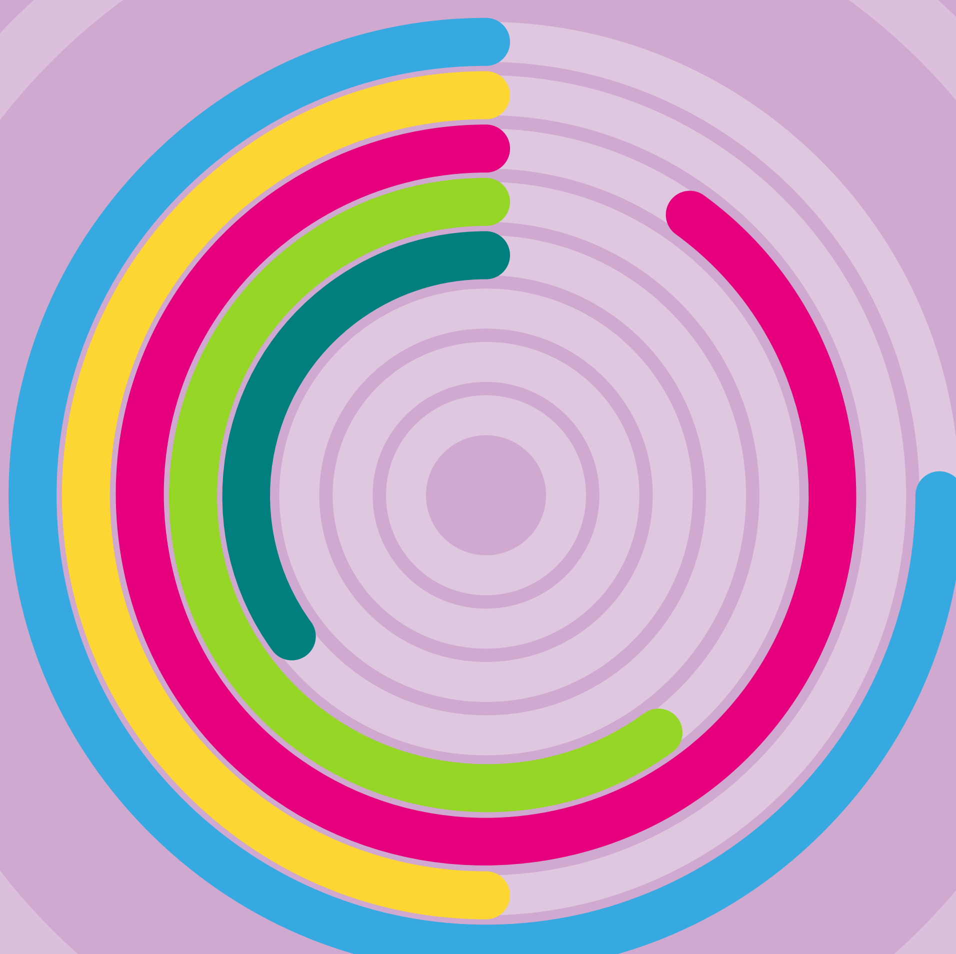
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΘΡΗΣΚΕΥΜΑΤΩΝ ΚΑΙ ΑΘΛΗΤΙΣΜΟΥ
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

ΡΕΚΟΥΜΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ
ΚΑΤΣΑΠΑΣ ΛΑΜΠΡΟΣ
ΚΟΥΜΑΝΤΟΣ ΝΙΚΟΛΑΟΣ
ΡΕΚΟΥΜΗ ΕΛΕΝΗ

ΣΤΟΧΑΣΤΙΚΑ ΜΑΘΗΜΑΤΙΚΑ



Α΄ ΛΥΚΕΙΟΥ

ΙΝΣΤΙΤΟΥΤΟ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΕΚΔΟΣΕΩΝ
«ΔΙΟΦΑΝΤΟΣ»

Α΄ ΛΥΚΕΙΟΥ

ΣΤΟΧΑΣΤΙΚΑ ΜΑΘΗΜΑΤΙΚΑ

Α΄ ΛΥΚΕΙΟΥ

Επιστημονική Επιτροπή Αξιολόγησης

Συντονιστής / Αξιολογητής	**Φουσκάκης Δημήτριος** Εν ενεργεία μέλος Διδακτικού Ερευνητικού Προσωπικού Πανεπιστημίου
Αξιολογητής	**Βενάρδος Παντελής** Εν ενεργεία Εκπαιδευτικός
Αξιολογητής	**Παναγιώτου Κωνσταντίνος** Εν ενεργεία Εκπαιδευτικός
Τεχνικός Εμπειρογνώμονας	**Παπατσιμούλη Μαρία** Πτυχιούχος Πληροφορικής
Επικουρικός Εμπειρογνώμονας	**Γαλάτη Αικατερίνη** Πτυχιούχος Τεχνολογίας Γραφικών Τεχνών
Υπεύθυνη του μαθήματος/γνωστικού αντικειμένου στο πλαίσιο της Πράξης	**Ειρήνη Γεωργάκη**, Σύμβουλος Α΄ ΙΕΠ, μέλος της Επιστημονικής Ομάδας Έργου (ΕΟΕ) της Πράξης

Πράξη με τίτλο: «Συγγραφή, Αξιολόγηση και Ένταξη διδακτικών βιβλίων στο Μητρώο Διδακτικών Βιβλίων και στην Ψηφιακή Βιβλιοθήκη Διδακτικών Βιβλίων» με κωδικό ΟΠΣ 6010165 στο Πρόγραμμα «Ανθρώπινο Δυναμικό και Κοινωνική Συνοχή» 2021-2027

ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

Σπυρίδων Δουκάκης

Πρόεδρος του Δ.Σ. του Ινστιτούτου Εκπαιδευτικής Πολιτικής

Υπεύθυνη Πράξης

Πολυξένη Μπίλλα

Σύμβουλος Α΄ του Ινστιτούτου Εκπαιδευτικής Πολιτικής

Προϊσταμένη Τμήματος Β΄ Προγραμμάτων Σπουδών και Εκπαιδευτικού Υλικού

Αναπληρώτρια Υπεύθυνη Πράξης

Άννα-Αικατερίνη Λυκούρη

Σύμβουλος Α΄ του Ινστιτούτου Εκπαιδευτικής Πολιτικής

«Με τη συγχρηματοδότηση της Ευρωπαϊκής Ένωσης»

και το Πρόγραμμα «Ανθρώπινο Δυναμικό και Κοινωνική Συνοχή»

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΘΡΗΣΚΕΥΜΑΤΩΝ ΚΑΙ ΑΘΛΗΤΙΣΜΟΥ
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

ΡΕΚΟΥΜΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ
ΚΑΤΣΑΠΑΣ ΛΑΜΠΡΟΣ
ΚΟΥΜΑΝΤΟΣ ΝΙΚΟΛΑΟΣ
ΡΕΚΟΥΜΗ ΕΛΕΝΗ

ΣΤΟΧΑΣΤΙΚΑ ΜΑΘΗΜΑΤΙΚΑ

Α΄ ΛΥΚΕΙΟΥ

ΙΝΣΤΙΤΟΥΤΟ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΕΚΔΟΣΕΩΝ
«ΔΙΟΦΑΝΤΟΣ»

ΣΥΓΓΡΑΦΕΙΣ

ΡΕΚΟΥΜΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ
ΚΑΤΣΑΠΑΣ ΛΑΜΠΡΟΣ
ΚΟΥΜΑΝΤΟΣ ΝΙΚΟΛΑΟΣ
ΡΕΚΟΥΜΗ ΕΛΕΝΗ

ΣΧΕΔΙΑΣΜΟΣ ΒΙΒΛΙΟΥ

CONCEPT MANIAX

Περιεχόμενα

Η Ταυτότητα του βιβλίου

Ψηφιακό υλικό Εκπαιδευτικού

Εισαγωγή

Το παρόν σχολικό εγχειρίδιο με τίτλο Στοχαστικά Μαθηματικά απευθύνεται στους μαθητές και τις μαθήτριες της Α΄τάξης των Λυκείων της χώρας μας. Είναι σχεδιασμένο και γραμμένο σύμφωνα με τις οδηγίες του ΙΕΠ και όλα, όσα ορίζονται στο νέο πρόγραμμα σπουδών (ΠΣ), έρχεται δε να υπηρετήσει και να υλοποιήσει τα προσδοκώμενα μαθησιακά αποτελέσματα (ΠΜΑ) που περιγράφονται αναλυτικά σε αυτό.

Τα στοχαστικά μαθηματικά, δηλαδή η σύζευξη στατιστικής και πιθανοτήτων, είναι στην ουσία τα μαθηματικά της ζωής και αυτό προσδίδει στον κλάδο δύο ιδιαιτερότητες. Η πρώτη είναι ότι από τη μια γίνονται από τους μαθητές/τριες σχετικά πιο εύκολα κατανοητά μια και οι αλγεβρικές τεχνικές που απαιτούνται είναι περιορισμένες και από την άλλη διότι άπτονται της καθημερινής τους εμπειρίας και των ενδιαφερόντων τους. Η δεύτερη είναι ότι ως μαθηματικός κλάδος αυτός θεμελιώνεται επαγωγικά και βρίσκει τεράστια εφαρμογή σε πολλούς τομείς της Οικονομίας, των κοινωνικών αλλά, κυρίως, των θετικών επιστημών. Τα Στοχαστικά Μαθηματικά είναι ένα σύγχρονο πολιτισμικό αγαθό, καθώς παρέχουν τα εργαλεία για την ανάλυση δεδομένων και την εξαγωγή πληροφοριών από αυτά. Ο πολίτης κατακλύζεται καθημερινά από στατιστικές πληροφορίες. Στις δημοσκοπήσεις, στις οικονομικές αναλύσεις, στις επιστημονικές ανακοινώσεις, στις κοινωνικές μελέτες χρησιμοποιούνται στατιστικά διαγράμματα και δείκτες ώστε να παρουσιαστούν, να αναλυθούν και να ερμηνευτούν τα δεδομένα που έχουν συλλεχθεί με σκοπό την λήψη αποφάσεων. Υπό τη σκοπιά της Επαγωγικής Στατιστικής, χρησιμοποιώντας ελέγχους υποθέσεων καθώς και στατιστικά μοντέλα, είναι εφικτό να βγουν συμπεράσματα για τον υπό μελέτη πληθυσμό. Στην ύλη που καλύπτουμε στη Α Λυκείου, δεν χρησιμοποιούμε τέτοια επαγωγικά εργαλεία· οι μαθητές/-τριες ενδέχεται να δουν τέτοια εργαλεία σε μελλοντικές τους σπουδές.

Η Θεωρία των Πιθανοτήτων παρέχει το μαθηματικό πλαίσιο μέσα στο οποίο μπορούμε να κατασκευάσουμε στοχαστικά μοντέλα για πολύπλοκα φαινόμενα του πραγματικού κόσμου και να αναλύσουμε αυτά τα μοντέλα για να εξερευνήσουμε τις συνέπειές τους. Η πιθανότητα ενός ενδεχομένου, η πεποίθησή μας για το πόσο βέβαιο θεωρούμε να συμβεί το ενδεχόμενο, αποκτά συγκεκριμένο μαθηματικό νόημα και αναπτύσσεται σε ένα σύνολο από κανόνες λογισμού που συνδέουν τις πιθανότητες ενδεχομένων και μας επιτρέπουν να υπολογίζουμε πιθανότητες σύνθετων ενδεχομένων από εκείνες απλούστερων. Έτσι, η Θεωρία Πιθανοτήτων δεν παρέχει απλά τη μαθηματική θεμελίωση των στατιστικών μεθόδων αλλά έχει εξελιχθεί σε έναν κεντρικό κλάδο των σύγχρονων Μαθηματικών με συνδέσεις και προβολές σε πολλά πεδία των Θεωρητικών και Εφαρμοσμένων Μαθηματικών.

Ορισμένα δομικά χαρακτηριστικά που η συγγραφική ομάδα θεωρεί σημαντικά και αποτελούσαν κατά κάποιον τρόπο πυξίδα σε όλη την πορεία συγγραφής, είναι το βιβλίο σε αντιστοίχιση με το ΠΣ και ως κύριος συνδετικός κρίκος των συντελεστών μάθησης σε μια αίθουσα, πρέπει:

- **Να υποστηρίζει** την ατομική, την κοινωνική και την πολιτισμική προσέγγιση στη μάθηση των μαθηματικών, να τις αντιμετωπίζει ως συμπληρωματικές και σε συνεχή αλληλεπίδραση και να εστιάζει στο βασικό δίπολο μάθηση και μαθητής, μάθηση και διδασκαλία. Επιδιώκει στο να καταστήσει τους/τις μαθητές/-τριες ικανούς/-ες να συνδέουν έννοιες, να κατανοούν το πεδίο εφαρμογής με τους περιορισμούς μιας μαθηματικής έννοιας,κυρίως όμως να αναζητούν και να επινοούν λύσεις σε προβλήματα.

- **Να ενθαρρύνει** και να προτρέπει τους μαθητές/-τριες να να οργανώνουν τις γνώσεις τους με λογικό τρόπο και να μαθαίνουν νέες ιδέες που σχετίζονται με τις υπάρχουσες γνώσεις, έτσι ώστε όταν διατυπώνουν έναν μαθηματικό συλλογισμό, να είναι σε θέση να κατανοήσουν, να αξιολογήσουν και να διακρίνουν τις κεντρικές ιδέες που συγκροτούν την μαθηματική επιχειρηματολογία.
- **Να παρέχει** μέσα από τις δραστηριότητες που περιλαμβάνει τη δυνατότητα στον εκπαιδευτικό να δημιουργεί ένα πλούσιο μαθησιακό περιβάλλον συνεργατικής οικοδόμησης της γνώσης με χρήσιμα επακόλουθα, όπως η αυτενέργεια, η αυτοπεποίθηση και η αυτοεκτίμηση. Να έχει στη διάθεσή του έργα που θα αναθέτει στην τάξη και θα καθορίζουν σε μεγάλο βαθμό το είδος της μαθηματικής δραστηριότητας που θα αναπτύξουν οι μαθητές/-τριες.
- **Να διασφαλίζει** ότι τα περιεχόμενα κάθε κεφαλαίου και κάθε ενότητας είναι συμβατά με τις αρχές και τους στόχους του νέου ΠΣ , όπως αυτοί εξειδικεύονται στα Προσδοκώμενα Μαθησιακά Αποτελέσματα (ΠΜΑ), τα οποία διατυπώνονται με ακρίβεια για να περιγράψουν τι αλλά και σε ποιο επίπεδο καλούνται να γνωρίζουν, να κατανοούν και να είναι σε θέση να κάνουν οι μαθητές/τριες.

Η συγγραφική ομάδα μελέτησε και υιοθέτησε πολλές από τις αρχές που διέπουν το νέο ΠΣ, όπως για παράδειγμα ότι το περιεχόμενο κάθε σχολικού βιβλίου πρέπει:

(i) Να αναπτύσσει το μαθηματικό περιεχόμενο λαμβάνοντας υπόψη: τη βασική αρχή της μάθησης μέσω διερεύνησης, την ανάπτυξη της ικανότητας όλων των μαθητών/-τριών ώστε να παίρνουν αποφάσεις με ατομική, συλλογική σκέψη, επικοινωνία, αλληλεπίδραση και συνεργασία.

(ii) Να μοντελοποιούν πραγματικές καταστάσεις.

(iii) Να εμπλέκουν τους/τις μαθητές/-τριες σε μαθηματικές πρακτικές οι οποίες αναπτύσσουν τον συλλογισμό, την μοντελοποίηση, την επικοινωνία και τον αναστοχασμό και ενδυναμώνουν τη μάθηση των μαθηματικών.

(iv) Να υποστηρίζουν κοινωνικο-πολιτισμικές και κοινωνικο-συναισθηματικές πρακτικές αλλά και διεργασίες όπως η επικοινωνία, η αλληλεπίδραση, ο λόγος , η συμπερίληψη, η ανάπτυξη μαθηματικής ταυτότητας μάθησης, η ανάπτυξη κριτικής επίγνωσης του τρόπου χρήσης των μαθηματικών, η κατανόηση της σχέσης μαθηματικών και πολιτισμού, η μαθηματική εγγραμματοσύνη.

Ένα σημαντικό στοιχείο των νέων βιβλίων είναι ότι στο τέλος κάθε Κεφαλαίου υπάρχει μια διδακτική ενότητα ανακεφαλαίωσης στην οποία αναπτύσσονται έργα, με τα οποία ο/η εκπαιδευτικός ελέγχει την επίτευξη των ΠΜΑ της θεματικής ενότητας και διαμορφώνει συνολική άποψη για την μαθησιακή πορεία των μαθητών και μαθητριών ατομικά και συλλογικά. Η διδακτική ενότητα ανακεφαλαίωσης περιλαμβάνει επιπλέον έργα, τα οποία προσφέρουν τη δυνατότητα περαιτέρω αναζητήσεων και διευρύνσεων στους μαθητές και τις μαθήτριες, τόσο ατομικά όσο και συλλογικά, με στόχο την εμβάθυνση στην κατανόηση του συνόλου των ΠΜΑ της θεματικής ενότητας με συνθετικό τρόπο.

Τέλος, με την πεποίθηση ότι το βιβλίο αυτό ανταποκρίνεται ικανοποιητικά στα κριτήρια ενός σύγχρονου σχολικού εγχειριδίου, το αφιερώνουμε στους μαθητές και τις μαθήτριες αλλά και στους εκπαιδευτικούς, που θα το ζωντανέψουν και θα το αξιοποιήσουν στο μέγιστο βαθμό με τις παιδαγωγικές αρετές τους.

Οι συγγραφείς

Τα Προσδοκώμενα Μαθησιακά Αποτελέσματα (ΠΜΑ)

ΣΤΑΤΙΣΤΙΚΗ

Διαχείριση δεδομένων

1. Διατυπώνουν ερωτήματα που αφορούν σχέσεις εξάρτησης μεταξύ ενός ποσοτικού και ενός κατηγορικού χαρακτηριστικού του πληθυσμού.
2. Κατασκευάζουν πολλαπλά θηκογράμματα, υπολογίζοντας και οριακές τιμές, για να περιγράψουν τις τιμές ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.

Μέτρα θέσης και μεταβλητότητας

1. Περιγράφουν και προσδιορίζουν τη διασπορά και την τυπική απόκλιση ποσοτικών δεδομένων χρησιμοποιώντας τετραγωνικές και απόλυτες αποκλίσεις.
2. Διερευνούν πώς επηρεάζεται η διασπορά και το ενδοτεταρτημοριακό εύρος από την ύπαρξη απόμακρων τιμών.
3. Περιγράφουν και προσδιορίζουν τη μέση τιμή και τη διάμεσο, καθώς και τη διασπορά, την τυπική απόκλιση και το ενδοτεταρτημοριακό εύρος των τιμών ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.
4. Περιγράφουν και προσδιορίζουν το συντελεστή μεταβλητότητας των τιμών ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.
5. Αναγνωρίζουν τη χρησιμότητα του συντελεστή μεταβλητότητας στη σύγκριση μεταβλητότητας ποσοτικών δεδομένων διαφορετικών μονάδων μέτρησης.
6. Επιλέγουν κατάλληλα μέτρα θέσης και μέτρα μεταβλητότητας ποσοτικών δεδομένων ανάλογα με την ύπαρξη απόμακρων τιμών.

Σχέσεις εξάρτησης μεταξύ δύο μεταβλητών

1. Με τη βοήθεια των θηκογραμμάτων κάνουν συγκρίσεις και εξάγουν συμπεράσματα σχετικά με τα μέτρα θέσης και μεταβλητότητας που έχουν οι τιμές του ποσοτικού χαρακτηριστικού σε κάθε στάθμη του κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.
2. Ανακαλύπτουν και εξηγούν με παραδείγματα ότι ένα ποσοτικό και ένα κατηγορικό χαρακτηριστικό δε διέπονται απαραίτητα από μια σχέση αιτίας-αιτιατού. του ποσοτικού χαρακτηριστικού σε κάθε στάθμη του κατηγορικού χαρακτηριστικού, εξάγουν συμπεράσματα και τα παρουσιάζουν στην ολομέλεια του τμήματος.

ΠΙΘΑΝΟΤΗΤΕΣ

Πειράματα τύχης και Πιθανότητες

1. Περιγράφουν πειράματα τύχης και αναγνωρίζουν τη χρησιμότητα των πιθανοθεωρητικών μοντέλων για τη μελέτη πολύπλοκων φαινομένων.
2. Μεταγράφουν ενδεχόμενα και σχέσεις ενδεχομένων που είναι διατυπωμένες σε φυσική γλώσσα, στη γλώσσα των συνόλων και αντίστροφα.
3. Περιγράφουν πειράματα τύχης και με μη ισοπίθανα απλά ενδεχόμενα.
4. Διατυπώνουν υποθέσεις για τους κανόνες που αναμένεται να ισχύουν στο λογισμό των πιθανοτήτων.
5. Διατυπώνουν τον γενικό (αξιωματικό) ορισμό πιθανότητας για έναν πεπερασμένο δειγματικό χώρο.
6. Αναγνωρίζουν διαφορές και συνδέσεις μεταξύ γενικού και κλασικού ορισμού πιθανότητας.
7. Αποδεικνύουν τους κανόνες λογισμού πιθανοτήτων στο πλαίσιο του αξιωματικού ορισμού πιθανότητας.
8. Λύνουν προβλήματα χρησιμοποιώντας τους κανόνες λογισμού πιθανοτήτων.

Οι συμβολισμοί του βιβλίου

Αριθμός κεφαλαίου και ενότητας

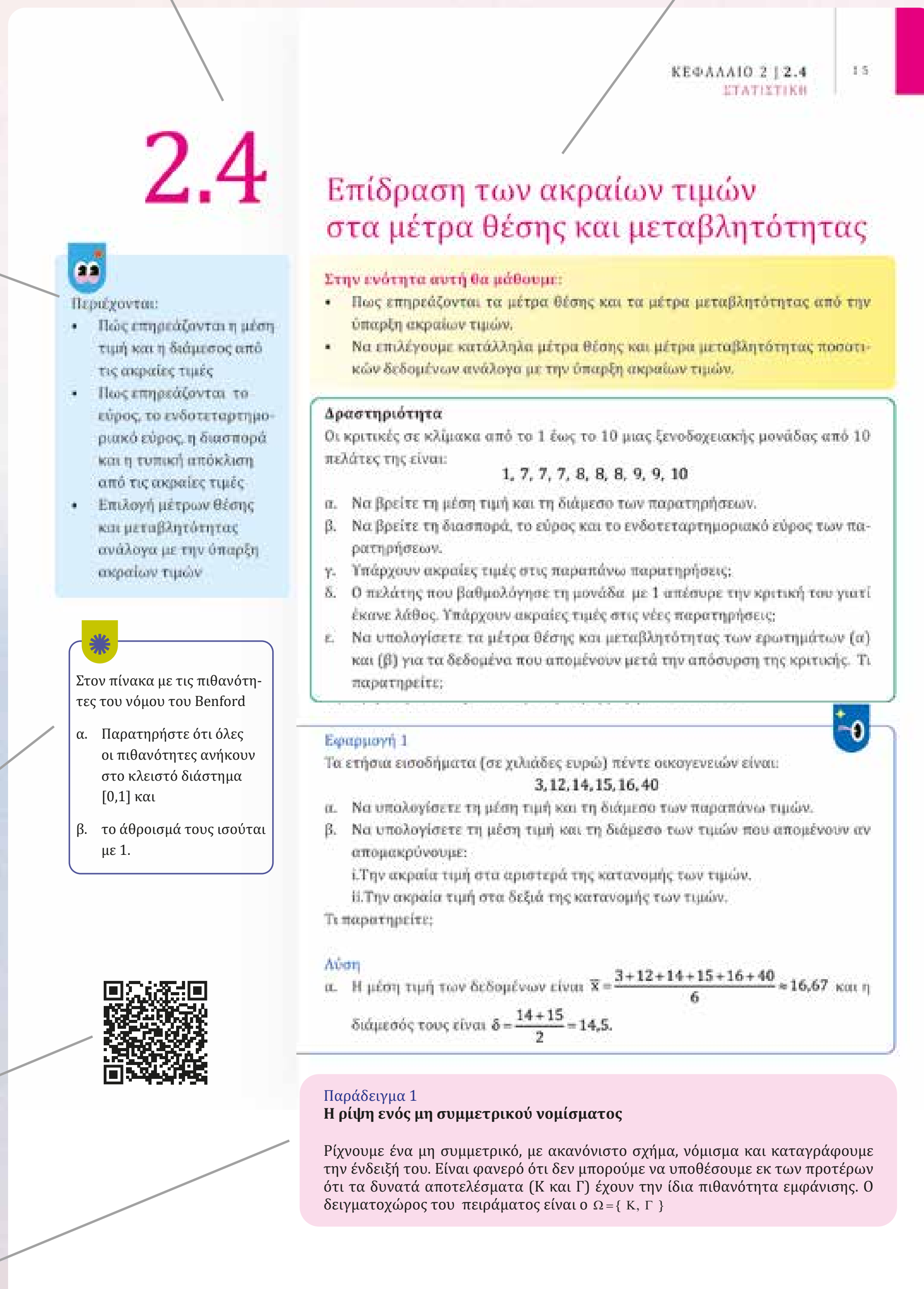
Τίτλος ενότητας

Περιεχόμενα Ενότητας

Παρατήρηση Σχόλιο Υπενθύμιση

Κωδικός QR

Παράδειγμα



ΚΕΦΑΛΑΙΟ 2 | 2.4
ΣΤΑΤΙΣΤΙΚΗ

15

2.4

Περιέχονται:

- Πώς επηρεάζονται η μέση τιμή και η διάμεσος από τις ακραίες τιμές
- Πως επηρεάζονται το εύρος, το ενδοτεταρτημοριακό εύρος, η διασπορά και η τυπική απόκλιση από τις ακραίες τιμές
- Επιλογή μέτρων θέσης και μεταβλητότητας ανάλογα με την ύπαρξη ακραίων τιμών

Στον πίνακα με τις πιθανότητες του νόμου του Benford

α. Παρατηρήστε ότι όλες οι πιθανότητες ανήκουν στο κλειστό διάστημα [0,1] και

β. το άθροισμά τους ισούται με 1.

Επίδραση των ακραίων τιμών στα μέτρα θέσης και μεταβλητότητας

Στην ενότητα αυτή θα μάθουμε:

- Πως επηρεάζονται τα μέτρα θέσης και τα μέτρα μεταβλητότητας από την ύπαρξη ακραίων τιμών.
- Να επιλέγουμε κατάλληλα μέτρα θέσης και μέτρα μεταβλητότητας ποσοτικών δεδομένων ανάλογα με την ύπαρξη ακραίων τιμών.

Δραστηριότητα

Οι κριτικές σε κλίμακα από το 1 έως το 10 μιας ξενοδοχειακής μονάδας από 10 πελάτες της είναι:

1, 7, 7, 7, 8, 8, 8, 9, 9, 10

α. Να βρείτε τη μέση τιμή και τη διάμεσο των παρατηρήσεων.

