐποχικότητας.

ΠΙΝΑΚΑΣ 7.4.3.

Ἔτος	Ι	Φ	Μ	Α	Μ	Ι
1975	91,7	90,9	103,5	97,6	104,3	100,1
1976	93,7	89,8	101,6	96,1	103,1	102,3
1977	99,8	96	104,2	107,2	107,2	102,2
Ἄθροισμα	285,2	276,7	309,3	296,4	314,6	304,6
Μέσος ὅρος	95,1	92,2	103,1	98,8	104,9	101,5

Ι	Α	Σ	Ο	Ν	Δ
105,2	85,8	106	106,8	103,5	104,3
105,5	84,4	105,5	111,7	105,5	100,8
110,2	86,3	107,2	108,7	77,4	103,5
320,9	256,5	318,7	327,2	286,4	308,6
107	85,5	106,2	109,1	95,5	102,9

Οἱ μέσοι ὅροι πού βρίσκονται στήν τελευταία γραμμή τοῦ Πίνακα 7.4.3 παριστάνουν τούς ζητούμενους ***δείκτες ἐποχικότητας*** καί ἐκφράζουν τίς τιμές τῶν διαφόρων μηνῶν ὡς ποσοστά τῆς μέσης μηνιαίας τιμῆς.

Ἔτσι π.χ. ὁ δείκτης 95,1% τοῦ παραπάνω πίνακα σημαίνει ὅτι οἱ πωλήσεις κατά τόν Ἰανουάριο εἶναι κάθε ἔτος κατά 5% μικρότερες τῆς μέσης μηνιαίας πωλήσεως τοῦ ἔτους.

7.5 Ἀσκήσεις.

1. Στόν παρακάτω πίνακα δίνονται οἱ γεννήσεις σέ μιά πόλη κατά τό 1971 ἀνά 10.000 κατοίκους της:

Μῆνες	Ι	Φ	Μ	Α	Μ	Ι	Ι	Α	Σ	Ο	Ν	Δ
Γεννήσεις	24,5	22,3	21,8	20,8	19,8	19,0	18,9	19,0	19,2	18,8	19	20,0

Νά γίνει τό χρονοδιάγραμμα τοῦ πίνακα.

2. Στόν παρακάτω πίνακα δίνεται ἡ παραγωγή ἑνός προϊόντος Α (σέ χιλιάδες τόννους) ἀπό ἕνα ἐργοστάσιο κατά τή χρονική περίοδο 1970 — 1978;

Ἔτος	1970	1971	1972	1973	1974	1975	1976	1977	1978
Παραγωγή (y)	10	12	15	14	17	18	18	20	22

Νά βρεθεῖ ἡ ἐξίσωση τῆς εὐθείας τάσεως (μέ τή μέθοδο τῶν ἐλαχίστων τετραγώνων).

3. Στήν παρακάτω χρονολογική σειρά πού ἀναφέρεται στή χρονική περίοδο 1966 — 1974:

Ἔτος	1966	1967	1968	1969	1970	1971	1972	1973	1974
Υ	80	101	106	93	115	126	111	133	105

νά βρεθεῖ ἡ εὐθεία τάσεως καί οἱ διαφορές τῶν τιμῶν τῆς χρονολογικῆς σειρᾶς ἀπό τίς ἀντίστοιχες τιμές πού δίνει ἡ εὐθεία τάσεως.

4. Στόν παρακάτω πίνακα δίνεται ἡ παραγωγή ἑνός προϊόντος σέ χιλιάδες τόννους κατά τά ἔτη 1973 — 1978:

Ἔτος	y
1973	30
1974	33
1975	40
1976	29
1977	36
1978	44

Νά βρεθεῖ ἡ εὐθεία τάσεως καί νά ἐκτιμηθεῖ ἡ παραγωγή τοῦ προϊόντος τοῦ ἔτους 1970.

5. Δίνεται ἡ παρακάτω χρονολογική σειρά:

Ἔτος	y
1970	5
1971	8
1972	12
1973	15
1974	20
	60

Νά βρεθεῖ ἡ ἐξίσωση τῆς εὐθείας τάσεως καί νά γίνει ἐκτίμηση τῆς τιμῆς τῆς Y γιά τό ἔτος 1972 καί πρόβλεψη γιά τό ἔτος 1982.

6. Στόν παρακάτω πίνακα δίνονται τά μηνιαῖα ἀκαθάριστα ἔσοδα (σέ χιλιάδες δρχ) μιᾶς ἐπιχειρήσεως κατά τά ἔτη 1975 — 1978:

Μῆνες

Ἔτος	Ι	Φ	Μ	Α	Μ	Ι	Ι	Α	Σ	Ο	Ν	Δ
1973	126	142	142	137	125	121	113	106	109	114	131	153
1974	174	179	159	158	154	140	125	121	122	131	148	162
1975	187	196	179	171	167	159	144	139	139	149	172	190

Νά ὑπολογισθοῦν οἱ δεῖκτες ἐποχικότητας.

