ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΕΚΠΑΙΔΕΥΣΗΣ ΕΝΗΛΙΚΩΝ
ΙΝΣΤΙΤΟΥΤΟ ΔΙΑΡΚΟΥΣ ΕΚΠΑΙΔΕΥΣΗΣ ΕΝΗΛΙΚΩΝ

βασικές γνώσεις μαθηματικών-στατιστικής

Στατιστική: Εφαρμογές στη διοίκηση-οικονομία- επιχειρήσεις

ΚΕΝΤΡΑ ΕΚΠΑΙΔΕΥΣΗΣ ΕΝΗΛΙΚΩΝ

ΓΓΕΕ

ΙΔΕΚΕ

ΕΛΛΑΔΑ 2008
Ανάπτυξη παντού. Ανάπτυξη για όλους.

ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΕΑΕΚ

ΕΥΡΩΠΑΪΚΗ ΕΝΩΣΗ
ΣΥΓΧΡΗΜΑΤΟΔΟΤΗΣΗ
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Η ΠΑΙΔΕΙΑ ΣΤΗΝ ΚΟΡΥΦΗ
Επιχειρησιακό Πρόγραμμα
Εκπαίδευσης και Αρχικής
Επαγγελματικής Κατάρτισης

Επιστημονική Ευθύνη	Νικόλαος Ανδρεδάκης, Ομότιμος Καθηγητής Παν. Αθηνών
Συγγραφή	Παναγιώτης Μαμαλής, Θέμις Καψή, Ευάγγελος Τόλης, Στέλιος Μιχαήλογλου, Γιάννης Πρίντεζης

*Το παρόν εκπαιδευτικό υλικό παράχθηκε στο πλαίσιο του Έργου **«Κέντρα Εκπαίδευσης Ενηλίκων ΙΙ»**, το οποίο εντάσσεται στο **Ε.Π.Ε.Α.Ε.Κ. ΙΙ** του **ΥΠ.Ε.Π.Θ,**, Μέτρο 1.1. Ενέργεια 1.1.2.Β. και συγχρηματοδοτείται από την **Ευρωπαϊκή Ένωση (Ε.Κ.Τ.).***

ΕΛΛΑΔΑ 2008
Ανάπτυξη παντού. Ανάπτυξη για όλους.

ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΕΑΕΚ

ΕΥΡΩΠΑΪΚΗ ΕΝΩΣΗ
ΣΥΓΧΡΗΜΑΤΟΔΟΤΗΣΗ
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Η ΠΑΙΔΕΙΑ ΣΤΗΝ ΚΟΡΥΦΗ
Επιχειρησιακό Πρόγραμμα
Εκπαίδευσης και Αρχικής
Επαγγελματικής Κατάρτισης

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ ΣΤΗ ΣΤΑΤΙΣΤΙΚΗ

Βασικές έννοιες:
- status
- συστηματική συλλογή στοιχείων
- βιομηχανία, εμπόριο
- στάδια στατιστικής έρευνας

Στόχος του μαθήματος: Εισαγωγή στην έννοια της στατιστικής.

Προσδοκώμενα αποτελέσματα: μέσα από σύντομη ιστορική αναδρομή της επιστήμης της στατιστικής ο εκπαιδευόμενος να αντιληφθεί την αναγκαιότητά της επιστήμης αυτής.

Ο όρος «Στατιστική» χρησιμοποιείται, για να δηλώσει αριθμητικές πληροφορίες. Οι στατιστικές μελέτες παρουσιάζουν τα αποτελέσματα της συστηματικής συλλογής ενός συνόλου αριθμητικών πληροφοριών. Η λέξη «Στατιστική» μπορεί να προέρχεται από την αρχαία ελληνική λέξη στατίζω που σημαίνει (τοποθετώ, ταξινομώ, συμπεραίνω) ή από τη λατινική λέξη «status» που σημαίνει (πολιτεία, κράτος).

Η αρχαιότερη ίσως συλλογή στοιχείων θεωρείται η απογραφή πληθυσμού που έγινε το 2238 π.Χ. στην Κίνα από τον αυτοκράτορα Yao. Αργότερα στοιχειώδεις απογραφές φαίνεται να έχουν πραγματοποιηθεί από τους Αιγυπτίους και τους Πέρσες. Στην αρχαιότητα, η συγκέντρωση στατιστικών στοιχείων είχε ως στόχο τον εντοπισμό των πολιτών που είχαν υποχρέωση: να υπηρετήσουν ως πολεμιστές ή να πληρώσουν φόρο.

Μάλιστα, η συστηματική συλλογή στοιχείων απασχόλησε και τους κατοίκους διαφόρων χωρών της Ευρώπης. Ο μεγάλος αριθμός θανάτων που οφειλόταν σε πολέμους, λιμοκτονίες, επιδημικές ασθένειες και διάφορες άλλες αιτίες είχε επιπτώσεις στον πληθυσμό και στην οικονομία. Το 1620 ο Άγγλος Graunt σε μια δειγματοληπτική έρευνα που έκανε σεοικογένειες του Λονδίνου διαπίστωσε ότι σε κάθε 88 άτομα υπήρχαν 3 θάνατοι. Χρησιμοποιώντας τους επίσημους καταλόγους, οι οποίοι έδιναν 13.200 θανάτους το 1620, εκτίμησε ότι ο πληθυσμός του Λονδίνου το έτος αυτό κυμαίονταν στα 387.200 άτομα.

Από το 16ο έως το 19ο αιώνα, η αλματώδης ανάπτυξη του εμπορίου ώθησε τις ηγεσίες των διαφόρων κρατών στη μελέτη οικονομικών δεδομένων, όπως την παραγωγικότητα των βιομηχανιών, το εξαγωγικό εμπόριο, κ.τ.λ. Ενώ αρχικά η Στατιστική ασχολείτο μόνο με την παράθεση τεραστίων αριθμητικών πινάκων, σήμερα μια στατιστική έρευνα χωρίζεται στα εξής στάδια:

- **Συλλογή δεδομένων**
- **Έλεγχος δεδομένων** (καταμέτρηση, διάταξη, διόρθωση λαθών, συμπλήρωση ελλιπών στοιχείων).
- **Παρουσίαση δεδομένων** (πίνακες, διαγράμματα)
- **Επεξεργασία των δεδομένων** με σκοπό την εξαγωγή χρήσιμων συμπερασμάτων.

Ο κλάδος της Στατιστικής που ασχολείται με τα δύο πρώτα στάδια λέγεται **σχεδιασμός πειραμάτων** (experimental design). Ο κλάδος που ασχολείται με το τρίτο στάδιο είναι η **περιγραφική στατιστική** (descriptive statistics).

Και τέλος η **επαγωγική στατιστική** περιλαμβάνει τις μεθόδους με τις οποίες γίνεται η προσέγγιση των χαρακτηριστικών ενός μεγάλου συνόλου δεδομένων, από την μελέτη των χαρακτηριστικών ενός μικρού υποσυνόλου των δεδομένων.

Σε κάθε χώρα έχουν δημιουργηθεί αυτοτελείς στατιστικοί οργανισμοί με σκοπό τον αποτελεσματικό συνονισμό όλων των στατιστικών εργασιών. Τέτοιος οργανισμός στην Ελλάδα είναι Ε.Σ.Υ.Ε. (Εθνική Στατιστική Υπηρεσία) (http://www.statistics.gr/). Οι στατιστικές που πραγματοποιεί η Ε.Σ.Υ.Ε. είναι μηνιαίες, τριμηνιαίες, ετήσιες, ανά 5ετία και ανά 10ετία, και καλύπτουν όλους σχεδόν τους τομείς δραστηριότητας. Πληθυσμιακά στοιχεία, στοιχεία απασχόλησης και ανεργίας, στοιχεία που αφορούν την υγεία, την κοινωνική ασφάλιση, την παιδεία, την παραγωγική διαδικασία, τις τιμές, το εθνικό εισόδημα κ.τ.λ.

Βασικός χρήστης των στατιστικών και δεικτών που καταρτίζει η Ε.Σ.Υ.Ε. είναι το Κράτος, που χρησιμοποιεί τα στοιχεία αυτά για να χαράξει τις επιμέρους πολιτικές στους διάφορους τομείς. Η Ευρωπαϊκή Ένωση με τη βοήθεια του ευρωπαϊκού στατιστικού γραφείου EUROSTAT (http://europa.eu.int/comm/eurostat/) χρειάζεται τα επιμέρους στοιχεία των κρατών-μελών, για να συνθέσει τις ευρωπαϊκές στατιστικές. Επίσης, υπάρχουν και διεθνείς οργανισμοί όπως η UNESCO με το στατιστικό της ινστιτούτο (http://www.uis.unesco.org) με αντικείμενο τη συλλογή, παρουσίαση και την επεξεργασία αριθμητικών πληροφοριών για τα επί μέρους κράτη και τις μεταξύ τους οικονομικές σχέσεις.

Βασικές έννοιες της Στατιστικής έχουν εισχωρήσει και ενσωματωθεί σε όλες τις επιστήμες (Ανθρωπιστικές, Κοινωνικές, Οικονομικές, Τεχνολογικές, Επιστήμες Υγείας, κ.ά.). Η ανάλυση σύνθετων φαινομένων απαιτεί από τους στατιστικούς τη χρησιμοποίηση των καταλληλότερων στατιστικών μεθόδων στις μεγάλες απαιτήσεις των παραπάνω επιστημονικών χώρων.

Θα εξετάσουμε στην συνέχεια του βιβλίου τα κυριότερα κεφάλαια των Μαθηματικών που αφορούν τη Στατιστική των οποίων η γνώση, εκτός από απαραίτητη, προσφέρει στον εκπαιδευόμενο μία βαθύτερη κατανόηση της στατιστικής ως θεμελιωμένης επιστήμης.

2. ΣΥΛΛΟΓΗ ΣΤΟΙΧΕΙΩΝ

Βασικές έννοιες:
- στατιστική μονάδα
- πληθυσμός
- μεταβλητές
- παρατήρηση

Στόχος του μαθήματος:
Εισαγωγή στις παραπάνω έννοιες που αφορούν τη συλλογή δεδομένων και τη στατιστική

Προσδοκώμενα αποτελέσματα: ο εκπαιδευόμενος μαθαίνει τις μεθόδους συλλογής στοχείων και είναι σε θέση να διαχωρίζει τις τυχαίες μεταβλητές.

2.1. ΣΤΑΤΙΣΤΙΚΗ ΜΟΝΑΔΑ - ΣΤΑΤΙΣΤΙΚΟΣ ΠΛΗΘΥΣΜΟΣ

Η **στατιστική μονάδα** μπορεί να είναι ένα αντικείμενο, ένα άτομο, μια εταιρεία, ένα ίδρυμα ή κάποιο γεγονός (εκλογική αναμέτρηση) και γενικά είναι αυτό από το οποίο λαμβάνουμε τις πληροφορίες που επιθυμούμε να επεξεργαστούμε και να αναλύσουμε στατιστικά.

Οι στατιστικές μονάδες μπορεί να είναι **απλές** (ένα άτομο, ένα αντικείμενο, μια μέρα) είναι, όμως, δυνατό να είναι και **συνθέτες** και να αποτελούνται από περισσότερα αντικείμενα ή πρόσωπα, όπως η οικογένεια, η μηνιαία ή ετήσια παραγωγή μιας βιομηχανίας κ.λ.π.

Το σύνολο των στατιστικών μονάδων, των οποίων επιθυμούμε τη μελέτη ενός ή περισσοτέρων συγκεκριμένων χαρακτηριστικών, ονομάζεται **πληθυσμός ή στατιστικός πληθυσμός**.

Ένα από τα βασικά στοιχεία που πρέπει να ορίζουμε για τον πληθυσμό είναι τα όριά του, δηλαδή ποιες είναι ακριβώς οι στατιστικές μονάδες του πληθυσμού που θα μελετηθούν.

Για παράδειγμα, αν θελήσει να μελετήσει ορισμένα χαρακτηριστικά των κατοίκων της Ραφήνας, δεν θα πρέπει να κάνει την έρευνα στο λιμάνι μια Παρασκευή απόγευμα ή Σάββατο πρωί, ή ένα Σαββατοκύριακο τους θερινούς μήνες. Δεν αρκεί, λοιπόν, μόνο ο προσδιορισμός των γεωγραφικών ορίων μιας πόλης.

Ο στατιστικός πληθυσμός μπορεί να είναι άπειρος, όπως η παραγωγή ενός προϊόντος, οι γεννήσεις βρεφών σε μία πόλη ή πεπερασμένος όπως οι αφίξεις και αναχωρήσεις των αεροσκαφών μια συγκεκριμένη μέρα στο αεροδρόμιο Ελ. Βενιζέλος.

2.2. ΜΕΘΟΔΟΙ ΣΥΛΛΟΓΗΣ ΣΤΟΙΧΕΙΩΝ

Οι διάφορες μέθοδοι συλλογής στατιστικών στοιχείων συνεχίζονται σε δύο μεγάλες ομάδες: τις απογραφές και τις δειγματοληπτικές έρευνες.

Απογραφή ονομάζεται η διαδικασία με την οποία συλλέγονται οι παρατηρήσεις όλων των μονάδων ενός πληθυσμού σε μια συγκεκριμένη χρονική στιγμή.

Μπορεί να έχουμε τις:

- Δημογραφικές απογραφές στις οποίες συλλέγονται στοιχεία σχετικά με το φύλο, την ηλικία, το επάγγελμα, κ.τ.λ.
- Οικονομικές απογραφές στις οποίες συγκεντρώνονται στοιχεία σχετικάμε την οικονομική κατάσταση.
- Βιομηχανικές απογραφές, στις οποίες συλλέγουμε πληροφορίες σχετικά με την οικονομική δραστηριότητα των βιομηχανιών, το αριθμό των απασχολούμενων, το επίπεδο μηχανοργάνωσης, κ.τ.λ.
- Γεωργικές απογραφές, στις οποίες συγκεντρώνουμε στοιχεία που αφορούν τις εκτάσεις που καλλιεργούνται, το είδος της γεωργικ΄ς παραγωγής, τον αριθμό των γεωργικών μηχανημάτων, κ.τ.λ.

Στην Ελλάδα η Ε.Σ.Υ.Ε. πραγματοποιεί απογραφή του πληθυσμού κάθε 10 χρόνια (τελευταία τον Μάρτιο του 2001).

Τα μειονεκτήματα των απογραφών είναι:

- **Το μεγάλο κόστος**, καθώς χρηάζεται ειδική προεργασία και μεγάλο αριθμό απογραφέων.
- **Τη μη ύπαρξη πολλών εξειδικευμένων ατόμων**, που έχει ως συνέπεια την συγκέντρωση εσφαλμένων στοιχείων τα οποία μπορεί να δώσουν λανθασμένη εικόνα της διόρθωσης του πληθυσμού.
- **Τη μη επίκαιρη έκδοση των αποτελεσμάτων**, λόγω του μεγάλου όγκου των πληροφοριών.

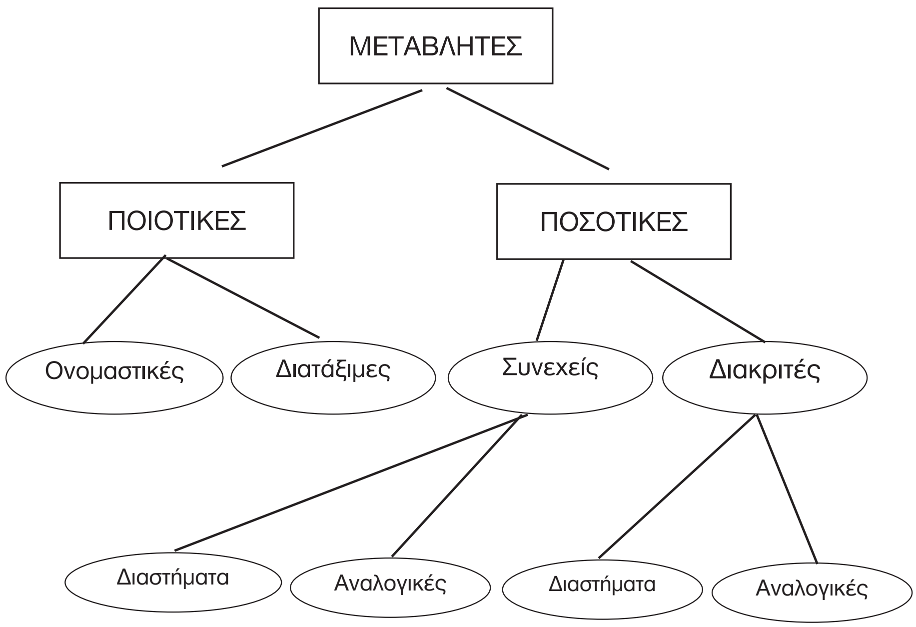
Δειγματοληψία είναι απογραφή ορισμένων συγκεκριμένων χαρακτηριστικών ενός τμήματος του πληθυσμού. Το τμήμα του πληθυσμού που απογράφεται ονομάζεται δείγμα. Η δειγματοληψία πραγματοποιείται σε συγκεκριμένα στάδια ανάλογα με τη μέθοδο που επιλέγουμε. Εκτενέστερη αναφορά πραγματοποιείται στο 10^{ο} κεφάλαιο, όπου εξετάζεται αποκλειστικά η έννοια της δειγματολειψίας.

Τα συμπεράσματα, που θα προκύψουν από την μελέτη ενός δείγματος, θα ισχύουν με ικανοποιητική ακρίβεια για ολόκληρο τον πληθυσμό, αν το δείγμα είναι **αντιπροσωπευτικό,** δηλαδή αν έχει επιλεγεί κατά τέτοιο τρόπο, ώστε κάθε μονάδα του πληθυσμού να έχει την ίδια δυνατότητα να επιλεγεί.

Ενδεικτικό είναι το ακόλουθο παράδειγμα στις επιλογές των Η.Π.Α. το 1936. Το περιοδικό Literary Digest χρησιμοποιώντας δείγμα 2.400.000 ατόμων πρόβλεψε νίκη του Landon με ποσοστό 57%. Αντίθετα, το δημοσκοπικό γραφείο του G. Gallup χρησιμοποιώντας δείγμα 50.000 ατόμων πρόβλεψε το σωστό αποτέλεσμα που ήταν νίκη των Roosvelt με ποσοστό 62%.

Οι αρχές και οι μεθόδοι επιλογής του δείγματος είναι αντικείμενο της **Δειγματοληψίας**.

Διάγραμμα ταξινόμησης μεταβλητών



2.3. ΜΕΤΑΒΛΗΤΕΣ

Τυχαία μεταβλητή ή απλά μεταβλητή ονομάζουμε οποιοδήποτε χαρακτηριστικό ως προς το οποίο εξετάζουμε ένα πληθυσμό και το οποίο παίρνει μια ή περισσότερες διαφορετικές τιμές. Για παράδειγμα το φύλο, η ηλικία, το ύψος, η τιμή του πετρελαίου σε διαφορετικές χρονικές στιγμές.

Οι μεταβλητές συμβολίζονται με κεφαλαία γράμματα Χ, Ψ, Ζ και ο όρος **παρατηρηθείσα τιμή ή παρατήρηση** χρησιμοποιείται για την αριθμητική ή άλλη συμβολική της έκφραση.

Οι μεταβλητές διακρίνονται σε δύο μεγάλες κατηγορίες **Ποιοτικές** και **Ποσοτικές**.

Ποιοτικές μεταβλητές είναι αυτές που αναφέρονται σε κάποιο ποιοτικό χαρακτηριστικό και οι τιμές τους δεν είναι αριθμητικές Για παράδειγμα επίπεδο εκπαίδευσης, μητρική γλώσσα, βιοτικό επίπεδο, κ.τ.λ.

Επιπλέον οι ποιοτικές μεταβλητές χωρίζονται σε:

- **Ονομαστικές** μεταβλητές, οι οποίες επιδέχονται μόνο αυθαίρετη κατάταξη όπως π.χ. φύλο, θρησκεία, κ.τ.λ.
- **Διατάξιμες** μεταβλητές, οι οποίες επιδέχονται μέτρηση ανωτέρου επιπέδου που επιτρέπει την ιεράρχησή τους, όπως π.χ. χαρακτηρισμός πτυχίου (άριστα, λίαν καλώς, καλώς), σοβαρότητας μιας ασθένειας (ήπια, μέτρια, σοβαρή), της γνώμης για κάποιο μέτρο (διαφωνώ πλήρως, διαφωνώ σε κάποια σημεία, συμφωνώ, συμφωνώ πλήρως).

Ποσοτικές είναι οι μεταβλητές των οποίων οι τιμές είναι αριθμητικές και επιδέχονται μέτρηση. Για παράδειγμα το εισόδημα, το βάρος, το ύψος, ο αριθμός των παιδιών μιας οικογένειας.

Οι ποσοτικές μεταβλητές διακρίνονται σε δύο κατηγορίες, τις:

- **Διακριτές**, οι οποίες παίρνουν μόνο «μεμονωμένες» αριθμητικές τιμές όπως για παράδειγμα, το νούμερο των ανδρικών υποδημάτων, ο αριθμός των παιδιών μιας οικογένειας, ο αριθμός των ελαττωμάτων ενός προϊόντος.
- **Συνεχείς** οι οποίες μπορούν να πάρουν οποιαδήποτε τιμή μέσα από ένα συνεχές διάστημα, όπως για παράδειγμα το βάρος, το ύψος, η διάρκεια μιας τηλεφωνικής συνδιάλεξης.

Παρατήρηση: Ο διαχωρισμός μεταξύ διακριτών και συνεχών μεταβλητών είναι δύσκολος στην πράξη λόγω των περιορισμών που επιβάλουν τα όργανα μέτρησης. Έτσι για π.χ. το βάρος ενός συνόλου ατόμων επειδή μετριέται με ακρίβεια γραμμαρίου μας δίνει διακριτές τιμές, όπως 63, 512 Kg 67, 383 Kg κ.τ.λ.

Οι ποσοτικές μεταβλητές διακρίνονται επιπλέον και σε δύο άλλες κατηγορίες, τις:

- Μεταβλητές διαστήματος, όπως π.χ. βαθμός πτυχίου (5, 10).
- Αναλογικές μεταβλητές, όπως για π.χ. χιλιομετρική απόσταση.

Εφαρμογή 1

Τα αποτελέσματα των εξετάσεων των φοιτητών του Μαθηματικού τμήματος στο μάθημα της Στατιστικής ήταν τα ακόλουθα: 2, 3, 3, 4, 4, 5, 7, 5, 5, 9. Να βρεθεί:

i) Ποιος είναι ο πληθυσμός;

ii) Ποια είναι τα άτομά;

iii) Ποιες είναι οι παρατηρήσεις;

iv) Ποια είναι η μεταβλητή και σε ποια κατηγορία ανήκει;

v) Ποιες είναι οι τιμές των μεταβλητών;

ΛΥΣΗ

i) Ο πληθυσμός είναι οι 10 φοιτητές του Μαθηματικού τμήματος.

ii) Κάθε φοιτητής είναι ένα άτομο.

iii) Οι παρατηρήσεις είναι: 2, 3, 3, 4, 4, 5, 7, 5, 5, 9.

iv) Η μεταβλητή είναι «ο βαθμός στη Στατιστική» η οποία είναι ποσοτική, διακριτή και μεταβλητή διαστήματος.

v) Οι τιμές της μεταβλητής είναι 2, 3, 4, 5, 7, 9.

Εφαρμογή 2

Εξετάζουμε τους κατοίκους μιας πόλης ως προς τα παρακάτω χαρακτηριστικά:

i) Φύλο ii) ύψος iii) μορφωτικό επίπεδο iv) εισόδημα v) θρήσκευμα

Να χαρακτηρίσετε τις παραπάνω μεταβλητές.

ΛΥΣΗ

i) Το «Φύλο» είναι ποιοτική ονομαστική μεταβλητή.

ii) Το «ύψος» είναι ποσοτική συνεχής μεταβλητή διαστήματος

iii) Το «μορφωτικό επίπεδο» είναι ποιοτικά διατάξιμη μεταβλητή.

iv) Το «εισόδημα» είναι ποσοτική συνεχής αναλογική μεταβλητή.

v) Το «θρήσκευμα» είναι ποιοτική ονομαστική μεταβλητή.

ΕΞΑΣΚΗΣΗ

1. Να αντιστοιχίσετε τις παρακάτω μεταβλητές της στήλης Α με τις παρατηρηθείσες τιμές στη στήλη Β.

Στήλη Α	Στήλη Β
Μεταβλητή	Παρατηρηθείσα τιμή
Θερμοκρασία περιβάλλοντος	Ρουμανικά
Μητρική γλώσσα	Καλή
Κατάσταση υγείας	19^0 C
Χαρακτηρισιμός πτυχίου	Εποχιακή
Εργασιακή κατάσταση	Απόφοιτος Τ.Ε.Ι.
Τύπος απασχόλησης	Καλώς

2. Να συνδέσετε με μία γραμμή κάθε μεταβλητήτης στήλης Α με τον αντίστοιχο χαρακτρισμό της στη στήλη Β.

Στήλη Α	Στήλη Β
Μηνιαίο εισόδημα	Ποιοτική, διατάξιμη
Μονάδες εισαγωγής	
σε ΑΕΙ-ΤΕΙ	Ποσοτική, διακριτική, διαστήματος
Βιοτικό επίπεδο	Ποσοτική, διακριτή, αναλογική
Θρήσκευμα	Ποσοτική, συνεχής, αναλογική
Αριθμός δωματίων κατοικίας	Ποιοτική, συνεχής, διαστήματος
Χρόνια σπουδών	Ποιοτική, ονομαστική

3. Σε μια δειγματοληπτική έρευνα για το βάρος των εμπορευμάτων μιας αποθήκης φρούτων, βρέθηκαν 10 κιβώτια με τα εξής κιλά:

   15, 23, 30, 15, 14, 10, 18, 10, 9, 20

   Να βρείτε: i) Ποιος είναι ο πληθυσμός;

   ii) Ποια είναι η μεταβλητή και σε ποια κατηγορία ανήκει;

   iii) Ποιες είναι οι παρατηρήσεις;

   iv) Ποιες είναι οι τιμές των μεταβλητών;

4. Για να βρούμε την άποψη των κατοίκων μιας χώρας για την ένταξη της χώρας στην Ο.Ν.Ε., παίρνουμε δείγμα 1000 ατόμων. Ποιος είναι ο καλύτερος κατά τη γνώμη σας τρόπος επιλογής του δείγματος;

1. Εξετάζουμε μόνο ενηλίκους.
2. Εξετάζουμε μόνο άτομα από επαρχιακές πόλεις.
3. Εξετάζουμε μόνο άτομα υψηλού μορφωτικού επιπέδου.
4. εξετάζουμε άτομα από διαφορετικές περιοχές και ηλικίες.
5. Εξετάζουμε ενήλικες με ποσόστωση σε μορφωτικό επίπεδο αντίστοιχη της συνολικής διαστρωμάτωσης του πληθυσμού και από διαφορετικές περιοχές.

5. Από ένα σύνολο 100 μαθητών (60 αγόρια και 40 κορίτσια) επιλέγουμε ένα δείγμα 15 μαθητών (9 αγόρια και 6 κορίτσια). Είναι το δείγμα αντιπροσωπευτικό;
Δικαιολογήστε την απάντησή σας.

6. Να ορίσετε από μια μεταβλητή για καθένα από τα παρακάτω δείγματα, ως προς την οποία μπορούμε να τα μελετήσουμε: α) οι τηλεθεατές σε μια κοινότητα, β) οι φίλαθλοι μιας ομάδας, γ) οι μαθητές μιας τάξης του Λυκείου.

7. Υποθέτουμε ότι ο πρόεδρος μιας ποδοσφαιρικής ομάδας θέλει να μάθει την γνώμη των φιλάθλων της ομάδας για τον τόπο κατασκευής ενός νέου γηπέδου. Αν ρωτήσει 150 άτομα που παρακολουθούν τον αγώνα μπάσκετ της ομάδας του ίδιου σωματίου, θα μπορούσε να αποκτήσει μια καλή ιδέα για την γνώμη των φιλάθλων; Εξηγήστε την απάντησή σας χρησιμοποιώντας τους όρους «πληθυσμός» και «δείγμα».

Η εκμάθηση των βασικών εννοιών της Στατιστικής γίνεται πιο ενδιαφέρουσα, όταν μπορούμε να αργαστούμε με πληροφορίες που συλλέγονται από εμάς. Η απλούστερη συλλογή δεδομένων μπορεί να αφορά πληροφορίες για τους συμμαθητές σας. Ξεκινήστε την έρευνα σας και κρατήστε τα δεδομένα που θα συλλέξετε. Στο τέλος κάθε παραγράφου θα υπάρχουν ερωτήσεις που θα αφορούν τα δεδομένα αυτά.

ΔΡΑΣΤΗΡΙΟΤΗΤΑ

Κάθε μαθητής να κατασκευάσει ένα έντυπο ερωτηματολογίου, να μοιραστεί σε κάθε συμμαθητή του και να συλλέξει απαντήσεις, ανώνυμα, που μπορεί να αφορούν:
- Φύλο
- Ηλικία
- Ύψος (σε cm)
- Βάρος (σε kg)
- Χρώμα μαλλιών
- Χρώμα ματιών
- Είναι δεξιόχειρας, αριστερόχειρας ή αμφίχειρας
- Αριθμό των αδελφών

Προσθέστε ο καθένας από εσάς τρεις τουλάχιστον ερωτήσεις για μεταβλητές που σας ενδιαφέρουν. Κατασκευάστε το ερωτηματολόγιο με τέτοιο τρόπο που να είναι ελκυστικός για τους συμμαθητές σας να απαντήσουν και εύκολος για εσάς στη συλλογή δεδομένων.

Σύνοψη

Μία απλή απαρίθμηση των εφαρμογών της Στατιστικής που είναι βασικά μια εφαρμοσμένη επιστήμη, αποδεικνύει ότι αυτή χρησιμοποιείται σε όλους σχεδόν τους τομείς της ανθρώπινης δραστηριότητας. Στην παράγραφο αυτή ο εκπαιδευόμενος μαθαίνει για τη βασική έννοια της Στατιστικής που είναι η Δειγματοληψία καθώς και για τα απαραίτητα εργαλεία της προκειμένου να πραγματοποιηθεί που είναι ο πληθυσμός και οι μεταβλητές.

Βιβλιογραφία / Internet

«*Στατιστική*», Α. Πέτρος Κιόχος

«*Πιθανότητες και Στατιστική*», Spiegel M. R., μετάφραση Σ. Περσίδη, Αθήνα

«Στατιστική» Δρακάτου Κ., Αθήνα, 1984

«Στατιστική» Κάκουλλου Θ., Αθήνα, 1972

«Στατιστική» Φραγκάκη Χαρ., Θεσσαλονίκη, 1985

«Πιθανότητες» Κάκουλλου, Αθήνα

www.statistics.gov.uk: *επίσημο site της Μ. Βρετανίας για την κοινωνία, την αγορά εργασιάς, την οικονομία. (Home of official U.K. statistics)*

www.en. wikipedia.org/wiki/statistics: *ο όρος statistics στην ηλεκτρονική εγκυκλοπαίδεια Wikipedia.*

ΟΔΗΓΟΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΜΕΛΕΤΗ

«*Μαθηματικά και Στοιχεία Στατιστικής*», Εκδόσεις Σταμούλης, Γιώργος Καλαϊτζής

Στο βιβλίο αυτό ο ενδιαφερόμενος μπορεί να βρει βασικές γνώσεις Μαθηματικών απαραίτητες στην επιστήμη της Στατιστικής, υπάρχουν 415 ασκήσεις, 152 ερωτήσεις, σύνοψη, ανάπτυξη καθώς και ερωτήσεις κατανόησης, ώστε να υπάρχει η καλύτερη δυνατή κατανόηση βασικών εννοιών.

«*Περιγραφική Στατιστική*», Βασ. Μπένος

Ο ενδιαφερόμενος για θέματα Στατιστικής μπορεί στις σελίδες του βιβλίου να βρει τα βασικά θέματα που αφορούν της επιστήμη της Στατιστικής όπως πληθυσμός, δείγμα., μέτρα θέσης και διασποράς, είδη αριθμοδείκτων, πλαισιωμένα με εφαρμογές για την καλύτερη κατανόηση τους.

3. ΠΑΡΟΥΣΙΑΣΗ ΣΤΟΙΧΕΙΩΝ

Βασικές έννοιες:

- κατανομή
- συχνότητα
- σχετική συχνότητα
- αθροιστική συχνότητα
- αθροιστική σχετική συχνότητα
- πίνακες - διαγράμματα
- ομαδοποίηση

Στόχος του μαθήματος:

Εισαγωγή και εξάσκηση στην παρουσίαση των στατιστικών δεδομένων.

Προσδοκώμενα αποτελέσματα: ο εκπαιδευόμενος εξοικειώνεται με τους τρόπους παρουσίασης στατιστικών στοιχείων, μαθαίνει να τους κατασκευάζει και να τους ερμηνεύει.

Εισαγωγικές παρατηρήσεις: μετά τη συλλογή των δεδομένων, όπως αυτή περιγράφηκε στο προηγούμενο κεφάλαιο, το επόμενο βήμα είναι η παρουσίαση και οπτικοποίηση των στατιστικών δεδομένων. Αναγκαία, λοιπόν, είναι η κατασκευή συνοπτικών πινάκων ή γραφικών παραστάσεων με τέτοιο τρόπο, ώστε να μπορεί ο εκπαιδευόμενος να αναλύει εύκολα τα στοιχεία και να εξάγει άμεσα χρήσιμα συμπεράστατα.

3.1. ΚΑΤΑΝΟΜΗ ΣΥΧΝΟΤΗΤΩΝ

Από την εξέταση ενός δείγματος 50 οικογενειών, ως προς τον αριθμό των παιδιών τους, προέκυψαν τα παρακάτω δεδομένα:

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

Αν μας ζητηθεί να δώσουμε απαντήσεις στα εξής ερωτήματα:

i) Πόσες οικογένειες έχουν δύο παιδιά;

ii) Πόσες οικογένειες δεν έχουν παιδιά;

iii) Ποιος αριθμός παιδιών είναι ο πιο συνηθισμένος;

Οι απαντήσεις στα ερωτήματα αυτά μπορούν να βρεθούν από την αρχική λίστα των δεδομένων που έχουμε. Δεν είναι εύκολο όμως να δώσουμε άμεσα απαντήσεις. Αν μάλιστα το δείγμα ήταν αρκετό μεγαλύτερο, τότε θα χρειαζόμασταν αρκετή ώρα για να απαντήσουμε με ακρίβεια στα παρακάτω ερωτήματα.

Είναι αναγκαία λοιπόν η κατασκευή στατιστικών πινάκων. Στους στατιστικούς πίνακες οι πληροφορίες τοποθετούνται σε γραμμές και στήλες με τέτοιο τρόπο ώστε να διευκολύνεται η σύγκριση των στοιχείων που παρουσιάζονται και η εξαγωγή συμπερασμάτων για το δείγμα του πληθυσμού που μελετούμε.

Οι πίνακες διακρίνονται σε:

ΓΕΝΙΚΟΥΣ ΠΙΝΑΚΕΣ	ΕΙΔΙΚΟΥΣ ΠΙΝΑΚΕΣ
που περιέχουν όλες τις πληροφρίες της έρευνας, τις περισσότερες φορές με αρκετά λεπτομερή στοιχεία	που είναι συνοπτικοί, σαφείς και αναφέρονται σε συγκεκριμένες πληροφορίες της έρευνας, χρησιμοποιώντας ορισμένες φορές κάποια στοιχεία από τους γενικούς πίνακες.

Κάθε πίνακας πρέπει να έχει τα παρακάτω στοιχεία:

- **Τον τίτλο**, ο οποίος δηλώνει συνοπτικά το περιεχόμενο τον πίνακα.
- **Τις επικεφαλίδες**, γραμμών και στηλών που δείχνουν συνοπτικά τη φύση και τις μονάδες μέτρησης των δεδομένων
- **Το κύριο σώμα**, που περιέχει διαχωρισμένα μέσα στις γραμμές και στις στήλες τα στατιστικά δεδομένα.
- **Την πηγή**, που αναγράφεται στο κάτω μέρος του πίνακα και δείχνει την προέλευση των στατιστικών στοιχείων.

Δύο παραδείγματα τέτοιων πινάκων δίνονται παρακάτω:

Χ. ΜΕΤΑΦΟΡΕΣ – ΕΠΙΚΟΙΝΩΝΙΕΣ – ΤΟΥΡΙΣΜΟΣ

Χα. Χερσαίες μεταφορές

	1971	1981	1991	2001	2002	2003
Επιβάτες *(χιλιάδες)*:						
Σιδηροδρόμων	13.256	10.388	12.253	13.935	14.288	8.004
Λεωφορείων	1.221.419	1.020.325	902.978	927.354	...	...
Αστικών γραμμών	*1.061.830*	*846.252*	*755.533*	*795.601*	...	...
Υπεραστικών »	*159.589*	*174.073*	*147.445*	*131.753*	...	...
Εμπορεύματα σιδηροδρόμων *(χιλ. τόνοι)*	3.358	2.995	3.542	2.784	2.027	2.592
Κυκλοφορούντα οχήματα	**414.046**	**1.477.214**	**2.888.009**	***5.389.996**	***5.693.008**	***5.967.610**
Επιβατηγά	226.893	912.385	1.777.484	3.423.704	3.646.069	3.839.549
Λεωφορεία	10.546	17.367	22.080	27.115	27.247	27.139
Φορτηγά	107.361	441.081	792.770	1.085.811	1.109.137	1.131.027
Μοτοσυκλέτες	69.246	106.381	295.675	853.366	910.555	969.895
Οδικά τροχαία ατυχήματα	**17.950**	**19.841**	**20.764**	**19.671**	**16.810**	**15.751**
Παθόντα πρόσωπα	25.801	29.061	30.739	25.216	24.093	22.342
Από αυτά, νεκροί	*943*	*1.354*	*1.790*	*1.880*	*1.634*	*1.605*

Πηγή: ΕΣΥΕ

V. ΓΕΩΡΓΙΑ – ΚΤΗΝΟΤΡΟΦΙΑ – ΔΑΣΗ – ΑΛΙΕΙΑ

Va. Εκμεταλλεύσεις, αγροτεμάχια, χρησιμοποιούμενη γεωργική έκταση, μέση έκταση αγροτεμαχίου

*(Προσωρινά στοιχεία της **Έρευνας** Διάρθρωσης Γεωργικών και Κτηνοτροφικών εκμεταλλεύσεων, 2003)*

εκτάσεις σε χιλ. στρέμματα

Περιφέρειες	Εκμεταλλεύσεις		Αγροτεμάχια	Χρησιμοποιούμενη γεωργική έκταση	Μέση έκταση αγροτεμαχίου
	Σύνολο	Με χρησιμοποιούμενη γεωργική έκταση			
Σύνολο Ελλάδος	**824.464**	**818.467**	**5.402.117**	**39.678**	**7,34**
Ανατολική Μακεδονία και Θράκη	64.948	63.841	585.585	3.915	6,92
Κεντρική Μακεδονία	117.596	116.469	892.127	6.954	7,79
Δυτική Μακεδονία	30.137	29.648	379.431	2.456	6,47
Θεσσαλία	79.721	78.204	390.109	4.206	10,78
Ήπειρος	42.974	42.390	186.876	1.205	6,45
Ιόνιοι Νήσοι	31.843	31.840	177.990	1.126	6,32
Δυτική Ελλάς	93.629	93.322	382.520	3.448	9,01
Στερεά Ελλάς	79.994	79.762	489.119	3.885	7,94
Πελοπόννησος	104.864	104.757	648.540	4.118	6,35
Αττική	26.369	26.044	117.010	628	5,37
Βόρειο Αιγαίο	33.126	33.111	261.508	1.978	7,57
Νότιο Αιγαίο	23.975	23.869	156.415	1.266	8,09
Κρήτη	95.287	95.219	754.888	4.493	5,95

Σημείωση: Τυχόν διαφορές μεταξύ των αθροισμάτων και των συνόλων οφείλονται σε στρογγυλοποιήσεις.

Πηγή: ΕΣΥΕ

Για να κατασκευάζουμε τον πίνακα συχνοτήτων για τον αριθμό των παιδιών που έχουν οι 50 οικογένειες που δώσαμε στην αρχή της παραγράφου, εργαζόμαστε ως εξής:

α) Στην πρώτη στήλη γράφουμε κατά σειρά μεγέθους τους αριθμούς των παιδιών που έχουν οι οικογένειες.

β) Στην δεύτερη στήλη «Διαλογή» σημειώνουμε με μία γραμμή (Ι) για κάθε αντίστοιχο αριθμό που συναντάμε, σχηματίζοντας πεντάδες (~~IIII~~)

γ) Στην τρίτη στήλη «Συχνότητα» γράφουμε τον φυσικό αριθμό που βρίσκουμε από την καταμέτρηση των γραμμών της δεύτερης στήλης.

Ο φυσικός αυτός αριθμός που φανερώνει πόσες φορές εμφανίζεται η τιμή x_i της μεταβλητής Χ στο σύνολο των παρατηρήσεων ονομάζεται συχνότητα ή απόλυτη συχνότητα και συμβολίζεται με v_i. Οπότε ο πίνακας κατανομής συχνοτήτων είναι:

Πίνακας 3

Αριθμός παιδιών x_i	Διαλογή	Συχνότητα v_i
0	~~IIII~~ III	8
1	~~IIII~~ ~~IIII~~ ~~IIII~~ II	17
2	IIII ~~IIII~~ III	13
3	~~IIII~~ II	7
4	~~IIII~~	5
	Σύνολο	**50**

Το άθροισμα των συχνοτήτων των τιμών μιας μεταβλητής Χ ενός δείγματος μεγέθους ν είναι ίσο με το μέγεθος του δείγματος.

Δηλαδή: $\boxed{v_1 + v_2 + v_3 + \ldots + v_\kappa = v}$

3.2. ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ

Από το προηγούμενο δείγμα των 50 οικογενειών διαπιστώσαμε ότι 17 οικογένειες έχουν 1 παιδί. Οι 17 οικογένειες αποτελούν τα $\frac{17}{50} = 0{,}34$ του δείγματος.

Ο αριθμός αυτός που συνήθως εκφράζεται επί τοις % ονομάζεται σχετική συχνότητα. Μπορούμε να πούμε ότι το 34% των 50 οικογενειών έχει 1 παιδί. Σχετική συχνότητα f_i της τιμής x_i είναι ο αριθμός, που προκύπτει αν διαιρέσουμε τη συχνότητα v_i με το μέγεθος του δείγματος.

Δηλαδή: $\boxed{f_i = \frac{v_i}{v},\ i = 1, 2, \ldots, \kappa}$

Για τη σχετική συχνότητα f_i, $i = 1, 2, \ldots, \kappa$ με $\kappa \leq v$, ισχύουν:

- $0 \leq f_i \leq 1$ (αφού 0£ vi £ v)
- $f_1 + f_2 + \ldots + f_\kappa = 1$

Οι σχετικές συχνότητες fi των τιμών x_i μιας μεταβλητής X, εκφράζονται συνήθως επί τοις εκατό και συμβολίζονται με f_i%.

Οπότε για τον πίνακα 3 έχουμε τον ακόλουθο πίνακα κατανομής σχετικών συχνοτήτων:

Πίνακας 4

Παρατήρηση:

Αριθμός παιδιών (x_i)	Συχνότητα v_i	Σχετική συχνότητα f_i	Σχετική συχνότητα f_i%
0	8	0,16	16
1	17	0,34	34
2	13	0,26	26
3	7	0,14	14
4	5	0,10	10
ΣΥΝΟΛΟ	**50**	**1**	**100**

Όταν το πλήθος ν των παρατηρήσεων είναι διαιρέτης ή πολλαπλάσιο του 100, μπορούμε να υπολογίσουμε τις σχετικές συχνότητες f_i, χωρίς να κάνουμε την διαίρεση, αλλά να προσπαθούμε με κατάλληλους πολλαπλασιασμούς ή διαιρέσεις να τρέπουμε το κλάσμα σε κλάσμα με παρανομαστή το 100. επίσης αν το ν δεν είναι διαιρέτης ή πολλαπλάσιο του 100, ενδέχεται να είναι τέτοιος αριθμός που με απλοποίηση να γίνει διαιρέτης του 100 και στη συνέχεια εργαζόμαστε όπως αρχικά.

Π.χ. 1: Να βρεθούν οι σχετικές συχνότητες:

x_i	v_i	f_i	f_i%
5	6	0,30	30
7	8	0,40	40
9	4	0,20	20
11	2	0,10	10
Σύνολο	**20**	**1**	**100**

$$\text{Εχουμε}: f_1 = \frac{v_1}{v} = \frac{6}{20} \overset{\cdot 5}{=} \frac{30}{100} = 0,30$$

$$f_2 = \frac{v_2}{v} = \frac{8}{20} \overset{\cdot 5}{=} \frac{40}{100} = 0,40$$

$$f_3 = \frac{v_3}{v} = \frac{4}{20} \overset{\cdot 5}{=} \frac{20}{100} = 0,20$$

$$f_4 = \frac{v_4}{v} = \frac{2}{20} \overset{\cdot 5}{=} \frac{10}{100} = 0,10$$

π.χ. 2: Να βρεθούν οι σχετικές συχνότητες

x_i	v_i	f_i	$f_i\%$
1	24	0,32	32
2	18	0,24	24
3	21	0,28	28
4	12	0,16	16
Σύνολο	**75**	**1**	**100**

$$\text{Έχουμε}: f_1 = \frac{v_1}{v} = \frac{24}{75} \overset{:3}{=} \frac{8}{25} \overset{\cdot 4}{=} \frac{32}{100} = 0,32$$

$$f_2 = \frac{v_2}{v} = \frac{18}{75} \overset{:3}{=} \frac{6}{25} \overset{\cdot 4}{=} \frac{24}{100} = 0,24$$

$$f_3 = \frac{v_3}{v} = \frac{21}{75} \overset{:3}{=} \frac{7}{25} \overset{\cdot 4}{=} \frac{28}{100} = 0,28$$

$$f_4 = \frac{v_4}{v} = \frac{12}{75} \overset{:3}{=} \frac{4}{25} \overset{\cdot 4}{=} \frac{16}{100} = 0,16$$

3.3. ΑΘΡΟΙΣΤΙΚΗ ΚΑΙ ΣΧΕΤΙΚΗ ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ

Σε ποσοτικές μεταβλητές, εκτός από τη συχνότητα f_i, χρησιμοποιούμε και τις αθροιστικές συχνότητες (N_i) και τις σχετικές αθροιστικές συχνότητες (F_i).

- Ονομάζουμε αθροιστική συχνότητα (N_i) της τιμής x_i την τιμή που εκφράζει το πλήθος των παρατηρήσεων του δείγματος, που είναι μικρότερες ή ίσες του x_i. Έχουμε:

- $N_i = v_i + v_2 + ... + v_i$, $\quad i = 1, 2, ..., κ$
- $N_i = N_{i-1} + v_i$

- Σχετική αθροιστική συχνότητα (F_i) είναι το ποσοστό των παρατηρήσεων του δείγματος που είναι μικρότερες ή ίσες της x_i.
  Ισχύει:

- $F_i = f_i + f_2 + ... + f_i$, $\quad i = 1, 2, ..., κ$
- $F_i = F_{i-1} + f_i$

Αν εκφράζεται επί τοις εκατό, τότε $F_i\% = f_1\% + f_2\% + ... + f_i\%$.

Εφαρμογή 1

Σε ένα τμήμα 25 μαθητών της Α΄ Λυκείου δόθηκε ένα τεστ Μαθηματικών και προέκυψαν τα παρακάτω αποτελέσματα:

11	17	9	16	5
19	13	15	16	15
19	17	10	20	16
9	13	9	5	19
19	15	17	16	10

α) Να κατασκευαστεί πίνακας συχνοτήτων απολύτων και αθροιστικών.

β) i) Πόσοι μαθητές είχαν βαθμό τουλάχιστον 15;

ii) Πόσοι μαθητές είχαν βαθμό μεγαλύτερο από 13;

iii) Τι ποσοστό μαθητών είναι κάτω από βάση (10);

iv) Τι ποσοστό μαθητών είναι πάνω από 16;

v) Τι ποσοστό μαθητών έχει βαθμό μεταξύ του 15 και 19;

vi) Τι εκφράζει η αθροιστική συχνότητα N_5;

vii) Τι εκφράζει η σχετική αθροιστική συχνότητα $F_3\%$.

ΛΥΣΗ

α) «Πίνακας κατανομής συχνοτήτων και αθροιστικών συχνοτήτων απολύτων και σχετικών».

Βαθμοί x_i	Διαλογή	Συχνότητα v_i	Σχετ. συχν. f_i	Αθροιστ. Συχν. $f_i\%$	Αθρ. Σχετ. Συχν. N_i	Αθρ. Σχετ. Συχν. F_i	Αθρ. Σχετ. Συχν. $F_i\%$
5	II	2	0,08	8	2	0,08	8
9	III	3	0,12	12	5	0,20	20
10	II	2	0,08	8	7	0,28	28
11	I	1	0,04	4	8	0,32	32
13	II	2	0,08	8	10	0,40	40
15	III	3	0,12	12	13	0,52	52
16	IIII	4	0,16	16	17	0,68	68
17	III	3	0,12	12	20	0,80	80
19	IIII	4	0,16	16	24	0,96	96
20	I	1	0,04	4	25	1	100
Σύνολο		**25**	**1**	**100**			

β) i) Βαθμό τουλάχιστον 15 έχουν $v - N_{i-1} = 25 - N_5 = 25 - 10 = 15$ μαθητές ή $v_6 + v_7 + v_8 + v_9 + v_{10} = 3 + 4 + 3 + 4 + 1 = 15$.

ii) Η έκφραση «βαθμό μεγαλύτερο από 13» ισοδυναμεί, σύμφωνα με τον πίνακα, με την έκφραση «βαθμό τουλάχιστον 15», που είναι 15 μαθητές.

iii) Το ποσοστό των μαθητών που είναι κάτω από τη βάση είναι το ποσοστό των μαθητών που έχει «βαθμό το πολύ 9» και είναι $F_2\% = 20\%$.

iv) Το ποσοστό των μαθητών που έχει βαθμολογία πάνω από 16 είναι το άθροισμα των ποσοστών των μαθητών με βαθμολογίες 17, 19, 20 δηλαδή $f_8\% + f_9\% + f_{10}\% = 32\%$ ή το ποσοστό των μαθητών που έχουν «βαθμολογία τουλάχιστον 17» και είναι $100\% - F_7\% = 100\% - 68\% = 32\%$.

v) Το ποσοστό των μαθητών που έχει βαθμό μεταξύ του 15 και 19 είναι το άθροισμα των ποσοστών των μαθητών που έχουν βαθμούς 16 και 17 δηλαδή $f_7\% + f_8\% = 16 + 12 = 28\%$

vi) Η αθροιστική συχνότητα N_5 = 10 δείχνει ότι 10 από τους 25 μαθητές βαθμολογήθηκαν με βαθμό μέχρι και 13.

vii) Η σχετική συχνότητα f_3% δείχνει ότι το 28% των μαθητών βαθμολογήθηκε με βαθμό μέχρι και 10.

3.4. ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΚΑΤΑΝΟΜΗΣ ΣΥΧΝΟΤΗΤΩΝ

Οι στατιστικοί πίνακες, παρά την πληρότητα την οποία παρουσιάζουν και την ακρίβεια των πληροφοριών που περιέχουν, είναι σχεδόν πάντοτε χρήσιμο οι πληροφορίες που περιέχουν να παρασταθούν με μορφή διαγραμμάτων ή γραφικών παραστάσεων.

Με αυτό τον τρόπο επιτυγχάνεται μια εποπτική αντίληψη του φαινομένου και επιτρέπεται ο τονισμός των κύριων χαρακτηριστικών του, αδιαφορώντας για τις λεπτομέρειες, που τις περισσότερες φορές δεν έχουν και μεγάλη σημασία.

Τα διαγράμματα πρέπει να έχουν τίτλο που να είναι σύντομος και σαφής και αναγράφεται συνήθως στο κάτω μέρος τους. Κατά μήκος των αξόνων των διαγραμμάτων πρέπει να σημειώνονται οι κλίμακες των τιμών των μεγεθών που απεικονίζονται. Όταν είναι αναγκαίο, θα πρέπει κάτω από το διάγραμμα να αναγράφονται οι τυχόν υποσημειώσεις για διευκρινήσεις ή συμπληρωματικές επεξηγήσεις των μεγεθών που απεικονίζονται. Τέλος, πρέπει να αναφέρεται και η πηγή από την οποία πήραμε τα αριθμητικά δεδομένα των μεγεθών.

Τα στατιστικά δεδομένα που συγκεντρώνουν σ' ένα πίνακα συχνοτήτων μπορούν να παρουσιαστούν με τη μορφή γραφικών παραστάσεων ή διαγραμμάτων.

Από τις γραφικές παραστάσεις ή τα διαγράμματα μπορούμε ν' αναλήσουμε συνοπτικά διάφορες χρήσιμές πληροφορίες. Με τα διαγράμματα διευκολύνεται η σύγκριση μεταξύ ομοειδών στοιχείων για τα ίδια ή διαφορετικά χαρακτηριστικά.

Τα συριότερα είδη γραφικών παραστάσεων είναι:

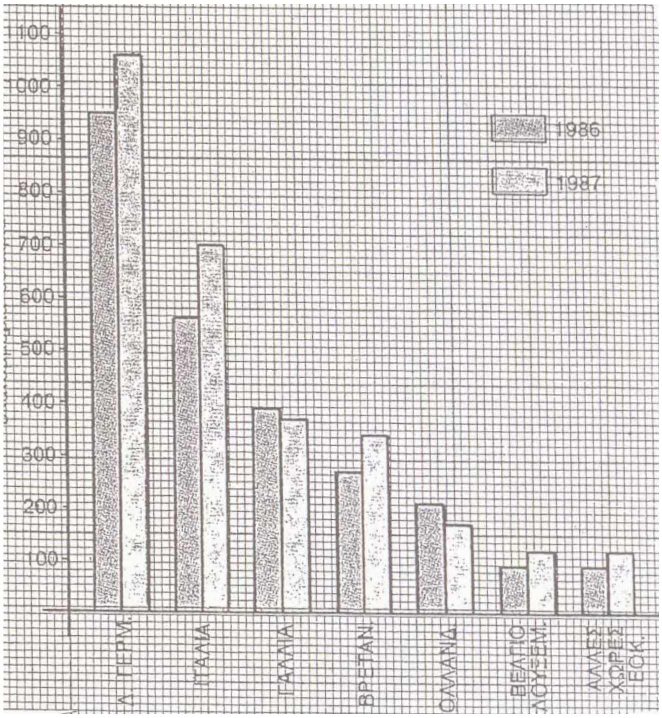
- Ραβδόγραμμα
- Διαγραμμα συχνοτήτων
- Κυκλικό δράγρμμα
- Χρονόγραμμα

α) Ραβδόγραμμα

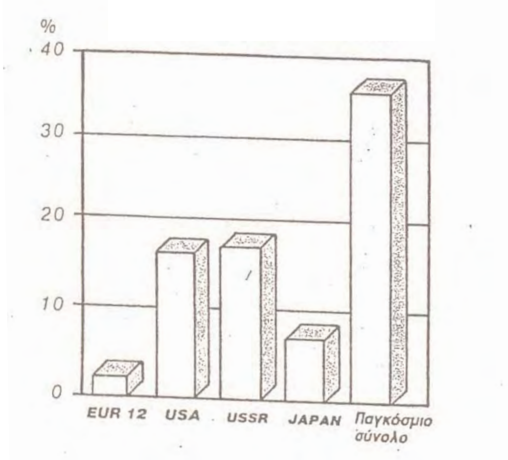
Το ραβδόγραμμα χρησιμοποιείται συνήθως όταν η μεταβλητή Χ είναι **ποιοτική**.

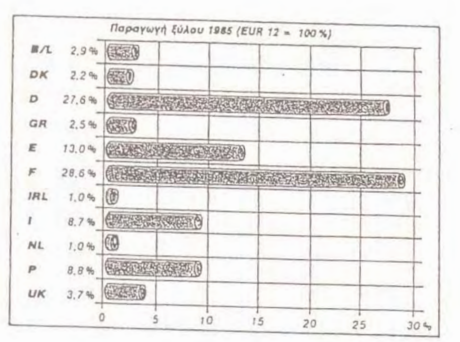
Αποτελείται από ορθογώνιες στήλες που οι βάσεις τους έχουν κοινό μήκος που επιλέγεται αυθαίρετα είναι όμως τέτοιο ώστε να εξασφαλίζονται κενά μεταξύ δύο διαδοχικών ορθογωνίων. Οι στήλες υψώνονται πάνω σε ορθογώνιο ή κατακόρυφο άξονα. Το ύψος κάθε στήλης είναι ίσο με την συχνότητα (vi) ή τη σχετική συχνότητα fi της αντίστοιχης τιμής της μεταβλητής. Όταν θέλουμε να κάνουμε σύγκριση των αντί-στοιχων τιμών της ίδιας μεταβλητής Χ για δύο διαφορετικά δείγματα, τότε κατα-σκευάζουμε δίπλα ορθογώνια για την ίδια τιμή της μεταβλητής Χ, ένα για το κάθε δείγμα.

Εξαγωγές προς ΕΟΚ
τα έτη 1986, 1987 (Ιαν.-Σεπτ.)



Αύξηση πληθυσμού 1985-2005





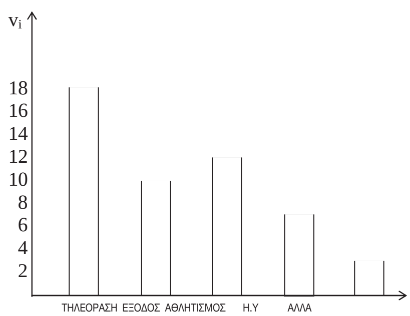
Εφαρμογή 2

Σε μια έρευνα για το πώς διαθέτουν 50 μαθητές τον ελεύθερο χρόνο τους πήραμε τον ακόλουθο πίνακα κατανομής συχνοτήτων.

x_Z	Συχνότητα vi	Σχετική Συχνότητα fi	Σχετική Συχνότητα F_i%
ΤΗΛΕΟΡΑΣΗ	17	0,34	34
ΕΞΟΔΟΣ	10	0,20	20
ΑΘΛΗΤΙΣΜΟΣ	13	0,26	26
Η.Υ.	7	0,14	14
ΑΛΛΑ	3	0,03	6
ΣΥΝΟΛΑ	**50**	**1**	**100**

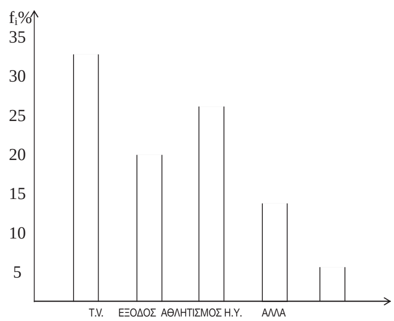
Να κατασκευαστούν τα ραβδογράμματα συχνοτήτων και σχετικών συχνοτήτων.

ΛΥΣΗ



Ραβδόγραμμα συχνοτήτων της μεταβλητής

«ΔΙΑΘΕΣΗ ΕΛΕΥΘΕΡΟΥ ΧΡΟΝΟΥ»



Ραβδόγραμμα σχετ. συχνοτήτων % της μεταβλητής Χ:

«ΔΙΑΘΕΣΗ ΕΛΕΥΘΕΡΟΥ ΧΡΟΝΟΥ»

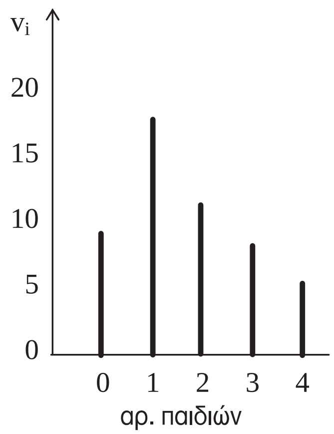
β) Διάγραμμα συχνοτήτων

Στην περίπτωση **ποσοτικών μεταβλητών** αυτή του ραβδογράμματος χρησιμοποιείται το διάγραμμα συχνοτήτων και σχετικών συχνοτήτων.

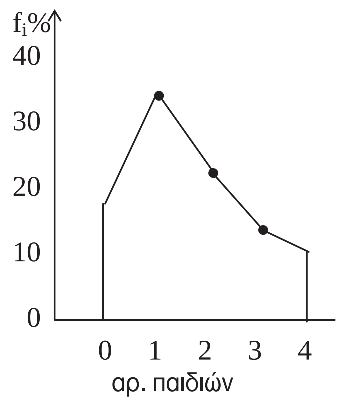
Αποτελείται από ευθύγραμμα τμήματα κάθετα στον άξονα της μεταβλητής Χ, ένα για κάθε τιμή x_i. Οι τιμές της μεταβλητής τοποθετούνται στο διάγραμμα σε αύξουσα σειρά. Το ύψος κάθε ευθύγραμμου τμήματος είναι ίσο με την αντίστοιχη συχνότητα vi ή με την αντίστοιχη σχετική συχνότητα fi. Αν ενώσουμε τα σημεία (x_i, v_i), i = 1, 2, ..., κ με διαδοχικά ευθύγραμμα τμήματα δημιουργείται μια πολυγωνική γραμμή που ονομάζεται πολύγωνο συχνοτήτων. Ομοίως αν ενώσουμε τα σημεία (x_i, f_i) δημιουργείται το πολύγωνο σχετικών συχνοτήτων. Με τον ίδιο τρόπο κατασκευάζονται και τα διαγράμματα των αθροιστικών συχνοτήτων.

Αν συμπληρώσουμε τον πίνακα 4 με τις αθροιστικές συχνότητες, έχουμε:

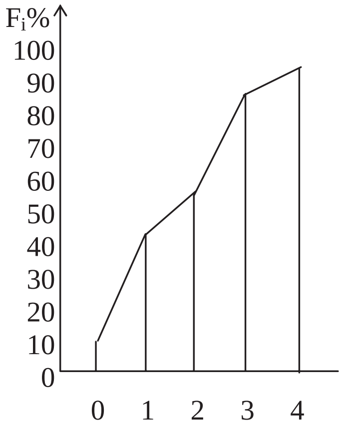
Αρ. παιδιών x_i	v_i	f_i	$f_i\%$	N_i	F_i	$F_i\%$
0	8	0,16	16	8	0,16	16
1	17	0,34	34	25	0,50	50
2	13	0,26	26	38	0,76	76
3	7	0,14	14	45	0,90	90
4	5	0,10	10	50	1	100
Σύνολο		**1**	**100**			



Διάγραμμα συχνοτήτων (v_i)



Πολύγωνο σχετ. συχνοτήτων $f_i\%$



Διάγραμμα και πολύγωνο σχετ. αθροιστικών συχνοτήτων F_i%

γ) Κυκλικό διάγραμμα

Το κυκλικό διάγραμμα χρησιμοποιείται για την γραφική παράσταση ποιοτικών ή ποσοτικών μεταβλητών όταν οι τιμές της μεταβλητής X είναι σχετικά λίγες.

Το κυκλικό διάγραμμα είναι ένας κυκλικός δίσκος χωρισμένος σε κ κυκλικούς τομείς (όσες και οι τιμές της μεταβλητής X). Το εμβαδόν E_i κάθε κυκλικού τομέα είναι ανάλογο προς τις αντίστοιχες συχνότητες v_i ή τις σχετικές συχνότητες f_i. Τα τόξα αi που αντιστοιχούν σε κάθε κυκλικό τομέα δίνονται από τη σχέση:

$$\alpha 1=360^o \cdot \frac{v_i}{v} = 360^o f_1, \qquad i=1,2,\ldots,\kappa$$

Εφαρμογή 3

Σε μια έρευνα που έγινε σε 100 μαθητές της Γ΄ Λυκείου ενός σχολείου, οι 40 μαθητές επέλεξαν τις θετικές επιστήμες, 20 μαθητές τις οικονομικές, 35 μαθητές τις θεωρητικές επιστήμες και 5 μαθητές τις ιατρικές επιστήμες. Να κατασκευαστεί το κυκλικό διάγραμμα σχετικών συχνοτήτων.

ΛΥΣΗ

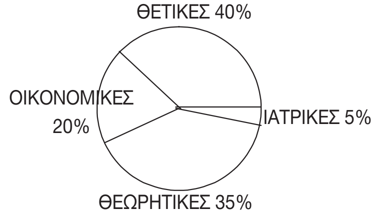
Για την κατασκευή του κυκλικού διαγράμματος βρίσκουμε πρώτα τις επίκεντρες γωνίες αξιοποιώντας τον τίτλο: αi = 360ο. Οπότε:

Θετικές: $\alpha_1 = 360^o \ \frac{40}{100} = 144^o$

Οικονομικές: $\alpha_2 = 360^o \ \frac{35}{100} = 72^o$

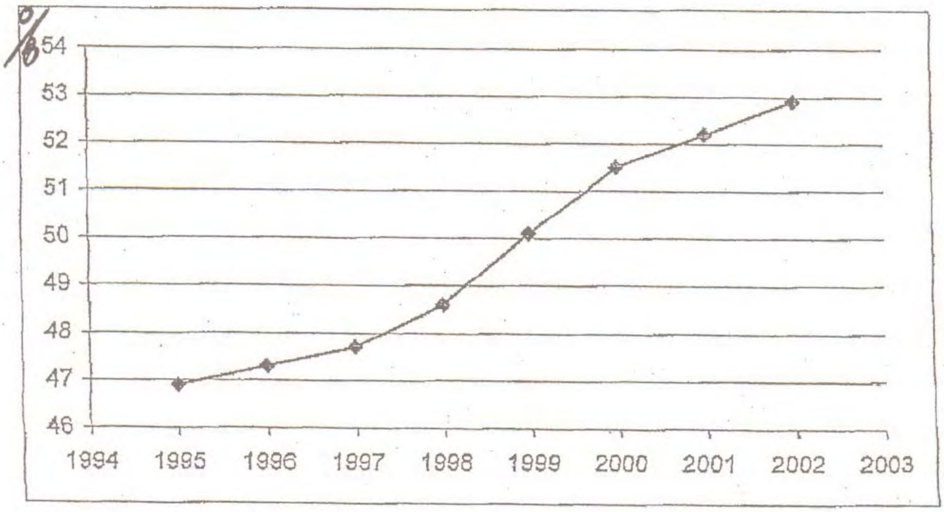
Θεωρητικές: $\alpha_3 = 360^o \ \frac{5}{100} = 126^o$

Ιατρικές: $\alpha_4 = 360^o \ \frac{20}{100} = 18^o$



δ) Χρονόγραμμα

Όταν θέλουμε να παρακολουθήσουμε την διαχρονική εξέλιξη διαφόρων μεγεθών, τότε κατασκευάζουμε τη γραφική παράσταση στην οποία στον οριζόντιο άξονα λαμβάνουμε ισομήκη διαδοχικά τμήματα, καθένα από τα οποία αντιστοιχεί στη μονάδα του χρησιμοποιούμενου χρόνου, και στο κατακόρυφο άξονα παίρνουμε κλίμακα η οποία πρέπει να καλύπτει τις τιμές της μεταβλητής.