β. Να βρείτε τη διασπορά, το εύρος και το ενδοτεταρτημοριακό εύρος των παρατηρήσεων.

γ. Υπάρχουν ακραίες τιμές στις παραπάνω παρατηρήσεις;

δ. Ο πελάτης που βαθμολόγησε τη μονάδα με 1 απέσυρε την κριτική του γιατί έκανε λάθος. Υπάρχουν ακραίες τιμές στις νέες παρατηρήσεις;

ε. Να υπολογίσετε τα μέτρα θέσης και μεταβλητότητας των ερωτημάτων (α) και (β) για τα δεδομένα που απομένουν μετά την απόσυρση της κριτικής. Τι παρατηρείτε;

Εφαρμογή 1

Τα ετήσια εισοδήματα (σε χιλιάδες ευρώ) πέντε οικογενειών είναι:

3, 12, 14, 15, 16, 40

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των παραπάνω τιμών.

β. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των τιμών που απομένουν αν απομακρύνουμε:

i. Την ακραία τιμή στα αριστερά της κατανομής των τιμών.

ii. Την ακραία τιμή στα δεξιά της κατανομής των τιμών.

Τι παρατηρείτε;

Λύση

α. Η μέση τιμή των δεδομένων είναι $\bar{x} = \frac{3+12+14+15+16+40}{6} \approx 16{,}67$ και η διάμεσός τους είναι $\delta = \frac{14+15}{2} = 14{,}5$.

Παράδειγμα 1

Η ρίψη ενός μη συμμετρικού νομίσματος

Ρίχνουμε ένα μη συμμετρικό, με ακανόνιστο σχήμα, νόμισμα και καταγράφουμε την ένδειξή του. Είναι φανερό ότι δεν μπορούμε να υποθέσουμε εκ των προτέρων ότι τα δυνατά αποτελέσματα (Κ και Γ) έχουν την ίδια πιθανότητα εμφάνισης. Ο δειγματοχώρος του πειράματος είναι ο Ω={ Κ, Γ }

Ερωτήσεις και ασκήσεις με διαβάθμιση

Ερωτήσεις κατανόησης

Ασκήσεις εμπέδωσης

Ασκήσεις ανάπτυξης

Ασκήσεις δημιουργικής προσέγγισης

Ασκήσεις και Προβλήματα

1 Να χαρακτηρίσετε καθεμία από τις επόμενες προτάσεις, ως αληθή ή ψευδή:

α. $P(A')=1-P(A)$.

β. Για δύο οποιαδήποτε ενδεχόμενα A και B ενός δειγματοχώρου Ω ισχύει: $P(A\cup B)=P(A)+P(B)$.

γ. Αν $P(A\cap B)=0$ τότε $P(A\cup B)=P(A)+P(B)$.

δ. Είναι $P(A)\leq P(A\cup B)$

ε. Είναι $P(B-A)=P(B)-P(A\cap B)$

2 Τα σύνολα A και B είναι ενδεχόμενα ενός δειγματοχώρου Ω, με $P(A)=P(B)=0,6$
Μπορούμε να συμπεράνουμε, ότι τα ενδεχόμενα είναι ίσα;

Να υπολογίσετε τις παρακάτω πιθανότητες

$$P(A),\ P(B) \text{ και } P(A\cap B).$$

8 Αν τα σύνολα και είναι ενδεχόμενα ενός δειγματοχώρου με να αποδείξετε ότι:

$$P(A\cup B)=P(B)$$

9 Έστω τα ενδεχόμενα A και B ενός δειγματοχώρου Ω με

$$P(A')=0,5 \text{ και } P(A\cup B)=0,7$$

Να βρείτε την πιθανότητα να πραγματοποιηθεί το B και να μη πραγματοποιηθεί το A.

Αριθμομηχανή

Ορισμός

Ορισμός (Αξιωματικός ορισμός της πιθανότητας)

Έστω Ω ένας πεπερασμένος δειγματοχώρος. Δεχόμαστε τα επόμενα αξιώματα:

1. Η πιθανότητα ενός ενδεχομένου A του Ω συμβολίζεται $P(A)$ και είναι ένας μη αρνητικός πραγματικός αριθμός, δηλαδή $P(A)\geq 0$.
2. Για τον δειγματοχώρο Ω ισχύει $P(\Omega)=1$.
3. Αν A και B είναι δύο, οποιαδήποτε, ασυμβίβαστα ενδεχόμενα του Ω τότε

$$P(A\cup B)=P(A)+P(B).$$

Σημείωση

Σημείωση 1: Το τρίτο αξίωμα ονομάζεται Απλός Προσθετικός Νόμος των πιθανοτήτων και γενικεύεται για περισσότερα των δύο ενδεχομένων, τα οποία είναι ανά δύο ξένα μεταξύ τους. Για παράδειγμα, αν τα ενδεχόμενα A, B και Γ είναι ανά δύο ξένα, τότε $P(A\cup B\cup \Gamma)=P(A)+P(B)+P(\Gamma)$.

Πρόταση: Έστω $\Omega=\{\omega_1,\omega_2,...,\omega_\nu\}$ ένας πεπερασμένος δειγματοχώρος.

1. Η πιθανότητα του κενού συνόλου είναι μηδέν: $P(\varnothing)=0$.
2. Για οποιοδήποτε ενδεχόμενο $A=\{\alpha_1,\alpha_2,...,\alpha_\kappa\}\neq\varnothing$ του Ω, ισχύει:

$$P(A)=P(\alpha_1)+P(\alpha_2)+...+P(\alpha_\kappa).$$

3. Το άθροισμα των πιθανοτήτων των απλών ενδεχομένων του Ω είναι 1:

$$P(\omega_1)+P(\omega_2)+...+P(\omega_\nu)=1.$$

Πρόταση
Απόδειξη

Απόδειξη

1. Είναι $\varnothing\cap\varnothing=\varnothing$, δηλαδή το κενό σύνολο είναι ασυμβίβαστο με τον εαυτό του. Επίσης $\varnothing\cup\varnothing=\varnothing$. Από το τρίτο αξίωμα έχουμε:

$$P(\varnothing\cup\varnothing)=P(\varnothing) \text{ ή } P(\varnothing)+P(\varnothing)=P(\varnothing) \text{ ή } P(\varnothing)=0$$

Προσπαθήστε να λύσετε τις ασκήσεις 1, 6, 7 και 10

Παραπομπή

ΚΕΦΑΛΑΙΟ 1^{ο}

Διαχείριση δεδομένων

Σύντομη γνωριμία

Κύριο αντικείμενο έρευνας της Στατιστικής επιστήμης είναι η διατύπωση ερευνητικών ερωτημάτων, η συλλογή δεδομένων, η ανάλυση δεδομένων και η ερμηνεία αποτελεσμάτων. Τα ερευνητικά ερωτήματα αποτελούν την αφετηρία για τη διεξαγωγή στατιστικών ερευνών. Τέτοια ερωτήματα στοχεύουν στην κατανόηση της μεταβλητότητας των δεδομένων και απαιτούν τη συλλογή δεδομένων για να απαντηθούν. Η σωστή διαχείριση των δεδομένων αποτελεί μια ουσιαστική διαδικασία για την κατανόηση και την αξιοποίηση των πληροφοριών που κατακλύζουν τον σύγχρονο άνθρωπο. Η αναπαράσταση των δεδομένων μέσω γραφημάτων επιτρέπει την ανάλυση των πληροφοριών και την ανίχνευση προτύπων και σχέσεων. Χωρίς την οπτική αναπαράσταση, η επεξεργασία των δεδομένων θα ήταν περιορισμένη και δύσκολη, καθώς η απλή περιγραφή τους μέσω αριθμητικών τιμών δεν θα μπορούσε να αποκαλύψει πλήρως τις πληροφορίες που αυτά περιέχουν. Οι περιγραφικοί δείκτες καθώς και οι γραφικές αναπαραστάσεις, μας δίνουν μια συνοπτική παρουσίαση των «τιμών» του δείγματος, με τη βοήθεια των οποίων μπορούμε να βγάλουμε κάποια χρήσιμα συμπεράσματα μόνο για τις εν λόγω «τιμές», χωρίς όμως αυτό να σημαίνει πως αναγκαστικά τα ίδια συμπεράσματα διέπουν και τον πληθυσμό.

Περιεχόμενα

John Wilder Tukey
1915 - 2000
Αμερικανός μαθηματικός και στατιστικολόγος.
Συνέβαλε ουσιαστικά στην ανάπτυξη της επιστήμης των δεδομένων καθώς πρωτοστάτησε σε πολλές από τις βασικές αρχές της.
Συνειδητοποίησε νωρίς τη σημασία της επιστήμης των υπολογιστών στην στατιστική ανάλυση δεδομένων. Εισήγαγε για πρώτη φορά το θηκόγραμμα (boxplot) το 1977 στο βιβλίο του ''Exploratory Data Analysis''.
Του αποδίδεται η δημιουργία του όρου bit και η πρώτη δημοσιευμένη χρήση της λέξης λογισμικό.

1.1

Περιέχονται:
- Βασικές έννοιες
- Σχέσεις εξάρτησης μεταξύ μεταβλητών

Λέξεις κλειδιά:
Μεταβλητές
Σχέσεις εξάρτησης

Μεταβλητές

Σχέσεις εξάρτησης μεταξύ ποσοτικού και κατηγορικού χαρακτηριστικού ενός πληθυσμού

Στην ενότητα αυτή θα μάθουμε:
Να διατυπώνουμε ερωτήματα που αφορούν σχέσεις εξάρτησης μεταξύ ενός ποσοτικού και ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.

Βασικές έννοιες

Υπενθυμίζουμε τις επόμενες βασικές έννοιες της Στατιστικής που συναντήσαμε στο Γυμνάσιο.
Πληθυσμός: Είναι το σύνολο το οποίο θέλουμε να μελετήσουμε – εξετάσουμε σε μια στατιστική έρευνα. Τα στοιχεία του πληθυσμού ονομάζονται μονάδες ή άτομα.
Μεταβλητή: Είναι ένα χαρακτηριστικό του υπό μελέτη πληθυσμού το οποίο μπορεί να μεταβάλλεται από μονάδα σε μονάδα του πληθυσμού. Συμβολίζεται με ένα κεφαλαίο γράμμα, όπως X, Y, Z κτλ.
Οι μεταβλητές χωρίζονται σε δύο κατηγορίες:
Ποσοτικές: Είναι οι μεταβλητές που παίρνουν μόνο αριθμητικές τιμές και διακρίνονται περαιτέρω σε:

- συνεχείς, οι οποίες μπορούν να πάρουν τιμές σε ένα διάστημα πραγματικών αριθμών, ακόμα και αν οι τιμές τους εμφανίζονται αρκετά συχνά ως ακέραιοι αριθμοί λόγω αυθαίρετης στρογγυλοποίησης που γίνεται όπως το ύψος ενός ανθρώπου ή το βάρος του κτλ.
- διακριτές, οι οποίες μπορούν να πάρουν διακεκριμένες τιμές, ακέραιες ή δεκαδικές, όπως ο αριθμός των παιδιών μιας οικογένειας, το πλήθος των μαθητών ενός σχολείου κτλ.

Η κρίσιμη διάκριση είναι ότι μεταξύ δύο τιμών μιας διακριτής μεταβλητής δεν υπάρχει καμία άλλη τιμή της, ενώ μια συνεχής μεταβλητή μπορεί να πάρει οποιαδήποτε τιμή μεταξύ δύο τιμών της.
Κατηγορικές (ή ποιοτικές): Είναι οι μεταβλητές οι οποίες δεν είναι ποσοτικές, δηλαδή δεν παίρνουν αριθμητικές τιμές αλλά μπορούν να ταξινομηθούν σε στάθμες (ή κατηγορίες). Διακρίνονται σε:

- ονομαστικές, όταν οι παρατηρήσεις είναι κατηγορίες που η σειρά τους δεν έχει καμία σημασία, για παράδειγμα το χρώμα των μαλλιών (ξανθό, καστανό, κόκκινο, μαύρο, λευκό κτλ), το μέσο μεταφοράς (τραίνο, λεωφορείο κτλ), το επάγγελμα (αγρότης, γιατρός κτλ).
- διατάξιμες, όταν οι παρατηρήσεις είναι κατηγορίες που η σειρά τους έχει σημασία. Παραδείγματα διατάξιμων μεταβλητών είναι: η σοβαρότητα μιας ασθένειας (καμία, ήπια, μέτρια, σοβαρή κτλ), η συμμετοχή σε κάποια δραστηριότητα (ποτέ, σπάνια, συχνά, πολύ συχνά κτλ).

Δείγμα: Είναι ένα μέρος (υποσύνολο) του πληθυσμού το οποίο αντιπροσωπεύει τον πληθυσμό και εξετάζεται ως προς ένα ή περισσότερα χαρακτηριστικά του (μεταβλητές).

Προσπαθήστε να λύσετε τις ασκήσεις 1 και 2

Σχέσεις εξάρτησης μεταξύ μεταβλητών

Ένα σημαντικό κομμάτι της Στατιστικής αφορά την ανάλυση των σχέσεων εξάρτησης μεταξύ δύο μεταβλητών. Τα δυνατά ζεύγη μεταβλητών τα οποία μπορούμε να εξετάσουμε είναι τα εξής:

- ποσοτική - ποσοτική
- κατηγορική - κατηγορική
- ποσοτική - κατηγορική

Το πρώτο μέρος του παρόντος βιβλίου έχει ως αντικείμενο μελέτης τις σχέσεις εξάρτησης μεταξύ μίας ποσοτικής μεταβλητής και μίας κατηγορικής μεταβλητής. Οι άλλες δύο περιπτώσεις θα μελετηθούν στις επόμενες τάξεις.

Λέμε ότι υπάρχει σχέση εξάρτησης ανάμεσα σε δύο μεταβλητές, όταν η μεταβολή της μίας συνοδεύεται από συστηματική και όχι τυχαία μεταβολή της άλλης. Για να μελετήσουμε σχέσεις εξάρτησης, πρέπει να αναγνωρίσουμε τον τύπο των μεταβλητών που έχουμε και να κατανοήσουμε ποια είναι η μεταβλητή με την οποία γίνονται οι συγκρίσεις και ποια είναι η μεταβλητή που εξηγεί τις αλλαγές στην πρώτη. Ειδικότερα, για το ζεύγος «ποσοτική - κατηγορική» θα εξετάσουμε αν οι τιμές της ποσοτικής μεταβλητής (εξαρτημένη μεταβλητή) διαφέρουν μεταξύ των κατηγοριών της κατηγορικής μεταβλητής (ανεξάρτητη μεταβλητή).
Για την καλύτερη κατανόηση των παραπάνω ας δούμε τα δύο επόμενα παραδείγματα.

	Άντρες	Γυναίκες
ΜΙΣΘΟΣ	1050	1020
	1100	1150
	1100	960
	1150	1400
	1120	950
	1600	1050
	1200	1100
	980	950
	1600	1070
	1070	1000

Πίνακας 1: Ύψος μηνιαίου μισθού ανά φύλλο σε δείγμα 20 εργαζομένων της επιχείρησης.

Παράδειγμα 1

Επιλέξαμε τυχαία 10 άντρες και 10 γυναίκες εργαζόμενους σε μια μικρομεσαία επιχείρηση και τους ζητήσαμε να μας αναφέρουν το ύψος του μηνιαίου μισθού τους (σε ευρώ). Οι απαντήσεις που πήραμε φαίνονται στον Πίνακα 1.
Εδώ, η κατηγορική μεταβλητή είναι το φύλο και έχει δύο στάθμες (άντρας-γυναίκα) και η ποσοτική μεταβλητή είναι ο μισθός (ευρώ). Για να εξερευνήσουμε τη σχέση μεταξύ των δύο μεταβλητών, μπορούμε να ξεκινήσουμε με τη διατύπωση ερωτημάτων όπως:

α. Υπάρχει διαφορά ως προς τον μέσο μισθό μεταξύ ανδρών και γυναικών που εργάζονται στην επιχείρηση;
β. Ποιο από τα δύο φύλα παρουσιάζει μεγαλύτερη μεταβλητότητα ως προς το ύψος του μηνιαίου μισθού;
γ. Υπάρχουν πολύ υψηλοί ή πολύ χαμηλοί μισθοί που αξίζει να διερευνηθούν;
δ. Επηρεάζεται η διαφορά στους μισθούς μεταξύ ανδρών και γυναικών από μεταβλητές που δεν συμπεριλαμβάνονται στα στοιχεία του πίνακα, όπως η εκπαίδευση, η εργασιακή εμπειρία, οι ώρες εργασίας κτλ;

Εδώ, η μεταβλητή με την οποία κάνουμε συγκρίσεις είναι ο μισθός ενώ η μεταβλητή που εξηγεί τις διαφορές στις τιμές της πρώτης είναι το φύλο.

Παράδειγμα 2

Ας υποθέσουμε ότι θέλουμε να διερευνήσουμε αν υπάρχει σχέση εξάρτησης της κατηγορικής μεταβλητής «Είδος Διατροφής», η οποία περιέχει τις στάθμες «φυτική, ζωική, μικτή», με την ποσοτική μεταβλητή «Βάρος ανθρώπων σε κιλά». Το βασικό ερώτημα που θα θέλαμε να απαντήσουμε είναι αν το βάρος ενός ατόμου σχετίζεται με την επιλογή συγκεκριμένου είδους διατροφής. Συγκεκριμένα, θα μας ενδιέφερε να εξετάσουμε αν το χαμηλότερο βάρος συνδέεται με κάποιο συγκεκριμένο διατροφικό πρότυπο, όπως η φυτική διατροφή, ή αν η αύξηση του βάρους σχετίζεται περισσότερο με τη ζωική ή μικτή διατροφή. Επίσης, ενδιαφέρον θα είχε να διερευνήσουμε αν υπάρχουν άτομα που, παρά το υψηλό βάρος τους, ακολουθούν φυτική διατροφή ή, αντίθετα, άτομα με χαμηλό βάρος που επιλέγουν κυρίως ζωική διατροφή, και ποιοι παράγοντες εξηγούν αυτές τις εξαιρέσεις.

Στην περίπτωση αυτή, η μεταβλητή με την οποία κάνουμε συγκρίσεις είναι το βάρος ενώ η μεταβλητή που εξηγεί τις διαφορές είναι το είδος διατροφής.

Προσπαθήστε να λύσετε την άσκηση 3

Μια σχέση εξάρτησης μπορεί να φανεί στα δεδομένα ενός δείγματος π.χ. μέσω γραφημάτων. Ωστόσο, ό,τι παρατηρείται στο δείγμα δεν σημαίνει απαραίτητα ότι ισχύει και στον πληθυσμό. Μπορεί η παρατηρούμενη σχέση να είναι τυχαία ή να οφείλεται στο συγκεκριμένο δείγμα που επιλέχθηκε. Στις επόμενες παραγράφους θα αναπτύξουμε νέες τεχνικές και έννοιες για να μπορέσουμε να απαντήσουμε σε πολλά από τα παραπάνω ερωτήματα.

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Εξετάζουμε τα αυτοκίνητα που διέρχονται από το κέντρο μιας πόλης, μια συγκεκριμένη Δευτέρα, ως προς το χρώμα (άσπρο, μαύρο, κίτρινο πράσινο μπλε, άλλο), το φύλο του/της οδηγού (άρρεν ή θήλυ), τον αριθμό των επιβατών και το τελευταίο ψηφίο του αριθμού κυκλοφορίας (μονό ή ζυγό).

α. Ποιος είναι ο πληθυσμός και ποιο το δείγμα της έρευνάς μας;

β. Ποιες είναι οι μεταβλητές της έρευνάς μας; Ποιες από αυτές είναι ποσοτικές; Ποιες είναι κατηγορικές και ποιες είναι οι στάθμες τους;

2 Για τις παρακάτω μεταβλητές, αναγνωρίστε τον τύπο τους και κατηγοριοποιήστε τις ως εξής: Ποσοτική συνεχής, Ποσοτική διακριτή, Κατηγορική ονομαστική, Κατηγορική διατάξιμη.

α. Η ηλικία ενός μαθητή

β. Ο αριθμός των παιδιών σε μια οικογένεια

γ. Το φύλο ενός μαθητή

δ. Η βαθμολογία ενός μαθητή σε μια εξέταση

ε. Η σειρά κατάταξης των μαθητών σε έναν αγώνα δρόμου

στ. Επίπεδο επιτυχίας (Επιτυχία ή Αποτυχία)

ζ. Χρόνος προετοιμασίας για μια γραπτή εξέταση

η. Επίπεδο ευτυχίας (ευχαριστημένος, ικανοποιημένος, αδιάφορος)

θ. Χρώμα αυτοκινήτου (μπλε, κόκκινο, ασημί, μαύρο)

ι. Ο αριθμός των βιβλίων που διαβάζει ένας μαθητής κάθε μήνα

3 Να διατυπώσετε κατάλληλα ερωτήματα για τις πιθανές σχέσεις εξάρτησης των παρακάτω ζευγών μεταβλητών:

α. Κατηγορική μεταβλητή: Κατηγορία οχήματος (Μικρό, Μεσαίο, Μεγάλο).
Ποσοτική μεταβλητή: Κατανάλωση καυσίμου.

β. Κατηγορική μεταβλητή: Περιοχή κατοικίας (Αστική, Ημιαστική, Αγροτική).
Ποσοτική μεταβλητή: Εισόδημα.

γ. Κατηγορική μεταβλητή: Τύπος μαθήματος (Δια ζώσης, Διαδικτυακό).
Ποσοτική μεταβλητή: Βαθμολογία.

δ. Κατηγορική μεταβλητή: Τύπος κινητού τηλεφώνου (μάρκα Α, μάρκα Β).
Ποσοτική μεταβλητή: Αριθμός εγκατεστημένων εφαρμογών.

ε. Κατηγορική μεταβλητή: Τύπος του αθλήματος (ποδόσφαιρο, μπάσκετ, κολύμβηση).
Ποσοτική μεταβλητή: Αριθμός προπονήσεων ανά εβδομάδα.

1.2 Πολλαπλά Θηκογράμματα

Περιέχονται:
- Το Θηκόγραμμα με οριακές τιμές
- Κατασκευή του θηκογράμματος με οριακές και ακραίες τιμές
- Η κατανομή των παρατηρήσεων στο θηκόγραμμα
- Περιγραφή των τιμών ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού με τη βοήθεια θηκογραμμάτων
- Ανακαλύπτοντας τη Συμμετρία και την Ασυμμετρία της κατανομής των παρατηρήσεων μέσω θηκογραμμάτων

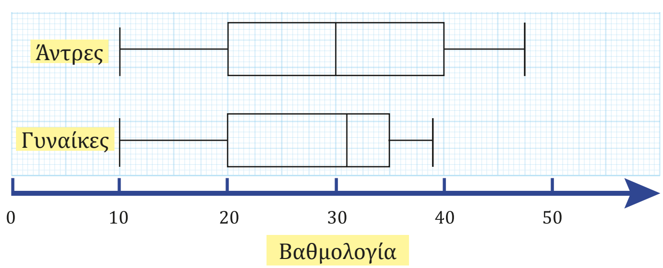
Λέξεις κλειδιά:
Θηκόγραμμα
Οριακές τιμές
Ακραίες τιμές

Περίληψη 5 αριθμών

Στην ενότητα αυτή θα μάθουμε:
- Να κατασκευάζουμε πολλαπλά θηκογράμματα, υπολογίζοντας και οριακές τιμές.
- Να χρησιμοποιούμε τα πολλαπλά θηκογράμματα για να περιγράψουμε τις τιμές ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.

Δραστηριότητα

Στα θηκογράμματα του επόμενου σχήματος, διακρίνονται οι βαθμοί που έλαβαν 30 άντρες και 30 γυναίκες σε ένα τεστ αξιολόγησης για μία συγκεκριμένη θέση εργασίας.



Προσπαθήστε να απαντήσετε στις επόμενες ερωτήσεις:

α. Ποια είναι η διάμεσος, το εύρος και το ενδοτεταρτημοριακό εύρος των βαθμών των αντρών;

β. Ποια είναι η διάμεσος, το εύρος και το ενδοτεταρτημοριακό εύρος των βαθμών των γυναικών;

γ. Σε τι εξυπηρετεί η ταυτόχρονη αναπαράσταση δύο θηκογραμμάτων;

Το Θηκόγραμμα με οριακές τιμές

Στο Γυμνάσιο μάθαμε να κατασκευάζουμε θηκογράμματα χρησιμοποιώντας την «περίληψη των πέντε αριθμών», δηλαδή την ελάχιστη τιμή (min), το πρώτο τεταρτημόριο (Q_1), τη διάμεσο (δ) ή (Q_2), το τρίτο τεταρτημόριο (Q_3) και την μέγιστη τιμή (max). Η διαφορά $Q_3 - Q_1$ ονομάζεται ενδοτεταρτημοριακό εύρος (interquartile range) και θα τη συμβολίζουμε με Q, δηλαδή

$$Q = Q_3 - Q_1$$

Παράδειγμα 1

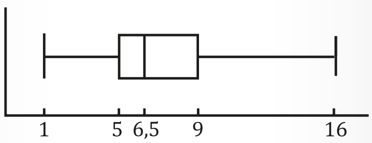
Δίνονται τα δεδομένα: 2, 5, 9, 15, 7, 1, 6, 16, 9, 5.

Αρχικά τα διατάσσουμε σε αύξουσα σειρά:

1, 2, 5, 5, 6, 7, 9, 9, 15, 16

Έχουμε ελάχιστη τιμή το 1 και μέγιστη τιμή το 16.

Επειδή το πλήθος των παρατηρήσεων είναι 10, άρτιο, η διάμεσος είναι η μέση τιμή της 5ης και 6ης παρατήρησης. Επομένως $\delta = \frac{6+7}{2} = 6{,}5$.



Το θηκόγραμμα, όπως προκύπτει, από την "περίληψη των πέντε αριθμών", χωρίς οριακές και ακραίες τιμές.

Το πρώτο τεταρτημόριο είναι η διάμεσος των πρώτων 5 παρατηρήσεων, οι οποίες βρίσκονται αριστερά της διαμέσου. Επομένως $Q_1 = 5$. Το τρίτο τεταρτημόριο είναι η διάμεσος των τελευταίων 5 παρατηρήσεων, οι οποίες βρίσκονται δεξιά της διαμέσου. Επομένως $Q_3 = 9$.

Στη συνέχεια, για ένα σύνολο παρατηρήσεων, εισάγουμε τις παρακάτω έννοιες:

- Κάτω φράγμα, ονομάζουμε τον αριθμό $Q_1 - 1{,}5Q$.
- Άνω φράγμα, ονομάζουμε τον αριθμό $Q_3 + 1{,}5Q$.
- Κάτω οριακή τιμή (L), ονομάζουμε την μικρότερη παρατήρηση που ανήκει στο διάστημα $[Q_1 - 1{,}5Q,\ Q_3 + 1{,}5Q]$.

Με άλλα λόγια, κάτω οριακή τιμή είναι η μικρότερη παρατήρηση που είναι μεγαλύτερη ή ίση από το κάτω φράγμα.