7. Μέ τή μέθοδο τῶν δύο μέσων σημείων νά ὑπολογισθεῖ ἡ εὐθεία τάσεως στίς παρακάτω χρονολογικές σειρές:

α)

Ἔτος	Y
1965	2500
1966	2700
1967	3100
1968	3700
1969	4000
1970	5500
1971	6300
1972	6900
1973	7800
1974	8500

β)

Ἔτος	Y
1968	50,5
1969	70,0
1970	115,0
1971	100,6
1972	93,6
1973	231,6
1974	392,8
1975	264,8
1976	148,5
1977	274,5

8. Στόν παρακάτω πίνακα δίνεται ἡ παραγωγή ἑνός προϊόντος A (σέ ἑκατομμύρια kg) ἀπό ἕνα ἐργοστάσιο κατά τή χρονική περίοδο 1973 — 1978:

Ἔτος	1973	1974	1975	1976	1977	1978
Παραγωγή	2	3	2,5	3,8	4	4,2

α) Νά βρεθεῖ ἡ ἐξίσωση τῆς εὐθείας τάσεως μέ τή μέθοδο τῶν ἐλαχίστων τετραγώνων. β) Νά ὑπολογισθοῦν οἱ διαφορές τῆς χρονολογικῆς σειρᾶς ἀπό τίς ἀντίστοιχες τιμές πού δίνει ἡ εὐθεία τάσεως. γ) Νά γίνει τό χρονοδιάγραμμα.

9. Στόν παρακάτω πίνακα δίνεται ἡ κυκλοφορία τῶν τραπεζογραμματίων κατά τά ἔτη 1957 — 1965:

Έτος	Τραπεζογρ. σέ έκ. δρ.
1957	973
1958	1202
1959	1959
1960	1887
1961	2199
1962	2476
1963	3503
1964	3888
1965	4951

Νά βρεθεῖ ἡ εὐθεία τάσεως $y = \alpha + \beta x$ καί νά γίνει ἐκτίμηση γιά τό ἔτος 1960 μέ βάση τήν εὐθεία τάσεως.

10. Στόν παρακάτω πίνακα δίνεται ἡ μηνιαία παραγωγή (σέ χιλιάδες kg) ἑνός προϊόντος Β ἀπό ἕνα ἐργοστάσιο κατά τά ἔτη 1975 — 1978.

Μῆνες

Έτος	Ι	Φ	Μ	Α	Μ	Ι	Ι	Α	Σ	Ο	Ν	Δ
1975	104	102	121	134	156	154	136	119	95	92	91	109
1976	108	104	125	129	158	152	123	102	92	95	93	106
1977	115	114	130	135	152	149	128	110	92	93	92	103
1978	119	113	132	130	151	145	127	98	87	92	90	107

Νά ὑπολογισθοῦν οἱ δεῖκτες ἐποχικότητας.

ΚΕΦΑΛΑΙΟ ΟΓΔΟΟ

ΑΡΙΘΜΟΔΕΙΚΤΕΣ

8.1 Γενικά.

Οἱ ***ἀριθμοδεῖκτες*** ἤ καί ἁπλῶς ***δεῖκτες*** εἶναι καθαροί ἀριθμοί πού μποροῦν νά ἐκφράσουν τή μεταβολή μιᾶς μεταβλητῆς μεταξύ δύο χρονικῶν στιγμῶν (χρονολογικοί ἀριθμοδεῖκτες) ἤ μεταξύ δύο γεωγραφικῶν περιοχῶν (γεωγραφικοί ἀριθμοδεῖκτες) ἤ γενικότερα μεταξύ δύο καταστάσεων.

Ἔτσι, π.χ. μέ τούς ἀριθμοδεῖκτες μποροῦμε νά συγκρίνομε τό κόστος διατροφῆς μιᾶς χώρας σέ δύο διαφορετικές χρονικές στιγμές ἤ τήν παραγωγή σιδήρου ἑνός ὁρισμένου ἔτους σέ δύο διαφορετικές χῶρες, κ.ο.κ.

Ἐδῶ, θά περιορισθοῦμε στούς χρονολογικούς ἀριθμοδεῖκτες, οἱ ὁποῖοι χρησιμοποιοῦνται σέ πολλούς τομεῖς τῆς ἀνθρώπινης δραστηριότητας καί ἰδιαίτερα στήν οἰκονομία, ὅπου μέ τίς πληροφορίες πού συγκεντρώνουν οἱ διάφοροι κρατικοί καί ἰδιωτικοί ὀργανισμοί, καταρτίζονται δεῖκτες γιά τήν παραγωγή καί τήν κατανάλωση ἀγαθῶν, γιά τό ὕψος τῶν ἡμερομισθίων, γιά τίς εἰσαγωγές καί ἐξαγωγές κ.τ.λ.

Οἱ πιό γνωστοί ἀριθμοδεῖκτες εἶναι ***ὁ δείκτης τοῦ κόστους ζωῆς*** καί ὁ ***τιμάριθμος*** πού ἐκφράζει τή μεταβολή τῶν τιμῶν τῶν διαφόρων ἀγαθῶν.

Ἕνας (χρονολογικός) ἀριθμοδείκτης, θά λέγεται ***ἰδιαίτερος ἀριθμοδείκτης.*** Ὅταν ἀναφέρεται σέ μία μεμονωμένη μεταβλητή (ὅπως ὁ δείκτης παραγωγῆς πατάτας, ὁ δείκτης ἡμερομισθίου ἀνειδίκευτου ἐργάτη κ.λπ.) ἐνῶ, ὅταν ἀναφέρεται σέ περισσότερες μεταβλητές (ὅπως ὁ δείκτης τιμῶν γεωργικῶν προϊόντων, ὁ δείκτης κόστους ζωῆς κ.τ.λ.) θά λέγεται ***συνθετικός δείκτης.***

8.2 Ἰδιαίτεροι ἀριθμοδεῖκτες.