Αναλογία εργαζομένων γυναικών ως προς το σύνολο των γυναικών στις χώρες ζώνης του Ευρώ (1995-2002).

3.5. ΟΜΑΔΟΠΟΙΗΣΗ ΠΑΡΑΤΗΡΗΣΕΩΝ

Ζυγίσαμε 55 μαθητές του Δημοτικού και το βάρος τους σε κιλά έδωσε τους παρακάτω αριθμούς:

53,5	63,2	56,7	51,4	51,3	53,8	53,9	52,9	63,8	53,9	52,7
54,4	53,6	51,4	52,4	53,6	53,1	57,5	56,9	57,8	62,7	53,9
53,3	59,7	52,0	54,4	53,3	58,7	53,9	61,6	55,6	59,0	53,4
51,4	53,5	52,8	60,4	61,9	62,3	52,9	52,8	51,5	54,1	55,9
53,1	58,7	52,8	53,8	63,0	53,0	56,4	53,2	55,9	57,0	58,2

Αν ζητηθεί να κατασκευάζουμε πίνακα κατανομής συχνοτήτων, τότε θα είχαμε πάρα πολλές διαφορετικές τιμές που τις συναντάμε μια μόνο φορά, και δεν θα διέφερε σημαντικά ο πίνακας από την καταμέτρηση που έχουμε πιο πάνω. Στην περίπτωση αυτή κάνουμε **ομαδοποίηση** των παρατηρήσεων.

Χωρίζουμε τα δεδομένα σε μικρό πλήθος ομάδων που λέγονται κλάσεις έτσι, ώστε κάθε τιμή να ανήκει σε μια κλάση. Το πλήθος των κλάσεων καθορίζεται αυθαίρετα, αλλά καθορίζεται με τέτοιο τρόπο, ώστε η πρώτη κλάση να περιλαμβάνει τη μικρότερη παρατήρηση και η τελευταία κλάση να περιλαμβάνει τη μεγαλύτερη παρατήρηση.

Οι κλάσεις είναι διαδοχικά διαστήματα, συνήθως με το ίδιο πλάτος, χωρίς όμως να αποκλείεται η περίπτωση οι κλάσεις να έχουν και διαφορετικό πλάτος. Το πλήθος κ των κλάσεων γίνεται βάσει συγκεκριμένου κανόνα από τον ερευνητή, μπορούμε όμως να χρησιμοποιήσουμε ενδεικτικά τον παρακάτω πίνακα:

ΜΕΓΕΘΟΣ ΔΕΙΓΜΑΤΟΣ	ΑΡΙΘΜΟΣ ΚΛΑΣΕΩΣ
μικρότερο του 20	5
20-50	6
50-100	7
100-200	8
200-400	9
400-700	10

Όταν θέλουμε να ομαδοποιήσουμε σε ισοπλατείς κλάσεις, βρίσκουμε το εύρος του δείγματος, δηλαδή τη διαφορά της μικρότερης από τη μεγαλύτερη παρατήρηση. Στα αριθμητικά δεδομένα του βάρους των 55 μαθητών είναι:

R=63,8-51,3=12,5, διαιρούμε με το πλήθος κ=7 των κλάσεων,

οπότε 12,5:7=1,786. Επιλέγουμε ως πλάτος κλάσης 1,8 (στρογγυλοποιώντας πάντα προς τα πάνω έτσι, ώστε να περιλαμβάνονται όλες οι τιμές στο προκαθορισμένο αριθμό κλάσεων).

Έτσι σχηματίζονται οι κλάσεις 51 , 3 - 53 ,1 , 53,1-54,9 , ...

Μια παρατήρηση που συμπίπτει με το άνω άκρο μιας κλάσης θα τοποθετηθεί στη διαλογή στην αμέσως επόμενη κλάση. Για παράδειγμα, η παρατήρηση 53,1 θα τοποθετηθεί στην κλάση 53,1 - 54,9.

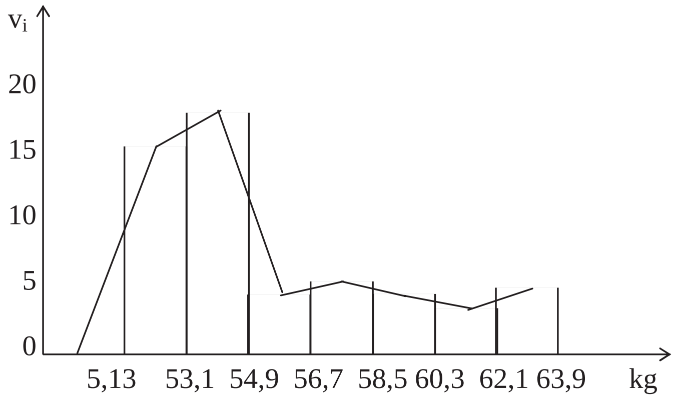
Αν έχουμε μια κλάση α - β, τότε ο αριθμός $x_i = \frac{\alpha+\beta}{2}$ λέγεται κεντρική τιμή της κλάσης και θα αντιπροσωπεύει τις παρατηρήσεις αυτής της κλάσης σε ορισμένα μέτρα που θα θέσουμε παρακάτω.

Για τα δεδομένα του βάρους των 55 μαθητών έχουμε:

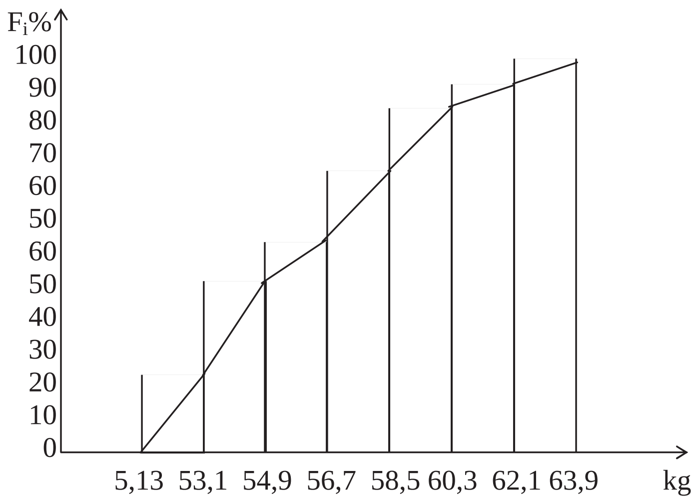
Κλάσεις	Κεντρική τιμή x_i	Συχνότητα v_i	Σχετ, Συχν. f_i%	Αθρ. Συχν. N_i	Αθροι. Σχετ. Συχν.
5,13 – 53,1	52,2	14	25,5	14	25,5
53,1 – 54,9	54,0	19	34,5	33	60,0
54,9 – 56,7	55,8	4	7,3	37	67,3
56,7 – 58,5	57,6	6	10,9	43	78,2
58,5 – 60,3	59,4	4	7,3	47	85,5
60,3 – 62,1	61,2	3	5,4	50	90,9
62,1 – 63, 9	63,0	5	9,1	55	100
		55	**100**		

Η γραφική παράσταση μιας ομαδοποιημένης κατανομής γίνεται με το ιστόγραμμα. Για να κατασκευάζουμε ιστόγραμμα και πολύγωνο συχνοτήτων ή σχετικών συχνοτήτων, τοποθετούμε τις κλάσεις στον οριζόντιο άξονα και δημιουργούμε ορθογώνια που έχουν πλάτος ίσο με το πλάτος της κλάσης και ύψος ίσο με v_i ή f_i% αντίστοιχα. Αν κατασκευάζουμε και δύο ακόμα υποθετικές κλάσεις, μια πριν την πρώτη κλάση και μια μετά την τελευταία κλάση με συχνότητες ή σχετικές συχνότητες μηδέν και θεωρήσουμε το μέσα των βάσεων των ορθογωνίων, τότε αν ενώσουμε αυτά θα πάρουμε το πολύγωνο συχνοτήτων ή σχετικών συχνοτήτων.

Για την προηγούμενη άσκηση το ιστόγραμμα και πολύγωνο συχνοτήτων είναι:



Για να κατασκευάσουμε το ιστόγραμμα και πολύγωνο αθροιστικών συχνοτήτων ή σχετικών αθροιστικών συχνοτήτων, τοποθετούμε τις κλάσεις στον οριζόντιο άξονα όπως πριν και δημιουργούμε ορθογώνια που έχουν ύψος ίσο με την αθροιστική συχνότητα N_i ή τη σχετική αθροιστική συχνότητα F_i%. Αν ενώσουμε τα δεξιά άκρα των άνω βάσεων, δημιουργούμε το πολύγωνο αθροιστικών συχνοτήτων. Έτσι το ιστόγραμμα και πολύγωνο της σχετικής αθροιστικής συχνότητας της προηγούμενης άσκησης είναι:



ΕΞΑΣΚΗΣΗ

1. Έχουμε τα δεδομένα

0, 0, 0, 1, 1, 2, 2, 2, 2, 3, 5, 5, 5, 5, 5, 6, 6, 7, 7, 7.

Να κατασκευάσετε πίνακα κατανομής συχνοτήτων και αθροιστικών συχνοτήτων και να υπολογίσετε τις τιμές v_2, f_3, $f_2\%$, N_3, $F_5\%$.

2. Τα χρήματα σε ευρώ που ξοδεύουν 25 μαθητές της Γ΄ Γυμνασίου κατά τη διάρκεια της εβδομάδας, δίνονται παρακάτω:

5 8 6 8 7 9 10 5 8 7 8 7 6

6 8 9 10 10 5 6 7 7 9 5 9

α) Να κατασκευάσετε πίνακα κατανομής συχνοτήτων

β) Να βρείτε πόσοι μαθητές ξοδεύουν:

i) το πολύ 7 ευρώ.

ii) τουλάχιστον 8 ευρώ.

iii) ακριβώς 6 ευρώ.

γ) Να βρείτε το ποσοστό των μαθητών που ξοδεύουν:

i) τουλάχιστον 9 ευρώ.

ii) το πολύ 8 ευρώ.

3. Στον παρακάτω πίνακα κατά την διάρκεια της εκτύπωσης χάθηκαν κάποια δεδομένα. Να συμπληρωθεί ο πίνακας.

x_i	v_i	f_i	$f_i\%$	N_i	F_i
10	4				
11				10	
12				15	60
13					
14		0,12			

4. Ο παρακάτω πίνακας παρουσιάζει τον αριθμό των δωματίων που έχουν τα διαμερίσματα ενός οικιστικού συγκροτήματος.

i) Να συμπληρώσετε τον πίνακα.

ii) Να κατασκευάσετε πίνακα σχετικών συχνοτήτων και σχετικών αθροιστικών συχνοτήτων.

iii) Να βρείτε πόσα διαμερίσματα έχουν:

α) το πολύ 4 διαμερίσματα.

β) τουλάχιστον 4 διαμερίσματα.

iv) Τι ποσοστό διαμερισμάτων έχει τουλάχιστον 3 δωμάτια;

Δωμάτια x_i	Δωμάτια v_i	N_i
1	6	
2	9	
3		36
4		
5	12	66
6	9	
ΣΥΝΟΛΟ		

5. Να συμπληρώσετε τον πίνακα που παρουσιάζει τους μετεξεταστέους μαθητές της Β΄ Λυκείου των σχολείων μιας περιφέρειας και στη συνέχει να γίνει το ραβδόγραμμα συχνοτήτων.

ΜΑΘΗΜΑ	v_i	F_i%
Αρχαία	5	
Φυσική		16
Λογοτεχνία	2	
Γαλλικά	4	
Άλγεβρα	20	40
Γεωμετρία		
ΣΥΝΟΛΟ		

6. Μια εταιρεία δημοσκοπήσεων έκανε έρευνα σε ένα δείγμα πολιτών για τις κομματικές προτιμήσεις τους και τα αποτελέσματα δίνονται στον διπλανό πίνακα.

i) Να βρείτε το μέγεθος του δείγματος

ii) Να βρείτε πόσοι προτίμησαν το κάθε κόμμα

iii) Να σχεδιάσετε το ραβδόγραμμα σχετικών συχνοτήτων.

ΚΟΜΜΑ	v_i	f_i
Α		0,3
Β	70	0,14
Γ		0,22
Δ		
ΣΥΝΟΛΟ		

7. Σε μια πόλη της Βόρειας Ελλάδας καταγράψαμε τις μέγιστες ημερήσιες θερμοκρασίες κατά το μήνα Σεπτέμβριο και πήραμε τα παρακάτω αποτελέσματα σε βαθμούς °C:

21 23 24 20 23 25 22 23 23 21

22 21 23 23 24 24 21 20 23 21

21 23 23 21 20 23 24 23 22 24

i) Να κατασκευάσετε πίνακα συχνοτήτων και αθροιστικών συχνοτήτων

ii) Πόσες μέρες η θερμοκρασία ήταν:

α) τουλάχιστον 23°C

β) το πολύ 23°C

γ) τουλάχιστον 22°C και το πολύ 24°C

iii) Να κατασκευάσετε το διάγραμμα και το πολύγωνο συχνοτήτων.

8. Εξετάζουμε 25 οικογένειες μιας πολυκατοικίας ως προς τον αριθμό των παιδιών τους και παίρνουμε το παρακάτω αποτέλεσμα:

2 1 2 1 3

3 0 3 2 2

1 4 1 3 2

2 1 0 1 1

1 2 3 0 4

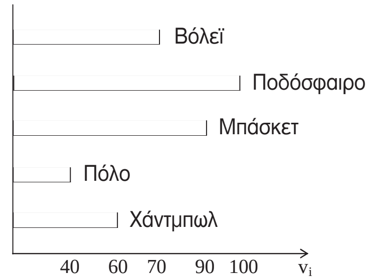
i) Να κατασκευάσετε πίνακα συχνοτήτων και αθροιστικών συχνοτήτων.

ii) Να κατασκευαστούν τα διαγράμματα συχνοτήτων και σχετικών συχνοτήτων καθώς και το πολύγωνο σχετικών συχνοτήτων.

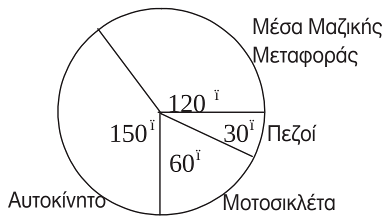
iii) Να κατασκευαστεί το διάγραμμα και το πολύγωνο αθροιστικών συχνοτήτων.

iv) Ο Δήμος της περιοχής αποφάσισε να δώσει 1000 ευρώ σε κάθε παιδί από το 28% των οικογενειών που έχουν τα περισσότερα παιδιά. Πόσα χρήματα θα δαπανηθούν και πόσες οικογένειες θα πάρουν το επίδομα;

9. Να μετατρέψετε το διπλανό ραβδόγραμμα συχνοτήτων των αθλημάτων που προτιμούν οι νέοι μιας περιοχής σε κυκλικό διάγραμμα.

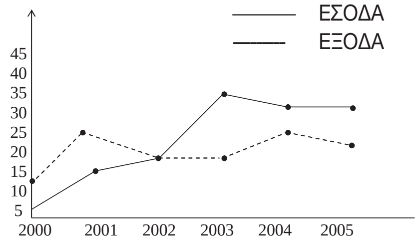


10. Το διπλανό κυκλικό διάγραμμα παρουσιάζει το μεταφορικό μέσο που χρησιμοποιούν 1.440 εργαζόμενοι μιας περιοχής της Αθήνας για να μεταβούν στην εργασία τους.



i) Πόσοι εργαζόμενοι πηγαίνουν στην εργασία τους με τα πόδια;
ii) Ποιο ποσοστό εργαζομένων χρησιμοποιεί μοτοσικλέτα;
iii) Να μετατρέψετε το κυκλικό διάγραμμα σε ραβδόγραμμα συχνοτήτων.

11. Στο παρακάτω χρονόγραμμα παρουσιάζονται τα έξοδα μιας εταιρείας σε εκατομμύρια ευρώ από το έτος 2000 έως το 2005.
i) Μετά από ποια χρονιά η εταιρεία αρχίζει να έχει έσοδα;
ii) Ποια χρονιά η εταιρεία είχε τη μεγαλύτερη ζημιά και ποια χρονιά είχε το μεγαλύτερο κέρδος;
iii) Να περιγράψετε την οικονομική κατάσταση της εταιρείας το έτος 2003.



iv) Να κατασκευάσετε συγκριτικό ραβδόγραμμα εσόδων – εξόδων.
v) Ποια είναι η συνολική οικονομική κατάσταση της εταιρείας από το 2000 και μετά;

12. Παρακάτω δίνονται οι εβδομαδιαίες αποδοχές σε ευρώ 50 υπαλλήλων μιας εταιρείας:

180	110	170	190	195	100	195	195	130	130
200	300	200	300	200	250	250	280	285	285
270	270	300	330	270	275	200	250	250	330
360	360	360	380	330	335	340	370	380	380
400	450	380	380	380	450	500	450	370	370

i) Να γίνει ομαδοποίηση των παρατηρήσεων σε 4 κλάσεις ίσου πλάτους και να κατασκευαστεί ο αντίστοιχος πίνακας v_i, f_i%, N_i, F_i%
ii) Να κατασκευάσετε τα ιστογράμματα και τα αντίστοιχα πολύγωνα συχνοτήτων v_i και αθροιστικών συχνοτήτων F_i%
iii) Να εκτιμήσετε το ποσοστό των υπαλλήλων που οι εβδομαδιαίες αποδοχές τους είναι πάνω από 350 ευρώ.

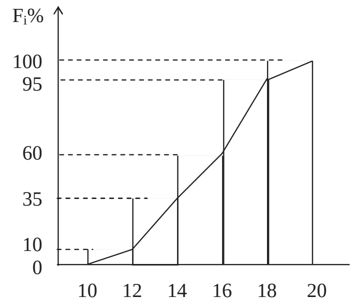
13. Στο παρακάτω σχήμα δίνεται το πολύγωνο αθροιστικών σχετικών συχνοτήτων του βαθμού του απολυτηρίου των 100 μαθητών της Γ΄ Λυκείου ενός σχολείου.

Να βρεθεί:

i) Το πλήθος των μαθητών της κάθε κλάσης.

ii) Το ποσοστό των μαθητών με βαθμό από 16 έως 20.

iii) Να κατασκευάσετε το πολύγωνο συχνοτήτων.



14. Το βάρος των αποσκευών καθενός εκ των 80 επιβατών μιας πτήσης είναι τουλάχιστον 11 κιλά, αλλά μικρότερο από 26 κιλά. Γνωρίζουμε ότι 8 επιβάτες έχουν αποσκευές με βάρος μικρότερο από 17 κιλά, 48 επιβάτες έχουν αποσκευές με βάρος μικρότερο από 20 κιλά και 15% των επιβατών έχει αποσκευές με βάρος τουλάχιστον 23 κιλά.

i) Να παραστήσετε τα δεδομένα σε πίνακα συχνοτήτων.

ii) Κάθε επιβάτης δικαιούται να μεταφέρει αποσκευές με βάρος μικρότερο των 20 κιλών, διαφορετικά έχει πρόσθετη οικονομική επιβάρυνση. Να βρείτε το ποσοστό των επιβατών της πτήσης που έχει πρόσθετη οικονομική επιβάρυνση

iii) Να βρεθούν οι γωνίες των αντίστοιχων κυκλικών τομέων του κυκλικού διαγράμματος σχετικών συχνοτήτων, για τα δεδομένα του προβλήματος.

ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ

1. Κατασκευάστε ένα γενικό πίνακα για όλα τα δεδομένα που συγκεντρώσατε με τα ερωτηματολόγια και στη συνέχεια να κατασκευάσετε πίνακα συχνοτήτων, σχετικών συχνοτήτων και αθροιστικών συχνοτήτων με τα αντίστοιχα διαγράμματα ανάλογα με το είδος της μεταβλητής. Για συντομία μπορείτε να χωριστείτε και σε ολιγομελείς ομάδες και στο τέλος να συγκρίνετε μεταξύ σας τα αποτελέσματα κάθε ομάδας.
2. Χωριστείτε σε ολιγομελείς ομάδες και ανατρέξτε στο διαδίκτυο στο site της Ε.Σ.Υ.Ε. (http://www.statistics.gr/) και συγκεντρώστε στοιχεία (πίνακες και διαγράμματα) που αφορούν στοιχεία από την απογραφή του πληθυσμού στις 18 Μαρτίου 2001, για διάφορες, μεταβλητές της επιλογής σας. Να παρουσιάσετε και αποτελέσματα στη τάξη σας.

Σύνοψη

Μετά τη συλλογή δεδομένων θα πρέπει αυτά να παρουσιάζονται και να οπτικοποιούνται αυτό επιτυγχάνεται με τη βοήθεια των Στατισκικών πινάκων

• <u>Στατισικοί πίνακες</u> → Γενικοί πίνακες
→ Ειδικοί πίνακες

Μετά τη συγκέντρωση των στατιστικών στοιχείων που αναφέρονται σε μια μόνο ιδιοτήτα ενός πληθυσμού το πιο σημαντικό βήμα είναι η κατάλληλη κατάταξη και η συστηματική ομαδοποίηση των τιμών της μεταβλητής με αντικειμενικό σκοπό την παρουσίαση των στοιχείων με τέτοιο τρόπο ώστε να δίνει πληροφορίες για την δομή του πληθυσμού που είναι προς μελέτη, και να διευκολύνεται το έργο της στατιστικής ανάλυσης και εξαγωγής χρήσιμων συμπερασμάτων. Η ομαδοποίηση των τιμών της μεταβλητής χρησιμοποιούμε ειδικές κατατάξεις που τις ομομάζουμε κατανομές συχνοτήτων ή πίνακες συχνοτήτων.

Για την κατασκευή αυτών των πινάκων χρησιμοποιούμε τις:

- Συχνότητες
- Σχετικές συχνότητες
- Αθριστικές συχνότητες

Σε περίπτωση μεγάλου αριθμού παρατηρήσεων αξιοποιούμε την μέθοδο της ομαδοποίησης των παρατηρήσεων.

- Γραφικές παραστάσεις

Οι γραφικές παραστάσεις είναι το καλύτερο μέσο μιας στατιστικής παρουσίασης γιατί δίνουν στους αφηρημένους αριθμούς μια συγκεκριμένη μορφή που μας διευκολύνει να έχουμε μια άμεση αντίληψη της μορφής του φαινομένου που θέλουμε να μελετήσουμε. Υπάρχουν πολλές κατηγορίες διαγραμμάτων μεταξύ αυτών είναι:

- Ραβδόγραμμα
- Διάγραμμα συχνοτήτων
- Κυκλικό διάγραμμα
- Χρονοδιάγραμμα

Η γραφική απεικόνηση μιας ομαδοποιημένη κατανομής γίνεται με τη βοήθεια των ιστογραμμάτων συχνοτήτων ή σχετικών συχνοτήτων καθώς και με τα αντίστοιχα πολύγωνα

Βιβλιογραφία

«*Εισαγωγή στη Στατιστική*», Αθανασοπούλου Δ.
«*Εισαγωγή στη Στατιστική*», Μέρος Ι, ΙΙ, Δαμιανού Χ. και Κούτρα Μ.
«*Στατιστική θεωρία και εφαρμογές*», Κάκουλλου Θ.

4. ΑΝΑΛΥΣΗ ΣΤΟΙΧΕΙΩΝ

Βασικές έννοιες:

- μέσος
- διάμεσος
- εκατοστημόρια - τεταρτημόρια
- εύρος
- ενδοτεταρτημοριακό εύρος
- διακύμανση
- τυπική απόκλιση
- συντελεστής μεταβλητότητας

Στόχος του μαθήματος:
Δυνατότητα αναπαράστασης των δεδομένων και συμπερασμάτα βάσεις των μέτρων θέσεων και διασποράς

Προσδοκόμενα αποτελέσματα: ο εκπαιδευόμενος είναι σε θέση να επιλέξει το καταλληλότερο μέτρο θέσης ή διασποράς σε σχέση με το φαινόμενο που εξετάζει.

Εισαγωγικές παρατηρήσεις: ανάλογα με τη μορφή του πληθυσμού που εξετάζουμε, άλλωτε μας επιτρέπεται να χρησιμοποιήσουμε ενδεικτικά για τον πληθυσμό ένα μέτρο θέσης (όπως η μέση τιμή) και άλλωτε ένα μέτρο διασποράς (όπως η διακύμανση).

4.1. ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ

Όταν εξετάζουμε κάποια δεδομένα, κρίνεται σκόπιμο κάποιες φορές να υπολογίσουμε και ορισμένα μέτρα, που έχουν σαν στόχο να περιγράψουν με περιληπτικό τρόπο τα βασικά χαρακτηριστικά τους.

Τα μέτρα θέσης δίνουν περιληπτικά την θέση των δεδομένων πάνω στον άξονα των πραγματικών αριθμών.

Προσδιορίζουν ένα κεντρικό σημείο στο οποίο τείνουν να συγκεντρώνονται τα δεδομένα. Τέτοια μέτρα είναι ο αριθμητικός μέσος, η διάμεσος, η επικρατούσα τιμή, τα εκατοστημόρια και τα τεταρτημόρια.

Τα μέτρα διασποράς δίνουν περιληπτικά το «άπλωμα» των δεδομένων πάνω στην ευθεία των πραγματικών αριθμών. Τέτοια μέτρα είναι το εύρος, η διακύμανση, η τυπική απόκλιση, ο συντελεστής μεταβολής και η ενδοτεταρτημοριακή απόκλιση.

4.2. ΑΡΙΘΜΗΤΙΚΟΣ ΜΕΣΟΣ (ΜΕΤΡΟ ΘΕΣΗΣ)

Τρεις οικογένειες καταγράφουν τα χρήματα σε ευρώ, δαπάνησαν για 10 ημέρες για την διατροφή τους και έχουμε τα παρακάτω δεδομένα:

Α΄ οικογένεια: 30 22 12 27 45 70 12 11 10 51

Β΄ οικογένεια: 35 28 17 32 78 19 15 36 25 15

Γ΄ οικογένεια: 45 47 52 22 11 12 14 17 38 22

Αν θέλουμε να απαντήσουμε στο ερώτημα: «ποια οικογένεια έχει δαπανήσει τα περισσότερα», μπορούμε να το βρούμε προσθέτοντας τα αντίστοιχα ποσά κάθε οικογένειας.

Αν όμως μας ζητηθεί να εκτιμήσουμε πόσα χρήματα θα χρειαστεί κάθε οικογένεια ολόκληρο το μήνα για τη διατροφή της, θα πρέπει να εκτιμήσουμε πόσο δαπανά κάθε οικογένεια κατά μέσο όρο την ημέρα.

Ο αριθμητικός μέσος μιας μεταβλητής Χ, ο οποίος συμβολίζεται με $\bar{x}$, ορίζεται ως το πηλίκο του αθροίσματος των τιμών δια του πλήθους των παρατηρήσεων.

Δηλαδή: $\boxed{\bar{x} = \dfrac{t_1 + t_2 + \ldots + t_\nu}{\nu}}$

Οπότε ο μέσος όρος της ημερήσιας δαπάνης των 3 οικογενειών είναι:

Α΄οικογένεια: $\bar{x}_1 = \dfrac{30+22+12+27+45+70+12+11+10+51}{10} = \dfrac{290}{10} = 29$

Β΄οικογένεια: $\bar{x}_2 = \dfrac{35+28+17+32+78+19+15+36+25+15}{10} = \dfrac{300}{10} = 30$

Γ΄οικογένεια: $\bar{x}_3 = \dfrac{45+47+52+22+11+12+14+17+38+22}{10} = \dfrac{280}{10} = 28$

Οπότε η Α΄ οικογένεια θα χρειαστεί: 29.30 = 870 ευρώ το μήνα.

η Β΄ οικογένεια θα χρειαστεί: 30.30 = 900 ευρώ το μήνα.

η Γ΄ οικογένεια θα χρειαστεί 28.30 = 840 ευρώ το μήνα.

Όταν οι παρατηρήσεις $x_1, x_2, \ldots, x_\kappa$ έχουν συχνότητες $\nu_1, \nu_2, \ldots, \nu_\kappa$ αντίστοιχα ή όταν έχουμε ομαδοποιημένες παρατηρήσεις με κεντρικές τιμές κάθε κλάσης $x_1, x_2, \ldots, x_\kappa$, τότε ο αριθμητικός μέσος δίνεται από τη σχέση:

$$\boxed{\bar{x} = \frac{x_1\nu_1 + x_2\nu_2 + \ldots + x_\kappa\nu_\kappa}{\nu}}$$

Όταν δεν γνωρίζουμε συχνότητες νi λλά ξέρουμε τις σχετικές συχνότητες f_i, i =1, 2,..., κ η μέση τιμή δίνεται από τη σχέση:

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

Εφαρμογή 1

Ο παρακάτω πίνακας δίνει τις ώρες που παρακολούθησαν τηλεόραση οι ένοικοι 20 διαμερισμάτων μέσα σε μια εβδομάδα. Να βρεθεί ο μέσος χρόνος παρακολούθησης.

Κλάσεις	v_i
0 – 3	4
3 – 6	7
6 – 9	6
9 – 12	3
ΣΥΝΟΛΟ	**20**

ΛΥΣΗ

Έχουμε:

Κλάσεις	x_i	v_i	$x_i v_i$
0 – 3	1,5	4	6
3 – 6	4,5	7	31,5
6 – 9	7,5	6	45
9 – 12	10,5	3	31,5
ΣΥΝΟΛΟ		**20**	**114**

Άρα: $\bar{x} = \frac{114}{20} = 5,7$ ώρες

Σε περίπτωση που δίνεται διαφορετική βαρύτητα στις τιμές $x_1, x_2, \ldots, x_\kappa$ ενός συνόλου δεδομένων, τότε αντί του αριθμητικού μέσου χρησιμοποιούμε το **σταθμικό μέσο**. Επομένως αν οι τιμές $x_1, x_2, \ldots, x_\kappa$ έχουν συντελεστή βαρύτητας $w_1, w_2, \ldots, w_\kappa$, αντίστοιχα τότε ο σταθμικός μέσος βρίσκεται από τη σχέση:

$$\boxed{\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \ldots + x_\kappa w_\kappa}{w_1 + w_2 + \ldots + w_\kappa}}$$

Εφαρμογή 2

Τα 5 τμήματα Θετικής και Τεχνολογικής κατεύθυνσης της Γ΄ τάξης ενός Λυκείου έγραψαν το ίδιο διαγώνισμα στα Μαθηματικά . Οι μέσοι όροι κάθε τμήματος σε βαθμολογία με άριστα το 100 ήταν 83, 88, 79, 65 και 72 αντίστοιχα. Αν κάθε τμήμα από τα προηγούμενα είχε 16, 18, 20, 21 και 23 μαθητές αντίστοιχα, να βρείτε το μέσο όρο όλης της τάξης.

ΛΥΣΗ

Χρησιμοποιώντας τον τύπο του σταθμικού μέσου έχουμε:

Παρατήρηση: $\bar{x} = \frac{83 \cdot 16 + 88 \cdot 18 + 79 \cdot 20 + 65 \cdot 21 + 72 \cdot 23}{16 + 18 + 20 + 21 + 23} = \frac{7513}{98} = 76,6$

Αν χρησιμοποιήσουμε τον αριθμητικό μέσο των 5 μέσων όρων δηλαδη $\frac{83 + 88 + 79 + 65 + 72}{5} = 77,4$ θα οδηγούμασταν σε λανθασμένη εκτίμηση του μέσου όρου.

4.3. ΔΙΑΜΕΣΟΣ (ΜΕΤΡΟ ΘΕΣΗΣ)

Ο κύριος Τάσος πήγε σε ένα κατάστημα κινητής τηλεφωνίας για να πάρει με έκπτωση μια νέα συσκευή κινητού τηλεφώνου. Είχε μαζί του τους μηνιαίους λογαριασμούς των τελευταίων 7 μηνών που ήταν: 11, 8, 9, 8, 7, 52, 10 ευρώ αντίστοιχα. Ζήτησε από τον υπάλληλο να υπολογίσει τον μέσο όρο αυτών των λογαριασμών για να υπολογίσει το ποσό της έκπτωσης και ο υπάλληλος του είπε ότι αυτός ο τρόπος δεν είναι σωστός.

Παρατηρούμε ότι η μέση τιμή τους είναι:

$$\bar{x} = \frac{11+8+9+8+7+52+10}{7} = \frac{105}{7} = 15$$

7 8 8 9 10 11 **15** 52

η οποία υπερβαίνει τις έξι από τις επτά παρατηρήσεις, αφού επηρεάζεται (έλκεται) από την ακραία τιμή 52.

Επομένως δεν μπορούμε να ισχυριστούμε ότι η μέση τιμή είναι το «κέντρο» των παρατηρήσεων.

Για να προσδιορίσουμε ένα είδους «κέντρο» των παρατηρήσεων τις διατά σουμε σε αύξουσα σειρά.

Και διαπιστώνουμε ότι η μεσαία παρατήρηση είναι το 9.

7 8 8 9 10 11 52

Αυτό αποτελεί ένα κέντρο κοντά στην πραγματικότητα, αφού δεν επηρεάζεται από την ακραία παρατήρηση 52 και ονομάζεται διάμεσος των παρατηρήσεων.

Διάμεσος (δ) ενός δείγματος παρατηρήσεων οι οποίες έχουν διαταχθεί σε αύξουσα σειρά ορίζεται ως η μεσαία παρατήρηση όταν το πλήθος ν είναι περιττός αριθμός ή το ημιάθροισμα των δύο μεσαίων παρατηρήσεων όταν το ν είναι άρτιος αριθμός. Έτσι για παράδειγμα η διάμεσος των παρατηρήσεων

1, 5, 8, 9, 16, 32 είναι $δ = \frac{8+9}{2} = 8,5$

$\boxed{δ=8,5}$

Όταν οι παρατηρήσεις είναι πολλές για να υπολογίσουμε την μεσαία ή τις δύο μεσαίες παρατηρήσεις κατασκευάζουμε πίνακα αθροιστικών συχνοτήτων (N_i).

Αν ο ν είναι περιττός η μεσαία παρατήρηση είναι $t_{\frac{ν+1}{2}}$, ενώ όταν ο ν είναι άρτιος οι δύο μεσαίες παρατηρήσεις είναι $t_{\frac{ν}{2}}$ και $t_{\frac{ν}{2}+1}$.

Εφαρμογή 3

Δίνονται τα δεδομένα:

x_i	1	2	3	4	5
$ν_i$	43	87	125	38	16

Να βρεθεί η διάμεσος.

ΛΥΣΗ

Κατασκευάζουμε πίνακα αθροιστικών συχνοτήτων (N_i).

x_i	v_i	N_i
1	43	43
2	87	130
3	125	255
4	38	293
5	16	309
	Σύνολο	**309**

Οπότε η διάμεσος είναι $\delta = t_{155} = 3$.

Εφαρμογή 4

Η μεταβλητή Χ παρουσίασε σ' ένα δείγμα τις τιμές 5, 6, 7, 8, 9 με αντίστοιχες συχνότητες 50, 90, 70, 40, 30. Να υπολογισθεί η διάμεσος του δείγματος.

ΛΥΣΗ

Κατασκευάζουμε τον πίνακα αθροιστικών συχνοτήτων.

Επειδή το πλήθος των παρατηρήσεων είναι ν = 280 (άρτιος), τότε οι δύο μεσαίες παρατηρήσεις είναι

$t_{\frac{\nu}{2}} = t_{\frac{280}{2}} = t_{140} = 6$

και η επόμενή της $t_{\frac{\nu}{2}+1} = t_{140+1} = t_{141} = 7$ οπότε $\delta = \dfrac{6+7}{2} = 6{,}5$.

Όταν τα δεδομένα μιας μεταβλητής Χ είναι ομαδοποιημένες σε κλάσεις, τότε η διάμεσος **υπολογίζεται γραφικά** ακολουθώντας τα παρακάτω βήματα:

- Κατασκευάζουμε το ιστόγραμμα και το πολύγωνο των σχετικών αθροιστικών συχνοτήτων.
- Από το σημείο του κατακόρυφου άξονα που αντιστοιχεί στο 50% φέρουμε παράλληλη στον άξονα x΄x μέχρι να συναντήσει τη πολυγωνική γραμμή.
- Από το σημείο τομείς φέρνουμε κάθετη στον άξονα x΄x. Το ίχνος της καθέτου είναι και το σημείο στο οποίο αντιστοιχεί η διάμεσος.

Εφαρμογή 5

Έστω η μεταβλητή Χ: «βάρος 100 μαθητών» ομαδοποιημένων στο διπλανό πίνακα:

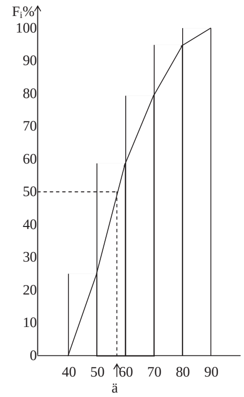
Κλάσεις	40-50	50-60	60-70	70-80	80-90
Συχνότητα	25	35	20	15	5

Να υπολογιστεί η διάμεσος.

ΛΥΣΗ

Κατασκευάζουμε τον πίνακα με στήλες v_i, f_i%, F_i%, και στη συνέχεια κατασκευάζουμε το ιστόγραμμα και το πολύγωνο των σχετικών αθροιστικών συχνοτήτων. Οπότε:

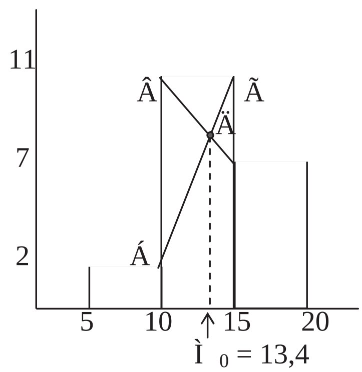
Κλάσεις	v_i	f_i	F_i%
40-50	25	25	25
50-60	35	35	60
60-70	20	20	80
70-80	15	15	95
80-90	5	5	100
Σύνολο	**100**	**100**	



Οπότε η διάμεσος είναι περίπου 57 κιλά.

Επικρατούσα τιμή (M_0) ονομάζουμε την παρατήρηση με την μεγαλύτερη συχνότητα. Η επικρατούσα τιμή ορίζεται και σε ποιοτικές και σε ποσοτικές μεταβλητές. Η επικρατούσα τιμή μπορεί να μην είναι μοναδική. Όταν έχουμε δύο παρατηρήσεις με την ίδια μέγιστη συχνότητα, τότε η κατανομή λέγεται δικόρυφη, ενώ όταν έχουμε πολλές κορυφές λέγεται πολυκόρυφη.

Σε ομαδοποιημένες κατανομές η επικρατούσα τιμή προσδιορίζεται γραφικά από το ιστόγραμμα συχνοτήτων.



4.4. ΕΚΑΤΟΣΤΗΜΟΡΙΑ - ΤΕΤΑΡΤΗΜΟΡΙΑ (ΜΕΤΡΟ ΘΕΣΗΣ)

Έχουμε ορίσει τη διάμεσο (δ) έτσι, ώστε το πολύ των 50% των παρατηρήσεων να είναι μικρότερες της διαμέσου και το πολύ το 50% να είναι μεγαλύτερος της δ. Με ανάλογο τρόπο ορίζουμε τα εκατοστημόρια ($P_κ$) με κ = 1, 2, 3, ..., 99.

Ορίζουμε ως $P_κ$ εκατοστημόριο την τιμή για την οποία το πολύ κ% των παρατηρήσεων είναι μικρότερο του $P_κ$ και το πολύ (100 - κ)% των παρατηρήσεων είναι μεγαλύτερο από το $P_κ$.

Ειδική περίπτωση των εκατοστημορίων αποτελούν τα P_{25}, P_{50}, P_{75} που λέγονται τεταρτημόρια και συμβολίζονται με Q_1, Q_2, Q_3 αντίστοιχα. Για να υπολογίσουμε τα τεταρτημόρια, έχουμε τα παρακάτω βήματα:

- Βρίσκουμε τη διάμεσο (δ) που είναι Q_2
- Βρίσκουμε για καθεμία από τις ομάδες των παρατηρήσεων που βρίσκονται αριστερά και δεξιά από τη διάμεσο δ την νέα διάμεσο και έτσι προσδιορίζουμε τα Q_1 και Q_3.

Ο υπολογισμός των εκατοστημορίων και τεταρτημορίων σε ομαδοποιημένες κατανομές γίνεται, όπως και στη διάμεσο, από το πολύγωνο αθροιστικών σχετικών συχνοτήτων.

Εφαρμογή 6

Να υπολογίσετε τα τεταρτημόρια Q_1, Q_2, Q_3 των παρατηρήσεων:

3, 5, 7, 9, 14, 15, 17, 19, 28.

ΛΥΣΗ

Επειδή το πλήθος ν = 9 (περιττό), τότε δ = 14, άρα $Q_2 = 14$.

Η αριστερή ομάδα είναι 3, 5, 7, 9, με διάμεσο $\delta = \dfrac{5+7}{2} = 6 = Q_1$,

ενώ η δεξιά ομάδα είναι 15, 17, 19, 28, με $\delta = \dfrac{17+19}{2} = 18 = Q_3$.

4.5. ΕΥΡΟΣ (ΜΕΤΡΟ ΔΙΑΣΠΟΡΑΣ)

Το εύρος (R) της μεταβολής μιας μεταβλητής ορίζεται ως η διαφορά της μικρότερης από τη μεγαλύτερη τιμή. Σε ομαδοποιημένα δεδομένα ως εύρος ορίζεται η διαφορά του κατώτερου ορίου της πρώτης κλάσης από το ανώτερο όριο της τελευταίας κλάσης.

4.6 ΕΝΔΟΤΕΤΑΡΤΗΜΟΡΙΑΚΟ ΕΥΡΟΣ (Q) (ΜΕΤΡΟ ΔΙΑΣΠΟΡΑΣ)

Το ενδοτεταρτημοριακό εύρος της μεταβολής μιας μεταβλητής ορίζεται ως η διαφορά του πρώτου τεταρτημορίου Q_1 από το τρίτο τεταρτημόριο Q_3, δηλαδή:

R=Μεγαλύτερη παρατήρηση – Μικρότερη παρατήρηση

Το ενδοτεταρτημοριακό εύρος έχει σημαντικό ρόλο σε μια κατανομή, διότι στο διάστημα Q_1, Q_3 περιλαμβάνεται το 50% των τιμών της μεταβλητής X. (75% - 25% = 50%)

$$\boxed{Q = Q_3 - Q_1}$$

Αυτό έχει σαν αποτέλεσμα όσο μικρότερη είναι η διαφορά $Q_3 - Q_1$, τόσο μικρότερη είναι και η διασπορά των τιμών της μεταβλητής.

4.7. ΔΙΑΚΥΜΑΝΣΗ - ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (s^2)

Ένα εργοστάσιο παραγωγής τυποποιημένων γλυκισμάτων παράγει τα ίδια γλυκίσματα με δύο μηχανές παραγωγής Α και Β. Ένας υπάλληλος, προκειμένου να κάνει δειγματοληπτικό έλεγχο για το βάρος των γλυκισμάτων πήρε 9 γλυκίσματα που παράχθηκαν από καθένα από τα μηχανήματα και τα ζύγισε.

Πήρε τα παρακάτω αποτελέσματα:

Α΄:	4,95	5,15	5	4,80	4,85	5,10	4,90	5,20	5,05
Β΄:	4,95	5,30	4,70	5,70	5,85	5	4,30	4,15	5,05

σε γραμμάρια.

Η μέση τιμή και των δύο δειγμάτων είναι ίδια: $\bar{x}_A = \bar{x}_B = 5$ γραμμάρια.

Επίσης η διάμεσος και των δύο δειγμάτων είναι $\delta_A = \delta_B = 5$.

Οπότε η μέση τιμή και η διάμεσος δεν μας δίνουν καμία διαφορά μεταξύ των δύο δειγμάτων.

Είναι φανερό ότι οι τιμές του Β΄ δείγματος είναι πιο διεσπαρμένες; Σχετικά με τη μέση τιμή και τη διάμεσο. Αν πάρουμε σε κάθε δείγμα τις αποκλίσεις κάθε τιμής από τη μέση τιμή, τότε έχουμε:

Α΄: (4,95 - 5) + (5,15 - 5) + (5 - 5) + (4,8 -5) +(4,85 - 5) + (5,1 - 5) + (4,9 - 5) + (5,2 - 5) + (5,05 - 5) = 0

Β΄: (4,95 - 5) + (5,3 - 5) + (4,7 - 5) + (5,7 - 5) +(5,8 - 5) + (5 - 5) +(4,3 - 5) + (4,15 - 5) + (5,05 - 5) = –0,05

Ούτε όμως και αυτό μας οδηγεί σε κάποιο συμπέρασμα.

Αν θεωρήσουμε όμως τις τιμές $\left(t_i - \bar{x}\right)^2$ και πάρουμε το μέσο όρο αυτών, έχουμε:

Α΄: $$\frac{(4.95-5)^2+(5.15-5)^2+(5-5)^2+(4.8-5)^2+(4.85-5)^2+(5.1-5)^2+(4.9-5)^2+(5.2-5)^2+(5.05-5)^2}{9} = \frac{0.15}{9} = 0.016$$

Β΄: $$\frac{(4.95-5)^2+(5.3-5)^2+(4.7-5)^2+(5.7-5)^2+(5.85-5)^2+(5-5)^2+(4.3-5)^2+(4.15-5)^2+(5.05-5)^2}{9} = \frac{2.61}{9} = 0.29$$

Παρατηρούμε ότι η διασπορά των τιμών που δίνει το Β΄ μηχάνημα είναι μεγαλύτερη.

Το μέσο όρο των παρατηρήσεων από τη μέση τιμή τους τον ονομάζουμε διακύμανση ή διασπορά και συμβολίζεται με s^2. Είναι:

$$\boxed{s^2 \frac{\left(t_1-\bar{x}\right)^2+\left(t_2-\bar{x}\right)^2+\ldots+\left(t_\nu-\bar{x}\right)^2}{\nu}}$$

Σε περίπτωση που οι παρατηρήσεις είναι κατανεμημένες σε πίνακα συχνοτήτων ή όταν έχουμε ομαδοποιημένες σε κλάσεις παρατηρήσεις στις οποίες θεωρούμε ως x_i την κεντρική τιμή κάθε κλάσης, τότε:

$$\boxed{s^2 = \frac{\nu_1\left(x_1-\bar{x}\right)^2+\nu_2\left(x_2-\bar{x}\right)^2+\ldots+\nu_\kappa\left(x_\kappa-\bar{x}\right)^2}{\nu}}$$

Οι τιμές των παρατηρήσεων στο προηγούμενο παράδειγμα είναι σε γραμμάρια, ενώ η διασπορά s^2 που υπολογίσαμε εμφανίζεται σε (γραμμάρια)^2.

Γι' αυτό το λόγο ορίζουμε την τυπική απόκλιση s, που είναι η τετραγωνική ρίζα της διακύμανσης, δηλαδή: $\boxed{s = \sqrt{s^2}}$

Η τυπική απόκλιση έχει το πλεονέκτημα ότι μετράται με τις ίδιες μονάδες, που μετριούνται οι παρατηρήσεις και δηλώνει την κατά μέσο όρο απόσταση της κάθε τιμής x_i από τη μέση τιμή $\bar{x}$. Όσο μεγαλύτερη είναι η τυπική απόκλιση, τόσο μεγαλύτερη είναι η απόσταση των παρατηρήσεων από τη μέση τιμή.

Αν σε ένα δείγμα η τυπική απόκλιση είναι μηδέν, τότε όλες οι παρατηρήσεις είναι ίσες με τη μέση τιμή.

Εφαρμογή 7

Οι εισπράξεις 100 πρατηρίων μιας εταιρείας καυσίμων (σε χιλιάδες ευρώ) την περίοδο των Χριστουγέννων του 2006 δίνονται στον παρακάτω πίνακα:

Κλάσεις	4-6	6-8	8-10	10-12	12-14	14-16	16-18
Συχνότητα	7	13	17	18	29	11	5

Να υπολογίσετε τη διακύμανση και την τυπική απόκλιση των εισπράξεων.

ΛΥΣΗ

Θα κατασκευάσουμε πίνακα που θα περιέχει στήλες $x_i v_i$ (πρέπει πρώτα να υπολογίσουμε τη μέση τιμή), $x_i - \bar{x}$, $(x_i - \bar{x})^2$ και $(x_i - \bar{x})^2 v_i$.

Έχουμε:

Κλάσεις σε χιλ ευρώ	Κεντ. Κλάσεις x_i	v_i	$x_i v_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$v_i(x_i - \bar{x})^2$
4-6	5	7	35	-6	36	252
6-8	7	13	91	-4	16	208
8-10	9	17	153	-2	4	68
10-12	11	18	198	0	0	0
12-14	13	29	377	2	4	116
14-16	15	11	165	4	16	176
16-18	17	5	85	6	36	180
ΣΥΝΟΛΑ		100	1104			1000

Οπότε η μέση τιμή $\bar{x} = \frac{1104}{100} = 11$

Οπότε η διακύμανση είναι: $s^2 = \frac{1000}{100} = 10$

και η τυπική απόκλιση $S = \sqrt{s^2} = \sqrt{10} = 3,16$.

4.8. ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΤΑΒΟΛΗΣ CV (ΜΕΤΡΟ ΔΙΑΣΠΟΡΑΣ)

Είδαμε προηγουμένως ότι όταν οι παρατηρήσεις σε δύο σύνολα εκφράζονται στις ίδιες μονάδες μέτρησης και έχουν το ίδιο ή περίπου ίδιο αριθμητικό μέσο, τότε η σύγκριση της μεταβλητότητας μπορεί να γίνει με την σύγκριση των τυπικών τους αποκλίσεων.

Όταν οι παρατηρήσεις στα δύο σύνολα εκφράζονται στις ίδιες μονάδες, αλλά ο αριθμητικός μέσος τους διαφέρει σημαντικά, τότε η σύγκριση της μεταβλητότητας μπορεί να γίνει με το συντελεστή μεταβλητότητας CV.

Ο συντελεστής μεταβολής ορίζεται από τη σχέση:

$$\boxed{CV = \frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}} = \frac{s}{\bar{x}}}$$

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

Ο συντελεστής μεταβολής είναι ανεξάρτητος από τις μονάδες μέτρησης, εκφράζεται σε ποσοστό και είναι ένα μέτρο που εκφράζει τη σχετική διασπορά των τιμών και όχι την απόλυτη διασπορά.

Ένα δείγμα ονομάζεται ομοιογενές όταν CV $\leq$ 10%.

Ο CV είναι αξιόπιστος μόνο όταν οι τιμές x_i είναι ομόσημες και η μέση τιμή $\bar{x}$ δεν είναι κοντά στο 0.

Εφαρμογή 8

Ένα προϊόν πωλείται σε 10 διαφορετικά καταστήματα στις παρακάτω τιμές σε ευρώ:

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

i) Να υπολογίσετε τη μέση τιμή, τη διάμεσο και την επικρατούσα τιμή

ii) Να υπολογίσετε το εύρος, την τυπική απόκλιση και τον συντελεστή μεταβολής

iii) Αν οι τιμές του προϊόντος σε όλα τα καταστήματα υποστούν έκπτωση 10%, να εξετάσετε αν θα μεταβληθεί ο συντελεστής μεταβολής.

ΛΥΣΗ

i) Έχουμε:

x_i	v_i	N_i	$x_i v_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$v_i(x_i - \bar{x})^2$
8	1	1	8	-5	25	25
0	1	2	9	-4	16	16
10	1	3	10	-3	9	9
13	2	5	26	0	0	0
14	2	7	28	1	1	2
15	1	8	15	2	4	4
16	1	9	16	3	9	9
18	1	10	18	5	25	25
Σύνολα	**10**		**130**			**90**

οπότε η μέση τιμή $\bar{x} = \frac{130}{10} = 13$ ευρώ.

Η διάμεσος δ είναι: δ = $\frac{13+14}{2}=13,5$.

Υπάρχουν δύο επικρατούσες τιμές 13, 14.

ii) Για το εύρος R, έχουμε: R = 18 – 8 = 10.

Η διακύμανση $s^2 = \frac{90}{10} = 9$, οπότε η τυπική απόκλιση s = 3.

Ο συντελεστής μεταβολής είναι $CV = \frac{s}{\bar{x}} = \frac{3}{13} \simeq 23\%$.

iii) Έστω y_i οι τιμές που θα προκύψουν από την έκπτωση 10% στα καταστήματα.

Τότε $y_i = x_i - \frac{10}{100} x_i = 0,9\,x_i$, όπου x_i οι αρχικές τιμές i = 1, 2, ..., 8.

Η μέση τιμή $\bar{y}$ θα είναι μειωμένη κατά 10% της αρχικής μέσης τιμής, δηλαδή $\bar{y} = 0,9\bar{x}$. Επίσης θα διαπιστώσουμε ότι και η τυπική απόκλιση $s_y = 0,9\,s_x$.

Οπότε $CV_y = \frac{s_y}{\bar{y}} = \frac{0,9s_x}{0,9\bar{x}} = \frac{s_x}{\bar{x}} = CV_x$, δηλαδή ο συντελεστής μεταβολής δεν μεταβάλλεται.

Παρατήρηση

- Όταν οι τιμές x_i ενός δείγματος πολλαπλασιάζονται με κάποιο αριθμό, τότε αντίστοιχα πολλαπλασιάζονται με τον ίδιο αριθμό η μέση τιμή και η τυπική απόκλιση.
- Όταν στις τις x_i μιας μεταβλητής προσθέσουμε κάποιο αριθμό, τότε η νέα μέση τιμή προκύπτει με πρόσθεση του ίδιου αριθμού, ενώ η τυπική απόκλιση παραμένει αμετάβλητη.

ΕΞΑΣΚΗΣΗ

1. Οι βαθμοί ενός μαθητή στα μαθήματά του κατά το Α΄ τετράμηνο ήταν:

16, 15, 14, 13, 19, 15, 18, 12, 13, 11.

Να βρεθεί η μέση βαθμολογία του μαθητή.

2. Το μέσο ημερομίσθιο 30 εργατών μιας κατασκευαστικής εταιρείας είναι 30 ευρώ. Οι οκτώ είναι ειδικευμένοι και έχουν μέσο ημερομίσθιο 57,5 ευρώ. Ποιο είναι το ημερομίσθιο των ανειδίκευτων εργατών;

3. Σε μια κατανομή που έπρεπε να έχει 30 παρατηρήσεις καταχωρήθηκαν από λάθος δύο ακόμα παρατηρήσεις με αντίστοιχες τιμές 20 και 40. αν η μέση τιμή που βρέθηκε είναι 50, να υπολογιστεί η σωστή μέση τιμή.

4. Ένα δείγμα εργαζομένων μιας εταιρείας εξετάστηκε ως προς το χρόνο (σε ώρες) υπερωριακής απασχόλησης κατά τη διάρκεια ενός μηνός και προέκυψε ο διπλανός πίνακας.

Ώρες υπερωριακής απασχόλησης Κλάσεις	Αθροιστική συχνότητα N_i
0 – 2	5
2 – 4	15
4 – 6	20
6 – 8	35
8 – 10	40

Να βρείτε:

i) Το μέγεθος του δείγματος.

ii) Τις συχνότητες και τις σχετικές συχνότητες των κλάσεων.

iii) Τη μέση τιμή.

5. Ο παρακάτω πίνακας δίνει την ποσοστιαία κατανομή των νοικο κυριών αστικών περιοχών ως προς τον αριθμό των δωματίων. Να βρεθεί η μέση τιμή της μεταβλητής Χ: αριθμός δωματίων.

Αριθμός δωματίων x_i	Αναλογία νοικορ. f_i%
1	12
2	28
3	32
4	20
5	6
6	2
ΣΥΝΟΛΟ	100

6. Για τον έλεγχο της κατανομής καυσίμων (ίδιου τύπου) δύο αυτοκινήτων Α και Β μετρήθηκε η κατανάλωσή τους σε έξι διαδρομές για το αυτοκίνητο Α και σε πέντε διαδρομές για το αυτοκίνητο Β. Η κατανάλωση του αυτοκινήτου Α (σε λίτρα ανά 100 χιλιόμετρα) ήταν για τις έξι διαδρομές 9, 6, 7, 9, 9, 8 ενώ η κατανάλωση στις πέντε διαδρομές για το αυτοκίνητο Β ήταν 8, 10, 7, 8, 12.

i) Να υπολογίσετε τη μέση τιμή και τη διάμεσο των μετρήσεων που αφορούν το αυτοκίνητο Α.

ii) Να υπολογίσετε τη μέση τιμή και την διάμεσο που αφορούν το αυτοκίνητο Β.

iii) Αν ένας πωλητής ήθελε να χρησιμοποιήσει τα πιο πάνω δεδομένα για να πείσει ένα υποψήφιο αγοραστή να αγοράσει το αυτοκίνητο Α και όχι το Β, ποιο μέτρο θέσης (μέση τιμή ή διάμεσο) θα χρησιμοποιούσε; Αν αντίστροφα ήθελε να πείσει τον υποψήφιο αγοραστή να αγοράσει το αυτοκίνητο Β και όχι το Α, ποιο μέτρο θέσης θα χρησιμοποιούσε;

7. Ο παρακάτω πίνακας δίνει τις εισπράξεις σε (δεκ. χιλ. €), που έγιναν από τους αντιπροσώπους μιας εταιρείας αυτοκινήτων κατά τη διάρκεια ενός μήνα:

Κλάσεις	1-3	3-5	5-7	7-9	9-11
Αριθ. αντιπρ.	14	6	12	10	8

i) Να συμπληρώσετε πίνακα συχνοτήτων απολύτων και σχετικών

ii) Να υπολογίσετε την μέση τιμή και τη διάμεσο δ.

8. Μια τάξη έχει 75 μαθητές. Η βαθμολογία τους στα Μαθηματικά κυμαίνεται από 10 - 20. από αυτούς 39 μαθητές έχουν βαθμολογία κάτω από 14, 15 μαθητές κάτω από 12,9 μαθητές πάνω από 18 και 21 μαθητές πάνω από 16.

i) Να παραστήσετε τα δεδομένα σε πίνακα συχνοτήτων.

ii) Να υπολογίσετε τη μέση τιμή.

iii) Να υπολογίσετε τη διάμεσο.

iv) Να βρεθεί πάνω από ποια βαθμολογία βρίσκεται το 20% των μαθητών.

9. Στον διπλανό πίνακα δίνονται τα αποτελέσματα της ρίψης ενός ζαριού 40 φορές. Να βρείτε τους φυσικούς αριθμούς α, β ώστε να υπάρχουν δύο επικρατούσες τιμές.

x_i	v_i
1	5
2	6
3	α
4	β
5	7
6	9

10. Μια εταιρία απασχολεί 20 εργαζομένους εκ των οποίων οι 10 εργάζονται στο τμήμα Α και οι 10 στο τμήμα Β. Η μέση τιμή των μηνιαίων μισθών του τμήματος Α είναι 720€ και ο μεγαλύτερος μισθός του τμήματος είναι 900€. Οι μισθοί των εργαζομένων στο τμήμα Β είναι: 950, 900, 1060, 980, 920, 945, 930, 900, 940. Να βρείτε:

i) Το άθροισμα των μηνιαίων μισθών του τμήματος Α

ii) Τη μέση τιμή και την επικρατούσα τομή του τμήματος Β

iii) Τη μέση τιμή και τη διάμεσο των μισθών όλων των εργαζομένων στην εταιρεία.

11. Οι εισπράξεις (σε χιλιάδες ευρώ) ενός δείγματος δέκα υποκαταστημάτων μιας εμπορικής επιχείρησης κατά τη διάρκεια ενός μήνα ήταν:

50, 15, 15, 20, 15, 30, 15, 20, 50, 50

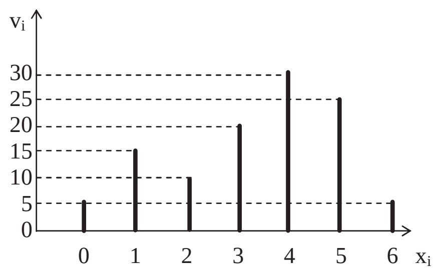
i) Να υπολογίσετε τη μέση τιμή   των εισπράξεων

ii) Να συμπληρώσετε τον παρακάτω πίνακα:

Εισπράξεις (σε χιλ. €) x_i	v_i	f_i	$x_i - \bar{x}$	$\left(x_i - \bar{x}\right)^2$	$v_i\left(x_i - \bar{x}\right)^2$
15					
20					
30					
50					
Σύνολο					

iii) Να υπολογίσετε την διακύμανση και την τυπική απόκλιση των εισπράξεων.

12. Στις εξετάσεις ενός ξενόγλωσσου διπλώματος υπαγο ρεύτηκε σε 110 μαθητές ένα κεί μενο. Κατά την διόρθωση των γραπτών ως προς τα ορθογραφικά λάθη προέκυψε το διπλανό διά γραμμα. Να υπολογίσετε τη μέση τιμή, την επικρατούσα τιμή, την διασπορά και την τυπική απόκλιση των λαθών.



13. Η εξέταση 10 μαθητών στο μάθημα της Στατιστικής έδωσε του εξής βαθμούς:

11, 3, 7, 5, 16, 14, 11, 10, 11, 12

Να βρείτε:

i) τη διάμεσο

ii) τη μέση τιμή

iii) την επικρατούσα τιμή

iv) το εύρος

v) τη διακύμανση της παραπάνω βαθμολογίας.

14. Η βαθμολογία ενός μαθητή σε 10 μαθήματα είναι:

13, 9, 8, 7, 10, 15, 12, 11, 17, 18

Να υπολογίσετε:

i) τη μέση τιμή

ii) τη διακύμανση s_2

iii) την τυπική απόκλιση S

iv) τη διάμεσο δ

v) τα τεταρτημόρια Q_1, Q_2, Q_3

vi) το ενδοτεταρτημοριακό εύρος των βαθμών

vii) το εύρος R

viii) το συντελεστή μεταβολής CV

15. Μια ομάδα ανθρώπων έχει μέσο βάρος 73 κιλά με τυπική απόκλιση 4 κιλά, ενώ το μέσο ύψος τους είναι 170cm με τυπική απόκλιση 7cm. Να εξετάσετε ποια κατανομή του βάρους ή του ύψους παρουσιάζει μεγαλύτερη ομοιογένεια.

16. Μια πολυεθνική εταιρεία απασχολεί υπαλλήλους στις Η.Π.Α. και στην Ευρώπη. Οι μηνιαίοι μισθοί των υπαλλήλων στις Η.Π.Α. έχουν μέση τιμή = 1000 δολάρια και τυπική απόκλιση SA = 125 δολάρια. Οι μηνιαίοι μισθοί των υπαλλήλων στην Ευρώπη έχουν μέση τιμή = 800 ευρώ και τυπική απόκλιση SE = 90 ευρώ.
i) Να βρείτε σε ποιο μέρος παρουσιάζεται μεγαλύτερη ομοιογένεια μισθών.
ii) Η εταιρεία αποφασίζει στις Η.Π.Α. να αυξήσει τον μηνιαίο μισθό κάθε υπαλλήλου κατά 250 δολάρια, ενώ στην Ευρώπη αποφασίζει να αυξήσει τον μηνιαίο μισθό κάθε υπαλλήλου κατά 20%. Να βρείτε την νέα μέση τιμή και την νέα τυπική απόκλιση των μηνιαίων μισθών και στις δύο ηπείρους.
iii) Σε ποιο μέρος παρουσιάζεται μεγαλύτερη ομοιογένεια μισθών μετά τις αυξήσεις;

17. Από την απογραφή πληθυσμού της 18ης Μαρτίου 2001 προέκυψε ο παρακάτω πίνακας:
Πίνακας 5. Μόνιμος πληθυσμός κατά ομάδες ηλικιών

ΗΛΙΚΙΑ	ΣΥΧΝΟΤΗΤΑ
[0,5)	529399
[5,10)	545105
[10,15)	586395
[15,20)	726174
[20,25)	835463
[25,30)	847427
[30,35)	869932
[3540)	783413
[40,45)	781943)
[45,50)	713975
[50,55)	687349
[55,60)	560215
[60,65)	640074
[65,70)	623245
[70,75)	545018
[75,80)	328918
[80,85)	188193
[85-90)	105194
[90-95)	29500
[95-100)	5467
[100-105)	1698

Πηγή: Ε.Σ.Υ.Ε (Απογραφή πληθυσμού της 18ης Μαρτίου 2001)

α) Να κατασκευαστεί πίνακας σχετικών συχνοτήτων και απόλυτων αθροι στικών, σχετικών αθροιστικών συχνοτήτων, το κατάλληλο διάγραμμα και να βρεθούν η μέση τιμή, η διάμεσος, το ενδοτεταρτημοριακό πλάτος και η διασπορά του πληθυσμού.

ΔΡΑΣΤΗΡΙΟΤΗΤΑ

Να υπολογίσετε τα μέτρα θέσης και διασποράς, για όσες μεταβλητές είναι δυνατόν, με τα στοιχεία που έχετε συγκεντρώσει από τους υπόλοιπους εκπαιδευόμενους στις προηγούμενες δραστηριότητες.

Σύνοψη

Οι πίνακες κατανομής συχνοτήτων δεν επιτρέπουν την εύκολη συγκριση. Για τον λόγο αυτό αντικαθίστανται από ορισμένους αντιπροσωπευτικούς αριθμούς που ανομάζονται στατιστικές παράμετροι και χαρακτηρίζουν τη θέση τη διασπορά και τη μορφολογία του πληθυσμού που ερευνούμε και διευκολύνουν πάρα πολύ στις συγκρίσεις ομοειδών ομάδων.

- **Στατιστικές παράμετροι**
  - **Μέτρα θέσης**
    - αριθμητικός μέσος
    - διάμεσος
    - εκατοστημόρια
  - **Μέτρα διασποράς**
    - εύρος
    - ενδοτεταρτημόριακό
    - διακύμανση
    - συντελεστής μεταβολής

Μέτρα θέσης: Οι παράμετροι αυτές ορόζουν το σημείο γύρω από το οποίο συσσωρεύονται οι δάφορες τιμές ενός πληθυσμού και βρίσκονται στο κέντρο μίας σειράς παρατηρήσεων. Αντικειμενικός σκοπός αυτών των παραμέτρων είνατι η αντιπροσώπευση του πληθυσμού από την άποψη μιας μεταβλητής με τον πιο απλό τρόπο.