- Άνω οριακή τιμή (U), ονομάζουμε την μεγαλύτερη παρατήρηση που ανήκει στο διάστημα $[Q_1 - 1{,}5Q,\ Q_3 + 1{,}5Q]$.

Με άλλα λόγια, άνω οριακή τιμή είναι η μεγαλύτερη παρατήρηση που είναι μικρότερη ή ίση από το άνω φράγμα.

- Ακραία ή απόμακρη τιμή (outlier), ονομάζουμε κάθε παρατήρηση που βρίσκεται εκτός του διαστήματος $[Q_1 - 1{,}5Q,\ Q_3 + 1{,}5Q]$.

Επομένως, οι ακραίες τιμές, όταν υπάρχουν, είναι οι παρατηρήσεις που είναι μικρότερες από το κάτω φράγμα (άρα και από την κάτω οριακή τιμή) ή μεγαλύτερες από το άνω φράγμα (άρα και από την άνω οριακή τιμή).

Είναι φανερό ότι ισχύει:

$[L,\ U] \subseteq [Q_1 - 1{,}5Q,\ Q_3 + 1{,}5Q]$

Παράδειγμα 2

Θα υπολογίσουμε τις οριακές τιμές καθώς και τις ακραίες τιμές των δεδομένων του προηγούμενου παραδείγματος. Τα διατεταγμένα δεδομένα είναι:

$$1,\ 2,\ 5,\ 5,\ 6,\ 7,\ 9,\ 9,\ 15,\ 16$$

Βρήκαμε $Q_1 = 5$ και $Q_3 = 9$. Το ενδοτεταρτημοριακό εύρος είναι

$$Q = Q_3 - Q_1 = 9 - 5 = 4$$

Το κάτω φράγμα είναι το $Q_1 - 1{,}5Q = 5 - 1{,}5 \cdot 4 = -1$ και το άνω φράγμα είναι το $Q_3 + 1{,}5Q = 9 + 1{,}5 \cdot 4 = 15$.

Είναι $[Q_1 - 1{,}5Q,\ Q_3 + 1{,}5Q] = [-1, 15]$.

Η μικρότερη παρατήρηση που ανήκει στο διάστημα αυτό είναι η κάτω οριακή τιμή L. Επομένως, $L = 1$. Η μεγαλύτερη παρατήρηση που ανήκει στο διάστημα αυτό είναι η άνω οριακή τιμή U. Επομένως, $U = 15$.

Ακραίες τιμές, είναι οι παρατηρήσεις που βρίσκονται εκτός του διαστήματος $[-1, 15]$. Επομένως, η παρατήρηση 16 είναι η μόνη ακραία τιμή.

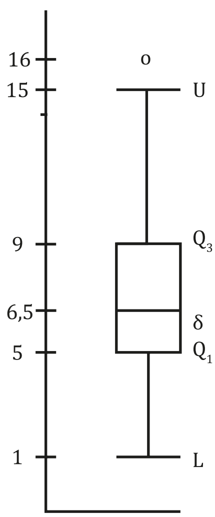
Μέθοδοι υπολογισμού των Q_1 και Q_3

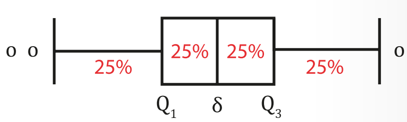
Χαρακτηριστικοί αριθμοί θηκογράμματος

Προσπαθήστε να λύσετε τις ασκήσεις 2, 4, 5 και 6

Κατασκευή του θηκογράμματος με οριακές και ακραίες τιμές

Στο θηκόγραμμα σημειώνουμε τους αριθμούς L, Q_1, δ, Q_3 και U, καθώς και τις ακραίες τιμές (εφόσον υπάρχουν). Στο επόμενο παράδειγμα περιγράφεται η διαδικασία κατασκευής ενός τέτοιου θηκογράμματος.





Παράδειγμα 3

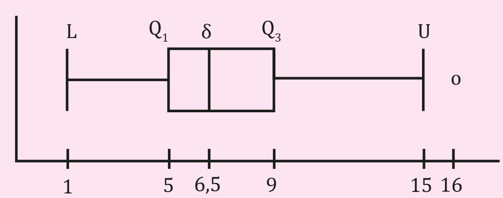
Θα κατασκευάσουμε το θηκόγραμμα των δεδομένων των προηγούμενων παραδειγμάτων:

1, 2, 5, 5, 6, 7, 9, 9, 15, 16

Βρήκαμε: Q_1=5, Q_3=9, L=1, U=15 και ακραία τιμή το 16.

- Στον οριζόντιο άξονα τοποθετούμε τις τιμές 1, 5, 6,5, 9, 15, 16.
- Στη συνέχεια σχεδιάζουμε το ορθογώνιο του θηκογράμματος. Η αριστερή του πλευρά βρίσκεται στο $Q_1=5$ και η δεξιά του στο $Q_3=9$. Το ύψος του είναι αυθαίρετο.
- Η γραμμή στο εσωτερικό του ορθογωνίου δείχνει τη διάμεσο δ=6,5.
- Δεξιά του ορθογωνίου σχεδιάζουμε μια γραμμή της οποίας η κατακόρυφη γραμμή δείχνει την άνω οριακή τιμή 15. Αριστερά του ορθογωνίου σχεδιάζουμε μια γραμμή της οποίας η κατακόρυφη γραμμή δείχνει την κάτω οριακή τιμή 1. Οι γραμμές αυτές ονομάζονται απολήξεις.
- Τέλος, σημειώνουμε την ακραία τιμή 16.

Στο επόμενο σχήμα βλέπουμε το θηκόγραμμα σε οριζόντια σχεδίαση. Στο διπλανό σχήμα είναι το ίδιο θηκόγραμμα σε κατακόρυφη σχεδίαση.



Η κατανομή των παρατηρήσεων στο θηκόγραμμα

Οι αριθμοί Q_1, δ και Q_3 χωρίζουν κάθε θηκόγραμμα σε τέσσερα μέρη, όπως φαίνεται στο διπλανό σχήμα. Σε κάθε μέρος ανήκει, περίπου, το 25% των τιμών των δεδομένων. Ειδικότερα, στο διάστημα $[Q_1, Q_3]$ ανήκουν, περίπου, οι μισές τιμές, δηλαδή το 50%.

Προσπαθήστε να λύσετε τις ασκήσεις 1 και 3

Περιγραφή των τιμών ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού με τη βοήθεια θηκογραμμάτων

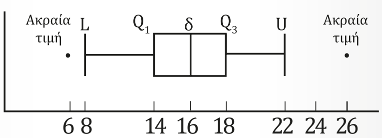
Τα πολλαπλά θηκογράμματα είναι μια χρήσιμη μέθοδος για να περιγράψουμε τις τιμές ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού και να συγκρίνουμε τις κατανομές τους. Η κατασκευή είναι απλή· σχεδιάζουμε τα θηκογράμματα για τις στάθμες, που μας ενδιαφέρουν, στο ίδιο σύστημα. Η διαδικασία αυτή φαίνεται στην επόμενη εφαρμογή.

Εφαρμογή

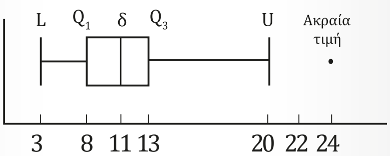
Για να μελετήσουμε την παρουσία στον αέρα μιας ουσίας σε mg/m^3 κάναμε μετρήσεις σε δύο πόλεις Α και Β (Πίνακας 1). Να κατασκευαστούν στο ίδιο σύστημα τα θηκογράμματα για τα δεδομένα των δύο πόλεων.

Πόλη Α	6 8 14 14 15 15 16 16 17 17 17 18 19 22 26
Πόλη Β	3 5 5 8 9 10 10 11 12 12 13 13 15 20 24

Πίνακας 1: Μετρήσεις μιας ουσίας σε mg/m^3 στις πόλεις Α και Β



Σχήμα 1:Θηκόγραμμα για τα δεδομένα της πόλης Α



Σχήμα 2: Θηκόγραμμα για τα δεδομένα της πόλης Β

Πολλαπλά θηκογράμματα

Θηκόγραμμα οδηγός

Λύση

Κατασκευή θηκογράμματος για τα δεδομένα της πόλης Α

Τα δεδομένα για την πόλη Α είναι διατεταγμένα σε αύξουσα σειρά:

6 8 14 14 15 15 16 16 17 17 17 18 19 22 26

Είναι δ=16, $Q_1 = 14$ και $Q_3 = 18$.

Υπολογισμός οριακών και ακραίων τιμών

- Το ενδοτεταρτημοριακό εύρος είναι $Q = Q_3 - Q_1 = 18 - 14 = 4$.
- Το κάτω φράγμα είναι $Q_1 - 1{,}5(Q_3 - Q_1) = 14 - 1{,}5 \cdot 4 = 8$.
- Το άνω φράγμα είναι $Q_3 + 1{,}5(Q_3 - Q_1) = 18 + 1{,}5 \cdot 4 = 24$.
- Είναι $[Q_1 - 1{,}5Q,\ Q_3 + 1{,}5Q] = [8, 24]$. Επομένως, $L = 8$ και $U = 22$.
- Έχουμε δύο ακραίες τιμές, το 6 και το 26.
- Το θηκόγραμμα φαίνεται στο διπλανό σχήμα 1.

Κατασκευή θηκογράμματος για τα δεδομένα της πόλης Β

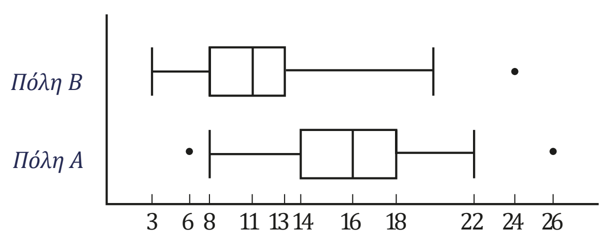
3 5 5 8 9 10 10 11 12 12 13 13 15 20 24

Εργαζόμενοι ομοίως βρίσκουμε:

- δ=11, $Q_1 = 8$, $Q_3 = 13$ και $Q = 5$.
- Το κάτω φράγμα είναι το 0,5 και άνω φράγμα το 20,5.
- $L = 3$ και $U = 20$.
- Ακραία τιμή είναι το 24.

Το θηκόγραμμα φαίνεται στο διπλανό σχήμα 2.

Προκύπτει τελικά το επόμενο διάγραμμα:



Σχήμα 3: Θηκογράμματα για τα δεδομένα των πόλεων Α και Β

Η διάμεσος των δεδομένων της πόλης Α είναι 16 mg/m^3, που σημαίνει ότι περίπου το 50% των παρατηρήσεων είναι μικρότερο ή ίσο από την τιμή 16 και περίπου το 50% είναι μεγαλύτερο ή ίσο από την τιμή 16. Το ενδοτεταρτημοριακό εύρος, δηλαδή η απόσταση μεταξύ του πρώτου τεταρτημορίου (14) και του τρίτου τεταρτημορίου (18), είναι 4mg/m^3. Αυτό σημαίνει ότι περίπου το 50% των παρατηρήσεων συγκεντρώνονται στο διάστημα [14, 18] που έχει μήκος 4. Υπάρχουν επίσης δύο ακραίες τιμές.

Παρόμοιες παρατηρήσεις μπορούμε να κάνουμε για τα δεδομένα της πόλης Β παρατηρώντας το δεύτερο θηκόγραμμα.

Ανακαλύπτοντας τη Συμμετρία και την Ασυμμετρία της κατανομής των παρατηρήσεων μέσω θηκογραμμάτων

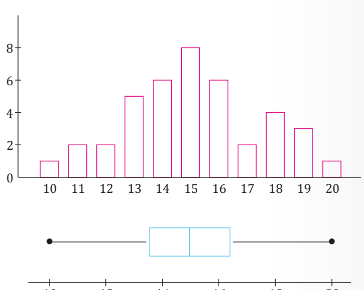
Τα θηκογράμματα είναι χρήσιμα εργαλεία για την οπτική παρουσίαση της συμμετρίας ή της ασυμμετρίας μιας κατανομής παρατηρήσεων.

Μια συμμετρική κατανομή αποτυπώνεται σε θηκόγραμμα με τη διάμεσο να βρίσκεται στο κέντρο του ορθογωνίου, και τις απολήξεις να εκτείνονται σε ίσες περίπου αποστάσεις προς τις δύο κατευθύνσεις (Σχήμα 1).

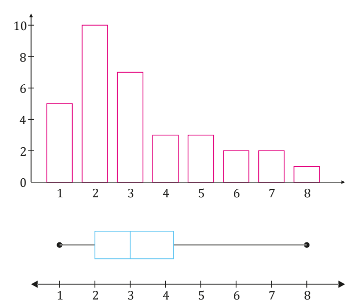
Αντίθετα, η ασυμμετρία γίνεται εμφανής όταν η διάμεσος μετατοπίζεται προς τη μία πλευρά του ορθογωνίου, ή όταν μία από τις απολήξεις είναι αισθητά μεγαλύτερη της άλλης.

Όταν οι παρατηρήσεις συγκεντρώνονται κοντά στην ελάχιστη τιμή και έχουμε λίγες μεγάλες τιμές, λέμε ότι η κατανομή των παρατηρήσεων παρουσιάζει θετική ασυμμετρία (Σχήμα 2).

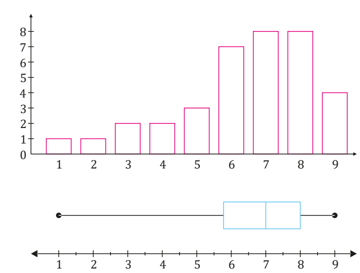
Όταν οι παρατηρήσεις συγκεντρώνονται κοντά στη μέγιστη τιμή και έχουμε λίγες μικρές τιμές, λέμε ότι η κατανομή των παρατηρήσεων παρουσιάζει αρνητική ασυμμετρία (Σχήμα 3).



Σχ. 1



Σχ. 2



Σχ. 3

Προσπαθήστε να λύσετε τις ασκήσεις 7, 8 και 9

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Να χαρακτηρίσετε καθεμία από τις επόμενες προτάσεις, ως αληθή ή ψευδή:

α. Οι ακραίες τιμές, όταν υπάρχουν, είναι οι τιμές που βρίσκονται εντός του διαστήματος
$$[Q_1-1,5Q,\ \ Q_3+1,5Q]$$

β. Η άνω οριακή τιμή είναι μικρότερη ή ίση του
$$Q_3+1,5Q.$$

γ. Το ορθογώνιο του οριζόντιου θηκογράμματος έχει αυθαίρετο ύψος.

δ. Η διάμεσος είναι μεγαλύτερη ή ίση από την κάτω οριακή τιμή.

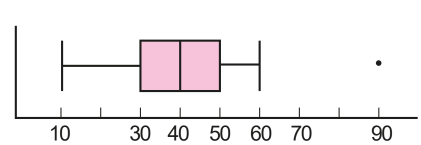
ε. Περίπου 75% των παρατηρήσεων βρίσκονται αριστερά από το Q_3.

2 Σε ένα σύνολο παρατηρήσεων υπολογίσαμε το διάστημα
$$[Q_1-1,5Q,\ \ Q_3+1,5Q]=[10,15].$$

α. Ποια είναι η κάτω οριακή τιμή, αν η μικρότερη παρατήρηση είναι το 10,2;

β. Ποια είναι η άνω οριακή τιμή αν οι τρεις μεγαλύτερες παρατηρήσεις είναι οι αριθμοί
$$14{,}8, \quad 14{,}9, \quad \text{και} \quad 15{,}2;$$

3 Δίνεται το παρακάτω θηκόγραμμα.



Να απαντήσετε στις επόμενες ερωτήσεις:

α. Ποια είναι τα τεταρτημόρια Q_1 και Q_3;

β. Πώς λέγονται οι τιμές 10 και 60;

γ. Είναι η διάμεσος δ=40;

δ. Υπάρχει ακραία τιμή και ποια είναι αυτή;

ε. Είναι το 10 η μικρότερη τιμή των παρατηρήσεων;

Ασκήσεις

4 Δίνονται οι παρατηρήσεις:

2, 2, 3, 5, 7, 7, 8, 9, 9, 12

α. Να βρείτε τους αριθμούς Q_1, Q_3 και $Q = Q_3 - Q_1$.

β. Να βρείτε τους αριθμούς $Q_1 - 1{,}5Q$ και $Q_3 + 1{,}5Q$.

γ. Να βρείτε τις οριακές τιμές των παρατηρήσεων.

δ. Να βρείτε τις ακραίες τιμές των παρατηρήσεων (αν υπάρχουν).

5 Σε ένα σύνολο παρατηρήσεων που περιέχει τους αριθμούς -8 και 32 βρήκαμε:

$$Q_1 = 5 \text{ και } Q_3 = 15$$

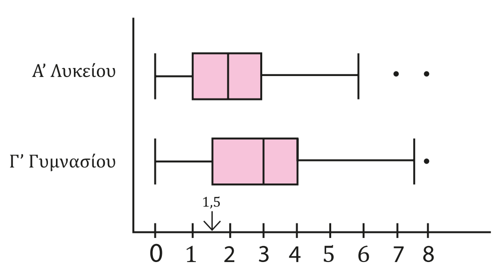
Να εξετάσετε αν οι τιμές -8 και 32 είναι ακραίες.

6 Μια ομάδα παρατηρήσεων έχει $Q_1 = 20$ και $Q = 2$.

α. Να βρείτε το Q_3.

β. Αν η τιμή 17 περιέχεται στο σύνολο των παρατηρήσεων, να εξετάσετε αν είναι οριακή ή ακραία τιμή.

7 Τα παρακάτω θηκογράμματα δίνουν τα αποτελέσματα μετρήσεων δύο ομάδων μαθητών, της Γ' Γυμνασίου και της Α' Λυκείου, που αφορούν το χρόνο (σε ώρες) καθημερινής μελέτης των μαθημάτων του σχολείου.



Να συμπληρώσετε τον παρακάτω πίνακα:

	Q_1	δ	Q_3	Q	L	U	Ακραίες Τιμές
Γ Γυμν.							
Α Λυκ.							

8 Στον επόμενο πίνακα παρουσιάζονται οι βαθμοί 11 μαθητών της Α' Λυκείου και 15 μαθητών της Β' Λυκείου στο μάθημα των Μαθηματικών.

Α Λυκείου	14 15 15 16 16 17 17 18 18 19 20
Β Λυκείου	12 15 15 16 16 17 17 17 18 18 18 18 19 19 20

α. Να εξετάσετε αν υπάρχουν ακραίες τιμές στις δύο ομάδες.

β. Να κατασκευάσετε στο ίδιο σύστημα δύο θηκογράμματα για τα δεδομένα του πίνακα.

9 Ο παρακάτω πίνακας αποτελεί την περίληψη των πέντε αριθμών για τη διάρκεια (σε λεπτά) των τηλεφωνικών κλήσεων του προηγούμενου μήνα δύο ομάδων παιδιών Α και Β.

	min	Q_1	δ	Q_3	max
Ομάδα Α	1	3	4	8	13
Ομάδα Β	2	4	5	10	18

α. Υπάρχουν ακραίες τιμές στα δεδομένα;

β. Να κατασκευάσετε στο ίδιο σύστημα δύο θηκογράμματα για τα παραπάνω δεδομένα και να τα σχολιάσετε ως προς το σχήμα τους (συμμετρία-ασυμμετρία).

1.3 Ανακεφαλαίωση

Στην ενότητα αυτή:
Περιλαμβάνονται έργα για περαιτέρω αναζητήσεις και διερευνήσεις των μαθητών/μαθητριών, με στόχο τη διεύρυνση της μαθησιακής διαδικασίας που οδηγεί στα Προσδοκώμενα Μαθησιακά Αποτελέσματα του κεφαλαίου.

Συλλογή δεδομένων
Συλλέγοντας δεδομένα μπορείτε να αποκτήσετε ενδιαφέρουσες και χρήσιμες πληροφορίες για διάφορα θέματα. Η απλούστερη συλλογή δεδομένων μπορεί να αφορά πληροφορίες για τους συμμαθητές σας. Δημιουργήστε ένα ή περισσότερα ερωτηματολόγια και ζητήστε από τους συμμαθητές σας να τα συμπληρώσουν. Η δημιουργία ερωτηματολογίου και γενικότερα η συλλογή δεδομένων είναι διεργασίες που απαιτούν προσεκτικό σχεδιασμό ώστε να διασφαλιστεί η ποιότητα και η αξιοπιστία των δεδομένων. Κατά την κατασκευή ενός ερωτηματολογίου, είναι κρίσιμο να διατυπώνονται σαφείς και αμερόληπτες ερωτήσεις, αποφεύγοντας τη χρήση πολύπλοκων όρων ή φράσεων που μπορεί να δημιουργήσουν σύγχυση. Αυτή η διαδικασία απαιτεί καλή κατανόηση των κατηγορικών και ποσοτικών μεταβλητών, καθώς και την ικανότητα δημιουργίας σαφώς διατυπωμένων ερωτημάτων. Ενδεικτικά, κάποιες προτάσεις για τη συλλογή δεδομένων από κατηγορικές και ποσοτικές μεταβλητές αναφέρονται παρακάτω.

Κατηγορικές μεταβλητές:
Συλλέξτε δεδομένα από τους συμμαθητές σας σχετικά με το φύλο τους, το χρώμα μαλλιών τους, το χρώμα ματιών τους αλλά και με τις προτιμήσεις τους σε διάφορα θέματα όπως το αγαπημένο τους σχολικό μάθημα, το άθλημα που τους αρέσει πιο πολύ, τις μουσικές τους προτιμήσεις, την αγαπημένη τους εποχή του χρόνου κ.α.

Ποσοτικές μεταβλητές:
Συλλέξτε δεδομένα από τους συμμαθητές σας σχετικά με το ύψος τους, το βάρος τους, πόσα αδέρφια έχουν, τον χρόνο που αφιερώνουν καθημερινά σε ψυχαγωγικές δραστηριότητες, τον χρόνο που αφιερώνουν καθημερινά στο διαδίκτυο, τον χρόνο καθημερινής μελέτης των σχολικών τους μαθημάτων, τις ώρες που αφιερώνουν εβδομαδιαία στον αθλητισμό, τις ώρες βραδινού ύπνου κ.α.

Ερευνητικός σχεδιασμός: Σε αυτό το στάδιο (προηγείται της συλλογής δεδομένων), θα πρέπει να ορίσετε σαφώς το ερευνητικό ερώτημα που επιθυμείτε να απαντήσετε, συνδυάζοντας μια κατηγορική με μια ποσοτική μεταβλητή. Για παράδειγμα, μπορείτε να εξετάσετε αν, κατά μέσο όρο, ο χρόνος καθημερινής μελέτης των σχολικών μαθημάτων (ποσοτική μεταβλητή) διαφοροποιείται στις στάθμες της κατηγορικής μεταβλητής ''Αγαπημένο σχολικό μάθημα''.

Ερευνητικές προτάσεις
Πρόταση 1: **Ενεργειακά ποτά και βραδινός ύπνος**
Συλλέξτε δεδομένα από τους συμμαθητές σας που να αφορούν στην λήψη ενερ-

γειακών ποτών μηνιαία βάση (κατηγορική μεταβλητή: Καθόλου, Σπάνια, Συχνά, Καθημερινά) και τις ώρες που κοιμούνται το βράδυ (ποσοτική μεταβλητή). Η συγκέντρωση αυτών των δεδομένων θα σας επιτρέψει να εντοπίσετε πιθανές σχέσεις εξάρτησης μεταξύ της κατανάλωσης ενεργειακών ποτών και ωρών ύπνου.

Πρόταση 2: **Πρωινό και επίδοση στο σχολείο**
Συλλέξτε δεδομένα από τους συμμαθητές σας που να αφορούν στην κατανάλωση πρωινού σε μηνιαία βάση (κατηγορική μεταβλητή: Καθόλου, Σπάνια, Συχνά, Καθημερινά) και τις επιδόσεις τους στο σχολείο (ποσοτική μεταβλητή: π.χ. βαθμός προαγωγής στη Γ' Γυμνασίου). Η συγκέντρωση αυτών των δεδομένων θα σας επιτρέψει να εντοπίσετε πιθανές σχέσεις εξάρτησης μεταξύ της πρωινής διατροφής και των σχολικών επιδόσεων.

Πρόταση 3: **Αίσθηση του χρόνου**
Χωριστείτε σε δύο ή περισσότερες ομάδες με βάση κάποιο ποιοτικό κριτήριο (π.χ. αγόρια και κορίτσια) και σχεδιάστε μια έρευνα για να ανακαλύψετε πόσο επιτυχημένη είναι η κάθε ομάδα στο να προβλέπει τη διάρκεια του ενός λεπτού. Η ακρίβεια στο να προσδιορίζουμε τον χρόνο που έχει περάσει μπορεί να είναι σημαντική (π.χ. άμεση δράση σε έκτακτες καταστάσεις όπως ιατρικά επείγοντα περιστατικά ή ατυχήματα, προγραμματισμός καθημερινών δραστηριοτήτων κτλ). Στο τέλος του μαθήματος θα είστε σε θέση να αξιολογείστε τα αποτελέσματα που θα προκύψουν από αυτή την έρευνα.

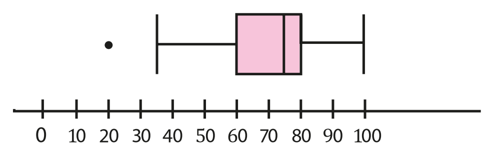
Μελέτη θηκογράμματος -Πληροφορίες που αντλούμε από ένα θηκόγραμμα

1 Να απαντήσετε στις επόμενες ερωτήσεις:
α. Ποιες μεταβλητές λέγονται ποσοτικές και ποιες κατηγορικές (ποιοτικές);
β. Σε ποιες κατηγορίες χωρίζονται οι ποσοτικές μεταβλητές;
γ. Έστω ένα σύνολο ποσοτικών δεδομένων. Πώς βρίσκουμε:
i. το άνω φράγμα και την άνω οριακή τιμή τους;
ii. το κάτω φράγμα και την κάτω οριακή τιμή τους;
iii. τις ακραίες ή αλλιώς τις απόμακρες τιμές, όταν υπάρχουν;

Μελέτη θηκογράμματος

2 Να χαρακτηρίσετε, ως αληθή ή ψευδή καθεμία από τις επόμενες προτάσεις:
α. Όταν εξετάζουμε έναν ανθρώπινο πληθυσμό, ως προς το φύλο των ατόμων του, η μεταβλητή είναι το «φύλο» και οι στάθμες της είναι «άρρεν» και «θήλυ».
β. Το άνω φράγμα ενός θηκογράμματος είναι το $Q_3 + 1{,}5(Q_3 - Q_1)$.
γ. Η άνω οριακή τιμή είναι μεγαλύτερη από το άνω φράγμα.
δ. Οι ακραίες τιμές βρίσκονται εκτός του διαστήματος $[Q_1 - 1{,}5Q, Q_3 + 1{,}5Q]$.
ε. Στο διάστημα $[Q_1, Q_3]$ βρίσκονται περίπου το 25% των παρατηρήσεων.