Ὁ πιό συνηθισμένος ἀριθμοδείκτης γιά μία μεταβλητή Y εἶναι ἡ ***σχετική τιμή της,*** ἡ ὁποία ὁρίζεται ὡς ἑξῆς:

— Ὀνομάζομε y_0 τήν τιμή τῆς Y σέ μία ὁρισμένη «ἀρχική» χρονική στιγμή. Ἡ χρονική αὐτή στιγμή θά σημειώνεται $T = 0$ καί θά λέγεται ***βάση*** ἤ ***ἀρχή τῶν χρόνων.***

— Ὀνομάζομε y_1 τήν τιμή τῆς Y σέ μιά ἄλλη χρονική στιγμή $T = t$.

— Σέ κάθε χρονική στιγμή t, σχηματίζομε τό λόγο y_1/y_0, ὁ ὁποῖος λέγεται ***σχετική τιμή τῆς y*** στή χρονική στιγμή t ὡς πρός βάση τήν $T = 0$ καί θά σημειώνεται $Y_{1/0}$, δηλαδή:

$$Y_{1/0} = \frac{y_1}{y_0}$$

Ἔτσι λοιπόν, ἡ σχετική τιμή τῆς Y, σέ κάθε χρονική στιγμή t, εἶναι τό μέτρο τῆς τιμῆς τῆς Y μέ μονάδα μετρήσεως τό y_0. Ἄν π.χ. ἔχομε $Y_{2/0} = 0{,}973$, αὐτό σημαίνει ὅτι ἡ τιμή τῆς Y στή χρονική στιγμή T = t, εἶναι τό 0,973 τῆς τιμῆς πού εἶχε ἡ Y κατά τή χρονική στιγμή T = 0 καί συνεπῶς εἶναι $y_1 = (0{,}973)\, y_0$. Εἶναι φανερό ὅτι ἡ σχετική τιμή τῆς βάσεως εἶναι 1. Συνήθως, ἡ σχετική τιμή ἐκφράζεται σέ ἑκατοστιαία ἀναλογία, δηλαδή παίρνομε:

$$Y_{1/0} = \frac{y_1}{y_0} \cdot 100$$

καί στήν περίπτωση αὐτή, ἡ σχετική τιμή τῆς βάσεως εἶναι 100. Ἔτσι, ὅταν γράφομε π.χ. $Y_{1/0} = 97{,}3\%$, καταλαβαίνομε ὅτι ἡ τιμή y_1 εἶναι τά 97,3% τῆς τιμῆς y_0 καί λέγεται συνήθως ***σχετικός δείκτης*** τῆς Y.

Μερικοί τέτοιοι βασικοί ἀριθμοδεῖκτες στήν Οἰκονομία εἶναι:

α) Ὁ σχετικός δείκτης τιμῆς ἑνός ἀγαθοῦ.

Ἄν ὀνομάσομε **P** τήν τιμή ἑνός ὁρισμένου ἀγαθοῦ καί σημειώσομε μέ P_0 τήν τιμή τῆς χρονικῆς στιγμῆς τῆς βάσεως T = 0 καί μέ P_1 τήν τιμή του σέ μιά ἄλλη χρονική στιγμή, τότε ἡ σχετική τιμή τῆς **P** σέ χρονική στιγμή t θά εἶναι:

$$\boxed{P_{1/0} = \frac{P_1}{P_0} \cdot 100} \qquad (8.1)$$

καί θά λέγεται ***σχετικός δείκτης τιμῆς τοῦ ἀγαθοῦ.***

Παράδειγμα.

Ἄν ἡ τιμή λιανικῆς πωλήσεως μιᾶς μεγάλης φιάλης πορτοκαλάδας κατά τά ἔτη 1976 καί 1978 ἦταν 12 καί 14 δρχ. ἀντίστοιχα, ὁ σχετικός δείκτης τῆς τιμῆς τῆς φιάλης τό 1978 (μέ βάση τό ἔτος 1976) ἦταν:

$$P_{1/0} = \frac{P_1}{P_2} \times 100 = \frac{14}{12} \times 100 = 127\%.$$

Αὐτό σημαίνει ὅτι τό 1978 ἡ τιμή μιᾶς μεγάλης φιάλης πορτοκαλάδας ἦταν τό 127% τῆς τιμῆς της τοῦ 1976, δηλαδή αὐξήθηκε κατά 27%.

β) Ὁ σχετικός δείκτης ποσότητας ἑνός ἀγαθοῦ.

Ἄν ὀνομάσομε Q τήν ποσότητα ἑνός ὁρισμένου ἀγαθοῦ καί σημειώσομε μέ q_0 τήν ποσότητά του κατά τή χρονική στιγμή τῆς βάσεως T = 0 καί μέ q_1 τήν ποσότητά του σέ μία ἄλλη χρονική στιγμή t, τότε, ἡ σχετική ποσότητα τοῦ ἀγαθοῦ τή χρονική στιγμή t θά εἶναι:

$$\boxed{Q_{1/0} = \frac{q_1}{q_0} \times 100} \qquad (8.2)$$

καί θά λέγεται ***σχετικός δείκτης ποσότητας τοῦ ἀγαθοῦ.***

Παράδειγμα.