Μέτρα διασποράς: Η αντιπροσώπευση ενός πληθυσμού με τα μέτρα θέσης έχει αξία με δεδομένο ότι ο πληθυσμός παρουσιάζει μεγάλη ομοιογένεια. Αν ο πληθυσμός παρουσιάζει μεγάλη ανομοιογένεια, τότε τα μέτρα θέσης θα πρέπει να μη χρησιμοποιούνται αντ’ αυτών να χρησιμοποιούνται τα μέτρα διασποράς που μας δίνουν το βαθμό συγκέντρωσης η διασπορά των τιμών της μεταβλητής.

Βιβλιογραφία / Internet

«Εισαγωγή στη στατιστική» Μέρος 1, Δαμιανού Χ., Κούτρας Μ.
«Statistics, Teach yourself»
«Στατιστική», *Κιόχος Α. Π.*
«Πιθανότητες και στατιστική», Spiegel, Schaum’ s outline Series www.asset-ahalysis.com

5. ΑΡΙΘΜΟΔΕΙΚΤΕΣ

Βασικές έννοιες:
- αριθμοδείκτης
- σχετικές τιμές
- σχετικές ποσότητες
- ιδιαίτεροι αριθμοδείκτες
- συνθετικοί αριθμοδείκτες

Στόχος του μαθήματος: η ανάδειξ των αριθμοδεικτών ως χρηστικών εργαλείων σε πολλούς τομείς της ανθρώπινης δραστηριότητας.

Προσδοκώμενα αποτελέσματα: ο εκπαιδευόμενος να κατανοήσει ότι οι αριθμοδείκτες αποτελούν ένα εργαλείο, η χρήση του οποίου μας δίνει τη δυνατότητα να εξάγουμε χρήσιμα συμπεράσματα σε διάφορα οικονομικά προβλήματα,

Εισαγωγικές παρατηρήσεις: οι αριθμοδείκτες είναι στατιστικά μέτρα, που μας δείχνουν τις μεταβολές μίας μεταβλητής ή ομάδας μεταβλητών, που σχετίζονται μεταξύ τους μεταξύ δύο χρονικών περιόδων, δύο περιοχών και γενικά μεταξύ δύο καταστάσεων.

Η χρήση των αριθμοδεικτών αποτελεί μία από τις πλέον διαδεδομένες και δυναμικές μεθόδους χρηματοοικονομικής ανάλυσης.

Με τη βοήθεια των αριθμοδεικτών μπορούμε να συγκρίνουμε π.χ. το κόστος διατροφής της Ελλάδας κατά την περίοδο 2005, με το κόστος διατροφής το 1980. Ή να συγκρίνουμε την παραγωγή των γαλακτοκομικών προϊόντων μίας χώρας με την παραγωγή μίας άλλης χώρας, την ίδια περίοδο.

Στην πρώτη περίπτωση έχουμε χρονολογικούς δείκτες, ενώ στη δεύτερη ονομάζονται Γεωγραφικοί δείκτες.

Ο λόγος που οδήγησε στην καθιέρωση της χρήσης των αριθμοδεικτών προέρχεται από την ανάγκη να γίνεται άμεσα αντιληπτή η πραγματική αξία και η σπουδαιότητα των απόλυτων μεγεθών. Π.χ. αν η εταιρία Α παρουσιάζει κέρδη 200.000€ σε μία συγκεκριμένη χρήση, αυτό ως αποτέλεσμα μπορεί να θεωρείται ικανοποιητικό, ίσως όμως και όχι.

Αν το απόλυτο αυτό μέγεθος το συσχετίσουμε με τα ίδια κεφάλαια της επιχείρησης, π.χ. 1.000.000€, τότε έχουμε μία αποδοτικότητα των ίδιων κεφαλαίων της τάξης του 20%.

Ο αριθμοδείκτης αυτός (20%) έχει για τον ενδιαφερόμενο πολύ μεγαλύτερη σημασία από ό,τι το απόλυτο ύψος των κερδών 200.000€.

Τους αριθμοδείκτες μπορούμε να τους χωρίσουμε σε δύο μεγάλες κατηγορίες:

ΑΡΙΘΜΟΔΕΙΚΤΕΣ

Που μας δίνουν τις μεταβολές των μεταβλητών στο χρόνο
ΧΡΟΝΟΛΟΓΙΚΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ

ΙΔΙΑΙΤΕΡΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ	ΣΥΝΘΕΤΙΚΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ
Μας δείχνουν τις μεταβολές μίας μόνο μεταβλητής μεταξύ δύο χρονικών περιόδων	Αυτοί εκφράζουν τις μεταβολές πολλών μεταβλητών του ίδιου φαινομένου με συνδυασμό των ιδιαίτερων δεικτών των επί μέρους μεταβλητών

Στη χρηματοοικονομική ανάλυση οι κυριότερες κατηγορίες αριθμοδεικτών είναι:

- Αριθμοδείκτες ρευστότητας
- Αριθμοδείκτες δραστηριότητας
- Αριθμοδείκτες αποδοτικότητας
- Αριθμοδείκτες διάρθρωσης κεφαλαίου και βιωσιμότητας
- Επενδυτικοί αριθμοδείκτες

5.1. ΙΔΙΑΙΤΕΡΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ

Ονομάζουμε σχετική τιμή το λόγο της τιμής ενός αγαθού σε μία χρονική περίοδο, προς την τιμή του ίδιου αγαθού σε μία άλλη χρονική περίοδο που ονομάζεται περίοδος βάσης.

Συνήθως, με p_0 συμβολίζεται η τιμή ενός αγαθού στην περίοδο 0, που θεωρείται ως βάση και p_1 η τιμή ενός αγαθού σε μία άλλη περίοδο 1.

Τότε, ως σχετική τιμή στην περίοδο 1 με βάση την περίοδο 0 ορίζεται ως το πηλίκο $\frac{p_1}{p_0}$ 100, που συμβολίζεται με $P_{1/0}$ δηλ. $P_{1/0}=\frac{p_1}{p_0}\cdot 100\%$.

Προσοχή!! Η σχετική τιμή που αντιστοιχεί στην περίοδο βάσης να είναι 100.

Εφαρμογή 1

Η τιμή πώλησης σοκολάτας το 2001 και το 2005 ήταν 1,2€ και 1,8€ αντίστοιχα. Αν πάρουμε το 2001 ως έτος βάσης. Να υπολογίσετε τη σχετική τιμή .

ΛΥΣΗ

Η σχετική τιμή είναι $P_{1/10}=\frac{p_1}{p_0}$ $100=\frac{1,8}{1,2}$ 100=150%.

Δηλαδή το 2005 με βάση το 2001 η τιμή της σοκολάτας ανέβηκε (αυξήθηκε) 50%.

Πολλές φορές, αντί να συγκρίνουμε τιμές, μπορούμε να συγκρίνουμε ποσότητες ή όγκους και στις περιπτώσεις αυτές μπορούμε πάλι να υπολογίσουμε τις σχετικές ποσότητες.

Συμβολίζουμε με την ποσότητα (ή V_0 τον όγκο) ενός αγαθού κατά την περίοδο 0 και με την ποσότητα (ή V_1 τον όγκο) του ίδιου αγαθού στην περίοδο 1. Τότε ο σχετικός δείκτης ποσοτήτων (όγκου) με βάση την περίοδο 0 ορίζεται όπως και προηγούμενα σαν το πηλίκο $\frac{q_1}{q_0}$ 100 $\rightarrow$ $Q_{1/10}$,

Δηλαδή $\boxed{Q_{1/10}\ \frac{q_1}{q_0}\cdot 100\%}$

Μπορούμε να δούμε και τη μεταβολή της αξίας και να ορίσουμε τη σχετική αξία.

Αν υποθέσουμε ότι p_0 είναι η τιμή ενός αγαθού την περίοδο 0 q_0 είναι η ποσότητα του αγαθού την ίδια περίοδο

Τότε το γινόμενο $p_0q_0=u_0$ είναι η αξία του αγαθού την περίοδο 0, ενώ $P_1Q_1=u_1$, είναι η αξία του αγαθού την περίοδο 1.

Τότε η σχετική αξία του αγαθού για την περίοδο 1 με βάση την περίοδο 0 είναι:

Δηλ. $U_{1/0} = \frac{p_1q_1}{p_0q_0} = \frac{p_1}{p_0}\cdot\frac{q_1}{q_0} = P_{1/10}\cdot Q_{1/10}$

Εφαρμογή 2

Δίνονται οι τιμές σε (€) και οι ποσότητες που καταναλώθηκαν σε kg, ενός αγαθού για τις χρονιές 2005 και 2006.

ΕΤΟΣ	ΤΙΜΗ	ΠΟΣΟΤΗΤΑ
2005	3,5	160
2006	4,2	165

Να υπολογίσετε το σχετικό δείκτη αξίας για το 2006 με βάση το έτος 2005.

ΛΥΣΗ

Γνωρίζουμε ότι ο σχετικός δείκτης αξίας δίνεται από τον τύπο:

$$U_{1/0} = \frac{p_1 q}{p_0 q_0} = \frac{3,5\cdot 160}{4,2\cdot 165}\ 100\% = 80,8\%$$

5.2. ΣΥΝΘΕΤΙΚΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ

Στην καθημερινότητά μας, συνήθως, δεν μας ενδιαφέρει πώς υπολογίζεται ο δείκτης ενός μόνο αγαθού μεταξύ δύο χρονικών περιόδων, αλλά το πώς μεταβάλλονται μεγάλες ομάδες των αγαθών αυτών.

Στην περίπτωση αυτή χρησιμοποιούμε τους συνθετικούς αριθμοδείκτες.

ΣΥΝΘΕΤΙΚΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ

Απλοί συνθετικοί αριθμοδείκτες

Σταθμικοί συνθετικοί αριθμοδείκτες

Στην περίπτωση που θεωρούμε ότι οι τιμές των αγαθών έχουν συντελεστή στάθμισης τη μονάδα, δηλαδή ότι όλα τα αγαθά έχουν την ίδια βαρύτητα, τότε χρησιμοποιούμε τους απλούς συνθετικούς δείκτες. Για τον υπολογισμό τους έχουμε δύο μεθόδους:

ΜΕΘΟΔΟΣ ΤΩΝ ΣΥΝΟΛΙΚΩΝ ΤΙΜΩΝ

Αν συμβολίσουμε με $p_0^{(1)}, p_0^{(2)}, \ldots p_0^{(v)}$ τις τιμές των αγαθών 1,2,3,...v στην περίοδο βάσης 0 και $p_1^{(1)}, p_1^{(2)}, \ldots p_1^{(v)}$ τις τιμές των αντίστοιχων προϊόντων στην περίοδο 1, τότε ο δείκτης των συνολικών τιμών βρίσκεται ως εξής:

ΜΕΙΟΝΕΚΤΗΜΑΤΑ

$$P_{1/0} = \frac{p_1^{(1)} + p_1^{(2)} + \ldots p_1^{(v)}}{p_0^{(1)} + p_0^{(2)} + \ldots p_0^{(v)}} = \frac{\sum p_1}{\sum p_0}$$

- Όλα τα αγαθά δεν έχουν ίδια βαρύτητα
- Δεν παίρνει υπόψη του τις ποσότητες των αγαθών

ΜΕΘΟΔΟΣ ΤΟΥ ΑΡΙΘΜΗΤΙΚΟΥ ΜΕΣΟΥ ΤΩΝ ΣΧΕΤΙΚΩΝ ΤΙΜΩΝ

Αν έχουμε ν αγαθά και διαιρέσουμε το άθροισμα των σχετικών τιμών ενός αγαθού με το ν.
Τότε προκύπτει άμεσος αριθμητικός των σχετικών τιμών.
Δηλαδή:

ΜΕΙΟΝΕΚΤΗΜΑΤΑ

$$P_{1/0} = \frac{\frac{p_1^{(1)}}{p_0^{(1)}} + \frac{p_1^{(2)}}{p_0^{(2)}} + \ldots \frac{p_1^{(v)}}{p_0^{(v)}}}{v} = \frac{\sum \left(\frac{p_1}{p_0} \right)}{v}$$

- Οι μονάδες που χρησιμοποιούνται για τον χαρακτηρισμό των τιμών κιλά, τόνοι κλπ., επηρεάζουν το δείκτη
- Δίνει ίση βαρύτητα σε όλα τα αγαθά, ανεξάρτητα της σπουδαιότητάς τους.

Εφαρμογή 3

Στον παρακάτω πίνακα δίνονται σε € οι τιμές τεσσάρων προϊόντων για τα έτη 2005 και 2006. Να υπολογίσετε:

i) Το δείκτη συνολικών τιμών.

ii) Μέσο αριθμητικό των σχετικών τιμών.

ΑΓΑΘΑ	ΤΙΜΕΣ		ΣΧΕΤΙΚΕΣ ΤΙΜΕΣ
	2005	2006	
	p_0	p_1	p_1/p_0
Α	2	4	2
Β	5	7	1,41
Γ	3	9	3
Δ	10	15	1,5
ΣΥΝΟΛΟ	**20**	**35**	**7,91**

ΛΥΣΗ

i) Ο δείκτης συνολικών τιμών θα είναι $P_{1/0} = \frac{\sum p_1}{\sum p_0} = \frac{35}{20} = 1,75$ ή 175%.

Δηλαδή οι τιμές το έτος 2006 είναι 75% μεγαλύτερες εκείνων του 2005.

ii) Ο μέσος αριθμητικός των σχετικών τιμών θα είναι:

$$P_{1/0} = \frac{1}{v}\sum\left(\frac{p_1}{p_0}\right) = \frac{1}{4}7,91 = 1,977 \quad \text{ή } 197,7\%$$

5.3. ΣΤΑΘΜΙΚΟΙ ΣΥΝΘΕΤΙΚΟΙ ΑΡΙΘΜΟΔΕΙΚΤΕΣ

Για να αποφύγουμε τα μειονεκτήματα των απλών συνδετικών δεικτών, χρησιμοποιούμε τους σταθμικούς συνθετικούς αριθμοδείκτες. Με αυτούς σταθμίζουμε τις τιμές κάθε αγαθού με ένα συντελεστή στάθμισης, που συνήθως είναι η ποσότητα, ο όγκος του αγαθού σε μία συγκεκριμένη χρονική περίοδο.

Οι πλέον σημαντικοί σταθμικοί συνθετικοί αριθμοδείκτες είναι αυτοί του τιμαρίθμου.

ΤΥΠΟΙ ΤΙΜΑΡΙΘΜΩΝ

Αν ως συντελεστές στάθμισης χρησιμοποιούνται οι ποσότητες της περιόδου Βάσης, τότε η μέθοδος για την κατάρτιση του δείκτη τιμών ονομάζεται δείκτης Luspeyres:

$$P_{1/0} = \frac{\sum p_1 q_0}{\sum p_0 q_0}$$

Στο δείκτη αυτό θεωρούμε ότι οι ποσότητες που αγοράσαμε στην περίοδο βάσης 0 είναι ίδιες και κατά την περίοδο 1.

Αν ως συντελεστές στάθμισης χρησιμοποιούνται οι ποσότητες της περιόδου 1 (τρέχουσα περίοδος) τότε η μέθοδος για την κατάρτιση του δείκτη τιμών ονομάζεται δείκτης Paasche:

$$P_{1/0} = \frac{\sum p_1 q_1}{\sum p_0 q_1}$$

Στο δείκτη αυτό θεωρούμε ότι οι ποσότητες που αγοράσαμε στην περίοδο 1 είναι οι ίδιες και κατά την περίοδο βάσης 0.

Εφαρμογή 4

Στον διπλανό πίνακα μας δίνουν τις τιμές σε € και τις μονάδες που έχουν παραχθεί σε χιλιάδες για τα προϊόντα Α, Β, από ένα εργοστάσιο για τα έτη 2003 και 2005. Έτος βάσης το 2003=100

	2003		2005	
ΠΡΟΪΟΝΤΑ	p_0	q_0	p_1	q_1
Α	20	6	24	4
Β	8	100	8	160

Να βρείτε:

i) Δείκτη Luspeyres

ii) Δείκτη Paasche

ΛΥΣΗ

Κατασκευάζουμε τον πίνακα αριθμητικών υπολογισμών.

ΠΡΟΪΟΝΤΑ	p_0	q_0	p_1	q_1	p_0q_0	p_0q_0	p_1q_0	p_1q_1
A	20	24	6	4	120	80	144	96
B	8	8	100	160	800	1.280	800	1.280
ΣΥΝΟΛΑ					**920**	**1.360**	**944**	**1.376**

Ο δείκτης τιμών κατά Luspeyres είναι:

$$P_{1/0} = \frac{\sum p_1 q_0}{\sum p_0 q_0} = \frac{944}{920} = 1,02 = 102\%$$

Ο δείκτης τιμών κατά Paasche είναι:

$$P_{1/0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{1376}{1360} = 1,01 = 101\%$$

Βασικές ομάδες Αριθμοδεικτών στην ανάλυση Λογιστικών καταστάσεων

Η στατική θεώρηση των δεικτών μιας επιχείρησης έχει μικρή πρακτική σημασία. Οι δείκτες πρέπει να εξετάζονται διαχρονικά, για να διαπιστωθεί η τάση τους και συγκριτικά με το μέσο όρο του κλάδου ή ορισμένες φορές, καλύτερα με τον μέσο όρο των πλησιέστερων από πλευράς μεγέθους και αντικειμένου, μονάδων του κλάδου.

Επισημαίνεται ότι για την αξιολόγηση επιχειρήσεων σε τομείς όπως ο τραπεζιτικός, ασφαλιστικός, χρηματιστηριακός και γενικά σε τομείς εξειδικευμένους ή με ειδικό καθεστώς εποπτείας, απαιτείται και η χρησιμοποίηση εξειδικευμένων δεικτών, πέρα από τους κλασικούς δείκτες, που χρησιμοποιούνται από τη βιομηχανία, το εμπόριο και τους τομείς υπηρεσιών.

Από την οπτική γωνία προσέγγισης της επιχείρησης εξαρτάται και η ομάδα δεικτών που χρησιμοποιείται για την ανάλυση.

Οι επενδυτές επικεντρώνουν την προσοχή τους στους δείκτες αποδοτικότητας και τις αξίες της καθαρής θέσης της επιχείρησης, ενώ οι πιστωτές στους δείκτες ρευστότητας, δανειακής επιβάρυνσης και επάρκειας του cash-flow.

Οι αριθμοδείκτες αποτελούν προσπάθεια ποσοτικής αξιολόγησης δύο βασικών κριτηρίων χρηματοδότησης:

- Του κριτηρίου κεφαλαιακής επάρκειας
- Της επάρκειας ταμειακών ροών.

Η μεταβολή ενός δείκτη επιδέχεται πολλές φορές διάφορες ερμηνείες. Όπως, για παράδειγμα, η μείωση ταχύτητας κυκλοφορίας των αποθεμάτων δεν αποτελεί μόνο αποτέλεσμα προβλημάτων στις πωλήσεις, αλλά ενδέχεται να αντανακλά πολιτική αποθεματοποιήσεων ενόψει αυξήσεων τιμών των πρώτων υλών.

Στη συνέχεια θα αναφερθούμε σε βασικές κατηγορίες αριθμοδεικτών, που χρησιμοποιούνται για την ανάλυση της πιστοληπτικής ικανότητας μίας επιχείρησης.

5.4. ΑΡΙΘΜΟΔΕΙΚΤΕΣ ΡΕΥΣΤΟΤΗΤΩΝ

Οι δείκτες ρευστότητας χρησιμοποιούνται στην αξιολόγηση της ικανότητας της επιχείρησης να ανταποκρίνεται στις βραχυπρόθεσμες υποχρεώσεις της.

Η επιχείρηση πρέπει να έχει τη δυνατότητα:

- Να εξοφλεί τους βραχυπρόθεσμους πιστωτές
- Να καταβάλει τους μισθούς, τους οφειλόμενους τόκους
- Να διατηρεί μια υγιή πιστοληπτική ικανότητα

Από τους αριθμοδείκτες ρευστότητας, οι κυριότεροι είναι:

1. Αριθμοδείκτης έμμεσης ρευστότητας (current ratio ή γενικής ρευστότητας ή λυκλοφοριακής ρευστότητας ή κεφαλαίου κίνησης)

Ο αριθμοδείκτης αυτός είναι ο πλέον χρησιμοποιούμενος δείκτης. Ο αριθμοδείκτης γενικής ρευστότητας προκύπτει αν διαιρέσουμε το σύνολο των κυκλοφοριακών στοιχείων μίας εταιρίας με το σύνολο των βραχυχρόνιων υποχρεώσεών της.

$$\text{Αριθμοδείκτης έμμεσης ρευστότητας} = \frac{\text{Κυκλοφορούν Ενεργητικό}}{\text{Βραχυπρόθεσμες υποχρεώσεις}}$$

Το κυκλοφορούν ενεργητικό περιλαμβάνει τα αποθέματα, τους εισπρακτέους λογαριασμούς, τα χρεόγραφα και τα χρηματικά διαθέσιμα, ενώ οι βραχυπρόθεσμες υποχρεώσεις αποτελούνται από τους πληρωτέους λογαριασμούς, τα πληρωτέα γραμμάτια, τις βραχυπρόθεσμες υποχρεώσεις, όπως βραχυχρόνια τραπεζικά δάνεια, οφειλόμενους τόκους κ.α.

Όσο μεγαλύτερος είναι ο αριθμοδείκτης γενικής ρευστότητας, τόσο καλύτερη από πλευράς ρευστότητας είναι η θέση της συγκεκριμένης επιχείρησης.

Γενικά, ένας αριθμοδείκτης γενικής ρευστότητας κοντά στο 2 θεωρείται συνήθως ικανοποιητικός για μία βιομηχανική ή εμπορική επιχείρηση.

Αν θέλουμε να συγκρίνουμε δύο επιχειρήσεις ως προς τη δυνατότητα που έχουν να ανταποκρίνονται στις τρέχουσες υποχρεώσεις τους, συνήθως υποθέτουμε ότι εκείνη που έχει το μεγαλύτερο κεφάλαιο κίνησης είναι αυτή που έχει και την καλύτερη ρευστότητα.

Αυτό όμως δεν είναι απαραίτητο, γιατί το μέτρο της ρευστότητας αποτελεί μία σχέση και δεν αποτελεί μόνο μία διαφορά μεταξύ του κυκλοφορούντος ενεργητικού και των τρεχουσών υποχρεώσεων της επιχείρησης.

Εφαρμογή 5

	Επιχειρήσεις A	Επιχειρήσεις B
Κυκλοφορούν ενεργητικό μείον	1.000.000 €	900.000 €
βραχυπρόθεσμες υποχρεώσεις	600.000 €	500.000.€
	400.000 €	400.000 €
Κεφάλαιο κίνησης		
Αριθμοδείκτης γενικής ρευστότητας	1,66	1,8

Από την εφαρμογή βλέπουμε ότι το κεφάλαιο κίνησης είναι το ίδιο στις δύο επιχειρήσεις. Η επιχείρηση Β παρέχει μεγαλύτερη ασφάλεια στους πιστωτές της από την Α. Δηλαδή, μία μονάδα βραχυχρόνιων υποχρεώσεων της εταιρίας Β καλύπτεται από 1,8 μονάδες κυκλοφορούντων ενεργητικών, έναντι 1,66 των αντίστοιχων αριθμοδεικτών της εταιρίας Β.

2. Αριθμοδείκτης άμεσης ρευστότητας (quick ration)

Ο δείκτης άμεσης ρευστότητας υπολογίζεται με το λόγο του κυκλοφορούντος ενεργητικού, πλην των αποθεμάτων προς τις βραχυπρόθεσμες υποχρεώσεις.

Η τιμή του δείκτη άμεσης ρευστότητας θεωρείται ικανοποιητική όταν είναι τουλάχιστον ίση με τη μονάδα.

$$\text{Αριθμοδείκτης άμεσης ρευστότητας} = \frac{\text{Κυκλοφορούν Ενεργητικό – Απόθεμα}}{\text{Βραχυπρόθεσμες υποχρεώσεις}}$$

3. Αριθμοδείκτης ταμειακής ρευστότητας

Είναι ο αριθμοδείκτης που μας δίνει την εικόνα της επάρκειας και όχι των μετρητών στην επιχείρηση, σε σχέση με τις τρέχουσες λειτουργικές ανάγκες.

$$\text{Αριθμοδείκτης ταμειακής ρευστότητας} = \frac{\text{Διαθέσιμο Ενεργητικό}}{\text{Ληξιπρόθεσμες υποχρεώσεις}}$$

Διαθέσιμο ενεργητικό: μετρητά, στο ταμείο οι καταθέσεις όψεως, επιταγές κ.λ.π.

Ο αριθμοδείκτης αυτός μας ενημερώνει για το πόσες φορές τα διαθέσιμα περιουσιακά στοιχεία μίας επιχείρησης καλύπτουν τις ληξιπρόθεσμες υποχρεώσεις της.

Σε γενικές γραμμές, η συνεχής πτώση των δεικτών ρευστότητας και η διαμόρφωσή τους σε επίπεδο μικρότερο του δείκτη του σχετικού κλάδου, αποτελεί ένα σημάδι κινδύνου για την επιχείρηση.

Παράλληλα όμως, υπάρχει και η άλλη εκδοχή:

- Μήπως αυτή η εξέλιξη σημαίνει ότι η εταιρία βελτίωσε τη διαδικασία μετατροπής των αποθεμάτων της σε πωλήσεις και ρευστό;
- Ή μπορεί να υποτεθεί ότι η εταιρία αναγκάζεται να ρευστοποιήσει τα αποθέματά της για την εξόφληση υποχρεώσεων. Ποια η ποιότητα και εμπορευσιμότητα των αποθεμάτων της;

Για να απαντηθούν τα πιο πάνω ερωτήματα, οι δείκτες ρευστότητας αξιολογούνται σε συνδυασμό και με τους δείκτες δραστηριότητας.

5.5. ΑΡΙΘΜΟΔΕΙΚΤΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ (ΚΥΚΛΟΦΟΡΙΑΚΗΣ ΤΑΧΥΤΗΤΑΣ)

Η χρησιμοποίηση ορισμένων αριθμοδεικτών κυκλοφοριακής ταχύτητας βοηθάει να προσδιορίσουμε το βαθμό μετατροπής ορισμένων περιουσιακών στοιχείων (αποθέματα, απαιτήσεις) σε ρευστό.

Γενικά, το ποσοστό των αποθεμάτων που διατηρεί μία επιχείρηση πρέπει να σχετίζεται πάντοτε με το ύψος πωλήσεών της. Μερικοί από τους πιο σημαντικούς αριθμοδείκτες κυκλοφοριακής ταχύτητας, είναι:

1. Αριθμοδείκτης ταχύτητας είσπραξης

Ο αριθμοδείκτης ταχύτητας είσπραξης απαιτήσεων δείχνει πόσες φορές, κατά μέσο όρο, εισπράττονται κατά τη διάρκεια της λογιστικής χρήσεως.

Οι απαιτήσεις της επιχείρησης:

$$\text{Αριθμοδείκτης ταχύτητας εισπράξεων απαιτήσεων} = \frac{\text{Καθαρές πωλήσεις}}{\text{Μέσος όρος απαιτήσεων}}$$

2. Αριθμοδείκτης ταχύτητας κυκλοφορίας αποθεμάτων

Ο δείκτης ταχύτητας κυκλοφορίας αποθεμάτων ορίζεται ως:

$$\text{Αριθμοδείκτης ταχύτητας κυκλοφορίας αποθεμάτων} = \frac{\text{Καθαρές πωλήσεις}}{\text{Μέσο απόθεμα προϊόντων}}$$

Αν διαιρέσουμε τον αριθμό των ημερών του έτους (365) με τον αριθμοδείκτη ταχύτητας κυκλοφορίας αποθεμάτων, τότε θα έχουμε σε αριθμό ημερών το χρόνο που παρέμειναν τα αποθέματα στην επιχείρηση μέχρι να πωληθούν.

Ο αριθμοδείκτης αυτός δείχνει την ικανότητα της επιχείρησης να μεγιστοποιεί την παραγωγική «διαδικασία» – «κύκλωμα» = πρώτες ύλες
→ εμπόρευμα → αποθήκη → πώληση.

Όσο μεγαλύτερος είναι ο αριθμοδείκτης ταχύτητας κυκλοφορίας αποθεμάτων, τόσο πιο αποτελεσματικά λειτουργεί η επιχείρηση.

3. <u>Αριθμοδείκτης ταχύτητας κυκλοφορίας κεφαλαίου κίνησης</u>

Ο δείκτης ταχύτητας κυκλοφορίας κεφαλαίου κίνησης ορίζεται ως :

$$\text{Δείκτης ταχύτητας κυκλοφορίας κεφαλαίου κίνησης} = \frac{\text{Καθαρές πωλήσεις}}{\text{Καθαρό κεφάλαιο κίνησης}}$$

Καθαρό κεφάλαιο κίνησης = σύνολο κυκλοφορίας ενεργητικού – βραχυπρόθεσμες υποχρεώσεις

Όσο αυξάνουν οι πωλήσεις μίας εταιρίας, τόσο περισσότερα κεφάλαια κινήσεως απαιτούνται για αποθέματα και για αυξημένες πιστώσεις προς τους πελάτες της.

Με την χρήση του αριθμοδείκτη αυτού μπορούμε να καταλάβουμε ποιο είναι το ύψος των πωλήσεων που επιτεύχθηκε από κάθε μονάδα καθαρού κεφαλαίου κίνησης και αν η εταιρία διατηρεί μεγάλα κεφάλαια κίνησης σε σχέση με τις πωλήσεις της.

Αριθμοδείκτες αποδοτικότητας

Οι δείκτες αυτοί χρησιμοποιούνται για την αξιολόγηση της ικανότητας της επιχείρησης να πραγματοποιεί κέρδη. Δηλαδή, οι δείκτες αυτοί αξιολογούν το βαθμό κερδοφορίας της επιχείρησης.

Βασικοί αριθμοδείκτες αποδοτικότητας είναι:

1. Αριθμοδείκτης καθαρού περιθωρίου ή καθαρού κέρδους

Ο αριθμοδείκτης αυτός, δείχνει το ποσοστό του καθαρού κέρδους που επιτυγχάνει μία επιχείρηση από τις πωλήσεις της.

Ο αριθμοδείκτης ορίζεται ως:

$$\text{Αριθμοδείκτης καθαρού κέρδους} = \frac{\text{Καθαρά κέρδη χρήσης}}{\text{Καθαρές πωλήσεις100}} \cdot 100$$

Όσο ο αριθμοδείκτης καθαρού κέρδους είναι μεγαλύτερος, τόσο πιο επικερδής είναι η επιχείρηση.

2. Αριθμοδείκτης αποδοτικότητας επένδυσης

Ο αριθμοδείκτης αυτός, μετράει την απόδοση των συνολικών περιουσιακών στοιχείων μίας επιχείρησης και αποτελεί είδος ελέγχου της διοικήσεώς της.

Ο αριθμοδείκτης ορίζεται ως:

$$\text{Αριθμοδείκτης αποδοτικότητας επένδυσης} = \frac{\text{Καθαρά κέρδη χρήσης}}{\text{Σύνολο ενεργητικού}} \cdot 100$$

Η αποτελεσματικότητα λειτουργίας μίας επιχείρησης, που δείχνει ο αριθμοδείκτης αυτός, μας δείχνει την ικανότητά της να μπορεί να επιβιώσει οικονομικά και να προσελκύει κεφάλαια που προσφέρονται για επένδυση, δίνοντας την κατάλληλη ανταμοιβή.

3. Αριθμοδείκτης αποδοτικότητας ίδιων κεφαλαίων

Ο αριθμοδείκτης αυτός απεικονίζει την κερδοφόρα δυναμικότητα της επιχείρησης και μας παρέχει ένδειξη του κατά πόσο πραγματοποιήθηκε ο στόχος πραγματοποίησης ενός ικανοποιητικού αποτελέσματος.

Ο αριθμοδείκτης ορίζεται ως εξής:

$$\text{Αριθμοδείκτης αποδοτικότητας ιδίων κεφαλαίων} = \frac{\text{Καθαρά κέρδη χρήσης}}{\text{Σύνολο ιδίων κεφαλαίων}} \cdot 100$$

ΙΣΟΛΟΓΙΣΜΟΣ ΕΤΑΙΡΙΑΣ
XXXXX A.E.

ΕΝΕΡΓΗΤΙΚΟ	2005
Διαθέσιμα (ΜΕΤΡΗΤΑ)	8.125
Λογαριασμοί εισπρακτέοι	87.600
Αποθέματα	160.875
Σύνολο κυκλοφορούντος ενεργητικού	256.600
Οικόπεδα	40.100
Μηχανήματα	35.900
Λοιπά πάγια	2.500
Άθροισμα	78.500
Σύνολο πάγιου ενεργητικού	335.100

ΠΑΘΗΤΙΚΟ	
Γραμμάτια πληρωτέα στην Τράπεζα	30.850
Γραμμάτια πληρωτέα στους προμηθευτές	12.150
Λογαριασμοί πληρωτέοι	33.000
Δεδουλευμένες υποχρεώσεις	15.500
ΣΥΝΟΛΟ ΒΡΑΧΥΠΡΟΘΕΣΜΩΝ ΥΠΟΧΡΕΩΣΕΩΝ	**91.500**
ΣΥΝΟΛΟ ΜΑΚΡΟΠΡΟΘΕΣΜΩΝ ΥΠΟΧΡΕΩΣΕΩΝ	**14.500**
ΜΕΤΟΧΙΚΟ ΚΕΦΑΛΑΙΟ	120.700)+
Αποθεματικά	108.400
ΣΥΝΟΛΟ ΙΔΙΩΝ ΚΕΦΑΛΑΙΩΝ (ΚΑΘΑΡΗ ΘΕΣΗ)	229.100
ΣΥΝΟΛΟ ΠΑΘΗΤΙΚΟΥ	**335.100**

ΛΟΓΑΡΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΧΡΗΣΗΣ ΤΗΣ ΕΤΑΙΡΙΑΣ ΧΧΧ

	(σε χιλ. ευρώ) 2005
Έσοδα πωλήσεων (καθαρά)	960.825
Κόστος πωληθέντων προϊόντων	795.725
Μικτό κέρδος	165.100
Έξοδα Διοίκησης / Διάθεσης	71.025
Κέρδη Εκμετάλλευσης	94.075
Χρηματοοικονομικά έξοδα	50.025
Κέρδη χρήσης προ φόρων	44.050
Φόροι πληρωτέοι	10.150
	33.900

Να υπολογίσετε τους χρηματοοικονομικούς δείκτες αποδοτικότητας και ρευστότητας της επιχείρησης για το έτος 2005.

Α. Αριθμοδείκτες αποδοτικότητας

i.

$$\text{Αριθμοδείκτης καθαρού κέρδους} = \frac{\text{Καθαρά κέρδη χρήσης}}{\text{Καθαρές πωλήσεις}} \cdot 100 = \frac{33.900}{960.825} \cdot 100 = 3,52\%$$

ii.

$$\text{Αριθμοδείκτης αποδοτικότητας επένδυσης} = \frac{\text{Καθαρά κέρδη χρήσης}}{\text{Σύνολο ενεργητικού}} \cdot 100 = \frac{33.900}{335.100} \cdot 100 = 10,11\%$$

iii.

$$\text{Αριθμοδείκτης αποδοτικότητας ιδίων κεφαλαίων} = \frac{\text{Καθαρά κέρδη χρήσης}}{\text{Σύνολο ιδίων κεφαλαίων}} \cdot 100 = \frac{33.900}{229.100} \cdot 100 = 14,79\%$$

Β. Αριθμοδείκτες ρευστότητας

i.

$$\text{Αριθμοδείκτης γενικής ρευστότητας} = \frac{\text{Κυκλοφορούν ενεργητικό}}{\text{Βραχυπρόθεσμες υποχρεώσεις}} = \frac{256.600}{91.500} = 2,8$$

ii.

$$\text{Αριθμοδείκτης άμεσης ρευστότητας} = \frac{\text{Κυκλοφορούν ενεργητικό -Απόθεμα}}{\text{Βραχυπρόθεσμες υποχρεώσεις}} = \frac{256.600 - 108.400}{91.500} = 1,6\%$$

Εφαρμογή 6

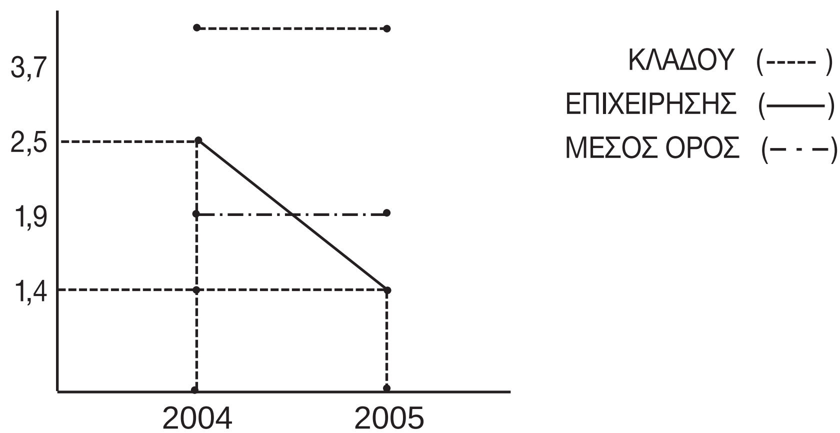
Στον παρακάτω πίνακα έχουν παρασταθεί οι χρηματοοικονομικοί δείκτες αποδοτικότητας της επιχείρησης ΧΑΑ για τα έτη 2004-2005, καθώς και οι αντίστοιχοι των επιχειρήσεων του κλάδου μ.ο.

i. Να απεικονίσετε διαγραμματικά τις τάσεις των αριθμοδεικτών του κλάδου και της επιχείρησης (ίδιο σχήμα)

ii. Να σχολιάσετε τα αποτελέσματα.

ΔΕΙΚΤΕΣ	2004	2005	Μ.Ο	ΔΕΙΚΤΕΣ ΚΛΑΔΟΥ Μ.Ο
Αριθμ. Καθαρού κέρδους	2,5	1,4	1,9%	3,7%
Αριθμ. Αποδοτικότητας και επένδυσης	6,8	2,6	4,7%	10%
Αριθμ. Αποδοτικότητας ιδίων κεφαλαίων	9,5	5,5	7,5%	15%

Α.



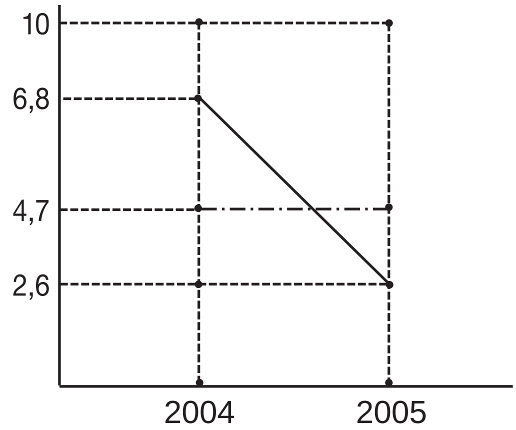
Παρατηρούμε ότι ο αριθμοδείκτης καθαρού κέρδους είναι χαμηλότερος από τον αντίστοιχο μέσο δείκτη του κλάδου.

Το χαμηλό αυτό ποσοστό μπορεί να σημαίνει ότι:

i. Οι τιμές των προϊόντων είναι χαμηλές.

ii. Το κόστος παραγωγής είναι υψηλό.

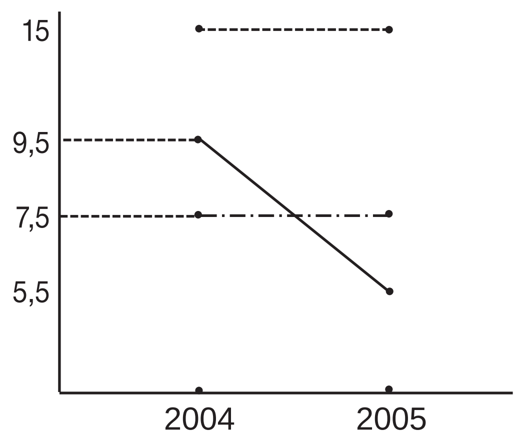
Β.



Παρατηρούμε ότι η αποδοτικότητα των συνολικών επενδυμένων κεφαλαίων, καθώς και η αποδοτικότητα των ίδιων κεφαλαίων είναι σημαντικά χαμηλότερες σε σύγκριση με τις αντίστοιχες μέσες αποδοτικότητες του κλάδου.

Και ο τρίτος αριθμοδείκτης αποδοτικότητας της επιχείρησης είναι σημαντικά κατώτερος σε σχέση με το μέσο όρο του κλάδου των επιχειρήσεων.

Γ.



1. Απαντήστε με (Σ) Σωστό ή (Λ) Λάθος, στις επόμενες ερωτήσεις:

	Σωστό	Λάθος
i. Αν συγκρίνουμε το κόστος διατροφής στην Ελλάδα με αυτόν της Γαλλίας το 2005 έχουμε ένα γεωγραφικό δείκτη.	☐	☐
ii. Οι ιδιαίτεροι αριθμοδείκτες μας δείχνουν τις μεταβολές μεταβλητών μεταξύ δύο χρονικών περιόδων.	☐	☐
iii. Το $P_{1/0} = \frac{P_1}{P_0}100$.	☐	☐
iv. Η σχετική αξία ορίζεται ως $U_{1/0} = \frac{q_{1/0}}{p_{1/0}}$.	☐	☐
v. Οι σταθμικοί συνθετικοί δείκτες ανήκουν στους ιδιαίτερους αριθμοδείκτες.	☐	☐
vi. Ο αριθμοδείκτης γενικής ρευστότητας ορίζεται ως: $\frac{\text{Βραχυπρόθεσμες υποχρεώσεις}}{\text{Κυκλοφορούν ενεργητικό}}$	☐	☐
vii. Μεταξύ δύο εταιριών, εκείνη που έχει το μεγαλύτερο κεφάλαιο κίνησης είναι αυτή που έχει τη μεγαλύτερη ρευστότητα.	☐	☐
viii. Ο αριθμοδείκτης ταχύτητας είσπραξης ανήκει στους αριθμοδείκτες αποδοτικότητας.	☐	☐

2. Η τιμή πώλησης μίας τυρόπιτας το 2002 ήταν 1,4€ και το 2005 ήταν 1,7€. Αν πάρουμε το 2002 ως έτος βάσης, να υπολογίσετε τη σχετική τιμή.

3. Δίνονται οι τιμές σε € και οι ποσότητες που πουλήθηκαν ενός αγαθού για τις χρονιές 2005-2006.

ΕΤΟΣ	ΤΙΜΗ	ΠΟΣΟΤΗΤΑ
2005	7	180
2006	8,5	200

Να υπολογίσετε το σχετικό δείκτη αξίας για το έτος 2006.

4. Στον παρακάτω πίνακα δίνονται οι τιμές σε € και οι ποσότητες που πουλήθηκαν. Έχει χαθεί η ποσότητα που πουλήθηκε το 2001. Αν ο σχετικός δείκτης αξίας είναι 70%, να βρεθεί η ποσότητα.

ΕΤΟΣ	ΤΙΜΗ	ΠΟΣΟΤΗΤΑ
2001	7	?
2006	8,5	100

5. Στον παρακάτω πίνακα δίνονται σε € οι τιμές τριών προϊόντων για τα έτη 2003 και 2004.
   i. Να υπολογίσετε τον δείκτη συνολικών τιμών.
   ii. Μέσο αριθμητικό των σχετικών τιμών.

ΑΓΑΘΑ	ΤΙΜΕΣ		ΣΧΕΤΙΚΕΣ ΤΙΜΕΣ
	2003	2004	
	p_0	p_1	p_1/p_0
Α	3	12	
Β	7	17,5	
Γ	2	3	3
ΣΥΝΟΛΟ			

6. Η εταιρία ΧΧΧ παράγει τρία προϊόντα Α,Β,Γ. Οι τιμές και οι ποσότητες που παρήχθησαν, φαίνονται στον παρακάτω πίνακα. Αν θεωρηθεί ως έτος βάσης το 2004, να βρεθούν:

i. Ο δείκτης τιμών του Laspeyres.

ii. Ο δείκτης τιμών του Paasche.

ΠΡΟΪΟΝΤΑ	2004		2005	
	p_0	q_0	p_1	q_1
Α	10	200	11	150
Β	15	150	17	180
Γ	7	110	10	140

7. Η εταιρία Α έχει κυκλοφορούν ενεργητικό 1.500.000 € και βραχυπρόθεσμες υποχρεώσεις 1.100.000€, ενώ η εταιρία Β έχει κυκλοφορούν ενεργητικό 500.000€ και βραχυπρόθεσμες υποχρεώσεις 35.000€. Ποια βρίσκεται σε καλύτερη θέση για να ανταποκρίνεται στις τρέχουσες δραστηριότητές της;

8. Δίνονται οι ισολογισμοί της εταιρίας «ALPHA» (σε χιλιάδες ευρώ) για το έτος 2005, καθώς και ο λογαριασμός αποτελεσμάτων χρήσης.

ΙΣΟΛΟΓΙΣΜΟΣ

ΕΝΕΡΓΗΤΙΚΟ

Διαθέσιμα (ΜΕΤΡΗΤΑ)	86.700	
Λογαριασμοί εισπρακτέοι	8.925	+
Αποθέματα	159.375	
Σύνολο κυκλοφορούντος ενεργητικού	**255.000**	
Οικόπεδα / κτίρια	36.975	+
Μηχανήματα	40.800	
Λοιπά πάγια στοιχεία	5.500	
	83.325	
Σύνολο πάγιου ενεργητικού	....	

ΠΑΘΗΤΙΚΟ

Γραμμάτια πληρωτέα στην Τράπεζα	12.112	
Γραμμάτια πληρωτέα στους προμηθευτές	36.338	
Λογαριασμοί πληρωτέοι	31.875	+
Δεδουλευμένες υποχρεώσεις	20.850	
ΣΥΝΟΛΟ ΒΡΑΧΥΠΡΟΘΕΣΜΩΝ ΥΠΟΧΡΕΩΣΕΩΝ	**101.175**	
ΣΥΝΟΛΟ ΜΑΚΡΟΠΡΟΘΕΣΜΩΝ ΥΠΟΧΡΕΩΣΕΩΝ	12.750	
ΜΕΤΟΧΙΚΟ ΚΕΦΑΛΑΙΟ	115.750	
ΑΠΟΘΕΜΑΤΙΚΑ	....	
ΣΥΝΟΛΟ ΙΔΙΩΝ ΚΕΦΑΛΑΙΩΝ (ΚΑΘΑΡΗ ΘΕΣΗ)	224.750	
ΣΥΝΟΛΟ ΠΑΘΗΤΙΚΟΥ	....	

ΛΟΓΑΡΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΧΡΗΣΗΣ

Έσοδα πωλήσεων (καθαρά)	860.425	−
Κόστος πωληθέντων προϊόντων	695.925	
Μικτό κέρδος	**164.500**	
Έξοδα Διοίκησης	71.025	
Κέρδη Εκμετάλλευσης	....	
Χρηματοοικονομικά έξοδα	55.875	
Κέρδη χρήσης προ φόρων	38.000	−
Φόροι πληρωτέοι	12.500	
	25.500	

i. Να συμπληρωθούν τα κενά που λόγω βλάβης του εκτυπωτή δεν τυπώθηκαν.

ii. Να προσδιορίσετε τους αριθμοδείκτες ρευστότητας, αποδοτικότητας, καθώς και τους δείκτες ταχύτητας κυκλοφορίας κεφαλαίου.

ΔΡΑΣΤΗΡΙΟΤΗΤΑ

Κατασκευάστε ή βρείτε έναν δημοσιευμένο ισολογισμό και συγκρίνετε την πορεία της εταιρίας με αυτήν της άσκησης 8.

Σύνοψη

Οι αριθμοδείκτες είναι κάποια στατιστικά μέτρα που μας δείχνουν τις μεταβολές που υφίσταται μια μεταβλητή ή ομάδα μεταβλητων που σχετίζονται μεταξύ τους μεταξύ δύο χρονικών περιόδων.
Οι αριθμοδείκτες αποτελούν μια δυναμική μέθοδο χρηματοποκονομικής ανάλυσης. Με τους αριθμοδείκτες γίνεται άμεσα αντιληπτή η πραγματική αξία και η σπυδαιότητα τω απόλυτων μεγεθών.

- **Αριθμοδείκτες**
  - Χρονολογικοί
    - Ιδιαίτεροι
    - Σύνθετοι
  - Χρηματοοικονομική ανάλυση
    - Ρευστότητας
    - Δραστηριότητας
    - Αποδοτικότητας
    - Διαρθρώσεις κεφαλαίων
    - Επενδύσων

Βιβλιογραφία / Internet

«*Στατιστική*», Κιόχος Π., Interbooks
«*Περιγραφική Στατιστική*», Βασ. Μπένος
«Αριθμοδείκτες», Τζωρτζόπουλος Π., 1991
«Index Numbers», Mudgelt B.D.
www.forum-training.gr/ari8modeiktes.html: *στο site αυτό βρίσκεται τυπολόγιο με το σύνολο των αριθμών*
http://homepages.pathfinder.gr/ageorg/128.htm: *αριθμοδείκτες και τρόπος υπολογισμού τους.*

ΟΔΗΓΟΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΜΕΛΕΤΗ

«*Στατιστική*», Εκδόσεις Μακεδονικές, Όθωνα Παπαδήμα
Στο βιβλίο αυτό θα υπάρχουν θέματα που αφορούν βασικές γνώσεις Στατιστικής, όπως συλλογή και παρουσίαση στατιστικών στοιχείων, θέματα στα μέτρα θέσης και διασποράς, στους αριθμοδείκτες και την παλινδρόμηση.

6. ΣΥΝΔΥΑΣΤΙΚΗ

Βασικές έννοιες:
- βασική αρχή απαρίθμησης
- διατάξεις
- μεταθέσεις
- συνδυασμοί

Στόχος του μαθήματος: ο εκπαιδευόμενος να κατανοήσει ότι η συνδυαστική αποτελεί μέρος της καθημερινότητάς μας, των επιλογών και ότι είναι άμεση συνδεδεμένη ως έννοια με την έννοια της πιθανότητας.

Προσδοκώμενα αποτελέσματα: ο εκπαιδευόμενος να εκοικειωθεί με την έννοια του συνόλου και την απαρίθμηση αυτού. Σαν αποτέλεσμα θα μπορεί να χειρίζεται τις έννοιες της συνδυαστικής και στην συνέχεια να εισαχθεί στην έννοια της πιθανότητας.

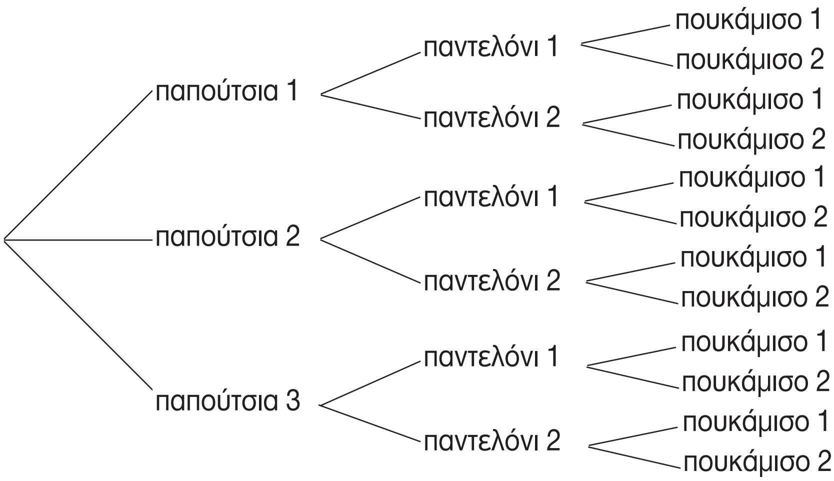
Εισαγωγικές παρατηρήσεις: Η συνδυαστική είναι ένας κλάδος των Μαθηματικών με άμεση σχέση με καθημερινά προβλήματα. Εξάλλου οι συνδυασμοί, οι διατάξεις και οι μεταθέσεις είναι βασικές έννοιες της που χρησιμοποιούνται ευρύτατα και στην καθημερινή μας ομιλία. Επίσης η συνδυαστική σχετίζεται με την απαρίθμηση συνόλων ή υποσυνόλων ευρύτερων συνόλων, διαδικασία που μας επιτρέπει να υπολογίσουμε την πιθανότητα να συμβούν συγκεκριμένα ενδεχόμενα, που μας ενδιαφέρουν.

6.1. ΤΙ ΕΙΝΑΙ ΣΥΝΔΥΑΣΤΙΚΗ - ΒΑΣΙΚΗ ΑΡΧΗ ΤΗΣ ΑΠΑΡΙΘΜΗΣΗΣ

Ο προπονητής ποδοσφαιρικής ομάδας του Ανίκητου, έχει στη διάθεσή του 3 τερματοφύλακες, 6 αμυντικούς, 5 μέσους και 3 επιθετικούς. Δεδομένου ότι η ομάδα παίζει με το σύστημα 4,3,3, με πόσους μπορεί να διαμορφώσει την ενδεκάδα;

Πολλές φορές μας δίνουν ένα σύνολο αντικειμένων και μας ζητούν να σχηματίσουμε με αυτά ομάδες, που να διαφέρουν μεταξύ τους, είτε κατά τη φύση του αντικειμένου που περιέχουν, είτε κατά τη θέση των αντικειμένων μέσα στην ομάδα.. Συνδυαστική είναι ο κλάδος των μαθηματικών που μας βοηθάει να υπολογίσουμε με πόσους τρόπους μπορούμε να φτιάξουμε τέτοιες ομάδες. Στο τέλος του κεφαλαίου θα είμαστε σε θέση να απαντήσουμε στο πρόβλημα της διαμόρφωσης της ενδεκάδας.

Ας δούμε ένα πιο απλό πρόβλημα. Κάποιος έχει στην ντουλάπα του 3 ζευγάρια παπούτσια, 2 παντελόνια και 2 πουκάμισα. Με πόσους τρόπους μπορεί να ντυθεί, δηλαδή πόσες διαφορετικές εμφανίσεις διαθέτει; Μπορεί να διαλέξει παπούτσια με 3 τρόπους και για κάθε επιλογή παπουτσιών μπορεί στη συνέχεια να διαλέξει παντελόνι με 2 τρόπους. Έχει λοιπόν $3 \cdot 2=6$ τρόπους επιλογής παπουτσιών - παντελονιού. Για κάθε έναν από αυτούς μπορεί να επιλέγει πουκάμισο με δύο τρόπους, δηλαδή συνολικά έχει $3 \cdot 2 \cdot 2=12$ τρόπους να ντυθεί. Οι τρόποι αυτοί φαίνονται στο παρακάτω διάγραμμα, που λόγω της εμφάνισής του, το ονομάζουμε «δενδροδιάγραμμα».



Γενικά, μπορούμε να διατυπώσουμε την επόμενη **βασική αρχή της απαρίθμησης**.

> Αν μία διαδικασία μπορεί να πραγματοποιηθεί σε ν βήματα και το πρώτο βήμα μπορεί να πραγματοποιηθεί με $κ_1$ τρόπους, το δεύτερο βήμα με $κ_2$ τρόπους, τότε η διαδικασία αυτή μπορεί να πραγματοποιηθεί με $κ_1 \cdot κ_2 \cdot \ldots \cdot κ_ν$ τρόπους.

Γι' αυτό, στο προηγούμενο παράδειγμα που η επιλογή παπουτσιών γίνεται με 3 τρόπους, η επιλογή παντελονιού γίνεται με 2 τρόπους και η επιλογή πουκαμίσου γίνεται με 2 τρόπους, ο άνθρωπος έχει συνολικά $3 \cdot 2 \cdot 2=12$ τρόπους να ντυθεί.

Στα επόμενα, θα συναντήσουμε συχνά γινόμενα της μορφής $1 \cdot 2 \cdot 3 \cdot \ldots \cdot ν$ που συμφωνούμε να συμβολίζουμε με $ν!$ και να διαβάζουμε «ν παραγωγικό».

Για παράδειγμα, $1 \cdot 2 \cdot 3 \cdot 4 \cdot 5=120$. Συμφωνούμε επίσης ότι $1!=1$ και $0!=1$. Το τελευταίο φαίνεται λίγο περίεργο, αλλά είναι ο μόνος τρόπος ώστε οι τύποι που θα μάθουμε παρακάτω να ισχύουν σε κάθε περίπτωση.

Αν και όλα σχεδόν τα προβλήματα συνδυαστικής μπορούν να λυθούν με την αρχή της απαρίθμησης, θα μάθουμε αργότερα πολλούς τύπους που μας διευκολύνουν.

ΕΞΑΣΚΗΣΗ

1. Κάποιος μπορεί να ταξιδέψει από τη Θεσσαλονίκη στην Αθήνα με τρένο, λεωφορείο ή αεροπλάνο και από την Αθήνα στο Ηράκλειο με πλοίο ή αεροπλάνο. Με πόσους τρόπους μπορεί να ταξιδέψει από τη Θεσσαλονίκη στο Ηράκλειο;
2. Στο παιχνίδι του ΠΡΟΠΟ, μπορούμε να προβλέψουμε τον νικητή καθενός από τους αγώνες με 3 τρόπους (1,Χ ή 2). Με πόσους τρόπους μπορούμε να συμπληρώσουμε τις 13 (ή 14 για το σούπερ δεκατριάρι) θέσεις μίας στήλης;

3. Ποιο είναι το πλήθος των πενταψήφιων αριθμών που διαβάζονται ίδιοι από δεξιά και από αριστερά, π.χ. 12821;
4. Έχουμε στη διάθεσή μας 3 χρώματα και θέλουμε να χρωματίσουμε μία ταινία με 5 διαδοχικά τετραγωνάκια, ώστε δύο γειτονικά τετράγωνα να μην έχουν το ίδιο χρώμα. Με πόσους τρόπους μπορούμε να το κάνουμε αυτό;
5. Οι πινακίδες κυκλοφορίας των ελληνικών αυτοκινήτων αποτελούνται από 3 γράμματα, που είναι κοινά στο ελληνικό και στο λατινικό αλφάβητο και ένα τετραψήφιο αριθμό που δεν μπορεί βέβαια να αρχίζει από 0. Πόσες τέτοιες πινακίδες μπορούμε να φτιάξουμε;

6.2. ΔΙΑΤΑΞΕΙΣ

Το Διοικητικό Συμβούλιο ενός συλλόγου αποτελείται από 7 άτομα. Συνεδριάζουν για να εκλέξουν πρόεδρο, γραμματέα και ταμία. Με πόσους τρόπους μπορεί να γίνει η εκλογή; Μπορούν να εκλέξουν πρόεδρο με 7 τρόπους (όσα είναι τα μέλη του ΔΣ) γραμματέα με 6 τρόπους (όσα είναι τα μέλη που απομένουν μετά την εκλογή προέδρου) και ταμία με 5 τρόπους (όσα είναι τα μέλη που μένουν μετά την εκλογή των δύο πρώτων). Σύμφωνα με την αρχή της απαρίθμησης, η εκλογή μπορεί να γίνει με $7 \cdot 6 \cdot 5=210$ τρόπους. Κάθε ένας από αυτούς τους τρόπους λέγεται **διάταξη χωρίς επανάληψη** των 7 ανά 3.

Διάταξη χωρίς επανάληψη των ν στοιχείων ενός συνόλου ανά κ (με $\kappa \leq \nu$) λέγεται καθένας από τους διαφορετικούς τρόπους με τους οποίους μπορούμε να πάρουμε κ διαφορετικά στοιχεία από τα ν και να τα βάλουμε σε μία σειρά.

Το πλήθος των διατάξεων χωρίς επανάληψη συμβολίζεται με $(\nu)_\kappa$ και αν εργαστούμε όπως στο προηγούμενο παράδειγμα, βρίσκουμε ότι είναι ίσο με $\nu(\nu-1)(\nu-2)\ldots(\nu-\kappa+1)$.

Το γινόμενο αυτό γράφεται και:

$$\nu(\nu-1)\ldots(\nu-\kappa+1)=\frac{\nu(\nu-1)\ldots(\nu-\kappa+1)(\nu-\kappa)\ldots 3\cdot 2\cdot 1}{(\nu-\kappa)\ldots 3\cdot 2\cdot 1}=\frac{\nu!}{(\nu-\kappa)}$$

Έτσι τελικά έχουμε $\boxed{(\nu)_\kappa=\frac{\nu!}{(\nu-\kappa)!}(\mu\varepsilon\ \kappa\leq\ \nu)}$.

Εφαρμογή 1

Πόσα ζευγάρια με διαφορετικά στοιχεία μπορούμε να φτιάξουμε με τα στοιχεία του συνόλου Α={α,β,γ,δ}; Ποια θα είναι αυτά τα ζευγάρια;

Το πλήθος των ζευγαριών θα είναι $(4)_2=4\cdot 3=12$.

Τα ζευγάρια αυτά θα είναι (α,β), (α,γ), (α,δ), (β,α), (β,γ), (β,δ), (γ,α), (γ,β), (γ,δ), (δ,α), (δ,β), (δ,γ).

Εφαρμογή 2

Σε μία πόλη που έχει 8 ξενοδοχεία, φθάνουν 5 τουρίστες και διαλέγουμε από ένα διαφορετικό ξενοδοχείο ο καθένας. Με πόσους τρόπους μπορούν να μοιραστούν οι τουρίστες στα ξενοδοχεία;

Πρόκειται για διατάξεις χωρίς επανάληψη των 8 στοιχείων ανά 5, δηλαδή $(8)_5=8\cdot7\cdot6\cdot5\cdot4=6.720$.

Εφαρμογή 3

Οι διατάξεις 5 στοιχείων ανά κ είναι 20. Να βρείτε το κ.

$$(5)_\kappa = 20 \Leftrightarrow \frac{5!}{(5-\kappa)!} = 20 \Leftrightarrow \frac{120}{(5-\kappa)!} = 20 \Leftrightarrow (5-\kappa)! = \frac{120}{20} = 6 \Leftrightarrow$$
$$(5-\kappa)! = 3! \Leftrightarrow 5-\kappa = 3 \Leftrightarrow \kappa = 2$$

Ας υποθέσουμε τώρα ότι έχουμε μία κάλπη με 7 σφαιρίδια αριθμημένα από 1 μέχρι 7. Επιλέγουμε τυχαία ένα σφαιρίδιο, σημειώνουμε το νούμερο και ξαναβάζουμε το σφαιρίδιο στην κάλπη. Η διαδικασία επαναλαμβάνεται άλλες δύο φορές, ώστε να σχηματίσουμε ένα τριψήφιο αριθμό. Πόσους τέτοιους τριψήφιους αριθμούς μπορούμε να σχηματίσουμε; Σύμφωνα με την αρχή της απαρίθμησης και εφόσον κάθε βήμα της διαδικασίας γίνεται με 7 τρόπους, συνολικά έχουμε $7^3=343$ τέτοιους τριψήφιους αριθμούς. Κάθε ένας από αυτούς τους τρόπους ονομάζεται **διάταξη με επανάληψη** των 7 στοιχείων ανά 3.

> Διάταξη με επανάληψη των ν στοιχείων ενός συνόλου ανά κ λέγεται καθένας από τους διαφορετικούς τρόπους με τους οποίους μπορούμε να πάρουμε κ στοιχεία από τα ν και να τα βάλουμε σε μία σειρά .

Το πλήθος των διατάξεων με επανάληψη συμβολίζεται με $\boxed{[\nu]_\kappa=\nu^\kappa}$.

Εφαρμογή 4

Πόσα ζευγάρια μπορούμε να φτιάξουμε με τα στοιχεία του συνόλου Α={α,β,γ,δ}; Ποια θα είναι αυτά τα ζευγάρια; (Συγκρίνετε το αποτέλεσμα με την εφαρμογή 1).

Το πλήθος των ζευγαριών θα είναι: $[4]_2=4^2=16$.
Τα ζευγάρια αυτά θα είναι (α,α), (α,β), (α,γ), (α,δ), (β,α), (β,β), (β,γ), (β,δ), (γ,α), (γ,β), (γ,γ), (γ,δ), (δ,α), (δ,β), (δ,γ), (δ,δ). Αντίθετα με την εφαρμογή 1, εδώ επιτρέπονται ζευγάρια με ίδια στοιχεία.

Εφαρμογή 5

Σε μία πόλη που έχει 8 ξενοδοχεία, φθάνουν 5 τουρίστες και διαλέγουν από ένα ξενοδοχείο ο καθένας, αλλά αντίθετα από την εφαρμογή 2, μπορούν να μείνουν περισσότεροι από ένας σε κάθε ξενοδοχείο. Με πόσους τρόπους μπορούν να μοιραστούν οι τουρίστες στα ξενοδοχεία;

Πρόκειται για διατάξεις με επανάληψη των 8 στοιχείων ανά 5, δηλαδή $[8]_5=8^5=32.768$.

1. Αντιστοιχίστε τα στοιχεία της πρώτης στήλης με τα στοιχεία της δεύτερης.

Α	Β
	1
$(3)_2$	6
$(2)_3$	8
$[3]_2$	9
$[2]_3$	12
	Δεν ορίζεται

2. Αντιστοίχισε τα στοιχεία της πρώτης στήλης με τα στοιχεία της δεύτερης.

Α	Β
	0
$(\nu)_0$	1
$(\nu)_1$	ν
$(\nu)_{\nu-1}$	$\nu^2-\nu$
$(\nu)_\nu$	$(\nu-1)!$
$(\nu)_2$	$\nu!$
	Δεν ορίζεται

3. Με πόσους τρόπους μπορούν να καθίσουν 4 άτομα σε 6 καρέκλες; Αν η πρώτη καρέκλα έχει κρατηθεί για τον επίσημο προσκεκλημένο, με πόσους τρόπους μπορούν τα 4 άτομα να καθίσουν στις υπόλοιπες καρέκλες;

4. Με πόσους τρόπους μπορούμε να βάλουμε 3 πουκάμισα σε 5 συρτάρια αν:
α) σε κάθε συρτάρι μπορεί να μπει το πολύ ένα πουκάμισο;
β) σε κάθε συρτάρι μπορούν να μπουν όσα πουκάμισα θέλουμε;

5. Ένα σύστημα ΠΡΟΠΟ έχει 4 διπλές και 3 τριπλές παραλλαγές. Από πόσες στήλες αποτελείται;

6. Πόσους τετραψήφιους αριθμούς μπορούμε να φτιάξουμε με τα ψηφία 1,2,3,6,7,8,9 αν:
α) τα ψηφία πρέπει να είναι διαφορετικά μεταξύ τους;
β) ο αριθμός μπορεί να περιέχει ίδια στοιχεία;

7. Πόσους τετραψήφιους αριθμούς με διαφορετικά ψηφία μπορούμε να φτιάξουμε (εννοείται όχι 0 στην πρώτη θέση);

8. Πόσους τετραψήφιους αριθμούς μπορούμε να φτιάξουμε, ώστε τα δύο πρώτα ψηφία να είναι 1,2,3,4 ή 5 και τα δύο τελευταία 6,7,8,9 ή 0 αν:
   α) δεν επιτρέπονται ίδια ψηφία;
   β) επιτρέπονται ίδια ψηφία;

9. Σε έναν αγώνα δρόμου λαμβάνουν μέρος 9 αθλητές. Πόσες είναι οι διαφορετικές τριάδες που μπορούν να εμφανιστούν στις 3 πρώτες θέσεις;

10 Να βρείτε το ν, αν οι διατάξεις χωρίς επαναλήψεις των ν στοιχείων ανά 4 είναι 42 φορές περισσότερες από τις διατάξεις χωρίς επαναλήψεις των ν στοιχείων ανά 2.