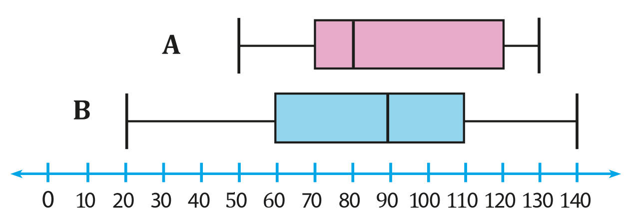
3 Το παρακάτω θηκόγραμμα δείχνει τις βαθμολογίες, στην κλίμακα 0-100, των μαθητών/τριών ενός τμήματος Α΄ Λυκείου σε ένα διαγώνισμα Αγγλικών.



α. Να βρείτε τις τιμές των τεταρτημόριων Q_1, Q_3 και της διαμέσου δ.
β. Να βρείτε το άνω και το κάτω φράγμα.
γ. Να βρείτε την άνω και την κάτω οριακή τιμή.
δ. Να εξετάσετε αν είναι αληθείς οι επόμενοι ισχυρισμοί:
i.Περίπου το 50% των βαθμών είναι μικρότεροι του 75.
ii.Στο διάστημα [60, 80] ανήκει περίπου το ένα τέταρτο των βαθμών.
iii.Περίπου το 25% των βαθμών είναι μεγαλύτεροι του 80.
iv.Το θηκόγραμμα παρουσιάζει αρνητική ασυμμετρία.

Μελέτη θηκογράμματος

4 Η ποσότητα R σε mg ανά κυβικό μέτρο αέρα μιας ρυπογόνου ουσίας σε 50 σημεία της πόλης Α και σε 50 σημεία της πόλης Β φαίνεται στα παρακάτω θηκογράμματα.



α. Να συμπληρώσετε τον πίνακα.

	min	Q_1	δ	Q_3	max	Q
Πόλη Α						
Πόλη Β						

β. Ποιο είναι το εύρος των ρύπων σε κάθε πόλη;
γ. Σε πόσα τουλάχιστον σημεία της πόλης (Β) οι ρύποι είναι μεταξύ 60 και $110 mg/m^3$;
δ. Σε πόσα το πολύ σημεία της πόλης (Α) οι ρύποι είναι μικρότεροι από $80 mg/m^3$;

5 Δύο σύνολα αριθμητικών δεδομένων Α και Β έχουν ίσες διαμέσους, ίδιο εύρος και ίδιο ενδοτεταρτημοριακό εύρος. Είναι δυνατόν τα θηκογράμματά τους να διαφέρουν; Να αιτιολογήσετε την απάντησή σας.

ΚΕΦΑΛΑΙΟ 2ο

Μέτρα θέσης και μεταβλητότητας

Σύντομη γνωριμία

Τα μέτρα θέσης και μεταβλητότητας αποτελούν μια σύνοψη των δεδομένων και παρέχουν σημαντικές πληροφορίες για την κατανομή τους. Η μέση τιμή και η διάμεσος μας δίνουν μια εικόνα για τη θέση του κέντρου των δεδομένων. Η διασπορά, η τυπική απόκλιση και το ενδοτεταρτημοριακό εύρος, μας δίνουν πληροφορίες για το βαθμό συγκέντρωσης των δεδομένων γύρω από το κέντρο τους. Είναι, επίσης, χρήσιμα εργαλεία για τη σύγκριση δεδομένων που προέρχονται από διαφορετικές ομάδες ή πληθυσμούς. Αυτό μας επιτρέπει να εξάγουμε συμπεράσματα σχετικά με τις διαφορές και τις ομοιότητες μεταξύ των συγκρινόμενων ομάδων. Τέλος, βοηθούν στην αναγνώριση ασυνήθιστων-ακραίων τιμών στα δεδομένα. Η κατανόηση των ιδιοτήτων τους είναι αναγκαία για τη σωστή επιλογή στατιστικής μεθόδου προκειμένου τα αποτελέσματα να είναι αξιόπιστα.

Περιεχόμενα

Karl Pearson
(1857 –1936)

Άγγλος μαθηματικός και στατιστικολόγος.
Του αποδίδεται η ίδρυση του κλάδου της Μαθηματικής Στατιστικής. Ίδρυσε το πρώτο τμήμα Στατιστικής στον κόσμο στο University College London το 1911.
Το έργο του καλύπτει ευρύ φάσμα εφαρμογών της Στατιστικής στα πεδία της Βιολογίας, της Επιδημιολογίας, της Ανθρωπομετρίας, της Ιατρικής, της Ψυχολογίας και της Κοινωνικής Ιστορίας.
Πολλές από τις γνωστές στατιστικές μεθόδους που χρησιμοποιούνται σήμερα οφείλονται σε αυτόν. Του αποδίδεται η δημιουργία του όρου τυπική απόκλιση.

2.1 Μέτρα Μεταβλητότητας

Περιέχονται:
- Μέση απόλυτη απόκλιση
- Διασπορά - Τυπική απόκλιση

Λέξεις κλειδιά:
Μέση απόλυτη απόκλιση
Διασπορά
Τυπική απόκλιση

Στην ενότητα αυτή θα μάθουμε:
Να περιγράφουμε και να προσδιορίζουμε την μέση απόλυτη απόκλιση, τη διασπορά και την τυπική απόκλιση ποσοτικών δεδομένων χρησιμοποιώντας απόλυτες και τετραγωνικές αποκλίσεις.

Δραστηριότητα
Μετρήσαμε τη θερμοκρασία (σε C^o) στο κέντρο μιας πόλης την πρώτη μέρα κάθε μήνα και ώρα 12.00 μ.μ., τους πρώτους πέντε μήνες του έτους 2024 και πήραμε τα παρακάτω αποτελέσματα:

	Ιαν	Φεβ	Μάρ	Απρ	Μάι
Θερμοκρασία	$x_1 = 5C^o$	$x_2 = 8C^o$	$x_3 = 14C^o$	$x_4 = 17C^o$	$x_5 = 26C^o$

α. Να υπολογίσετε τη μέση θερμοκρασία $\bar{x}$ της πόλης για τους παραπάνω μήνες.
β. Να συμπληρώσετε τον επόμενο πίνακα:

	Ιαν	Φεβ	Μάρ	Απρ	Μάι
x_i	$x_1 = 5C^o$	$x_2 = 8C^o$	$x_3 = 14C^o$	$x_4 = 17C^o$	$x_5 = 26C^o$
$\lvert x_i - \bar{x} \rvert$					
$(x_i - \bar{x})^2$					

γ. Να υπολογίσετε τη μέση τιμή των τιμών $\lvert x_i - \bar{x} \rvert$ της δεύτερης γραμμής και των τιμών $(x_i - \bar{x})^2$ της τρίτης γραμμής. Τι περιγράφουν αυτές οι μέσες τιμές;

Μέση απόλυτη απόκλιση - Διασπορά - Τυπική απόκλιση

Για την περιγραφή της μεταβλητότητας ενός συνόλου παρατηρήσεων έχουμε χρησιμοποιήσει μέχρι τώρα το εύρος (R) και το ενδοτεταρτημοριακό εύρος (Q). Στην ενότητα αυτή θα γνωρίσουμε τρία επιπλέον σημαντικά μέτρα μεταβλητότητας, τα οποία είναι η μέση απόλυτη απόκλιση (mean absolute difference), η μέση τετραγωνική απόκλιση ή αλλιώς διασπορά ή διακύμανση (variance) και η τυπική απόκλιση (standard deviation).

Στην προηγούμενη δραστηριότητα υπολογίσαμε τις παρακάτω μέσες τιμές:

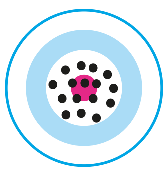
$$\frac{|x_1 - \bar{x}| + \ldots + |x_5 - \bar{x}|}{5} \quad \text{και} \quad \frac{(x_1 - \bar{x})^2 + \ldots + (x_5 - \bar{x})^2}{5}$$

Στη συνέχεια τη διαφορά $x_i - \bar{x}$ θα την ονομάζουμε απόκλιση της τιμής x_i από την μέση τιμή $\bar{x}$, ενώ την απόλυτη τιμή $\lvert x_i - \bar{x} \rvert$ θα την ονομάζουμε απόλυτη απόκλιση της τιμής x_i από την μέση τιμή $\overline{x}$. Το τετράγωνο $(x_i - \overline{x})^2$ θα το ονομάζουμε τετραγωνική απόκλιση της τιμής x_i από την μέση τιμή $\overline{x}$. Επομένως, οι μέσες τιμές που υπολογίσαμε μπορούμε να πούμε ότι είναι οι:

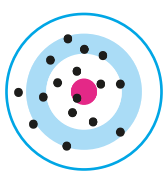
Μέση απόλυτη απόκλιση (ή MAD) $= \dfrac{|x_1 - \overline{x}| + \ldots + |x_5 - \overline{x}|}{5}$

Μέση τετραγωνική απόκλιση (ή διασπορά s^2) $= \dfrac{(x_1 - \overline{x})^2 + \ldots + (x_5 - \overline{x})^2}{5}$

Παρατηρούμε ότι οι δύο αυτές μέσες τιμές (MAD και s^2) καθορίζονται από τις αποστάσεις των τιμών των παρατηρήσεων από την μέση τιμή τους $\overline{x}$. Συγκεκριμένα, επειδή $|x_i - \overline{x}|^2 = (x_i - \overline{x})^2$, καθορίζονται από τις απόλυτες αποκλίσεις των τιμών. Όσο πιο μεγάλες είναι οι απόλυτες αποκλίσεις, τόσο πιο μεγάλες είναι η μέση απόλυτη απόκλιση MAD και η διασπορά s^2.



Μικρή μεταβλητότητα



Μεγάλη μεταβλητότητα

Το διπλανό σχήμα, με τον στόχο, αποδίδει οπτικά την έννοια της διασποράς και της MAD. Όσο πιο κοντά στο κέντρο του στόχου είναι οι βολές, τόσο πιο μικρές είναι η MAD και η διασπορά τους γύρω από αυτό. Πάλι, όσο πιο μακριά από το κέντρο του στόχου είναι οι βολές, τόσο πιο μεγάλες είναι η MAD και η διασπορά τους γύρω από αυτό.

Μπορούμε τώρα να δώσουμε τους παρακάτω ορισμούς:

Ορισμοί

Έστω ένα σύνολο ν παρατηρήσεων $x_1, x_2, \ldots, x_\nu$ με μέση τιμή $\overline{x}$.

Μέση απόλυτη απόκλιση (συμβολικά MAD- Mean Absolute Difference) ονομάζεται η μέση τιμή των απόλυτων αποκλίσεων:

$$MAD = \frac{|x_1 - \overline{x}| + \ldots + |x_\nu - \overline{x}|}{\nu}$$

Διασπορά (συμβολικά s^2) ή μέση τετραγωνική απόκλιση ονομάζεται η μέση τιμή των τετραγωνικών αποκλίσεων:

$$s^2 = \frac{(x_1 - \overline{x})^2 + \ldots + (x_\nu - \overline{x})^2}{\nu}$$

Τυπική απόκλιση ονομάζεται η τετραγωνική ρίζα της διασποράς:

$$s = \sqrt{s^2} = \sqrt{\frac{(x_1 - \overline{x})^2 + \ldots + (x_\nu - \overline{x})^2}{\nu}}$$

Μέτρα μεταβλητότητας

Γενικά: Η μέση απόλυτη απόκλιση και η τυπική απόκλιση περιγράφουν το πόσο μακριά, κατά μέσο όρο, βρίσκονται οι παρατηρήσεις από τη μέση τιμή τους.

Σημείωση 1: Παρατηρούμε ότι η μέση απόλυτη απόκλιση και η τυπική απόκλιση εκφράζονται και οι δύο στην ίδια μονάδα μέτρησης με τις παρατηρήσεις. Η διασπορά εκφράζεται στη μονάδα μέτρησης των παρατηρήσεων υψωμένη στο τετράγωνο.

Σημείωση 2: Η μέση απόλυτη απόκλιση, η διασπορά και η τυπική απόκλιση παίρνουν μη αρνητικές τιμές και είναι ίσες με μηδέν όταν οι παρατηρήσεις είναι ίσες μεταξύ τους αλλά και αντίστροφα.

Μετασχηματισμός δεδομένων-Πρόσθεση

Μετασχηματισμός δεδομένων-Πολλαπλασιασμός

Σημείωση 3. **Αποδεικνύεται ότι:**

- Αν σε κάποιες παρατηρήσεις με μέση τιμή $\bar{x}$, μέση απόλυτη απόκλιση MAD και τυπική απόκλιση s, προσθέσουμε τον ίδιο αριθμό α, τότε οι παρατηρήσεις που προκύπτουν έχουν μέση τιμή $\bar{x}+\alpha$ και ίδιες μέση απόλυτη απόκλιση και τυπική απόκλιση s.
- Αν τις παρατηρήσεις μιας ομάδας δεδομένων με μέση τιμή $\bar{x}$, μέση απόλυτη απόκλιση MAD και τυπική απόκλιση s τις πολλαπλασιάσουμε με τον ίδιο αριθμό α, τότε οι παρατηρήσεις που προκύπτουν έχουν μέση τιμή $\alpha\bar{x}$, μέση απόλυτη απόκλιση $|\alpha|$MAD και τυπική απόκλιση $|\alpha|s$. (βλέπε εφαρμογή 2 και για τις δύο περιπτώσεις)

Παράδειγμα 1
Να βρείτε την μέση απόλυτη απόκλιση, τη διασπορά και τη τυπική απόκλιση των παρατηρήσεων: −5, −2, 4, 7, 16

Λύση
Η μέση τιμή είναι

$$\bar{x}=\frac{-5-2+4+7+16}{5}=4$$

Έχουμε κατά σειρά:

$$MAD=\frac{|-5-4|+|-2-4|+|4-4|+|7-4|+|16-4|}{5}=\frac{9+6+0+3+12}{5}=6$$

$$s^2=\frac{(-5-4)^2+(-2-4)^2+(4-4)^2+(7-4)^2+(16-4)^2}{5}$$

$$=\frac{9^2+6^2+0^2+3^2+12^2}{5}=\frac{81+36+9+144}{5}=\frac{270}{5}=54$$

και $s=\sqrt{s^2}=\sqrt{54}\approx 7{,}35$

Προσπαθήστε να λύσετε τις ασκήσεις 1, 2, 3 και 4

Μέτρα μεταβλητότητας

Μέτρα μεταβλητότητας

Παράδειγμα 2
Θεωρούμε τις επόμενες ομάδες παρατηρήσεων:
Α ομάδα: 5, 5, 6, 6, 6, 8, 9, 9, 9,
Β ομάδα: 1, 1, 3, 3, 6, 10, 10, 14, 15
Γ ομάδα: 7, 7, 7, 7, 7, 7, 7, 7, 7

α. Να βρείτε την μέση τιμή τους.
β. Να εκτιμήσετε ποια ομάδα έχει μικρότερη και ποια μεγαλύτερη διασπορά.

Λύση
α. Κατά σειρά είναι:

$$\bar{x}_A=\frac{5+5+6+6+6+8+9+9+9}{9}=7$$

$$\bar{x}_B=\frac{1+1+3+3+6+10+10+14+15}{9}=7$$

και προφανώς $\bar{x}_\Gamma=7$. Παρατηρούμε ότι οι ομάδες έχουν ίδια μέση τιμή.

Όταν οι παρατηρήσεις είναι ίσες μεταξύ τους έχουν διασπορά μηδέν.

β. Παρότι οι τρείς ομάδες δεδομένων έχουν ίδια μέση τιμή διαφέρουν σημαντικά μεταξύ τους. Είναι φανερό ότι οι παρατηρήσεις της τρίτης ομάδας, οι οποίες είναι ίσες μεταξύ τους, έχουν διασπορά μηδέν. Οι παρατηρήσεις της δεύτερης ομάδας βρίσκονται «πιο μακριά» από την μέση τιμή τους, σε σχέση με τις παρατηρήσεις της πρώτης ομάδας που βρίσκονται «πιο κοντά» στη μέση τιμή τους. Επομένως, η δεύτερη ομάδα έχει μεγαλύτερη διασπορά από την πρώτη. Άρα μεγαλύτερη διασπορά έχει η ομάδα Β και μικρότερη η ομάδα Γ.

Προσπαθήστε να λύσετε την άσκηση 5

Εφαρμογή 1

Ρωτήσαμε 5 μαθητές και μαθήτριες πόσες ώρες, κατά μέσο όρο, αφιερώνουν την ημέρα στα κοινωνικά δίκτυα. Οι απαντήσεις που πήραμε είναι οι εξής:

0, 1, 3, 1, 0.

α. Να υπολογίσετε τη μέση τιμή των παρατηρήσεων.

β. Να υπολογίσετε τη μέση απόλυτη απόκλιση, τη διασπορά και την τυπική απόκλιση των ωρών αυτών.

Λύση

α. Η μέση τιμή των απαντήσεων είναι: $\bar{x}=\dfrac{0+1+3+1+0}{5}=\dfrac{5}{5}=1.$

β. Συμπληρώνουμε (για λόγους ευκολίας) τον επόμενο πίνακα:

x_i	$x_1=0$	$x_2=1$	$x_3=3$	$x_4=1$	$x_5=0$
$\lvert x_i-\bar{x}\rvert$	1	0	2	0	1
$(x_i-\bar{x})^2$	1	0	4	0	1

Από τις τιμές του παραπάνω πίνακα προκύπτει:

$$MAD=\frac{\lvert x_1-\bar{x}\rvert+\ldots+\lvert x_5-\bar{x}\rvert}{5}=\frac{1+0+2+0+1}{5}=\frac{4}{5}\ \text{ώρες}$$

$$s^2=\frac{(x_1-\bar{x})^2+\ldots.+(x_5-\bar{x})^2}{5}=\frac{1+0+4+0+1}{5}=\frac{6}{5}\ (\text{ώρες})^2$$

Η τυπική απόκλιση των δεδομένων είναι $s=\sqrt{s^2}=\sqrt{\dfrac{6}{5}}\approx 1{,}1$ ώρες.

Εφαρμογή 2

Με δεδομένο ότι οι παρατηρήσεις 1, 2, 3, 4, 5 έχουν μέση τιμή $\bar{x}=3,$ μέση απόλυτη απόκλιση $MAD=\dfrac{6}{5}$ και τυπική απόκλιση $s=\sqrt{2},$ να βρείτε τη μέση τιμή, τη μέση απόλυτη απόκλιση και την τυπική απόκλιση των παρατηρήσεων:

α. 5, 6, 7, 8, 9 β. $-5,-10,-15,-20,-25$

Λύση

α. Οι παρατηρήσεις 5, 6, 7, 8, 9 προκύπτουν από τις 1, 2, 3, 4, 5 αν προσθέσουμε σε καθεμία τον αριθμό 4. Επομένως, έχουν μέση τιμή $\bar{x}+4=3+4=7$, ίδια μέση απόλυτη απόκλιση $MAD=\frac{6}{5}$ και τυπική απόκλιση $s=\sqrt{2}$.

β. Οι παρατηρήσεις $-5,-10,-15,-20,-25$ προκύπτουν από τις 1, 2, 3, 4, 5 αν τις πολλαπλασιάσουμε με τον αριθμό -5.
Επομένως έχουν μέση τιμή $-5\bar{x}=-5\cdot 3=-15$, μέση απόλυτη απόκλιση $MAD=|-5|\frac{6}{5}=6$ και τυπική απόκλιση $|-5|s=5\sqrt{2}$.

Προσπαθήστε να λύσετε τις ασκήσεις 6 έως 12

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Να απαντήσετε στις ερωτήσεις:

α. Τι ονομάζουμε μέση τετραγωνική απόκλιση;

β. Ποια είναι η απόλυτη απόκλιση και ποια η τετραγωνική απόκλιση της τιμής 2 από την μέση τιμή $\bar{x}=5$;

γ. Ποια από τις ομάδες παρατηρήσεων Α και Β έχει μεγαλύτερη διασπορά;
Ομάδα Α:1, 2, 10, 18, 19 Ομάδα Β: 8, 9, 10, 11, 12

δ. Πώς ορίζεται η τυπική απόκλιση;

2 Χωρίς να κάνετε υπολογισμούς να εξηγήσετε γιατί η μέση απόλυτη απόκλιση, η διασπορά και η τυπική απόκλιση των παρακάτω ομάδων παρατηρήσεων είναι ίσες.
Ομάδα Α: 1, 2, 3, 4, 5
Ομάδα Β: $-1,\ -2,\ -3,\ -4,\ -5$

Ασκήσεις

3 Ρωτήσαμε 6 ασθενείς ενός νοσοκομείου πόσα ποτήρια νερό, κατά μέσο όρο, πίνουν καθημερινά και πήραμε τις επόμενες απαντήσεις:

5, 8, 0, 1, 2, 2

Να βρείτε την μέση απόλυτη απόκλιση, τη διασπορά και τη τυπική απόκλιση των παραπάνω παρατηρήσεων.

4 Να βρεθεί η μέση τιμή, η μέση απόλυτη απόκλιση, η διασπορά και η τυπική απόκλιση για τα δεδομένα του παρακάτω πίνακα.

Τιμή	Συχνότητα
3	1
4	3
5	3
6	1

5 Χωρίς να κάνετε υπολογισμούς να εκτιμήσετε ποια από τις επόμενες ομάδες παρατηρήσεων έχει μικρότερη και ποια μεγαλύτερη διασπορά.

Ομάδα 1: 1, 2, 3, 4, 5 Ομάδα 2: 2, 2, 3, 4, 4

6 Σε ένα τηλεοπτικό διάλειμμα προβλήθηκαν 12 διαφημίσεις με μέση διάρκεια 16 sec και τυπική απόκλιση 3 sec. Να βρείτε την μέση διάρκεια και την τυπική απόκλιση της διάρκειας των διαφημίσεων στις περιπτώσεις, όπου:

α. Ο χρόνος κάθε διαφήμισης μειωθεί κατά 2 sec.

β. Ο χρόνος κάθε διαφήμισης αυξηθεί κατά 10%.

7 Ο χρόνος εργασίας 5 υπαλλήλων, σε ώρες, είναι:

7, 8, 7, 9, 7

α. Να βρείτε την διασπορά και την τυπική απόκλιση των ωρών εργασίας.

β. Αν κάθε υπάλληλος κάνει 1,5 ώρες υπερωρίες θα αλλάξει η διασπορά των ωρών εργασίας;

8 Οι πόντοι που σημείωσαν 5 παίκτες μπάσκετ με βολές δύο πόντων ήταν:

20, 14, 18, 16, 12

α. Να βρείτε την διασπορά των πόντων αυτών.

β. Αν οι ίδιες βολές ήταν τριών πόντων πόσους πόντους θα είχε σημειώσει κάθε παίκτης και ποια θα ήταν η διασπορά τους;

9 Τα βιβλία που αγόρασαν δέκα πελάτες ενός βιβλιοπωλείου ήταν:

2, 1, 3, 4, 1, 2, 4, 5, 2, 6

α. Να βρείτε τη μέση απόλυτη απόκλιση των παραπάνω δεδομένων.

β. Το βιβλιοπωλείο προσέφερε σε κάθε πελάτη, ως δώρο, δύο επιπλέον βιβλία. Ποια είναι η νέα μέση απόλυτη απόκλιση του αριθμού των βιβλίων που πήραν οι πελάτες;

10 Τα χρήματα που είχαν μαζί τους 5 μαθητές σε ευρώ ήταν:

5, 10, 15, 20, 10

α. Να βρείτε την τυπική απόκλιση των χρημάτων των μαθητών.

β. Κάθε ένας από τους μαθητές αυτούς έδωσε από 2 ευρώ σε έναν έρανο. Ποια είναι η τυπική απόκλιση των χρημάτων που απέμειναν στους μαθητές;

11 Οι ηλικίες των μαθητών μιας τάξης έχουν μέση τιμή 15 χρόνια και 3 μήνες και διασπορά 4 μήνες. Να βρείτε τη μέση τιμή και την τυπική απόκλιση της ηλικίας των μαθητών ύστερα από 4 χρόνια.

12 Για τρεις αριθμούς x, y, z ισχύουν:

$$x+y+z=3 \quad \text{και} \quad x^2+y^2+z^2=3$$

Να βρεθεί η μέση τιμή, η διασπορά και η τυπική απόκλιση των αριθμών.

Μια απροσδόκητη εφαρμογή της Στατιστικής
Το πρόβλημα των γερμανικών αρμάτων μάχης
(German tank problem)

Κατά τη διάρκεια του Β' Παγκοσμίου Πολέμου, οι συμμαχικές δυνάμεις βρέθηκαν αντιμέτωπες με μια πρόκληση άκρως σημαντική:

Με ποιον τρόπο θα μπορούσαν να εκτιμήσουν τον μηνιαίο αριθμό παραγωγής αρμάτων μάχης (τανκ) των Γερμανών, στο πλαίσιο ενός πολέμου όπου η πρόσβαση σε ακριβείς πληροφορίες ήταν περιορισμένη;

Οι Σύμμαχοι βρέθηκαν αντιμέτωποι με την επιτακτική ανάγκη για αξιόπιστες εκτιμήσεις των γερμανικών δυνατοτήτων, απέναντι στον περιορισμένο όγκο πληροφοριών που είχαν στη διάθεσή τους, λόγω των πολεμικών συνθηκών.

Επέλεξαν να υιοθετήσουν στατιστικές μεθόδους και κατόρθωσαν να προσεγγίσουν με ικανοποιητική ακρίβεια τον αριθμό των γερμανικών τεθωρακισμένων, βασιζόμενοι σε δεδομένα που προέκυπταν από την παρατήρηση των σειριακών αριθμών των τεθωρακισμένων που καταστράφηκαν ή καταλήφθηκαν στο πεδίο της μάχης.

Με αυτόν τον τρόπο, αντιμετώπισαν με επιτυχία την έλλειψη πληροφοριών και προσέφεραν μια σημαντική συμβολή στη στρατηγική αντίληψη και την πολεμική τακτική της εποχής.

2.2 Μέτρα θέσης και μεταβλητότητας ποσοτικού χαρακτηριστικού στις στάθμες κατηγορικού χαρακτηριστικού

Λέξεις κλειδιά:
Μέτρα στις στάθμες κατηγορικού χαρακτηριστικού

Στην ενότητα αυτή θα μάθουμε:
Να περιγράφουμε και να προσδιορίζουμε τη μέση τιμή και τη διάμεσο, καθώς και τη διασπορά, την τυπική απόκλιση και το ενδοτεταρτημοριακό εύρος των τιμών ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.

Ημέρα	ATM_1	ATM_2
Δευτέρα	20	30
Τρίτη	18	22
Τετάρτη	25	33
Πέμπτη	13	17
Παρασκευή	37	39
Σάββατο	47	42
Κυριακή	15	20
Σύνολο	175	203

Δραστηριότητα
Στον διπλανό πίνακα παρουσιάζεται ο αριθμός των αναλήψεων που έγιναν σε δύο γειτονικά μηχανήματα αυτόματης ανάληψης (ATM) μιας τράπεζας κατά τη διάρκεια μιας εβδομάδας.