Ἂν ἡ παραγωγή πατάτας (σέ τόννους) στό Νομό Θεσπρωτίας κατά τά ἔτη 1976 καί 1977 ἦταν 4200 καί 4800 ἀντίστοιχα. Ὁ σχετικός δείκτης ποσότητας παραγωγῆς πατάτας τό 1977 (μέ βάση τό ἔτος 1976) ἦταν:

$$Q_{1/0} = \frac{4800}{4200} \times 100 = 114\%.$$

Αὐτό σημαίνει ὅτι ἡ παραγωγή πατάτας κατά τό 1977 αὐξήθηκε, σέ σύγκριση μέ τό ἔτος 1976 κατά 14%.

γ) Ὁ σχετικός δείκτης ἀξίας ἑνός ἀγαθοῦ.

Ἂν ὀνομάσομε V τήν ἀξία ἑνός ὁρισμένου ἀγαθοῦ καί σημειώσομε μέ U_0 τήν ἀξία του στή χρονική στιγμή τῆς βάσεως T = 0 καί μέ U_1 τήν ἀξία του σέ μία ἄλλη χρονική στιγμή t, τότε ἡ σχετική ἀξία τοῦ ἀγαθοῦ στή χρονική στιγμή t θά εἶναι:

$$V_{1/0} = \frac{U_1}{U_0} \times 100$$

καί θά λέγεται ***σχετικός δείκτης ἀξίας*** τοῦ ἀγαθοῦ.

Ἂν Q εἶναι ἡ ποιότητα τοῦ ἀγαθοῦ καί **P** ἡ τιμή του, τότε ἔχομε V = **P** . Q. Ἔτσι, θά εἶναι $U_1 = P_1 \cdot q_1$ καί $U_0 = P_0 \cdot q_0$ καί συνεπῶς ὁ σχετικός δείκτης ἀξίας τοῦ ἀγαθοῦ γράφεται:

$$\boxed{V_{1/0} = \frac{P_1 \cdot q_1}{P_0 \cdot q_0} \times 100} \qquad (8.3)$$

Παράδειγμα.

Ὁ Πίνακας 8.2.1 δίνει τίς τιμές (σέ δρχ) καί τίς ποσότητες (σέ kg) ἑνός ἀγαθοῦ κατά τά ἔτη 1975 καί 1977.

ΠΙΝΑΚΑΣ 8.2.1:

Ἔτη	Τιμές	Ποσότητες
1975	32	162
1977	41	168

Ὁ σχετικός δείκτης ἀξίας τοῦ ἀγαθοῦ τό 1977 (μέ βάση τό ἔτος 1975) εἶναι:

$$V_{1/0} = \frac{P_1 \cdot q_1}{P_0 \cdot q_0} \times 100 = \frac{41 \times 168}{32 \times 162} \times 100 = \frac{6888}{5184} \times 100 = 133\%.$$

Ἂν ἐκφράσομε τούς δεῖκτες σέ ἁπλά πηλίκα (καί ὄχι σέ ἑκατοστιαῖες ἀναλογίες) θά ἔχομε:

$$V_{1/0} = \frac{U_1}{U_0} = \frac{P_0 \cdot q_0}{P_1 \cdot q_1} = \frac{P_0}{P_1} \cdot \frac{q_0}{q_1} = P_{1/0} \cdot Q_{1/0},$$

δηλαδή ὁ σχετικός δείκτης ἀξίας θά εἶναι ἴσος μέ τό γινόμενο τῶν σχετικῶν δεικτῶν τῆς τιμῆς καί τῆς ποσότητας ἑνός ἀγαθοῦ.

8.3 Συνθετικοί ἀριθμοδείκτες.

Στά παραπάνω ἀσχοληθήκαμε εἰδικά μέ τή μεταβολή τῆς τιμῆς ἤ τῆς ἀξίας ἑνός μεμονωμένου ἀγαθοῦ μεταξύ δύο χρονικῶν περιόδων. Στήν καθημερινή μας ὅμως ζωή, ἐνδιαφερόμαστε κυρίως γιά τίς μεταβολές πού παρουσιάζουν οἱ τιμές ἤ οἱ ἀξίες ὅλων τῶν ἀγαθῶν μιᾶς ὁμάδας, ἡ ὁποία ἀποτελεῖται ἀπό ὁμοειδή ἀγαθά. Ἔτσι, π.χ. γιά τόν ὑπολογισμό τοῦ κόστους ζωῆς δέν μᾶς ἀρκεῖ μόνο ἡ μεταβολή τῆς τιμῆς τοῦ ψωμιοῦ, ἀλλά μᾶς ἐνδιαφέρουν οἱ μεταβολές μιᾶς ὁμάδας ἀγαθῶν (ψωμιοῦ, κρέατος, λαχανικῶν, κ.λπ.). Σέ μία τέτοια περίπτωση ὑπολογίζομε ἕνα συνθετικό δείκτη, ὁ ὁποῖος εἶναι τό μέτρο μεταβολῆς τῶν τιμῶν ἤ τῶν ἀξιῶν ὅλων μαζί τῶν ἀγαθῶν τῆς ὁμάδας.