11. Να βρείτε το ν, αν οι διατάξεις χωρίς επανάληψη των ν στοιχείων ανά 3 είναι 120.

12. Να βρείτε το κ, αν οι διατάξεις χωρίς επανάληψη των 7 στοιχείων ανά κ είναι 210.

6.3. ΜΕΤΑΘΕΣΕΙΣ

Έχουμε μία ομάδα 5 ατόμων και θέλουμε να τους παρατάξουμε σε μία σειρά. Με πόσους τρόπους μπορεί να γίνει αυτό; Είναι φανερό ότι πρόκειται για διατάξεις χωρίς επανάληψη των 5 ανά 5, συνεπώς αυτό γίνεται με $(5)_5 = \frac{5!}{(5-5)!} = \frac{5!}{0!} = 5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ τρόπους.

Κάθε ένας από αυτούς ονομάζεται και <u>μετάθεση</u> των 5 στοιχείων.

> Μετάθεση των ν στοιχείων ενός συνόλου λέγεται καθένας από τους διαφορετικούς τρόπους με τους οποίους μπορούμε να βάλουμε τα ν αυτά στοιχεία σε μία σειρά.

Το πλήθος των μεταθέσεων αυτών συμβολίζεται με και είναι ίσο με $(\nu)_\nu = \frac{\nu!}{(\nu-\nu)!} = \frac{\nu!}{0!} = \frac{\nu!}{1} = \nu!$

Έτσι καταλήγουμε στον τύπο $\boxed{P_\nu = \nu!}$.

Εφαρμογή 1

Θέλουμε να βάλουμε σε μία σειρά τον Πέτρο, το Νίκο και την Ελένη. Με πόσους τρόπους μπορούμε να το κάνουμε; Ποιοι είναι αυτοί οι τρόποι;

Πρόκειται για μεταθέσεις 3 στοιχείων, άρα υπάρχουν $P_3=3!=3\cdot2\cdot1=6$ τρόποι. Αυτοί είναι οι εξής: ΕΝΠ, ΕΠΝ, ΝΕΠ, ΝΠΕ, ΠΕΝ, ΠΝΕ (για ευκολία έχει γραφεί μόνο το αρχικό γράμμα κάθε ονόματος).

Εφαρμογή 2

Πόσους τετραψήφιους αριθμούς μπορούμε να φτιάξουμε χρησιμοποιώντας τα ψηφία 1,3,7,8 από μία φορά το καθένα; Πόσοι από τους αριθμούς αυτούς δεν αρχίζουν από 1;

Πρόκειται για μεταθέσεις 4 στοιχείων, άρα το πλήθος τους είναι: $P_4=4!=4\cdot3\cdot2\cdot1=24$. Για να βρούμε πόσοι αριθμοί δεν αρχίζουν από 1, αρκεί να βρούμε πόσοι είναι αυτοί που αρχίζουν από 1 και να τους αφαιρέσουμε. Οι αριθμοί που αρχίζουν από 1, είναι όσες και οι μεταθέσεις των άλλων τριών ψηφίων στις υπόλοιπες θέσεις, δηλαδή $P_3=3!=3\cdot2\cdot1=6$. Έτσι, οι αριθμοί που δεν αρχίζουν από 1 είναι $24-6 = 8$.

Ας υποθέσουμε τώρα ότι έχουμε 3 μπλε σφαίρες, 2 πράσινες και 1 κόκκινη και θέλουμε να τα βάλουμε σε μία σειρά, για παράδειγμα «μμπκμπ». Με πόσους τρόπους μπορεί να γίνει αυτό; Αν οι 6 σφαίρες ήταν διαφορετικές μεταξύ τους, τότε θα είχαμε $P_6=6!=6\cdot5\cdot4\cdot3\cdot2\cdot1=720$. Τώρα όμως οι 3 μπλε σφαίρες είναι ίδιες και αν τις αλλάξουμε θέσεις μεταξύ τους δεν δημιουργούμε μια διαφορετική σειρά. Οι μπλε σφαίρες μπορούν να αλλάξουν θέσεις μεταξύ τους με $P_3=3!=3\cdot2\cdot1=6$τρόπους. Όμοια οι 2 πράσινες σφαίρες αλλάζουν θέσεις μεταξύ τους με$P_2=2!=2\cdot1=2$ τρόπους. Έτσι, τελικά οι τρόποι με τους οποίους
μπορούμε να τοποθετήσουμε τις σφαίρες είναι $\dfrac{6!}{3!2!}=\dfrac{720}{6\cdot2}=60$ τρόποι.

Κάθε τέτοιος τρόπος τοποθέτησης είναι μία μετάθεση με επανάληψη.

Αν έχουμε ένα σύνολο ν στοιχείων από τα οποία α_1 είναι ίδια μεταξύ τους, α_2 επίσης είναι ίδια μεταξύ τους,..., α_κ είναι ίδια μεταξύ τους, τότε κάθε τοποθέτηση των ν στοιχείων σε μία σειρά λέγεται μετάθεση με επανάληψη.

Το πλήθος των δυνατών αυτών μεταθέσεων είναι $\boxed{\dfrac{\nu!}{\alpha_1!\alpha_2!\ldots\alpha_\kappa!}}$.

Εφαρμογή 3

Πόσες διαφορετικές λέξεις μπορούμε να φτιάξουμε με αναγραμματισμός της λέξης ΣΤΑΤΙΣΤΙΚΗ; (Δεν μας ενδιαφέρει αν οι λέξεις έχουν νόημα).

Έχουμε στη διάθεσή μας 10 στοιχεία, μέσα στα οποία εμφανίζεται 3 φορές το Τ, 2 φορές το Σ και 2 φορές το Ι. Το πλήθος των διαφορετικών λέξεων θα είναι ίσο με το πλήθος των μεταθέσεων με επανάληψη, δηλαδή:

$$\frac{10!}{3!2!2!}=\frac{10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2}{3\cdot2\cdot2\cdot2}=151.200$$

Σημείωση: Όταν οι αριθμοί που εμφανίζονται στις πράξεις είναι μεγάλοι, πρέπει πρώτα να κάνουμε όλες τις δυνατές απλοποιήσεις. Διαφορετικά, οι ενδιάμεσοι αριθμοί που προκύπτουν είναι τόσο μεγάλοι, που μπορεί να μην υπολογίζονται ούτε με κομπιουτεράκι!

1. Επιλέξτε τη σωστή απάντηση:
   α. Πόσες είναι μεταθέσεις 4 στοιχείων;
      Α. 4 Β. 6 Γ.12 Δ. 24 Ε. 120
   β. Πόσες είναι οι μεταθέσεις 4 στοιχείων από τα οποία δύο είναι ίδια μεταξύ τους;
      Α. 4 Β. 6 Γ.12 Δ. 24 Ε. 120
   γ. Η διαφορά είναι ίση με :
      Α. 2 Β. 3 Γ. 4 Δ. 6 Ε. 12
2. Πόσες διαφορετικές λέξεις μπορούμε να φτιάξουμε με αναγραμματισμό της λέξης:
   α) ΔΟΚΙΜΗ
   β) ΜΑΘΗΜΑΤΙΚΑ
3. Πόσοι αριθμοί μεταξύ του 3.000 και του 4.000 περιέχουν τα ψηφία 2,3,4,5;
4. Έχουμε στη διάθεσή μας 3 κόκκινες και 2 πράσινες σφαίρες. Με πόσους τρόπους μπορούμε να τις τοποθετήσουμε στη σειρά;
5. Μία ομάδα μαθητών αποτελείται από 4 αγόρια και 3 κορίτσια. Με πόσους τρόπους μπορούμε να τους παρατάξουμε στη σειρά, αν:
   α. δεν μας ενδιαφέρει ο τρόπος παράταξης
   β. θέλουμε τα κορίτσια μπροστά και τα αγόρια πίσω
   γ. θέλουμε τα κορίτσια μαζί και τα αγόρια μαζί.
6. Δύο ανδρόγυνα και δύο ανύπαντροι φίλοι τους (συνολικά 6 άτομα), έχουν κλείσει τις 6 κεντρικές θέσεις μίας σειράς στο θέατρο. Με πόσους τρόπους μπορούν να καθίσουν, αν:
   α. δεν μας ενδιαφέρει η σειρά
   β. τα ανδρόγυνα πρέπει να καθίσουν ο ένας δίπλα στον άλλο.
7. Μία αυτόματη κλειδαριά έχει έναν τετραψήφιο κωδικό με διαφορετικά ψηφία. Πόσους κωδικούς πρέπει να δοκιμάσει ο ιδιοκτήτης που θυμάται τα τέσσερα ψηφία, αλλά :
   α. δεν θυμάται καθόλου τη σωστή σειρά
   β. θυμάται μόνο πιο είναι το πρώτο ψηφίο
   γ. θυμάται μόνο πιο είναι το πρώτο και πιο είναι το τελευταίο ψηφίο.
8. Δείξτε ότι:
   α) $P_\nu=(\nu)_{\nu-1}$
   β) $(10)_4=P_7$
   γ) $(6)_3=P_5$

6.4. ΣΥΝΔΥΑΣΜΟΙ

Αν θέλουμε από μία ομάδα 7 ατόμων να επιλέγουμε μία τριμελή επιτροπή, με πόσους τρόπους μπορεί να γίνει αυτό; Αν επρόκειτο για διατάξεις θα είχαμε $(7)_3=7{\cdot}6{\cdot}5=210$ τρόπους. Όμως δεν πρόκειται για διατάξεις, γιατί δεν μας ενδιαφέρει η σειρά ανάμεσα στα μέλη της επιτροπής, δηλαδή μία τριάδα «αβγ» θα είναι ίδια με την «βγα» και με όλες αυτές που προκύπτουν από τη μετάθεση των τριών στοιχείων, Αυτές οι τριάδες είναι. Έτσι, το πραγματικό πλήθος των τρόπων με τους οποίους μπορεί να γίνει η επιλογή της τριμελούς επιτροπής είναι $\frac{210}{6}=35$.

Κάθε τέτοια τριάδα που δεν έχει συγκεκριμένη σειρά, είναι ένας **συνδυασμός χωρίς επανάληψη** των 7 στοιχείων αντί 3.

> Συνδυασμός χωρίς επανάληψη των ν στοιχείων ενός συνόλου Α ανά κ $(\mu\varepsilon\ \kappa \le \nu)$ λέγεται κάθε υποσύνολο του Α με κ στοιχεία, δηλαδή κάθε ομάδα κ διαφορετικών στοιχείων του Α, χωρίς να μας ενδιαφέρει η σειρά των στοιχείων.

Το πλήθος των συνδυασμών χωρίς επανάληψη ν στοιχείων ανά κ συμβολίζεται με $\binom{\nu}{\kappa}$ και, όπως είδαμε στο προηγούμενο παράδειγμα, είναι ίσο με:

$$\binom{\nu}{\kappa}=\frac{(\nu)\kappa}{P_\kappa}=\frac{\nu(\nu-1)(\nu-2)\ldots(\nu-\kappa+1)}{\kappa(\kappa-1)\ldots2\cdot1}=\frac{\nu(\nu-1)\ldots(\nu-\kappa+1)(\nu-\kappa)\ldots2\cdot1}{\kappa(\kappa-1)\ldots2\cdot1\cdot(\nu-\kappa)\ldots2\cdot1}=\frac{\nu!}{\kappa!(\nu-\kappa)!}$$

Έτσι καταλήγουμε στο $\boxed{\binom{\nu}{\kappa}=\frac{\nu!}{\kappa!(\nu-\kappa)!}\quad(\mu\varepsilon\ \kappa\le\nu)}$.

Εφαρμογή 1

Αν Α={α,β,γ,δ,ε}, πόσα υποσύνολα με 3 στοιχεία μπορούμε να φτιάξουμε και ποια είναι αυτά;

Το πλήθος των υποσυνόλων είναι $\binom{5}{3}=\frac{5!}{3!2!}=\frac{3!\cdot4\cdot5}{3!\cdot2}=10$.

Τα υποσύνολα αυτά είναι:
{α,β,γ},{α,β,δ},{α,β,ε},{α,γ,δ},{α,γ,ε},{α,δ,ε},{β,γ,δ},{β,γ,ε},{β,δ,ε},{γ,δ,ε}

Εφαρμογή 2

Με πόσους τρόπους μπορούμε να χωρίσουμε μία ομάδα 10 ατόμων σε τρεις υποομάδες των 5,3 και 2 ατόμων;

Η επιλογή των ατόμων της πρώτης ομάδας γίνεται με $\binom{10}{5}$ τρόπους.

Η επιλογή των μελών της δεύτερης ομάδας γίνεται από τους υπόλοιπους 5 με $\binom{5}{3}$ τρόπους και τέλος η επιλογή των μελών της τρίτης ομάδας γίνεται με $\binom{2}{2}$ τρόπους.

Έτσι, συνολικά έχουμε: $\binom{10}{5}\binom{5}{3}\binom{2}{2}=\frac{10!}{5!5!}\cdot\frac{5!}{3!2!}\cdot\frac{2!}{2!0!}=\frac{5!\cdot6\cdot7\cdot8\cdot9\cdot10}{5!\cdot3\cdot2\cdot2}=2.520$ τρόπους επιλογής των μελών της κάθε ομάδας.

Το αποτέλεσμα δεν αλλάζει αν αλλάξουμε τη σειρά των ομάδων, για παράδειγμα $\binom{10}{3}\binom{2}{2}\binom{5}{5}=2.520$

Όμως οι τρεις ομάδες μπορούν να μετατεθούν μεταξύ τουςμε 3! = 6 τρόπους και από μία τέτοια μετάθεση δεν προκύπτει διαφορετικός χωρισμός. Τελικά λοιπόν έχουμε $\frac{2.520}{6}=420$ τρόπους να φτιάξουμε τις ομάδες.

Εφαρμογή 3

Δείξτε ότι : α) $\binom{\nu}{1}=\nu$, β) $\binom{\nu}{\nu}=1$, γ) $\binom{\nu}{\kappa}=\binom{\nu}{\nu-\kappa}$.

α) $\binom{\nu}{1}=\frac{\nu!}{1!(\nu-1)!}=\frac{(\nu-1)!\nu}{(\nu-1)!}=\nu$

β) $\binom{\nu}{\nu}=\frac{\nu!}{\nu!0!}=\frac{\nu!}{\nu!}=1$

γ) $\binom{\nu}{\kappa}=\frac{\nu!}{\kappa!(\nu-\kappa)!}=\frac{\nu!}{(\nu-\kappa)!(\nu-\kappa+\kappa)!}=\frac{\nu!}{(\nu-\kappa)!(\nu-(\nu-\kappa))!}=\binom{\nu}{\nu-\kappa}$

Αν έχουμε τρία είδη σφαιρών, κόκκινες, πράσινες και μπλε και μπορούμε να επιλέξουμε δύο, με ποιους τρόπους μπορούμε να το κάνουμε; Εύκολα μπορούμε να καταγράψουμε όλα τα αποτελέσματα, γιατί είναι λίγα. Αυτά είναι λοιπόν κκ, κπ, κμ, πп, πμ, μμ.

Κάθε ένα από αυτά τα αποτελέσματα είναι ένας συνδυασμός με επανάληψη 3 στοιχείων ανά 2.

Συνδυασμός με επανάληψη ν στοιχείων ενός συνόλου Α ανά κ, είναι κάθε ομάδα κ στοιχείων του Α, όχι κατ' ανάγκη διαφορετικών μεταξύ τους και χωρίς να μας ενδιαφέρει η σειρά.

Το πλήθος των συνδυασμών με επανάληψη ν στοιχείων ανά κ, συμβολίζεται με $\begin{bmatrix}\nu\\ \kappa\end{bmatrix}$ και αποδεικνύεται ότι είναι ίσο με :

Παρατηρούμε ότι, όπως και στις διατάξεις με επανάληψη, το κ μπορεί να είναι μεγαλύτερο από το ν.

$$\boxed{\begin{bmatrix}\nu\\ \kappa\end{bmatrix}=\binom{\nu+\kappa-1}{\kappa}=\frac{(\nu+\kappa-1)!}{\kappa!(\nu-1)!}}$$

Εφαρμογή 4

Πόσα είναι τα διαφορετικά αποτελέσματα όταν ρίχνουμε δύο ζάρια;

Όταν ρίχνουμε ένα ζάρι, έχουμε 6 διαφορετικά αποτελέσματα. Όταν ρίχνουμε δύο ζάρια θα έχουμε συνδυασμούς αυτών των 6 αποτελεσμάτων ανά 2, αλλά μπορούμε να έχουμε και το ίδιο αποτέλεσμα δύο φορές, π.χ. δύο εξάρια. Έτσι, έχουμε συνδυασμούς με επανάληψη και το πλήθος τους είναι:

$$\begin{bmatrix} 6 \\ 2 \end{bmatrix} = \frac{(6+2-1)!}{2!(6-1)!} = \frac{7!}{2!5!} = \frac{7\cdot 6}{2} = 21$$

Ας προσπαθήσουμε τώρα να απαντήσουμε το πρόβλημα του προπονητή που είδαμε στην αρχή του κεφαλαίου. Ο προπονητής έχει στη διάθεσή του 3 τερματοφύλακες, 6 αμυντικούς, 5 μέσους και 3 επιθετικούς. Αφού η ομάδα παίζει με σύστημα 4,3,3, ο προπονητής πρέπει να διαλέξει 1 τερματοφύλακα από

τους 3 (αυτό γίνεται με $\binom{3}{1}$ τρόπους), 4 αμυντικούς από τους 6 (με $\binom{6}{4}$ τρόπους), 3 μέσους από τους 5 (με $\binom{3}{3}$ τρόπους) και 3 επιθετικούς από τους 3 (με $\binom{5}{3}$ τρόπους).

Έτσι, σύμφωνα με την αρχή της απαρίθμησης, έχει τρόπους να συνθέσει τη εντεκάδα. (Δύσκολη δουλειά να είσαι προπονητής!).

ΕΞΑΣΚΗΣΗ

1. Αντιστοίχησε τα στοιχεία της πρώτης στήλης με τα στοιχεία της δεύτερης.

Α	Β
$\binom{4}{2}$	5
$\binom{2}{4}$	6
$\begin{bmatrix} 4 \\ 2 \end{bmatrix}$	12
$\begin{bmatrix} 2 \\ 4 \end{bmatrix}$	15
	30
	Δεν ορίζεται

2. Αντιστοίχησε τα στοιχεία της πρώτης στήλης με τα στοιχεία της δεύτερης.

Α	Β
$\binom{\nu}{0}$	0
$\binom{\nu}{1}$	1
$\binom{\nu}{\nu-1}$	$\nu-1$
$\binom{\nu}{\nu}$	ν
$\binom{\nu}{2}$	$\nu^2-\nu$
	$\frac{\nu^2-\nu}{2}$
	Δεν ορίζεται

3. Σε ένα τουρνουά σκακιού, πήραν μέρος 6 παίκτες και ο καθένας αντιμετώπισε από μία φορά όλους τους άλλους. Πόσοι αγώνες έγιναν;

4. Σε μία δεξίωση, ο κάθε καλεσμένος αντάλλαξε μία χειραψία με κάθε έναν από τους υπόλοιπους κα συνολικά έγιναν 45 χειραψίες. Πόσοι ήταν οι καλεσμένοι;

5. Ένας σύλλογος έχει 20 μέλη. Με πόσους τρόπους μπορεί να εκλέξει:
   α. μία τριμελή διοικητική επιτροπή
   β. πρόεδρο, γραμματέα και ταμία.

6. Σε ένα δελτίο LOTTO έχουν σημειωθεί 8 αριθμοί. Πόσες ομάδες έχουν παιχτεί στο δελτίο αυτό;

7. Σε μία εκδήλωση που παρευρίσκονται 15 άτομα, πρόκειται να διανεμηθούν με κλήρωση 5 δώρα. Με πόσους τρόπους μπορεί να γίνει αυτό;
8. Αν έχουμε στη διάθεσή μας κέρματα των 5,10,20 και 50 λεπτών, με πόσους τρόπους μπορούμε να φτιάξουμε μία ομάδα 6 νομισμάτων;

9. Πόσα είναι τα δυνατά αποτελέσματα, όταν :
   α. ρίχνουμε ένα ζάρι τρεις φορές
   β. ρίχνουμε τρία ζάρια ταυτόχρονα.

10. Μία τράπουλα έχει 52 χαρτιά. Σε ένα παιχνίδι με δύο παίκτες, πρέπει ο καθένας να πάρει 6 χαρτιά. Με πόσους τρόπους μπορεί να γίνει αυτό;

11. Πόσους αριθμούς με διαφορετικά ψηφία μπορούμε να σχηματίσουμε αν πάρουμε 3 από τα ψηφία 1,2,3,4 και 2 από τα 5,6,7,8,9.

12. Μία τάξη αποτελείται από 6 μαθητές και 9 μαθήτριες. Θέλουμε να εκλέξουμε μία τετραμελή αντιπροσωπεία. Με πόσους τρόπους μπορεί να γίνει η επιλογή, αν:

α. ο αριθμός των μαθητών πρέπει να είναι ίσος με τον αριθμό των μαθητριών

β. η αντιπροσωπεία πρέπει να περιέχει μαθητές και των δύο φύλων

γ. η αντιπροσωπεία πρέπει να περιέχει τουλάχιστον δύο μαθήτριες.

13. Δείξτε ότι:

α. $\binom{\nu}{\kappa} = \binom{\nu-1}{\kappa-1} + \binom{\nu-1}{\kappa}$

β. $\binom{\kappa+\mu}{\kappa}\binom{\nu}{\kappa+\mu} = \binom{\nu}{\kappa}\binom{\nu-\kappa}{\mu}$

γ. $\kappa\binom{\nu}{\kappa} = \nu\binom{\nu-1}{\kappa-1}$

14. Να βρείτε τα ν και κ, αν οι διατάξεις ν στοιχείων ανά κ, είναι 120 και οι συνδυασμοί ν στοιχείων ανά κ είναι 20.

Σύνοψη

Όταν μας δίνουν ένα σύνολο που αποτελείται από ν αντικείμενα (στοιχεία) και θέλουμε να σχηματίσουμε με αυτή διαφορερικές μεταξύ τους ομάδες που η καθεμία να διαφέρουν από τις άλλες κατά τη θέση των αντικειμένων η κατά τη φύση της. Στην περίπτωση που το πλήθος τω αντικειμένων είναι πολύ μεγάλο ο προσδιορισμός του πλήθους των ομάδων είαν δύσκολο να βρεθεί με απλό πρακτικό τρόπο για το λόγο αυτό χρησιμοποιούμε τη συνδυαστική ανάλυση.

- Συνδιαστική ανάλυση → Βασική αρχή απαρίθμησης
- Συνδιαστική ανάλυση → Διατάξεις
- Συνδιαστική ανάλυση → Μεταθέσεις
- Συνδιαστική ανάλυση → Συνδιασμοί

Τα παραπάνω είναι χρήσιμα στη θεωρία λογισμού πιθανοτήτων. Ο εκπαιδευόμενος θα είναι σε θέση να απαριθμήση σύνολα που προκύπτουν σε πραγματικά προβλήματα και θα μπορεί έτσι να υπολογίσει την πιθανότητα να συμβούν αντίστοιχα ενδεχόμενα.

Βιβλιογραφία / Internet

«Applied Combinatories», Tucker A.C., 1990

«Combinatorial theory», Hall M., 1967

«Εισαγωγή στη Συνδυαστική», Πανάρετος, 1995

«Συνδυαστική απαρίθμιση», Μωυσιάδης Χρόνης

«Συνδυαστική», Θ. Ν. Καζαντζής, Μαθητική βιβλιοθήκη Βαφειάδης

«Συνδυαστική», Χαράλαμπου Α. Χαραλαμπίδη, Συμμετρία

www.gomath.com: *site για βοήθεια στα μαθηματικά, που περιέχει αλγόριθμους επίλυσης εξισώσεων, γεωμετρία, εξηγήσεις για μαθηματικά σύμβολα, πίνακες με μαθηματικές πράξεις κ.ά.*

http://combinators.scientist.gr: *στο site υπάρχουν βασικές θεωρίες για μεταθέσεις και διατάξεις, συνδυασμοί με ή χωρίς επανάληψη, δημιουργία συνδυασμών καθώς και αρκετά παραδείγματα.*

ΟΔΗΓΟΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΜΕΛΕΤΗ

«Λογισμός πιθανοτήτων», Εκδόσεις Σταμούλης, Αποστ. Κιόχος

Στο βιβλίο αναπτύσσονται η θεωρία της Συνδυαστικής η έννοια της πιθανότητας. Θέματα που αφορούν τυχαίες και θεωρητικές κατανομές με πληθώρα παραδειγμάτων και λυμένων εφαρμογών.

7. ΠΙΘΑΝΟΤΗΤΕΣ

Βασικές έννοιες:
- δειγματικός χώρος
- ενδεχόμενο
- πράξεις με σύνολα

Στόχος του μαθήματος: ο εκπαιδευόμενος να μπορεί να υπολογίζει με ακρίβεια και ορθότητα την πιθανότητα να πραγματοποιηθεί ένα ενδεχόμενο, και αυτό ακόμη και σε σύνθετα προβλήματα. Ενδιάμεσος στόχος είναι, προκειμένου να συμβεί αυτό, να εντοπίζει βεβαίως τα σύνολα αυτά στα προβλήματα που του παρουσιάζονται.

Προσδοκώμενα αποτελέσματα: Καθώς ο εκπαιδευόμενος πετύχει τον παραπάνω στόχο, θα είναι έτοιμος να περάσει στο επόμενο κεφάλαιο, των τυχαίων μεταβλητών, όπου τα όσα έμαθε στο παρόν κεφάλαιο, θα λειτουργήσουν ως βάση – εργαλείο για την εις βάθος μελέτη της έννοιας της πιθανότητας, όπως χρησιμοποιείται σήμερα.

Εισαγωγικές παρατηρήσεις: Με την έννοια της πιθανότητας είμαστε εξοικειωμένοι από την καθημερινότητά μας, είτε αφορά τη ρίψη των ζαριών στο τάβλι, είτε την αποτελεσματικότητα μιας θεραπείας.
Προφανώς στη δεύτερη περίπτωση το θέμα είναι σοβαρό. Άρα η πιθανότητα έχει καθοριστικό ρόλο στη ζωή μας.
Τι σημαίνει, όμως, για τα μαθηματικά η πιθανότητα;
Οι bookers δίνουν αγώνες του αγγλικού ποδοσφαίρου 2 προς 1, ή στα μαθηματικά 2/1. Έχουμε 98% πιθανότητα να είναι αποτελεσματική μία θεραπεία. Δηλαδή 98 στους 100 ασθενείς θεραπεύονται ή στις 100 περιπτώσεις, οι 98 μας ευνοούν (ευνοϊκές). Δηλαδή χονδρικά θα λέγαμε ότι διαιρούμε τις ευνοϊκές περιπτώσεις προς όλες τις περιπτώσεις και προκύπτει η πιθανότητα. Αυτό στα μαθηματικά διατυπώνεται μέσω του ακόλουθου ορισμού της πιθανότητας:
P(A): πιθανότητα του ενδεχομένου που μας ενδιαφέρει (π.χ. αποτελεσματικότητα θεραπείας
N(Ω): σύνολο των περιπτώσεων που έχουμε (π.χ. όλοι οι ασθενείς)

7.1. ΔΕΙΓΜΑΤΙΚΟΣ ΧΩΡΟΣ – ΕΝΔΕΧΟΜΕΝΑ

Γνωρίζουμε από τη Φυσική ότι, αν ζεστάνουμε ένα δοχείο με απεσταγμένο νερό στη συνηθισμένη ατμοσφαιρική πίεση (760 mmHg), αυτό θα βράσει στους 100^{o} C. Επίσης γνωρίζουμε ότι αν αφήσουμε ένα μικρό αντικείμενο να πέσει ελεύθερα από ύψος h, αυτό θα φθάσει στο έδαφος σε χρόνο $t=\sqrt{\frac{2h}{g}}$.

Όσες φορές και αν επαναλάβουμε τα παραπάνω πειράματα, θα έχουμε το ίδιο αποτέλεσμα. Ένα πείραμα κατά το οποίο η γνώση των συνθηκών υπό τις οποίες πραγματοποιείται καθορίζει πλήρως το αποτέλεσμα λέγεται αιτιοκρατικό.

Υπάρχουν όμως και πειράματα κατά τα οποία δεν μπορεί να γίνει ακριβής πρόβλεψη του αποτελέσματος. Αν καταγράψουμε τη θερμοκρασία στις 12 το μεσημέρι κάθε μέρας για ένα μήνα, τα αποτελέσματα μπορεί να βρίσκονται μέσα σε κάποια λογικά πλαίσια, αλλά δεν μπορούμε να τα ξέρουμε εκ των προτέρων. Επίσης, αν μετρήσουμε το χρόνο που χρειαζόμαστε κάθε πρωί για να πάμε στη δουλειά μας με το αυτοκίνητο, δεν είναι κάθε φορά ο ίδιος, ακόμα και αν φροντίσουμε να φύγουμε την ίδια ώρα και ακολουθήσουμε την ίδια διαδρομή γιατί εξαρτάται από ποια φανάρια θα μας πιάσουν, τι κίνηση θα έχει, αν θα μας συμβεί κάποιο ατύχημα κ.λ.π. Ένα πείραμα το οποίο μπορεί να επαναληφθεί όσες φορές θέλουμε κάτω από τις ίδιες ακριβώς συνθήκες και του οποίου δεν μπορούμε να προβλέψουμε το αποτέλεσμα, λέγεται **πείραμα τύχης**.

Πειράματα τύχης, είναι για παράδειγμα:

– το ρίξιμο ενός ζαριού και η καταγραφή του αποτελέσματος.
– Η κλήρωση του JOKER για τον καθορισμό των αριθμών που κερδίζουν.
– Σταματάμε τυχαία άτομα στο δρόμο και καταγράφουμε την ηλικία τους.
– Καταγράφουμε τις μετρήσεις της AGB για την τηλεθέαση ενός προγράμματος κατά τη διάρκεια ενός μήνα.
– Βγάζουμε ένα χαρτί από μία τράπουλα και κοιτάζουμε ποιο είναι.

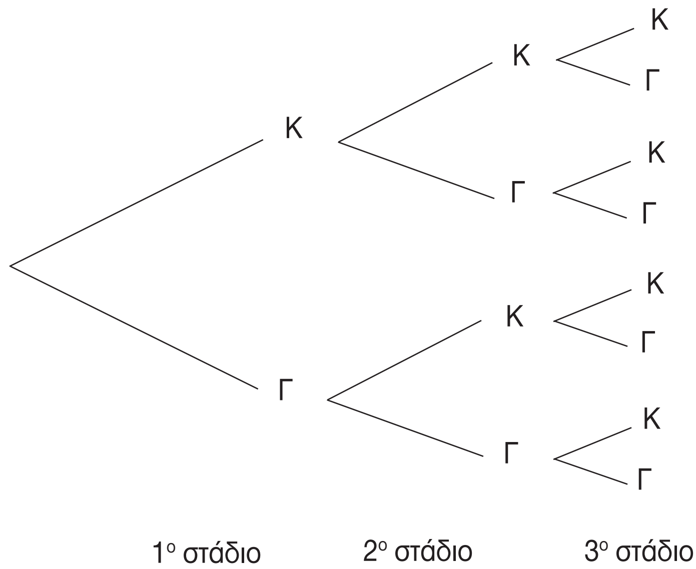
Το σύνολο όλων των δυνατών αποτελεσμάτων ενός πειράματος τύχης, λέγεται **δειγματικός χώρος** του πειράματος και συμβολίζεται με Ω. Όταν ρίχνουμε ένα ζάρι ο δειγματικός χώρος του πειράματος είναι Ω={1,2,3,4,5,6}. Όταν επιλέγουμε τυχαία ένα τηλεφώνημα και καταγράφουμε τη διάρκειά του, ο δειγματικός χώρος είναι $\Omega=(0,+\infty)$

Το πλήθος των στοιχείων του δειγματικού χώρου, μπορεί να υπολογιστεί με τη βοήθεια των κανόνων της συνδυαστικής. Η αναλυτική καταγραφή όμως των στοιχείων του, απαιτεί μεγάλη προσοχή, ιδιαίτερα αν αυτά είναι πολλά. Σ' αυτό μπορεί να μας βοηθήσει το λεγόμενο δενδροδιάγραμμα.

Εφαρμογή 1

Ρίχνουμε ένα νόμισμα τρεις φορές και καταγράφουμε το αποτέλεσμα Κ (κεφαλή) ή Γ (γράμματα). Ποιος είναι ο δειγματικός χώρος του πειράματος;

Επειδή πρόκειται για διατάξεις με επανάληψη, το πλήθος των στοιχείων του Ω θα είναι: $[2]_3=2^3=8$. Για να βρούμε ποια είναι τα στοιχεία του, σημειώνουμε στο παρακάτω δενδροδιάγραμμα τα αποτελέσματα κάθε σταδίου του πειράματος:



Τα στοιχεία του δειγματικού χώρου περιγράφονται από τις διακλαδώσεις του δένδρου, δηλαδή από όλες τις δυνατές διαδρομές από τη ρίζα μέχρι την άκρη κάθε κλαδιού. Έτσι έχουμε Ω = {ΚΚΚ, ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΚΚ, ΓΚΓ, ΓΓΚ, ΓΓΓ}.

Κάθε υποσύνολο του δειγματικού χώρου Ω ενός πειράματος τύχης, ονομάζεται **ενδεχόμενο**. Για παράδειγμα, όταν ρίχνουμε ένα ζάρι, τα σύνολα Α = {5}, Β = {2,6}, Γ = {1,3,5}, Δ = {1,2,4,6} είναι ενδεχόμενα. Ένα ενδεχόμενο μπορεί να είναι **απλό** αν έχει μόνο ένα στοιχείο ή

Σύνθετο αν έχει περισσότερα στοιχεία. Στο προηγούμενο παράδειγμα, το ενδεχόμενο Α είναι απλό, ενώ τα Β,Γ,Δ είναι σύνθετα. Λέμε ότι ένα ενδεχόμενο **πραγματοποιείται**, όταν το αποτέλεσμα του πειράματος ανήκει στο δενδροδιάγραμμα αυτό, δηλαδή αν ρίξουμε ένα ζάρι και το αποτέλεσμα είναι 6, τότε πραγματοποιούνται τα Β και Δ, ενώ δεν πραγματοποιούνται τα Α και Γ. Τα στοιχεί ενός ενδεχόμενου τα λέμε και **ευνοϊκές περιπτώσεις** για την πραγματοποίησή του, δηλαδή το ενδεχόμενο Β έχει δύο ευνοϊκές περιπτώσεις γιατί πραγματοποιείται αν το αποτέλεσμα του πειράματος είναι 2 ή 6.

Υποσύνολο του δειγματικού χώρου Ω θεωρείται και το ίδιο το Ω και συνεπώς ολόκληρο το Ω είναι ένα ενδεχόμενο που ονομάζεται **βέβαιο** ενδεχόμενο, γιατί πραγματοποιείται πάντοτε οποιοδήποτε και να είναι το αποτέλεσμα του πειράματος τύχης. Υποσύνολο του Ω και συνεπώς ενδεχόμενο, θεωρείται και το κενό σύνολο που δεν περιέχει κανένα στοιχείο.

Το σύνολο αυτό το συμβολίζουμε με $\varnothing$ ή {} και το ονομάζουμε **αδύνατο** ενδεχόμενο, γιατί δεν πραγματοποιείται ποτέ, όποιο και αν είναι το αποτέλεσμα του πειράματος τύχης.

Το πλήθος των στοιχείων ενός ενδεχόμενου Α, ονομάζεται **πληθάριθμος** του ενδεχόμενου και συμβολίζεται με Ν(Α). Σύμφωνα με τα παραδείγματα που αναφέρθηκαν παραπάνω για το ρίξιμο ενός ζαριού, έχουμε Ν(Α)=1, Ν(Β)=2, Ν(Ω)=6,Ν($\varnothing$)=0.

Εφαρμογή 2

Όταν ρίχνουμε ένα ζάρι τρεις φορές (βλέπε εφαρμογή 1), να καταγράψετε αναλυτικά τα ενδεχόμενα:

Α. να έχουμε 2 φορές Κεφαλή

Β. να έχουμε τουλάχιστον 2 φορές Κεφαλή

Γ. να έχουμε το πολύ 2 φορές Κεφαλή

Δ. να έχουμε ίσο αριθμό Κεφαλών και Γραμμάτων

Ο δειγματικός χώρος Ω, είναι ο αναφερόμενος στην εφαρμογή 1.

Α. το ενδεχόμενο είναι Α = {ΚΚΓ, ΚΓΚ, ΓΚΚ}

Β. το ενδεχόμενο πραγματοποιείται αν έχουμε 2 ή 3 φορές Κεφαλή και συνεπώς Β = {ΚΚΚ, ΚΚΓ, ΚΓΚ, ΓΚΚ}.

Γ. το ενδεχόμενο πραγματοποιείται αν έχουμε 2 ή 1 ή καμία Κεφαλή και συνεπώς Γ = {ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΚΚ, ΓΚΓ, ΓΓΚ, ΓΓΓ}.

Δ. το ενδεχόμενο δεν πραγματοποιείται ποτέ, γι' αυτό Δ = $\varnothing$.

Εφαρμογή 3

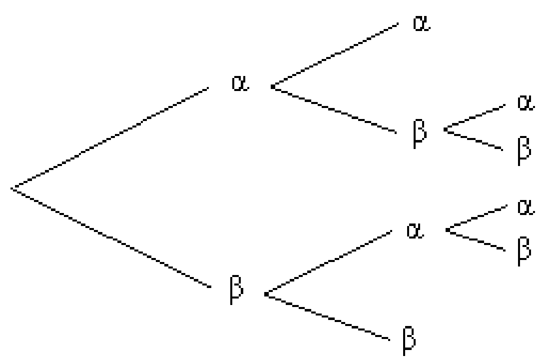
Ένας αγώνας βόλεϊ γυναικών τελειώνει ότι μία από τις δύο ομάδες κέρδισε 2 σετ. Κάθε τέτοιος αγώνας είναι ένα πείραμα τύχης. Να καταγράψετε το δειγματικό χώρο του πειράματος και τα ενδεχόμενα:

Α. να κερδίσει η πρώτη ομάδα

Β. να παιχτούν συνολικά 3 σετ.

Γ. να τελειώσει το ματς ισόπαλο.

Συμβολίζουμε μα α τη νίκη της πρώτης ομάδας σε κάποιο σετ και αντίστοιχα με β τη νίκη της δεύτερης ομάδας. Ο δειγματικός χώρος του πειράματος μπορεί να περιγραφεί από το επόμενο δενδροδιάγραμμα.



Έτσι έχουμε Ω = {αα, αβα, αββ, βαα, βαβ, ββ}.

Α. Η πρώτη ομάδα κερδίζει όταν τα α είναι περισσότερα από τα β, δηλαδή Α = {αα, αβα, βαα}

Β. Είναι φανερό ότι Β = {αβα, αββ, βαα, βαβ}.

Γ. Ένα παιχνίδι βόλεϊ δεν τελειώνει ποτέ ισόπαλο, άρα Γ = $\varnothing$.

ΕΞΑΣΚΗΣΗ

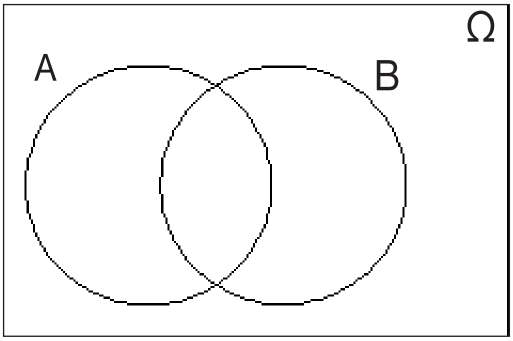
1. Μία οικογένεια από την Αθήνα, αποφασίζει να κάνει διακοπές στην Θεσσαλονίκη ή στην Κρήτη. Στη Θεσσαλονίκη μπορεί να πάει με αυτοκίνητο, τρένο ή αεροπλάνο, ενώ στην Κρήτη με πλοίο ή αεροπλάνο. Αν η επιλογή του τόπου και του μέσου γίνει τυχαία, να βρείτε το δειγματικό χώρο του πειράματος και τα ενδεχόμενα

α) η οικογένεια να πάει διακοπές στη Θεσσαλονίκη,

β) η οικογένεια να πάει διακοπές με αεροπλάνο.

2. Σε ένα κουτί, έχουμε μία κόκκινη, μία πράσινη και μία μπλε σφαίρα. Βγάζουμε διαδοχικά δύο σφαίρες χωρίς να ξαναβάλουμε την πρώτη μέσα στο κουτί.
   i. Ποιος είναι ο δειγματικός χώρος του πειράματος;
   ii. Ποιο είναι το ενδεχόμενο να βγάλουμε τουλάχιστον μία κόκκινη σφαίρα;
   iii. Ποιο είναι το ενδεχόμενο να μην βγάλουμε πράσινη σφαίρα;
   iv. Ποια είναι τα ενδεχόμενα να βγάλουμε ομοιόχρωμες σφαίρες;
3. Απαντήστε τα ερωτήματα της προηγούμενης άσκησης, αν ξαναβάλουμε μέσα στο κουτί την πρώτη σφαίρα πριν βγάλουμε τη δεύτερη.
4. Ρίχνουμε ένα ζάρι δύο φορές. Να βρείτε:
   i. Το δειγματικό χώρο του πειράματος.
   ii. Το ενδεχόμενο το αποτέλεσμα της πρώτης ρίψης να είναι μικρότερο από το αποτέλεσμα της δεύτερης.
   iii. Το ενδεχόμενο το άθροισμα των δύο ενδείξεων να είναι ζυγό και μικρότερο του 8.
5. Σε ένα κουτί έχουμε τρεις σφαίρες αριθμημένες 1,2,3. Επιλέγουμε μία - μία τις σφαίρες μέχρι να βγει ο αριθμός 3. Το πείραμα σταματάει μετά την τρίτη προσπάθεια. Να βρείτε το δειγματικό χώρο του πειράματος και τα ενδεχόμενα Α: να σταματήσουμε στη δεύτερη προσπάθεια και Β: να εμφανιστεί ο αριθμός 3, αν:
   i. Δεν ξαναβάλουμε μέσα στο κουτί την επιλεγμένη σφαίρα.
   ii. Ξαναβάλουμε μέσα στο κουτί την επιλεγμένη σφαίρα.
6. Σε ένα κιβώτιο υπάρχουν 20 λάμπες, από τις οποίες οι μισές είναι χαλασμένες. Δοκιμάσαμε μία - μία τις λάμπες μέχρι να βρούμε μία καλή (εννοείται ότι δεν ξαναβάζουμε στο κιβώτιο τις χαλασμένες). Να βρείτε το δειγματικό χώρο του πειράματος και το ενδεχόμενο να δοκιμάσουμε 5 λάμπες.

7.2. ΠΡΑΞΕΙΣ ΜΕ ΕΝΔΕΧΟΜΕΝΑ

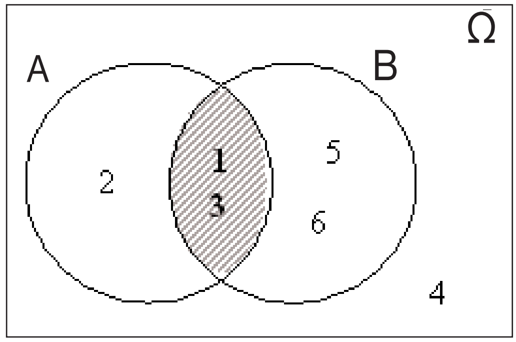


Το δειγματικό χώρο και τα ενδεχόμενα εφόσον είναι σύνολα, μπορούμε να τα παραστήσουμε με διαγράμματα Venn. Συμβολίζουμε δηλαδή το δειγματικό χώρο Ω με ένα ορθογώνιο παραλληλόγραμμο και κάθε ενδεχόμενο με μία κλειστή καμπύλη στο εσωτερικό του Ω.

Μεταξύ των ενδεχομένων μπορούμε να ορίσουμε τις γνωστές πράξεις μεταξύ των συνόλων από τις οποίες προκύπτουν νέα ενδεχόμενα.

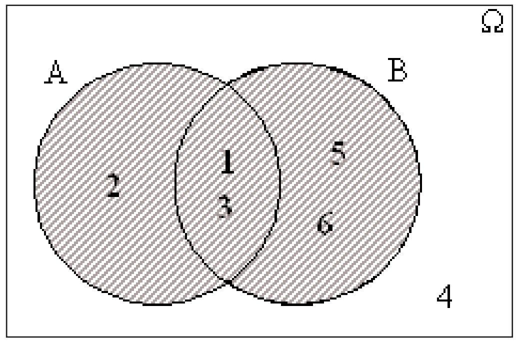
Στη συνέχεια, θα χρησιμοποιήσουμε σαν παράδειγμα το δειγματικό χώρο που προκύπτει από το ρίξιμο ενός ζαριού και τα ενδεχόμενα Α={1,2,3} και Β={1,3,5,6}.

Τομή δύο ενδεχομένων Α και Β είναι το ενδεχόμενο που πραγματοποιείται αν πραγματοποιηθούν και τα δύο ενδεχόμενα Α και Β συγχρόνως, δηλαδή αποτελείται από τα κοινά στοιχεία Α και Β. Την τομή των Α και Β τη συμβολίζουμε με $A \cap B$.



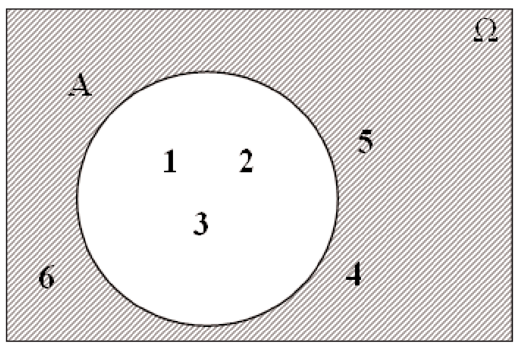
$A \cap B$ ={1,3}

Ένωση δύο ενδεχομένων Α και Β είναι το ενδεχόμενο που πραγματοποιείται αν πραγματοποιηθεί τουλάχιστον ένα από τα Α, Β, δηλαδή αποτελείται από τα στοιχεία που ανήκουν τουλάχιστον σε ένα από τα Α, Β. Την ένωση των Α και Β τη συμβολίζουμε με $A \cup B$.



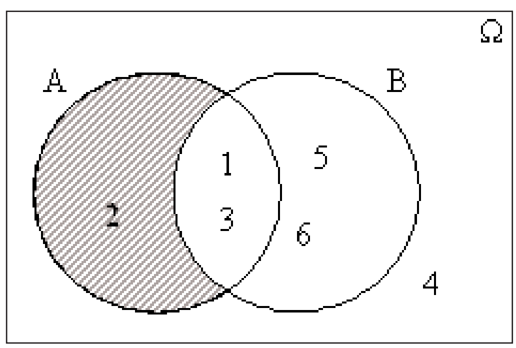
$A \cup B$ ={1,2,3,5,6}

Συμπλήρωμα του ενδεχομένου Α είναι το ενδεχόμενο που πραγματοποιείται όταν δεν πραγματοποιείται το Α, δηλαδή αποτελείται από τα στοιχεία του Ω που δεν ανήκουν στο Α. Το συμπλήρωμα του Α το συμβολίζουμε με Α΄.



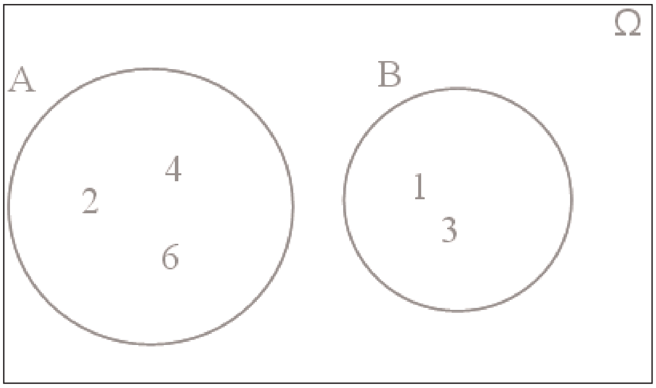
Α΄={4,5,6}

Διαφορά του ενδεχομένου Β από το ενδεχόμενο Α, είναι το ενδεχόμενο που πραγματοποιείται όταν πραγματοποιείται το Α αλλά όχι το Β, δηλαδή αποτελείται από τα στοιχεία του Α που δεν ανήκουν στο Β. Τη διαφορά αυτή τη συμβολίζουμε με Α-Β.



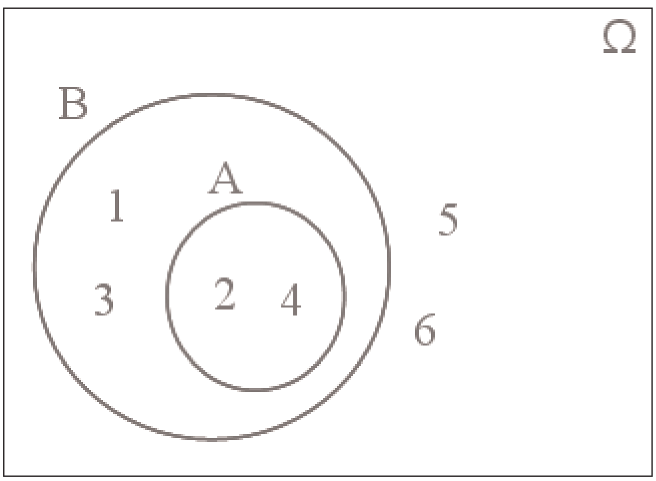
Α–Β={2}

Ας υποθέσουμε τώρα ότι Α = {2,4,6} και Β = {1,3}. Δύο ενδεχόμενα λέγονται **ασυμβίβαστα** ή **ξένα**, αν η πραγματοποίηση του ενός αποκλείει την πραγματοποίηση του άλλου. Στην περίπτωση αυτή, τα ενδεχόμενα δεν έχουν κοινά στοιχεία, δηλαδή $A \cap B = \varnothing$.



Α,Β ασυμβίβαστα

Τέλος, λέμε ότι το ενδεχόμενο Α **συνεπάγεται** το ενδεχόμενο Β αν η πραγματοποίηση του Α επιβάλλει την πραγματοποίηση του Β, δηλαδή το Α είναι υποσύνολο του Β (ένα σύνολο Α λέμε ότι είναι υποσύνολο του Β, αν το Α περιέχεται μέσα στο Β, δηλαδή κάθε στοιχείο του Α είναι και στοιχείο του Β). Τότε συμβολίζουμε $A \subseteq B$. Αυτό συμβαίνει για παράδειγμα αν Α = {2,4} και Β = {1,2,3,4}.



$A \subseteq B$

Ιδιότητες των πράξεων

1. Για τα σύνολα Α και Α΄ ισχύουν:

$$(A')' = A$$
$$A \cap A' = \varnothing$$
$$A \cup A' = \Omega$$

Είναι προφανείς από τον ορισμό του συμπληρώματος

2. Η τομή συνόλων έχει την αντιμεταθετική και προσεταιριστική ιδιότητα:

$$\boxed{\begin{gathered} A \cap B = B \cap A \\ (A \cap B) \cap \Gamma = A \cap (B \cap \Gamma) \end{gathered}}$$

Αποδεικνύεται εύκολα από ορισμό της πράξης. Επειδή ισχύει η προσεταιριστική, συνήθως γράφουμε $A \cap B \cap \Gamma$ χωρίς παρενθέσεις και μπορούμε να επεκτείνουμε την πράξη για όσα σύνολα θέλουμε, δηλαδή να έχουμε $A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_\nu$.

3. Η ένωση συνόλων έχει την αντιμεταθετική και προσεταιριστική ιδιότητα: Αποδεικνύεται εύκολα από τον ορισμό της πράξης. Λόγω της προσεταιριστικής, συνήθως γράφουμε $A \cup B \cup \Gamma$ και μπορούμε να επεκτείνουμε την

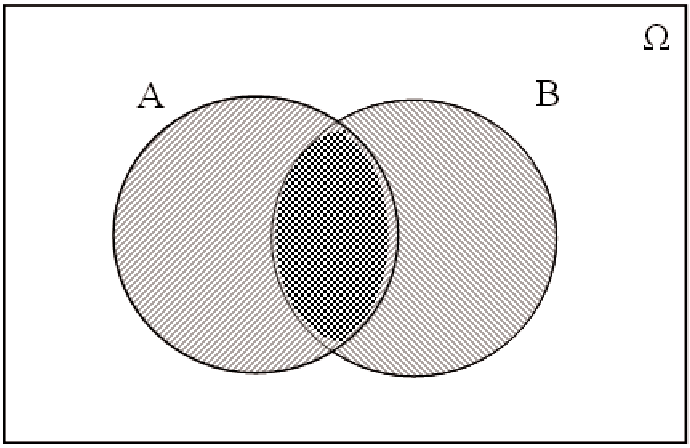
$$\boxed{\begin{gathered} A \cup B = B \cup A \\ (A \cup B) \cup \Gamma = A \cup (B \cup \Gamma) \end{gathered}}$$

πράξη για όσα σύνολα θέλουμε $A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_\nu$.

4. Αν $A \subseteq B$, τότε $A \cap B = A$ και $A \cup B = B$. Είναι προφανές από τον ορισμό των πράξεων και του υποσυνόλου.

5. Ισχύει πάντοτε $\boxed{N(A \cup B) = N(A) + N(B) - N(A \cap B)}$. Αν προσθέσουμε το πλήθος των στοιχείων

των ενδεχομένων Α και Β, τότε έχουμε μετρήσει δύο φορές το πλήθος των στοιχείων του $A \cap B$.

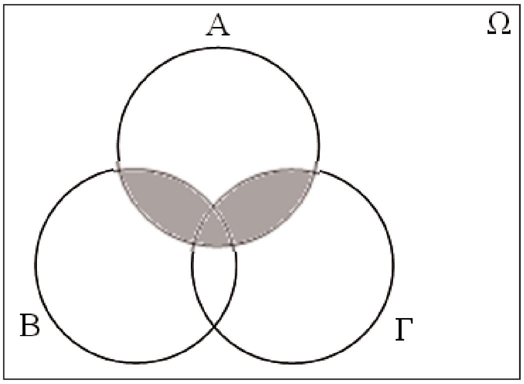


Για να υπολογίσουμε λοιπόν σωστά το πλήθος των στοιχείων του $A \cup B$ πρέπει να αφαιρέσουμε από το άθροισμα το πλήθος των στοιχείων που έχουν υπολογιστεί δύο φορές.

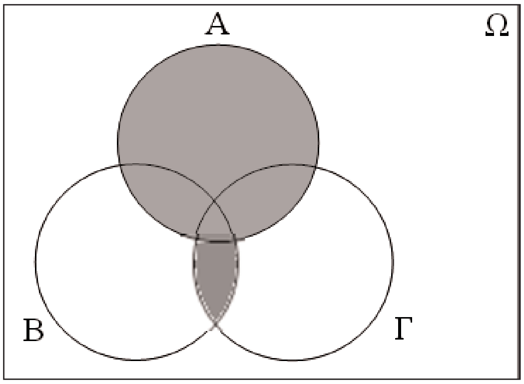
6. Αν Α,Β είναι ασυμβίβαστα, τότε $N(A \cup B) = N(A) + N(B)$, γιατί τότε $A \cap B = \varnothing$ και ο προηγούμενος τύπος μας δίνει:

7. Ισχύει η επιμεριστική ιδιότητα της ένωσης ως προς την τομή και αντίστροφα:

$$N(A \cup B) = N(A) + N(B) - N(A \cap B) = N(A) + N(B) - 0 = N(A) + N(B)$$

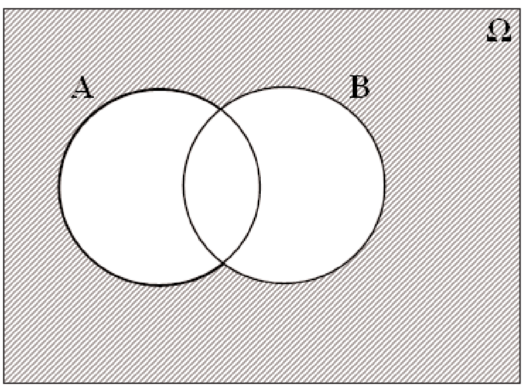


$$A \cap (B \cup \Gamma) = (A \cap B) \cup (A \cap \Gamma)$$

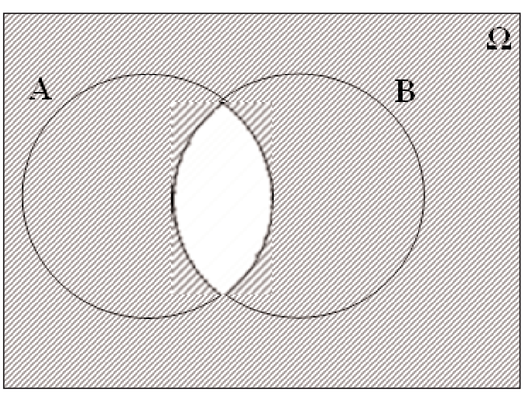


$$A \cup (B \cap \Gamma) = (A \cup B) \cap (A \cup \Gamma)$$

8. Ισχύουν οι λεγόμενοι τύποι του de Morgan, δηλαδή:



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



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

Στον πίνακα που ακολουθεί, αναφέρονται μερικές συνηθισμένες εκφράσεις και τα ενδεχόμενα στα οποία αντιστοιχούν:

Δεν πραγματοποιείται το Α	A'
Πραγματοποιούνται το Α και το Β	$A \cap B$
Πραγματοποιείται το Α ή το Β	$A \cup B$
Πραγματοποιείται μόνο το Α	$A - B$
Δεν πραγματοποιείται ούτε το Α ούτε το Β	$(A \cup B)'$
Πραγματοποιείται μόνο ένα από τα Α και Β	$(A - B) \cup (B - A)$ ή $(A \cup B) - (A \cap B)$

Εφαρμογή 4

Μία οικογένεια έχει τρία παιδιά. Παριστάνουμε με Α το ενδεχόμενο να έχει ακριβώς δύο αγόρια και με Β το ενδεχόμενο το πρώτο παιδί να είναι κορίτσι. Να βρείτε τα ενδεχόμενα A', B', $A \cap B$, $A \cup B$, Α–Β, Β–Α.

Ο δειγματικός χώρος μπορεί να βρεθεί με τρόπο όμοιο με την εφαρμογή 1 της προηγούμενης παραγράφου και είναι: Ω = {ΑΑΑ, ΑΑΚ, ΑΚΑ, ΑΚΚ, ΚΑΑ, ΚΑΚ, ΚΚΑ, ΚΚΚ}. Τα δύο αρχικά ενδεχόμενα είναι Α = {ΑΑΚ, ΑΚΑ, ΚΑΑ} και Β = {ΚΑΑ, ΚΑΚ, ΚΑΚ, ΚΚΑ, ΚΚΚ}. Έτσι, έχουμε:

Α΄ = {ΑΑΑ, ΑΚΚ, ΚΑΚ, ΚΚΑ, ΚΚΚ}

Β΄ = {ΑΑΑ, ΑΑΚ, ΑΚΑ, ΑΚΚ,}

$A \cap B$={ΚΑΑ}

$A \cup B$={ΑΑΚ,ΑΚΑ,ΚΑΑ,ΚΑΚ,ΚΚΑ,ΚΚΚ}

Α – Β = {ΑΑΚ, ΑΚΑ}

Β – Α = {ΚΑΚ, ΚΚΑ, ΚΚΚ}

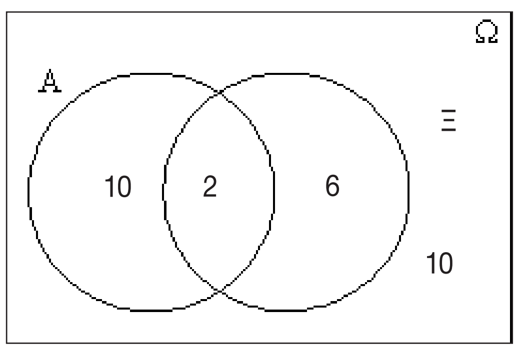
Εφαρμογή 5

Σε μία τάξη, υπάρχουν 28 μαθητές, από τους οποίους 12 είναι αγόρια. Επίσης, υπάρχουν 8 άτομα με ξανθά μαλλιά, από τα οποία 2 είναι αγόρια. Αν επιλέξουμε τυχαία ένα άτομο, να βρείτε το πλήθος των στοιχείων των ενδεχομένων:

Α. να είναι αγόρι και να μην έχει ξανθά μαλλιά

Β. να είναι κορίτσι και να μην έχει ξανθά μαλλιά

Γ. να είναι αγόρι ή να έχει ξανθά μαλλιά



Για να απαντήσουμε ένα τέτοιο ερώτημα, ο καλύτερος τρόπος είναι να σχεδιάσουμε ένα διάγραμμα Venn και να προσπαθήσουμε να συμπληρώσουμε το κάθε τμήμα του με το σωστό αριθμό. Έστω Α το ενδεχόμενο ότι είναι αγόρι και Ξ το ενδεχόμενο να έχει ξανθά μαλλιά. Από την εκφώνηση υπάρχουν 2 αγόρια με ξανθά μαλλιά και ανήκουν στο $A \cap \Xi$, 10 αγόρια που δεν έχουν ξανθά μαλλιά και ανήκουν στο Α–Ξ, 8–2=6 ξανθά κορίτσια που ανήκουν στο Ξ – Α και 28 – (10 + 2 + 6)=10 άτομα που ανήκουν στο $(A \cup \Xi)'$.

Α. Το πρώτο ερώτημα έχει ήδη απαντηθεί, αφού το ζητούμενο ενδεχόμενο είναι το Α–Ξ και είδαμε ότι Ν(Α–Ξ) = 10.

Β. Το δεύτερο ενδεχόμενο σημαίνει ότι δεν έχουμε αγόρι, ούτε άτομο με ξανθά μαλλιά, άρα αναφερόμαστε στο $(A \cup \Xi)'$ και $N((A \cup \Xi)') = 10$.

Εφαρμογή 6

Δείξτε ότι Α–Β= $A \cap B'$.

Το ενδεχόμενο Α–Β πραγματοποιείται όταν πραγματοποιείται το Α αλλά όχι το Β, δηλαδή όταν πραγματοποιείται το Α και δεν πραγματοποιείται το Β, δηλαδή όταν πραγματοποιείται το Α και Β΄, δηλαδή όταν πραγματοποιείται το $A \cap B'$ (το ίδιο αποτέλεσμα μπορεί να αποδειχθεί και με τήρηση της θεωρίας συνόλων).

ΕΞΑΣΚΗΣΗ

1. Αντιστοίχισε τα στοιχεία της πρώτης στήλης με τα στοιχεία της δεύτερης.

Α	Β
$A \cap \varnothing$ •	• $\varnothing$
$A \cup \varnothing$ •	• A
$A \cap \Omega$ •	• Ω
$A \cup \Omega$ •	• A'
$\Omega - A$ •	

2. Αν Α = {α,β,γ} και Β = {β,δ,ε} να βρείτε τα $A \cap B$, $A \cup B$, Α–Β, Β–Α.

3. Σε κάθε μία από τις παρακάτω περιπτώσεις εξετάστε αν τα ενδεχόμενα είναι ασυμβίβαστα:
   ι) Ρίχνουμε ένα ζάρι. Α Το ενδεχόμενο να φέρουμε το πολύ 3 και Β το ενδεχόμενο να φέρουμε ζυγό αποτέλεσμα.
   ιι) Επιλέγουμε τυχαία ένα άτομο. Α το ενδεχόμενο να έχει γεννηθεί στην Ελλάδα και Β το ενδεχόμενο να έχει ξένη υπηκοότητα.
   ιιι) Επιλέγουμε τυχαία ένα άτομο. Α το ενδεχόμενο να είναι ανήλικος και Β το ενδεχόμενο να ψηφίζει στις εκλογές.
   ιv) Επιλέγουμε ένα φύλλο από μία τράπουλα. Α το ενδεχόμενο να είναι σπαθί και Β το ενδεχόμενο να είναι άσσος.
   v) Επιλέγουμε τυχαία ένα μοντέλο τηλεόρασης. Α το ενδεχόμενο να κατασκευάστηκε στην Ευρώπη και Β το ενδεχόμενο να κατασκευάστηκε στην Κίνα.

4. Από μία συνήθη τράπουλα 52 φύλλων, επιλέγουμε ένα. Α το ενδεχόμενο να είναι κόκκινο και Β το ενδεχόμενο να είναι φιγούρα. Να περιγράψετε λεκτικά και να βρείτε το πλήθος των στοιχείων των ενδεχομένων:
$$A \cap B,\ A \cup B,\ A',\ B',\ A' \cap B,\ A' \cup B,\ A' \cap B',\ A' \cup B' .$$

5. Σε ένα γραφείο εργάζονται 7 άτομα που φοράνε γυαλιά και 5 που δεν φοράνε. Από αυτούς που φοράνε γυαλιά, οι 3 είναι άνδρες και από αυτούς που δεν φοράνε, οι 3 είναι γυναίκες. Επιλέγουμε ένα άτομο τυχαία. Να βρείτε το πλήθος των στοιχείων των ενδεχομένων:
   α) να είναι άνδρας και να μην φοράει γυαλιά,
   β) να είναι γυναίκα και να φοράει γυαλιά.

6. Αν Α,Β,Γ τρία ενδεχόμενα ενός δειγματικού χώρου Ω, να παραστήσετε με διάγραμμα και να περιγράψετε με τύπο τα ενδεχόμενα:
   α) να πραγματοποιηθούν και τα τρία,
   β) να πραγματοποιηθούν ακριβώς δύο από αυτά.

7. Να δείξετε ότι:
   α) $A\cap(B\cap\Gamma)'=(A\cap B')\cup(A\cap\Gamma')$,
   β) $(A'\cup B')'=A\cap B$.