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο του αριθμού των αναλήψεων σε κάθε ATM.
β. Να υπολογίσετε τη διασπορά, την τυπική απόκλιση και το ενδοτεταρτημοριακό εύρος των αναλήψεων σε κάθε ATM.
γ. Αν η τράπεζα αποφάσισε, λόγω κόστους, να αποσύρει ένα από τα δύο ATM, ποιο θα τη συμβουλεύατε να αποσύρει και γιατί;

Στην ενότητα αυτή θα χρησιμοποιήσουμε τα μέτρα θέσης και μεταβλητότητας που είδαμε στις προηγούμενες ενότητες για να βγάλουμε συμπεράσματα ως προς τις διαφορές ομάδων δεδομένων.

Παράδειγμα
Ας μελετήσουμε, για παράδειγμα, τα δεδομένα της δραστηριότητας. Υπολογίζουμε τα κυριότερα μέτρα θέσης και μεταβλητότητας για τον αριθμό αναλήψεων σε κάθε ATM.

Για το ATM_1 έχουμε τα δεδομένα: 13, 15, 18, 20, 25, 37, 47. Είναι

$$\delta = 20,\quad Q_1 = 15,\quad Q_3 = 37,\quad Q = 37 - 15 = 22$$

$$\bar{x}_1 = \frac{13+15+18+20+25+37+47}{7} = \frac{175}{7} = 25$$

$$s_1^2 = \frac{(13-25)^2 + (15-25)^2 + \ldots + (47-25)^2}{7} \approx 135{,}14$$

$$s_1 = \sqrt{135{,}14} \approx 11{,}62$$

Για το ATM_2 έχουμε τα δεδομένα: 17, 20, 22, 30, 33, 39, 42. Είναι

$$\delta = 30,\quad Q_1 = 20,\quad Q_3 = 39,\quad Q = 39 - 20 = 19$$

$$\bar{x}_2 = \frac{17+20+22+30+33+39+42}{7} = \frac{203}{7} = 29$$

$$s_2^2 = \frac{(17-29)^2+(20-29)^2+\ldots+(42-29)^2}{7} = 80$$

$$s_2 = \sqrt{80} \approx 8{,}94$$

Αν θέλαμε να προτείνουμε στην τράπεζα να αποσύρει κάποιο από τα δύο ΑΤΜ αυτό θα ήταν το ATM_1 λαμβάνοντας, ως κριτήριο, τη μέση επισκεψιμότητα του ATM_2 από τους πελάτες της τράπεζας. Επιπλέον, ο αριθμός των αναλήψεων στο ATM_2 παρουσιάζει μικρότερη διασπορά και ενδοτεταρτημοριακό εύρος, που σημαίνει ότι η αβεβαιότητα ως προς τον μέσο αριθμό των πελατών είναι μικρότερη.

Προσπαθήσετε να λύσετε τις ασκήσεις 1 και 2

	Χρόνος αντίδρασης						
Α	6	7	6	5	7	6	5
Β	7	8	9	8	10	9	5

Εφαρμογή

Μια φαρμακευτική εταιρεία έχει κατασκευάσει δύο αναισθητικά φάρμακα, το Α και το Β. Για να υπολογίσει το χρόνο αντίδρασης ενός οργανισμού σε κάθε φάρμακο, δηλαδή το χρόνο που χρειάζεται ο οργανισμός για να περιέλθει σε αναισθησία, χορήγησε σε 7 πειραματόζωα την ίδια δόση του αναισθητικού Α. Έπειτα από κάποιο εύλογο χρονικό διάστημα, χορήγησε στα ίδια πειραματόζωα την ίδια δόση του αναισθητικού Β. Τα αποτελέσματα για τους χρόνους αντίδρασης (σε sec) δίνονται στον διπλανό πίνακα.

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των χρόνων αντίδρασης σε κάθε κατηγορία φαρμάκου.

β. Να μελετήσετε ως προς την μεταβλητότητα τους χρόνους αντίδρασης των δύο φαρμάκων.

γ. Υπολογίζοντας την τυπική απόκλιση να αναφέρετε πόσο, κατά μέσο όρο, απέχουν οι χρόνοι αντίδρασης από τη μέση τιμής τους σε κάθε κατηγορία φαρμάκου;

Λύση

α. Για τη μέση τιμή των χρόνων αντίδρασης σε κάθε φάρμακο έχουμε:

$$\bar{x}_A = \frac{6+7+6+5+7+6+5}{7} = \frac{42}{7} = 6 \quad \text{και} \quad \bar{x}_B = \frac{7+8+9+8+10+9+5}{7} = \frac{56}{7} = 8$$

Τα διατεταγμένα, σε αύξουσα σειρά, δεδομένα είναι:

Φάρμακο Α: 5, 5, 6, 6, 6, 7, 7 **Φάρμακο Β**: 5, 7, 8, 8, 9, 9, 10

Η διάμεσος των χρόνων για το φάρμακο Α είναι $\delta_A=6$ και για το φάρμακο Β είναι $\delta_B=8$.

β. Για τη διασπορά των χρόνων αντίδρασης έχουμε:

$$s_A^2 = \frac{(6-6)^2+(7-6)^2+(6-6)^2+(5-6)^2+(7-6)^2+(6-6)^2+(5-6)^2}{7} = \frac{4}{7}$$

$$s_B^2 = \frac{(7-8)^2+(8-8)^2+(9-8)^2+(8-8)^2+(10-8)^2+(9-8)^2+(5-8)^2}{7} = \frac{16}{7}$$

Επίσης, από τα διατεταγμένα δεδομένα βρίσκουμε:

$$Q_A = 7-5 = 2 \quad \text{και} \quad Q_B = 9-7 = 2$$

γ. Οι χρόνοι αντίδρασης για το φάρμακο Α απέχουν από τη μέση τιμή τους, κατά μέσο όρο, $\sqrt{s_A^2} = s_A \approx 0{,}76$ λεπτά. Για το φάρμακο Β, αντίστοιχα, βρίσκουμε $\sqrt{s_B^2} = s_B \approx 1{,}51$ λεπτά.

Προσπαθήσετε να λύσετε τις ασκήσεις 3 έως 6

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Δύο ομάδες παιδιών διαγωνίστηκαν πολλές φορές στο παιχνίδι της παντομίμας. Στον παρακάτω πίνακα παρουσιάζονται τα στατιστικά για τους χρόνους (σε λεπτά) που χρειάστηκαν οι δύο ομάδες για να ολοκληρώσουν το παιχνίδι. Να χαρακτηρίσετε καθεμία από τις επόμενες προτάσεις, ως αληθή ή ψευδή:

Μέτρα θέσης και μεταβλητότητας	Ομάδα Α	Ομάδα Β
$\bar{x}$	10	9
δ	8	7
s	4	3
Q	4	3

α. Για την ομάδα Β: Το «μεσαίο» 50% των χρόνων είναι 4 λεπτά.
β. Περίπου οι μισοί χρόνοι της ομάδας Α είναι μεγαλύτεροι από τη μέση τιμή τους.
γ. Οι χρόνοι της ομάδας Β απέχουν, κατά μέσο όρο, 9 λεπτά από τη μέση τιμή τους.
δ. Περίπου οι μισοί χρόνοι της ομάδας Β βρίσκονται αριστερά της δ=7.

Ασκήσεις

2 Στον πίνακα φαίνονται οι απαντήσεις που πήραμε από 5 άντρες και 6 γυναίκες στο ερώτημα για το πόσα μηνύματα έχουν στείλει από το κινητό τους την τελευταία εβδομάδα.

Άντρες	10	7	5	8	10	
Γυναίκες	4	12	11	14	1	6

α. Να δείξετε ότι η μέση τιμή του πλήθους των μηνυμάτων είναι ίδια για τα δύο φύλα.
β. Παρατηρώντας τα δεδομένα και χωρίς να κάνετε υπολογισμούς, να πείτε ποιο από τα δύο φύλα παρουσιάζει μεγαλύτερη μεταβλητότητα ως προς το πλήθος των μηνυμάτων που έχει στείλει.
γ. Να υπολογίσετε το ενδοτεταρτημοριακό εύρος Q, τη διασπορά s^2 και την τυπική απόκλιση s του πλήθους των μηνυμάτων κάθε φύλου.
δ. Να αξιολογήσετε την απάντηση σας στο ερώτημα (β), βασιζόμενοι στα δεδομένα που προέκυψαν από το ερώτημα (γ).

3 Δύο παιδιά, ο Μιχάλης και η Φωτεινή, συμπλήρωσαν από πέντε σταυρόλεξα. Οι χρόνοι (σε λεπτά) που χρειάστηκαν για να συμπληρώσουν 5 ίδια σταυρόλεξα, δίνονται στον παρακάτω πίνακα.

Μιχάλης	5	7	6	4	8
Φωτεινή	7	6	4	3	5

α. Να υπολογίσετε τη μέση τιμή του χρόνου συμπλήρωσης σταυρόλεξου κάθε παιδιού.
β. Υπολογίζοντας την τυπική απόκλιση να αναφέρετε πόσο, κατά μέσο όρο, απέχουν οι χρόνοι συμπλήρωσης από τη μέση τιμή τους για κάθε παιδί;

4 Στον παρακάτω πίνακα φαίνεται ο αριθμός αμυγδάλων που περιέχονται σε καθένα από 10, τυχαία επιλεγμένες, συσκευασίες (των 50gr) παραγωγής της εταιρείας Α και, αντίστοιχα, σε 10, τυχαία επιλεγμένες, συσκευασίες (των 50gr) της εταιρείας Β.

ΕΤΑΙΡΕΙΑ Α			ΕΤΑΙΡΕΙΑ Β		
30	32	30	34	31	32
31	31	33	30	31	31
32	31	30	33	33	34
30			34		

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο του αριθμού των αμυγδάλων στις δύο συσκευασίες.

β. Να υπολογίσετε το ενδοτεταρτημοριακό εύρος Q, τη διασπορά s^2 και την τυπική απόκλιση s του αριθμού των αμυγδάλων στις δύο συσκευασίες.

γ. Πώς μπορούν οι πληροφορίες που πήρατε από τα προηγούμενα ερωτήματα να επηρεάσουν τις αποφάσεις σας σχετικά με την επιλογή μιας εταιρείας για την αγορά αμυγδάλων;

5 Στον παρακάτω πίνακα φαίνονται οι απαντήσεις 7 παιδιών που αξιολόγησαν τη γεύση των παγωτών Α και Β στην κλίμακα 1 έως 10.

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των βαθμών του παγωτού Α και του παγωτού Β.

β. Να υπολογίσετε το ενδοτεταρτημοριακό εύρος Q, τη διασπορά s^2 και την τυπική απόκλιση s των βαθμών των δύο παγωτών.

ΠΑΓΩΤΟ Α	ΠΑΓΩΤΟ Β
3	5
5	1
10	4
1	6
7	5
6	6
5	5

γ. Ποιο παγωτό φαίνεται να απέσπασε τις πιο θετικές αξιολογήσεις από τα παιδιά, κατά μέσο όρο; Ποιο παγωτό φαίνεται να έχει μικρότερη μεταβλητότητα στις αξιολογήσεις των παιδιών, με βάση το ενδοτεταρτημοριακό εύρος και την τυπική απόκλιση;

6 Στον παρακάτω πίνακα παρουσιάζονται οι πόντοι των παικτών δύο ομάδων μπάσκετ που διαγωνίστηκαν στον τελικό κυπέλου.

α. Πόσους πόντους, κατά μέσο όρο, πέτυχε κάθε παίχτης της ομάδας Α και πόσους της Β; Ποια είναι η διάμεσος των πόντων σε κάθε ομάδα;

β. Να υπολογίσετε το ενδοτεταρτομοριακό εύρος Q, τη διασπορά s^2 και την τυπική απόκλιση s των πόντων κάθε ομάδας.

γ. Ποιοι παίκτες της ομάδας Β ξεχώρισαν στο παιχνίδι; Ποιες πτυχές της απόδοσης της ομάδας Α θα μπορούσαν να βελτιωθούν για τη μελλοντική της απόδοση;

Παίκτης	Ομάδα Α	Ομάδα Β
1	0	0
2	2	0
3	10	15
4	0	10
5	0	0
6	22	25
7	6	32
8	30	0
9	4	12
10	6	0
11	7	6
12	5	0
Σύνολο	92	100

2.3 Μέτρα σχετικής μεταβλητότητας

Περιέχονται
- Συντελεστής μεταβλητότητας
- Σύγκριση ομάδων τιμών ως προς την ομοιογένεια
- Σύγκριση ως προς την μεταβλητότητα ομάδων τιμών με διαφορετική μονάδα μέτρησης
- Περιγραφή του συντελεστή μεταβλητότητας ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού

Λέξεις κλειδιά:
Συντελεστής μεταβλητότητας
Ομοιογένεια

Συντελεστής μεταβλητότητας

Στην ενότητα αυτή θα μάθουμε:
- Την έννοια του συντελεστή μεταβλητότητας και τη χρησιμότητά του στη σύγκριση της μεταβλητότητας ομάδων τιμών με διαφορετικούς μέσους ή με διαφορετική μονάδα μέτρησης.
- Να προσδιορίζουμε και να περιγράφουμε τον συντελεστή μεταβλητότητας ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.

Συντελεστής μεταβλητότητας

Τα μέτρα διασποράς που συναντήσαμε μέχρι τώρα, όπως το εύρος, το ενδοτεταρτημοριακό εύρος, η διασπορά, η τυπική απόκλιση και η μέση απόλυτη απόκλιση δεν μπορούν να χρησιμοποιηθούν για συγκρίσεις ομάδων τιμών οι οποίες εκφράζονται σε διαφορετική μονάδα μέτρησης ή έχουν διαφορετικές μέσες τιμές. Στην περίπτωση αυτή χρήσιμα είναι τα μέτρα σχετικής μεταβλητότητας με σημαντικότερο τον συντελεστή μεταβλητότητας (coefficient of variation), ο οποίος συμβολίζεται με CV και είναι ανεξάρτητος από τις μονάδες μέτρησης.

Ορισμός

Αν ένα σύνολο αριθμητικών δεδομένων έχει μέση τιμή $\bar{x} \neq 0$ και τυπική απόκλιση s τότε ορίζουμε:

$$CV = \frac{s}{|\bar{x}|} \quad \text{ή} \quad CV = \frac{s}{|\bar{x}|}100\%$$

Αν $\bar{x} = 0$ τότε δεν ορίζουμε συντελεστή μεταβλητότητας.

Αν δύο σύνολα δεδομένων A και B έχουν συντελεστές μεταβλητότητας CV_A και CV_B αντίστοιχα, με $CV_A < CV_B$, τότε λέμε ότι το σύνολο A είναι πιο ομοιογενές από το σύνολο B και εννοούμε ότι η σχετική μεταβλητότητα του A είναι μικρότερη αυτής του B.

Τέλος, ένα σύνολο δεδομένων θα λέγεται ομοιογενές, όταν έχει συντελεστή μεταβλητότητας $CV \leq 10\%$.

Προσπαθήστε να λύσετε τις ασκήσεις 1 και 2

Σύγκριση ομάδων τιμών ως προς την ομοιογένεια

Στην επόμενη εφαρμογή βλέπουμε, ότι για να συγκρίνουμε, ως προς την ομοιογένεια, δύο ομάδες δεδομένων, χρησιμοποιούμε τον συντελεστή μεταβλητότητας.

	Δείγμα A	Δείγμα B
$\bar{x}$	80	50
s	10	4

Εφαρμογή 1

Για δύο δείγματα δύο διαφορετικών πληθυσμών προέκυψαν τα δεδομένα του διπλανού πίνακα (τα δεδομένα έχουν ίδια μονάδα μέτρησης).

α. Ποιο από τα δύο δείγματα έχει μεγαλύτερη ομοιογένεια;

β. Είναι τα δείγματα ομοιογενή;

Λύση

α. Είναι

$$\frac{s_A}{\bar{x}_A} = \frac{10}{80} = 0{,}125 \quad \text{και} \quad \frac{s_B}{\bar{x}_B} = \frac{4}{50} = 0{,}08.$$

Άρα

$$CV_A = 12{,}5\% \quad \text{και} \quad CV_B = 8\%$$

Επειδή $CV_B < CV_A$ το δείγμα Β είναι πιο ομοιογενές από το δείγμα Α.

β. Επειδή $CV_B < 10\%$ και $CV_A > 10\%$ το δείγμα Α δεν είναι ομοιογενές, ενώ το δείγμα Β είναι ομοιογενές.

Συντελεστής μεταβλητότητας

Προσπαθήστε να λύσετε τις ασκήσεις 3, 4, 5, 7 και 9

Σύγκριση ως προς την μεταβλητότητα ομάδων τιμών με διαφορετική μονάδα μέτρησης

Ο συντελεστής μεταβλητότητας μπορεί, επίσης, να χρησιμοποιηθεί για τη σύγκριση της μεταβλητότητας ομάδων δεδομένων με διαφορετική μονάδα μέτρησης όπως φαίνεται στην επόμενη εφαρμογή:

Εφαρμογή 2

Εξετάσαμε έναν πληθυσμό ανθρώπων, ως προς το ύψος (σε cm) και το βάρος τους (σε kg) και προέκυψαν τα αποτελέσματα του διπλανού πίνακα. Να συγκρίνετε τα δεδομένα του βάρους και του ύψους, ως προς την ομοιογένεια.

	Βάρος	Ύψος
$\bar{x}$	85	180
s	10	12

Λύση

Θα χρησιμοποιήσουμε τον συντελεστή μεταβλητότητας, που είναι ανεξάρτητος από τις μονάδες μέτρησης. Έχουμε:

$$\text{Βάρος: } CV_1 = \frac{10}{85} \cdot 100\% \approx 11{,}76\% \quad \text{και} \quad \text{Ύψος: } CV_2 = \frac{12}{180} \cdot 100\% \approx 6{,}7\%.$$

Συμπεραίνουμε ότι υπάρχει μεγαλύτερη ομοιογένεια στο ύψος απ' ότι στο βάρος του δείγματος.

Συντελεστής μεταβλητότητας

Από τα παραπάνω συμπεραίνουμε ότι:

Ο συντελεστής μεταβλητότητας μπορεί να χρησιμοποιηθεί:

- Ως μέτρο σύγκρισης της μεταβλητότητας δύο ή περισσοτέρων ομάδων ποσοτικών δεδομένων με διαφορετικές μέσες τιμές ή με διαφορετικές μονάδες μέτρησης.
- Ως μέτρο της ομοιογένειας μιας ομάδας ποσοτικών δεδομένων.

Προσπαθήστε να λύσετε τις ασκήσεις 6 και 8

Περιγραφή του συντελεστή μεταβλητότητας ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη ενός κατηγορικού χαρακτηριστικού

Εφαρμογή 3

Μια μαθήτρια θέλει να μελετήσει ως προς τη μεταβλητότητα των θερμίδων τα αναψυκτικά τύπου light και τα αναψυκτικά που δεν είναι light. Για το σκοπό αυτό, συνέλεξε ένα δείγμα πέντε αναψυκτικών από κάθε κατηγορία και έφτιαξε τον διπλανό πίνακα:

	Θερμίδες kcal/100g				
light	1	2	4	4	1
όχι light	40	50	45	35	60

α. Να υπολογίσετε τη μέση τιμή, τη διασπορά και την τυπική απόκλιση των παρατηρήσεων για κάθε τύπο αναψυκτικού.

β. Να προτείνεται ένα κατάλληλο μέτρο μεταβλητότητας για τη σύγκριση, ως προς την ομοιογένεια των θερμίδων, των δύο αναψυκτικών.

γ. Να υπολογίσετε τους συντελεστές μεταβλητότητας κάθε κατηγορίας αναψυκτικού και να βρείτε ποια κατηγορία αναψυκτικού παρουσιάζει μικρότερη ομοιογένεια.

Λύση

α. Για τα αναψυκτικά τύπου light βρίσκουμε:

$$\bar{x}_1 = \frac{1+2+4+4+1}{5} = 2{,}4\ \text{kcal/100g}$$

$$s_1^2 = \frac{(1-2{,}4)^2+(2-2{,}4)^2+(4-2{,}4)^2+(4-2{,}4)^2+(1-2{,}4)^2}{5} = 1{,}84\ (\text{kcal/100g})^2$$

$$s_1 = \sqrt{s_1^2} = \sqrt{1{,}84} \approx 1{,}36\ \text{kcal/100g}$$

Για τα αναψυκτικά που δεν είναι light βρίσκουμε:

$$\bar{x}_2 = \frac{40+50+45+35+60}{5} = 46\,\text{kcal/100g}$$

$$s_2^2 = \frac{(40-44)^2+(50-44)^2+(45-44)^2+(35-44)^2+(60-44)^2}{5} = 74\ (\text{kcal/100g})^2$$

$$s_2 = \sqrt{s_1^2} = \sqrt{74} \approx 8{,}6\ \text{kcal/100g}$$

β. Από το προηγούμενο ερώτημα βλέπουμε ότι οι μέσες τιμές των θερμίδων των δύο κατηγοριών διαφέρουν σε μεγάλο βαθμό. Επομένως, δεν μπορούμε να συγκρίνουμε άμεσα τις τυπικές αποκλίσεις (ή τις διασπορές) κάθε κατηγορίας για να συγκρίνουμε τις μεταβλητότητες των τιμών κάθε κατηγορίας. Ένα κατάλληλο μέτρο είναι ο συντελεστής μεταβλητότητας CV.

γ. Από τα αποτελέσματα του πρώτου ερωτήματος έχουμε:

$$\text{light: } CV_1 = \frac{s_1}{\bar{x}_1}\cdot 100\% = 56{,}7\% \quad \text{και} \quad \text{όχι light: } CV_2 = \frac{s_2}{\bar{x}_2}\cdot 100\% = 18{,}7\%.$$

Μπορούμε τώρα να ισχυριστούμε ότι τα αναψυκτικά τύπου light παρουσιάζουν μικρότερη ομοιογένεια, διότι $CV_1 > CV_2$.

Προσπαθήστε να λύσετε τις ασκήσεις 10 και 11

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Να χαρακτηρίσετε, ως αληθή ή ψευδή καθεμία από τις επόμενες προτάσεις:

α. Ο συντελεστής μεταβλητότητας είναι μέτρο θέσης.

β. Μεταξύ δύο συνόλων παρατηρήσεων μεγαλύτερη ομοιογένεια έχει αυτό με τον μεγαλύτερο συντελεστή μεταβλητότητας.

γ. Ομοιογενές είναι ένα δείγμα με $CV \leq 10\%$.

δ. Ο συντελεστής μεταβλητότητας χρησιμοποιείται για να συγκρίνουμε ως προς την μεταβλητότητα ομάδες παρατηρήσεων με διαφορετική μονάδα μέτρησης.

ε. Αν ένα δείγμα τιμών έχει $\bar{x} = -4$ και $s = 8$, τότε έχει $CV = 2$.

2 Να απαντήσετε στις ερωτήσεις:

α. Αν αλλάξουμε τη μονάδα μέτρησης ενός συνόλου παρατηρήσεων τότε ο CV θα αλλάξει;

β. Ποιο από τα γνωστά σας μέτρα μπορεί να χρησιμοποιηθεί ως μέτρο της ομοιογένειας μιας ομάδας ποσοτικών δεδομένων;

γ. Πότε μια ομάδα δεδομένων θεωρείται ομοιογενής;

δ. Είναι το σύνολο των δεδομένων 5, 5, 5, 5, 5, 5 ομοιογενές;

Ασκήσεις

3 Η μέση τιμή και η τυπική απόκλιση ενός συνόλου παρατηρήσεων είναι 10 και 2 αντίστοιχα. Να υπολογίσετε τον CV.

4 Ένα δείγμα παρατηρήσεων έχει μέση τιμή 5 και συντελεστή μεταβλητότητας 40%. Να υπολογίστε την τυπική απόκλιση του δείγματος.

5 Ένας πληθυσμός έχει τυπική απόκλιση 10 και συντελεστή μεταβλητότητας 5%. Να υπολογίσετε τη μέση τιμή του πληθυσμού. Είναι ο πληθυσμός ομοιογενής;

6 Η ετήσια μέση τιμή και τυπική απόκλιση των καταθέσεων (σε ευρώ) του κ. Γιώργου και του κ. Δημήτρη (σε δολάρια) φαίνονται στον παρακάτω πίνακα:

	κ. Γιώργος	κ. Δημήτρης
Μέσος όρος καταθέσεων	$\bar{x}_A = 1000€$	$\bar{x}_B = 20000\$$
Τυπική απόκλιση	$s_A = 70€$	$s_B = 80\$$

α. Μπορούμε να ισχυριστούμε ότι η μεταβλητότητα των καταθέσεων του κ. Δημήτρη είναι μεγαλύτερη από αυτή του κ. Γιώργου επειδή $s_B > s_A$;

β. Οι καταθέσεις ποιου κυρίου παρουσίασαν μεγαλύτερη σχετική μεταβλητότητα;

7 Η μέση τιμή και η διασπορά των βαθμών 20 παιδιών μιας τάξης Α' Λυκείου στα μαθήματα των Μαθηματικών και της Χημείας δίνονται στον επόμενο πίνακα:

Μάθημα	Μέση τιμή $\bar{x}$	Διασπορά s^2
Μαθηματικά	15	9
Χημεία	17	4

Σε ποιο από τα δύο μαθήματα η τάξη παρουσιάζει μεγαλύτερη ομοιογένεια βαθμών;

8 Καθεμία από τις παρακάτω ομάδες παρατηρήσεων έχουν μέση τιμή 10.

Ομάδα 1: 0kg, 5 kg, 10 kg, 15 kg, 20 kg

Ομάδα 2: 0cm, 1 cm, 10 cm, 19 cm, 20 cm

Ομάδα 3: 0sec, 9 sec, 10 sec, 11 sec, 20 sec

Χωρίς να κάνετε υπολογισμούς, να βρείτε ποια ομάδα παρουσιάζει τη μεγαλύτερη ομοιογένεια και ποια τη μικρότερη.

9 Ο χρόνος (σε λεπτά) που χρειάστηκε ένα δείγμα 5 μαθητών και μαθητριών για να λύσουν μια άσκηση Στατιστικής είναι: 5, 10, 8, 2, 10. Να υπολογίσετε τον CV. Είναι το δείγμα ομοιογενές;

10 Στον παρακάτω πίνακα φαίνονται οι αποδόσεις δύο παικτών Α και Β στο άθλημα της σκοποβολής. Κάθε αριθμός εκφράζει την απόσταση της βολής από το κέντρο του στόχου (σε cm).