Σέ ὅλα λοιπόν τά ἑπόμενα θά ὑποθέτομε ὅτι ἔχομε μία ὁμάδα ἀπό ν ἀγαθά $A_1^{(1)}$, $A_2^{(2)}$, ...$A_\nu^{(\nu)}$ καί θά ὀνομάζομε ἀντίστοιχα:

- $P^{(1)}, P^{(2)}, \ldots, P^{(\nu)}$ τίς τιμές τους.
- $Q^{(1)}, Q^{(2)}, \ldots, Q^{(\nu)}$ τίς ποσότητές τους.
- $V^{(1)}, V^{(2)}, \ldots, V^{(\nu)}$ τίς ἀξίες τους.

Ἐπίσης, μέ τά ἀντίστοιχα μικρά γράμματα θά σημειώνομε τίς ὁρισμένες τιμές τῶν παραπάνω μεγεθῶν καί μάλιστα, ὅταν ἀναφέρονται στή βάση T = 0, θά ἔχουν δείκτη 0, ἐνῶ, ὅταν ἀναφέρονται σέ μιά χρονική στιγμή t, θά ἔχουν δείκτη 1. Π.χ. μέ $P_0^{(i)}$, $q_0^{(i)}$, $U_0^{(i)}$ θά σημειώνεται ἡ τιμή, ἡ ποσότητα καί ἡ ἀξία τοῦ ἀγαθοῦ $p^{(i)}$ στή βάση T = 0, καί μέ $P_1^{(i)}$, $q_1^{(i)}$, $U_1^{(i)}$ θά σημειώνονται τά ἴδια μεγέθη στή χρονική στιγμή t.

α) Ὁ ἁπλός δείκτης συνολικῶν τιμῶν.

Θά ὀνομάζεται ἔτσι ὁ ἀριθμός $P_{1/0}$, ὁ ὁποῖος ὁρίζεται ἀπό τήν ἰσότητα:

$$P_{1/0} = \frac{P_1^{(1)} + P_1^{(2)} + P_1^{(3)} + \ldots + P_1^{(\nu)}}{P_0^{(1)} + P_0^{(2)} + P_0^{(3)} + \ldots + P_0^{(\nu)}} = \frac{\sum_i P_1^{(i)}}{\sum_i P_0^{(i)}} \qquad (8.4)$$

Βλέπομε δηλαδή ὅτι ὁ δείκτης αὐτός εἶναι ὁ λόγος τοῦ ἀθροίσματος τῶν τιμῶν πού ἔχουν τά ν ἀγαθά κατά τή χρονική στιγμή t, πρός τό ἄθροισμα τῶν τιμῶν τους, στή χρονική στιγμή τῆς βάσεως.

Παράδειγμα.

ΠΙΝΑΚΑΣ 8.3.1.

Ἀγαθά	Τιμές ἀγαθῶν	
	1975	1977
	P_0	P_1
Α	24	30
Β	30	36
Γ	36	39
Δ	12	18
Ἄθροισμα	102	123

Ὁ Πίνακας 8.3.1 δίνει τίς μέσες τιμές χονδρικῆς πωλήσεως τεσσάρων ἀγαθῶν κατά τά ἔτη 1975 καί 1977.

Ὁ ἁπλός δείκτης συνολικῶν τιμῶν τοῦ ἔτους 1977 (μέ βάση τό ἔτος 1977) θά εἶναι:

$$P_{1/0} = \frac{\sum_{i} P_1^{(i)}}{\sum_{i} P_0^{(i)}} = \frac{123}{102} = 1{,}206 = 120{,}6\%$$

Αὐτό σημαίνει ὅτι οἱ τιμές χονδρικῆς πωλήσεως τό 1977 εἶναι κατά 20,6% μεγαλύτερες ἀπό τίς τιμές τοῦ 1975.

Ὁ παραπάνω δείκτης παρουσιάζει τά ἑξῆς μειονεκτήματα:

— ***Ἀγνοεῖ τίς ποσότητες τῶν ἀγαθῶν.***

— ***Ἐπηρεάζεται ἀπό τίς διαφορετικές μονάδες μετρήσεως τῶν ἀγαθῶν.***

— ***Θεωρεῖ ὅτι ὅλα τά ἀγαθά ἔχουν τήν ἴδια βαρύτητα.***

β) Ὁ δείκτης τῆς μέσης σχετικῆς τιμῆς.

Θά ὀνομάζεται ἔτσι ὁ ἀριθμός $P_{1/0}$, ὁ ὁποῖος δίνεται ἀπό τήν ἰσότητα:

$$P_{1/0} = \frac{1}{v}\left[\frac{P_1^{(1)}}{P_0^{(1)}} + \frac{P_1^{(2)}}{P_0^{(2)}} + \ldots + \frac{P_1^{(v)}}{P_0^{(v)}}\right] = \frac{1}{v}\sum_{i}\left(\frac{P_1^{(i)}}{P_0^{(i)}}\right) \quad (8.5)$$

δηλαδή ὁ ἀριθμός, ὁ ὁποῖος εἶναι μέσος ἀριθμητικός τῶν σχετικῶν τιμῶν ὅλων τῶν ἀγαθῶν.

Παράδειγμα.

Ὁ δείκτης τῶν μέσων σχετικῶν τιμῶν στόν πίνακα τοῦ προηγούμενου παραδείγματος εἶναι:

$$P_{1/0} = \frac{1}{4}\left[\frac{30}{24} + \frac{36}{30} + \frac{39}{36} + \frac{18}{12}\right] = \frac{5{,}03}{4} = 1{,}258 = 125{,}8\%$$

Αὐτό σημαίνει ὅτι ἡ τιμή χονδρικῆς πωλήσεως τό 1977 ἦταν κατά 25,8% μεγαλύτερη ἀπό τήν τιμή τοῦ 1975.