8. Να δείξετε ότι:
   α) $N(A')=N(\Omega)-N(A)$,
   β) $N(A-B)=N(A)-N(A\cap B)$

7.3. Η ΕΝΝΟΙΑ ΤΗΣ ΠΙΘΑΝΟΤΗΤΑΣ

Για να μελετήσουμε ένα ενδεχόμενο Α μπορούμε να εκτελέσουμε το πείραμα τύχης ν φορές και να καταγράψουμε τη συχνότητα εμφάνισης του ενδεχομένου Α. Ας υποθέσουμε ότι ρίχνουμε ένα ζάρι και μας ενδιαφέρει το ενδεχόμενο Α, το αποτέλεσμα να είναι 2.

Εκτελούμε το πείραμα ν φορές και συμβολίζουμε με ν_A τη συχνότητα εμφάνισης του ενδεχομένου Α όπως φαίνεται στον ακόλουθο πίνακα:

ΠΙΝΑΚΑΣ ΣΥΧΝΟΤΗΤΩΝ		
ν	ν_A	f_A
10	2	0,2
20	3	0,15
30	7	0,233
50	8	0,16
100	17	0,17
150	25	0,166
200	33	0,165
500	84	0,168

Παρατηρούμε ότι η σχετική συχνότητα $f_A = \frac{\nu_A}{\nu}$ του ενδεχομένου δεν είναι σταθερή όταν το ν μεταβάλλεται.

Όμως, όσο το ν μεγαλώνει, παρατηρούμε ότι το τείνει να σταθεροποιηθεί γύρω από μία σταθερή τιμή. Ο κανόνας αυτός λέγεται **νόμος των μεγάλων αριθμών**. Η τιμή προς την οποία τείνει η σχετική συχνότητα όταν το ν αυξάνει, μας δείχνει πόσες φορές πρέπει να αναμένουμε λογικά την πραγματοποίηση του ενδεχομένου Α, δηλαδή πόσο «πιθανό» είναι να πραγματοποιηθεί το ενδεχόμενο Α. Την τιμή αυτή του f_A την ονομάζουμε **στατιστική πιθανότητα** του ενδεχομένου Α και είναι πάντα θετική και μικρότερη της μονάδας.

Στηριζόμενοι στα προηγούμενα συμπεράσματα, μπορούμε να δώσουμε τον παρακάτω **αξιωματικό ορισμό της πιθανότητας** που οφείλεται στο Ρώσο μαθηματικό Colmogorov και έχει ιδιότητες ανάλογες με τη σχετική συχνότητα.

Αν $\Omega = \{\omega_1, \omega_2, \dots \omega\nu\}$ είναι ένας δειγματικός χώρος, σε κάθε απλό ενδεχόμενο $\{\omega\kappa\}$ αντιστοιχίζουμε έναν αριθμό που τον συμβολίζουμε με $P(\omega\kappa)$ και τον ονομάζουμε πιθανότητα του ενδεχομένου $\{\omega\kappa\}$, έτσι ώστε να ισχύουν:

- $P(\omega_\kappa) \geq 0$
- $P(\omega_1)+P(\omega_2)+\dots+P(\omega_\nu)=1$

Ως πιθανότητα του ενδεχομένου $A = \{\omega_1, \omega_2, \dots \omega_\kappa\}$ ορίζουμε το άθροισμα $P(\omega_1)+P(\omega_2)+\dots+P(\omega_\kappa)$και ως πιθανότητα του αδύνατου ενδεχομένου ορίζουμε $P(\varnothing)=0$.

Από τον ορισμό αυτό, είναι φανερό ότι $0 \leq P(A) \leq 1$ για οποιοδήποτε ενδεχόμενο Α και $P(\Omega) = 1$.

Ας ξαναγυρίσουμε τώρα στο πείραμα με το ρίξιμο του ζαριού. Ο δειγματικός χώρος είναι $\Omega=\{1,2,3,4,5,6\}$. Αν το ζάρι δεν είναι «πειραγμένο», τότε θα πρέπει όλα τα αποτελέσματα να έχουν την ίδια πιθανότητα, να είναι όπως λέμε **ισοπίθανα**. Σύμφωνα λοιπόν με τον παραπάνω ορισμό:

$P(1)+P(2)+P(3)+P(4)+P(5)+P(6)=1 \Leftrightarrow P(2)= \frac{1}{6}$ και βεβαίως:

P(1)=P(2)=P(3)=P(4)=P(5)=P(6)= $\frac{1}{6}$. Όπως βλέπουμε, η τιμή αυτή $\frac{1}{6}$ =0.1666 είναι η τιμή προς την οποία τείνει το f_A στις δοκιμές του πειράματος που κάναμε. Αν θέλουμε να υπολογίσουμε την πιθανότητα του ενδεχομένου Β, το αποτέλεσμα να είναι ζυγό, τότε έχουμε Β={2,4,6} και P(B)=P(2)+P(4)+P(6)= $\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{3}{6}$, δηλαδή είναι ίση με το πηλίκο του πλήθους των στοιχείων του Β προς το πλήθος των στοιχείων του δειγματικού χώρου Ω.

Με τον ίδιο τρόπο, μπορούμε να δουλέψουμε αν $\Omega=\{\omega_1, \omega_2, \dots \omega_\nu\}$. Τότε, με την προϋπόθεση ότι τα απλά ενδεχόμενα είναι ισοπίθανα, η πιθανότητα ν
καθενός είναι $\frac{1}{\nu}$ και αν $A=\{\omega_1,\omega_2,\dots\omega_\kappa\}$ έχουμε P(A)= $\frac{\kappa}{\nu}=\frac{N(A)}{N(\Omega)}$.

Έτσι καταλήγουμε στον επόμενο **κλασικό ορισμό της πιθανότητας.**

Αν ο δειγματικός χώρος Ω αποτελείται από ισοπίθανα απλά ενδεχόμενα, τότε η πιθανότητα ενός ενδεχομένου Α είναι ίση με:

$$P(A)=\frac{\text{πλήθος ευνοϊκών περιπτώσεων}}{\text{πλήθος δυνατών περιπτώσεων}}=\frac{N(A)}{N(\Omega)}$$

Ο ορισμός αυτός λέγεται κλασικός, γιατί ιστορικά είναι ο πρώτος που δόθηκε από τον Laplace. Εύκολα μπορεί να διαπιστώσει κανείς, ότι ικανοποιεί όλες τις προϋποθέσεις του αξιωματικού ορισμού. Οι πληθάριθμοι του δειγματικού χώρου Ω και του ενδεχομένου Α συνήθως υπολογίζονται με τους κανόνες της συνδυαστικής.

Εφαρμογή 7

Αν ρίξουμε ένα ζάρι δύο φορές, ποια είναι η πιθανότητα να εμφανιστεί το ίδιο νούμερο;

Τα στοιχεία του δειγματικού χώρου είναι διατάξεις με επανάληψη και το πλήθος τους είναι $[6]_2=6^2=36$. Το ενδεχόμενο που μας ενδιαφέρει είναι Α={11,22,33,44,55,66} με N(A)=6.

Άρα, $P(A)=\frac{N(A)}{N(\Omega)}=\frac{6}{36}=\frac{1}{6}$.

Εφαρμογή 8

Μέσα σε ένα κουτί έχουμε 6 άσπρες και 4 μαύρες μπάλες.

Α) αν βγάλουμε μία μπάλα, ποια είναι η πιθανότητα να είναι άσπρη;

Β) αν βγάλουμε δύο μπάλες μαζί, ποια είναι η πιθανότητα να είναι άσπρες;

Γ) αν βγάλουμε δύο μπάλες μαζί, ποια είναι η πιθανότητα να είναι μία άσπρη και μία μαύρη;

ΛΥΣΗ

Α) Στην πρώτη περίπτωση, το πλήθος των στοιχείων του Ω είναι 10 και το πλήθος των ευνοϊκών περιπτώσεων είναι 6, άρα: $P(A)=\frac{6}{10}=\frac{3}{5}$

Β) Στη δεύτερη περίπτωση, το πλήθος των στοιχείων του Ω είναι $\binom{10}{2}=\frac{10!}{2!8!}=\frac{10\cdot 9}{2}=45$.

Το πλήθος των στοιχείων του ενδεχομένου Β είναι $\binom{6}{2}=\frac{6!}{2!4!}=\frac{6\cdot 5}{2}=15$.
Έτσι, $P(B)=\frac{15}{45}=\frac{1}{3}$.

Γ) Στην τρίτη περίπτωση, μπορούμε να επιλέξουμε την άσπρη μπάλα με $\binom{6}{1}=6$ τρόπους και τη μαύρη με $\binom{4}{1}=4$ τρόπους.

Το πλήθος των στοιχείων του ενδεχομένου Γ θα είναι: $\binom{6}{1}\binom{4}{1}=6\cdot 4=24$ και θα έχουμε $P(\Gamma)=\frac{24}{45}=\frac{8}{15}$.

ΕΞΑΣΚΗΣΗ

1. Η πιθανότητα να φέρουμε αριθμό μικρότερο του 5 ρίχνοντας ένα ζάρι είναι:

Α. $\frac{1}{6}$ Β. $\frac{1}{5}$ Γ. $\frac{1}{3}$ Δ. $\frac{5}{6}$ Ε. 1

2. Η πιθανότητα να βγάλουμε μία κόκκινη μπάλα από ένα κουτί που έχει 3 κόκκινες και 5 πράσινες μπάλες, είναι:

Α. $\frac{1}{3}$ Β. $\frac{1}{5}$ Γ. $\frac{3}{5}$ Δ. $\frac{5}{8}$ Ε. $\frac{3}{8}$

3. Η πιθανότητα να εμφανιστεί τουλάχιστον μία φορά «κεφαλή »αν ρίξουμε ένα νόμισμα δύο φορές, είναι:

Α. $\frac{1}{2}$ Β. $\frac{3}{4}$ Γ. $\frac{1}{4}$ Δ. 1 Ε. 0

4. Ένα άτομο παίζει ρουλέτα και ποντάρει στο κόκκινο. Ποια είναι η πιθανότητα να κερδίσει; (Η ρουλέτα έχει αριθμούς από 1 μέχρι 36 μοιρασμένους ίσα σε κόκκινους και μαύρους και τον αριθμό 0 που δεν έχει χρώμα.

Α. $\frac{18}{37}$ Β. $\frac{1}{2}$ Γ. $\frac{1}{18}$ Δ. $\frac{18}{19}$ Ε. $\frac{1}{36}$

5. Φτιάχνουμε τετραψήφιους αριθμούς επιλέγοντας τυχαία ψηφία από τα 1,2,3,4,5. Ποια είναι η πιθανότητα ένας τέτοιος αριθμός να έχει όλα τα ψηφία του διαφορετικά;

6. Από μία ομάδα 3 ανδρών και 5 γυναικών, επιλέγουμε τυχαία μία πενταμελή αντιπροσωπεία. Ποια είναι η πιθανότητα να αποτελείται:
α) από 2 άνδρες και 3 γυναίκες,
β) μόνο από γυναίκες,
γ) μόνο από άνδρες.

7. Επιλέγουμε τυχαία δύο αριθμούς από το 1 μέχρι το 20. Ποια είναι η πιθανότητα:
α) το άθροισμά τους να είναι μονός αριθμός,
β) το γινόμενό τους να είναι μονός αριθμός.

8. Αν παίξουμε 7 αριθμούς σε ένα δελτίο LOTTO, ποια είναι η πιθανότητα να πιάσουμε:
α) εξάρι β) πεντάρι γ) τεσσάρι

9. Μία ηλεκτρονική κλειδαριά λειτουργεί με ένα τριψήφιο κωδικό. Αν δοθεί ο σωστός κωδικός ανοίγει, αλλά αν κανένα ψηφίο δεν είναι σωστό, τότε χτυπάει ο συναγερμός:
α) Ο ιδιοκτήτης θυμάται τα 3 ψηφία του κωδικού, αλλά έχει ξεχάσει τη σωστή σειρά. Αν επιλέξει τυχαία ένα κωδικό, ποια είναι η πιθανότητα να ανοίξει η κλειδαριά και ποια είναι η πιθανότητα να χτυπήσει ο συναγερμός.
β) Ένα διαρρήκτης που δεν ξέρει τίποτα για τον κωδικό, επιλέγει έναν τυχαία. Ποια είναι η πιθανότητα να χτυπήσει ο συναγερμός;

10. Δίνεται ο δειγματικός χώρος $\Omega = \{\omega_1,\omega_2,\omega_3,\omega_4\}$ και τα ενδεχόμενα
$A = \{\omega_2,\omega_3\}$ και $B = \{\omega_2,\omega_4\}$. Αν $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ και $P(\omega_3) = \frac{1}{4}$, να βρείτε το $P(\omega_1)$.

11. Δίνεται ο δειγματικός χώρος $\Omega = \{\omega_1,\omega_2,\omega_3,\}$. Αν το ενδεχόμενο ω_2 έχει διπλάσια πιθανότητα από το ω1 και το ω3 έχει πιθανότητα ίση με το άθροισμα των πιθανοτήτων των δύο άλλων, να βρείτε την πιθανότητα του καθενός.

12. Σε ένα κουτί υπάρχουν 6 κόκκινες και άγνωστος αριθμός πράσινες μπάλες. Αν η πιθανότητα να επιλέξουμε μία πράσινη μπάλα είναι $\frac{2}{5}$, να βρείτε πόσες είναι οι πράσινες μπάλες.

13. Ένα κουτί περιέχει 12 άσπρες, x κόκκινες και y πράσινες μπάλες. Αν επιλέξουμε μία τυχαία, η πιθανότητα να είναι κόκκινη είναι $\frac{1}{2}$ και η πιθανότητα να είναι πράσινη είναι $\frac{1}{3}$. Να βρείτε τα x και y.

7.4. ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΠΙΘΑΝΟΤΗΤΩΝ

Για τις πιθανότητες των ενδεχομένων ενός δειγματικού χώρου, ισχύουν οι επόμενες ιδιότητες που είναι γνωστές σαν κανόνες λογισμού των πιθανοτήτων.

1. Αν Α, Β ασυμβίβαστα ενδεχόμενα, τότε:

$$\boxed{P(A \cup B) = P(A) + P(B)}$$

Απόδειξη: Αν Α = $\{\alpha_1, \alpha_2, \ldots \alpha_\nu\}$ και Β = $\{\beta_1, \beta_2, \ldots \beta_\mu\}$ τότε $A \cup B$ $=\{\alpha_1, \alpha_2, \ldots \alpha_\nu, \beta_1, \beta_2, \ldots \beta_\mu\}$. Έτσι, $P(A \cup B) = P(\alpha_1)+P(\alpha_2)+\ldots+P(\alpha_\nu)+P(\beta_1)+P(\beta_2)+\ldots+P(\beta_\mu)=P(A)+P(B)$. Ο κανόνας επεκτείνεται ....

2.Για τα συμπληρωματικά ενδεχόμενα, ισχύει ότι:

$$\boxed{P(A')=1-P(A)}$$

Απόδειξη: Επειδή Α και Α΄ είναι ασυμβίβαστα, έχουμε:

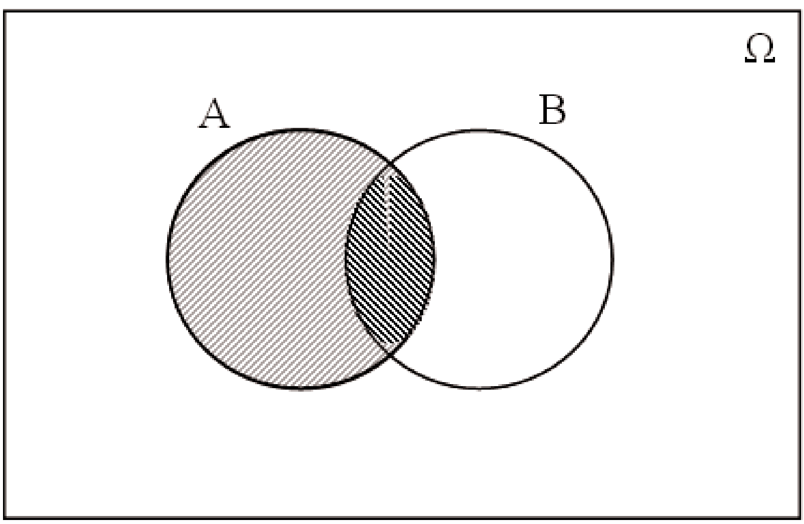
$$P(A \cup A') = P(A) + P(A') \Leftrightarrow P(\Omega) = P(A) + P(A') \Leftrightarrow 1 = P(A) + P(A') \Leftrightarrow P(A') = 1 - P(A)$$

3. Για οποιαδήποτε ενδεχόμενα Α, Β, ισχύει ότι:

$$\boxed{P(A-B)=P(A)-P(A \cap B)}$$

Απόδειξη: Επειδή $A \cup B = (A - B) \cup (A \cap B) \cup (B - A)$ και τα $A - B$, $A \cap B$, $B - A$ είναι ασυμβίβαστα, έχουμε:

$$\begin{aligned} P(A \cup B) &= P(A-B)+P(A \cap B)+P(B-A)= \\ &=P(A)-\cancel{P(A \cap B)}+\cancel{P(A \cap B)}+P(B)-P(A \cap B)= \\ &=P(A) + P(B)-P(A \cap B) \end{aligned}$$



4.Για οποιαδήποτε ενδεχόμενα Α, Β ισχύει ότι:

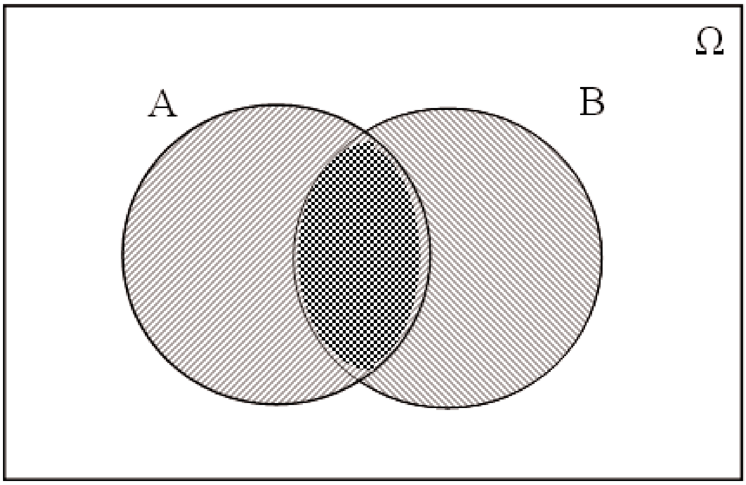
$$\boxed{P(A\cup B)=P(A)+P(B)-P(A\cap B)}$$

Απόδειξη: Επειδή $A \cup B = (A-B) \cup (A \cap B) \cup (B-A)$και $A-B$, $A \cap B$, $B-A$ ασυμβίβαστα ανά δύο, έχουμε:

$P(A \cup B) = P(A-B)+P(A\cap B)+P(B-A)=$

$=P(A)-P(A \cap B)+P(A \cap B)+P(B)-P(A \cap B)=$

$=P(A) + (B)-P(A \cap B)$

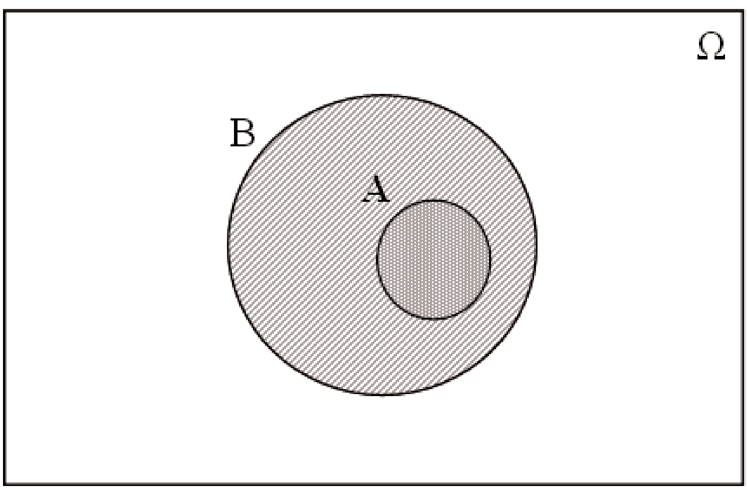


Ο κανόνας αυτός είναι πιο γενικός από τον κανόνα 1, αφού ισχύει για οποιαδήποτε ενδεχόμενα Α, Β και είναι γνωστός σαν **προσθετικός νόμος των πιθανοτήτων.**

5. Αν $A\subseteq B$, τότε $P(A) \leq P(B)$.

Απόδειξη: $P(B-A)=P(B)-P(A\cap B) \Leftrightarrow P(B-A)=P(B)-P(A) \Leftrightarrow P(A)+P(B-A)=P(B)$

Άρα $P(A) \leq P(B)$.



Εφαρμογή 9

Για δύο ενδεχόμενα Α και Β ισχύει ότι $P(A)=0,6$, $P(B)=0,3$ και $P(A \cap B)=0,2$. Να βρείτε την πιθανότητα:

α) να πραγματοποιηθεί το Α ή το Β,

β) να πραγματοποιηθεί μόνο ένα από τα Α και Β

γ) να μην πραγματοποιηθεί κανένα από τα Α και Β

ΛΥΣΗ

α) Το ζητούμενο ενδεχόμενο είναι το $A \cup B$ και έχουμε:

$$P(A \cup B)=P(A)+P(B)-P(A \cap B)=0,6+0,3-0,2=0,7$$

β) Το ζητούμενο ενδεχόμενο είναι το $(A-B) \cup (B-A)$ και έχουμε:

$$P\big((A-B) \cup (B-A)\big)=P(A-B)+P(B-A)=P(A)-P(A \cap B)+P(B)-P(A \cap B)=$$
$$=P(A)+P(B)-2P(A \cap B)=0,6+0,3-2\cdot 0,2=0,5$$

γ) Το ζητούμενο ενδεχόμενο είναι το $(A \cup B)'$ και έχουμε:

$$P\big((A \cup B)'\big)=1-P(A \cup B)=1-0,7=0,3$$

Εφαρμογή 10

Ένα εργοστάσιο κατασκευάζει 50 εξαρτήματα την ώρα, από τα οποία, 2 είναι ελαττωματικά. Ένας ελεγκτής ελέγχει τυχαία τρία εξαρτήματα ανά ώρα. Ποια η πιθανότητα να βρει ένα τουλάχιστον ελαττωματικό;

ΛΥΣΗ

Σε αρκετές περιπτώσεις είναι πιο εύκολος ο υπολογισμός της πιθανότητας του συμπληρωματικού ενδεχομένου, ιδιαίτερα αν το αρχικό ενδεχόμενο χωρίζεται σε μικρότερα υποσύνολα. Εδώ για παράδειγμα, το ενδεχόμενο Α: βρίσκει τουλάχιστον ένα ελαττωματικό, χωρίζεται σε A^1: βρίσκει ένα ελαττωματικό και A^2: βρίσκει δύο ελαττωματικά. Είναι ευκολότερο να υπολογίσουμε την πιθανότητα του A': δεν βρίσκει κανένα ελαττωματικό.

Το πλήθος των στοιχείων του Ω είναι $\binom{50}{3}=\frac{50\cdot 49\cdot 48}{3\cdot 2}=19.600$.

Το πλήθος των στοιχείων του A' είναι $\binom{48}{3}=\frac{48\cdot 47\cdot 46}{3\cdot 2}=17.296$, γιατί τα μη ελαττωματικά εξαρτήματα είναι 48 και από αυτά, πρέπει να επιλέξει τα τρία, ώστε να μην διαπιστώσει πρόβλημα. Έτσι έχουμε:

$$P(A')=\frac{17.296}{19.600}\approx 0,88$$ και $P(A)=1-P(A')=1-0,88=0,12$.

Εφαρμογή 11

Για δύο ενδεχόμενα ενός δειγματικού χώρου Ω ισχύει $P(A)=0,4$ και $P(B)=0,8$.

α) Να εξετάσετε αν τα Α και Β είναι ασυμβίβαστα

β) Να αποδείξετε ότι $0,2 \le P(A \cap B) \le 0,4$

ΛΥΣΗ

α) Αν Α, Β ήταν ασυμβίβαστα, τότε, $P(A \cap B)=P(A)+P(B)=0{,}4+0{,}8=1{,}2$ άτοπο, γιατί $P(A \cup B) \leq 1$. Άρα, τα Α, Β δεν είναι ασυμβίβαστα.

β) Επειδή $A \cap B \subseteq A$ έχουμε $P(A \cap B) \leq P(A) \Leftrightarrow P(A \cap B) \leq 0{,}4$. Επίσης, $P(A \cup B) \leq 1 \Leftrightarrow$ $P(A)+P(B)-P(A \cap B) \leq 1 \Leftrightarrow P(A \cap B) \geq P(A)+P(B)-1$ $\Leftrightarrow P(A \cap B) \geq 0{,}4+0{,}8-1 \Leftrightarrow P(A \cap B) \geq 0{,}2$

ΕΞΑΣΚΗΣΗ

1. Χαρακτηρίστε τις παρακάτω προτάσεις σαν Σωστές (Σ) ή Λανθασμένες (Λ):

	Σωστό	Λάθος
α. $P(A)=1-P(A')$	☐	☐
β. $P(A \cup B)=P(A)+P(B)$	☐	☐
γ. $P(A')=-P(A)$	☐	☐
δ. $P(B-A)=P(B)-P(A \cap B)$	☐	☐
ε. $P(A-B)+P(B-A)=P(A \cup B)-P(A \cap B)$	☐	☐

2. Σε ένα δειγματικό χώρο Ω, ισχύουν $P(A')=\frac{3}{4}$, $P(A \cup B)=\frac{4}{5}$, και $P(A \cap B)=\frac{1}{5}$.

Να βρείτε τις P(A), P(B), $P(A \cap B')$, $P(A \cup B')$.

3. Σε ένα δειγματικό χώρο Ω ισχύουν $P(A \cup B)=0{,}8$, $P(A \cap B)=0{,}2$ και P(A)=P(B). Να βρείτε τις P(A), $P(A' \cap B')$, $P(B\text{-}A)$, $P(A\text{-}B')$.

4. Σε ένα δειγματικό χώρο Ω ισχύει $\frac{P(A)}{P(A')}=\frac{2}{3}$. Να βρείτε την P(A).

5. Σε μία έρευνα βρέθηκε ότι το 90% των σπιτιών έχουν τηλεόραση, το 95% έχουν ραδιόφωνο και το 88% έχουν και τα δύο. Αν επιλέξουμε τυχαία ένα σπίτι, ποια είναι η πιθανότητα:
α. να έχει τηλεόραση ή ραδιόφωνο
β. να έχει μόνο τηλεόραση
γ. να μην έχει τίποτα από τα δύο.

6. Σε έναν αγώνα, η πιθανότητα να κερδίσει ο Γιάννης είναι 50%, ενώ η πιθανότητα να μην κερδίσει ο Κώστας είναι 70%. Ποια είναι η πιθανότητα:
α. να κερδίσει ο Γιάννης ή ο Κώστας,
β. να μην κερδίσει κανένας από τους δύο.

7. Σε ένα κουτί, έχουμε 5 άσπρα και 4 μαύρα μπαλάκια. Επιλέγουμε ταυτόχρονα τέσσερα μπαλάκια. Να υπολογίσετε την πιθανότητα να επιλέξουμε μπαλάκια και από τα δύο χρώματα.

8. Σε μία ομάδα 10 ατόμων, ποια είναι η πιθανότητα να βρεθούν τουλάχιστον δύο άτομα που να έχουν γενέθλια την ίδια μέρα;

9. Μία παρέα 8 ατόμων, κάθονται τυχαία στις 8 θέσεις μίας σειράς στο θέατρο. Ποια η πιθανότητα ο Γιώργος να καθίσει δίπλα στη Μαρία;

10. Ο υπάλληλος του γκαράζ μπερδεύει τα κλειδιά τριών αυτοκινήτων και τα δίνει στους κατόχους των αυτοκινήτων τυχαία. Να βρείτε τις πιθανότητες:
   α. κάθε οδηγός να πάρει το δικό τους κλειδί,
   β. μόνο ένας οδηγός να πάρει το δικό τους κλειδί,
   γ. κανένας να μην πάρει το δικό του κλειδί.

11. Σε ένα δειγματικό χώρο Ω ισχύουν P(A)=0,4 και P(B)=0,5. Δείξτε ότι $0,5 \le P(A \cup B) \le 0,9$.

12. Σε ένα δειγματικό χώρο Ω ισχύουν P(A) = P(B) = 0,6. Δείξτε ότι.

13. Αν Α, Β ενδεχόμενα ενός δειγματικού χώρου Ω, δείξτε ότι: $0,2 \le P(A \cap B) \le 0,6$

   α. $P(A \cup B) \le P(A) + P(B)$

   β. $P(A) - P(B') \le P(A \cap B)$

14. Αν Α, Β, Γ ενδεχόμενα ενός δειγματικού χώρου Ω, δείξτε ότι:
$$P(A \cup B \cup \Gamma) = P(A) + P(B) + P(\Gamma) - P(A \cap B) - P(A \cap \Gamma) - P(B \cap \Gamma) + P(A \cap B \cap \Gamma)$$

7.5. ΔΕΣΜΕΥΜΕΝΗ ΠΙΘΑΝΟΤΗΤΑ

Η πιθανότητα να πραγματοποιηθεί ένα ενδεχόμενο Β όταν είναι γνωστό ότι έχει πραγματοποιηθεί ένα άλλο ενδεχόμενο $A \neq \varnothing$, ονομάζεται **δεσμευμένη πιθανότητα** του Β δεδομένου του Α και συμβολίζεται με P(B\A). Αποδεικνύεται ότι P(B\A) $= \dfrac{P(A \cap B)}{P(A)}$.

Με τον ίδιο τρόπο, έχουμε $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$ και καταλήγουμε ότι:

$$\boxed{P(A \cap B) = P(A) \cdot P(B \setminus A) = P(B) \cdot P(A \setminus B)}$$

Η σχέση αυτή ονομάζεται **πολλαπλασιαστικός** νόμος των πιθανοτήτων και μπορεί να γενικευτεί:

$$\boxed{P(A_1 \cap A_2 \cap \ldots \cap A_\nu) = P(A_1)P(A_2 / A_1)P(A_3 / A_1 \cap A_2)\ldots P(A_\nu / A_1 \cap A_2 \ldots A_\nu)}$$

Δύο ενδεχόμενα A και B λέγονται **ανεξάρτητα** αν και μόνο αν,

$$\boxed{P(A \cap B) = P(A) \cdot P(B)}$$

Στην περίπτωση αυτή, αν A,B $\neq \varnothing$ από τον πολλαπλασιαστικό νόμο έχουμε ότι P(A/B)=P(A) και P(B/A)=P(B), δηλαδή η πραγματοποίηση του ενός δεν εξαρτάται από την πραγματοποίηση του άλλου.

Για παράδειγμα, ας υποθέσουμε ότι έχουμε μία οικογένεια με 3 παιδιά. Ο δειγματικός χώρος είναι Ω = {ααα, αακ, ακα, ακκ, καα, κακ, κκα, κκκ}. Να βρεθεί η πιθανότητα η οικογένεια να έχει ένα μόνο κορίτσι, δεδομένου ότι το πρώτο παιδί είναι αγόρι.

Συμβολίζουμε A: το πρώτο παιδί είναι αγόρι, δηλαδή A={ααα, αακ, ακα, ακκ}και B: έχει ακριβώς ένα κορίτσι, δηλαδή B={αακ, ακα, καα}. Τότε έχουμε $A \cap B =$ {ακα, αακ} και η ζητούμενη πιθανότητα είναι:

$$P(B \setminus A) = \frac{P(A \cap B)}{P(A)} = \frac{2/8}{4/8} = \frac{2}{4} = \frac{1}{2}$$

Στο ίδιο παράδειγμα, αν Γ: τα δύο τελευταία παιδιά είναι κορίτσια, να εξετάσετε αν τα ενδεχόμενα A και Γ είναι ανεξάρτητα. Έχουμε Γ={ακκ, κκκ} και $A \cap \Gamma$ ={ακκ}. Έτσι, $P(A) \cdot P(\Gamma) = \frac{4}{8} \cdot \frac{2}{8} = \frac{1}{8} = P(A \cap \Gamma)$ άρα τα A και Γ είναι ανεξάρτητα.

Εφαρμογή 12

Αν A, B ασυμβίβαστα και μη κενά, τότε A, B είναι εξαρτημένα.
Πράγματι, $A \cap B = \varnothing$ άρα $P(A \cap B) = 0$. Όμως, $P(A) \cdot P(B) \neq \varnothing$, άρα, $P(A \cap B) \neq P(A) \cdot P(B)$ δηλαδή τα A, B είναι εξαρτημένα.

Εφαρμογή 13

Σε ένα κουτί, υπάρχουν 6 άσπρα και 4 μαύρα μπαλάκια. Βγάζουμε διαδοχικά δύο μπαλάκια. Ποια είναι η πιθανότητα το πρώτο να είναι μαύρο και το δεύτερο άσπρο, αν:

α. ξαναβάζουμε μέσα το πρώτο μπαλάκι

β. δεν ξαναβάζουμε μέσα το πρώτο μπαλάκι.

α. Στην πρώτη περίπτωση, είναι φανερό ότι τα ενδεχόμενα A: πρώτο μπαλάκι μαύρο και B: δεύτερο μπαλάκι άσπρο, είναι ανεξάρτητα, γιατί η πραγματοποίηση του ενός δεν επηρεάζει την πραγματοποίηση του άλλου. Έτσι, έχουμε:

$$P(A \cap B) = P(A) \cdot P(B) = \frac{4}{10} \cdot \frac{6}{10} = \frac{24}{100} = \frac{6}{25}$$

β. Στη δεύτερη περίπτωση, τα δύο ενδεχόμενα δεν είναι ανεξάρτητα, γιατί το χρώμα που θα έχει το πρώτο μπαλάκι επηρεάζει την αναλογία των χρωμάτων των υπολοίπων. Έτσι έχουμε:

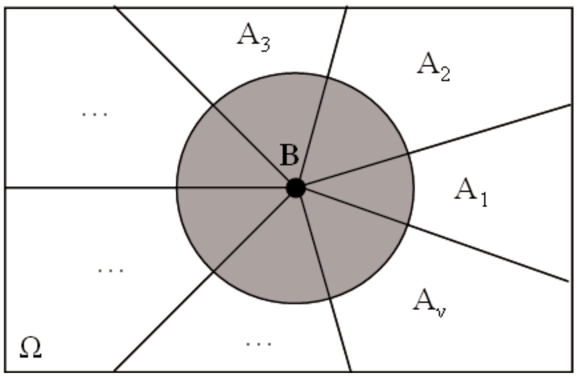
$$P(A \cap B) = P(A) \cdot P(B \setminus A) = \frac{4}{10} \cdot \frac{6}{9} = \frac{24}{90} = \frac{4}{15}$$

Αν ο δειγματικός χώρος Ω μπορεί να αναλυθεί σε ν το πλήθος ενδεχόμενα, που ανά δύο είναι ασυμβίβαστα και $A_1 \cup A_2 \cup \ldots \cup A_\nu = \Omega$, τότε ένα οποιοδήποτε ενδεχόμενο Β μπορεί να γραφτεί

$B = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_\nu)$ και συνεπώς:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_\nu) = P(A_1) + P(B \setminus A_1) + P(A_2) + P(B \setminus A_2) + \ldots + P(A_\nu) + P(B \setminus A_\nu)$$

Ο τύπος αυτός είναι γνωστός σαν τύπος της ολικής πιθανότητας. Συχνά παίρνουμε A1=A και A2=A΄, οπότε P(B)=P(A)P(B/A)+P(A΄)P(B/A΄).



Εφαρμογή 14

Ένα εργοστάσιο έχει δύο γραμμές παραγωγής. Η γραμμή Ι παράγει το 60% των συνολικών προϊόντων, αλλά ένα 5% της παραγωγής της είναι προβληματικό. Η γραμμή ΙΙ παράγει το 40% των συνολικών προϊόντων και μόνο 2% της παραγωγής της είναι προβληματικό. Αν επιλέξουμε τυχαία ένα προϊόν του εργοστασίου, ποια είναι η πιθανότητα να είναι προβληματικό;

ΛΥΣΗ

Έστω Β: προϊόν προβληματικό, Α1: το προϊόν είναι από τη γραμμή Ι και Α2: το προϊόν είναι από τη γραμμή ΙΙ. Τότε $A_1 \cap A_2 = \varnothing$ και $\boxed{A_1 \cup A_2 = B}$, άρα μπορούμε να χρησιμοποιήσουμε τον τύπο της ολικής πιθανότητας και έχουμε:

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) = \frac{60}{100} \cdot \frac{5}{100} + \frac{40}{100} \cdot \frac{2}{100} = \frac{30}{1000} + \frac{8}{1000} = \frac{38}{1000} = 3{,}8\%$$

Αν Α, Β δύο ενδεχόμενα του δειγματικού χώρου Ω, τότε:

$$\boxed{P(B \setminus A) = \frac{P(A \cup B)}{P(A)} = \frac{P(B) \cdot P(A/B)}{P(A)}}$$

Ο τύπος αυτός είναι γνωστός σαν **τύπος του Bayes**.

Εφαρμογή 15

Ένας μαθητής απαντάει σε ερωτήσεις πολλαπλής επιλογής που έχουν 4 απαντήσεις η κάθε μία. Ο μαθητής είναι αρκετά καλά διαβασμένος και γνωρίζει το 70% των ερωτήσεων. Σε όσες δεν γνωρίζει απαντάει τυχαία:

α. ποιο είναι το ποσοστό των σωστών απαντήσεων που δίνει ο μαθητής,

β. αν ο μαθητής έχει απαντήσει σωστά σε μία ερώτηση, ποια είναι η πιθανότητα να γνωρίζει την απάντηση.

ΛΥΣΗ

α. Η έκφραση «Ποιο είναι το ποσοστό των σωστών απαντήσεων», ισοδυναμεί με την «Ποια είναι η πιθανότητα να απαντήσει σωστά σε μία ερώτηση»(στατιστικός ορισμός της πιθανότητας). Έστω λοιπόν Σ το ενδεχόμενο να απαντήσει σωστά και Γ το ενδεχόμενο να γνωρίζει την απάντηση. Έχουμε Ρ(Γ) = 70%, Ρ(Σ/Γ) = 100% (αν γνωρίζει την απάντηση προφανώς επιλέγει σωστά) και Ρ(Σ/Γ΄)=25% (αφού αν δεν γνωρίζει την απάντηση επιλέγει τυχαία μία από τις τέσσερις που προσφέρονται). Έτσι, σύμφωνα με τον τύπο της ολικής πιθανότητας

$$P(\Sigma) = P(\Gamma)\cdot P(\Sigma/\Gamma) + P(\Gamma')P(\Sigma/\Gamma') = \frac{70}{100}\cdot\frac{100}{100} + \frac{30}{100}\cdot\frac{25}{100} = \frac{775}{1000} = 77,5\%$$

β. Από τον τύπο του Bayes έχουμε:

$$P(\Gamma/\Sigma) = \frac{P(\Sigma/\Gamma)P(\Gamma)}{P(\Sigma)} = \frac{\frac{100}{100}\cdot\frac{70}{100}}{\frac{775}{100}} = \frac{700}{775} \approx 0,9 \text{ ή } 90\%$$

Όταν ένα πείραμα τύχης εκτελείται σε διαδοχικά βήματα και τα αποτελέσματα καθενός δεν είναι ισοπίθανα, τότε ένας εύκολος τρόπος υπολογισμού της πιθανότητας κάθε ενδεχομένου είναι η κατασκευή ενός δενδροδιαγράμματος, η αναγραφή της πιθανότητας σε κάθε «κλαδί»και η χρήση του πολλαπλασιαστικού νόμου.

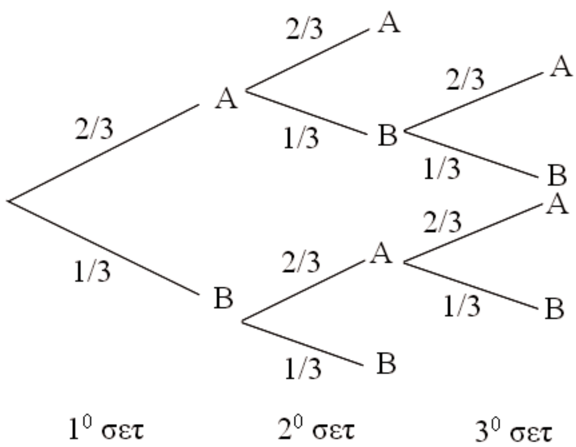
Εφαρμογή 16

Σε έναν αγώνα τένις κερδίζει όποιος κερδίσει δύο σετ. Ο παίκτης Α έχει πιθανότητα $\frac{2}{3}$ να κερδίσει ένα σετ από τον παίκτη Β. Να βρείτε την πιθανότητα:

α. ο παίκτης Α να κερδίσει τον αγώνα,

β. να παιχτούν μόνο δύο σετ,

γ. ο παίκτης Β να κερδίσει πρώτο σετ, δεδομένου ότι κέρδισε τον αγώνα.



$$P(AA)=\frac{2}{3}\cdot\frac{2}{3}=\frac{4}{9}$$

$$P(ABA)=\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{4}{27}$$

$$P(ABB)=\frac{2}{27}$$

$$P(BAA)=\frac{4}{27}$$

$$P(BAB)=\frac{2}{27}$$

$$P(BB)=\frac{1}{9}$$

α. Η πιθανότητα να κερδίσει ο πρώτος παίκτης είναι:

$$P(A)=P(AA)+P(ABA)+P(BAA)=\frac{4}{9}+\frac{4}{27}+\frac{4}{27}=\frac{20}{27}$$

β. Η πιθανότητα να παιχτούν μόνο δύο σετ είναι: $P(\Gamma)=\frac{4}{9}+\frac{1}{9}=\frac{5}{9}$.

γ. Η πιθανότητα να κερδίσει ο δεύτερος παίκτης είναι:

$P(B)=1-P(A)=\frac{7}{27}$. Αν Δ είναι το ενδεχόμενο ο δεύτερος παίκτης να κερδίσει το πρώτο σετ,

τότε $P(B\cap\Delta)=\frac{1}{9}+\frac{2}{27}=\frac{5}{27}$.

Έτσι, $P(\Delta/B)=\dfrac{P(B\cap\Delta)}{P(B)}=\dfrac{5/27}{7/27}=\dfrac{5}{7}$.

ΕΞΑΣΚΗΣΗ

1. Χαρακτηρίστε Σωστές (Σ) ή Λανθασμένες (Λ) τις παρακάτω προτάσεις:

	Σωστό	Λάθος
α. Ισχύει πάντοτε $P(A \cap B) = P(A) \cdot P(B)$	☐	☐
β. Ισχύε $P(A/B) = \dfrac{P(A \cap B)}{P(A)}$	☐	☐
γ. Ισχύει $P(A)P(B) = P(B) \cdot P(A/B)$	☐	☐
δ. Δύο ασυμβίβαστα ενδεχόμενα είναι ανεξάρτητα	☐	☐
ε. Αν Α, Β ασυμβίβαστα τότε P(A/B)=P(B/A)=0	☐	☐

2. Επιλέξτε τη σωστή απάντηση:

α. Αν μία οικογένεια έχει ήδη δύο αγόρια, τότε η πιθανότητα το τρίτο παιδί να είναι αγόρι, είναι:

Α. $\frac{1}{2}$ Β. $\frac{1}{4}$ Γ. $\frac{1}{8}$ Δ. $\frac{3}{8}$ Ε. $\frac{7}{8}$

β. Αν έχουμε ρίξει ένα ζάρι τρεις φορές και δεν έχουμε φέρει 6, τότε η πιθανότητα να φέρουμε 6 την επόμενη ζαριά είναι:

Α. $\frac{1}{1296}$ Β. $\frac{1}{216}$ Γ. $\frac{1}{36}$ Δ. $\frac{1}{6}$ Ε. $\frac{1}{3}$

γ. Ρίχνει κάποιος ένα ζάρι και αναγγέλλει ότι έφερε μονό αριθμό. Ποια είναι η πιθανότητα να είναι 3

Α. $\frac{1}{2}$ Β. $\frac{1}{3}$ Γ. $\frac{1}{6}$ Δ. $\frac{1}{12}$ Ε. $\frac{1}{18}$

δ. Ρίχνουμε ταυτόχρονα ένα κέρμα και ένα ζάρι. Ποια είναι η πιθανότητα το κέρμα να φέρει «γράμματα»και το ζάρι να φέρει 4;

Α. $\frac{1}{3}$ Β. $\frac{2}{3}$ Γ. $\frac{2}{7}$ Δ. $\frac{1}{7}$ Ε. $\frac{1}{12}$

3. Για δύο ενδεχόμενα Α, Β ισχύουν $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{4}$ και $P(A \setminus B) = \frac{4}{5}$.

Να υπολογίσετε τις πιθανότητες $P(A \cap B)$, $P(B/A)$, $P(A \cup B)$.

4. Για δύο ενδεχόμενα Α, Β ισχύουν $P(A \cup B) = \frac{5}{6}$, $P(A/B) = \frac{1}{4}$, $P(B') = \frac{1}{3}$.

Να υπολογίσετε τις πιθανότητες $P(A \cap B)$, $P(B/A)$, $P(A' \cup B)$.

5. Για δύο ανεξάρτητα ενδεχόμενα Α, Β ισχύουν Ρ(Α)=0,3 Ρ(Α Β)=0,12. Να υπολογίσετε τις πιθανότητες Ρ(Β), $P(A \cup B)$, Ρ(Α/Β).

6. Για δύο ενδεχόμενα Α, Β ισχύουν $P(A) = \frac{1}{3}$ και $P(A \setminus B) = \frac{2}{3}$.

   α. Τα ενδεχόμενα Α, Β είναι ανεξάρτητα;
   β. Τα ενδεχόμενα Α, Β είναι ασυμβίβαστα.

7. Ένα δικινητήριο αεροπλάνο μπορεί να πετάξει όταν λειτουργεί τουλάχιστον η μία μηχανή. Οι δύο μηχανές λειτουργούν ανεξάρτητα και η πιθανότητα βλάβης της κάθε μίας είναι 0,02%. Ποια είναι η πιθανότητα το αεροπλάνο να φθάσει με ασφάλεια στον προορισμό του;

8. Σε μία επιχείρηση το 70% των εργαζόμενων είναι άνδρες. Από τους άνδρες καπνίζει το 40% και από τις γυναίκες το 25%. Αν επιλέξουμε τυχαία ένα άτομο που καπνίζει, ποια είναι η πιθανότητα να είναι γυναίκα.

9. Σε ένα κιβώτιο υπάρχουν 20 ασφάλειες από τις οποίες η μία είναι καμένη. Δοκιμάζουμε τις ασφάλειες μία - μία μέχρι να βρούμε την καμένη. Ποια είναι η πιθανότητα να τη βρούμε:
   α. με την πρώτη δοκιμή,
   β. με την πέμπτη δοκιμή,
   γ. το πολύ με πέντε δοκιμές.

10. Μία βιομηχανία κατασκευάζει τηλεοράσεις από τις οποίες 10% είναι ελαττωματικές. Μία τηλεόραση πριν βγει από το εργοστάσιο περνάει από έλεγχο που εντοπίζει το 90% των ελαττωματικών. Ποια είναι η πιθανότητα:
    α. μία τηλεόραση είναι ελαττωματική αλλά περνάει τον έλεγχο,
    β. μία τηλεόραση που φθάνει στο εμπόριο να είναι εντάξει.

11. Ένα κουτί περιέχει 6 κόκκινες και 4 πράσινες σφαίρες. Επιλέγουμε τυχαία μία και την τοποθετούμε σε ένα δεύτερο που έχει ήδη 3 κόκκινες και 6 πράσινες. Τέλος, επιλέγουμε μία σφαίρα από το δεύτερο κουτί. Ποια είναι η πιθανότητα:
    α. να είναι κόκκινη η δεύτερη σφαίρα,
    β. να είναι πράσινη η πρώτη σφαίρα, δεδομένου ότι η δεύτερη ήταν κόκκινη.

12. Γνωρίζουμε ότι το 50% του πληθυσμού πάσχει από μία συγκεκριμένη ασθένεια. Ένα καινούργιο τεστ διάγνωσης έχει πιθανότητα 2% να βγει θετικό, ενώ το άτομο είναι υγιές και 8% να βγει αρνητικό, ενώ το άτομο πάσχει. Ένα άτομο έκανε το τεστ και βγήκε θετικό. Ποια είναι η πιθανότητα να πάσχει πράγματι από τη συγκεκριμένη ασθένεια;

13. Τρεις μαθητές λύνουν μία άσκηση με πιθανότητα $\frac{1}{2}$, $\frac{1}{3}$, $\frac{3}{5}$ αντιστοιχα.

    Παραδίδουν ανώνυμα τη λύση και ο καθηγητής βρίσκει δύο σωστές απαντήσεις. Ποια είναι η πιθανότητα κάθε μαθητή να έχει δώσει τη λανθασμένη απάντηση;

14. Δείξτε ότι δύο ανεξάρτητα μη κενά ενδεχόμενα δεν μπορεί να είναι ασυμβίβαστα.

15. Αν τα ενδεχόμενα Α, Β είναι ανεξάρτητα δείξτε ότι είναι ανεξάρτητα και τα ενδεχόμενα
α) Α΄ και Β,
β) Α΄ και Β΄.

16. Δείξτε ότι:
α. $P(A \setminus B) + P(A' \setminus B) = 1$

β. $P(B / A') = \dfrac{P(B) - P(B \cap A)}{1 - P(A)}$

17. Σε ένα λούνα-παρκ ο ιδιοκτήτης σκέφτηκε το εξής παιχνίδι. Πήρε δύο κουτιά και στο καθένα έβαλε 5 άσπρες και 5 μαύρες σφαίρες. Ο πελάτης επιλέγει τυχαία ένα κουτί, στη συνέχεια μία σφαίρα από αυτό και κερδίζει αν είναι άσπρη. Δείξτε ότι η πιθανότητα να κερδίσει ο πελάτης είναι $\frac{1}{2}$.

Κάποιος πελάτης όμως πρότεινε στον ιδιοκτήτη να αλλάξει όπως θέλει την κατανομή των σφαιρών στα κουτιά και ο ιδιοκτήτης δέχτηκε πιστεύοντας ότι οι πιθανότητες δεν θα αλλάξουν, αφού πάντα ο αριθμός των άσπρων σφαιρών θα είναι ίσος με τον αριθμό των μαύρων. Υπάρχει τρόπος ο πελάτης να μεγαλώσει την πιθανότητα να κερδίσει και ποια είναι η μέγιστη τιμή που μπορεί να πετύχει;

Σύνοψη

Ο λογισμός των πιθανοτήτων ασχολείται με την αναζήτηση και τη μελέτη των νόμων της τύχης με τους οποίους προσπαθούμε να ερμηνεύσουμε τα τυχαία εκείνα γεγονότα που παρουσιάζουν το χαρακτηριστικό της αβεβαιότητα.

Βασικές έννοες στο λογισμό πιθανοτήτων αποτελούν:

• Ο δειγματικός χώρος Ω / Πείραμα τύχης
• Ο πληθάριθμος
• Το ενδεχόμενο → Απλό
  → Σύνθετο
  → Βέβαιο
  → Αδύνατο
• Πράξεις με ενδεχόμενα → Ενδεχόμενα ασυμβίμαστα ή ξένα μεταξύ τους
  → Συμπληρωματικά ή αντίθετα
  → Τομή / ένωση ενδεχομένων
• Η έννοια της πιθανότητας → Αξιωματικός ορισμός
  → Κλασσικός ορισμός
• Ιδιότητες πιθανοτήτων (Λογισμός) → $P(A') = 1 - P(A)$
  → $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  → $P(A - B) = P(A) - P(A \cup B)$

- Δεσμευμένη πιθανότητα ⟶ $P(B/A) = \frac{P(A \cup B)}{P(A)}$

- Ανεξάρτητη ενδεχόμενα ⟶ $P(A \cap B) = P(A) \cdot P(B)$

Βιβλιογραφία / Internet

«Πιθανότητες και Στατιστική», Spiegel, MrGraw-Hill, ΕΣΠΙ

«Θεωρία Πιθανοτήτων και εφαρμογές», Χαραλαμπίδη Α.Χ.

«An Introduction to probability theory and its Application», Fellezw

«Εισαγωγικά Μαθήματα Θεωρίας Πιθανοτήτων», Τζιαφέτο Γ.

«Θεωρία Πιθανοτήτων», Αθανασοπούλου Δ.

www.answers.com/topic/probability: *ορισμός και ερμηνεία της πιθανότητας με links σε sites με μαθήματα στατιστικής και αντίστοιχο software.*

www/bymath.com: *θεωρία, προβλήματα, βοήθεια, tests και συμβουλές για στατιστική κ.ά. σε επίπεδο που καλύπτει τις γνώσεις γυμνασίου - λυκείου.*

Οδηγός για περεταίρω μελέτη

***«Ασκήσεις Πιθανοτήτων»*, Χαράλαμπος Χαραλαμπίου**

Στο βιβλίο αυτό ο ενδιαφερόμενος μπορεί να βρει μια ανασκόπηση της βασικής θεωρίας αλλά και πολλές ασκήσεις λυμένες ώστε να συμβάλλει σε μεγάλο βαθμό στην κατανόηση των σχερικών εννοιών και της τεχνικης συναγωγής συμπερασμάτων.

8. ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ

Βασικές έννοιες:
- διακριτές τυχαίες μεταβλητές
- συνάρτηση πιθανότητας διακριτής τυχαίας μεταβλητής
- συνάρτηση κατανομής διακριτής τυχαίας μεταβλητής
- αναμενόμενη τιμή διακριτής τυχαίας μεταβλητής
- διακύμανση διακριτής τυχαίας μεταβλητής
- τυπική απόκλιση διακριτής τυχαίας μεταβλητής
- συνεχείς τυχαίες μεταβλητές
- συναρτήση πυκνότητας πιθανοτήτας συνεχούς τυχαίας μεταβλητής
- συναρτήση κατανομής συνεχούς τυχαίας μεταβλητής
- μέση τιμή συνεχούς τυχαίας μεταβλητής

Στόχος του μαθήματος: Η εισαγωγή στην έννοια της τυχαίας μεταβλητής και η σύνδεσή σας με την έννοια της πιθανότητας, η οποία έχει ήδη εξεταστεί. Επίσης ο διαχωρισμός των τυχαίων μεταβλητών σε διακριτές και συνεχείς σε σχέση με το εκάστοτε πείραμα τύχης

Προσδοκώμενα αποτελέσματα: Ο εκπαιδευόμενος μαθένει να χειρίζεται τις τυχαίες μεταβλητές σε προβλήματα που του δίνονται. Παράλληλα τις εφαρμόζει χρησιμοποιώντας επιπλέον έννοιες που τις αφορούν, όπως συναρτήσεις πιθανοτήτας, κατανομής, διακύμανση και τυπική απόκλιση.

Εισαγωγικές παρατηρήσεις: Οι τυχαίες μεταβλητές μας δίνουν την πλήρη μαθηματικοποίηση του πειράματος τύχης πυ εκτελούμε

Αυτό πραγματοποιείται με την αντιστοίχιση κάθε ενδεχομένου ή αποτελέσματος σε συγκεκριμένο αριθμό και τη μελέτη του φαινομένου, που μας ενδιαφέρει, βάσει μεγεθών και συναρτήσεων, που αφορούν τους αριθμούς αυτούς, δηλαδή τις τιμές που παίρνουν οι τυχαίες μεταβλητές.

8.1. ΔΙΑΚΡΙΤΗ ΤΥΧΑΙΑ ΜΕΤΑΒΛΗΤΗ

Όταν εκτελούμε ένα τυχαίο πείραμα, όπως να ρίξουμε ένα ζάρι ή ένα νόμισμα, έχουμε και τα αντίστοιχα αποτελέσματα. Τα αποτελέσματα αυτά, άλλοτε είναι αριθμοί και άλλοτε όχι. Σε περίπτωση που δεν είναι αριθμοί τα αντιστοιχούμε σε αριθμούς, προκειμένου να τα χειριστούμε καλύτερα.

Στη συνέχεια χρησιμοποιούμε την αντιστοίχηση αυτή για τη μελέτη του πειράματος σε σχέση και με την πιθανότητα να συμβεί κάθε ενδεχόμενο.

Στα παραδείγματα που ακολουθούν γίνεται:

- Αντιστοίχηση των μη αριθμητικών αποτελεσμάτων σε αριθμούς.
- Σύνδεση με την πιθανότητα του κάθε ενδεχόμενου.

Εφαρμογή 1

Ρίχνουμε ένα ζάρι μία φορά. Ol πιθανές ενδείξεις, δηλαδή τα αποτελέσματα που μπορούμε να έχουμε, είναι:

$$\Omega = \{1,2,3,4,5,6\}$$

Συνολικά έχουμε 6 πιθανά αποτελέσματα. Άρα ο δειγματικός χώρος αποτελείται από 6 ενδεχόμενα ή $N(\Omega) = 6$.

Αν A_1 η πιθανότητα να εμφανιστεί η ένδειξη 1, A_2 η ένδειξη 2 και ούτω καθεξής, τότε η πιθανότητα να εμφανιστεί η ένδειξη 1 είναι:

$$P(A_1) = \frac{N(A_1)}{N(\Omega)} = \frac{1}{6}$$

και ομοίως υπολογίζουμε ότι για τις υπόλοιπες ενδείξεις 2,3,4,5 και 6, οι αντίστοιχες πιθανότητες εμφάνισής τους, είναι:

$$P(A_2) = \frac{N(A_2)}{N(\Omega)} = \frac{1}{6}$$

$$P(A_3) = \frac{N(A_3)}{N(\Omega)} = \frac{1}{6}$$

$$P(A_4) = \frac{N(A_4)}{N(\Omega)} = \frac{1}{6}$$

$$P(A_5) = \frac{N(A_5)}{N(\Omega)} = \frac{1}{6}$$

$$P(A_6) = \frac{N(A_6)}{N(\Omega)} = \frac{1}{6}$$

Στο σημείο αυτό για συντομία, αντί να λέμε «ένδειξη του ζαριού», θα χρησιμοποιήσουμε το γράμμα X. Όταν δηλαδή, η ένδειξη είναι 1 θα λέμε $X = 1$, όταν είναι 2 θα λέμε $X = 2$ και λοιπά.

Δηλαδή, η μεταβλητή X παίρνει τιμές από το σύνολο {1,2,3,4,5,6}.

Ή όπως φαίνεται στον πίνακα:

Ενδείξεις ζαριού	Τιμές X
1	1
2	2
3	3
4	4
5	5
6	6

Και χρησιμοποιώντας τη μεταβλητή X για τις πιθανότητες που βρήκαμε προηγουμένως, θα έχουμε:

$$P[X=1]=\frac{1}{6}$$

δηλαδή, η πιθανότητα το X να είναι 1 (δηλαδή η ένδειξη του ζαριού να είναι 1) ισούται με και όμοια για κάθε ένδειξη:

$$P[X=2]=\frac{1}{6}$$

$$P[X=3]=\frac{1}{6}$$

$$P[X=4]=\frac{1}{6}$$

$$P[X=5]=\frac{1}{6}$$

$$P[X=6]=\frac{1}{6}$$

Εφαρμογή 2

Ρίχνοντας ένα νόμισμα μία φορά, τα πιθανά αποτελέσματα είναι κορόνα και γράμματα. Ο δειγματικός χώρος είναι ο εξής:

$$\Omega = \{K, \Gamma\}$$

Οπότε, για τον δειγματικό χώρο είναι $N(\Omega) = 2$.

Παρατηρούμε ότι στην περίπτωση αυτή, τα αποτελέσματα δεν είναι αριθμοί όπως στο ζάρι. Πρέπει να κάνουμε ένα επιπλέον βήμα αντιστοιχώντας τα αποτελέσματα σε αριθμούς. Θεωρούμε ως X την εμφάνιση της ένδειξης κορόνα (K).

Οπότε έχουμε τον πίνακα:

Ένδειξη νομίσματος	Τιμές X
K	1
Γ	0

Οπότε:

Για την πιθανότητα να εμφανιστεί η ένδειξη Γράμματα (Γ)

$$P(\Gamma)=P[X=0]= \frac{1}{6}$$

Για την πιθανότητα να εμφανιστεί η ένδειξη Κορόνα (Κ)

$$P(K)=P[X=0]= \frac{1}{6}$$

Εφαρμογή 3

Ρίχνοντας ένα νόμισμα δύο φορές τα πιθανά αποτελέσματα και πάλι δεν είναι αριθμοί, όπως με το ζάρι. Ο δειγματικός χώρος, δηλαδή όλα τα δυνατά ενδεχόμενα, στην περίπτωση αυτή είναι το εξής σύνολο:

$$\Omega = \{(K,K), (K,\Gamma), (\Gamma,K), (\Gamma,\Gamma)\}$$

όπου Κ η ένδειξη κορόνα και Γ η ένδειξη γράμματα.

Οπότε, όλα τα δυνατά ενδεχόμενα είναι 4, δηλαδή $N(\Omega)=4$.

Για να αντιστοιχίσουμε αποτελέσματα σε πιθανότητες στην περίπτωση αυτή απαιτείται και πάλι το επιπλέον βήμα: να αντιστοιχίσουμε τα αποτελέσματα σε αριθμούς. Αν ονομάσουμε Χ τον αριθμό κεφαλών που εμφανίζονται στις δύο ρίψεις του νομίσματος, θα έχουμε:

Ενδείξεις 2 ρίψεων νομίσματος	Τιμές που παίρνει η Χ
(Γ, Γ)	0
(Γ, Κ), (Κ, Γ)	1
(Κ, Κ)	2

Οπότε θα έχουμε:

Για την πιθανότητα να εμφανιστεί η ένδειξη Γράμματα (Γ) δύο φορές:

$$P((\Gamma,\Gamma)) = P[X=0] = \frac{1}{4}$$

Για την πιθανότητα να εμφανιστεί η ένδειξη Γράμματα (Γ) μία φορά και η ένδειξη Κορόνα (Κ) μία φορά:

$$P((\Gamma,K),(K,\Gamma)) = P[X=1] = \frac{2}{4} = \frac{1}{2}$$

Για την πιθανότητα να εμφανιστεί η ένδειξη Κορόνα (Κ) δύο φορές:

$$P((K,K)) = P[X=2] = \frac{1}{4}$$

Εφαρμογή 4

Ρίχνουμε το ζάρι μία φορά και ενδιαφερόμαστε να δούμε όχι ποια είναι η ένδειξη, αλλά αν είναι μονός ή ζυγός ο αριθμός που φέραμε. Οπότε, θα θεωρήσουμε την μεταβλητή Χ ως εξής:

Ένδειξη ζαριού	Τιμή Χ
2,4,6	0
1,3,5	1

Οπότε θα έχουμε:

Για την πιθανότητα να φέρουμε ζυγό αριθμό:

$$P[X=0]=P((2,4,6))=P(2)+P(4)+P(6)=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{3}{6}=\frac{1}{2}$$

Για την πιθανότητα να φέρουμε μονό αριθμό:

$$P[X=1]=P((1,3,5))=P(1)+P(3)+P(5)=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{3}{6}=\frac{1}{2}$$

Επομένως:

- Στα παραπάνω παραδείγματα, τα πειράματα ήταν τυχαία, δηλαδή το ζάρι, το νόμισμα ήταν αμερόληπτα και κάθε ένδειξη είχε την ίδια πιθανότητα εμφάνισης με τις υπόλοιπες.
- Επιπλέον, χρησιμοποιήθηκε η Χ για να αντιστοιχίσουμε κάθε πιθανό αποτέλεσμα σε έναν αριθμό. Η Χ λέγεται **τυχαία μεταβλητή.**
- Στη συνέχεια, θα συμβολίζουμε με Χ, Υ, Ζ κλπ τέτοιες αντιστοιχίσεις, θα τις λέμε τυχαίες μεταβλητές ή τ.μ.
- Θα συμβολίζουμε με x, y, z κλπ. Τις τιμές που παίρνουν οι τ.μ.
- Η τ.μ. Χ στα παραπάνω παραδείγματα, έπαιρνε τις τιμές 1,2,3... Σε αυτή την περίπτωση, η τ.μ. ονομάζεται **διακριτή.**

8.2. ΣΥΝΑΡΤΗΣΗ ΠΙΘΑΝΟΤΗΤΑΣ ΔΙΑΚΡΙΤΗΣ Τ.Μ.

Επιπλέον στα προηγούμενα παραδείγματα:

- Αντιστοιχίσαμε σε κάθε τιμή της τ.μ. Χ την αντίστοιχη πιθανότητα, χρησιμοποιώντας την $P[X=x_j]$, την οποία ονομάζουμε **συνάρτηση πιθανότητας.**
- Η συνάρτηση πιθανότητας για συντομία συμβολίζεται με p_j αντί $P[X=x_j]$.

Εξετάζοντας τα παραδείγματα ως προς τη συνάρτηση πιθανότητας, είναι:

Εφαρμογή 5

Ρίψη ζαριού μία φορά.

Μας ενδιαφέρει ποιος αριθμός είναι η ένδειξη:

$P[X=1]=p_1=\frac{1}{6}$ $P[X=2]=p_2=\frac{1}{6}$

$P[X=3]=p_3=\frac{1}{6}$ $P[X=4]=p_4=\frac{1}{6}$

$P[X=5]=p_5=\frac{1}{6}$ $P[X=6]=p_6=\frac{1}{6}$

Επομένως:

$$\sum_{i=1}^{6} p_i = p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = \frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{6}{6}=1$$

Μπορούμε επίσης να παραστήσουμε σε πίνακα τα αποτελέσματα οπότε θα έχουμε:

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Εφαρμογή 6

Ρίψη ζαριού μία φορά.
Μας ενδιαφέρει ποια είναι η ένδειξη.

$P[X=0]=p_0=\frac{1}{2}$ $P[X=1]=p_1=\frac{1}{2}$

Επομένως:

$$\sum_{i=1}^{2} p_i = p_1 + p_2 = \frac{1}{2}+\frac{1}{2}=\frac{2}{2}=1$$

και σε πίνακα:

x_i	0	1
p_i	$\frac{1}{6}$	$\frac{1}{6}$

Εφαρμογή 7

Ρίψη νομίσματος δύο φορές.
Μας ενδιαφέρει ποια ένδειξη έχουμε κάθε φορά.

$P[X=0]=p_0=\frac{1}{4}$ $P[X=1]=p_1=\frac{1}{2}$ $P[X=2]=p_2=\frac{1}{4}$

Επομένως:

$$\sum_{i=1}^{3} p_i = p_1 + p_2 + p_3 = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = \frac{1}{4} + \frac{2}{4} + \frac{1}{4} = \frac{4}{4} = 1$$

και σε πίνακα:

x_i	0	1	2
p_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Εφαρμογή 8

Ρίψη ζαριού μία φορά.
Μας ενδιαφέρει αν η ένδειξη είναι μονός ή ζυγός αριθμός.