Παίκτης Α	2	5	7	2	2
Παίκτης Β	3	5	4	2	6

Οι βολές ποιου παίκτη παρουσιάζουν μεγαλύτερη ομοιογένεια ως προς την απόστασή τους από το στόχο;

2.4 Επίδραση των ακραίων τιμών στα μέτρα θέσης και μεταβλητότητας

Περιέχονται:
- Πώς επηρεάζονται η μέση τιμή και η διάμεσος από τις ακραίες τιμές
- Πώς επηρεάζονται το εύρος, το ενδοτεταρτημοριακό εύρος, η διασπορά και η τυπική απόκλιση από τις ακραίες τιμές
- Επιλογή μέτρων θέσης και μεταβλητότητας ανάλογα με την ύπαρξη ακραίων τιμών

Λέξεις κλειδιά:
Επίδραση ακραίων τιμών
Μέτρα θέσης
Μέτρα μεταβλητότητας

Επίδραση ακραίας τιμής στα μέτρα θέσης

Στην ενότητα αυτή θα μάθουμε:
- Πώς επηρεάζονται τα μέτρα θέσης και τα μέτρα μεταβλητότητας από την ύπαρξη ακραίων τιμών.
- Να επιλέγουμε κατάλληλα μέτρα θέσης και μέτρα μεταβλητότητας ποσοτικών δεδομένων ανάλογα με την ύπαρξη ακραίων τιμών.

Δραστηριότητα

Οι κριτικές σε κλίμακα από το 1 έως το 10 μιας ξενοδοχειακής μονάδας από 10 πελάτες της είναι:

1, 7, 7, 7, 8, 8, 8, 9, 9, 10

α. Να βρείτε τη μέση τιμή και τη διάμεσο των παρατηρήσεων.
β. Να βρείτε τη διασπορά, το εύρος και το ενδοτεταρτημοριακό εύρος των παρατηρήσεων.
γ. Υπάρχουν ακραίες τιμές στις παραπάνω παρατηρήσεις;
δ. Ο πελάτης που βαθμολόγησε τη μονάδα με 1 απέσυρε την κριτική του γιατί έκανε λάθος. Υπάρχουν ακραίες τιμές στις νέες παρατηρήσεις;
ε. Να υπολογίσετε τα μέτρα θέσης και μεταβλητότητας των ερωτημάτων (α) και (β) για τα δεδομένα που απομένουν μετά την απόσυρση της κριτικής. Τι παρατηρείτε;

Πώς επηρεάζονται η μέση τιμή και η διάμεσος από τις ακραίες τιμές

Στο Γυμνάσιο μάθαμε ότι αν από ένα σύνολο δεδομένων απομακρύνουμε τις ακραίες τιμές, η μέση τιμή και η διάμεσος δεν επηρεάζονται στον ίδιο βαθμό αλλά η μέση τιμή επηρεάζεται σε μεγαλύτερο βαθμό σε σχέση με τη διάμεσο.

Ας θυμηθούμε τα παραπάνω με την εφαρμογή που ακολουθεί.

Εφαρμογή 1

Τα ετήσια εισοδήματα (σε χιλιάδες ευρώ) πέντε οικογενειών είναι:

3, 12, 14, 15, 16, 40

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των παραπάνω τιμών.
β. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των τιμών που απομένουν αν απομακρύνουμε:
  i. Την ακραία τιμή στα αριστερά της κατανομής των τιμών.
  ii. Την ακραία τιμή στα δεξιά της κατανομής των τιμών.

Τι παρατηρείτε;

Λύση

α. Η μέση τιμή των δεδομένων είναι $\bar{x}=\dfrac{3+12+14+15+16+40}{6}\approx 16{,}67$ και η διάμεσός τους είναι $\delta=\dfrac{14+15}{2}=14{,}5$.

β. Είναι $Q_1 = 12$, $Q_3 = 16$ και $Q = 4$. Άρα $[Q_1 - 1{,}5Q_1, Q_3 + 1{,}5Q_3] = [6, 22]$.

βi. Η ακραία τιμή στα αριστερά της κατανομής των τιμών είναι η τιμή 3. Μετά την απομάκρυνση αυτής της τιμής τα δεδομένα είναι τα παρακάτω:

$$12, 14, 15, 16, 40$$

Η νέα μέση τιμή των δεδομένων είναι $\bar{x}_1 = \frac{12+14+15+16+40}{5} = 19{,}4$ και η διάμεσος είναι $\delta_1 = 15$.

βii. Η ακραία τιμή στα δεξιά της κατανομής των τιμών είναι η τιμή 40. Μετά την απομάκρυνση αυτής της τιμής τα δεδομένα είναι τα παρακάτω:

$$3, 12, 14, 15, 16$$

Η νέα μέση τιμή των δεδομένων είναι $\bar{x}_2 = \frac{3+12+14+15+16}{5} = 12$ και η διάμεσος είναι $\delta_2 = 14$.

Παρατηρούμε ότι:

- Όταν απομακρύναμε την ακραία τιμή στα αριστερά της κατανομής των τιμών η μέση τιμή αυξήθηκε.
- Όταν απομακρύναμε την ακραία τιμή στα δεξιά της κατανομής των τιμών η μέση τιμή μειώθηκε.
- Όταν απομακρύναμε μία ακραία τιμή, η μέση τιμή μεταβλήθηκε περισσότερο από τη διάμεσο.

Μάντεψε τη μέση τιμή

Προσπαθήστε να λύσετε τις ασκήσεις 3, 4 και 10

Πώς επηρεάζονται το εύρος, το ενδοτεταρτημοριακό εύρος, η διασπορά και η τυπική απόκλιση από τις ακραίες τιμές

Στην εφαρμογή που ακολουθεί εξετάζουμε την επίδραση των ακραίων τιμών στο εύρος, στο ενδοτεταρτημοριακό εύρος, στη διασπορά και στην τυπική απόκλιση ενός συνόλου δεδομένων.

Επίδραση ακραίας τιμής στα μέτρα μεταβλητότητας

Εφαρμογή 2

Οι ηλικίες 8 παιδιών και ενός ενήλικα συνοδού σε μια εορταστική εκδήλωση είναι:

$$9,\ 10,\ 10,\ 10,\ 11,\ 11,\ 12,\ 12,\ 50$$

α. Να υπολογίσετε το εύρος και ενδοτεταρτημοριακό εύρος των παρατηρήσεων.

β. Υπάρχει ακραία τιμή στις παρατηρήσεις;

γ. Να υπολογίσετε την διασπορά και την τυπική απόκλιση των παρατηρήσεων.

δ. Αφού απομακρύνετε την ακραία τιμή να υπολογίσετε το εύρος, το ενδοτεταρτημοριακό εύρος, τη διασπορά και την τυπική απόκλιση για τα δεδομένα που απέμειναν.

Τι παρατηρείτε;

Λύση

α. Από τα δεδομένα έχουμε:

$$R = 50 - 9 = 41,\ \ Q_1 = 10 \text{ και } Q_3 = 12$$

Το ενδοτεταρτημοριακό εύρος είναι $Q = Q_3 - Q_1 = 2$.

β. Στο διάστημα $[Q_1 - 1,5Q, Q_3 + 1,5Q] = [7,15]$ δεν ανήκει μόνο η τιμή 50 και επομένως αυτή είναι η μοναδική ακραία τιμή των παρατηρήσεων.

γ. Για τη διασπορά και την τυπική απόκλιση βρίσκουμε:

$$\bar{x} = \frac{9+10+10+10+11+11+12+12+50}{9} = 15$$

$$s^2 = \frac{(9-15)^2 + (10-15)^2 + \ldots + (50-15)^2}{9} = 154$$

$$s = \sqrt{s^2} = \sqrt{154} \approx 12,41$$

δ. Τα δεδομένα χωρίς την ακραία τιμή 50 είναι:

9, 10, 10, 10, 11, 11, 12, 12

Έχουμε:

$$R' = 12 - 9 = 3,\ \ Q_1' = 10 \text{ και } Q_3' = 11,5$$

Το ενδοτεταρτημοριακό εύρος είναι $Q' = Q_3' - Q_1' = 1,5$.

Για τη διασπορά και την τυπική απόκλιση βρίσκουμε:

$$\bar{x}' = \frac{9+10+10+10+11+11+12+12}{8} = 10,625$$

$$s'^2 = \frac{(9-10,625)^2 + (9-10,625)^2 + \ldots + (12-10,625)^2}{8} = 0,98$$

$$s' = \sqrt{s'^2} = \sqrt{0,98} \approx 0,99$$

Παρατηρούμε ότι:

- Όταν απομακρύναμε την ακραία τιμή το ενδοτεταρτημοριακό εύρος ελαττώθηκε ελάχιστα.
- Όταν απομακρύναμε την ακραία τιμή το εύρος, η διασπορά και η τυπική απόκλιση ελαττώθηκαν σημαντικά.

Προσπαθήστε να λύσετε τις ασκήσεις 2, 5, 7 και 8

Επιλογή μέτρων θέσης και μεταβλητότητας ανάλογα με την ύπαρξη ακραίων τιμών

Συνοψίζοντας τα αποτελέσματα των δύο προηγούμενων παραγράφων συμπεραίνουμε ότι η ύπαρξη ακραίων τιμών:

- επηρεάζει σε μεγαλύτερο βαθμό τη μέση τιμή απ'ότι τη διάμεσο και
- επηρεάζει σε μεγαλύτερο βαθμό τη διασπορά, την τυπική απόκλιση και το εύρος απ' ότι το ενδοτεταρτημοριακό εύρος.

Η διασπορά και η τυπική απόκλιση είναι μέτρα της απόκλισης των δεδομένων γύρω από τη μέση τιμή τους. Συνεπώς, έχει νόημα να χρησιμοποιούνται σε συνδυασμό με τη μέση τιμή ως μέτρο θέσης των δεδομένων. Αντίστοιχα, το ενδοτεταρτημοριακό εύρος έχει νόημα να χρησιμοποιηθεί σε συνδυασμό με τη διάμεσο.

Γενικά:

- Όταν έχουμε ένα σύνολο δεδομένων με ακραίες τιμές, είναι προτιμότερο να χρησιμοποιήσουμε τη διάμεσο και το ενδοτεταρτημοριακό εύρος για να περιγράψουμε τα δεδομένα.
- Όταν στα δεδομένα δεν περιέχονται ακραίες τιμές, προτιμούμε, συνήθως, τη μέση τιμή και την τυπική απόκλιση για την περιγραφή των δεδομένων, αφού για τον υπολογισμό τους λαμβάνουν υπόψη το σύνολο των δεδομένων.

Εφαρμογή 3

Παρακάτω δίνονται οι χρόνοι (σε λεπτά) που χρειάστηκαν 10 λεωφορεία του ΚΤΕΛ Αθηνών για να φτάσουν στη Θεσσαλονίκη, όπως αυτοί καταγράφηκαν στα αρχεία του ΚΤΕΛ.

365, 369, 361, 370, 375, 367, 426, 365, 371, 36

α. Υπάρχουν ακραίες τιμές στα δεδομένα και αν ναι, που νομίζετε ότι οφείλονται;

β. Αφού απομακρύνετε την αφύσικη ακραία τιμή, να υπολογίσετε τη μέση τιμή, τη διάμεσο, το εύρος, το ενδοτεταρτημοριακό εύρος και την τυπική απόκλιση των χρόνων:

- με την ακραία τιμή 426
- χωρίς την ακραία τιμή 426

γ. Αξιοποιώντας τα αποτελέσματα του προηγούμενου ερωτήματος να πείτε ποια μέτρα θέσης και μεταβλητότητας θεωρείτε κατάλληλα για να περιγράψουν τους χρόνους που καταγράφηκαν στα αρχεία του ΚΤΕΛ.

Λύση

α. Διατάσσουμε σε αύξουσα σειρά τους χρόνους:

36, 361, 365, 365, 367, 369, 370, 371, 375, 426

Παρατηρούμε ότι στους παραπάνω χρόνους υπάρχει μία πολύ μικρή τιμή, η τιμή 36, και μία σχετικά μεγάλη τιμή, η τιμή 426. Για την τιμή 36 είμαστε βέβαιοι ότι δεν ανταποκρίνεται σε πραγματικό χρόνο. Είναι αδύνατο ένα λεωφορείο να έκανε το δρομολόγιο Αθήνα-Θεσσαλονίκη σε 36 λεπτά. Θα πρέπει λοιπόν να την απομακρύνουμε από τα δεδομένα διότι οφείλετε, προφανώς, σε λάθος καταγραφή. Η τιμή 426 (7 ώρες και 6 λεπτά) μπορεί να προέκυψε από καθυστέρηση λόγω αυξημένης κίνησης αυτοκινήτων στο δρόμο ή από μηχανική βλάβη του λεωφορείου ή από πρόβλημα κάποιου επιβάτη που ανάγκασε το λεωφορείο να κάνει στάση κ.λπ. Η τιμή 426 είναι μια δυνατή τιμή για το συγκεκριμένο δρομολόγιο και επομένως, δεν υπάρχει λόγος για να την απομακρύνουμε.

β. Τα αποτελέσματα που προκύπτουν έπειτα από τους υπολογισμούς συνοψίζονται στους παρακάτω πίνακες.

$\bar{x}$	≈374,33
δ	369
R	65
Q	8
s	≈18,66

ΠΙΝΑΚΑΣ 1. Στατιστικά αρχικών δεδομένων χωρίς την τιμή 36.

$\bar{x}$	≈367,88
δ	368
R	14
Q	5,5
s	≈4,04

ΠΙΝΑΚΑΣ 2. Στατιστικά δεδομένων χωρίς τις τιμές 36 και 426.

γ. Από το προηγούμενο ερώτημα προκύπτουν οι παρακάτω παρατηρήσεις:

Μέση τιμή: Παρατηρούμε ότι η μέση τιμή $\bar{x} \approx 374{,}33$ είναι μεγαλύτερη από την πλειοψηφία των χρόνων (7 από τους 9 χρόνους είναι μικρότεροι της μέσης τιμής). Ο λόγος που συμβαίνει αυτό είναι η ύπαρξη του μεγάλου χρόνου 426 στα δεδομένα. Όταν αφαιρέσουμε την τιμή 426 γίνεται περίπου 367,88 λεπτά ($\approx 6{,}5$ λεπτά μικρότερη).

Διάμεσος: Η διάμεσος είναι $\delta = 369$. Όταν αφαιρέσουμε την τιμή 426 γίνεται 368 λεπτά, δηλαδή 1 λεπτό μικρότερη από την αρχική.

Εύρος: Το εύρος είναι $R = 65$. Όταν αφαιρέσουμε την τιμή 426 γίνεται 14 λεπτά, δηλαδή 51 λεπτά μικρότερο από το αρχικό.

Ενδοτεταρτημοριακό εύρος: Το ενδοτεταρτημοριακό εύρος είναι $Q = 8$. Όταν αφαιρέσουμε την τιμή 426 γίνεται 5,5 λεπτά, δηλαδή 2,5 λεπτά μικρότερο από το αρχικό.

Τυπική απόκλιση: Η τυπική απόκλιση είναι $s \approx 18{,}66$. Όταν αφαιρέσουμε την τιμή 426 γίνεται περίπου $4{,}04$, δηλαδή περίπου 14,6 λεπτά μικρότερη από την αρχική.

Από τα παραπάνω συμπεραίνουμε ότι η ακραία τιμή 426 επηρεάζει σε μεγαλύτερο βαθμό τη μέση τιμή, το εύρος και την τυπική απόκλιση των χρόνων απ'ότι τη διάμεσο και το ενδοτεταρτημοριακό εύρος.
Καταλληλότερα μέτρα θέσης και μεταβλητότητας για να περιγράψουν τα δεδομένα είναι η διάμεσος (369 λεπτά) και το ενδοτεταρτημοριακό εύρος ($Q = 8$ λεπτά)

Προσπαθήστε να λύσετε τις ασκήσεις 1, 6 και 9

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Να χαρακτηρίσετε ως αληθή ή ψευδή καθεμία από τις επόμενες προτάσεις:

α. Οι ακραίες τιμές έχουν μικρή επίδραση στη διάμεσο και στο ενδοτεταρτημοριακό εύρος.

β. Οι ακραίες τιμές επηρεάζουν σημαντικά τη μέση τιμή και τη διασπορά.

γ. Αν παραλείψουμε την ελάχιστη τιμή ενός συνόλου δεδομένων, το εύρος των δεδομένων θα αυξηθεί.

δ. Όταν υπάρχουν ακραίες τιμές στα δεδομένα η διάμεσος είναι καταλληλότερη από την μέση τιμή για την περιγραφή τους.

2 Ποια από τις ομάδες Α και Β θεωρείτε ότι έχει μικρότερη διασπορά και γιατί;

Α: 2, 4, 4, 5, 4   Β: 2, 4, 4, 5, 4, 39

3 Μετρήσαμε το ύψος των αγοριών μιας τάξης και βρήκαμε ότι η διάμεσος είναι 174cm και η μέση τιμή είναι 178cm. Που νομίζετε ότι οφείλετε αυτή η διαφορά;

Ασκήσεις

4 Οι χρόνοι, σε ώρες, χρήσης των κινητών τηλεφώνων 5 ατόμων είναι:

1, 1, 2, 2, 3, 7

α. Να εξετάσετε αν η τιμή 7 είναι ακραία τιμή.

β. Πώς θα μεταβληθεί η μέση τιμή και η διάμεσος όταν απομακρύνουμε την τιμή 7;

5 Οι θερμοκρασίες 7 ημερών σε μια πόλη ήταν:

$-6, -3, -3, -3, -2, -1$

α. Να βρείτε τη διασπορά και το εύρος των θερμοκρασιών.

β. Να σχολιάσετε πώς θα αλλάξει η τιμή της διασποράς και του εύρους αν απομακρύνουμε την τιμή -6 και να επαληθεύσετε τις εκτιμήσεις σας.

6 Τα ετήσια εισοδήματα (σε χιλιάδες ευρώ) πέντε οικογενειών είναι:

12, 15, 16, 50, 14

Ποια μέτρα θέσης και μεταβλητότητας θεωρείτε κατάλληλα για να περιγράψουν τα παραπάνω δεδομένα;

7 Το ύψος της βροχής (σε mm) 8 ημερών σε μια πόλη ήταν

0, 3, 3, 4, 4, 4, 5, 9

α. Να βρείτε το εύρος, το ενδοτεταρτημοριακό εύρος και τις ακραίες τιμές των παρατηρήσεων.

β. Να εκτιμήσετε πώς θα μεταβληθούν το εύρος και το ενδοτεταρτημοριακό εύρος των παρατηρήσεων που απομένουν αν απομακρύνουμε:

- την ακραία τιμή στα αριστερά της κατανομής των τιμών
- Την ακραία τιμή στα δεξιά της κατανομής των τιμών

και να επαληθεύσετε τις εκτιμήσεις σας.

8 Χρησιμοποιώντας τα δεδομένα της εφαρμογής 2 να εξετάσετε πώς θα μεταβληθεί η μέση απόλυτη απόκλιση των δεδομένων όταν απομακρύνουμε την ακραία τιμή.

9 Παρακάτω φαίνεται το βάρος (σε κιλά) 8 ενήλικων ανθρώπων:

Βάρος: 73, 67, 85, 8, 59, 142, 91, 87

α. Υπάρχουν ακραίες τιμές στα δεδομένα; Ποιες από αυτές θεωρείτε ότι πρέπει να απομακρυνθούν;

β. Να αφαιρέσετε την αφύσικη ακραία τιμή από τις παραπάνω παρατηρήσεις και να υπολογίσετε:

- τη μέση τιμή και τη διάμεσο των βαρών,
- το εύρος, το ενδοτεταρτημοριακό εύρος και την τυπική απόκλιση των βαρών.

γ. Ποιο μέτρο θέσης και ποιο μέτρο μεταβλητότητας θεωρείτε κατάλληλο για να περιγράψει τα δεδομένα χωρίς την αφύσικη ακραία τιμή.

10 Σε ορισμένα αγωνίσματα κατάβασης στο σκι, οι αθλητές λαμβάνουν τον μέσο των χρόνων τους από τρεις προσπάθειες. Θα προτιμούσατε ο μέσος να είναι η μέση τιμή ή η διάμεσος, αν συνήθως έχετε:

α. έναν πολύ κακό χρόνο και δύο μέτριους χρόνους;

β. έναν πολύ καλό χρόνο και δύο μέτριους χρόνους;

γ. δύο καλούς χρόνους και έναν μέτριο χρόνο;

δ. τρεις διαφορετικούς χρόνους, που απέχουν περίπου ίσα χρονικά διαστήματα;

2.5 Ανακεφαλαίωση

Στην ενότητα αυτή:
Περιλαμβάνονται έργα για περαιτέρω αναζητήσεις και διερευνήσεις των μαθητών/μαθητριών, με στόχο τη διεύρυνση της μαθησιακής διαδικασίας που οδηγεί στα Προσδοκώμενα Μαθησιακά Αποτελέσματα του κεφαλαίου.

1 Έστω ένα σύνολο Α αριθμητικών δεδομένων: $x_1, x_2, ..., x_\nu$.
Να χαρακτηρίσετε ως αληθή ή ψευδή καθεμία από τις επόμενες προτάσεις:

α. Απόκλιση της τιμής x_1 από την μέση τιμή $\bar{x}$ ονομάζουμε τη διαφορά $x_1 - \bar{x}$.

β. Τετραγωνική απόκλιση της τιμής x_1 από τη μέση τιμή $\bar{x}$ ονομάζουμε το τετράγωνο $(x_1 - \bar{x})^2$.

γ. Η διασπορά είναι η μέση τιμή των απόλυτων αποκλίσεων των αριθμών από τη μέση τιμή τους.

δ. Η τυπική απόκλιση παίρνει μόνο θετικές τιμές.

ε. Η διασπορά είναι μηδέν, όταν οι αριθμοί είναι ίσοι μεταξύ τους.

στ. Αν $s = -5\bar{x}$ τότε $CV = -5$.

ζ. Αν η διασπορά των αριθμών είναι $s^2 = 4$, τότε η τυπική απόκλιση είναι $s = -2$.

η. Η διάμεσος δεν επηρεάζεται από τις απόμακρες τιμές όσο η μέση τιμή.

θ. Αν υπάρχουν ακραίες τιμές, η διάμεσος περιγράφει καλύτερα τα δεδομένα από τη μέση τιμή.

ι. Το ενδοτεταρτημοριακό εύρος επηρεάζεται από τις ακραίες τιμές σε μεγαλύτερο βαθμό απ' ότι η διασπορά.

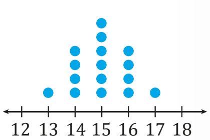
ια. Η διασπορά και η τυπική απόκλιση δεν επηρεάζονται σημαντικά από τις ακραίες τιμές.

ιβ. Αν $CV \leq 0{,}1$ τότε το Α είναι ομοιογενές.

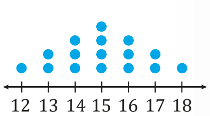
ιγ. Αν πολλαπλασιάσουμε τα στοιχεία του Α με τον αριθμό -2 η διασπορά θα τετραπλασιαστεί.

ιδ. Αν προσθέσουμε στα στοιχεία του Α τον αριθμό 3, η διάμεσος θα παραμείνει ίδια.

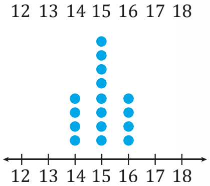
Α



Β



Γ



2 Τα διπλανά σημειογράμματα δείχνουν τις ηλικίες των παιδιών τριών διαφορετικών σχολικών ομάδων ποδοσφαίρου. Χωρίς να κάνετε υπολογισμούς:

α. Να συγκρίνετε τη μέση τιμή και τη διάμεσο των ηλικιών των παιδιών των τριών ομάδων.

β. Να συγκρίνετε την τυπική απόκλιση των ηλικιών των παιδιών των τριών ομάδων.

3 Μια εταιρεία έχει 7 υποκαταστήματα πώλησης των προϊόντων της σε περιοχές ισοδύναμες μεταξύ τους από πλευράς ευκαιριών και δυνατοτήτων. Η εταιρεία προσέλαβε νέο γενικό διευθυντή, ο οποίος την αναδιοργάνωσε και εφάρμοσε νέες επιστημονικές μεθόδους οργάνωσης, διοίκησης και πωλήσεων. Μετά τη συμπλήρωση μίας οικονομικής χρήσης ο νέος διευθυντής αποφασίζει να μελετήσει την πρόοδο της εταιρείας. Έτσι θα συγκρίνει τη μέση τιμή και την τυπική απόκλιση της νέας περιόδου με της περσινής περιόδου. Οι πωλήσεις (σε εκατομμύρια ευρώ) των 7 υποκαταστημάτων της φετινής χρήσης ήταν:

$$10,\ 12,\ 11,\ 11,\ 12,\ 13,\ 12$$

Την περσινή χρήση η μέση τιμή των πωλήσεων ήταν 11,2 και η τυπική απόκλιση 1 εκατομμύριο ευρώ. Να εξετασθεί εάν η επιχείρηση βελτίωσε την οικονομική της θέση, συγκρίνοντας τις μέσες τιμές και τις τυπικές αποκλίσεις των δύο χρήσεων.

4 Ο αριθμός των ψήφων που πήραν 11 υποψήφιοι για την ανάδειξη ενός πενταμελούς συμβουλίου είναι

$$50,\ 42,\ 100,\ 230,\ 82,\ 53,\ 109,\ 95,\ 121,\ 97,\ 31$$

α. Να υπολογίσετε τη μέση τιμή και τη διάμεσο των ψήφων και να ερμηνεύσετε τις τιμές τους.

β. Να εξετάσετε αν η τιμή 230 είναι ακραία τιμή.

Στη συνέχεια αφαιρούμε τη τιμή 230 από το δείγμα.

γ. Να υπολογίσετε τη μέση τιμή και τη διάμεσο του νέου συνόλου δεδομένων και να συγκρίνετε τις νέες τιμές με τις αρχικές των μέτρων του ερωτήματος (α). Τι παρατηρείτε; Πώς επηρεάζουν οι ακραίες τιμές, τα μέτρα θέσης;

δ. Να υπολογίσετε και να συγκρίνετε τη διασπορά καθώς και το ενδοτεταρτημοριακό εύρος στα δύο σύνολα δεδομένων (με και χωρίς τη τιμή 230). Πώς επηρεάζουν οι ακραίες τιμές τα μέτρα μεταβλητότητας;

Διατροφή Α	Διατροφή Β
3	4
-1	3
5	6
4	5
3	4
3	4
1	2
2	3
4	5

5 Για να ελέγξουμε την αποτελεσματικότητα ως προς την απώλεια βάρους δύο διαφορετικών τύπων διατροφής Α και Β, πήραμε ένα τυχαίο δείγμα 18 ανθρώπων και τους χωρίσαμε τυχαία σε δύο ομάδες. Την μία ομάδα την υποβάλλαμε στη διατροφή τύπου Α και την άλλη στη διατροφή τύπου Β. Η διάρκεια της έρευνας ήταν 2 μήνες. Στον διπλανό πίνακα παρουσιάζονται τα αποτελέσματα των μετρήσεων στο τέλος της έρευνας. Κάθε αριθμός εκφράζει την απώλεια βάρους (σε kg) του συγκεκριμένου ανθρώπου, π.χ. ο αριθμός 1 εκφράζει απώλεια ενός κιλού.

α. Πώς ερμηνεύεται την τιμή -1 του πίνακα;

β. Να βρεθεί η μέση τιμή και η διάμεσος κάθε ομάδας και να σχολιαστούν τα αποτελέσματα.