Ἄν καί ὁ δείκτης αὐτός δέν ἐπηρεάζεται ἀπό τίς μονάδες μετρήσεως τῶν ἀγαθῶν (ἀφοῦ εἶναι ἄθροισμα λόγων), ὅμως ἀγνοεῖ πάλι τίς ποιότητες τῶν ἀγαθῶν καί θεωρεῖ ὅτι ὅλα τά ἀγαθά ἔχουν τήν ἴδια βαρύτητα.

8.4 Σταθμικοί συνθετικοί ἀριθμοδεῖκτες.

Πολλές φορές, ὅταν θεωροῦμε μία ὁμάδα μεταβλητῶν, μᾶς ἐνδιαφέρει περισσότερο ἡ μεταβολή ἑνός ὁρισμένου ἀγαθοῦ, παρά ἡ μεταβολή ἑνός ἄλλου. Ἔτσι, π.χ. στόν ὑπολογισμό τοῦ κόστους ζωῆς, μᾶς ἐνδιαφέρει περισσότερο ἡ μεταβολή τῆς τιμῆς τοῦ ψωμιοῦ ἀπό τή μεταβολή τῆς τιμῆς τοῦ κρέατος ἤ τοῦ βουτύρου. Στήν περίπτωση αὐτή καί ἀνάλογα μέ τή σημασία πού ἔχουν τά διάφορα ἀγαθά, γιά τόν ὑπολογισμό τοῦ συνθετικοῦ ἀριθμοδείκτη, σταθμίζομε τίς τιμές τῶν ἀγαθῶν, δηλαδή πολλαπλασιάζομε τήν τιμή κάθε ἀγαθοῦ μέ ἕνα ***συντελεστή σταθμίσεως,*** ὁ ὁποῖος εἶναι, συνήθως, ἡ ποσότητα ἤ ὁ ὄγκος του, σέ μία ὁρισμένη χρονική περίοδο. Οἱ ἀριθμοδεῖκτες πού προκύπτουν μέ τόν τρόπο αὐτό, λέγονται ***σταθμικοί συνθετικοί ἀριθμοδεῖκτες*** καί οἱ κυριότεροι ἀπό αὐτούς εἶναι:

1. Σταθμικός δείκτης τιμῶν (τιμάριθμος).

Ὅταν ἐνδιαφερόμαστε γιά τίς μεταβολές τῶν τιμῶν μιᾶς ὁμάδας ἀγαθῶν, παίρνομε ὡς συντελεστή σταθμίσεως τῆς τιμῆς κάθε ἀγαθοῦ τήν ποσότητά του κατά τή χρονική στιγμή τῆς βάσεως. Τότε ὁ δείκτης συνολικῶν τιμῶν γίνεται:

$$\boxed{P_{1/0} = \frac{\sum_i P_1^{(i)} q_0^{(i)}}{\sum_i P_0^{(i)} q_0^{(i)}} = \frac{P_1^{(1)} q_0^{(1)} + P_1^{(2)} q_0^{(2)} + \ldots + P_1^{(V)} q_0^{(V)}}{P_0^{(1)} q_0^{(1)} + P_0^{(2)} q_0^{(2)} + \ldots + P_0^{(V)} q_0^{(V)}}} \tag{8.6}$$

καί λέγεται ***σταθμικός τιμάριθμος κατά Laspeyres.***

Ὁ τύπος τοῦ Laspeyres γράφεται ἀκόμη:

$$\boxed{P_{1/0} = \frac{\sum_i P_0^{(i)} q_0^{(i)} \cdot (P_1^{(i)}/P_0^{(i)})}{\sum_i P_0^{(i)} q_0^{(i)}} = \sum_i a_i \frac{P_1^{(i)}}{P_0^{(i)}}} \tag{8.7}$$

Ὅπου οἱ ἀριθμοί $a = \dfrac{P_0^{(i)} q_0^{(i)}}{\sum_i P_0^{(i)} q_0^{(i)}}$ εἶναι οἱ συντελεστές σταθμίσεως ὅπου ἀντιστοιχίζονται οἱ σχετικές τιμές τῶν μεταβλητῶν.

Μποροῦμε ἐπίσης νά πάρομε ὡς συντελεστή σταθμίσεως τῆς τιμῆς κάθε ἀγαθοῦ τήν ποσότητά του κατά τή χρονική στιγμή t, ὁπότε ὁ δείκτης συνολικῶν τιμῶν γίνεται:

$$\boxed{P_{1/0} = \frac{\sum_i P_1^{(i)} q_1^{(i)}}{\sum_i P_0^{(i)} q_1^{(i)}} = \frac{P_1^{(1)} q_1^{(1)} + P_1^{(2)} q_1^{(2)} + \ldots + P_1^{(V)} q_1^{(V)}}{P_0^{(1)} q_1^{(1)} + P_0^{(2)} q_1^{(2)} + \ldots + P_0^{(i)} q_1^{(V)}}} \tag{8.8}$$

καί ἀκόμη: ***λέγεται σταθμικός τιμάριθμος κατά Paasche.*** Ὁ τύπος αὐτός γράφεται

$$\boxed{P_{1/0} = \frac{\sum_i P_0^{(i)} q_0^{(i)} \left(\dfrac{P_1^{(i)}}{P_0^{(i)}}\right)}{\sum_i P_0^{(i)} q_1^{(i)}} = \sum_i a_i \frac{P_1^{(i)}}{P_0^{(i)}}} \tag{8.9}$$

όπου οί άριθμοί $a_i = \dfrac{P_0^{(i)} q_1^{(i)}}{\sum_i P_0^{(i)} q_1^{(i)}}$ είναι οί συντελεστές σταθμίσεως τών σχετικών τιμών πού άντιστοιχίζονται στίς σχετικές συχνότητες τών μεταβλητών.