$P[X=0]=p_0=\frac{1}{2}$ $P[X=1]=p_1=\frac{1}{2}$

Επομένως:

$$\sum_{i=1}^{2} p_i = p_1 + p_2 = \frac{1}{2} + \frac{1}{2} = \frac{2}{2} = 1$$

και σε πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Συμπεράσματα:

1. Σε κάθε περίπτωση, η p_i είναι θετική και μικρότερη του 1.
2. Το άθροισμα όλων των p_i ισούται με 1.

- Τα παραπάνω είναι αναμενόμενα αφού τα p_i είναι πιθανότητες.
- Με μαθηματικές σχέσεις, για να εκφράσουμε τις παραπάνω παρατηρήσεις, γράφουμε:
  1. $0 \le p_j \le 1$
  2. $\sum_{i=1}^{k} p_i = p_1 + p_2 + \ldots + p_k = 1$

8.3. ΣΥΝΑΡΤΗΣΗ ΚΑΤΑΝΟΜΗΣ ΔΙΑΚΡΙΤΗΣ Τ.Μ.

Αν θέλουμε να εξετάσουμε την πιθανότητα η τ.μ. Χ να παίρνει τιμές μικρότερες ή ίσες με ένα συγκεκριμένο αριθμό, χρειαζόμαστε μία νέα συνάρτηση που θα αντιστοιχεί κάθε τιμή x_i στην αντίστοιχη τέτοια πιθανότητα.

Η συνάρτηση η οποία αντιστοιχεί σε κάθε αριθμό x_i την $P[X \le x_i]$ λέγεται **συνάρτηση κατανομής** της διακριτής τ.μ. Χ που είναι μικρότερες ή ίσες της x_i και συμβολίζεται με $F(x_i)$.

Εφαρμογή 9

Αν στη ρίψη ζαριού μας ενδιαφέρει να βρούμε την πιθανότητα να φέρουμε ένδειξη μικρότερη ή ίση του 1, δηλαδή $P[X \le 1]$.

Δηλαδή θέλουμε την πιθανότητα η ένδειξη να είναι 1.

Οπότε, χρησιμοποιούμε τα δεδομένα του πίνακα:

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

και συμβολικά γράφουμε:

$$F(1) = P[X \le 1] = P[X = 1] = p_1 = \frac{1}{6}$$

Για ένδειξη μικρότερη του 3, θέλουμε την πιθανότητα $P[X < 3]$.

Δηλαδή θέλουμε την πιθανότητα η ένδειξη να είναι 1 ή 2.

Οπότε συμβολικά γράφουμε:

$$F(2) = P[X \le 2] = P[X = 1] + P[X = 2] = p_1 + p_2 = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

Για ένδειξη μικρότερη του 4, θέλουμε την πιθανότητα $P[X < 4]$.

Δηλαδή θέλουμε την πιθανότητα η ένδειξη να είναι 1 ή 2 ή 3.

Οπότε συμβολικά γράφουμε:

$$F(3) = P[X \le 3] = P[X = 1] + P[X = 2] + P[X = 3] = p_1 + p_2 + p_3 =$$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

Για ένδειξη μικρότερη του 5 θέλουμε την πιθανότητα $P[X < 5]$.

Δηλαδή θέλουμε την πιθανότητα η ένδειξη να είναι 1 ή 2 ή 3 ή 4.

Οπότε συμβολικά γράφουμε:

$$F(4) = P[X \le 4] = P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4] =$$

$$= p_1 + p_2 + p_3 + p_4 = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

Για ένδειξη μικρότερη του 6 θέλουμε την πιθανότητα $P[X < 6]$.

Δηλαδή, θέλουμε την πιθανότητα η ένδειξη να είναι 1 ή 2 ή 3 ή 4 ή 5.

Οπότε, συμβολικά γράφουμε:

$$F(5)=P[X\le 5]=P[X=1]+P[X=2]+P[X=3]+P[X=4]+P[X=5]=$$

$$=p_1+p_2+p_3+p_4+p_5=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{5}{6}$$

και αν θέλουμε την πιθανότητα να είναι η ένδειξη μικρότερη του 7, θα γράψουμε:

$$F(6)=P[X\le 6]=P[X=1]+P[X=2]+P[X=3]+P[X=4]+P[X=5]+$$

$$+P[X=6]=p_1+p_2+p_3+p_4+p_5+p_6=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{6}{6}=1$$

κάτι αναμενόμενο, αφού είναι βέβαιο ότι η ένδειξη θα είναι μικρότερη ή ίση του 6.

Το ίδιο θα βρίσκαμε και αν θέλαμε την πιθανότητα η ένδειξη να είναι μικρότερη του 7, 8, 9 κλπ.

Επίσης, αν θέλουμε την πιθανότητα η ένδειξη να είναι μικρότερη ή ίση του 0 θα έχουμε:

$$P[X\le 0]=0$$

Κάτι επίσης αναμενόμενο, αφού η ένδειξη του ζαριού θα είναι τουλάχιστον 1.

Παρατηρήσεις:

- Καθώς η ανώτατη τιμή του X μεγαλώνει, μεγαλώνει και η αντίστοιχη πιθανότητα. Για παράδειγμα, 3<5 και $P[X\le 3]<P[X\le 5]$.
- Επιπλέον, κάθε φορά για να βρούμε την $F(\rho)=P[X\le\rho]$ προσθέτουμε στην $F(\rho-1)=P[X\le\rho-1]$.

  Συγκεκριμένα:

$$F(2)=P[X\le 2]=P[X\le 1]+P[X=2]=F(1)+P[X=2]$$

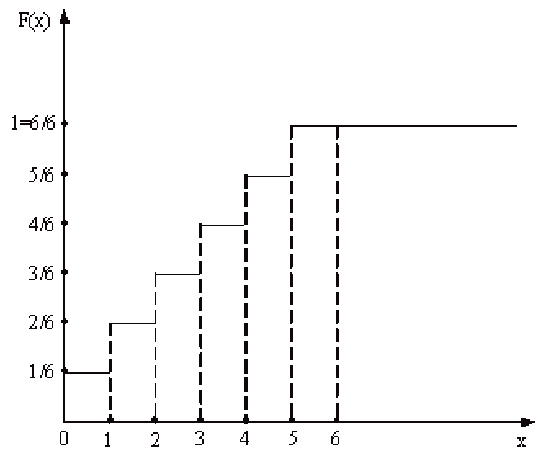
$$F(3)=P[X\le 3]=P[X\le 2]+P[X=3]=F(2)+P[X=3]$$

$$F(4)=P[X\le 4]=P[X\le 3]+P[X=4]=F(3)+P[X=4]$$

$$F(5)=P[X\le 5]=P[X\le 4]+P[X=5]=F(4)+P[X=5]$$

$$F(6)=P[X\le 6]=P[X\le 5]+P[X=6]=F(5)+P[X=6]$$

Κι αν θέλουμε να παραστήσουμε γραφικά τα παραπάνω:



Εφαρμογή 10

Στην περίπτωση ρίψης νομίσματος μία φορά από τον πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Έχουμε:

$P[X<0]=0$, αφού είναι αδύνατο το ενδεχόμενο η τ.μ. Χ να πάρει τιμή μικρότερη του 0. Ή 1 θα είναι (αν φέρουμε γράμματα) ή 0 (αν φέρουμε κορόνα).

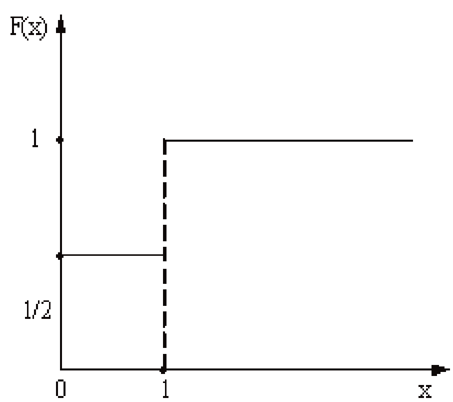
Επίσης,

$$F(0)=P[X\leq 0]=P[X=0]=\frac{1}{2}$$

και

$$F(1)=P[X\leq 1]=P[X=0]+P[X=1]=\frac{1}{2}+\frac{1}{2}=1$$

Σε διάγραμμα:



Εφαρμογή 11

Ρίχνοντας ένα νόμισμα δύο φορές και βάσει του πίνακα που βρήκαμε παραπάνω:

x_i	0	1	2
p_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Βρίσκουμε την πιθανότητα να φέρουμε λιγότερες από μία φορά κορόνα:

$$F(0) = P[X \leq 0] = P[X = 0] = \frac{1}{4}$$

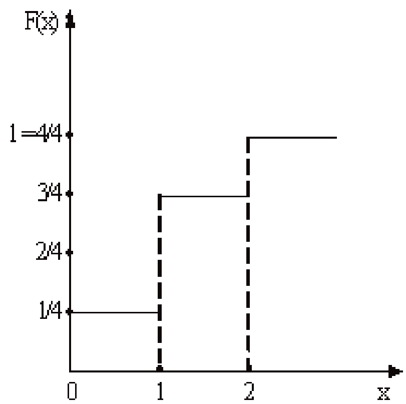
Το πολύ μία φορά κορόνα:

$$F(1) = P[X \leq 1] = P[X = 0] + P[X = 1] = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

Και το πολύ δύο φορές κορόνα:

$$F(2) = P[X \leq 2] = P[X = 0] + P[X = 1] + P[X = 2] = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = \frac{4}{4} = 1$$

Γραφικά έχουμε:



Εφαρμογή 12

Στην περίπτωση αυτή θα έχουμε τον πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

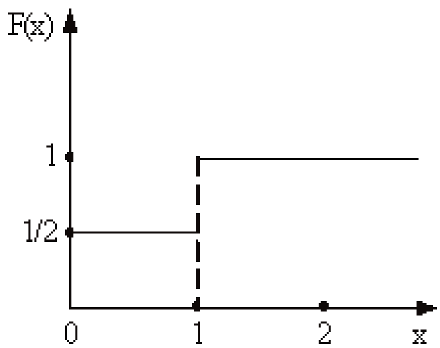
Και αντίστοιχα βρίσκουμε τις πιθανότητες:

$$F(0) = P[X \leq 0] = P[X = 0] = \frac{1}{2}$$

και

$$F(1) = P[X \leq 1] = P[X = 0] + P[X = 1] = \frac{1}{2} + \frac{1}{2} = 1$$

Οπότε γραφικά θα έχουμε:



8.4. ΑΝΑΜΕΝΟΜΕΝΗ ΤΙΜΗ ΔΙΑΚΡΙΤΗΣ Τ.Μ. Χ

Αν έχουμε μία διακριτή τυχαία μεταβλητή Χ που παίρνει τις τιμές:

$$x_1, x_2, \ldots, x_n$$

με αντίστοιχες πιθανότητες

$$p_1, p_2, \ldots, p_n$$

ονομάζουμε **μέση τιμή** ή αναμενόμενη ή μαθηματική ελπίδα διακριτής τ.μ. την ποσότητα:

$$\boxed{E(X) = p_1 x_1 + p_2 x_2 + \ldots + p_n x_n}$$

Ή χρησιμοποιώντας το συμβολισμό του αθροίσματος:

$$\boxed{E(X) = \sum_{i=1}^{n} p_i x_i}$$

Εκτός από το συμβολισμό «E(X)», για τη μέση τιμή χρησιμοποιείται και το «μ». Σε σχέση με τα προηγούμενα παραδείγματα:

Εφαρμογή 13

Έχουμε τον πίνακα:

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Οπότε υπολογίζουμε:

$$E(X) = \sum_{i=1}^{6} p_i x_i = p_1 x_1 + p_2 x_2 + \ldots + p_6 x_6 = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6 = \frac{21}{6}$$

Εφαρμογή 14

Βάσει του πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Έχουμε:

$$E(X)=\sum_{i=1}^{2}p_i x_i = p_1x_1+p_2x_2=\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 1=\frac{1}{2}$$

Εφαρμογή 15

Βάσει του πίνακα:

x_i	0	1	2
p_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Έχουμε:

$$E(X)=\sum_{i=1}^{3}p_i x_i = p_1x_1+p_2x_2+p_3x_3=\frac{1}{4}\cdot 0+\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2=0+\frac{1}{2}+\frac{2}{4}=\frac{2}{4}+\frac{2}{4}=\frac{4}{4}=1$$

Εφαρμογή 16

Βάσει του πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Έχουμε:

$$E(X)=\sum_{i=1}^{2}p_i x_i = p_1x_1+p_2x_2=\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 1=0+\frac{1}{2}=\frac{1}{2}$$

Σημείωση:
Η αναμενόμενη τιμή μίας διακριτής τ.μ. X δεν είναι απαραίτητα ακέραιος αριθμός και δεν είναι απαραίτητα μία από τις x1 δηλαδή μία από τις τιμές που παίρνει η διακριτή τ.μ. X.

8.5. ΔΙΑΚΥΜΑΝΣΗ ΔΙΑΚΡΙΤΗΣ Τ.Μ. Χ

Έστω ότι έχουμε μία τ.μ. Χ που παίρνει τις τιμές:

$$x_1, x_2, ..., x_n$$

με αντίστοιχες πιθανότητες

$$p_1, p_2, ..., p_n$$

Ονομάζουμε **διασπορά** διακριτής τ.μ. την ποσότητα:

$$\boxed{Var(X)=E(X^2)-\{E(X)\}^2}$$

Όπου

$$\boxed{E(X^2)=\sum_{i=1}^{n} p_i x_i^2}$$

Ή χρησιμοποιώντας το συμβολισμό του αθροίσματος:

$$\boxed{Var(X)=\sum_{i=1}^{n} p_i x_i^2-\left\{\sum_{i=1}^{n} p_i x_i\right\}^2}$$

Εκτός από το «Var(X)» για τη διακύμανση χρησιμοποιείται και ο συμβολισμός «σ^2».

Σε σχέση με τα προηγούμενα παραδείγματα:

Εφαρμογή 17

Έχουμε τον πίνακα:

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Οπότε υπολογίζουμε:

$$E(X)=\sum_{i=1}^{6} p_i x_i = p_1x_1+p_2x_2+...+p_6x_6=\frac{21}{6} \text{ οπότε } E(X)^2=\frac{121}{36}$$

και

$$E(X^2)=\sum_{i=1}^{6} p_i x_i^2=\frac{1}{6}\cdot 1^2+\frac{1}{6}\cdot 2^2+\frac{1}{6}\cdot 3^2+\frac{1}{6}\cdot 4^2+\frac{1}{6}\cdot 5^2+\frac{1}{6}\cdot 6^2=$$

$$\frac{1}{6}\cdot 1+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 9+\frac{1}{6}\cdot 16+\frac{1}{6}\cdot 25+\frac{1}{6}\cdot 36=\frac{1}{6}+\frac{4}{6}+\frac{9}{6}+\frac{16}{6}+\frac{25}{6}+\frac{36}{6}=\frac{91}{6}$$

Επομένως $Var(X)=E(X^2)-E(X)^2=\frac{81}{6}-\frac{121}{36}=\frac{365}{36}$.

Εφαρμογή 18

Βάσει του πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Είναι:

$$E(X)=\sum_{i=1}^{2} p_i x_i = \frac{1}{2}$$

Άρα $\{E(X)\}^2 = \frac{1}{4}$

και $E(X)=\sum_{i=1}^{2} p_i x_i = p_1 x_1 + p_2 x_2 = \frac{1}{2}\cdot 0+\frac{1}{2}\cdot 1=\frac{1}{2}$

Επομένως:

$$Var(X)=E(X^2)-\{E(X)\}^2=\frac{1}{2}-\frac{1}{2}=0$$

Εφαρμογή 19

Βάσει του πίνακα:

x_i	0	1	2
p_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Έχουμε:

$$E(X)=\sum_{i=1}^{3} p_i x_i = p_1 x_1 + p_2 x_2 + p_3 x_3 = \frac{1}{4}\cdot 0+\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2 = 0+\frac{1}{2}+\frac{2}{4}=\frac{2}{4}+\frac{2}{4}=\frac{4}{4}=1$$

άρα $\{E(X)\}^2=1$

και $E(X^2)=\sum_{i=1}^{3} p_i x_i^2 = \frac{1}{4}\cdot 0^2+\frac{1}{2}\cdot 1^2+\frac{1}{4}\cdot 2^2 = 0+\frac{1}{2}+\frac{4}{4}=\frac{1}{2}+\frac{2}{2}=\frac{3}{2}$

Επομένως:

$$Var(X)=E(X^2)-\{E(X)\}^2=\frac{3}{2}-1=\frac{1}{2}$$

Εφαρμογή 20

Βάσει του πίνακα:

x_i	0	1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Έχουμε:

$$E(X)=\sum_{i=1}^{2}p_i x_i = p_1 x_1 + p_2 x_2 = \frac{1}{2}\cdot 0+\frac{1}{2}\cdot 1=\frac{1}{2}$$

Άρα $\{E(X)\}^2=\frac{1}{4}$

και $E(X^2)=\sum_{i=1}^{2}p_i x_i^2=\frac{1}{2}\cdot 0^2+\frac{1}{2}\cdot 1^2=0+\frac{1}{2}=\frac{1}{2}$

Επομένως:

$$Var(X)=E(X^2)-\{E(X)\}^2=\frac{1}{2}-\frac{1}{2}=0$$

<u>Παρατηρήσεις</u>:

- Η διακύμανση δεν είναι ποτέ αρνητική.
- Οι ποσότητες $E[X^2]$ και $\{E[X]\}^2$ είναι διαφορετικές στη γενική περίπτωση.

8.6 .ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ ΔΙΑΚΡΙΤΗΣ ΤΥΧΑΙΑΣ ΜΕΤΑΒΛΗΤΗΣ

Ονομάζεται **τυπική απόκλιση** μίας διακριτής τ.μ. Χ η ποσότητα, $\sqrt{Var(X)}$ δηλαδή η τετραγωνική ρίζα της διασποράς της Χ.

Η τυπική απόκλιση συμβολίζεται με το γράμμα «σ».

Άσκηση

Για τη διακριτή τ.μ. Χ με συνάρτηση πιθανότητας

$$f(x)=\frac{1}{30}x_i^2,\quad x_i=1,2,3,4$$

Να γίνει ο πίνακας πιθανότητας και να βρεθούν οι:

Α. μέση τιμή

Β. διασπορά

Γ. τυπική απόκλιση

ΛΥΣΗ:

Έχουμε:

$$X=1 \quad \rightarrow \quad f(1)=p_1=\frac{1}{30}1^2=\frac{1}{30}$$

$$X=2 \quad \rightarrow \quad f(2)=p_2=\frac{1}{30}2^2=\frac{4}{30}$$

$$X=3 \quad \rightarrow \quad f(3)=p_3=\frac{1}{30}3^2=\frac{9}{30}$$

$$X=4 \quad \rightarrow \quad f(4)=p_4=\frac{1}{30}4^2=\frac{16}{30}$$

Άρα έχουμε τον πίνακα:

x_i	1	2	3	4
p_i	$\frac{1}{30}$	$\frac{4}{30}$	$\frac{9}{30}$	$\frac{16}{30}$

A. $E(X)=\sum_{i=1}^{4} p_i x_i = p_1x_1+p_2x_2+p_3x_3+p_4x_4=\frac{1}{30}1+\frac{4}{30}2+\frac{9}{30}3+\frac{16}{30}4=\frac{100}{30}=\frac{10}{3}$

B. $E(X)=\frac{10}{3}$ άρα $\left\{E(X)^2\right\}=\frac{100}{9}$

και $E(X^2)=\sum_{i=1}^{4} p_i x_i^2=\frac{1}{30}\cdot 1^2+\frac{4}{30}\cdot 2^2+\frac{9}{30}\cdot 3^2+\frac{16}{30}\cdot 4^2=\frac{1}{30}+\frac{16}{30}+\frac{81}{30}+\frac{256}{30}=\frac{354}{30}$

Επομένως:

$$Var(X)=E(X^2)-\{E(X)\}^2=\frac{354}{30}-\frac{100}{30}=\frac{254}{100}$$

Γ. η τυπική απόκλιση θα είναι στην περίπτωση αυτή:

$$\sigma=\sqrt{Var(X)}=\sqrt{\frac{254}{30}}$$

1. Για τη διακριτή τυχαία μεταβλητή X με συνάρτηση πιθανότητας:

$$f(x) = \frac{3x_i - 2}{35}, \quad x_i = 1, 2, 3, 4$$

α. Να γίνει ο πίνακας πιθανότητας.

β. Να βεβαιωθείτε ότι πρόκειται για συνάρτηση πιθανότητας (Υπόδειξη: αθροίστε για x_i = 1,2,3,4 και ελέγξτε αν το αποτέλεσμα είναι 1).

γ. Να βρεθεί η μέση τιμή.

δ. Να βρεθεί η διασπορά.

ε. Να βρεθεί η τυπική απόκλιση.

2. Για τη διακριτή τυχαία μεταβλητή δίνεται ο πίνακας πιθανότητας:

x_i	1	2	3	4
p_i	$\frac{2}{5}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{3}{10}$

Να εξετάσετε αν πρόκειται για συνάρτηση πιθανότητας και στη συνέχεια:

α. Να βρεθεί η μέση τιμή.

β. Να βρεθεί η διασπορά.

γ. Να βρεθεί η τυπική απόκλιση.

3. Ενδιαφερόμαστε για την ηλικία σε έτη, στην οποία έκαναν το πρώτο τους παιδί 16 μητέρες σε ένα συγκεκριμένο χωριό. Συγκεντρώσαμε τα αποτελέσματα και προέκυψε ο πίνακας:

25	24	28	25
20	29	20	24
28	29	24	22
22	21	24	25

α. Να γίνει ο πίνακας πιθανοτήτων.

β. Να βρεθεί η μέση τιμή.

γ. Να βρεθεί η διασπορά.

δ. Να βρεθεί η τυπική απόκλιση.

8.7. ΣΥΝΕΧΗΣ ΤΥΧΑΙΑ ΜΕΤΑΒΛΗΤΗ

Υπάρχουν περιπτώσεις στις οποίες η τυχαία μεταβλητή είναι για παράδειγμα ένα μέγεθος, όπως το ύψος των μαθητών της Α΄ Λυκείου, ο χρόνος καθυστέρησης των αεροπλάνων, η απόσταση που διανύουμε και άλλα. Είναι προφανές ότι αν θεωρήσουμε τυχαία μεταβλητή διακριτή, όπως στις προηγούμενες παραγράφους, δεν θα έχουμε όλες τις δυνατές τιμές του μεγέθους που εξετάζουμε. Αυτό γιατί η μεταβλητή αλλάζει διαρκώς και δεν παίρνει μόνο ακέραιες τιμές, αλλά και όλες τις ενδιάμεσές τους. Όπως λέμε στα Μαθηματικά, παίρνει τιμές στο διάστημα [α, β] αν θεωρήσουμε ότι κυμαίνεται από το α ως το β.

Μία τέτοια μεταβλητή ονομάζεται **συνεχής**.

Για τις συνεχείς μεταβλητές ορίζονται τα ίδια μέτρα, όπως και στις διακριτές τ.μ., αλλά που υπολογίζονται με διαφορετικό τρόπο.

Εφαρμογή 21

Μία εταιρία που προμηθεύει ζαχαροπλαστεία με αυγά, θέλει να χωρίσει τα αυγά σε συγκεκριμένα μεγέθη ανάλογα με το βάρος τους, για την καλύτερη εξυπηρέτηση των πελατών της. Οπότε, σε μετρήσεις που έγιναν καταλήξαμε στον εξής πίνακα που περιέχει τα βάρη 50 αυγών σε γραμμάρια:

51	72	70	52	75	59	58	63	62	59
65	64	58	52	61	63	61	55	61	61
73	65	62	63	56	60	68	57	64	60
66	61	63	53	59	70	67	66	54	64
74	68	68	55	75	69	54	58	71	63

Για να μιλήσουμε για τα δεδομένα που θα έχουμε δεν πρακτικό να έχουμε ως τυχαία μεταβλητή το βάρος κάθε αυγού, η οποία θα παίρνει όλες τις τιμές που περιέχει ο πίνακας. Για ευκολία, χωρίζουμε τα δεδομένα μας σε κλάσεις (ομάδες), ως εξής:

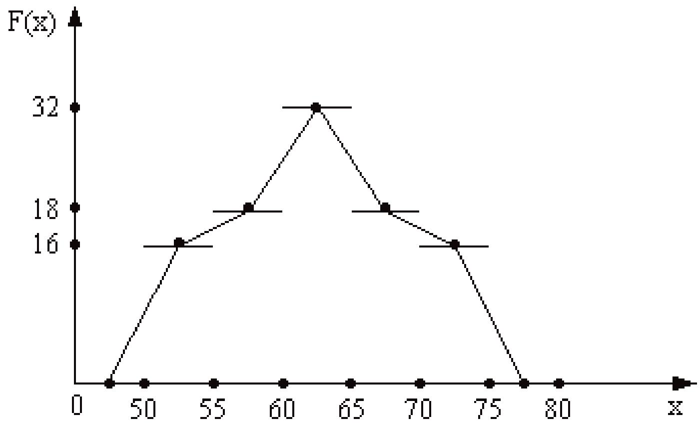
Βάρος αυγού
50 - 55
55 - 60
60 - 65
65 - 70
70 - 75

Αν θέλουμε, μπορούμε να έχουμε μεγαλύτερη ακρίβεια, χωρίζοντας τα δεδομένα μας σε κλάσεις με μικρότερο πλάτος. Η διαδικασία αύξησης του αριθμού των κλάσεων έχει νόημα όταν τα δεδομένα είναι αρκετά, ώστε να μην προκύπτουν κλάσεις με μηδενική συχνότητα!

Στη συνέχεια, υπολογίζουμε τις σχετικές συχνότητες και τις αθροιστικές συχνότητες, οπότε προκύπτει ο πίνακας κατανομής συχνοτήτων:

Βάρος αυγού	Συχνότητα
50 - 55	18%
55 - 60	32%
60 - 65	18%
65 - 70	16%
70 - 75	16%

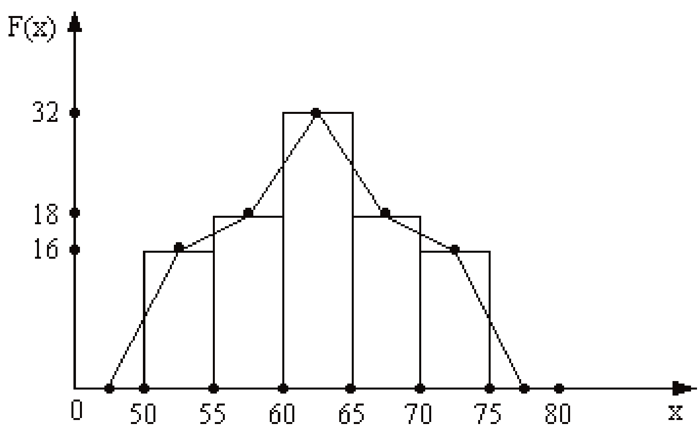
Για να έχουμε καλύτερη εικόνα της κατάστασης φτιάχνουμε το αντίστοιχο ιστόγραμμα:



Ιστόγραμμα Συχνοτήτων

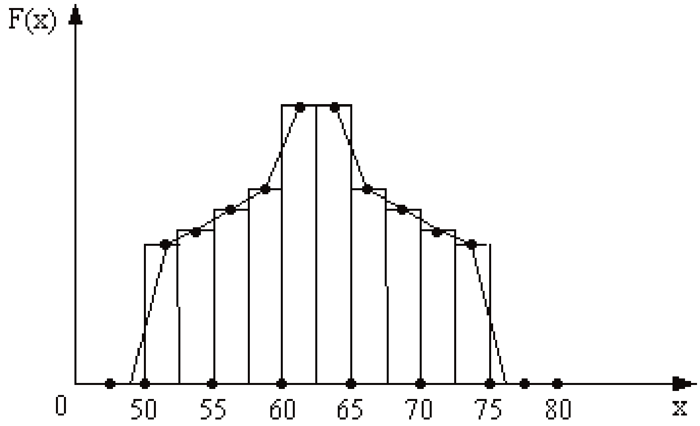
8.8. ΣΥΝΑΡΤΗΣΗ ΠΥΚΝΟΤΗΤΑΣ ΠΙΘΑΝΟΤΗΤΑΣ ΣΥΝΕΧΟΥΣ Τ.Μ.

Παρατηρώντας το ιστόγραμμα του παραδείγματος και τοποθετώντας σε αυτό τα αντίστοιχα παραλληλόγραμμα, έχουμε το προηγούμενο σχήμα:

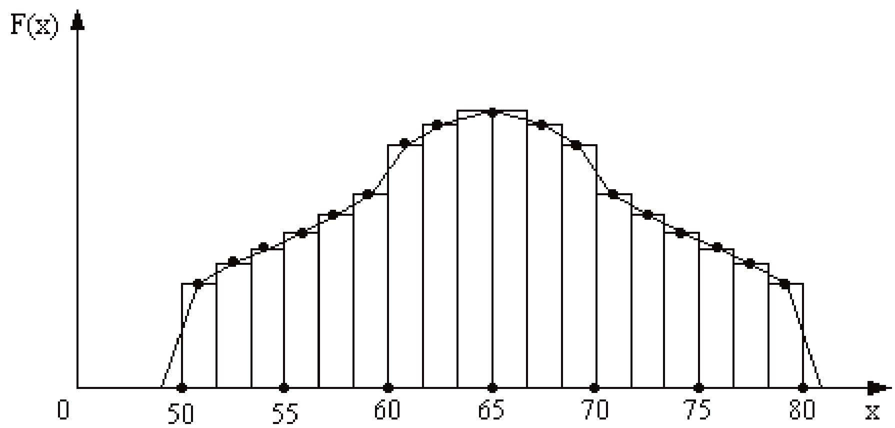


Ιστόγραμμα Συχνοτήτων

Αν τώρα αυξήσουμε τον αριθμό των κλάσεων, θα έχουμε:



Και αν συνεχίσουμε αυξάνοντας τον αριθμό των κλάσεων:



Βλέπουμε ότι καθώς το πλήθος των κλάσεων αυξάνεται, τόσο καλύτερα προσεγγίζεται η πραγματική κατάσταση και βελτιώνεται η εκτίμησή μας.

Κάτι τέτοιο βέβαια δε σημαίνει ότι οι κλάσεις πρέπει να είναι πάρα πολλές κάνοντας δύσκολη τη μελέτη του φαινομένου που μας ενδιαφέρει.

Επιπλέον, κάτι που πρέπει να παρατηρήσουμε γιατί είναι χρήσιμο, είναι ότι, το εμβαδόν κάθε ορθογωνίου είναι ίσο με f_i και άρα το εμβαδόν όλων των ορθογωνίων θα ισούται με 1.

Όμως, $f_i = \frac{n_i}{n}$ και $p_i = \frac{n_i}{n}$, άρα $f_i=p_i$.

Άρα, $\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} p_i$ (*).

Είναι όμως γνωστό ότι το άθροισμα όλων των πιθανοτήτων ισούται με 1, συμβολικά $\sum_{i=1}^{n} p_i = 1$ άρα από την (*) θα είναι και $\sum_{i=1}^{n} f_i = 1$, δηλαδή το εμβαδόν μεταξύ της καμπύλης συχνοτήτων και των κατακόρυφων ευθειών x = 50 και x = 75 είναι ίσο με 1.

Στα Μαθηματικά, ο υπολογισμός του εμβαδού μεταξύ μίας καμπύλης και δύο ευθειών, όπως αυτό που εμφανίζεται παραπάνω, γίνεται με τη βοήθεια του ολοκληρώματος.

Οπότε, έχουμε αντί για άθροισμα $\sum_{i=1}^{n} f(x_i)$, ολοκλήρωμα $\int_{\alpha}^{\beta} f(x)dx$ άρα $\int_{\alpha}^{\beta} f(x)dx = 1$.

Στην περίπτωση συνεχούς τ.μ. η συνάρτηση f ονομάζεται συνάρτηση πυκνότητας της πιθανότητας.

8.9. ΣΥΝΑΡΤΗΣΗ ΚΑΤΑΝΟΜΗΣ ΣΥΝΕΧΟΥΣ Τ.Μ.

Η πιθανότητα P{γ<X<δ} θα είναι από το ιστόγραμμα: ίση με το εμβαδόν ανάμεσα στην καμπύλη συχνοτήτων και τις κατακόρυφες ευθείες x = γ και x = δ.

Χρησιμοποιώντας ολοκλήρωμα, θα έχουμε:

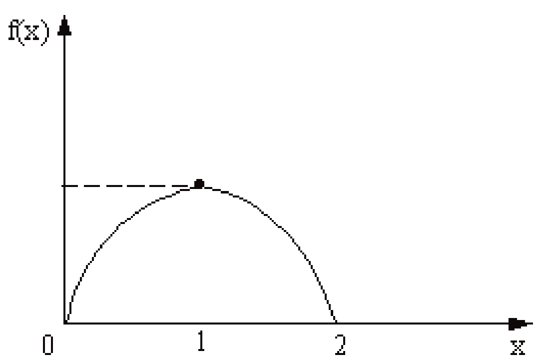
$$\boxed{P\{\gamma < X < \delta\} = \int_{\gamma}^{\delta} f(x)dx}$$

Εφαρμογή 22

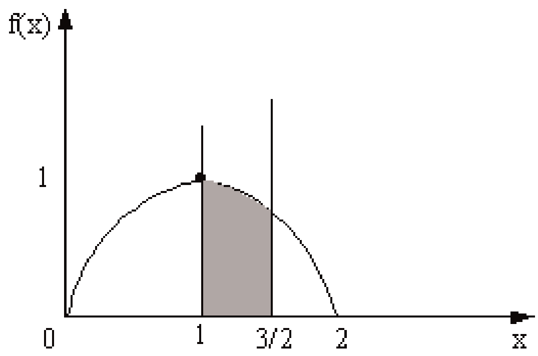
Έστω ότι δίνεται συνεχής τ.μ. με συνάρτηση πυκνότητας σταθερότητας:

$$f(x) = \begin{cases} \frac{3}{2}x - \frac{3}{4}x^2, & 0 \le x \le 2 \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

Τότε το αντίστοιχο διάγραμμα, θα είναι:



Και αν θέλουμε την πιθανότητα $P\left(1<x<\frac{3}{2}\right)$, θα θέλουμε το γραμμοσκιασμένο εμβαδό:



Οπότε, χρησιμοποιώντας ολοκλήρωμα, θα είχαμε:

$$P\left(1<x<\frac{3}{2}\right)=\int_{1}^{3/2} f(x)dx=\int_{1}^{3/2}\left(\frac{3}{2}x-\frac{3}{4}x^2\right)dx=\frac{3}{2}\int_{1}^{3/2} xdx-\frac{3}{4}\int_{1}^{3/2} x^2dx=$$

$$=\frac{3}{4}\left[x^2\right]_1^{3/2}-\frac{1}{4}\left[x^3\right]_1^{3/2}$$

$$\frac{3}{4}\cdot\left(\frac{9}{4}-1\right)-\frac{1}{4}\left(\frac{27}{8}-1\right)=\frac{3}{4}\cdot\left(\frac{9}{4}-\frac{4}{4}\right)-\frac{1}{4}\left(\frac{27}{8}-\frac{8}{8}\right)=$$

$$=\frac{3}{4}\cdot\frac{5}{4}-\frac{1}{4}\cdot\frac{19}{8}=\frac{15}{16}-\frac{19}{32}=\frac{30}{32}-\frac{19}{32}=\frac{11}{32}$$

8.10. ΜΕΣΗ ΤΙΜΗ ΣΥΝΕΧΟΥΣ Τ.Μ.

Χρησιμοποιώντας το ολοκλήρωμα στη συνεχή τ.μ. αντί του αθροίσματος της διακριτής τ.μ. αντίστοιχα για τη μέση τιμή συνεχούς τ.μ. θα είναι:

$$\boxed{E(x)=\int_{-\infty}^{+\infty} x\cdot f(x)dx}$$

ή στην πράξη $E(x)=\int_{\alpha}^{\beta} x\cdot f(x)dx$, όταν το x παίρνει τιμές α ως β, δηλαδή στο διάστημα [α, β].

Σε σχέση με την προηγούμενη εφαρμογή:

$$f(x)=\begin{cases}\frac{3}{2}x-\frac{3}{4}x^2, & 0\le x\le 2\\ 0, & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}$$

η μέση τιμή θα είναι:

$$E(x) = \int_0^2 x \cdot f(x)dx =$$

$$\int_0^2 x \cdot \left(\frac{3}{2}x - \frac{3}{4}x^2\right)dx =$$

$$\int_0^2 \left(\frac{3}{2}x^2 - \frac{3}{4}x^3\right)dx =$$

$$\frac{3}{2}\int_0^2 x^2 dx - \frac{3}{4}\int_0^2 x^3 dx =$$

$$\frac{1}{2}\left[x^3\right]_0^2 - \frac{3}{16}\left[x^4\right]_0^2 =$$

$$\frac{1}{2}(8-0) - \frac{3}{16}(16-0) =$$

$$\frac{1}{2}\cdot 8 - \frac{3}{16}\cdot 16$$

$$= 4 - 3$$

$$= 1$$

ΕΞΑΣΚΗΣΗ

1. Να εξετάσετε αν έχουμε διακριτή ή συνεχή τυχαία μεταβλητή σε κάθε μία από τις επόμενες περιπτώσεις:
   α. χρόνος αναμονής στην τράπεζα.
   β. αριθμός παιδιών ανά οικογένεια.
   γ. αριθμός επιτυχημένων χτυπημάτων πέναλτι.
   δ. ακυρώσεις αεροπορικών εισιτηρίων ανά ημέρα που δέχεται μία αεροπορική εταιρία.
   ε. διάρκεια τηλεφωνικών συνδιαλέξεων.
   στ. αριθμός ελαττωματικών λαμπτήρων ανά παρτίδα ενός εργοστασίου.
   ζ. αριθμός τυπογραφικών λαθών ανά σελίδα ενός βιβλίου.
   η. ηλικία σε έτη που γέννησαν το $1^{ο}$ τους παιδί 20 γυναίκες.

2. Να εξετάσετε αν η συνάρτηση $f(x) = \frac{1}{x}, \quad x=1,2,3$ είναι συνάρτηση πιθανότητας.

3. Να εξετάσετε αν η συνάρτηση f(x)=2−x, x=1,2 είναι συνάρτηση πιθανότητας.

4. Να βρεθεί η μέση τιμή της τ.μ. x με συνάρτηση πιθανότητας f(x)=2−x, x=1,2.

5. Να εξετάσετε αν η συνάρτηση $f(x) = -2, \quad x \in [0,2]$ είναι συνάρτηση πιθανότητας.

6. Ομοίως για τη συνάρτηση $f(x) = -2, \quad x \in [0,1]$

7. Να βρεθεί η μέση τιμή της συνάρτησης $f(x)=\frac{3}{2}x-\frac{3}{4}x^2, \quad x\in[0,2]$

8. Να βρεθεί ο αριθμός κ, αν η συνάρτηση:

$$f(x)=\begin{cases}\frac{1}{4}x+\kappa, & \alpha\nu \quad 0\le x\le 2\\ 0, & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}$$

είναι συνάρτηση πυκνότητας πιθανότητας.

9. Να βρεθεί η μέση τιμή της x, βάσει της συνάρτησης της άσκησης 8, για την τιμή του κ που βρήκατε.

10. Να σημειώσετε Σ (σωστό) ή Λ (λάθος) στις παρακάτω προτάσεις:

(α) Η συνάρτηση , είναι συνάρτηση πυκνότητας πιθανότητας. Σ Λ

(β) Η συνάρτηση f(x)=3, $x_1=1$, $x_2=2$, $x_3=3$ είναι συνάρτηση πιθανότητας. Σ Λ

(γ) Το γράφημα της συνάρτησης πιθανότητας διακριτής τυχαίας μεταβλητής είναι μία συνεχής καμπύλη. Σ Λ

(δ) Για τον υπολογισμό της μέσης τιμής συνεχούς τυχαίας μεταβλητής, χρησιμοποιούμε το άθροισμα. Σ Λ

11. Άν x συνεχής τυχαία μεταβλητή με

$$f(x)=\begin{cases}e^{-x}, & x>0\\ 0, & x\le 0\end{cases}$$

να βρεθεί η πιθανότητα x>2 και η πιθανότητα $-3<x\le 4$

12. Άν x συνεχής τυχαία μεταβλητή με

$$f(x)=\begin{cases}e^{-x}, & x>0\\ 0, & x\le 0\end{cases}$$

να βρεθεί η E(x) και η $E(x^2)$.

13. Άν x συνεχής τυχαία μεταβλητή με $f(x)=\begin{cases}4x^2, & 0\le x\le 1\\ 0, & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}$ να υπολογιστεί η E(x) και η $E(x^2)$.

Σύνοψη

Αν η φύση των στοιχείων των τυχαίων ενδεχόμενων είναι ποσοτική τότε τη στοιχειώδη ενδεχόμενη του δ. χώρου ενός πειράματος τύχτύχης εκφράζονται με αριθμούς όπως ρίψη ενός ζαριού. Όταν όμως τα στοιχειώδη ενδ. του δειγματικού χώρου εκφράζονται με αντίστοιχες τιμές ποιοτικών χαρακτηριστικών όπως στη ρίψη νομίσματος, τότε είναι χρήσιμο να παρασταθούν συμβατικά τα στοιχεία στου δ. χώρου με πραγματικούς αριθμούς.

- **Τυχαία μεταβλητή:** Όταν εξετάζουμε ένα δ.χ. Ω που αναφέρεται σ' ένα πείραμμα τύχης και αντιστοιχίζουμε σε κάθε αποτέλεσμα του δ. χώρου έναν πραγματικό αριθμό, ορίζουμε έτσι μια συνάρτηση που έχει πεδίο ορισμού το Ω και σύνολο τιμών ένα υποσύνολο των πραγματικών αριθμών. Η συνάρτηση αυτή ομονάζεται τυχαία μεταβλητή.
- **Συνάρτηση πιθανότητας:** Η πιθανότητα με την οποία η τυχαία μεταβλητή x μπορεί να λάβει την τιμή x_i $P[x = x_i]$
- **Συνάρτηση κατανομής:** Είναι η πιθανότητα να πάρει η τυχαία μεταβλητή x τιμή μικρότερη ή ίση από την τιμή x.
- **Συνεχής τυχαία μεταβλητή:** Όταν η τυχαία μεταβλητή x μπορεί να πάρει όλες τις τιμές στου διαστήματος [α,β].
- **Συνάρτηση κατανομής συνεχούς τυχαίες μεταβλητές:** Όταν το πεδίο τιμών της τυχαία μεταβλητής x είναι το σύνολο [α,β] τότε η συνάρτηση κατανομής τυχαίας μεταβλητής δίνεται από τη σχέση $P[\alpha<x<\beta] = \int_{\alpha}^{\beta} f(x)d(x)$

Βιβλιογραφία / Internet

«Λογισμός πιθανοτήτων», Κιόχος Π., 1989

«Πιθανότητες και Στατιστική», Spiegel, MrGraw-Hill, ΕΣΠΙ

«Στατιστική», Πέτρος Α. Κιόχος, INTERBOOKS

«Θεωρία Πιθανοτήτων και Εφαρμογές», Κοκολάκης Γ. Σπηλιώτης Ι., 1985

«Θεωρία Πιθανοτήτων και Εφαρμογές», τεύχος 1, Χαράλαμπου Α. Καραλαμπίδη, ΣΥΜΜΕΤΡΙΑ

«Θεωρία Πιθανοτήτων», Θεόφιλου Ν. Κάκκουλου

«The Elements of Probability theory and some of its application», Cramer H, New York, 1973

www.statcan.ca: *επίσημο site στατιστικής τον Καναδά.*

www.math.aegean.gr/courses/probabilities/P_KO6/index.htm: *site που ο ενδιαφερόμενος μπορεί να βρει τη βασική θεωρία και ασκήσεις τυχαίων μεταβλητών*

ΟΔΗΓΟΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΜΕΛΕΤΗ

«*Εισαγωγή στις πιθανότητες Θεωρία και εφαρμογές*», Εκδόσεις INTERBOOKS, Μάρκου Κούτρα

Στο βιβλίο αυτό αναπτύσσονται θέματα στο λογισμό πιθανοτήτων, στις τυχαίες μεταβλητές καθώς και στις διακριτές και συνεχές κατανομές υπάρχουν πολλές λυμένες εφαρμογές που βοηθούν τον ενδιαφερόμενο στην ακόμα καλύτερη κατανόηση του γνωσιακού αντικειμένου.

9. ΘΕΩΡΗΤΙΚΕΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ

Βασικές έννοιες:
- κατανομή Bernoulli
- διωνυμική κατανομή
- γεωμετρική κατανομή
- κατανομή Poisson
- υπεργεωμετρική κατανομή
- κανονική κατανομή
- εκθετική κατανομή
- ομοιόμορφη κατανομή
- κατανομή Γάμα

Στόχος του μαθήματος: να χειρίζεται ο εκπαιδευόμενος τις βασικές θεωρητικές κατανομές.

Προσδοκώμενα αποτελέσματα: ο εκπαιδευόμενος να είναι σε θέση να κρίνει ποια κατανομή θα χρησιμοποιήσει ανάλογα με το πρόβλημα που πρέπει να λύσει.

Εισαγωγικές παρατηρήσεις: Στην πράξη προσπαθούμε να προσεγγίσουμε και να προβλέψουμε φαινόμενα χρησιμοποιώντας συγκεκριμένες κατανομές τυχαίων μεταβλητών διακριτών ή συνεχών ανάλογα με την περίπτωση που εξετάζουμε. Αυτές τις κατανομές τις ονομάζουμε θεωρητικές κατανομές. Το πλεονέκτημα στην περίπτωση που χρησιμοποιούμε μία θεωρητική κατανομή είναι ότι οι παράμετροι (μέση τιμή, διασπορά) είναι γνωστές.

9.1. ΔΙΑΚΡΙΤΕΣ ΘΕΩΡΗΤΙΚΕΣ Τ.Μ.

9.1.1. Κατανομή Bernoulli

Εκτελούμε τυχαίο πείραμα και θεωρούμε p την πιθανότητα επιτυχίας και q την πιθανότητα αποτυχίας. Αντίστοιχα η διακριτή τ.μ. Χ ισούται με 1 στην επιτυχία και με 0 στην περίπτωση αποτυχίας.
Οπότε έχουμε τον πίνακα:

x_n	0	1
αντίστοιχη πιθανότητα	p	q

Παρατηρούμε ότι θα ισχύει $q + p = 1$
Τότε για τη συνάρτηση πιθανότητας θα είναι:

$$\boxed{P[X=x] = p^x q^{1-x}, \text{ για } x = 0, 1 \text{ και } q = 1 - p}$$

Τότε λέμε ότι η τ.μ. X ακολουθεί κατανομή Bernoulli
Επίσης για την κατανομή Bernoulli θα είναι:

$$\boxed{\mu = p}$$

Και

$$\boxed{\sigma^2 = pq}$$

Η ποσότητα p λέγεται παράμετρος της κατανομής και αρκεί για να προσδιορίσουμε την κατανομή και να υπολογίσουμε τις πιθανότητες της τ.μ. X.

Εφαρμογή 1

Έστω ότι η τ.μ. ακολουθεί κατανομή Bernoulli με παράμετρο p=0,3.
Τότε $P[X=x] = 0{,}3^x \cdot (1-0{,}3)^{1-x}$, για $x = 0,1$
Οπότε για x=0 έχουμε:

$$P[X=0] = 0{,}3^0 \cdot (1-0{,}3)^{1-0} = 0{,}3^0 \cdot (1-0{,}3)^{1-0} = 0{,}7$$

Και για x=1 έχουμε:

$$P[X=1] = 0{,}3^1 \cdot (1-0{,}3)^{1-1} = 0{,}3^1 \cdot (1-0{,}3)^{1-1} = 0{,}3$$

Η δοκιμή που εκτελούμε για να διαπιστώσουμε επιτυχία ή αποτυχία ονομάζεται δοκιμή Bernoulli.

Παραδείγματα Δοκιμών Bernoulli:

1ο: Ρίψη νομίσματος, θεωρώντας επιτυχία την ένδειξη κορόνα και αποτυχία την ένδειξη γράμματα ή αντίστροφα.

2ο: Συμμετοχή σε εξετάσεις με επιτυχία και αποτυχία τα όμοια στην εξέταση.

3ο: Απόφαση δικαστηρίου με επιτυχία την απόφαση αθώος και αποτυχία την απόφαση ένοχος ή αντίστροφα (για τον αντίδικο).

4ο: Ρίψη ζαριού θεωρώντας επιτυχία την ένδειξη 2 ή 4 ή 6 (ζυγό) και αποτυχία την ένδειξη 1 ή 3 ή 5 (μονό).

5ο: Η γέννηση παιδιού θεωρώντας επιτυχία το αγόρι και αποτυχία το κορίτσι ή αντίστροφα.

6ο: Τεστ θετικό ή αρνητικό για εγκυμοσύνη.

9.1.2. Διωνυμική κατανομή

Τα περισσότερα φαινόμενα μελετώνται με τη βοήθεια αυτής της κατανομής. Έστω ότι εξετάζουμε έναν πληθυσμό ως προς το αν έχει ή όχι ένα χαρακτηριστικό. Επιπλέον το αν παρουσιάζεται το χαρακτηριστικό σε κάποιο μέλος του πληθυσμού είναι ανεξάρτητο από το αν παρουσιάζεται σε οποιοδήποτε άλλο μέλος. Εκτελούμε δηλαδή δοκιμές Bernoulli. Τότε ο πληθυσμός χωρίζεται σε δύο ομάδες. Η πρώτη είναι αυτή των μελών που έχουν το χαρακτηριστικό και η δεύτερη αυτή που δεν το έχουν.

Ονομάζουμε p την πιθανότητα να παρουσιάζει μία μονάδα το χαρακτηριστικό και q την πιθανότητα να μην το παρουσιάζει. Ή αντίστοιχα επιτυχία και αποτυχία.

Στην περίπτωση που εκτελούμε περισσότερες από μία δοκιμές Bernoulli, έστω n, λέμε ότι η τ.μ. που ορίζεται ακολουθεί διωνυμική κατανομή και σημειώνουμε $X \sim B(n, p)$.

Για τη διωνυμική κατανομή ισχύει ότι:

$$\boxed{P[X=x] = \binom{n}{x} p^x q^{n-x},\ x = 0, 1, \dots, n}$$

Όπου

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{1 \cdot 2 \cdot 3 \dots \cdot (n-1) \cdot n}{[1 \cdot 2 \cdot 3 \dots \cdot (x-1) \cdot x] \cdot [1 \cdot 2 \cdot 3 \dots \cdot (n-x-1) \cdot (n-x)]}$$

Οι n και p ονομάζονται παράμετροι της διωνυμικής κατανομής.

Για τη διωνυμική κατανομή ισχύουν επίσης οι τύποι:

$\boxed{\mu = np}$ και $\boxed{\sigma^2 = npq}$

Εφαρμογή 2

Μία κάλπη περιέχει 7 άσπρα και 4 μαύρα σφαιρίδια. Εξάγουμε 6 σφαιρίδια με επανατοποθέτηση. Ποια η πιθανότητα να μην εμφανιστεί άσπρο σφαιρίδιο;

ΛΥΣΗ

Χρησιμοποιούμε την διωνυμική κατανομή γιατί σε κάθε εξαγωγή σφαιριδίου έχουμε την επιτυχία να είναι άσπρο και την αποτυχία να είναι μαύρο. Επίσης οι δοκιμές είναι μεταξύ τους ανεξάρτητες και σε κάθε δοκιμή οι πιθανότητες επιτυχίας και αποτυχίας παραμένουν σταθερές (αφού έχουμε επανατοποθέτηση των σφαιριδίων).

Η πιθανότητα στη διωνυμική κατανομή δίνεται από τον τύπο:

$$P[X=x] = \binom{n}{x} p^x q^{n-x}$$

Και στο συγκεκριμένο πείραμα θα είναι:

n =6 δοκιμές

x = 0 (κανένα άσπρο σφαιρίδιο)

$p = \frac{7}{10} = 0{,}7$ η πιθανότητα να βγάλω άσπρο σφαιρίδιο

$q = \frac{3}{10} = 0{,}3$ η πιθανότητα να βγάλω μαύρο σφαιρίδιο

άρα αντικαθιστώντας στον τύπο της διωνυμικής κατανομής θα είναι:

$$P[X=0] = \binom{n}{0} p^0 q^{n-0} = \binom{6}{0} 0{,}7^0 0{,}3^6 = 0{,}0072$$

Εφαρμογή 3

Ρίχνουμε ένα αμερόληπτο νόμισμα 10 φορές. Να υπολογιστεί η πιθανότητα να εμφανιστεί η ένδειξη κορόνα 3 φορές. Στη συνέχεια να υπολογιστεί η πιθανότητα να εμφανιστεί η ένδειξη κορόνα από 3 ως 7 φορές.

ΛΥΣΗ

Η ρίψη αμερόληπτου νομίσματος αποτελεί δοκιμή Bernoulli. Η επανάληψή της 10 φορές θα ακολουθεί ως φαινόμενο τη διωνυμική κατανομή, της οποίας η πιθανότητα γενικά δίνεται από τον τύπο:

$$P[X=x] = \binom{n}{x} p^x q^{n-x}$$

Βάσει του πειράματος που περιγράφεται θα είναι:

n = 10 (ρίψεις νομίσματος)

p = 0,5 (πιθανότητα να εμφανιστεί η ένδειξη κορόνα)

q = 0,5 (πιθανότητα να εμφανιστεί η ένδειξη γράμματα)

και x = 3 (να εμφανιστεί η ένδειξη κορόνα τρεις φορές)

οπότε θα έχουμε αντικαθιστώντας:

$$P[X=3] = \binom{n}{3} p^3 q^{n-3} = \binom{10}{3} 0{,}5^3 0{,}5^{10-3} = \frac{15}{178}$$

Για να υπολογίσουμε την πιθανότητα να εμφανιστεί κορόνα από 3 έως 7 φορές θα υπολογίσουμε την P[3<X<7] που δίνεται από τον τύπο:

$$P[3<X<7] = \sum_{i=3}^{7} p_i = [X=3] + P[X=4] + P[X=5] + P[X=6] + P[X=7]$$

Έχουμε:

n = 10 (ρίψεις νομίσματος)

p = 0,5 (πιθανότητα να εμφανιστεί η ένδειξη κορόνα)

q = 0,5 (πιθανότητα να εμφανιστεί η ένδειξη γράμματα)

Για x = 4:

$$P[X=4] = \binom{n}{4} p^4 q^{n-4} = \binom{10}{4} 0{,}5^4 0{,}5^{10-4} = \frac{105}{512}$$

Για x = 5:

$$P[X=5] = \binom{n}{5} p^5 q^{n-5} = \binom{10}{5} 0{,}5^5 0{,}5^{10-5} = \frac{176}{512}$$

Για x = 6:

$$P[X=6] = \binom{n}{6} p^6 q^{n-6} = \binom{10}{6} 0,5^6 0,5^{10-6} = \frac{105}{512}$$

Για x = 7:

$$P[X=7] = \binom{n}{7} p^7 q^{n-7} = \binom{10}{7} 0,5^7 0,5^{10-7} = \frac{15}{178}$$

Άρα είναι:

$$P[3<X<7] = \sum_{i=3}^{7} p_i = [X=3] + P[X=4] + P[X=5] + P[X=6] + P[X=7] =$$

$$= \frac{60}{512} + \frac{105}{512} + \frac{176}{512} + \frac{105}{512} + \frac{60}{512} = \frac{506}{512}$$

Εφαρμογή 4

Μία κάλπη περιέχει 10 σφαιρίδια 6 άσπρες και 4 μαύρες. Εξάγουμε πέντε σφαίρες με επανατοποθέτηση. Ποια η πιθανότητα:

1) να μην εμφανιστεί άσπρο σφαιρίδιο

2) να εμφανιστεί ένα άσπρο σφαιρίδιο

3) να εμφανιστούν δύο άσπρα σφαιρίδια

4) να εμφανιστούν τρία άσπρα σφαιρίδια

5) να εμφανιστούν τέσσερα άσπρα σφαιρίδια

6) να εμφανιστούν πέντε άσπρα σφαιρίδια

Να γίνει το αντίστοιχο διάγραμμα.

ΛΥΣΗ

1)

- ✔ έχουμε δοκιμή Bernoulli, αφού ένα σφαιρίδιο θα είναι άσπρο ή μαύρο ανεξαρτήτως του τι είναι τα υπόλοιπα σφαιρίδια
- ✔ θεωρούμε επιτυχία να είναι άσπρο και αποτυχία να είναι μαύρο
- ✔ Επειδή αναφερόμαστε συγκεκριμένα σε πέντε σφαιρίδια που θα εξάγουμε, οι δοκιμές Bernoulli είναι προκαθορισμένες ως προς τον αριθμό,δηλαδή πέντε

Επομένως το φαινόμενο ακολουθεί διωνυμική κατανομή με:

- ✔ $p = \frac{6}{10}$ (πιθανότητα επιτυχίας,δηλαδή εξαγωγή άσπρου σφαιριδίου)
- ✔ $q = \frac{4}{10}$ αφού $q = 1 - p = 1 - \frac{6}{10} = \frac{4}{10}$
- ✔ n = 5, αφού τόσα είναι συνολικά τα σφαιρίδια που εξάγονται
- ✔ x = 0, αφού θέλουμε να μην υπάρχει άσπρο σφαιρίδιο

Η συνάρτηση πιθανότητας της διωνυμικής κατανομής είναι

$$P[X=x] = \binom{n}{x} p^x q^{n-x}$$

Άρα αντικαθιστώντας θα έχουμε:

$$P[X=0] = \binom{n}{0} p^0 q^{n-0} = \binom{5}{0} \frac{6^0}{10} \left(\frac{4}{10}\right)^{5-0} = \frac{5!}{0!(5-0)!} \cdot \left(\frac{4}{10}\right)^5 = \frac{256}{25000}$$

2) θα ισχύουν τα του 1ου ερωτήματος αλλά με x = 1, αφού θέλουμε ένα άσπρο σφαιρίδιο άρα:

$$P[X=1] = \binom{n}{1} p^1 q^{n-1} = \binom{5}{1} \left(\frac{6}{10}\right)^1 \left(\frac{4}{10}\right)^{5-1} = \frac{5!}{1!(5-1)!} \cdot \frac{6}{10} \cdot \left(\frac{4}{10}\right)^4 = \frac{48}{625}$$

3) θα ισχύουν τα του 1ου ερωτήματος αλλά με x = 2, αφού θέλουμε δύο άσπρα σφαιρίδια άρα:

$$P[X=2] = \binom{n}{2} p^2 q^{n-2} = \binom{5}{2} \left(\frac{6}{10}\right)^2 \left(\frac{4}{10}\right)^{5-2} = \frac{5!}{2!(5-2)!} \cdot \left(\frac{6}{10}\right)^2 \cdot \left(\frac{4}{10}\right)^3 = \frac{36}{25000}$$

4) θα ισχύουν τα του 1ου ερωτήματος αλλά με x = 3, αφού θέλουμε τρία άσπρα σφαιρίδια άρα:

$$P[X=3] = \binom{n}{3} p^3 q^{n-3} = \binom{5}{3} \left(\frac{6}{10}\right)^3 \left(\frac{4}{10}\right)^{5-3} = \frac{5!}{3!(5-3)!} \cdot \left(\frac{6}{10}\right)^3 \cdot \left(\frac{4}{10}\right)^2 = \frac{216}{625}$$

5) θα ισχύουν τα του 1ου ερωτήματος αλλά με x = 4, αφού θέλουμε τέσσερα άσπρα σφαιρίδια άρα:

$$P[X=4] = \binom{n}{4} p^4 q^{n-4} = \binom{5}{4} \left(\frac{6}{10}\right)^4 \left(\frac{4}{10}\right)^{5-4} = \frac{5!}{4!(5-4)!} \cdot \left(\frac{6}{10}\right)^4 \cdot \left(\frac{4}{10}\right)^1 = \frac{162}{625}$$

6) θα ισχύουν τα του 1ου ερωτήματος αλλά με x = 5, αφού θέλουμε πέντε άσπρα σφαιρίδια άρα:

$$P[X=5] = \binom{n}{5} p^5 q^{n-5} = \binom{5}{5} \left(\frac{6}{10}\right)^5 \left(\frac{4}{10}\right)^{5-5} = \frac{5!}{5!(5-5)!} \cdot \left(\frac{6}{10}\right)^5 \cdot \left(\frac{4}{10}\right)^1 = \frac{243}{6250}$$

9.1.3. Γεωμετρική κατανομή

Στην περίπτωση της διωνυμικής κατανομής το ερώτημα ήταν «ποια η πιθανότητα x επιτυχιών σε n δοκιμές Bernoulli». Αν όμως το ερώτημα είναι «ποια η πιθανότητα να έχουμε επιτυχία στην δοκιμή x εκτελώντας πάλι n δοκιμές Bernoulli», δεν δίνει λύση η διωνυμική κατανομή, αλλά η Γεωμετρική.

Θεωρούμε:

n δοκιμές Bernoulli

p την πιθανότητα επιτυχίας

q την πιθανότητα αποτυχίας

και x η πρώτη επιτυχημένη δοκιμή.

Σημειώνεται ότι οι δοκιμές 1, 2, ..., x – 1 είναι αποτυχημένες δοκιμές.

Οπότε για τη γεωμετρική κατανομή θα είναι:

$$P[X=x] = P[\underbrace{AA..A}_{x-1 \text{ αποτυχίες}}\underbrace{E}_{1\text{η επιτυχία}}] = \underbrace{P(A)P(A)...P(A)}_{x-1}P(E) = \underbrace{qq...q}_{x-1} \cdot p = q^{x-1}p$$

δηλαδή

$$\boxed{P[X=x] = q^{x-1}p}$$

επιπλέον ισχύουν:

$$\boxed{\mu = \frac{1}{p}} \quad \text{και} \quad \boxed{\sigma^2 = \frac{q}{p^2}}$$

Η p είναι η παράμετρος της Γεωμετρικής κατανομής.

Λέμε ότι η X ακολουθεί γεωμετρική κατανομή και σημειώνουμε $X \sim \Gamma(p)$

Εφαρμογή 5

Ποια η πιθανότητα να φέρουμε κορόνα στην τέταρτη ρίψη αμερόληπτου νομίσματος.

ΛΥΣΗ

Αν Κ κεφάλι και Γ γράμματα τότε:

$P(K) = \frac{1}{2}$ και $P(\Gamma) = \frac{1}{2}$

Άρα $p = q = \frac{1}{2}$

Και x=4, αφού στην 4η δοκιμή θα πετύχουμε γράμματα

Η τ.μ. X συμβολίζει την δοκιμή όπου εμφανίζεται η πρώτη επιτυχία και ακολουθεί γεωμετρική κατανομή με $p = \frac{1}{2}$ οπότε είναι:

$$P[X=4] = p^{4-1} \cdot q = \left(\frac{1}{2}\right)^{4-1} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{3} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{4} = \frac{1}{16}$$

Σχόλιο:

Σύγκριση Διωνυμικής και Γεωμετρικής κατανομής:

Διωνυμική κατανομή	Γεωμετρική κατανομή
Προκαθορισμένος αριθμός δοκιμών Bernoulli (n)	Επανάληψη δοκιμών Bernoulli μέχρι την πρώτη επιτυχία

Εφαρμογή 6

Υποθέτουμε ότι η πιθανότητα να γεννηθεί κορίτσι σε μία οικογένεια είναι $\frac{1}{2}$.

Να βρεθούν:

1) η πιθανότητα σε πέντε παιδιά τα 3 να είναι κορίτσια

2) η πιθανότητα το τρίτο παιδί να είναι κορίτσι και τα δύο προηγούμενα αγόρια

ΛΥΣΗ

1)

- έχουμε δοκιμή Bernoulli, αφού ένα παιδί θα είναι αγόρι ή κορίτσι ανεξαρτήτως του τι είναι τα υπόλοιπα παιδιά
- θεωρούμε επιτυχία να είναι κορίτσι και αποτυχία να είναι αγόρι
- Επειδή αναφερόμαστε συγκεκριμένα σε πέντε παιδιά, οι δοκιμές Bernoulli είναι προκαθορισμένες ως προς τον αριθμό, δηλαδή πέντε

Επομένως το φαινόμενο ακολουθεί διωνυμική κατανομή με

- $p = \frac{1}{2}$ (πιθανότητα επιτυχίας,δηλαδή γέννησης κοριτσιού)
- $q = \frac{1}{2}$, αφού $q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$
- $x = 3$, αφού θέλουμε να γεννηθούν τρία κορίτσια
- $n = 5$, αφού τόσα είναι συνολικά τα παιδιά

Η συνάρτηση πιθανότητας της διωνυμικής κατανομής είναι

$$P[X=x] = \binom{n}{x} p^x q^{n-x}$$

Άρα αντικαθιστώντας θα έχουμε:

$$P[X=3] = \binom{n}{3} p^3 q^{n-3} = \binom{5}{3} \frac{1^3}{2} \frac{1^{5-3}}{2} = \frac{5!}{3!(5-3)!} \cdot \frac{1^5}{2} = \frac{5}{16}$$

2)

- έχουμε δοκιμή Bernoulli, αφού ένα παιδί θα είναι αγόρι ή κορίτσι ανεξαρτήτως του τι είναι τα υπόλοιπα παιδιά
- θεωρούμε επιτυχία να είναι κορίτσι και αποτυχία να είναι αγόρι
- Επειδή σταματάμε τις δοκιμές Bernoulli μόλις γεννηθεί κορίτσι, ο αριθμός των δοκιμών δεν είναι προκαθορισμένος.

Επομένως το φαινόμενο ακολουθεί γεωμετρική κατανομή

- $p = \frac{1}{2}$ (πιθανότητα επιτυχίας,δηλαδή γέννησης κοριτσιού)
- $q = \frac{1}{2}$, αφού $q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$
- $x = 3$, αφού θέλουμε το τρίτο παιδί να γεννηθεί κορίτσι

Η συνάρτηση πιθανότητας της γεωμετρικής κατανομής είναι

$$P[X = x] = p^x q$$

Άρα αντικαθιστώντας θα έχουμε:

$$P[X=3] = p^3 \cdot q = \left(\frac{1}{2}\right)^3 \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

9.1.4. Κατανομή Poisson

Έχει παρατηρηθεί ότι φαινόμενα όπως:

α) η εμφάνιση σεισμών,

β) η προσέλευση πελατών σε ταμείο,

γ) ο αριθμός ατυχημάτων σε συγκεκριμένο σημείο,

δ) ο αριθμός θανάτων σε συγκεκριμένη περιοχή,

ε) ο αριθμός των τηλεφωνικών κλήσεων και

στ) ο αριθμός των τυπογραφικών σφαλμάτων

όλα σε συγκεκριμένο χρονικό διάστημα ακολουθούν τη λεγόμενη κατανομή Poisson.