γ. Να υπολογίσετε τη διασπορά, την τυπική απόκλιση και το ενδοτεταρτημοριακό εύρος των τιμών κάθε ομάδας. Σε ποια από τις δύο ομάδες υπάρχει μεγαλύτερη μεταβλητότητα ως προς την απώλεια κιλών;

δ. Ποια από τις δύο διατροφές στο δείγμα, κατά μέσο όρο, έχει τη μεγαλύτερη απώλεια βάρους;

6 Η Μαρία πήρε βαθμό στο διαγώνισμα τετράμηνου στα Μαθηματικά 17 και παρατήρησε ότι όσοι συμμαθητές της είχαν μεγαλύτερη βαθμολογία από αυτήν άλλοι τόσοι είχαν μικρότερη βαθμολογία (Όλοι οι βαθμοί είναι ακέραιοι αριθμοί). Στο επόμενο μάθημα ο καθηγητής έδωσε στον Γιάννη, που απουσίαζε στο προηγούμενο μάθημα, τον βαθμό του που ήταν 16 και η Μαρία παρατήρησε ότι η μέση βαθμολογία της τάξης της δεν άλλαξε. Ακόμα υπολόγισε ότι οι βαθμολογίες της τάξης της είχαν συντελεστή μεταβλητότητας 12,5%.

α. Να βρείτε τη διάμεσο και τη μέση τιμή της βαθμολογίας της τάξης.

β. Να βρείτε την διασπορά της βαθμολογίας της τάξης.

γ. Επειδή στο διαγώνισμα μια ερώτηση πολλαπλής επιλογής ήταν πολύ δύσκολη και δεν την απάντησε κανένας μαθητής, ο καθηγητής είπε ότι θα πρόσθετε στη βαθμολογία κάθε μαθητή μισή μονάδα.

γ_1. Να υπολογίσετε τη νέα μέση τιμή, διάμεσο, τυπική απόκλιση και διασπορά των βαθμολογιών.

γ_2. Να εξετάστε αν οι νέες βαθμολογίες έχουν μεγαλύτερη ομοιογένεια.

7 Σε ένα εργοστάσιο η μέση τιμή των μισθών των υπαλλήλων στο τμήμα Α είναι 950 € και η τυπική απόκλιση 40 € ενώ στο τμήμα Β είναι 1050 € και 50 €, αντίστοιχα. Αν ο κάθε εργαζόμενος στο τμήμα Α πάρει αύξηση 60 € και στο τμήμα Β 10% να εξετάσετε σε ποιο τμήμα οι νέοι μισθοί έχουν μεγαλύτερη ομοιογένεια.

8 Επιλέξαμε τυχαία δύο δείγματα από αμοιβές υπαλλήλων δύο εταιρειών (ένα για κάθε εταιρεία), μίας αμερικάνικης και μίας ευρωπαϊκής. Παρακάτω δίνονται οι δειγματικές μέσες τιμές και οι δειγματικές τυπικές αποκλίσεις των δύο δειγμάτων.

Αμερικάνικη εταιρεία: $\bar{x}_A = 1100\$$ και $s_A = 120\$$

Ευρωπαϊκή εταιρεία: $\bar{x}_E = 1050€$ και $s_E = 70€$

α. Να εξηγήσετε γιατί η τυπική απόκλιση δεν είναι κατάλληλο μέτρο για τη σύγκριση της μεταβλητότητας των δύο εταιρειών. Ποιο μέτρο μεταβλητότητας θεωρείται ότι είναι κατάλληλο για τη σύγκριση;

β. Ποιο από τα δύο δείγματα παρουσιάζει μεγαλύτερη ομοιογένεια ως προς το ύψος των μισθών;

γ. Οι δύο εταιρείες ανακοίνωσαν αύξηση των μισθών των υπαλλήλων τους. Η αμερικάνικη εταιρεία θα κάνει αύξηση 100$ στο μισθό κάθε υπαλλήλου ενώ η ευρωπαϊκή θα αυξήσει κατά 10% τους μισθούς. Πώς θα επηρεαστούν οι τυπικές αποκλίσεις και οι συντελεστές μεταβλητότητας των μισθών των δύο εταιρειών μετά τις αυξήσεις;

9 Υποθέστε ότι ζητάτε πληροφορίες από έναν μεσίτη σχετικά με τις τιμές των σπιτιών σε δύο διαφορετικά, αλλά συγκρίσιμα, προάστια μιας πόλης. Ας τα ονομάσουμε προάστιο Α και προάστιο Β. Ο μεσίτης σας παρέχει τις εξής πληροφορίες που προέρχονται από ένα τυχαίο δείγμα 40 σπιτιών σε κάθε προάστιο:

- Τη μέση τιμή των τιμών των σπιτιών σε κάθε ένα από τα δύο προάστια.
- Την περίληψη των πέντε αριθμών των τιμών των σπιτιών σε κάθε προάστιο.

Όλες οι πληροφορίες που παρέχονται από τον μεσίτη δίνονται στους διπλανούς δύο πίνακες. Σημειώστε ότι όλες οι τιμές των σπιτιών είναι σε χιλιάδες ευρώ.

Προάστιο	Α	Β
$\bar{x}$	133,4	130,6
S	19,2	18,5

	min	Q_1	δ	Q_3	max
Α	87	115	120	131	578
Β	101	120	128	137	202

α. Εξηγήστε με ποιον τρόπο οι πληροφορίες που σας παρέχει ο μεσίτης μπορούν να σας βοηθήσουν να πάρετε απόφαση σχετικά με το προάστιο όπου θα ψάξετε για να αγοράσετε σπίτι.

β. Εξηγήστε γιατί θα μπορούσατε να παραπλανηθείτε αν κοιτάζατε μόνο τη μέση τιμή και την τυπική απόκλιση των τιμών των σπιτιών σε κάθε προάστιο έχοντας ως στόχο να ξοδέψετε όσο το δυνατό λιγότερα χρήματα για την αγορά σπιτιού.

ΚΕΦΑΛΑΙΟ 3ο

Σχέσεις Εξάρτησης μεταξύ δύο μεταβλητών

Σύντομη γνωριμία

Η χρήση θηκογραμμάτων, τα οποία μάθαμε πώς κατασκευάζονται στο πρώτο κεφάλαιο, αποδεικνύεται ιδιαιτέρως χρήσιμη στην ανάδειξη σχέσεων εξάρτησης μεταξύ κατηγορικών και ποσοτικών μεταβλητών. Μας παρέχουν μια εικόνα της κατανομής των τιμών της ποσοτικής μεταβλητής στις διάφορες κατηγορίες της κατηγορικής μεταβλητής, επιτρέποντας μας με αυτό τον τρόπο να εξετάσουμε τυχόν ομοιότητες και διαφορές μεταξύ των κατηγοριών.

Σε αυτό το κεφάλαιο, θα αξιοποιήσουμε τα θηκογράμματα για να συγκρίνουμε τις τιμές μιας ποσοτικής μεταβλητής στις διάφορες κατηγορίες μιας κατηγορικής μεταβλητής. Θα αναδείξουμε, επίσης, ένα συνηθισμένο λάθος που συνίσταται στην εσφαλμένη ερμηνεία της σχέσης εξάρτησης ως αιτιώδους σχέσης. Μέσω αυτής της ανάλυσης, θα καταδείξουμε τη σημασία της σωστής ερμηνείας των σχέσεων μεταξύ των μεταβλητών και της ανάγνωσης των αποτελεσμάτων με προσοχή προκειμένου να αποφευχθούν λανθασμένα συμπεράσματα.

Περιεχόμενα

Jerzy Neyman
(1894 – 1981)

Πολωνός μαθηματικός και στατιστικολόγος.
Δημοσίευσε πολλά βιβλία που ασχολούνται με πειράματα και στατιστικές, και επινόησε τον τρόπο με τον οποίο ο Οργανισμός Τροφίμων και Φαρμάκων δοκιμάζει τα φάρμακα σήμερα.
Το 1966 του απονεμήθηκε το Μετάλλιο Γκάι της Βασιλικής Εταιρείας Στατιστικής και το 1969 το Εθνικό Μετάλλιο Επιστημών των ΗΠΑ.

Ronald Aylmer Fisher
(1890 – 1962)

Βρετανός μαθηματικός, στατιστικολόγος, βιολόγος, γενετιστής και ακαδημαϊκός.
Για το έργο του στη στατιστική, χαρακτηρίστηκε ως μια ιδιοφυΐα, η οποία σχεδόν μόνη της δημιούργησε τα θεμέλια για τη σύγχρονη στατιστική επιστήμη.
Το 1925 δημοσίευσε το «Statistical Methods for Research Workers», ένα από τα σημαντικότερα βιβλία του 20ου αιώνα σχετικά με τις στατιστικές μεθόδους.

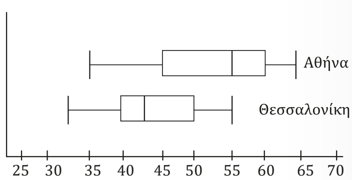
3.1 Σύγκριση μέτρων θέσης και μεταβλητότητας ποσοτικού χαρακτηριστικού στις σταθμές κατηγορικού χαρακτηριστικού με τη βοήθεια θηκογραμμάτων

Περιέχονται:
- Ερμηνεία πολλαπλών θηκογραμμάτων

Λέξεις κλειδιά:
Σύγκριση θηκογραμμάτων

Στην ενότητα αυτή θα μάθουμε:
Να ερμηνεύουμε τα πολλαπλά θηκογράμματα για να κάνουμε συγκρίσεις και να εξάγουμε συμπεράσματα σχετικά με τα μέτρα θέσης και μεταβλητότητας που έχουν οι τιμές ενός ποσοτικού χαρακτηριστικού σε κάθε στάθμη του κατηγορικού χαρακτηριστικού του υπό μελέτη πληθυσμού.

Δραστηριότητα
Με βάση τον κώδικα οδικής κυκλοφορίας (ΚΟΚ), το ανώτατο επιτρεπόμενο όριο ταχύτητας για τα αυτοκίνητα, σε κατοικημένες περιοχές, είναι 50 km/h. Για να ελέγξουμε την οδική συμπεριφορά των κατοίκων της Αθήνας και της Θεσσαλονίκης, συλλέξαμε αντιπροσωπευτικά δείγματα από κάθε πόλη καταγράφοντας τις ταχύτητες των αυτοκινήτων. Από την ανάλυση των αποτελεσμάτων προέκυψαν τα θηκογράμματα της διπλανής εικόνας.



α. Σε ποια από τα δύο δείγματα είναι πιο πιθανό να παρατηρήσουμε οδηγό που υπερβαίνει το όριο ταχύτητας;

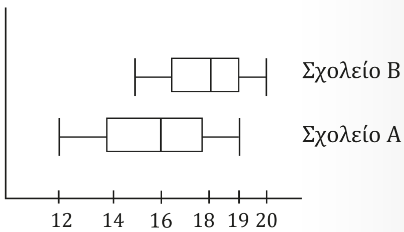
β. Σε ποιο δείγμα οι ταχύτητες παρουσιάζουν μεγαλύτερη μεταβλητότητα;

γ. Σε ποιο δείγμα θεωρείτε ότι οι κάτοικοι έχουν καλύτερη οδική συμπεριφορά;

Ερμηνεία πολλαπλών θηκογραμμάτων

Η χρήση των πολλαπλών θηκογραμμάτων είναι ένα εξαιρετικό εργαλείο για την οπτικοποίηση της κατανομής των τιμών μιας ποσοτικής μεταβλητής στις στάθμες μιας κατηγορικής. Για παράδειγμα, αν έχουμε τους βαθμούς των μαθητών μιας πόλης και θέλουμε να συγκρίνουμε την επίδοση τους στα διάφορα σχολεία, τότε με τα πολλαπλά θηκογράμματα εύκολα και γρήγορα μπορούμε να εντοπίσουμε και να αναδείξουμε τις διαφορές στις επιδόσεις των μαθητών μεταξύ των σχολείων.

Παράδειγμα
Τα θηκογράμματα της διπλανής εικόνας έχουν προκύψει από δείγμα βαθμών απόλυσης μαθητών/τριών δύο σχολείων της πόλης των Ιωαννίνων.
Με μια πρώτη ματιά, είναι φανερό ότι οι μαθητές/τριες του σχολείου Β έχουν πετύχει υψηλότερες βαθμολογίες από αυτούς/τες του σχολείου Α. Οι βαθμοί τους παρουσιάζουν μεγάλη συγκέντρωση στις υψηλές βαθμολογίες (17 έως 19), σε αντίθεση με τους μαθητές/τριες του σχολείου Α, των οποίων οι βαθμοί παρουσιάζουν μεγαλύτερη μεταβλητότητα από αυτούς του σχολείου Β, με το «κέντρο» τους να εντοπίζεται στους βαθμούς 14 έως 18. Φαίνεται πως στο δείγμα μας υπάρχει σχέση ανάμεσα στο σχολείο φοίτησης και στις βαθμολογίες των μαθητών/τριών.

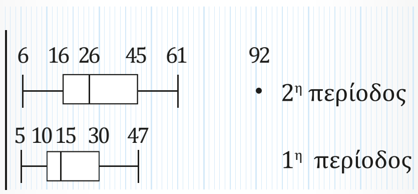


Προσπαθήστε να λύσετε τις ασκήσεις 1 και 2

Ας δούμε όμως μία αναλυτικότερη περιγραφή ενός πολλαπλού θηκογράμματος με την επόμενη εφαρμογή.

Εφαρμογή

Τα διπλανά θηκογράμματα προέκυψαν από τα δεδομένα ενός αεροδρομίου για τις καθυστερήσεις (σε λεπτά) των αεροπορικών πτήσεων (εξαιρουμένων των αναβολών πτήσεως) κατά την περίοδο πριν το καλοκαίρι (1η περίοδος) και κατά την καλοκαιρινή περίοδο (2η περίοδος) ενός έτους.



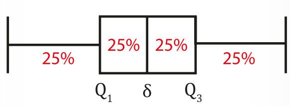
α. Να συγκρίνετε τις διαμέσους των δύο περιόδων. Να ερμηνεύσετε τη διαφορά.

β. Να βρείτε σε ποια από τις δύο περιόδους η διαφορά της κάτω από την άνω οριακή τιμή και το ενδοτεταρτημοριακό εύρος του χρόνου καθυστέρησης είναι μεγαλύτερα. Να ερμηνεύσετε τις διαφορές. Επιπλέον να συγκρίνετε τα θηκογράμματα, ως προς το σχήμα.

γ. Μπορούμε να ισχυριστούμε ότι ο χρόνος καθυστέρησης αυξήθηκε τη 2η περίοδο σε σύγκριση με αυτόν της 1ης περιόδου;

Θυμίζουμε ότι:

Οι αριθμοί Q_1, δ, Q_3 χωρίζουν κάθε θηκόγραμμα σε τέσσερα μέρη. Σε καθένα από αυτά ανήκει περίπου το 25% των παρατηρήσεων, όπως φαίνεται στο σχήμα.



Λύση

α. Η διάμεσος για την 1η περίοδο είναι 15 λεπτά, ενώ για τη 2η περίοδο είναι 26 λεπτά, δηλαδή 11 λεπτά μεγαλύτερη από αυτή της 1ης περιόδου. Άρα την 1η περίοδο οι μισές περίπου πτήσεις καθυστέρησαν το πολύ 15 λεπτά και αντίστοιχα τη 2η περίοδο οι μισές περίπου πτήσεις καθυστέρησαν το πολύ 26 λεπτά.

β. Για την 1η περίοδο έχουμε:

$$U - L = 47 - 5 = 42 \text{ λεπτά}$$
$$Q = 30 - 10 = 20 \text{ λεπτά}$$

Για την 2η περίοδο έχουμε:

$$U' - L' = 61 - 6 = 55 \text{ λεπτά}$$
$$Q' = 45 - 16 = 29 \text{ λεπτά}$$

Επομένως το εύρος και το ενδοτεταρτημοριακό εύρος είναι μεγαλύτερα τη 2η περίοδο.

Ερμηνεία: Η διαφορά U – L της 1ης περιόδου είναι 42 λεπτά και της 2ης περιόδου είναι 55 λεπτά. Επομένως την 2η περίοδο το σύνολο των καθυστερήσεων (πλην της ακραίας τιμής 92) παρουσιάζει μεγαλύτερη μεταβλητότητα από το αντίστοιχο της 1ης περιόδου. Επίσης, το ενδοτεταρτημοριακό εύρος της 1ης περιόδου είναι 20 λεπτά, ενώ της 2ης περιόδου είναι 9 λεπτά μεγαλύτερο. Επομένως το διάστημα $[Q_1, Q_3]$ είναι πιο πλατύ τη 2η περίοδο. Αυτό σημαίνει ότι το «μεσαίο» 50% των χρόνων καθυστέρησης της 2ης περιόδου παρουσιάζει μεγαλύτερη μεταβλητότητα από το αντίστοιχο της 1ης περιόδου. Τέλος, και τα δύο θηκογράμματα παρουσιάζουν θετική ασυμμετρία.

γ. Και οι πέντε χαρακτηριστικοί αριθμοί του θηκογράμματος της 2ης περιόδου βρίσκονται δεξιότερα των αντίστοιχων αριθμών της 1ης περιόδου. Μπορούμε, λοιπόν, να συμπεράνουμε ότι ο χρόνος καθυστέρησης αυξήθηκε κατά τη διάρκεια της 2ης περιόδου.

Σύγκριση θηκογραμμάτων

Προσπαθήστε να λύσετε τις ασκήσεις 3, 4, 5 και 6

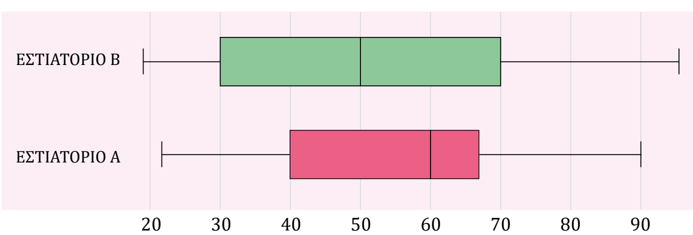
Γενικά, για να συγκρίνουμε θηκογράμματα, ακολουθούμε την παρακάτω διαδικασία.

- Συγκρίνουμε τις διαμέσους των θηκογραμμάτων.
- Συγκρίνουμε τα ενδοτεταρτημοριακά εύρη καθώς και τις διαφορές μεταξύ άνω και κάτω οριακών τιμών.
- Συγκρίνουμε τα θηκογράμματα ως προς το σχήμα (συμμετρία-ασυμμετρία).
- Ελέγχουμε αν υπάρχουν ακραίες τιμές.

Ασκήσεις και Προβλήματα

Ερωτήσεις

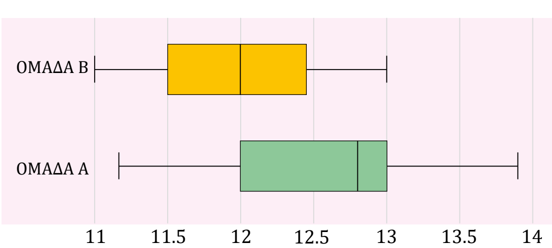
1 Καταγράψαμε τον αριθμό των πελατών ανά ημέρα και για μία εβδομάδα σε δύο γειτονικά εστιατόρια Α και Β και κατασκευάσαμε τα παρακάτω θηκογράμματα.



α. Σε ποιο από τα δύο εστιατόρια η διάμεσος του αριθμού των πελατών είναι μεγαλύτερη;

β. Σε ποιο εστιατόριο ο αριθμός των πελατών παρουσιάζει μεγαλύτερη μεταβλητότητα;

2 Στο παρακάτω πολλαπλό θηκόγραμμα παρουσιάζονται οι καλύτερες επιδόσεις δύο ομάδων αθλητριών σε έναν αγώνα δρόμου των 100m.



Να συμπληρώσετε τα κενά στις παρακάτω προτάσεις.

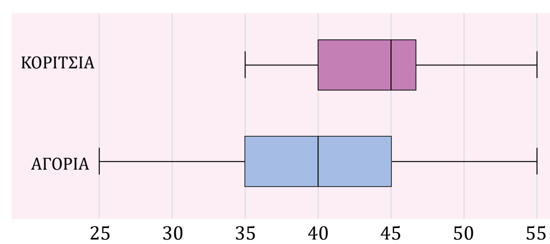
α. Γενικά, η ομάδα …… έχει πετύχει μικρότερους χρόνους ως προς τη διάμεση τιμή.

β. Οι χρόνοι της ομάδας …… παρουσιάζουν μεγαλύτερη μεταβλητότητα από αυτούς της ομάδας …… .

γ. Η κατανομή των χρόνων της ομάδας …… παρουσιάζει …… ασυμμετρία.

Ασκήσεις

3 Στα παρακάτω θηκογράμματα παρουσιάζεται ο αριθμός των ακόλουθων (followers) σε ένα δημοφιλές κοινωνικό δίκτυο για μία ομάδα αγοριών και μία ομάδα κοριτσιών.



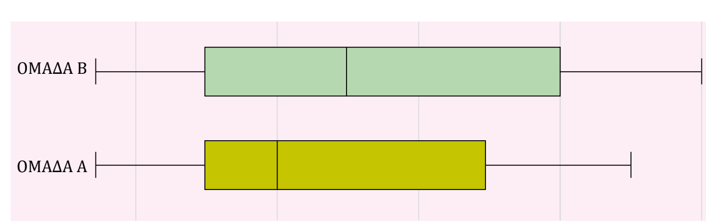
α. Ποια είναι η διαφορά των διαμέσων των δύο ομάδων;

β. Σε ποια ομάδα το εύρος του πλήθους των ακόλουθων είναι μεγαλύτερο;

γ. Ποιο είναι το ενδοτεταρτημοριακό εύρος κάθε ομάδας;

δ. Ποια είναι περίπου τα ποσοστά των παιδιών σε κάθε ομάδα που βρίσκονται δεξιά της τιμής 40;

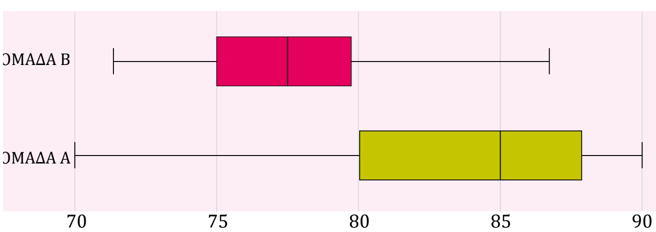
4 Στα παρακάτω θηκογράμματα φαίνονται τα αποτελέσματα κάποιων μετρήσεων για δύο ομάδες δεδομένων. Όλες οι μετρήσεις εκφράζονται με ακέραιους αριθμούς.



α. Ένα άτομο της ομάδας Α έχει τιμή μεγαλύτερη ή ίση από το 75% των τιμών της ομάδας Β. Ποιες είναι οι πιθανές τιμές που αφορούν το άτομο αυτό;

β. Ένα άτομο της ομάδας Β έχει τιμή μεγαλύτερη ή ίση από το 75% των τιμών της ομάδας Α. Ποιες είναι οι πιθανές τιμές που αφορούν το άτομο αυτό;

5 Τα θηκογράμματα του παρακάτω σχήματος προέκυψαν από τα στατιστικά δύο ομάδων μπάσκετ, Α και Β, για την αγωνιστική περίοδο του προηγούμενου έτους και παρουσιάζουν τους πόντους κάθε ομάδας στη λήξη των παιχνιδιών.

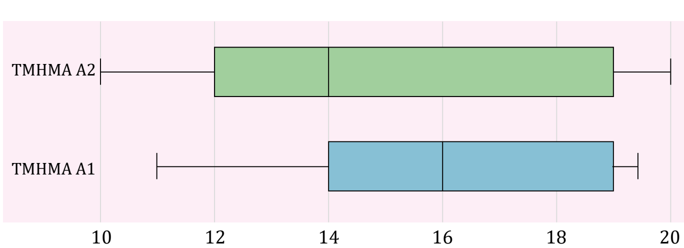


α. Ποια ομάδα παρουσιάζει μεγαλύτερη μεταβλητότητα ως προς του πόντους;

β. Σε ποια από τις δύο ομάδες φαίνεται να είναι πιο συμμετρική γύρω από τη διάμεσο η κατανομή των πόντων;

γ. Μπορούμε να συμπεράνουμε από τα παραπάνω θηκογράμματα ποια από τις δύο ομάδες πήρε καλύτερη θέση στην τελική κατάταξη στο τέλος της αγωνιστικής περιόδου;

Οι βαθμοί απολυτηρίου των μαθητών σε δύο τμήματα Α Λυκείου ενός σχολείου παρουσιάζονται στα παρακάτω θηκογράμματα.



α. Σε ποιο τμήμα παρατηρήθηκε ο υψηλότερος βαθμός και σε ποιο ο χαμηλότερος;

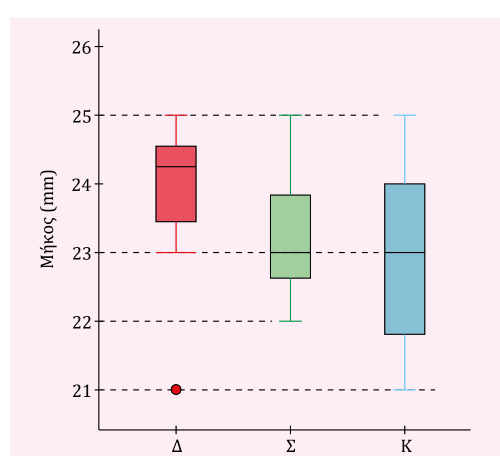
β. Ποιο τμήμα παρουσιάζει μεγαλύτερη μεταβλητότητα ως προς τους βαθμούς;

γ. Το σχολείο αποφάσισε να βραβεύσει τους μαθητές που έβγαλαν βαθμό τουλάχιστον 19. Αν το τμήμα $Α_1$ έχει 24 μαθητές και το τμήμα $Α_2$ έχει 28 μαθητές, πόσοι τουλάχιστον μαθητές θα βραβευθούν;

7 Το πτηνό κούκος είναι γνωστό λόγω του παρασιτισμού του στις φωλιές άλλων πτηνών. Ο θηλυκός κούκος γεννά τα αυγά του στις φωλιές άλλων πτηνών τα οποία ανατρέφουν τους νεοσσούς κούκους εις βάρος των δικών τους νεοσσών.

Συλλέχθηκε δείγμα από αυγά κούκου που βρέθηκαν σε φωλιές τριών διαφορετικών ειδών πουλιών: της Δενδρότσιχλας (Δ), του Σπουργιτιού (Σ) και της Κίσσας (Κ). Για κάθε αυγό μετρήθηκε το μήκος του σε χιλιοστά (mm). Χρησιμοποιήστε τα παρακάτω θηκογράμματα για να συγκρίνετε τα μήκη των αυγών του κούκου που βρέθηκαν στις φωλιές των τριών ειδών πουλιών.

Από τα δεδομένα του δείγματος, φαίνεται να υπάρχει κάποια σχέση ανάμεσα στο είδος του πουλιού και το μήκος των αυγών του κούκου που βρέθηκαν στις φωλιές του; Δικαιολογήστε την απάντησή σας.