2. Σταθμικοί δείκτες ποσοτήτων ή όγκων.

Όταν ένδιαφερόμαστε γιά τή μεταβολή τών ποσοτήτων μιάς όμάδας άγαθών, παίρνομε ώς συντελεστή σταθμίσεως τής ποσότητας κάθε άγαθοΰ τήν τιμή του ή κατά τή χρονική στιγμή τής βάσεως ή κατά τή χρονική στιγμή t. Έτσι, μπορούμε πάλι νά όρίσομε ***σταθμικό δείκτη όγκου*** ή μέ τόν τύπο τοΰ ***Laspeyres:***

$$\boxed{Q_{1/0} = \frac{\sum_i q_1^{(i)} P_0^{(i)}}{\sum_i q_0^{(i)} P_0^{(i)}}} \tag{8.10}$$

ή μέ τόν τύπο τοΰ ***Paasche:***

$$\boxed{Q_{1/0} = \frac{\sum_i q_1^{(i)} P_1^{(i)}}{\sum_i q_0^{(i)} P_1^{(i)}}} \tag{8.11}$$

3. Σταθμικοί δείκτες άξίας.

Όταν ένδιαφερόμαστε γιά τίς μεταβολές τών άξιών μιάς όμάδας μεταβλητών, παίρνομε ώς μέτρο τής συνολικής μεταβολής τους τόν άριθμό $V_{1/0}$ πού όρίζεται άπό τήν ίσότητα:

$$\boxed{V_{1/0} = \frac{\sum_i P_1^{(i)} q_1^{(i)}}{\sum_i P_0^{(i)} q_0^{(i)}}} \tag{8.12}$$

Αύτός θά λέγεται ***δείκτης συνολικής άξίας.*** Βλέπομε δηλαδή ότι ό δείκτης αύτός είναι ό λόγος τής συνολικής άξίας όλων τών άγαθών στήν τρέχουσα περίοδο t, πρός τή συνολική άξία τους στήν περίοδο βάσεως T = 0.

Παράδειγμα.

Ό Πίνακας 8.4.1 μάς δίνει τίς τιμές μονάδας τριών άγαθών Α, Β, Γ καί τίς ποσότητες πού έχουν παραχθεί (σέ χιλιάδες τόννους), άπό ένα έργοστάσιο κατά τά έτη 1972 καί 1975.

ΠΙΝΑΚΑΣ 8.4.1.

Άγαθά	1972		1975	
	P_0	q_0	P_1	q_1
Α	10	3	12	2
Β	4	50	4	80
Γ	100	2	180	1

Νά βρεθοῦν (μέ βάση τό ἔτος 1972):
α) Ὁ σταθμικός δείκτης τιμῶν κατά Laspeyres.
β) Ὁ σταθμικός δείκτης τιμῶν κατά Paasche.
γ) Ὁ δείκτης ὄγκου κατά Laspeyres.
δ) Ὁ δείκτης ὄγκου κατά Paasche.
ε) Ὁ δείκτης ἀξίας.

Λύση:

Σχηματίζομε τόν Πίνακα 8.4.2 ἀριθμητικῶν ὑπολογισμῶν:

ΠΙΝΑΚΑΣ 8.4.2.

Ἀγαθά	P_0	P_1	q_0	q_1	$P_0 q_0$	$P_0 q_1$	$P_1 q_0$	$P_1 q_1$
Α	10	12	3	2	30	20	36	24
Β	4	4	50	80	200	320	200	320
Γ	100	180	2	1	200	100	360	180
Ἄθροισμα					430	440	596	524

Ἔτσι, ἔχομε ἀμέσως:
α) Δείκτη τιμῶν κατά Laspeyres:

$$P_{1/0} = \frac{\sum_i P_1^{(i)} P_0^{(i)}}{\sum_i P_0^{(i)} q_0^{(i)}} = \frac{596}{430} = 1,386 = 138,6\%.$$

β) Δείκτη τιμῶν κατά Paasche:

$$P_{1/0} = \frac{\sum_i P_1^{(i)} q_1^{(i)}}{\sum_i P_0^{(i)} q_1^{(i)}} = \frac{524}{440} = 1,190 = 119\%.$$

γ) Δείκτη ὄγκου κατά Laspeyres:

$$Q_{1/0} = \frac{\sum_i q_1^{(i)} P_0^{(i)}}{\sum_i q_0^{(i)} P_0^{(i)}} = \frac{440}{430} = 1,023 = 102,3\%.$$

δ) Δείκτη ὄγκου κατά Paasche:

$$Q_{1/0} = \frac{\sum_i q_1^{(i)} P_1^{(i)}}{\sum_i P_1^{(i)} q_0^{(i)}} = \frac{524}{596} = 0,879 = 87,9\%.$$

ε) Δείκτη ἀξίας:

$$V_{1/0} = \frac{\sum_i P_1^{(i)} q_1^{(i)}}{\sum_i P_0^{(i)} q_0^{(i)}} = \frac{524}{430} = 1,219 = 121,9\%.$$

8.5 Ἀσκήσεις.