Γενικά θα λέγαμε ότι η κατανομή Poisson χρησιμοποιείται όταν:

- ✔ έχουμε ως τ.μ. μία διακριτή τ.μ. που μετράει τη συχνότητα εμφάνισης σπάνιων ενδεχομένων μέσα σε συγκεκριμένο χρονικό διάστημα
- ✔ η εμφάνιση κάποιου ενδεχομένου είναι ανεξάρτητη της εμφάνισης άλλου ενδεχομένου μέσα στο συγκεκριμένο χρονικό διάστημα όπου αναφερόμαστε
- ✔ η πιθανότητα να εμφανιστεί οποιοδήποτε ενδεχόμενο είναι ίση με την πιθανότητα να εμφανιστεί οποιοδήποτε άλλο ενδεχόμενο
- ✔ τα ενδεχόμενα εμφανίζονται μεμονωμένα, δηλαδή όχι ταυτόχρονα

Τότε λέμε ότι η Χ ακολουθεί κατανομή Poisson και σημειώνουμε ότι $X \sim P(\lambda)$

Στην περίπτωση της κατανομής Poisson η συνάρτηση πιθανότητας είναι:

$$\boxed{P[X = x] = \frac{\lambda^x e^{-\lambda}}{x!}, \text{ όπου } x = 0, 1, 2, \dots \text{ και } \lambda>0}$$

Το λ είναι η παράμετρος της κατανομής oisson.

Επίσης είναι:

$$\boxed{\mu = \lambda} \quad \text{και} \quad \boxed{\sigma^2 = \lambda}$$

Σημείωση: η ιδιότητα $\mu = \sigma^2 = \lambda$ είναι χαρακτηριστική της κατανομής Poisson.

Εφαρμογή 7

Σε 200 σελίδες ενός βιβλίου περιέχονται 25 τυπογραφικά λάθη. Αν υποθέτουμε ότι τα λάθη αυτά ακολουθούν κατανομή Bernoulli, να βρεθεί:

1) η κατανομή πυκνότητας θεωρώντας τυχαία μεταβλητή τον αριθμό λαθών σε μια σελίδα του βιβλίου

2) η πιθανότητα σε μία σελίδα να έχουμε 2 λάθη.

ΛΥΣΗ

1) Έχουμε $f(x) = P(X=x) = e^{-\lambda} \cdot \dfrac{\lambda^x}{x!}$, $x = 0, 1, 2, \ldots$ την κατανομή Bernoulli.

Η παράμετρος λ για το συγκεκριμένο παράδειγμα θα είναι

$$\lambda = \frac{25}{200}$$

$$\lambda = 0{,}125$$

Άρα $f(x) = e^{-0{,}125} \cdot \dfrac{0{,}125^x}{x!}$, $x = 0, 1, 2, \ldots$

2) Θέλουμε να βρούμε την $P(x=4)$

Έχουμε:

$$P(x=2) = f(2) = e^{-0{,}125} \cdot \frac{0{,}125^2}{2!}$$

δηλαδή $P(x=2) \cong 0{,}6\%$

Προσέγγιση διωνυμικής κατανομής από κατανομή Poisson:

Πολλές φορές η διωνυμική κατανομή προσεγγίζεται από την Poisson με $\lambda = np$, όπου λ η παράμετρος την Poisson και n, p οι παράμετροι της διωνυμικής κατανομής.

Προκειμένου να γίνει αυτή η προσέγγιση πρέπει

1) $n > 50$ και $p < 0.1$

2) όσο μικρότερη είναι η τιμή της p τόσο καλύτερη είναι η προσέγγιση

3) όσο μεγαλύτερο το n τόσο καλύτερη η προσέγγιση

Εφαρμογή 8

Αν το ποσοστό αυτοκτονίας σε μία χώρα είναι κάθε μήνα 3 άτομα στο 1.000.000, να βρεθεί η πιθανότητα σε μια πόλη των 2.000.000 κατοίκων

Α) να παρατηρηθούν 2 αυτοκτονίες σε ένα μήνα

Β) να παρατηρηθούν το πολύ 2 αυτοκτονίες σε ένα μήνα

ΛΥΣΗ

Πρόκειται για θανάτους σε συγκεκριμένο μέρος για συγκεκριμένο χρονικό διάστημα. Θεωρώντας το φαινόμενο ακολουθεί την κατανομή Poisson.

Η παράμετρος p θα είναι:

$$p = \frac{3}{10^6}$$

και

$$n = 2.000.000$$

άρα υπολογίζουμε την παράμετρο λ από τον τύπο $\lambda = np$, από όπου με αντικατάσταση θα έχουμε!

$$\lambda = \frac{3}{10^6} \cdot 2.000.000 = \frac{3}{10^6} \cdot 2 \cdot 10^6 = 6$$

Α) θέλουμε την πιθανότητα να έχουμε 2 αυτοκτονίες σε ένα μήνα δηλαδή την πιθανότητα P[X=2]

Επομένως από τον τύπο της συνάρτησης πιθανότητας της κατανομής Poisson:

$$P[X = x] = \frac{\lambda^x e^{-\lambda}}{x!}$$

αντικαθιστώντας έχουμε:

$$P[X = 2] = \frac{\lambda^2 e^{-\lambda}}{2!} = \frac{6^2 e^{-6}}{2!} = 18e^{-6}$$

Β) θέλουμε την πιθανότητα να παρατηρηθούν το πολύ δύο αυτοκτονίες, άρα την P

$$P[X \leq 2] = P(X = 0) + P(X = 1) + P(X = 2)$$

Με αντικατάσταση στον τύπο της συνάρτησης της πιθανότητας:

$$P[X = x] = \frac{\lambda^x e^{-\lambda}}{x!}$$

Για $X = 0$:

$$P[X = 0] = \frac{\lambda^0 e^{-\lambda}}{0!} = \frac{6^0 e^{-6}}{0!} = e^{-6}$$

Για $X = 1$:

$$P[X = 1] = \frac{\lambda^1 e^{-\lambda}}{1!} = \frac{6^1 e^{-6}}{1!} = 6e^{-6}$$

Για $X = 2$:

$P[X = 2] = 18e^{-6}$ (από το προηγούμενο ερώτημα)

Άρα:

$$P[X \leq 2] = P(X = 0) + P(X = 1) + P(X = 2) = e^{-6} + 6e^{-6} + 18e^{-6} = 25e^{-6}$$

9.1.5. Υπεργεωμετρική Κατανομή

Έστω ότι έχουμε μία κάλπη με συνολικά 10 σφαιρίδια. Από αυτά τα 4 είναι άσπρα και τα 6 μαύρα έτσι, ώστε $4 + 6 = 10$.

Υποθέτουμε ότι από την κάλπη αυτή βγάζουμε 3 σφαιρίδια χωρίς επανάθεση.

Η πιθανότητα από τα 3 σφαιρίδια να βγάλουμε x λευκά και 3–x μαύρα σφαιρίδια υπολογίζεται ως εξής:

$$P[X=x] = \frac{\binom{4}{x}\binom{6}{3-x}}{\binom{10}{3}}$$

Γενικά:

Έστω ότι έχουμε μία κάλπη με συνολικά N σφαιρίδια. Από αυτά τα N_1 είναι άσπρα και τα N_2 μαύρα έτσι, ώστε $N_1 + N_2 = N$.

Υποθέτουμε ότι από την κάλπη αυτή βγάζουμε n σφαιρίδια χωρίς επανάθεση.

Η πιθανότητα από τα n σφαιρίδια να βγάλουμε x λευκά και n–x μαύρα σφαιρίδια υπολογίζεται ως εξής:

$$\boxed{P[X=x] = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}}$$

Και λέμε ότι η τ.μ. X ακολουθεί την υπεργεωμετρική κατανομή.

Όπως φαίνεται και από τον τύπο της συνάρτησης πιθανότητας, μας ενδιαφέρει να γνωρίζουμε:

1) το μέγεθος N του πληθυσμού που ερευνούμε

2) το μέγεθος n του δείγματος

3) το ποσοστό $p = \frac{N_1}{N}$.

Επίσης για την υπεργεωμετρική κατανομή έχουμε:

$$\boxed{\mu = np} \quad \text{και} \quad \boxed{\sigma^2 = npq\,\frac{N-n}{N-1}}$$

Εφαρμογή 9

Το Δ.Σ. ενός συλλόγου αποτελείται από 7 γυναίκες και 8 άντρες. Για ένα συγκεκριμένο λόγο συγκροτείται μία επιτροπή 4 ατόμων. Ποια η πιθανότητα η επιτροπή να αποτελείται από 2 άντρες και 2 γυναίκες;

ΛΥΣΗ

Έχουμε συνολικά 7 + 8 = 15 άτομα και επιλέγουμε 4 από αυτά.

Τα άτομα χωρίζονται σε δύο κατηγορίες: άντρες και γυναίκες (όπως είχαμε άσπρα και μαύρα σφαιρίδια προηγουμένως)

Η διακριτή τ.μ. που εκφράζει τον αριθμό των γυναικών που θα συμπεριλαμβάνονται στην ομάδα των 4 ακολουθεί υπεργεωμετρική κατανομή με συνάρτηση πιθανότητας:

$$P[X=x] = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$$

Όπου $N = 15$, $N_1 = 7$, $N_2 = 8$ και $x = 2$

Επομένως με αντικατάσταση στον παραπάνω τύπο θα έχουμε:

$$P[X=x] = \frac{\binom{7}{2}\binom{8}{4-2}}{\binom{15}{4}} = \frac{\binom{7}{2}\binom{8}{2}}{\binom{15}{4}} =$$

$$= \frac{\frac{7!}{2!(7-2)!}\cdot\frac{8!}{2!(8-2)!}}{\frac{15!}{4!(15-4)!}} = \frac{28}{65}$$

Σημείωση
Σε περίπτωση που το δείγμα είναι πολύ μικρό σε σχέση με τον πληθυσμό, αντί της υπεργεωμετρικής, χρησιμοποιούμε τη διωνυμική κατανομή.

Ειδικές συνεχείς κατανομές:

Στη συνέχεια θα αναφερθούμε στις κυριότερες κατανομές συνεχών τυχαίων μεταβλητών οι οποίες έχουν τεράστια εφαρμογή σε πολλούς τομείς της καθημερινότητάς μας.

9.2.1. Κανονική Κατανομή

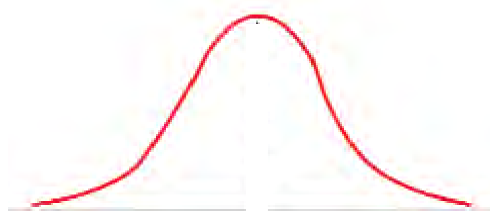
Η κανονική κατανομή είναι η σπουδαιότερη για τη στατιστική, τη δειγματοληψία και τις πιθανότητες κατανομή. Μελετήθηκε αρχικά από τον de Moivre και στη συνέχεια από τους Gauss και Laplace.
Έστω συνεχής τ.μ. που ακολουθεί κανονική κατανομή. Τότε η συνάρτηση πυκνότητας της πιθανότητας είναι:

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

όπου μ η μέση τιμή της κατανομής.

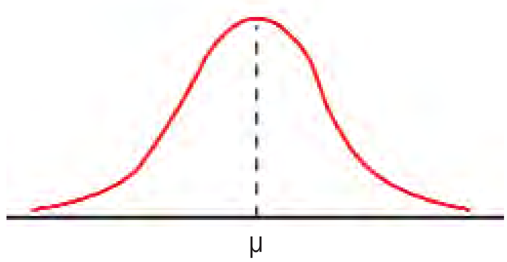
Παρατηρήσεις:

1) Η καμπύλη της κανονικής κατανομής έχει σχήμα καμπάνας:



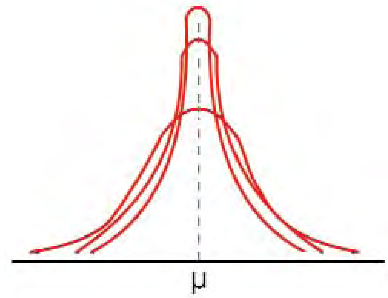
2) Το μ καθορίζει τη θέση της καμπύλης της κανονικής κατανομής και παίρνει τιμές από – ∞ μέχρι + ∞.

- Διατηρώντας σταθερό το σ και μεταβάλλοντας το μ θα έχουμε:

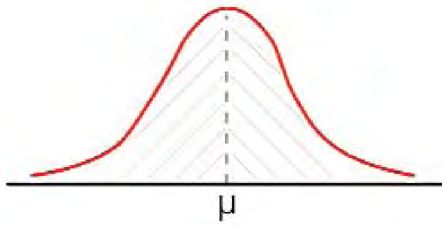


3) Το σ είναι η τυπική απόκλιση της τ.μ. Χ και καθορίζει το σχήμα της καμπύλης.

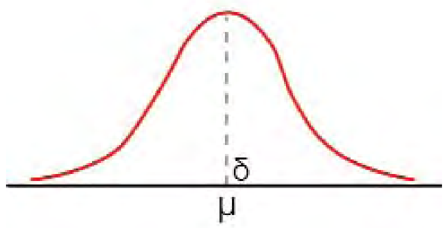
• Διατηρώντας σταθερό το μ και μεταβάλλοντας το σ θα έχουμε:



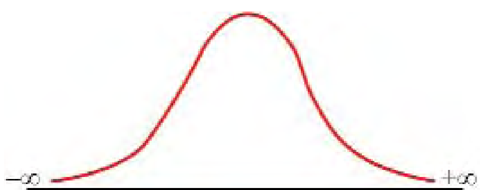
4) Η καμπύλη της κανονικής κατανομής είναι συμμετρική ως προς το μ



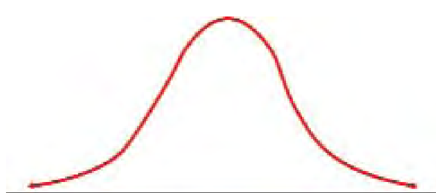
5) Η επικρατούσα τιμή και η διάμεσος ταυτίζονται λόγω συμμετρίας



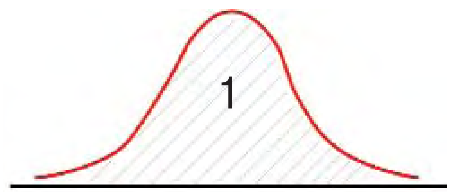
6) Η τ.μ. Χ παίρνει τιμές σε όλο το R, δηλαδή μπορεί να είναι οποιοσδήποτε πραγματικός αριθμός.



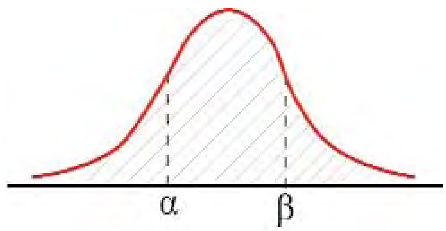
7) Η καμπύλη της κανονικής κατανομής δεν ακουμπά ποτέ τον οριζόντιο άξονα, ενώ τον πλησιάζει καθώς το x παίρνει πολύ μεγάλες ή πολύ μικρές τιμές. Στα μαθηματικά λέμε ότι ο οριζόντιος άξονας είναι στην περίπτωση αυτή ασύμπτωτος της καμπύλης, καθώς το x τείνει στο + ∞ ή – ∞ αντίστοιχα.



8) Το συνολικό εμβαδό που βρίσκεται ανάμεσα στην καμπύλη της κανονικής κατανομής και τον οριζόντιο άξονα των x είναι ίσο με τη μονάδα, καθώς ολόκληρο το εμβαδό αντιστοιχεί στην πιθανότητα να συμβεί οποιοδήποτε ενδεχόμενο ή η τ.μ. X να πάρει οποιαδήποτε πραγματική τιμή, κάτι που είναι βέβαιο.



9) Αν θέλουμε να εξετάσουμε την πιθανότητα η τ.μ. να βρίσκεται μεταξύ των τιμών α και β, θέλουμε την πιθανότητα $P[\alpha<X<\beta]$, η οποία είναι ίση με το εμβαδό που βρίσκεται ανάμεσα στην καμπύλη της κανονικής κατανομής, τον οριζόντιο άξονα και τους κατακόρυφους άξονες που περνάνε από τα α, β.



Αλλά είναι:

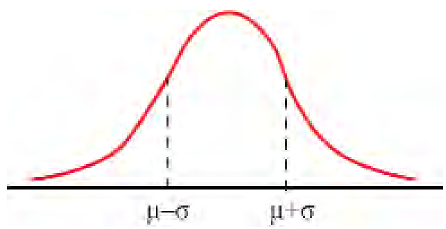
$P[\alpha<X<\beta] = E_1 - E_2 = F(\beta) - F(\alpha)$

Πιο συγκεκριμένα η εύρεση των πιθανοτήτων της μορφής $P[\alpha<X<\beta]$ γίνεται με τη βοήθεια πινάκων, όπως θα δούμε παρακάτω.

10) Η καμπύλη της κατανομής παρουσιάζει μέγιστο για:

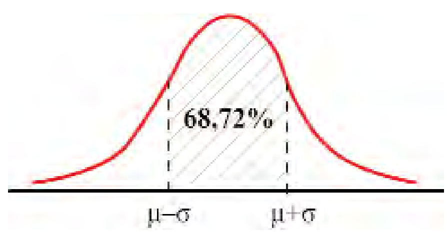
$$x = \frac{1}{\sigma\sqrt{2\pi}}$$

και σημεία καμπής για $x=\mu-\sigma$ και $x=\mu+\sigma$

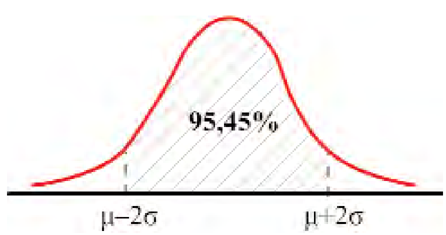


11)

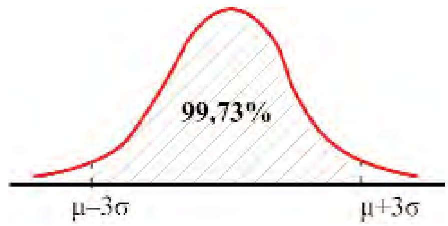
✔ σε απόσταση σ γύρω από το μ, ή αλλιώς όταν $\mu - \sigma < x < \mu + \sigma$ περιλαμβάνει το 68,72% του συνολικού εμβαδού της κανονικής καμπύλης



✔ σε απόσταση 2μ γύρω από το μ περιλαμβάνεται το 95,45% του συνολικού εμβαδού της κανονικής καμπύλης



✔ σε απόσταση 3μ γύρω από το μ περιλαμβάνεται το 99,73% του συνολικού εμβαδού της κανονικής κατανομής



Όπως φαίνεται από τα παραπάνω υπάρχουν άπειρες κανονικές κατανομές, αφού για κάθε ζεύγος διαφορετικών μ και σ η κατανομή αλλάζει.

Τυποποιημένη κανονική κατανομή:

Στην περίπτωση που η συνεχής τ.μ. X ακολουθεί κατανομή με παραμέτρους μ=0 και σ=1 λέμε ότι ακολουθεί τυποποιημένη κανονική κατανομή.

Ισχύει το εξής πολύ σημαντικό για τη στατιστική:

Αν η συνεχής τ.μ. X έχει μέση τιμή μ και διακύμανση σ^2, τότε η τ.μ. $Z = \dfrac{X-\mu}{\sigma}$ ακολουθεί τυποποιημένη κανονική κατανομή και σημειώνουμε $Z \sim N(0,1)$

Ή

Αν $X \sim N(\mu, \sigma^2)$ τότε $\dfrac{X-\mu}{\sigma} \sim N(0, 1)$

Αυτό στην πράξη σημαίνει ότι οποιαδήποτε κανονική κατανομή μας δώσουν, μπορούμε να την μετασχηματίσουμε σε τυποποιημένη κανονική κατανομή.

Παράδειγμα:

Δίνεται η τ.μ. X που ακολουθεί κανονική κατανομή με μ=2 και σ=3.

Τη μετασχηματίζουμε ως εξής:

Θεωρούμε την τ.μ. Z όπου $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-2}{3}$

Τότε η Z θα ακολουθεί τυποποιημένη κανονική κατανομή, δηλαδή κανονική κατανομή με μ=0 και σ=1

Για τη συνάρτηση πυκνότητας πιθανότητας της τυποποιημένης κανονικής αντικαθιστούμε στον τύπο $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ της κανονικής όπου μ το 0 και όπου σ το 1

και βρίσκουμε:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z}{2}\cdot 2} \quad , \text{με} -\infty < Z < +\infty$$

στην περίπτωση της κανονικής κατανομής είχαμε: P[α<X<β] = F(α), όπου F η συνάρτηση κατανομής της τ.μ. X

στην περίπτωση της τυποποιημένης κατανομής χρησιμοποιείται το γράμμα Φ και γράφουμε:

$$P[\alpha<Z<\beta] = \Phi(\beta) - \Phi(\alpha)$$

Οι τιμές της Φ δεν υπολογίζονται, αλλά δίνονται έτοιμες σε πίνακα (στο τέλος του βιβλίου)

Εφαρμογή 10

Κάτι ιδιαίτερα διαδεδομένο στην ψυχολογία τα τελευταία χρόνια είναι τα τεστ IQ, δηλαδή τεστ νοημοσύνης. Σε δείγμα 100 ατόμων κάναμε το ίδιο τεστ IQ και καταλήξαμε στον ακόλουθο πίνακα:

IQ	Αριθμός ατόμων
$IQ < 70$	5
$70 \leq IQ < 85$	10
$85 \leq IQ < 100$	38
$100 \leq IQ < 115$	30
$115 \leq IQ < 130$	13
$130 \leq IQ$	1
σύνολο	100

Τα αποτελέσματα ακολουθούν κανονική κατανομή με $\mu=100$ και $\sigma^2=225$.

Χρησιμοποιώντας την τυποποιημένη κανονική κατανομή, να βρείτε:

Α) την πιθανότητα κάποιο άτομο να έχει δείκτη IQ μικρότερο του 115

Β) την πιθανότητα κάποιο άτομο να έχει δείκτη IQ μικρότερο του 85

ΛΥΣΗ

Α) Θεωρούμε το δείκτη IQ ως τυχαία μεταβλητή X με $X \sim N(100, 225)$

Τότε

$$P(X \leq 115) = P\left(\frac{X-\mu}{\sigma} \leq \frac{115-\mu}{\sigma}\right) \text{ αντικαθιστούμε το μ και το σ}$$

$$= P\left(\frac{X-100}{15} \leq \frac{115-100}{15}\right) \text{ κάνουμε πράξεις}$$

$$= P\left(\frac{X-100}{15} \leq 1\right) \text{ θέτουμε } Z = \frac{115-100}{15}$$

$$= P\,(Z \leq 1)$$

$$= 0{,}84$$

Β) Θεωρούμε ομοίως με το πρώτο ερώτημα το δείκτη IQ ως τυχαία μεταβλητή X με $X \sim N(100, 225)$

Τότε

$$P(X \leq 115) = P\left(\frac{X-\mu}{\sigma} \leq \frac{85-\mu}{\sigma}\right) \text{ αντικαθιστούμε το μ και το σ}$$

$$= P\left(\frac{X-100}{15} \leq \frac{85-100}{15}\right) \text{ κάνουμε πράξεις}$$

$$= P\left(\frac{X-100}{15} \leq 1\right) \text{ θέτουμε } Z = \frac{X-100}{15}$$

$$= P\,(Z \leq -1)$$

Εφερμογή 11

Αν η τ.μ. X ακολουθεί κανονική κατανομή με μ=10 και σ=30 (X ~ N(10,900)), να βρεθεί η πιθανότητα $P[15 \le X \le 20]$

ΛΥΣΗ

Θεωρούμε την τ.μ. Z με $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-10}{30}$ (με αντικατάσταση μ=10 και σ=30)

η οποία ακολουθεί τυποποιημένη κανονική κατανομή (Z ~ N(0, 1)

Οπότε η ζητούμενη πιθανότητα γράφεται:

$$P(15 \le X \le 20) = P\left[\frac{15-10}{30} \le \frac{X-10}{30} \le \frac{20-10}{30}\right] =$$

$$= P\left[\frac{15-10}{30} \le Z \le \frac{20-10}{30}\right] =$$

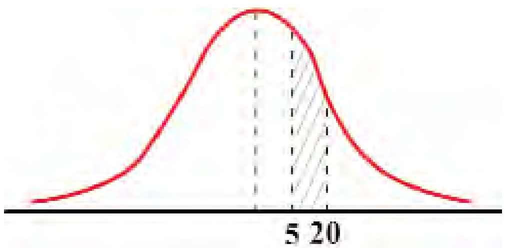
$$= P\,[0{,}333 \le Z \le 0{,}666] =$$

$$= \Phi\,(0{,}666) - \Phi\,(0{,}333)$$

κοιτάμε στον πίνακα τις τιμές Φ(0,666) και Φ(0,333) και με αντικατάσταση βρίσκουμε:

$P[5 < X < 20] = 0{,}745 - 0{,}63 = 0{,}115$ περίπου

Αντίστοιχα το γράφημα θα είναι:



Όπου το γραμμοσκιασμένο εμβαδό είναι η ζητούμενη πιθανότητα

Εφαρμογή 12

Αν η τ.μ. X ακολουθεί κανονική κατανομή με μ=100 και σ=30 (X ~ N(10,900)), να βρεθεί η πιθανότητα $P[150 \le X \le 200]$

ΛΥΣΗ

Θεωρούμε την τ.μ. Z με $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-100}{30}$ (με αντικατάσταση μ=10 και σ=30)

η οποία ακολουθεί τυποποιημένη κανονική κατανομή (Z ~ N(0, 1))

Οπότε η ζητούμενη πιθανότητα γράφεται:

$$P(150 < X < 200) = P\left[\frac{150-100}{30} \le \frac{X-100}{30} \le \frac{200-100}{30}\right] =$$

$$= P\left[\frac{150-100}{30} \le Z \le \frac{200-100}{30}\right] =$$

$$= P\,[1{,}666 \le Z \le 3{,}333] = \Phi\,(3{,}333) - \Phi\,(1{,}666)$$

κοιτάμε στον πίνακα τις τιμές Φ(1,666) και Φ(3,333) και με αντικατάσταση βρίσκουμε:

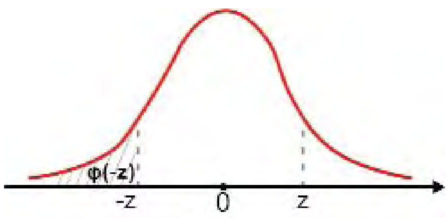
$$P[150 < X < 200] = 0{,}99 - 0{,}856 = 0{,}034 \text{ περίπου}$$

Παρατήρηση:

Από το γράφημα της τυποποιημένης κανονικής κατανομής παρατηρούμε ότι

$$\Phi(-z) = 1 - \Phi(z)$$

Κάτι χρήσιμο, καθώς στους πίνακες συνήθως μας δίνονται οι τιμές της Φ για θετικά z



Εφαρμογή 13

Αν η τ.μ. X ακολουθεί κανονική κατανομή με μ = 6 και σ = 2 (X ~ N(6, 4)), να βρεθεί η πιθανότητα P [4 < X < 10]

ΛΥΣΗ

Θεωρούμε την τ.μ. Z με $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-6}{2}$ (με αντικατάσταση μ=6 και σ=2)

η οποία ακολουθεί τυποποιημένη κανονική κατανομή (Z ~ N(0, 1))

Οπότε η ζητούμενη πιθανότητα γράφεται:

$$P[4 < X < 10] = P\left[\frac{4-6}{2} < \frac{X-6}{2} < \frac{10-6}{2}\right] =$$

$$= P\left[\frac{4-6}{2} < Z < \frac{10-6}{2}\right] =$$

$$= P[-1 < Z < 2] =$$

$$= \Phi(2) - \Phi(-1) =$$

$$= \Phi(2) - [1 - \Phi(1)] =$$

$$= \Phi(2) - 1 + \Phi(1)$$

κοιτάμε στον πίνακα τις τιμές Φ(1) και Φ(2) και με αντικατάσταση βρίσκουμε:

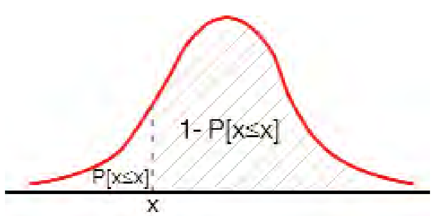
$$P[4 < X < 10] = 0{,}9772 - 0{,}8413 = 0{,}1359$$

Παρατήρηση:

Σε περιπτώσεις που θέλουμε να βρούμε πιθανότητα της μορφής P[X > x] έχουμε από το γράφημα ότι η ζητούμενη πιθανότητα είναι ίση με όλο το εμβαδό μείων την P[X ≤ x]

Δηλαδή:

$$P[X > x] = 1 - P[X \le x]$$



Εφαρμογή 14

Αν η τ.μ. X ακολουθεί κανονική κατανομή με μ = 100 και σ = 30 (X ~ N(10, 30)), να βρεθεί η πιθανότητα P[X > 110]

ΛΥΣΗ

Θεωρούμε την τ.μ. Ζ με $Z = \frac{X-\mu}{\sigma} = \frac{X-100}{30}$ (με αντικατάσταση μ=10 και σ=30)

η οποία ακολουθεί τυποποιημένη κανονική κατανομή (Z ~ N(0, 1))

Οπότε η ζητούμενη πιθανότητα γράφεται:

$$P[X > 110] = P\left[\frac{X-100}{30} > \frac{110-100}{30}\right] =$$

$$= P\left[Z > \frac{110-100}{30}\right] =$$

$$= P[\,Z > 0{,}333] = 1 - \Phi(3{,}333)$$

κοιτάμε στον πίνακα την τιμή Φ(0,333) και με αντικατάσταση βρίσκουμε:

P[X > 110] = 1 – 63 = 0,27 περίπου

Προσέγγιση διωνυμικής κατανομής από τυποποιημένη κανονική:

Αν μία τ.μ. ακολουθεί κατανομή Bernoulli με παραμέτρους ν, p και το ν είναι αρκετά μεγάλο (στην πράξη μεγαλύτερο του 30), ισχύει:

$$P(\alpha \le X \le \beta) = \Phi\left(\frac{\beta + \frac{1}{2} - \nu p}{\sqrt{\nu p(1-p)}}\right) - \Phi\left(\frac{\alpha + \frac{1}{2} - \nu p}{\sqrt{\nu p(1-p)}}\right)$$

Εφαρμογή 15

Έστω ότι μας ενδιαφέρει το ποσοστό p υποψηφίου στις εκλογές. Θεωρούμε την τυχαία μεταβλητή Χ που παίρνει την τιμή 1 αν ο ψηφοφόρος ψηφίσει τον υποψήφιο και 0 αν δεν τον ψηφίσει.

Αν το δείγμα μας (δηλαδή το πλήθος των ερωτώμενων ψηφοφόρων) είναι μεγαλύτερο του 30, μπορούμε να προσεγγίσουμε την κατανομή Bernoulli (που ακολουθεί το δείγμα μας) από την τυποποιημένη κανονική κατανομή βάσει του προηγούμενου τύπου.

9.2.2. Εκθετική κατανομή:

Η συνεχής τ.μ. Χ ακολουθεί εκθετική κατανομή όταν η συνάρτηση πυκνότητας της πιθανότητας δίνεται από την σχέση:

$$\boxed{f(x) = \begin{cases} \theta e^{\theta x}, & \text{αν } x > 0 \\ 0, & \text{αν } x \le 0 \end{cases}}$$

στην περίπτωση αυτή το θ είναι η παράμετρος της κατανομής. Επίσης για την εκθετική κατανομή ισχύουν:

$$\boxed{\mu = \frac{1}{\theta}}$$

και

$$\boxed{\sigma^2 = \frac{1}{\theta^2}}$$

9.2.3. Ομοιόμορφη κατανομή:

Η συνεχής τ.μ. X λέμε ότι ακολουθεί την ομοιόμορφη κατανομή αν η συνάρτηση πυκνότητας της πιθανότητας δίνεται από την σχέση:

$$f(x)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \text{αν } \alpha\le x\le\beta\\ 0 & \text{, αν } x<\alpha \text{ ή } x>\beta\end{cases}$$

Επίσης για την ομοιόμορφη κατανομή ισχύουν:

$$\mu=\frac{\alpha+\beta}{2}$$

και

$$\sigma^2=\frac{(\beta-\alpha)^2}{12}$$

9.2.4. Κατανομή Γάμμα:

Μία συνεχής τ.μ. θα λέμε ότι ακολουθεί κατανομή Γάμμα όταν η συνάρτηση πυκνότητας της πιθανότητας δίνεται από την σχέση:

$$f(x)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(v)}\, x^{v-1}e^{-x/\beta} & \text{, αν } x>0\\ 0 & \text{, αν } x\le 0\end{cases}$$

όπου α, β > 0 και Γ(ν) η συνάρτηση Γάμμα που δίνεται από συγκεκριμένη μαθηματική σχέση την οποία δεν θα χρησιμοποιήσουμε.
Επίσης για την κατανομή Γάμμα ισχύουν:

$$\mu=\alpha\beta$$

και

$$\sigma^2=\alpha\beta^2$$

Σημείωση
Υπάρχουν και άλλες πολύ σημαντικές κατανομές συνεχών μεταβλητών με τις οποίες θα ασχοληθούμε σε επόμενο κεφάλαιο, στα διαστήματα εμπιστοσύνης:

ΕΞΑΣΚΗΣΗ

1. Τα τυχαία σφάλματα μιας μέτρησης ακολουθούν κανονική κατανομή με μέση τιμή 2 και διασπορά 4.
   α) Να βρεθεί η πιθανότητα να έχουμε λιγότερα από 3 λάθη.
   β) Να βρεθεί η πιθανότητα να έχουμε περισσότερα από 1 λάθη.

2. Ο αριθμός X των ατόμων που έρχονται σε ένα κατάστημα στο διάστημα 8-9 π.μ. ακολουθεί την κατανομή Poisson με παράμετρο λ = 2.

α) Να βρεθεί η πιθανότητα στο κατάστημα να μπουν 3 πελάτες στο διάστημα 8-9 π.μ.
β) Να βρεθεί η πιθανότητα στο κατάστημα να μπουν το πολύ 3 πελάτες στο διάστημα 8-9 π.μ.
γ) Να βρεθεί η πιθανότητα στο κατάστημα να μπουν τουλάχιστον 2 πελάτες.

3. Έστω ότι ο αριθμός των εκπεμπομένων σωματιδίων α από ραδιενεργό πηγή σε χρόνο t ωρών, ακολουθεί την κατανομή Poisson με μέση τιμή. Να βρεθεί η πιθανότητα:
α) Να βρεθεί η συνάρτηση πιθανότητας όταν μας ενδιαφέρει ο αριθμός σωματιδίων που εκπέμπονται σε διάστημα 4 ωρών.
β) Να βρεθεί η πιθανότητα σε διάστημα 4 ωρών να γίνει εκπομπή 3 σωματιδίων.
γ) Να βρεθεί η πιθανότητα σε διάστημα 4 ωρών να γίνει εκπομπή το πολύ 3 σωματιδίων.

4. Παίκτης basket εκτελεί ελεύθερες βολές. Έχει ποσοστό επιτυχίας $p = 0,7$. Να βρεθεί η πιθανότητα να ευστοχήσει σε 8 από τις 10 βολές.

5. Αν το ποσοστό επιτυχίας στόχου είναι $p = 0,6$, να βρεθεί η πιθανότητα να πετύχουμε το στόχο στη 2η προσπάθεια.

6. Αν ένα δείγμα είναι από την κανονική κατανομή με μέσο $\mu = 12$ και διασπορά $\sigma^2 = 25$, να βρεθούν οι πιθανότητες:
α) $p(x \leq 12)$
β) $p(x > 11)$

7. Αν ένα δείγμα είναι από την κανονική κατανομή με μέσο $\mu = 4$ και διασπορά $\sigma^2 = 4$, να βρεθούν οι πιθανότητες:
α) $p(2 \leq X \leq 3)$
β) $p(4 \leq X \leq 6)$
γ) $p(X > 5)$

8. Δείγμα από την διωνυμική κατανομή μεγέθους n μας δίνει πληροφορία για την ύπαρξη ή όχι ενός συγκεκριμένου ενζύμου στον ανθρώπινο οργανισμό.
α) Τι μέγεθος πρέπει να έχει το δείγμα μας, ώστε να γίνει προσέγγιση από την τυποποιημένη, κανονική κατανομή;
β) Αν όντως γίνει η προσέγγιση από την τυποποιημένη κανονική κατανομή του δείγματος με $n = 40$ και $p = 0,7$ από τη διωνυμική κατανομή, να βρεθεί η πιθανότητα $p(3 \leq X \leq 7)$

Σύνοψη

Οι κατανομές διακρίνονται σε ασυνεχείς και συνεχείς. Τα πλεονεκτήματα των θεωρητικών κατανομών είναι τα εξής:

Το πληθος των παρατηρήσεων αντικαθίσταται μόνο από λίγες παραμέτρους και έτσι γίνεται ευκολότερη η σύγκριση των διαφόρων παρατηρήσεων. Οι θεωρητικές κατανομές έχουν γνωστές ιδιότητες.

- **Ασυνεχείς κατανομές**
  - Διωνυμική
  - Bernoulli
  - Poisson
  - Γεωμετρική
  - Υπεργεωμετρική
- **Συνεχείς κατανομές**
  - Κανονική → ε
  - Τυποποιημένες
  - Εκθετική
  - Ομοιομορφη κατανομή

Βιβλιογραφία / Internet

«Probability and Statistics», Aanon, 1997

«Introduction to Probability theory and its Application», Eller W.

«Modern Probability theory and its Application», Parzen E.

«Statistics», Graham Alan, teach yourself

«Ασκήσεις Θεωρίας Πpιθανοτήτων», Κακούλας Θ.

«Μαθήματα Θεωρίας Πιθανοτήτων και Στατιστικής», Κούνια Σ.

«Πιθανότητες και Στατιστική», Spiegel, McGraw-Hill, ΕΣΠΙ «Στατιστική», Πέτρος Α. Κιόχος, INTERBOOKS

ee.wikipedia.org/wiki (βικιπαιδεία): *στο site αυτό ο ενδιαφερόμενος μπορεί να βρει διάφορες πληροφορίες για τις κατανομές Bernoulli, Διωνυμική, Γεωμετρική κ.ά.*

Οδηγός για περεταίρω μελέτη

***«Πιθανότητες και Στατιστική»*, Spiegel, McGraw-Hill**

Στο βιβλίο αυτό ο ενδιαφερόμενος μπορεί να βρεί σε πλήρη ανάπτυξη θέματα που αφορούν τόσο τις τυχαίες όσο και τις θεωρητικές μεταβλητές, υπάρχουν πολλά λυμένα παραδείγματα και εφαρμογές που καλύπτουν όλο το φάσμα της θεματολογίας των πιθανοτήτων και των κατανομών.

***«Θεωρία Πιθανοτήτων»*, Κούτρας Μάρκος, 2004**

Στο βιβλίο οενδιαφερόμενος θα βρει θέματα που αφορούν την έννοια της πιθανότητας, των πεπερασμένων δειγματικών χώρων, της δεσμευμένης πιθανότητας κατανομών, διακριτών και τυχαίων μεταβλητών, συνεχών κατανομών καθώς και εφαρμογές για υπολόγια.

10. ΔΕΙΓΜΑΤΟΛΗΨΙΑ & ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Βασικές έννοιες:
- δειγματοληψία
- πληθυσμός, δείγμα
- απλή τυχαία δειγματοληψία
- δειγματοληψία με επανατοποθέτηση
- δειγματοληπτικές κατανομές
- διαστήματα εμπιστοσύνης

Στόχος του μαθήματος: να γίνει μία παρουσίαση του πως επιλέγεται δείγμα από τον πληθυσμό που μας ενδιαφέρει και πώς αυτό σχετίζεται με την αξιοπίστία της έρευνας που διενεργείται.

Προσδοκώμενα αποτελέσματα: Ο εκπαιδευόμενος να είναι σε θέση να προσδιορίσει το μέγεθος του δείγματος που θα επιλέξει, ανάλογα με το μέγεθος του πληθυσμού, την αξιοπιστία που θέλει να πετύχει, αλλά λαμβάνοντας υπ' όψην και άλλους παράγοντες, όπως κόστος.

Εισαγωγικές παρατηρήσεις: Η δειγματοληψία και τα διαστήματα εμπιστοσύνης αφορούν το πρακτικό μέρος μίας στατιστικής έρευνας. Το πώς και γιατί θα επιλεγεί το κατάλληλο δείγμα καθορίζει το αν το αποτέλεσμα που θα προκύψει θα είναι σωστό, αξιόπιστο. Στην περίπτωση αυτή το δείγμα θα είναι αντιπροσωπευτικό του πληθυσμού. Αξίζει να σημειωθέι ότι έχουν καταγραφεί στατιστικές έρευνες πάνω σε σημαντικά ζητήματα, οι οποίες στη συνέχεια αποδείχθηκαν αναξιόπιστες και ακυρώθηκαν, λόγω λανθασμένης εφαρμογής δειγματοληπτικής μεθόδου.

10.1. ΔΕΙΓΜΑΤΟΛΗΨΙΑ

Προκειμένου να βγουν ορισμένα συμπεράσματα σε τομείς όπως η Βιομηχανία, η Γεωργία, η Ιατρική, η Εκπαίδευση, η Ψυχολογία, η Κοινωνιολογία, γίνεται συλλογή δεδομένων βάσει των οποίων προκύπτουν κάποια συμπεράσματα.

Υπάρχουν δύο βασικές μέθοδοι συλλογής δεδομένων:

1) απογραφή

2) δειγματοληπτική έρευνα

Κατά την απογραφή καταγράφονται οι παρατηρήσεις από το σύνολο του πληθυσμού, ενώ κατά τη δειγματοληπτική έρευνα επιλέγεται ένα **μέρος του πληθυσμού (δείγμα)** και βάσει αυτού βγαίνουν συμπεράσματα για ολόκληρο τον πληθυσμό.

Από τη φύση των μεθόδων προκύπτει ότι η απογραφή έχει τα εξής μειονεκτήματα:

- **Μεγάλο Κόστος**

  Για να πραγματοποιηθεί μία απογραφή, χρειάζεται ειδική προεργασία, καθώς και μεγάλος αριθμός εκπαιδευμένων απογραφέων, ατόμων δηλαδή που θα συλλέξουν τα δεδομένα. Για την προετοιμασία αυτή, αλλά και για την ίδια τη διαδικασία της απογραφής, απαιτούνται πολλά έξοδα. Για το λόγο αυτό, η απογραφή του πληθυσμού της χώρας, για παράδειγμα, γίνεται κάθε δέκα χρόνια, ενώ η απογραφή βιομηχανιών και βιοτεχνιών κάθε πέντε χρόνια, αφού είναι πιο οικονομική.

- **Εξειδικευμένα Άτομα**

  Προκειμένου να γίνει μία απογραφή, χρειάζονται εξειδικευμένα άτομα, οι απογραφείς όπως ειπώθηκε, ώστε να αποφεύγονται προσωπικά σφάλματα. Η εκπαίδευση τέτοιων ατόμων δεν είναι πάντα δυνατή λόγω της έκτασης των απογραφών. Σαν συνέπεια, η χρήση μη εξειδικευμένων ατόμων στην απογραφή έχει ως αποτέλεσμα τη συγκέντρωση πληροφοριών που δεν δίνουν ορθή εικόνα του πληθυσμού.

- **Επικαιρότητα αποτελεσμάτων**

  Ακόμη ένα μειονέκτημα των απογραφών είναι ότι λόγω του μεγάλου όγκου των δεδομένων, η επεξεργασία καθυστερεί, άρα και η δημοσίευση των αποτελεσμάτων της απογραφής. Έτσι, πολλές φορές όταν δημοσιευτούν τα αποτελέσματα δεν ανταποκρίνονται στην τρέχουσα εικόνα του πληθυσμού.

- **Καταστροφή εξεταζόμενων αντικειμένων**

  Αν για παράδειγμα θέλουμε να εξετάσουμε τους ελαττωματικούς λαμπτήρες, θα πρέπει να καταστραφούν όλοι οι λαμπτήρες για να εξεταστούν με απογραφή, κάτι προφανώς ασύμφορο.

Προκειμένου να αποφευχθούν οι συνέπειες από τα μειονεκτήματα των απογραφών, εφαρμόζεται η μέθοδος της δειγματοληψίας, πραγματοποιούνται δηλαδή, δειγματοληπτικές έρευνες.

Καθημερινά πλέον σε εφημερίδες, τηλεόραση και ραδιόφωνο ανακοινώνονται τα αποτελέσματα ερευνών δειγματοληπτικών, δημοσκοπήσεων όπως είναι γνωστές. Μάλιστα, είναι τέτοια η συχνή αναφορά σε δημοσκοπήσεις, που η διενέργειά τους φαίνεται απλή υπόθεση. Παρ' όλα αυτά, ιστορικά λάθη έχουν δείξει ότι η διεξαγωγή μίας δειγματοληπτικής έρευνας απαιτεί συγκεκριμένες γνώσεις, οι οποίες αφορούν τη Στατιστική και συγκεκριμένα τον κλάδο της που ονομάζεται Δειγματοληψία. Η Δειγματοληψία ασχολείται ως επιστημονικός κλάδος με τις αρχές και τις μεθόδους συλλογής και ανάλυσης δεδομένων από συγκεκριμένους πληθυσμούς.

Πληθυσμός και δείγμα:

Το πρώτο βήμα της δειγματοληπτικής έρευνας είναι, να καταγράψουμε ακριβώς τον πληθυσμό για τον οποίο θέλουμε να βγάλουμε συμπεράσματα. Ο όρος πληθυσμός αναφέρεται στην ομάδα αντικειμένων, ατόμων ή παρατηρήσεων που μας ενδιαφέρει. Στη συνέχεια, δίνονται όροι οι οποίοι θα χρησιμοποιηθούν στη συνέχεια και διευκολύνουν τη συνεννόησή μας:

- **Στοιχείο (element)**

  Αν μετράμε το ύψος των μαθητών, ως στοιχείο θα θεωρήσουμε τον μαθητή. Αν θέλουμε να βρούμε πόσα παιδιά έχουν οι οικογένειες των εργαζομένων σε μία συγκεκριμένη εταιρία, σαν στοιχείο θα πάρουμε μία οικογένεια εργαζομένου στη συγκεκριμένη εταιρία. Δηλαδή, στοιχείο είναι η μονάδα ενός συνόλου όπου γίνεται μέτρηση ή παρατήρηση στα πλαίσια της δειγματοληπτικής έρευνας.

- **Πληθυσμός (population)**

  Είναι το σύνολο των στοιχείων για το οποίο μας ενδιαφέρει να βγάλουμε κάποια συμπεράσματα.

- **Δειγματοληπτικές μονάδες**

  Είναι ομάδες απλών στοιχείων, όπως οι διάφορες τάξεις σχολείων, οι οικογένειες των εργαζομένων σε ένα εργοστάσιο, τα νοικοκυριά μίας επαρχιακής πόλης.

  Σχόλιο: οι δειγματοληπτικές μονάδες μπορεί να μην συμπίπτουν με τα στοιχεία. Για παράδειγμα, η δειγματοληπτική μονάδα μπορεί να είναι μία οικογένεια και στοιχείο να είναι η μητέρα ή ο πατέρας ή κάποιο άλλο μέλος της οικογένειας (συνήθως με περιορισμούς όπως άνω των 18, εργαζόμενος ή όχι και άλλα).

- **Δειγματοληπτικό πλαίσιο (frame)**

  Συνήθως είναι μία λίστα που περιέχει όλα τα στοιχεία του πληθυσμού που ερευνούμε. Μπορεί δηλαδή να είναι ένας τηλεφωνικός κατάλογος ή μία λίστα από εγγεγραμμένους (μέλη) ενός συλλόγου.

  Το δειγματοληπτικό πλαίσιο πρέπει να περιέχει πληροφορίες σωστές, έγκυρες, να μην έχει διπλές εγγραφές και να είναι πλήρες, δηλαδή να μην λείπουν πληροφορίες.

- **Δείγμα**

  Είναι μία συλλογή δειγματοληπτικών μονάδων από το πλαίσιο. Για παράδειγμα παίρνουμε τον τηλεφωνικό κατάλογο και κάθε 10ο όνομα επιλέγεται να ερωτηθεί σχετικά με το ποιο ραδιοφωνικό σταθμό ακούει. Δεν θα ρωτήσουμε ολόκληρο τον πληθυσμό, αλλά επιλέγουμε τον 10^{ο}, 20^{ο}, 30^{ο},... από τον τηλεφωνικό κατάλογο. Αυτοί αποτελούν το δείγμα.

Το δείγμα θα χρησιμοποιηθεί για την εξαγωγή συμπερασμάτων για ολόκληρο τον πληθυσμό. Όπως είναι γνωστό, προκειμένου να προβλέψουμε ένα φαινόμενο, πρέπει να γνωρίζουμε την κατανομή που ακολουθεί. Για παράδειγμα, ένα δείγμα που ακολουθεί την διωνυμική κατανομή ονομάζεται κανονικό δείγμα, ένα δείγμα που ακολουθεί την διωνυμική κατανομή ονομάζεται διωνυμικό και ούτω καθ' εξής. Ανάλογα με την περίπτωση που αντιμετωπίζουμε, θα λέγαμε ότι είναι σχετικά εύκολο να βρούμε την κατανομή. Για παράδειγμα, αν έχουμε πρόβλεψη φύλου παιδιού που γεννιέται, θα καταφύγουμε σε διωνυμική ή γεωμετρική κατανομή, ανάλογα και με το ζητούμενο. Ποιες θα είναι όμως οι παράμετροι της κατανομής που θα επιλέξουμε;

Οι παράμετροι είναι καθοριστικές για τη μορφή που θα έχει η κατανομή που επιλέξαμε. Αρκεί να υπενθυμίσουμε την περίπτωση της καμπύλης της κανονικής κατανομής και το πώς άλλαζε σε σχέση με το μ (μέσο) ή το σ^2 (διασπορά).

Το θέμα είναι ότι οι παράμετροι δεν είναι γνωστές εκ των προτέρων. Για να μπορέσουμε να μελετήσουμε το φαινόμενο που μας απασχολεί, εφόσον οι παράμετροι δεν μας είναι γνωστές, θα τις «εκτιμήσουμε». Δηλαδή, θα προσπαθήσουμε να βρούμε περίπου ποιες είναι. Αυτό γίνεται με συγκεκριμένους τρόπους, ώστε τα αποτελέσματα που θα έχουμε να είναι έγκυρα. Με την εκτίμηση των παραμέτρων ασχολείται ο κλάδος της Στατιστικής που ονομάζεται «Εκτιμητική».

Η εκτίμηση παραμέτρων του πληθυσμού από τις πληροφορίες (απαντήσεις σε ερωτηματολόγια κλπ) που παίρνουμε από το δείγμα είναι ο στόχος της δειγματοληπτικής έρευνας. Πιο συγκεκριμένα μας ενδιαφέρουν οι ακόλουθες παράμετροι:

1) Πληθυσμιακό ολικό
   για παράδειγμα, το συνολικό εισόδημα των κατοίκων ενός χωριού, ο συνολικός αριθμός πελατών σε ένα κατάστημα μέσα σε ένα συγκεκριμένο χρονικό διάστημα, η συνολική ποσότητα λιπάσματος που χρησιμοποιείται σε καλλιέργειες μίας έκτασης.
2) Πληθυσμιακός μέσος
   για παράδειγμα, το μέσο ύψος των μαθητών ενός σχολείου, η μέση ημερήσια κατανάλωση βενζίνης μίας χώρας, ο μέσος αριθμός επισκεπτών ενός μουσείου κατά την τουριστική περίοδο.
3) Πληθυσμιακό ποσοστό
   για παράδειγμα, το ποσοστό των μαθητών που καπνίζουν, το ποσοστό των ανέργων μεταξύ 18 και 24 ετών, το ποσοστό των γυναικών που είναι εργαζόμενες μητέρες, το ποσοστό των πολιτών που είναι υπέρ ενός κυβερνητικού μέτρου.
4) Λόγος
   ο λόγος σημαίνει να συγκρίνουμε δύο χαρακτηριστικά και να πούμε για παράδειγμα ότι οι μαθητές είναι δεκαπλάσιοι των καθηγητών, δηλαδή ο λόγος των καθηγητών προς τους μαθητές είναι 1 προς 10 ή $\frac{1}{10}$.

Στάδια έρευνας:

Τα στάδια μίας έρευνας είναι καθοριστικά και πρέπει να γίνονται με μεγάλη προσοχή, ώστε να αποφύγουμε συμπεράσματα τα οποία δεν ισχύουν, συγκεκριμένα για να πραγματοποιηθεί μία δειγματοληπτική έρευνα έχουμε τα ακόλουθα στάδια:

Στάδιο 1ο: Σκοπός έρευνας

Προσδιορίζουμε ποιο ακριβώς είναι το πρόβλημα. Η έρευνα γίνεται για να δώσει λύση σε κάποιο θέμα που έχει τεθεί και χρειαζόμαστε να γνωρίζουμε επιπλέον στοιχεία γι' αυτό.

Στάδιο 2ο: Επιλογή τεχνικής

Αποφασίζουμε την κατάλληλη τεχνική ανάλογα με το πρόβλημα και τον πληθυσμό μας.

Στάδιο 3ο: Εκλογή και καθορισμός δείγματος

Σε ορισμένες περιπτώσεις ο πληθυσμός ταυτίζεται με το δείγμα. Για παράδειγμα, θέλουμε να δούμε το ποσοστό των καπνιστών μεταξύ 18 και 24 και κάνουμε έρευνα σε φοιτητές. Προφανώς, στην περίπτωση αυτή, έχουμε αποκλείσει τους νέους που δεν είναι φοιτητές. Σε άλλες περιπτώσεις θεωρούμε ότι έχουμε πληθυσμό οικογενειών και ρωτάμε ένα μέλος της οικογένειας. Πρέπει λοιπόν να καθορίσουμε ξεκάθαρα ποιος είναι ο πληθυσμός στον οποίο θα αναφερθούμε. Για να γίνει αυτό έχουμε κάποια κριτήρια:

1) Γεωγραφικά
   Μία έρευνα μπορεί να αναφέρεται σε μία συγκεκριμένη περιοχή, σε ένα νομό, ένα χωριό (κοινότητα), μία πόλη ή σε ολόκληρη την επικράτεια. Όταν για παράδειγμα μας ενδιαφέρει η γνώμη των κατοίκων μίας περιοχής για την δημιουργία αθλητικού κέντρου ή εμπορικού κέντρου στην περιοχή, είμαστε υποχρεωμένοι να καθορίσουμε αυστηρά τα όρια της περιοχής αυτής.

2) Ηλικία

Ανάλογα με την έρευνα συνήθως απευθυνόμαστε σε άτομα συγκεκριμένης ηλικίας. Για παράδειγμα, αν μας ενδιαφέρει ο ελεύθερος χρόνος των μαθητών του Γυμνασίου, περιοριζόμαστε σε ηλικίες 12-15 ετών. Αν μας ενδιαφέρει η πρόθεση ψήφου, θα έχουμε ως κάτω όριο ηλικίας τα 18 έτη.

3) Άλλα στοιχεία, όπως:

- Φύλο
- Οικογενειακή κατάσταση
- Εκπαίδευση
- Εθνικότητα
- Υπηκοότητα

4) Νοικοκυριό

Νοικοκυριό είναι:

- Δύο ή περισσότερα άτομα, τα οποία δεν χρειάζεται να είναι συγγενείς, αλλά που κατοικούν στο ίδιο σπίτι, αγοράζουν από κοινού τα απαραίτητα (για παράδειγμα δεν πηγαίνει ο καθένας μόνος του να αγοράσει το φαγητό του).
- Ένα άτομο που κατοικεί μόνο του ή με άλλα άτομα, αλλά δεν αγοράζει από κοινού με τα άλλα άτομα τα απαραίτητα (δηλαδή αγοράζει μόνο του το φαγητό του για παράδειγμα).

5) Διάφορα άλλα χαρακτηριστικά

Υπάρχουν χαρακτηριστικά που λαμβάνουμε υπόψη μας ανάλογα με τη συγκεκριμένη έρευνα που κάνουμε.

Για παράδειγμα, αν θέλουμε να εξετάσουμε το χαρτζιλίκι των μαθητών της Γ΄ τάξης της Λυκείου, θα πρέπει να περιοριστούμε βέβαια σε μαθητές της Γ΄ Λυκείου, αλλά και να λάβουμε υπόψη μας χαρακτηριστικά όπως η περιοχή όπου βρίσκεται το σχολείο, αν είναι ιδιωτικό ή όχι, αν είναι αθλητικό, μουσικό, γενικό Λύκειο κλπ.

Το επόμενο βήμα για να μελετήσουμε ένα φαινόμενο με τον τρόπο που περιγράψαμε παραπάνω είναι να επιλέξουμε το δείγμα. Η επιλογή του δείγματος μπορεί να γίνει με πολλούς τρόπους. Για παράδειγμα, αλφαβητικά από τον τηλεφωνικό κατάλογο, από πόρτα σε πόρτα, από το internet.

Πιο συγκεκριμένα, πριν διενεργηθεί μία δειγματοληπτική έρευνα, η ομάδα εργασίας, όπως λέγεται η ομάδα των επιστημόνων που θα ασχοληθεί με το θέμα, πρέπει να αποφασίσει σχετικά με το πλαίσιο της δειγματοληψίας.

Το πλαίσιο της δειγματοληψίας είναι ένας κατάλογος που περιέχει όλες τις μονάδες του πληθυσμού. Προφανώς ο κατάλογος αυτός πρέπει να μην περιέχει διπλές εγγραφές, να είναι ενημερωμένος και έγκυρος.

Στη συνέχεια, η ομάδα εργασίας φτιάχνει το ερωτηματολόγιο, ένα έντυπο το οποίο αποτελεί την πηγή των πληροφοριών ουσιαστικά, καθώς σε αυτό θα απαντήσουν οι ερωτώμενοι. Το ερωτηματολόγιο πρέπει να είναι σύντομο, απλό, σαφές, χωρίς αρνητικές ερωτήσεις, χωρίς ερωτήσεις που κατευθύνουν τον ερωτώμενο και με ερωτήματα σε λογική σειρά. Η συγγραφή ενός ερωτηματολογίου απαιτεί πολλές γνώσεις και μεγάλη προσοχή από την ομάδα εργασίας.

Προκειμένου να γίνει η έρευνα απαιτείται πλήθος ατόμων εκτός της ομάδας εργασίας που θα αναλάβουν να δώσουν το ερωτηματολόγιο και να συλλέξουν έτσι το δείγμα. Τα άτομα αυτά πρέπει να εκπαιδευτούν κατάλληλα, ώστε να γνωρίζουν το σκοπό της έρευνας και βάσει αυτού να εργαστούν.

Τέλος, πριν την έρευνα γίνεται μία τελική δοκιμαστική έρευνα, στόχος της οποίας είναι να δοκιμαστεί το ερωτηματολόγιο πριν δοθεί στα πλαίσια της κύριας έρευνας.

Το δείγμα επιλέγεται έτσι, ώστε να έχει τα ίδια περίπου χαρακτηριστικά με τον πληθυσμό. Εννοείται ότι προσπαθούμε να έχουν τα ίδια χαρακτηριστικά, αλλά αυτό δεν γίνεται, αφού δεν ρωτάμε τα ίδια άτομα για παράδειγμα. Αυτό το «περίπου» καθορίζεται από τους ερευνητές. Δηλαδή με πόσο καλή προσέγγιση (πλησίασμα) των πραγματικών χαρακτηριστικών είμαστε ικανοποιημένοι. Το πόσους από τον πληθυσμό θα ρωτήσουμε, δηλαδή ποιο θα είναι το μέγεθος του δείγματός μας και το πώς θα επιλέξουμε αυτούς τους συγκεκριμένους (επιλογή δείγματος) είναι επιλογές μεγάλης σημασίας για μία δειγματοληπτική έρευνα. Υπάρχουν περιπτώσεις που το δείγμα είναι μεγάλο, αλλά έχει επιλεγεί με τέτοιο τρόπο, ώστε δεν είναι αντιπροσωπευτικό.

Αν για παράδειγμα ενδιαφερόμαστε για το ποσοστό καπνιστών σε ηλικίες 18-24, είναι προτιμότερο να επιλέξουμε ως δείγμα 2000 νέους με υψηλό εισόδημα ή 1000 νέους, οι οποίοι θα προέρχονται από όλες τις κοινωνικές τάξεις; Προφανώς, στη δεύτερη περίπτωση αν και το δείγμα είναι μικρότερο, θα έχουμε καλύτερη εικόνα για τον πληθυσμό, καθώς αντιπροσωπεύει όλες τις κοινωνικές τάξεις. Είναι όπως λέμε «αντιπροσωπευτικό του πληθυσμού». Οπότε, αν και γενικά τα μεγάλα δείγματα είναι πιο αξιόπιστα, δεν αρκεί μόνο αυτό, πρέπει να έχουν επιλεγεί και με τον καλύτερο δυνατό τρόπο.

Γενικά, θα λέγαμε ότι ένα πολύ μεγάλο δείγμα είναι ασύμφορο, ενώ ένα πολύ μικρό μπορεί να οδηγήσει σε μεροληπτικές (δηλαδή όχι αντικειμενικά σωστές) εκτιμήσεις.

Η δειγματοληψία γίνεται με διαφορετικούς τρόπους ως προς την επιλογή του δείγματος. Στο ερώτημα λοιπόν «πως θα επιλέξω αυτούς που θα ρωτήσω», η απάντηση δίνεται ανάλογα με:

- Το μέγεθος του πληθυσμού
- Την ακρίβεια που θέλουμε να έχουμε
- Το χρόνο που έχουμε στη διάθεσή μας
- Τα χρήματα που έχουμε στη διάθεσή μας
- Τους καταλόγους που έχουμε στη διάθεσή μας
- Τη φύση του πληθυσμού, δηλαδή το αν μπορεί ο πληθυσμός να διαιρεθεί σε ομάδες οι οποίες είναι ομοιογενείς, δηλαδή περιέχουν περίπου το ίδιο ποσοστό από όλες τις ομάδες μονάδων του πληθυσμού.

Επειδή δεν έχουμε ολόκληρο τον πληθυσμό, είναι αναμενόμενο να κάνουμε κάποιο σφάλμα στην εκτίμηση των παραμέτρων που μας ενδιαφέρουν. Όσο πιο μικρό είναι το σφάλμα, τόσο καλύτερα για την έρευνά μας. Σε κάθε δειγματοληπτική έρευνα έχουμε λοιπόν, το «περιθώριο σφάλματος». Το περιθώριο σφάλματος είναι ακριβώς αυτό που λέει η λέξη, πόσο λάθος μπορούμε να κάνουμε, πόσο μπορεί να απέχει η πραγματική τιμή της παραμέτρου από αυτή που βρήκαμε από το δείγμα.

Επίσης, πρέπει στην αρχή της έρευνας να ορίσουμε την αξιοπιστία ή εμπιστοσύνη με την οποία θα γίνει η εκτίμηση. Η αξιοπιστία δείχνει το βαθμό στον οποίο τα αποτελέσματα που βρήκαμε είναι τα σωστά.

Συμβολικά:

d: σφάλμα

θ: παράμετρος του πληθυσμού που θέλουμε να εκτιμήσουμε

$\hat{\theta}$: εκτιμήτρια της θ, δηλαδή αποτέλεσμα που βρίσκουμε από το δείγμα

$1-\alpha$: αξιοπιστία

α: επίπεδο σημαντικότητας

τότε το περιθώριο σφάλματος d ορίζεται από τη σχέση:

$$P\left[\left|\theta-\hat{\theta}\right| \le d\right] \ge 1-a$$

όπου

$\left|\theta-\hat{\theta}\right| \le d$ σημαίνει ότι $-d \le \theta-\hat{\theta} \le d$ ή $\hat{\theta}-d \le \theta \le d+\hat{\theta}$

δηλαδή, ότι το θ βρίσκεται μεταξύ των τιμών $\hat{\theta}-d$ και $\hat{\theta}+d$ που είναι η τιμή που συμπεράναμε από το δείγμα συν πλην το σφάλμα με πιθανότητα (1-α)100%.

Συνήθως, σαν επίπεδο σημαντικότητας επιλέγεται το 0,01 ή το 0,05. Δηλαδή, είμαστε κατά 99% ή κατά 95% σίγουροι ότι η πραγματική τιμή της παραμέτρου βρίσκεται μεταξύ των $\hat{\theta}-d$ και $\hat{\theta}+d$.

Όπως φαίνεται, το κυριότερο μειονέκτημα μίας δειγματοληπτικής έρευνας σε σχέση με την απογραφή είναι αυτά τα σφάλματα. Όμως, με κατάλληλες μεθόδους είμαστε σε θέση να επιλέξουμε το δείγμα έτσι, ώστε να ελέγχουμε τα σφάλματα αυτά και να τα προβλέπουμε.

Στάδιο 4ο: Συλλογή δεδομένων

Η συλλογή των δεδομένων από το δείγμα γίνεται με διάφορους τρόπους, οι οποίοι αναφέρονται στη συνέχεια:

Α) Παρατήρηση (αυτόματα)

Χρησιμοποιείται σε περιπτώσεις που δεν έχουμε ερωτηματολόγιο, αλλά μας ενδιαφέρουν κάποιες μετρήσεις. Εφαρμογή βρίσκει σε επιστήμες όπως η Ανθρωπολογία, για την μελέτη μίας φυλής στην Αφρική με διαφορετικό τρόπο ζωής από τον δικό μας. Επίσης, παρατήρηση γίνεται στους δρόμους για τυχόν παραβάσεις του κοκ, για την πρόληψη ατυχημάτων ή για κυκλοφοριακές ρυθμίσεις προκειμένου να βελτιωθεί η κυκλοφορία.

Β) Προσωπική συνέντευξη

Η συνέντευξη αποτελεί έναν πολύ καλό τρόπο συλλογής στατιστικών στοιχείων σε επιστήμες, όπως η Διδακτική των Μαθηματικών. Στην περίπτωση της συνέντευξης η συμπλήρωση του ερωτηματολογίου γίνεται από ειδικά εκπαιδευμένα άτομα, που λέγονται ερευνητές ή ερευνήτριες.

Γ) Ταχυδρομική αποστολή του ερωτηματολογίου

Πρόκειται για μία μέθοδο, η οποία ουσιαστικά δε χρησιμοποιείται σήμερα, καθώς πολύ λίγοι ερωτώμενοι μπαίνουν στη διαδικασία να αποστείλουν το συμπληρωμένο ερωτηματολόγιο.

Δ) Τηλεφωνικά

Η τηλεφωνική συνέντευξη πλεονεκτεί ως προς το κόστος και το χρόνο, όμως υπάρχει ο κίνδυνος να μη ληφθούν υπόψη άτομα τα οποία δεν έχουν τηλέφωνα ή απουσιάζουν τη στιγμή που καλούνται. Επιπλέον, είναι πιθανό να μην καταλάβει ο ερωτώμενος τις ερωτήσεις που του θέτουμε από το τηλέφωνο.

10.2. ΜΕΘΟΔΟΙ ΔΙΕΝΕΡΓΕΙΑΣ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ

10.2.1. Απλή τυχαία δειγματοληψία

Το τυχαία αναφέρεται στο ότι κάθε μονάδα του πληθυσμού έχει την ίδια πιθανότητα να επιλεγεί στο δείγμα με όλες τις υπόλοιπες μονάδες. Στην περίπτωση αυτή, το δείγμα μας λέγεται ***τυχαίο δείγμα***.

Στην απλή τυχαία δειγματοληψία, όλες οι μονάδες επιλέγονται χωρίς επανάθεση. Αν για παράδειγμα ενδιαφερόμαστε για την πρόθεση ψήφου, κάθε ψηφοφόρο που ρωτάμε δεν τον ξαναρωτάμε.

Προκειμένου να πραγματοποιηθεί απλή τυχαία δειγματοληψία, πρέπει να χρησιμοποιήσουμε τη μέθοδο της κλήρωσης. Στην πράξη όμως χρησιμοποιούμε πίνακες τυχαίων αριθμών οι οποίοι είναι ως εξής:

ΤΥΧΑΙΟΙ ΑΡΙΘΜΟΙ

03 47 43 73 86	36 96 47 86 61	46 98 63 71 62	33 26 16 80 45	60 11 14 10 95
97 47 70 91 42	05 96 43 10 09	02 90 46 77 77	42 91 77 64 75	00 85 58 73 12
16 76 29 26 66	55 48 18 34 32	09 26 45 93 79	97 21 02 75 29	11 68 82 47 82
12 56 69 30 58	02 69 56 51 47	08 03 41 76 22	09 83 19 73 82	80 56 14 18 78
55 59 53 34 51	06 80 97 20 42	59 94 02 78 79	27 34 59 07 45	21 55 05 31 60
16 22 98 19 08	13 88 37 39 83	29 94 42 60 84	47 93 69 18 27	32 65 02 57 70
84 42 64 02 36	87 70 71 01 29	68 65 66 30 66	72 92 80 81 68	27 07 03 70 73
63 01 94 46 11	60 42 55 73 78	47 64 73 44 36	88 38 85 77 91	30 59 90 05 50
33 21 51 75 73	16 62 00 74 07	00 27 38 24 38	83 54 32 77 40	18 36 62 44 82
57 60 56 64 71	77 54 52 88 43	12 77 31 00 90	28 70 75 28 73	93 69 59 32 14
18 18 95 58 67	39 22 87 93 59	87 43 38 93 99	09 71 50 32 41	35 81 97 02 77
26 62 47 19 91	51 30 09 70 00	79 99 47 51 37	59 11 12 28 69	78 30 94 63 74
23 42 75 82 96	32 36 45 68 00	81 05 20 95 12	70 06 20 18 43	34 59 62 60 77
52 36 78 66 20	68 80 07 26 12	84 93 01 87 63	68 60 34 33 49	49 56 62 24 57
37 85 07 73 03	03 63 29 50 75	35 49 33 09 02	76 00 25 06 11	76 17 56 43 99
70 29 34 67 02	44 52 78 92 54	34 17 80 46 15	46 50 49 25 11	33 85 91 16 28
56 62 78 09 89	00 18 69 93 40	76 77 71 58 80	13 81 51 69 51	62 83 89 59 38
99 49 39 03 69	55 63 01 26 71	42 19 51 99 18	66 42 86 79 93	99 32 69 22 24
16 08 80 79 12	22 87 15 13 37	80 15 83 06 02	48 66 51 86 72	78 28 58 32 66
31 16 36 90 61	52 48 32 79 65	85 73 89 74 66	03 51 0 95 46	32 53 90 54 33
68 34 96 39 86	59 56 28 26 15	78 55 57 78 56	97 37 85 05 58	21 87 23 99 21
74 57 58 53 61	99 63 52 87 11	88 52 91 21 28	93 12 74 63 33	11 90 98 73 09
27 42 55 34 77	55 55 67 01 17	04 25 78 29 48	94 16 85 42 67	23 49 48 15 33
00 39 73 92 38	13 93 84 08 66	36 03 65 73 43	43 34 59 49 07	42 62 27 05 95
29 94 16 38 58	34 48 77 19 50	86 47 62 27 03	07 11 73 35 73	76 17 81 34 16
16 90 78 73 50	03 02 77 95 09	60 57 30 25 03	57 96 75 87 15	64 98 61 17 65
11 27 78 41 42	09 18 07 34 36	33 49 41 85 69	41 91 75 49 59	36 74 33 25 65
35 24 54 72 22	16 87 92 05 84	25 98 37 31 57	46 87 97 68 30	88 04 16 57 39
38 23 92 02 24	20 37 86 39 45	14 98 70 86 16	56 89 90 12 79	68 72 24 56 25
31 96 03 87 46	06 38 99 20 51	68 69 81 12 43	92 46 03 51 08	14 93 57 88 42
66 67 90 26 36	21 52 00 80 91	04 44 54 34 73	84 94 78 76 10	52 65 61 07 55
14 90 31 22 54	51 52 84 51 49	40 05 34 24 80	52 13 22 23 26	47 01 70 70 45
68 05 28 41 97	57 86 20 51 84	30 64 10 30 89	73 43 62 98 22	55 56 00 54 64
20 46 84 92 89	46 55 56 02 68	57 42 70 37 31	85 24 42 85 37	78 01 06 96 45
64 19 21 46 69	68 94 45 33 67	80 81 42 40 99	69 40 59 64 29	90 45 99 56 93
05 26 56 75 63	79 47 35 14 36	89 59 70 89 01	75 08 83 88 59	72 94 22 37 13
07 97 17 54 82	00 24 21 89 41	52 28 84 65 75	46 79 49 57 50	85 26 88 06 03
68 71 58 13 17	87 56 80 66 56	23 69 96 54 60	83 73 13 82 78	37 77 97 92 85
26 99 16 60 95	75 58 52 68 04	90 43 45 62 19	17 74 05 88 62	96 91 98 34 44
14 65 82 02 57	11 63 76 93 51	07 57 43 98 46	69 30 63 42 54	22 02 70 15 62
37 53 95 97 26	98 43 13 02 34	13 53 74 63 04	09 92 16 23 77	20 87 90 69 18
90 26 54 00 14	75 63 65 84 73	71 34 79 53 62	49 83 22 43 95	71 08 40 9 83
41 23 94 22 61	43 91 35 69 25	13 17 20 33 94	06 90 19 83 48	10 20 06 26 27
60 20 63 68 24	00 67 60 76 49	66 23 92 10 37	02 65 49 39 16	74 56 76 31 22
91 75 83 89 66	37 47 16 25 16	50 29 11 37 90	42 15 92 00 02	44 71 73 73 44
34 50 39 62 84	77 05 26 23 38	29 99 31 73 21	83 87 25 78 87	83 47 34 82 00
85 22 73 19 66	20 41 79 47 96	47 06 31 69 78	38 61 83 82 74	61 97 30 13 69
09 79 75 42 87	53 18 40 09 14	02 54 88 67 55	17 05 86 78 07	45 18 48 74 20
88 75 93 68 46	77 91 01 54 08	19 51 04 71 31	11 96 92 58 90	27 29 09 60 53
90 96 91 73 34	36 46 24 14 77	94 78 16 22 68	07 48 86 64 36	16 35 36 64 26

Εφαρμογή 1

Για να γίνει η επιλογή έχουμε:

Βήμα 1ο:

Αποφασίζουμε το μέγεθος του δείγματος. Έστω ότι θα πάρουμε δείγμα 15 ατόμων από πληθυσμό που έχει 90 άτομα.

Βήμα 2ο:

Παίρνουμε τον κατάλογο των τυχαίων αριθμών και επιλέγουμε το δείγμα αρχίζοντας από οποιοδήποτε αριθμό και συνεχίζοντας πάντα με τον ίδιο τρόπο, είτε οριζόντια είτε κατακόρυφα.

Θα ξεκινήσουμε από το 3 στην πρώτη γραμμή θα κινηθούμε οριζόντια.

Βήμα 3ο:

Καταγράφουμε τους αριθμούς που είναι μικρότεροι ή ίσοι με το πλήθος του πληθυσμού, δηλαδή στο παράδειγμά μας μικρότεροι ή ίσοι του 90.
Σημείωση: αν βρούμε δεύτερη φορά τον ίδιο αριθμό, δεν τον σημειώνουμε. Οπότε, ξεκινώντας από το 3 και κινούμενοι οριζόντια, καταγράφουμε τους αριθμούς:
3, 47, 43, 73, 86, 36, 61, 46, 63, 71, 62, 33, 26, 16, 80

Σχόλια:
Δεν καταγράψαμε τους αριθμούς 96, 98, γιατί είναι μεγαλύτεροι από το πλήθος των ατόμων του πληθυσμού (90).
Επίσης, δεν καταγράψαμε δεύτερη φορά τους αριθμούς 47 και 86.

Εφαρμογή 2

Έστω ότι θέλουμε δείγμα 15 ατόμων από πληθυσμό 400 ατόμων. Ακολουθούμε την ίδια διαδικασία, αλλά διαβάζουμε τριψήφιους αριθμούς.
Οπότε, ξεκινώντας από το 369 στην πρώτη γραμμή και στην έκτη στήλη και κινούμενοι κατακόρυφα θα καταγράψουμε τους αριθμούς:

369,	059,	026,	068,	138,	166,	392,	323,
036,	001,	228,	139,	030,	091,	168.	

10.2.2. Δειγματοληψία με Επανατοποθέτηση

Εκτός από τη μέθοδο της τυχαίας επιλογής δείγματος χρησιμοποιώντας πίνακες τυχαίων αριθμών, η επιλογή μπορεί να γίνει και με επανατοποθέτηση.

Επανατοποθέτηση σημαίνει ότι η μονάδα του πληθυσμού που επιλέγουμε ξανατοποθετείται στον πληθυσμό με ενδεχόμενη την επανεκλογή της.

10.2.3. Δειγματοληπτικές κατανομές

Στο σημείο αυτό, θα ασχοληθούμε με μία από τις κυριότερες παραμέτρους του πληθυσμού, τη μέση τιμή. Συμβολίζουμε με:

n: το μέγεθος του δείγματος

N: το μέγεθος του πληθυσμού

• Δειγματικός Μέσος

Κατανομή της δειγματικής μέσης τιμής ονομάζεται η κατανομή πιθανότητας της στατιστικής συνάρτησης $\overline{x}$ του δείγματός μας. Για την $\overline{x}$ ισχύουν:

$$\boxed{E\left(\overline{x}\right)=\mu} \qquad (1)$$

όπου μ η μέση τιμή του πληθυσμού

$$\boxed{E\left[\left(\overline{x}-\mu\right)^2\right]=\sigma_{\overline{x}}^2=\frac{\sigma^2}{n}} \qquad (2)$$

όπου σ^2 η διασπορά του πληθυσμού με επανατοποθέτηση και

$\sigma_{\overline{x}}^2=\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$, χωρίς επανατοποθέτηση.

Εφαρμογή 3

Ένας πληθυσμός περιλαμβάνει τους αριθμούς 3,4,6,8,9. Θεωρούμε όλα τα δυνατά δείγματα μεγέθους 2 που μπορούμε να πάρουμε με επανατοποθέτηση. Να υπολογιστεί η μέση τιμή του πληθυσμού.

ΛΥΣΗ

$$\mu=\frac{3+4+6+8+9}{5}=6$$

ΕΞΑΣΚΗΣΗ

1. Να βρεθεί η τυπική απόκλιση του πληθυσμού για το παραπάνω παράδειγμα.

2. Να βρεθεί η μέση τιμή της δειγματοληπτικής κατανομής της δειγματικής μέσης τιμής του παραπάνω παραδείγματος.

3. Να βρεθεί η τυπική απόκλιση της δειγματικής μέσης τιμής για το ίδιο παράδειγμα.

4. Να λυθούν τα προβλήματα (1), (2) και (3) σε σχέση με το παράδειγμα, αν υποθέσουμε ότι η δειγματοληψία δεν πραγματοποιείται με επανατοποθέτηση.

5. Το βάρος 6.000 μαθητών της Α΄ Λυκείου, έχει μέση τιμή 70 κιλά και τυπική απόκλιση 3,5 κιλά. Εάν ληφθεί

δείγμα 22 μαθητών, να υπολογιστεί :

(α) η αναμενόμενη τιμή της δειγματικής μέσης τιμής,

(β) η τυπική απόκλιση της δειγματικής μέσης τιμής,

αν η δειγματοληψία πραγματοποιείται με επανατοποθέτηση.

6. Να λυθεί η άσκηση (5), αν η δειγματοληψία πραγματοποιείται χωρίς επανατοποθέτηση.
7. Τα βάρη 2.500 πακέτων τσιγάρων ακολουθούν την κανονική κατανομή με μέση τιμή 21gr και τυπική απόκλιση 0,06gr.

   Να βρεθούν:

   (α) η μέση τιμή και

   (β) η τυπική απόκλιση των μέσω τιμών

   200 δειγμάτων με 32 πακέτα το καθένα, δεδομένου ότι πρόκειται για δειγματοληψία με επανατοποθέτηση.

8. Να λυθεί η άσκηση (7) για δειγματοληψία χωρίς επανατοποθέτηση.

9. Οι βαλίτσες των επιβατών εταιρίας αερογραμμών έχουν βάρη με μέση τιμή 18 κιλά και τυπική απόκλιση 4 κιλά. Εάν επιλεγούν 12 βαλίτσες, ποια η πιθανότητα να ζυγίζουν περισσότερο από 200 κιλά (αν προσθέσουμε τα βάρη τους).

10.3. ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Στην πράξη, όταν ένας πληθυσμός έχει μέγεθος μεγαλύτερο του 30, θεωρούμε ότι ακολουθεί την κανονική κατανομή. Στη συνέχεια θα αναφερθούμε σε κανονικούς πληθυσμούς.

Προσπαθώντας να εκτιμήσουμε παραμέτρους του πληθυσμού, είτε:

(α) δίνουμε συγκεκριμένες τιμές σ' αυτές, βάσει της δειγματοληπτικής έρευνας που έχουμε κάνει και έχουμε ***σημειακή εκτίμηση***, είτε:

(β) δίνουμε ένα διάστημα μέσα στο οποίο πιστεύουμε (βάσει της δειγματοληψίας μας) ότι βρίσκεται η τιμή της παραμέτρου που μας ενδιαφέρει και έχουμε εκτίμηση διαστήματος της παραμέτρου.

Πιο συγκεκριμένα, θα ασχοληθούμε με διαστήματα εμπιστοσύνης (δ.ε.) για το μέσο μ και τη διασπορά σ^2.

Α. Διάστημα εμπιστοσύνης για το μέσο μ, όταν η διασπορά σ^2 είναι γνωστή.

Η τυχαία μεταβλητή $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{\nu}}$ ακολουθεί την τυποποιημένη κατανομή N(0,1).

Άρα, $P\left[-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right] = 1 - \alpha$ και το δ.ε. για το μ, θα είναι το:

$$I = \left[\bar{x} - \frac{\sigma}{\sqrt{\nu}} Z_{\alpha/2},\ \bar{x} + \frac{\sigma}{\sqrt{\nu}} Z_{\alpha/2}\right]$$

δηλαδή, το μ θα βρίσκεται μεταξύ των τιμών $\bar{x}-\frac{\sigma}{\sqrt{\nu}}Z_{\alpha/2}$, $\bar{x}+\frac{\sigma}{\sqrt{\nu}}Z_{\alpha/2}$ με πιθανότητα (1 – α) · 100%.
Το πλάτος του διαστήματος εμπιστοσύνης για το μ, ισούται με $\frac{2\sigma}{\sqrt{\nu}}Z_{\alpha/2}$.

Και αν θέλουμε μικρότερο δ.ε., δηλαδή μεγαλύτερη ακρίβεια, έχουμε:

$$\frac{2\sigma}{\sqrt{\nu}}Z_{\alpha/2} \le d \Leftrightarrow \frac{2\sigma}{d}Z_{\alpha/2} \le \sqrt{\nu} \Leftrightarrow \left(\frac{2\sigma}{d}Z_{\alpha/2}\right)^2 \le \nu$$

ως προϋπόθεση από όπου θα βρούμε το κατάλληλο μέγεθος του δείγματος για να έχουμε τη ζητούμενη ακρίβεια.

Εφαρμογή 4

Ένα διεγερτικό φάρμακο ελέγχεται βάσει της επίδρασής του στην πίεση του αίματος. Οι πιέσεις 20 ατόμων μετρώνται πριν και μετά τη λήψη του φαρμάκου, οπότε λαμβάνονται οι διαφορές που φαίνονται στον ακόλουθο πίνακα:

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

Αν η διαφορά πιέσεων έχει βρεθεί από προηγούμενες μελέτες ότι ακολουθεί την κανονική κατανομή με διασπορά σ^2=25, να βρεθεί δ.ε. 95% της μέσης διαφοράς μ, της πίεσης του αίματος.

ΛΥΣΗ

Έστω $\bar{x}$ η μέση τιμή του δείγματος, την οποία και υπολογίζουμε:

$$\bar{x} = \frac{0-1-4-6+7+2+6+7+0+0+8+6-9-5+1+6-9-2+1+4}{20}$$

$$\bar{x} = 0,6$$

Επίσης, η $\bar{x}$ ακολουθεί την κανονική κατανομή Ν $\left(\mu, \frac{\sigma^2}{n}\right)$ (n=20, μέγεθος δείγματος).

Τότε η τυχαία μεταβλητή $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ ακολουθεί την τυποποιημένη κανονική κατανομή.

Αφού θέλουμε δ.ε. 95%, θα είναι:

$$1-\alpha = 0,95 \Rightarrow \alpha = 1-0,95 \Rightarrow \alpha = 0,05$$

Επομένως το ζητούμενο δ.ε. είναι το :

$$\left[\bar{x}-Z_{0,05/2}\frac{\sigma}{\sqrt{n}}, \ \bar{x}+Z_{0,05/2}\frac{\sigma}{\sqrt{n}}\right]$$

ή

$$\left[0,6-1,96\frac{5}{\sqrt{20}}, \ \ 0,6+1,96\frac{5}{\sqrt{20}}\right]$$

ή [–1,59, 2,79]

Δηλαδή, το μ (του πληθυσμού) βρίσκεται μεταξύ των τιμών -1,59 και 2,79 με πιθανότητα 95%.

Σημείωση:
Το κάτω άκρο του διαστήματος εμπιστοσύνης συμβολίζεται με L και το άνω άκρο με U. Οπότε, για το παράδειγμα παραπάνω, θα λέγαμε ότι:

L = -1,59 και U = 2,79

B. Διάστημα εμπιστοσύνης για το μέσο μ όταν η διασπορά σ^2 είναι άγνωστη.

Στις περισσότερες περιπτώσεις, η διακύμανση του πληθυσμού δεν είναι γνωστή, όπως υποθέσαμε στο 1, οπότε δεν μπορούμε να χρησιμοποιήσουμε την τυχαία μεταβλητή $Z = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$, γιατί περιέχει την άγνωστη παράμετρο σ.

Για να βρούμε ένα δ.ε. για το μ, εκτιμούμε την σ^2. Σαν εκτιμήτρια συνάρτηση της σ^2 χρησιμοποιούμε τη δειγματική διακύμανση S^2, με:

$$S^2 = \frac{\sum \left(x_i - \overline{x} \right)^2}{n - 1}$$

και χρησιμοποιούμε την τυχαία μεταβλητή

$$t = \frac{\overline{x} - \mu}{S / \sqrt{n}}$$

που ακολουθεί την κατανομή Student με v=n-1 βαθμούς ελευθερίας.
Άρα:

$$P\left[-t_{v,\alpha} < t < t_{v,\alpha} \right] = 1 - \alpha$$

και το δ.ε. για το μ θα είναι:

$$\left[\overline{x} - t_{v,\alpha} \frac{S}{\sqrt{n}}, \ \overline{x} + t_{v,\alpha} \frac{S}{\sqrt{n}} \right]$$

Σημείωση:
Αν το δείγμα είναι μεγάλο, τότε αντί της κατανομής Student, χρησιμοποιούμε την τυποποιημένη κανονική και το δ.ε. για το μέσο μ με σ^2 άγνωστο, γίνεται :

$$\left[\overline{x} - z_{\alpha/2} \frac{S}{\sqrt{n}}, \ \overline{x} + z_{\alpha/2} \frac{S}{\sqrt{n}} \right]$$

Εφαρμογή 5

Κάτι ιδιαίτερα διαδεδομένο τις τελευταίες δεκαετίες, είναι το test IQ, δηλαδή τεστ νοημοσύνης. Σε δείγμα 20 ατόμων κάναμε το ίδιο test IQ και οι μετρήσεις φαίνονται στον ακόλουθο πίνακα:

73	85	100	82	95
97	120	119	131	112
128	95	77	88	120
89	110	107	84	128

Υποθέτουμε ότι πρόκειται για κανονικό πληθυσμό. Να βρεθεί δ.ε. 95% για το μέσο μ.

ΛΥΣΗ

Υπολογίζουμε το $\overline{x}$:

$$\overline{x} = \frac{73+85+100+82+95+97+120+119+131+112+128+95+77+88+120+89+110+107+84+128}{20}$$

$$\overline{x} = 102$$

Υπολογίζουμε το S:

$$S^2 = \frac{(73-102)^2 + (85-102)^2 + \ldots + (84-102)^2 + (128-102)^2}{20-1}$$

$$S^2 = \frac{6030}{19}$$

$$S^2 = 317,36842$$

$$S = 17,8 \text{ περίπου}$$

Άρα το ζητούμενο δ.ε. για το μ, είναι:

$$\left[\overline{x} - t_{v,\alpha}\frac{S}{\sqrt{n}},\ \ \overline{x} + t_{v,\alpha}\frac{S}{\sqrt{n}}\right]$$

ή $$\left[102 - t_{19,\,0,05}\frac{17,8}{\sqrt{20}},\ \ 102 + t_{19,\,0,05}\frac{17,8}{\sqrt{20}}\right]$$

ή $$\left[102 - 2,093\frac{17,8}{\sqrt{20}},\ \ 102 + 2,093\frac{17,8}{\sqrt{20}}\right]$$

ή $$[102 - 8,33\ ,\ \ 102 + 8,33]$$

ή $$[93,67\ ,\ \ 110,33]$$

Σημείωση: χρησιμοποιήθηκε η Student κατανομή, αφού n=20<50.

ΕΞΑΣΚΗΣΗ

1. Η μέση τιμή και η τυπική απόκλιση των μέγιστων φορτίων που άντεξαν συρματόσχοινα συγκεκριμένου τύπου, είναι σε τόνους 12 και 0,4 αντίστοιχα. Να βρεθεί δ.ε. 95% για το μέσο μ αντοχής των συρματόσχοινων αυτών.

2. Να βρεθεί δ.ε. 99% για το παραπάνω πρόβλημα.

3. Επιθυμούμε να εκτιμήσουμε το μέσο βάρος των παραγόμενων γκρέιπφρουτ σε βιολογική καλλιέργεια. Ποιο πρέπει να είναι το απαιτούμενο μέγεθος του τυχαίου δείγματος που θα επιλέξουμε, ώστε με πιθανότητα 95% το σφάλμα από τη δειγματοληψία να είναι μικρότερο του 0,5 κιλού, δεδομένου ότι η διακύμανση είναι $\sigma^2=2{,}96$;

4. Να κατασκευαστεί δ.ε. για το μέσο μ κανονικού πληθυσμού αν $n=250$, $\mu=0{,}72642$ και $\sigma^2=0{,}0064$.

5. Να κατασκευαστεί δ.ε. για το μέσο μ κανονικού πληθυσμού, αν $n=250$, $\mu=0{,}72642$ και $S^2=0{,}0064$.

6. Η τυπική απόκλιση της ζωής ενός δείγματος λυχνιών είναι 49 ώρες. Πόσο μεγάλο πρέπει να είναι το δείγμα, ώστε να είμαστε 95% σίγουροι ότι το σφάλμα δεν υπερβαίνει τις 9 ώρες;

7. Να λυθεί το πρόβλημα (6), ώστε να είμαστε:
   (α) 99% σίγουροι,
   (β) 99,73% σίγουροι.

8. Αν δείγμα μεγέθους 20 από κανονικό πληθυσμό ακολουθεί την N(100, 225), να βρεθεί δ.ε. 95% για το μέσο μ του πληθυσμού.

9. Αν δείγμα μεγέθους 20 από κανονικό πληθυσμό έχει μέση τιμή 100 και $S^2=225$, να βρεθεί δ.ε. 95% για το μέσο μ του πληθυσμού.

10. Αν δείγμα μεγέθους 80 από κανονικό πληθυσμό έχει μέση τιμή 100 και $S^2=225$, να βρεθεί δ.ε. 95% για το μέσο μ του πληθυσμού.

Σύνοψη

Για τη συλλογή στατιστικών στοιχείων που αναφέρονται σε ορισμένες χαρακτηριστικές ιδιότητες ενός πληθυσμού εφαρμόζουμε διάφορες στατιστικές μεθόδους, οι σπουδαιότερες από τις οποίες είναι η απογραφή, η δειγματοληψία. Σε περίπτωση που εφαρμόζουμε τη μέθοδο της δειγματοληψίας που ουσιαστικά στηρίζεται στην προσπάθεια να γνωρίσουμε τις ιδιότητες ενός πληθυσμού εξετάζοντας μόνο ένα μικρό αριθμό στατιστικών μονάδων που το ονομάζουμε δείγμα.

Προκύπτει όμως το ερώτημα ποια είναι η αξία της τυχαίας δειγματοληψίας που βασίζεται σ' ένα τυχαίο δείγμα;

Η απάντηση είναι η εξής: ένα τυχαίο δείγμα είναι πιθανό να μας δώσει μία τελείως ανακριβή εικόνα του πληθυσμού που πήραμε για να μελετήσουμε τα περισότερα όμως τυχαία δείγματα που παίρνουμε από τον ίδιο πληθυσμό μπορεί να μας δώσουν μια εικόνα των χαρακτηριστικών ιδιοτήτων του πληθυσμού με ικανοποιητική ακρίβεια. Η μεθοδολογία με την οποία εξάγουμε συμπεράσματα για τις παραμέτρους ενός πληθυσμού με βάση τις περιορισμένες πληροφορίες ενός τυχαίου δείγματος ορίζει την επιστήμη της επαγωγικής στατιστικής. Ένας από ους τρόπους εκτίμησης των παραμέτρων του πληθυσμού αποτελούν τα διαστήματα εμπιστοσύνης. Το διάστημα εμπιστοσύνης σχετίζεται άμεσα με την επιλογή του δείγματος και μας δίνει πληροφορίες για το μέγεθος του δείγματος αναλογικά και με την αξιοπιστία που επιδιώκεται.

Βιβλιογραφία / Internet

«Μεθοδολογία δειγματοληψίας: Τεχνικές και εφαρμογές», Χαράλαμπος Χ. Δαμιανού

«Πιθανότητες και Στατιστική», Spiegel, Mc-Graw-Hill, ΕΣΠΙ

«Μεθοδολογία Εκπαιδευτικής Έρευνας», Κ. και Ε. Παπαναστασίου, Λευκωσία

«Στατιστική», Πέτρος Α. Κίοχος, INTERBOOKS

«The Elements of Probability theory and some of its application», Cramer H, New York, 1973

www.stats.gr: *site με πληροφορίες για δραστηριότητες μιας στατιστικής εταιρίας σε σχέση με επιστημονική έρευνα, εκπαίδευση και ανάλυση δεδομένων.*

11. ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Βασικές έννοιες:
- απλή γραμμική παλινδρόμηση
- μέθοδος ελαχίστων τετραγώνων
- ευθεία παλονδρόμηση
- εκτιμήτριες

Στόχος του μαθήματος: Ο εκπαιδευόμενος να μπορεί να δημιουργήσει μία μαθηματική έκφραση (εξίσωση) που να συνδέει δύο μεταβλητές και να έχει τη δυνατότητα να προσδιορίσει άμεσα την αλληλοσυσχέτιση αυτών των μεταβλητών. Βάσει του συγκεκριμένου διαγράμματος.

Προσδοκώμενα αποτελέσματα: Ο εκπαιδευόμενος να αναληφθεί την έννοια της προσέγγισης, όπως παρουσιάζεται στη γραμμική παλινδρόμηση και να την εκφράζει σε απλά προβλήματα.

Εισαγωγικές παρατηρήσεις: Οι μέθοδοι που παρατίθενται στο συγκεκριμένο κεφάλαιο αφορούν πρακτικές οι οποίες εφαρμόζονται καθημερινά σε πλήθος επιστημονικών πεδίων. Η συμβολή τους στην έρευνα είναι τέτοια που επιβάλλει την εκμαση τους από οποιονδήποτε που ασχολείται με τη Στατιστική. Βάσει διαγραμμάτων που περιγράφουν συγκεκριμένα φαινόμενα έχουμε την δυνατότητα χρησιμοποιόντας την παλινδρόμηση να γενικεύσουμε και να προβλέψουμε αποτελέσματα σε σχέση με πρακτικά ζητήματα.

Στα προβλήματα Στατιστικής που είδαμε έως τώρα, εξετάζουμε κάθε φορά ένα δείγμα ως προς μία μόνο μεταβλητή, π.χ. το βάρος των μαθητών, η επίδοση ενός μαθητή στα μαθήματα, οι βεβαιωθέντες θάνατοι από ναρκωτικά μία συγκεκριμένη χρονική περίοδο κ.α. Ιδιαίτερα ενδιαφέρουσα όμως είναι η μελέτη δύο ή περισσότερων μεταβλητών ταυτόχρονα, ώστε να μπορούμε να προσδιορίσουμε την αλληλοσυσχέτιση των μεταβλητών αυτών.

Για παράδειγμα:

(α) Τα κρούσματα άσθματος σε μία περιοχή και η ατμοσφαιρική ρύπανση της περιοχής αυτής. Όσο μεγαλύτερη είναι η ατμοσφαιρική ρύπανση (X) μίας περιοχής, τόσο περισσότερα είναι τα κρούσματα άσθματος (Y) στην περιοχή αυτή.

Στην περίπτωση αυτή λέμε ότι μεταξύ αυτών των μεταβλητών, υπάρχει θετική συσχέτιση.

(β) Η τιμή πώλησης ενός προϊόντος και ο αριθμός των πωλήσεών του. Όσο αυξάνεται η τιμή πώλησης (Χ) ενός προϊόντος, τόσο ελαττώνονται οι πωλήσεις του (Υ).

Στην περίπτωση αυτή λέμε ότι μεταξύ των μεταβλητών υπάρχει αρνητική συσχέτιση.

Στα προβλήματα αυτά, εστιάζουμε το ενδιαφέρον μας στον προσδιορισμό μίας εξίσωσης μεταξύ των μεταβλητών Χ και Ψ, ώστε να μπορούμε να προβλέψουμε τις τιμές της μεταβλητής Ψ, γνωρίζοντας τις μεταβολές της μεταβλητής Χ.

Έτσι λέμε ότι ενδιαφερόμαστε για την παλινδρόμηση της μεταβλητής Ψ πάνω στη Χ. Η μεταβλητή Χ ονομάζεται ανεξάρτητη μεταβλητή, ενώ η μεταβλητή Ψ εξαρτημένη.

Ο κλάδος της Στατιστικής που ασχολείται με αυτόν τον τομέα, ονομάζεται ανάλυση Παλινδρόμησης.

11.1. ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Η πιο απλή μορφή παλινδρόμησης είναι η γραμμική. Στη γραμμική παλινδρόμηση υπάρχει μία ανεξάρτητη μεταβλητή Χ και η εξαρτημένη μεταβλητή Ψ προσεγγίζεται ικανοποιητικά από μία γραμμική συνάρτηση του Χ.

Για παράδειγμα, ο παρακάτω πίνακας περιέχει το ποσοστό διοξειδίου του άνθρακα (Χ) στην ατμόσφαιρα σε διάφορες περιοχές και τα καταγεγραμμένα κρούσματα άσθματος (Ψ) στις περιοχές αυτές.

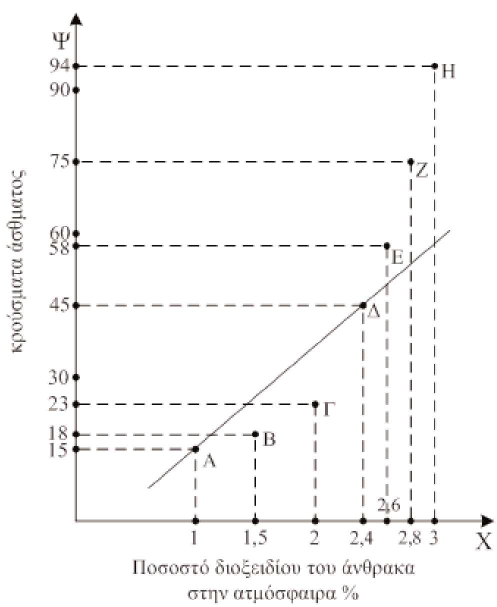
Πίνακας 1

ΠΟΛΗ	Α	Β	Γ	Δ	Ε	Ζ	Η
Διοξείδιο του άνθρακα στην ατμόσφαιρα %	1	1,5	2	2,4	2,6	2,8	3
Κρούσματα άσθματος Ψ	15	18	23	45	58	75	94

Αν παραστήσουμε τα ζεύγη (x,y) σε ένα ορθογώνιο σύστημα αξόνων, προκύπτει το διπλανό σχήμα που ονομάζεται διάγραμμα διασποράς. Αν τώρα με το μάτι βρούμε δύο σημεία του διαγράμματος που ενώνοντάς τα με ευθεία γραμμή, αυτή είναι όσο το δυνατόν πιο κοντά στα υπόλοιπα σημεία του διαγράμματος, τότε λέμε ότι έχουμε κατασκευάσει την ευθεία παλινδρόμηση της Ψ πάνω στη Χ. Η εξίσωση μίας ευθείας είναι της μορφής y=α+βx. Επειδή αυτή διέρχεται από τα σημεία (1,15) και (2,4,45), προκύπτει το παρακάτω σύστημα για τις παραμέτρους α, β:

$$\begin{cases} 15 = \alpha + \beta \\ 45 = \alpha + 2,4 \cdot \beta \end{cases}$$

Επιλύοντας το σύστημα προκύπτει: α= -10 και β=25. Δηλαδή, η ευθεία παλινδρόμησης έχει εξίσωση y= -10+25x

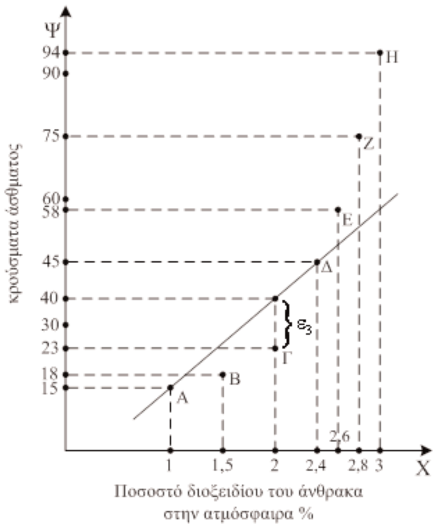


11.2. ΜΕΘΟΔΟΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Η προηγούμενη μέθοδος κατασκευής της ευθείας παλινδρόμησης της Ψ πάνω στη Χ έχει πολλά μειονεκτήματα.

Ας επανέλθουμε στο διάγραμμα διασποράς του προηγούμενου παραδείγματος. Είδαμε ότι η ευθεία παλινδρόμησης έχει εξίσωση y= -10+25x.

Με βάση αυτή την ευθεία, για την πόλη Γ προβλέπονται y=-10+25 2=40
κρούσματα άσθματος, ενώ γνωρίζουμε από τον πίνακα 1, ότι τα κρούσματα στην πόλη Γ είναι 23.



Υπάρχει δηλαδή ένα σφάλμα ε3=40-23=17 κρουσμάτων.

Η μέθοδος ελαχίστων τετραγώνων χρησιμοποιείται για τον προσδιορισμό των παραμέτρων α, β έτσι ώστε να γίνεται ελάχιστο το άθροισμα των τετραγώνων των σφαλμάτων ε_i.

Είναι $\varepsilon_i = y_i - y$, όπου y_i η πραγματική τιμή της μεταβλητής Ψ και y η εκτίμηση από την ευθεία y=α+βx.

Δηλαδή, $\varepsilon_i = y_i - \alpha - \beta x_i$. Άρα, ζητούμε την ελαχιστοποίηση του:

$$\sum_{i=1}^{\nu} \varepsilon_i^2 = \sum_{i=1}^{\nu} (y_i - \alpha - \beta x_i)^2$$

Οι ζητούμενες τιμές των παραμέτρων α, β ονομάζονται εκτιμήτριες ελαχίστων τετραγώνων και συμβολίζεται με $\hat{\alpha}$ (α καπέλο) και $\hat{\beta}$ (β καπέλο). Αποδεικνύεται ότι δίνονται από τις σχέσεις:

$$\boxed{\begin{array}{l} \hat{\beta} = \dfrac{\nu \cdot \sum_{i=1}^{\nu} x_i y_i - \left(\sum_{i=1}^{\nu} x_i\right) \cdot \left(\sum_{i=1}^{\nu} y_i\right)}{\nu \cdot \sum_{i=1}^{\nu} x_i^2 - \left(\sum_{i=1}^{\nu} x_i\right)^2} \\ \hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} \end{array}}$$

όπου $\bar{y} = \dfrac{\sum_{i=1}^{\nu} y_i}{\nu}$ και $\bar{x} = \dfrac{\sum_{i=1}^{\nu} x_i}{\nu}$.

Η ευθεία $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ ονομάζεται ευθεία παλινδρόμησης της Ψ πάνω στη Χ.

Αντικαθιστώντας τις τιμές του πίνακα 1, προκύπτει:

$$\hat{\alpha} = 38{,}47 \text{ και } \hat{\beta} = -37{,}39$$

Η ευθεία παλινδρόμησης του παραδείγματος έχει εξίσωση $\hat{y} = -37,39 + 38,47 \cdot x$. Από τη σχέση $\hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x}$, προκύπτει ότι η ευθεία ελαχίστων τετραγώνων διέρχεται από το σημείο $(\bar{x}, \bar{y})$ και έχει συντελεστή διεύθυνσης $\hat{\beta}$.

Παρατηρήσεις:

1. Από την ευθεία παλινδρόμησης για x=0 προκύπτει:

$$\hat{y} = -37,39 + 38,47 \cdot 0 = -37,39$$

Δηλαδή, αν η ατμόσφαιρα μίας πόλης δεν περιείχε διοξείδιο του άνθρακα, τότε θα υπήρχαν -37,39 κρούσματα άσθματος.

Το παράδοξο αυτό προκύπτει επειδή η τιμή 0 της ανεξάρτητης μεταβλητής βρίσκεται έξω από το διάστημα τιμών που παρέχουν τα δεδομένα του πίνακα 1.

Γενικά., η ευθεία παλινδρόμησης δίνει ορθές εκτιμήσεις μίας μεταβλητής, όταν οι τιμές της ανεξάρτητης μεταβλητής βρίσκονται εντός του διαστήματος τιμών που παρέχουν τα δεδομένα ή πολύ κοντά στις ακραίες τιμές του διαστήματος.

2. Για τις τιμές x_1 και $x_2=x_1+1$ που διαφέρουν κατά 1, από την ευθεία παλινδρόμησης προκύπτει:

$$y_2 - y_1 = \left(\hat{\alpha} + \hat{\beta}x_2\right) - \left(\hat{\alpha} + \hat{\beta}x_1\right) = \hat{\beta}(x_2 - x_1) = \hat{\beta}$$

Δηλαδή, το $\hat{\beta}$ παριστάνει τη μεταβλητή του Ψ όταν το Χ. Οπότε, όταν το Χ αυξηθεί κατά 1 μονάδα, το Ψ αυξάνεται κατά $\hat{\beta}$ όταν $\hat{\beta}$ >0 και ελαττώνεται κατά $\hat{\beta}$ όταν $\hat{\beta}$ <0.

ΑΣΚΗΣΕΙΣ

1. Στον παρακάτω πίνακα, δίνονται τα ύψη Χ (σε cm) και τα βάρη Ψ (kg) 10 μαθητών μίας τάξης ενός σχολείου.

Μαθητής	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ
Ύψος Χ	170	172	174	175	178	179	180	182	183	187
Βάρος Ψ	60	61	70	68	72	79	80	79	81	90

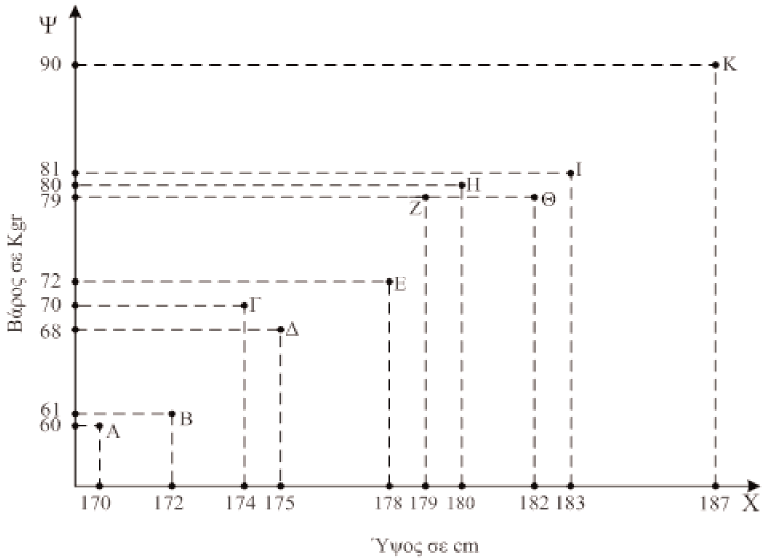
(α) Να γίνει το διάγραμμα διασποράς.

(β) Να υπολογιστούν οι εκτιμήτριες $\hat{\alpha}$, $\hat{\beta}$ και να χαραχθεί η ευθεία παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$.

(γ) Ποιο αναμένετε να είναι το βάρος μαθητή που έχει ύψος 185cm;

(δ) Με βάση την ευθεία παλινδρόμησης του βάρους Ψ πάνω στο ύψος Χ, μπορούμε να εκτιμήσουμε το ύψος ενός μαθητή που έχει βάρος 71kg;

(α)



(β) Για τον υπολογισμό των εκτιμητριών $\hat{\alpha}$, $\hat{\beta}$, κατασκευάζουμε τον παρακάτω πίνακα:

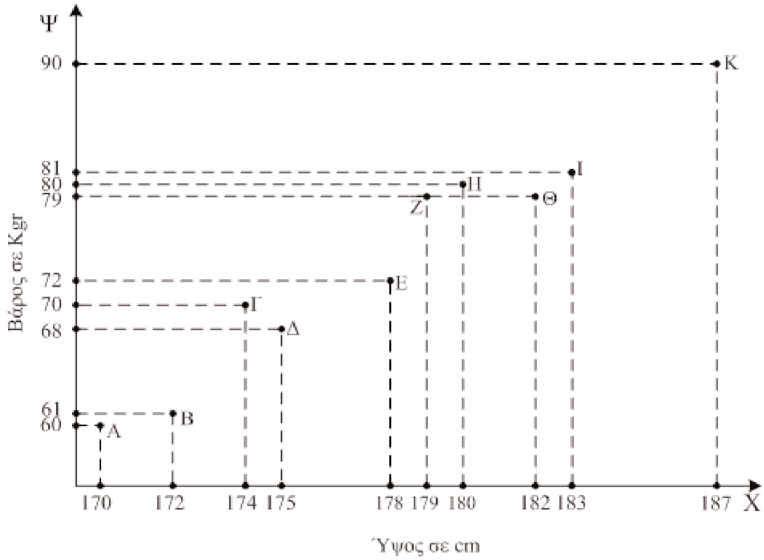
Μαθητής	Ύψος X	Βάρος Ψ	x^2	$x \cdot y$
Α	170	60	28.900	10.200
Β	172	61	29.584	10.492
Γ	174	70	30.276	12.180
Δ	175	68	30.625	11.900
Ε	178	72	31.684	12.816
Ζ	179	79	32.041	14.141
Η	180	80	32.400	14.400
Θ	182	79	33.124	14.378
Ι	183	81	33.489	14.823
Κ	187	90	34.869	16.830
	Σx=1.780	Σy=740	Σx^2=317.092	Σxy=132.160

Είναι $\bar{x} = \frac{\Sigma x}{10} = \frac{1780}{10} = 178$, $\bar{y} = \frac{\Sigma y}{10} = \frac{740}{10} = 74$,

$$\hat{\beta} = \frac{10 \cdot \Sigma x \cdot y - \Sigma x \cdot \Sigma y}{10 \cdot \Sigma x^2 - (\Sigma x)^2} = \frac{10 \cdot 132.160 - 1.780 \cdot 740}{10 \cdot 317.092 - (1.780)^2} = 1,75$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} = 74 - 1,75 \cdot 178 = -237,5$$

Η ευθεία παλινδρόμησης έχει εξίσωση $\hat{y} = -237,5 + 1,75 \cdot x$ Για να την κατασκευάσουμε θα βρούμε δύο σημεία της. Το ένα είναι το $(\bar{x}, \bar{y})$, δηλαδή το (178, 74). Το άλλο θα βρεθεί δίνοντας μία τυχαία τιμή στο X. Για x=180, είναι $\hat{y} = 77,5$, δηλαδή (180, 77,5).



(γ) Για x=185cm, είναι $\hat{y} = -237,5 + 1,75 \cdot 185 = 86,25\text{kgr}$. Δηλαδή, για μαθητή ύψους 185cm, εκτιμάται βάρος 86,25kgr.

(δ) Επειδή η ευθεία παλινδρόμησης του βάρους Ψ πάνω στο ύψος Χ δίνει εκτιμήσεις για το βάρος Ψ με ανεξάρτητη μεταβλητή το Χ, δεν μπορεί να χρησιμοποιηθεί για εκτίμηση του Χ όταν δίνεται το Ψ.

2. Στον παρακάτω πίνακα δίνονται η τιμή πώλησης Χ(€) το κιλό φράουλες, σε ένα μανάβικο και οι πωλήσεις Ψ (kgr) που πραγματοποιεί:

Τιμή κιλού Χ €	1,20	1,30	1,50	1,70	1,90	2
Πωλήσεις Ψ kgr	70	65	60	50	35	14

(α) Να βρείτε τις ευθείες παλινδρόμησης της Ψ πάνω στη Χ και της Χ πάνω στην Ψ.
(β) Να εκτιμήσετε τις πωλήσεις φράουλας που θα κάνει το μανάβικο όταν τις πουλά προς 1,40€ το κιλό.
(γ) Να εκτιμήσετε την τιμή πώλησης της φράουλας αν γνωρίζετε ότι ο μανάβης θέλει να διαθέσει 40kgr προς πώληση.

Απάντηση

(α) Για τον υπολογισμό των εκτιμητριών $\hat{\alpha}$, $\hat{\beta}$ της ευθείας παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ της Ψ πάνω στη Χ, κατασκευάζουμε τον παρακάτω πίνακα:

Τιμή κιλού Χ €	Πωλήσεις Ψ kgr	x^2	$x \cdot y$
1,20	70	1,44	84
1,30	65	1,69	84,5
1,50	60	2,50	90
1,70	50	2,89	85
1,90	35	3,61	66,5
2	14	4	28
Σx=9,60	Σy=294	Σx2=16,13	Σxy=438

Είναι: $\bar{x} = \frac{\Sigma x}{6} = \frac{9,60}{6} = 1,60$, $\bar{y} = \frac{\Sigma y}{6} = \frac{294}{6} = 49$

$$\hat{\beta} = \frac{6 \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{6 \cdot \Sigma x^2 - (\Sigma x)^2} = \frac{6 \cdot 438 - 9,60 \cdot 294}{6 \cdot 16,13 - (9,60)^2} = 42,1$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 49 - 42,1 \cdot 1,60 = -18,36$$

Η ευθεία παλινδρόμησης της Ψ πάνω στη Χ είναι:

$$\hat{y} = -18,36 + 42,1 \cdot x$$

Για να βρούμε την ευθεία παλινδρόμησης της Χ πάνω στην Ψ, πρέπει οι μεταβλητές Χ, Ψ να αντιστρέψουν τους ρόλους τους. Έτσι, η ευθεία παλινδρόμησης, στην περίπτωση αυτή, έχει εξίσωση $\hat{x} = \widehat{\alpha'} + \widehat{\beta'} \cdot y$, με :

$$\widehat{\beta'} = \frac{\nu \cdot \Sigma xy - (\Sigma x) \cdot (\Sigma y)}{\nu \cdot \Sigma y^2 - (\Sigma y)^2} \text{ και } \widehat{\alpha'} = \bar{x} - \widehat{\beta'}\bar{y}$$

Για τον υπολογισμό των εκτιμητριών $\widehat{\alpha'}$ και $\widehat{\beta'}$, κατασκευάζουμε τον παρακάτω πίνακα:

Χ	Ψ	y^2	$x \cdot y$
1,20	70	4.900	84
1,30	65	4.225	84,5
1,50	60	3.600	90
1,70	50	2.500	85
1,90	35	1.225	66,5
2	14	196	28
Σx=9,60	Σy=294	Σy^2=16.646	Σxy=438

Είναι: $\widehat{\beta'} = \dfrac{6 \cdot 438 - 9,60 \cdot 294}{6 \cdot 16.646 - (294)^2} = 0,014$ και $\widehat{\alpha'} = 1,60 - 0,014 \cdot 49 = 0,914$.

Επομένως, η ευθεία παλινδρόμησης της Χ πάνω στην Ψ έχει εξίσωση:

$$\hat{x} = 0,914 + 0,014 \cdot y$$

(β) για x=1,40€, από την ευθεία παλινδρόμησης $\hat{y} = -18,36 + 42,1 \cdot x$ προκύπτει:

$$\hat{y} = -18,36 + 42,1 \cdot 1,40 = 40,58\text{kgr}$$

(γ) Για y=40kgr, από την ευθεία παλινδρόμησης $\hat{x} = 0,914 + 0,014 \cdot y$, προκύπτει:

$$\hat{x} = 0,914 + 0,014 \cdot 40 = 1,474 \text{ €}$$

3. Σε μία πόλη υπάρχουν δύο αγροτικοί συνεταιρισμοί διαχείρισης βαμβακιού. Η παραγωγή Ψ (τόνοι) του 1^{ου} συνεταιρισμού, σε σχέση με την καλλιεργήσιμη έκταση Χ (στρέμματα) που διαχειρίζεται, εκτιμάται από την ευθεία παλινδρόμησης $\hat{y} = -1,5 + x$ και αντίστοιχα για τον δεύτερο συνεταιρισμό εκτιμάται από την ευθεία $\hat{y} = -3,5 + 1,5x$. Αν η συνολική έκταση που διαχειρίζονται οι δύο συνεταιρισμοί είναι 250 στρέμματα και για το 2006 εκτιμάται συνολική παραγωγή 300 τόνων βαμβάκι, να εκτιμήσετε την έκταση που διαχειρίζεται ο κάθε συνεταιρισμός.

Απάντηση

Έστω ότι ο 1^{ος} συνεταιρισμός διαχειρίζεται έκταση α στρεμμάτων, τότε ο 2^{ος} συνεταιρισμός διαχειρίζεται έκταση (250-α) στρεμμάτων.
Η παραγωγή βαμβακιού που εκτιμάται από τον 1^{ο} συνεταιρισμό για το 2006 είναι: $\hat{y} = -1,5 + \alpha$ τόνοι, ενώ για τον 2^{ο} συνεταιρισμό είναι: $\hat{y} = -3,5 + 1,5(250 - \alpha)$. Επειδή εκτιμάται συνολική παραγωγή 300 τόνων, έχουμε:

$$-1,5 + \alpha - 3,5 + 1,5 \cdot (250 - \alpha) = 300 \text{ ή}$$
$$-5 + \alpha + 375 - 1,5\alpha = 300 \text{ ή}$$
$$-0,5\alpha = -70 \text{ ή}$$
$$\alpha = 140 \text{ στρέμματα}$$

Δηλαδή ο 1^{ος} συνεταιρισμός διαχειρίζεται 140 στρέμματα και ο 2^{ος} 250-140=110 στρέμματα.

ΕΞΑΣΚΗΣΗ

1. Η εκτιμήτρια $\hat{\beta}$ της ευθείας παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ παριστάνει τη μεταβολή της μεταβλητής Ψ, όταν το Χ μεταβάλλεται κατά:

   Α. $\hat{\beta}$ μονάδες Β. $\hat{\alpha}$ μονάδες Γ. 1 μονάδα

   Δ. 10 μονάδες Ε. 5 μονάδες

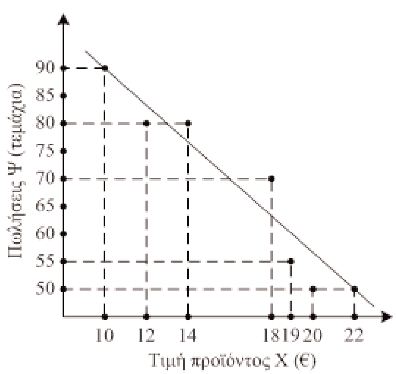
2. Με βάση την ευθεία παλινδρόμησης $\hat{y} = -8 + 2,5 \cdot x,\ x \in [5,20]$, η εκτίμηση για το y όταν x=8, είναι:

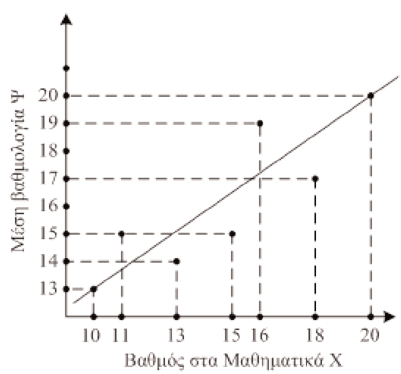
   Α. 8 Β. 9 Γ. 10 Δ. 11 Ε. 12

3. Ερωτήσεις του τύπου Σωστό - Λάθος:

	Σωστό	Λάθος
(α) Η ευθεία παλινδρόμησης της μεταβλητής Χ πάνω στην Ψ έχει εξίσωση: $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$	☐	☐
(β) Η ευθεία παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ διέρχεται από το σημείο $(\bar{x}, \bar{y})$	☐	☐
(γ) Η εκτιμήτρια $\hat{\alpha}$ της ευθείας παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ παριστάνει την τιμή της μεταβλητής Ψ όταν x=0.	☐	☐
δ) Όταν $\hat{\alpha} \neq 0$, η ευθεία $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ διέρχεται από την αρχή των αξόνων.	☐	☐
(ε) Η εκτιμήτρια $\hat{\beta}$ της ευθείας παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ παριστάνει τη μεταβλητή του Ψ όταν το Χ αυξηθεί κατά 1 μονάδα.	☐	☐

4. Στα παρακάτω διαγράμματα διασποράς, έχουν χαραχθεί «με το μάτι» οι ευθείες παλινδρόμησης. Να βρείτε τις εξισώσεις y = αx + β των ευθειών αυτών.





5. Για καθέναν από τους παρακάτω πίνακες τιμών, να βρείτε τις ευθείες παλινδρόμησης της μεταβλητής Ψ πάνω στη Χ και της Χ πάνω στην Ψ.

(α)

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

(β)

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

6. Οι επιδόσεις 5 μαθητών στα Μαθηματικά και τη Φυσική, δίνονται στον παρακάτω πίνακα:

Βαθμός στα Μαθηματικά Χ	11	13	15	18	19
Βαθμός στη Φυσική Ψ	10	13	17	19	20

(α) Να βρείτε τις εκτιμήτριες $\hat{\alpha}$, $\hat{\beta}$ και να χαράξετε στο διάγραμμα διασποράς την ευθεία παλινδρόμησης $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$.

(β) Να εκτιμήσετε το βαθμό στη Φυσική που έχει ένας μαθητής που η βαθμολογία του στα Μαθηματικά είναι 17.

7. Οι πωλήσεις Ψ (τεμάχια) ενός φορητού υπολογιστή σε σχέση με την τιμή πώλησης του Χ (€), σε ένα συγκεκριμένο κατάστημα, δίνονται στον παρακάτω πίνακα:

Τιμή πώλησης Χ €	1000	900	850	800	650
Πωλήσεις Ψ τεμάχια	5	8	13	22	42

(α) Να βρείτε την ευθεία παλινδρόμησης των πωλήσεων πάνω στην τιμή πώλησης του υπολογιστή.

(β) Να εκτιμήσετε τις πωλήσεις του φορητού υπολογιστή στο συγκεκριμένο κατάστημα, αν η τιμή πώλησής του είναι 700 €.

(γ) Να εκτιμήσετε πόσο αυξάνονται οι πωλήσεις για κάθε ευρώ που μειώνεται η τιμή πώλησης του υπολογιστή.

8. Αν για δύο μεταβλητές Χ, Ψ ισχύει:

$$\sum_{i=1}^{10} x_i = 40, \quad \sum_{i=1}^{10} y_i = 10, \quad \sum_{i=1}^{10} x_i y_i = 60, \quad \sum_{i=1}^{10} x_i^2 = 150, \quad \sum_{i=1}^{10} y_i^2 = 8$$

να βρείτε τις ευθείες παλινδρόμησης της Ψ πάνω στην Χ και της Χ πάνω στην Ψ.

9. Με βάση τον διπλανό πίνακα, να βρείτε την τιμή της παραμέτρου Κ, ώστε η ευθεία παλινδρόμησης του Ψ πάνω στο Χ, να έχει εξίσωση: $\hat{y}=\frac{1}{10}+\frac{1}{2}x$.

x	y
1	1
2	K
3	1
4	2K
5	3

10. Σε έρευνα που έγινε στους μαθητές ενός σχολείου για το πώς μεταβάλλεται ο βαθμός τους στη Φυσική σε σχέση με τα Μαθηματικά, διαπιστώθηκε ότι αν ο βαθμός στα Μαθηματικά αυξηθεί κατά 1 μονάδα, τότε ο βαθμός του ίδιου μαθητή στη Φυσική, αυξάνεται κατά 1,5 μονάδα. Αν η μέση βαθμολογία όλων των μαθητών του σχολείου στα Μαθηματικά είναι 13 και στη Φυσική 14, να εκτιμήσετε σύμφωνα με την ευθεία παλινδρόμησης, το βαθμό ενός μαθητή στη Φυσική, όταν ο βαθμός του στα Μαθηματικά είναι 15.

11. Μία εταιρία έχει δύο υποκαταστήματα Α, Β. Για τις πωλήσεις Ψ (τεμάχια) ενός προϊόντος σε σχέση με την τιμή Χ (€) πώλησής του, διαπιστώθηκε ότι για το κατάστημα Α οι πωλήσεις εκτιμώνται από την ευθεία παλινδρόμησης $\hat{y}=116-0,8\cdot x$ και για το κατάστημα Β από την ευθεία $\hat{y}=132-0,9\cdot x$. Αν η τιμή πώλησης του προϊόντος στο κατάστημα Α είναι 10% φθηνότερη από το κατάστημα Β και τα δύο καταστήματα μαζί πούλησαν 86 τεμάχια του προϊόντος, να εκτιμήσετε την τιμή πώλησης του προϊόντος στα δύο καταστήματα.

Σύνοψη

Στις προηγούμενες παραγράφους παρουσιάσαμε τα διάφορα προβλήματα της Στατιστικής εξετάζοντας κάθε φορά μια μόνο μεταβλητή χωριστά.

Στην παράγραφο αυτή εξετάσαμε τη συνάφεια και την αλληλεξάρτηση που υπάρχει μεταξύ δύο μεταβλητών. Η μελέτη της συσχέτισης των μεταβλητών, δηλαδή ο υπολογισμός των τιμών της μιας μεταβλητής από τις τιμές της άλλης γίνεται με τη λεγόμενη στατιστική πρόβλεψη ή ανάλυση παλινδρόμησης που η πιο απλή μορφή της είναι η ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ κατά την οποία υπάρχει μια ανεξάρτητη μεταβλητή Χ και η εξαρτημένη μεταβλητή y η οποία μπορεί να προσεγγιστεί ικανοποιητικά από μια γραμμική σχέση του x. (y = α + βx).

Γραμμική παλινδρόμηση της y πάνω στη x.

- Η πρόβλεψη μιας μελλοντικής εξέλιξης της μεταβλητής y στην περίπτωση που η x εκφράζει χρόνο ή η εκτίμηση της y για κάποια τιμή της x γίνεται με τη βοήθεια της γραμμής εξίσωσης παλινδρόμησης ŷ = Æ + ˆ x .
- Οι εκτιμήτριες Æ, ˆ της ευθείας ŷ = Æ + ˆ x προσδιορίζονται με τη μέθοδο ελαχίστων τετραγώνων.

ΚΑΤΑΝΟΜΗ POISSON

Τιμές της συνάρτησης πιθανότητας

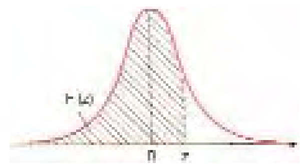
$$P\{X = x\} = \frac{e^{-\lambda}}{x!}\lambda^x$$

x	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$	$\lambda = 5$	$\lambda = 6$	$\lambda = 7$	$\lambda = 8$	$\lambda = 10$	$\lambda = 12$	$\lambda = 20$
0	0,3679	0,1353	0,0498	0,0183	0,0067	0,0025	0,0009	0,0003	0,0000	0,0000	0,0000
1	0,3679	0,2707	0,1494	0,0733	0,0337	0,0149	0,0064	0,0027	0,0005	0,0001	0,0000
2	0,1839	0,2707	0,2240	0,1465	0,0842	0,0446	0,0223	0,0107	0,0023	0,0004	0,0000
3	0,0613	0,1804	0,2240	0,1954	0,1404	0,0892	0,0521	0,0286	0,0076	0,0018	0,0000
4	0,0153	0,0902	0,1680	0,1954	0,1755	0,1339	0,0912	0,0573	0,0189	0,0058	0,0000
5	0,0031	0,0361	0,1008	0,1563	0,1755	0,1606	0,1277	0,0916	0,0378	0,0127	0,0001
6	0,0005	0,0120	0,0504	0,1042	0,1462	0,1606	0,1490	0,1221	0,0631	0,0255	0,0002
7	0,0001	0,0034	0,0216	0,0595	0,1044	0,1377	0,1490	0,1396	0,0901	0,0437	0,0005
8	0,0000	0,0009	0,0081	0,0298	0,0653	0,1033	0,1304	0,1396	0,1126	0,0655	0,0013
9	0,0000	0,0002	0,0027	0,0132	0,0363	0,0688	0,1014	0,1241	0,1251	0,0874	0,0029
10	0,0000	0,0000	0,0008	0,0053	0,0181	0,0413	0,0710	0,0993	0,1251	0,1048	0,0058
11	0,0000	0,0000	0,0002	0,0019	0,0082	0,0225	0,0452	0,0722	0,1137	0,1144	0,0106
12	0,0000	0,0000	0,0001	0,0006	0,0034	0,0113	0,0264	0,0481	0,0948	0,1144	0,0176
13	0,0000	0,0000	0,0000	0,0002	0,0013	0,0052	0,0142	0,0296	0,0729	0,1055	0,0271
14	0,0000	0,0000	0,0000	0,0001	0,0005	0,0022	0,0071	0,0169	0,0521	0,0905	0,0387
15	0,0000	0,0000	0,0000	0,0000	0,0002	0,0009	0,0033	0,0090	0,0347	0,0724	0,0516
16	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0014	0,0045	0,0217	0,0543	0,0646
17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0021	0,0128	0,0383	0,0760
18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0009	0,0071	0,0256	0,0844
19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0037	0,0161	0,0888
20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0019	0,0097	0,0888
21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009	0,0055	0,0846
22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0000	0,0769
23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0000	0,0669
24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0557
25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0446
26	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0343
27	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0254
28	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0181
29	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0125
30	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0083
31	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0054
32	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0034
33	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0020
34	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0012

ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Τιμές της συνάρτησης κατανομής *N* (0,1)

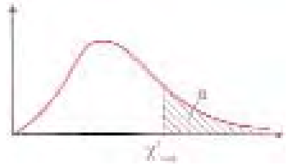
$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

z	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,8	4,0	4,5
F(z)	0,99865	0,99904	0,99931	0,99952	0,99966	0,99976	0,999841	0,999928	0,999968	0,999997

ΚΑΤΑΝΟΜΗ χ^2



Τιμές $\chi^2_{n,a}$ για τις οποίες:

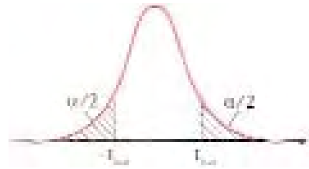
$$P\{\chi^2 > \chi^2_{n,a}\} = a$$

n	a								
	0,9990	0,975	0,950	0,900	0,100	0,050	0,025	0,010	0,001
1	0,0002	0,0010	0,0039	0,0158	2,71	3,84	5,02	6,63	10,83
2	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	13,82
3	0,12	0,22	0,35	0,58	6,25	7,81	9,35	11,34	16,27
4	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	18,47
5	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	20,52
6	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81	22,46
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,47	24,32
8	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	26,13
9	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	27,88
10	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	29,59
11	3,05	3,82	4,57	5,58	17,27	19,67	21,92	24,72	31,26
12	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	32,91
13	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	34,53
14	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	36,12
15	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	37,70
16	5,81	6,91	7,96	9,31	23,54	26,30	28,84	32,00	39,25
17	6,41	7,56	8,67	10,08	24,77	27,59	30,19	33,41	40,79
18	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,80	42,31
19	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	43,82
20	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	45,32
21	8,90	10,28	11,59	13,24	29,61	32,67	35,48	38,93	46,80
22	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29	48,27
23	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64	49,73
24	10,86	12,40	13,85	15,66	33,20	36,41	39,37	42,98	51,18
25	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31	52,62
26	12,20	13,84	15,38	17,29	35,56	38,88	41,92	45,64	54,05
27	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96	55,48
28	13,57	15,31	16,93	18,94	37,92	41,34	44,46	48,28	56,89
29	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59	58,30
30	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	59,70

ΚΑΤΑΝΟΜΗ t

Τιμές $t_{n,a}$ για τις οποίες:

$$P\left(|t| > t_{n,a}\right) = a$$



n	a												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291