1. Δίνονται παρακάτω οἱ τιμές καί οἱ ποσότητες ἑνός ἀγαθοῦ πού καταναλώθηκαν κατά τά ἔτη 1976, 1977 καί 1978:

Ἔτη	1976	1977	1978
Τιμές (σέ δρχ.)	42	46	54
Ποσότητες (σέ τόννους)	160	174	168

Νά ὑπολογισθοῦν οἱ σχετικές τιμές καί οἱ σχετικές ποσότητες γιά τά ἔτη 1977 καί 1978 μέ βάση τό ἔτος 1976 = 100.

2. Δίνεται παρακάτω ἡ τιμή τῆς πατάτας (σέ δρχ. κατά kg) κατά τά ἔτη 1972 — 1975:

Ἔτη	1972	1973	1974	1975
Τιμές	6	9	11	13

Νά ὑπολογισθοῦν οἱ σχετικές τιμές γιά τά ἔτη 1973, 1974, 1975 μέ βάση τό ἔτος 1972.

3. Δίνονται παρακάτω οἱ τιμές (σέ δρχ. κατά kg) πέντε ἀγαθῶν γιά τά ἔτη 1975 καί 1978:

Ἀγαθά	Τιμές	
	1975 P_0	1978 P_1
Α	12	15
Β	6	7
Γ	8	10
Δ	20	25
Ε	38	44

Νά ὑπολογισθεῖ ὁ δείκτης συνολικῶν τιμῶν γιά τό ἔτος 1978 μέ βάση τό ἔτος 1975.

4. Δίνονται οἱ τιμές (σέ δρχ.) τεσσάρων ἀγαθῶν γιά τά ἔτη 1976 καί 1977:

Ἀγαθά	Τιμές	
	1976 P_0	1977 P_1
Α	40	42
Β	100	115
Γ	250	270
Δ	265	280

Νά ὑπολογισθεῖ ὁ δείκτης τῆς μέσης σχετικῆς τιμῆς.

5. Τά παρακάτω δεδομένα δίνουν τίς τιμές (σέ δρχ. κατά τόννο) τριῶν βιομηχανικῶν προϊόντων γιά τά ἔτη 1974 καί 1976:

'Αγαθά	1974		1976	
	P_0	q_0	P_1	q_1
Α	2000	1000	2100	980
Β	620	100	680	120
Γ	440	50	520	120

Νά ὑπολογισθεῖ, μέ βάση τό ἔτος 1974, ὁ σταθμικός δείκτης τιμῶν κατά Laspeyres.

6. Δίνονται τά παρακάτω στοιχεῖα:

'Αγαθά	1972		1975	
	P_0	q_0	P_1	q_1
Α	5	10	7	8
Β	4	20	6	30
Γ	3	12	4	10

Νά ὑπολογισθεῖ, μέ βάση τό ἔτος 1972, ὁ σταθμικός δείκτης τιμῶν μέ τόν τύπο τοῦ Paasche.

7. Μέ βάση τά δεδομένα τοῦ παρακάτω πίνακα:

'Αγαθά	1976		1978	
	q_0	P_0	q_1	P_1
Α	2	25	1	30
Β	8	10	7	15
Γ	5	40	6	30
Δ	10	20	9	24

Νά ὑπολογισθεῖ, μέ βάση τό ἔτος 1976:
α) 'Ο δείκτης ὄγκου κατά Laspeyres.
β) 'Ο δείκτης ὄγκου κατά Paasche.
γ) 'Ο δείκτης ἀξίας.

8. Μέ βάση τά δεδομένα τοῦ παρακάτω πίνακα

'Αγαθά	1975		1976	
	q_0	P_0	q_1	P_1
Α	5	20	6	18
Β	10	30	8	36
Γ	20	10	25	15
Δ	2	50	4	60

Νά ὑπολογισθεῖ:
α) 'Ο μέσος ἀριθμητικός τῶν σχετικῶν τιμῶν.
β) 'Ο σταθμικός δείκτης τιμῶν κατά Laspeyres.
γ) 'Ο σταθμικός δείκτης τιμῶν κατά Paasche.
δ) 'Ο δείκτης ὄγκου κατά Laspeyres καί Paasche.
ε) 'Ο δείκτης ἀξίας.
Ἔτος βάσεως τό 1975.

ΣΤΑΤΙΣΤΙΚΗ

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

ΣΚΟΠΟΣ ΚΑΙ ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ

ΣΥΛΛΟΓΗ ΣΤΑΤΙΣΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ

ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

ΜΕΤΡΑ ΘΕΣΕΩΣ

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

ΚΕΦΑΛΑΙΟ ΕΚΤΟ

ΠΑΛΙΝΔΡΟΜΗΣΗ ΚΑΙ ΣΥΣΧΕΤΙΣΗ ΔΥΟ ΜΕΤΑΒΛΗΤΩΝ

ΚΕΦΑΛΑΙΟ ΕΒΔΟΜΟ

ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ

ΚΕΦΑΛΑΙΟ ΟΓΔΟΟ

ΑΡΙΘΜΟΔΕΙΚΤΕΣ