3.2 Αιτιότητα

Περιέχονται:
- Η έννοια της αιτιότητας
- Η αιτιότητα στη Στατιστική

Λέξεις κλειδιά:
Αίτιο - Αποτέλεσμα
Αιτιότητα

Στην ενότητα αυτή:
Θα μάθουμε να ανακαλύπτουμε και να εξηγούμε με παραδείγματα ότι ένα ποσοτικό και ένα κατηγορικό χαρακτηριστικό δε διέπονται απαραίτητα από μια σχέση αιτίου-αιτιατού.

Η έννοια της αιτιότητας

Η έννοια της αιτιότητας αναφέρεται στην αμοιβαία σχέση μεταξύ δύο καταστάσεων, όπου η μία είναι η αιτία και η άλλη το αποτέλεσμα (αιτιατό). Για παράδειγμα:
- Θερμαίνουμε το νερό και αυτό βράζει.
- Ασκούμε δύναμη σε ένα αντικείμενο και αυτό κινείται.
- Ανάβουμε τα φώτα σε ένα σκοτεινό δωμάτιο και αυτό φωτίζεται.
- Πιέζουμε με το πόδι μας το φρένο του αυτοκινήτου και η ταχύτητα του αυτοκινήτου μειώνεται.

Σε όλα τα παραπάνω παραδείγματα υπάρχει ένα αίτιο που προκαλεί ένα αποτέλεσμα. Η θέρμανση του νερού (αίτιο) αυξάνει την κινητική ενέργεια των μορίων του νερού και το νερό βράζει (αποτέλεσμα). Η εφαρμογή δύναμης (αίτιο) στο αντικείμενο του μεταφέρει ενέργεια και το κάνει να κινείται (αποτέλεσμα). Το άνοιγμα των φώτων (αίτιο) δημιουργεί φως, το οποίο διαχέεται στο δωμάτιο και εξαλείφει το σκοτάδι. Συνέπεια είναι η φωτεινότητα του χώρου (αποτέλεσμα). Η πίεση στο πεντάλ του φρένου (αίτιο) ενεργοποιεί το σύστημα φρένων, περιορίζοντας την κίνηση των τροχών και η ταχύτητα του αυτοκινήτου μειώνεται (αποτέλεσμα).

Η αιτιότητα προϋποθέτει μια σταθερή σχέση μεταξύ αιτίας και αιτιατού, όπου κάθε φορά που συμβαίνει μια συγκεκριμένη αιτία, το αντίστοιχο αποτέλεσμα είναι πάντα το ίδιο.

Η εξέταση της αιτιότητας ως έννοια ξεκινά από τα αρχαία χρόνια. Ο Αριστοτέλης (384-322 π.Χ.) έθεσε τις βάσεις μεταξύ αιτίου και αιτιατού. Για τον Αριστοτέλη, κάθε γεγονός διαθέτει μια αιτία και η αιτία προηγείται πάντα του αποτελέσματος.

Η αιτιότητα στη Στατιστική

Στο πλαίσιο της Στατιστικής, θα δούμε ότι μια σχέση εξάρτησης μεταξύ μιας ποσοτικής και μιας κατηγορικής μεταβλητής δεν συνεπάγεται απαραίτητα ότι οι δύο μεταβλητές συνδέονται με σχέση αιτίας και αποτελέσματος. Οι μεταβλητές μπορούν να επηρεάζονται από διάφορους παράγοντες, χωρίς να υπάρχει κατ' ανάγκη αιτιακή σχέση μεταξύ τους.

Παράδειγμα 1
Ας υποθέσουμε ότι έχουμε δεδομένα που δείχνουν ότι υπάρχει σχέση εξάρτησης μεταξύ της κατηγορικής μεταβλητής «Πωλήσεις παγωτού» (χαμηλές ή υψηλές σύμφωνα με κάποιο κριτήριο) και της ποσοτικής μεταβλητής «Αριθμός ηλιακών εγκαυμάτων». Συγκεκριμένα, όταν οι πωλήσεις παγωτού βρίσκονται σε χαμηλά επίπεδα, ο αριθμός ηλιακών εγκαυμάτων είναι μικρός και όταν οι πωλήσεις παγωτού είναι υψηλές, ο αριθμός ηλιακών εγκαυμάτων είναι μεγάλος.

Είναι προφανές ότι οι πωλήσεις παγωτού δεν μπορούν να επηρεάσουν τον αριθμό των ηλιακών εγκαυμάτων αλλά ούτε το αντίστροφο. Άλλοι παράγοντες, όπως οι καιρικές συνθήκες, μπορεί να επηρεάσουν τη σχέση ανάμεσα στις δύο μεταβλητές.

Παράδειγμα 2

Από δειγματοληπτική έρευνα που πραγματοποιήθηκε σε μια εταιρεία κινητής τηλεφωνίας, προέκυψε ότι υπάρχει σχέση εξάρτησης μεταξύ του τύπου εκπαίδευσης των υπαλλήλων (Δευτεροβάθμια, Τριτοβάθμια, Μεταπτυχιακό) και του αριθμού πωλήσεων συμβολαίων κινητής τηλεφωνίας. Η ποσοτική μεταβλητή είναι ο ‘’Αριθμός πωλήσεων’’, ενώ η κατηγορική μεταβλητή είναι ο ‘’Τύπος Εκπαίδευσης’’ των υπαλλήλων. Η ανάλυση έδειξε ότι οι υπάλληλοι που έχουν μεταπτυχιακό πραγματοποιούν κατά μέσο όρο περισσότερες πωλήσεις από τους υπαλλήλους με δευτεροβάθμια ή τριτοβάθμια εκπαίδευση. Ωστόσο, αυτή η σχέση εξάρτησης δεν σημαίνει απαραίτητα ότι το επίπεδο εκπαίδευσης είναι η αιτία για την αύξηση των πωλήσεων. Άλλοι παράγοντες, όπως η εμπειρία των υπαλλήλων ή η τοποθεσία του καταστήματος, μπορεί να επηρεάζουν τις πωλήσεις, δημιουργώντας εξάρτηση μεταξύ των μεταβλητών χωρίς να υπάρχει αιτιώδης σχέση. Επιπλέον, η παρατηρούμενη σχέση μπορεί να οφείλεται σε σύμπτωση.

Παράδειγμα 3

Από μια έρευνα που εκπονήθηκε για να εξετάσει τη σχέση μεταξύ της εθνικότητας των μαθητών που συμμετέχουν στη Διεθνή Μαθηματική Ολυμπιάδα (IMO) και των επιδόσεών τους σε αυτή, παρατηρήθηκε ότι οι μαθητές από διαφορετικές χώρες παρουσίασαν διαφοροποιήσεις στις επιδόσεις τους. Ωστόσο, δεν μπορούμε να συμπεράνουμε ότι η εθνικότητα είναι η αιτία αυτών των διαφορών. Πολλοί άλλοι παράγοντες, όπως η ποιότητα του εκπαιδευτικού συστήματος της κάθε χώρας, η προετοιμασία των μαθητών και οι διαθέσιμες εκπαιδευτικές ευκαιρίες, μπορεί να εξηγούν τη διαφορά στην απόδοση, και όχι η εθνικότητα καθαυτή.

Από τα παραπάνω συμπεραίνουμε ότι:

Μια σχέση εξάρτησης δεν σημαίνει απαραίτητα και αιτιότητα.
Όταν παρατηρούμε μια σχέση ανάμεσα σε δύο μεταβλητές, μπορεί να συμβαίνει ένα από τα εξής:

- Η μία μεταβλητή να είναι η αιτία των μεταβολών που παρατηρούμε στην άλλη.
- Υπάρχει ένας τρίτος παράγοντας (ή και περισσότεροι) που επηρεάζει και τις δύο μεταβλητές.
- Η σχέση να οφείλεται σε σύμπτωση: Η παρατηρούμενη σχέση στο δείγμα δεν αντικατοπτρίζει πραγματική σχέση στον πληθυσμό, αλλά είναι προϊόν της τύχης που υπάρχει όταν δουλεύουμε με δείγματα.

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Καταγράψαμε τον αριθμό των πελατών που επισκέφτηκαν τον προηγούμενο μήνα δύο καταστήματα παιχνιδιών, Α και Β, της ίδιας επωνυμίας και παρατηρήσαμε ότι το κατάστημα Α δέχεται γενικά περισσότερους πελάτες από το κατάστημα Β.

α. Υπάρχει σχέση εξάρτησης μεταξύ της κατηγορικής μεταβλητής «Κατάστημα» με την ποσοτική μεταβλητή «Αριθμός πελατών»;

β. Συνδέονται απαραίτητα οι δύο μεταβλητές με σχέση αιτίας-αιτιατού;

γ. Να διατυπώσετε μερικούς παράγοντες που μπορεί να επηρεάζουν τον αριθμό των πελατών στα δύο καταστήματα.

2 Καταγράψαμε τον αριθμό των ατυχημάτων με ποδήλατο που συμβαίνουν σε μια πόλη και την κατανάλωση ζεστών ροφημάτων (συχνή ή σπάνια) και παρατηρήσαμε ότι όταν ο αριθμός ατυχημάτων είναι υψηλός, η κατανάλωση ζεστών ροφημάτων είναι σπάνια ενώ όταν αριθμός ατυχημάτων είναι χαμηλός η κατανάλωση ζεστών ροφημάτων είναι συχνή.

α. Υπάρχει σχέση εξάρτησης μεταξύ της κατηγορικής μεταβλητής «Κατανάλωση ζεστών ροφημάτων» με την ποσοτική μεταβλητή «Αριθμός ατυχημάτων με ποδήλατο»;

β. Συνδέονται απαραίτητα οι δύο μεταβλητές με σχέση αιτίας-αιτιατού;

γ. Να διατυπώσετε μερικούς παράγοντες που μπορεί να επηρεάζουν τις δύο μεταβλητές.

3 Ας υποθέσουμε ότι έχουμε δεδομένα που δείχνουν ότι στο άθλημα του δρόμου 100m στον στίβο οι άνθρωποι με σκουρόχρωμα μάτια έχουν καλύτερες επιδόσεις από αυτούς με ανοικτόχρωμα μάτια.

α. Υπάρχει σχέση εξάρτησης μεταξύ του της κατηγορικής μεταβλητής «Χρώμα ματιών» με την ποσοτική μεταβλητή «Επίδοση στα 100m δρόμου»;

β. Μπορεί να θεωρηθεί ότι το χρώμα των ματιών είναι η αιτία για τις διαφορές στις επιδόσεις που παρατηρήσαμε μεταξύ των ανθρώπων με σκουρόχρωμα μάτια και αυτών με ανοιχτόχρωμα μάτια;

γ. Ποιοι παράγοντες μπορεί να επηρεάζουν τη σχέση μεταξύ του χρώματος των ματιών και της επίδοσης στα αθλήματα τρεξίματος στον στίβο;

3.3 Ανακεφαλαίωση

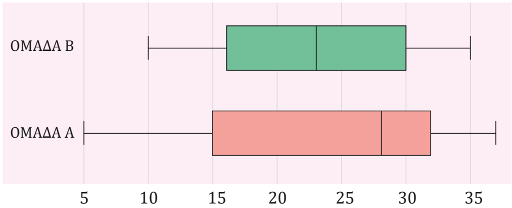
Στην ενότητα αυτή:
Περιλαμβάνονται έργα για περαιτέρω αναζητήσεις και διερευνήσεις των μαθητών/μαθητριών, με στόχο τη διεύρυνση της μαθησιακής διαδικασίας που οδηγεί στα Προσδοκώμενα Μαθησιακά Αποτελέσματα του κεφαλαίου.

Οδηγός επανάληψης-Στατιστική

Σύγκριση θηκογραμμάτων

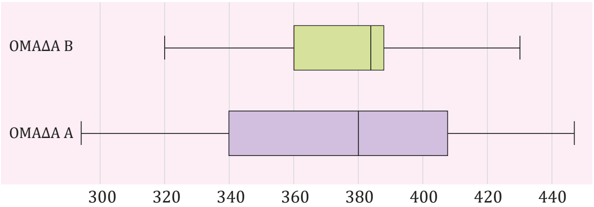
1 Ώρες που παρακολουθούν τηλεόραση τα παιδιά

Στα παρακάτω θηκογράμματα φαίνονται οι χρόνοι παρακολούθησης τηλεόρασης σε δύο ομάδες παιδιών Α και Β, ηλικίας 12 και 13, αντίστοιχα. Να συγκρίνετε τις δύο ομάδες ως προς τους χρόνους παρακολούθησης.



2 Στατιστική και βροχόπτωση

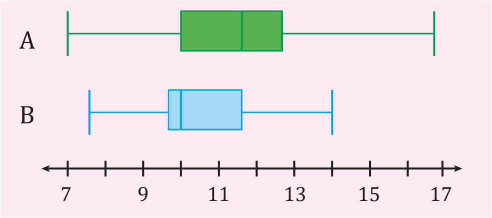
Στα θηκογράμματα του παρακάτω σχήματος παρουσιάζεται το ετήσιο ύψος βροχόπτωσης (σε mm) σε δύο πόλεις για τα έτη 2014-2024.



Χρησιμοποιώντας την πληροφορία που μας δίνουν τα θηκογράμματα να συγκρίνετε τις ετήσιες βροχοπτώσεις στις δύο πόλεις ως προς το μέγεθος και τη μεταβλητότητα τους.

3 Διάρκεια μπαταρίας

Τα παρακάτω θηκογράμματα περιγράφουν τη διάρκεια μπαταρίας (σε ώρες) δύο κινητών τηλεφώνων.

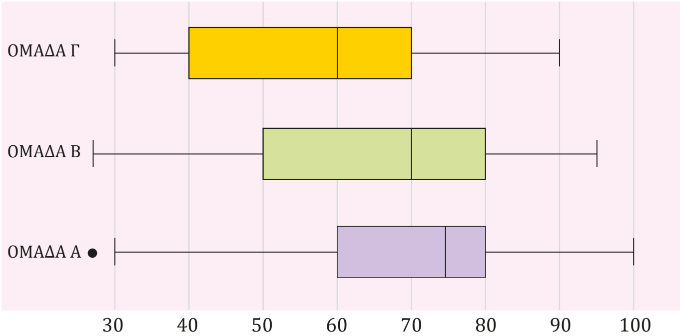


α. Να συγκρίνετε τα δύο σύνολα δεδομένων ως προς τη μεταβλητότητα.
β. Ποια από τις δύο μπαταρίες νομίζετε ότι έχει μεγαλύτερη τυπική απόκλιση;
γ. Αν χρειάζεστε ένα κινητό τηλέφωνο που έχει διάρκεια μπαταρίας τουλάχι-

στον 10 ώρες το μεγαλύτερο μέρος του χρόνου, ποιο από τα δύο κινητά είναι προτιμότερο να αγοράσετε;

4 **Σχολικό πρωτάθλημα**

Οι πόντοι που πέτυχαν τρεις ομάδες μπάσκετ σε 15 αγώνες του σχολικού πρωταθλήματος, φαίνονται στα παρακάτω θηκογράμματα.



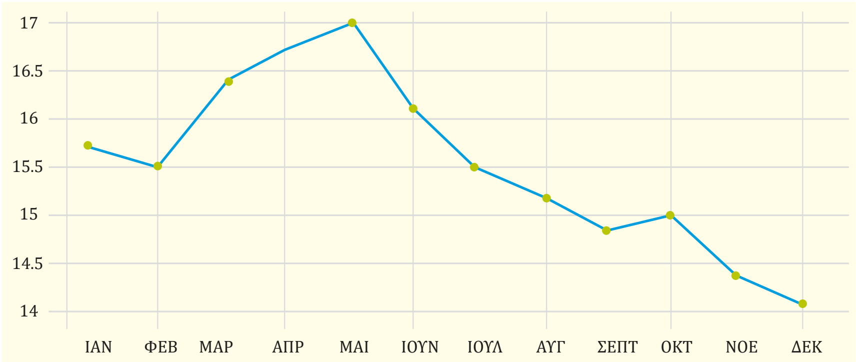
α. Να κατατάξετε τις ομάδες σε αύξουσα σειρά ανάλογα με:
i.την μέγιστη τιμή των πόντων.
ii.την τιμή της διαμέσου των πόντων.
iii.την τιμή του ενδοτεταρτημοριακού εύρους των πόντων.

β. Ποια από τις τρεις ομάδες νομίζετε ότι είχε καλύτερη θέση στην βαθμολογία και ποια την χειρότερη;

Αιτιότητα

5 **Πελαργοί και αριθμός γεννήσεων**

Σε πολλές ευρωπαϊκές χώρες οι γονείς λένε στα παιδιά τους ότι οι πελαργοί φέρνουν τα μωρά. Τα πουλιά αυτά καταφθάνουν στην Ευρώπη την άνοιξη (Μάρτιο έως Μάιο), επιστρέφοντας από νότιες περιοχές, όπως η Αφρική, όπου διαβιούν τον χειμώνα. Στο χρονοδιάγραμμα που ακολουθεί παρουσιάζεται ο αριθμός γεννήσεων, σε χιλιάδες, κατά τη διάρκεια ενός έτους σε μια ευρωπαϊκή χώρα.



Παρατηρούμε ότι υπήρξε αύξηση του αριθμού των γεννήσεων την άνοιξη του συγκεκριμένου έτους, δηλαδή μεταξύ Μαρτίου και Μαΐου.

α. Μπορούμε να συμπεράνουμε ότι η άφιξη των πελαργών προκαλεί αύξηση του αριθμού γεννήσεων ή το αντίστροφο;

β. Ποιοι παράγοντες νομίζετε ότι είναι οι αιτίες για τη σχέση εξάρτησης που παρατηρήθηκε μεταξύ των μεταβλητών «άφιξη πελαργών» και «αριθμός γεννήσεων»;

ΚΕΦΑΛΑΙΟ 4ο

Πιθανότητες

Η ιστορία των πιθανοτήτων-Προσωπικότητες και διάσημα προβλήματα

Σύντομη γνωριμία

Οι πιθανότητες είναι ένας σπουδαίος κλάδος των μαθηματικών που έχει τεράστιο ενδιαφέρον τόσο στο θεωρητικό του σκέλος όσο και στο χώρο των εφαρμογών. Στην καθημερινή ζωή αλλά και στο χώρο των επιστημών, συναντάμε πολύ συχνά την έννοια της πιθανότητας. Για παράδειγμα, ποια είναι η πιθανότητα να έχει μια οικογένεια με δυο παιδιά δύο αγόρια. Παρόμοια ενδιαφέροντα προβλήματα παρουσιάζονται σε τυχερά παιγνίδια με ζάρια, χαρτιά ή ρουλέτες και έχουν απασχολήσει κατά καιρούς σημαντικούς μαθηματικούς. Στο κεφάλαιο αυτό εισάγουμε την έννοια του πειράματος τύχης, του δειγματοχώρου, του ενδεχομένου, της πιθανότητας ενός ενδεχομένου και παρουσιάζουμε τους βασικούς κανόνες λογισμού των πιθανοτήτων.

Περιεχόμενα

Blaise Pascal
(1623 -1662)

Γάλλος μαθηματικός, φυσικός, συγγραφέας και φιλόσοφος, θεολόγος και στατιστικολόγος. Γνωστός από πληθώρα θεωρημάτων και αρχών που φέρουν το όνομά του. Με την εργασία του Traité du triangle arithmétique, που δημοσιεύτηκε το 1654, έθεσε τις βάσεις για τη συνδυαστική και την θεωρία πιθανοτήτων.

Jacob Bernoulli
(1655 -1705)

Βρετανός μαθηματικός, στατιστικολόγος, βιολόγος, γενετιστής και ακαδημαϊκός.
Για το έργο του στη στατιστική, χαρακτηρίστηκε ως μια ιδιοφυΐα, η οποία σχεδόν μόνη της δημιούργησε τα θεμέλια για τη σύγχρονη στατιστική επιστήμη.
Το 1925 δημοσίευσε το «Statistical Methods for Research Workers», ένα από τα σημαντικότερα βιβλία του 20ου αιώνα σχετικά με τις στατιστικές μεθόδους.

Galileo di Vincenzo Bonaiuti de Galile
(1564 – 1642)

Είναι πιο γνωστός με το όνομα Γαλιλαίος. Ιταλός φυσικός, μαθηματικός, αστρονόμος και φιλόσοφος με δράση σε όλους τους τομείς της ανθρώπινης διανόησης. Στη θεωρία των πιθανοτήτων, η οποία μας ενδιαφέρει εδώ, είναι γνωστό το πρόβλημα του Γαλιλαίου, όπου ήταν από τους πρώτους που υπολόγισε σωστά την πιθανότητα του ενδεχομένου να εμφανιστεί άθροισμα 9 με τρία ζάρια και ότι είναι μικρότερη από τη πιθανότητα του ενδεχομένου να εμφανιστεί άθροισμα 10.

4.1 Πειράματα Τύχης

Περιέχονται:
- Κατηγορίες πειραμάτων
- Περιγραφή πειραμάτων τύχης
- Πιθανοθεωρητικά Μοντέλα

Λέξεις κλειδιά:
Πείραμα τύχης
Δειγματοχώρος
Δειγματικός χώρος
Ενδεχόμενο
Πιθανοθεωρητικά μοντέλα

Στην ενότητα αυτή θα μάθουμε:
- Να περιγράφουμε πειράματα τύχης.
- Να αναγνωρίζουμε τη χρησιμότητα των πιθανοθεωρητικών μοντέλων.

Δραστηριότητα
Παρακάτω περιγράφονται διάφορα πειράματα. Σε ποια από αυτά μπορούμε να προβλέψουμε με βεβαιότητα το αποτέλεσμά τους και σε ποια όχι;
α. Ρίχνουμε ένα ζάρι και καταγράφουμε την ένδειξη της άνω έδρας του.
β. Καταγράφουμε το χρονικό διάστημα που μεσολαβεί μεταξύ δύο διαδοχικών εκλείψεων του ηλίου.
γ. Υπολογίζουμε το άθροισμα των γωνιών ενός τριγώνου.
δ. Καταγράφουμε τον αριθμό των τροχαίων ατυχημάτων σε μια πόλη, κατά τη διάρκεια ενός μήνα.
ε. Καταθέτουμε στην τράπεζα ένα συγκεκριμένο ποσό χρημάτων με σταθερό επιτόκιο και υπολογίζουμε τον τόκο ενός έτους.
στ. Επιλέγουμε με κλειστά μάτια ένα τραπουλόχαρτο, από μια τράπουλα των 52 φύλλων.

Κατηγορίες πειραμάτων

Πολλές φορές, όταν διεξάγουμε ένα πείραμα μπορούμε να προβλέψουμε με βεβαιότητα το αποτέλεσμά του. Σε αυτή την περίπτωση λέμε, ότι το πείραμα είναι αιτιοκρατικό ή ντετερμινιστικό ή προσδιοριστικό. Για παράδειγμα, τα πειράματα (β), (γ) και (ε), που περιγράφονται στη δραστηριότητα, είναι αιτιοκρατικά. Όταν ένα τέτοιο πείραμα επαναλαμβάνεται, κάτω από τις ίδιες συνθήκες, το αποτέλεσμα του είναι απολύτως προκαθορισμένο και μπορεί να προβλεφθεί με βεβαιότητα. Υπάρχουν όμως πειράματα, όπως τα (α), (δ) και (στ) της παραπάνω δραστηριότητας, στα οποία είναι αδύνατη οποιαδήποτε πρόβλεψη του αποτελέσματος. Αυτά τα πειράματα ονομάζονται τυχαία ή στοχαστικά.

Γενικότερα:
Τυχαίο πείραμα, ή αλλιώς πείραμα τύχης ή αλλιώς στοχαστικό πείραμα, ονομάζεται κάθε διαδικασία που εκτελείται (πείραμα) ή παρατηρείται (φαινόμενο) στην οποία το τελικό αποτέλεσμα δεν είναι γνωστό εκ των προτέρων.
Στα πειράματα τύχης γνωρίζουμε τα πιθανά αποτελέσματα που μπορούν να προκύψουν, ενώ κάτι που χαρακτηρίζει ένα πείραμα τύχης είναι ότι μπορεί να επαναληφθεί κάτω από τις ίδιες συνθήκες πολλές φορές.
Ορισμένα πειράματα, αν και βασίζονται σε αιτιοκρατικά μοντέλα, παρουσιάζουν μεγάλη αβεβαιότητα λόγω της πολυπλοκότητας των συνθηκών τους. Ένα παράδειγμα είναι το στρίψιμο ενός αμερόληπτου κέρματος: θεωρητικά, η έκβαση μπορεί να προβλεφθεί με τους νόμους της Φυσικής, αλλά η ευαισθησία στις αρχικές συνθήκες καθιστά τη μοντελοποίηση μη πρακτική. Έτσι, η Θεωρία

Πιθανοτήτων χρησιμοποιείται για να περιγράψει το φαινόμενο ως πείραμα τύχης. Μπορούμε λοιπόν να πούμε ότι «η τυχαιότητα μοντελοποιεί την πολυπλοκότητα».

Περιγραφή πειράματος τύχης

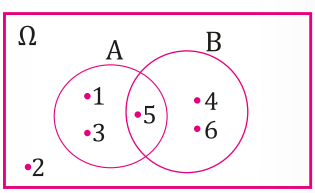
Περιγραφή πειραμάτων τύχης

Τα πειράματα τύχης τα οποία θα μελετήσουμε έχουν πεπερασμένο πλήθος δυνατών αποτελεσμάτων. Για να δούμε πώς περιγράφεται ένα πείραμα, ας πάρουμε για παράδειγμα τη ρίψη ενός ζαριού.
Τα δυνατά αποτελέσματα του πειράματος είναι οι φυσικοί αριθμοί από το 1 μέχρι το 6. Το σύνολο $\Omega=\{1,2,3,4,5,6\}$ έχει στοιχεία τα δυνατά αποτελέσματα και λέγεται δειγματοχώρος ή δειγματικός χώρος του πειράματος.

Κάθε υποσύνολο του Ω, όπως τα $\varnothing$, Ω, $\{6\}$, $\{1,2\}$, $\{2,4,6\}$, $\{1,3,5\}$, $\{4,6\}$ λέγεται ενδεχόμενο του Ω.
Τα μονομελή ενδεχόμενα $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$ λέγονται απλά ή στοιχειώδη ενδεχόμενα.

Ρίψη ζαριού

Τα υποσύνολα του Ω, που περιέχουν περισσότερα από ένα στοιχεία, όπως τα Ω, $\{1,2\}$, $\{2,4,6\}$, $\{1,3,5\}$, $\{4,6\}$ λέγονται σύνθετα ενδεχόμενα.
Αν, τώρα, ρίξουμε το ζάρι μια φορά και έρθει, ως αποτέλεσμα π.χ. το 6, τότε λέμε ότι πραγματοποιήθηκε κάθε ενδεχόμενο στο οποίο ανήκει το 6, όπως τα $\{6\}$, $\{2,4,6\}$, $\{4,6\}$.



Πολλές φορές χρησιμοποιούμε τα διαγράμματα Venn, για να παραστήσουμε διάφορα ενδεχόμενα. Στο διπλανό σχήμα παριστάνουμε τον δειγματοχώρο $\Omega=\{1,2,3,4,5,6\}$ και τα ενδεχόμενά του $A=\{1,3,5\}$ και $B=\{4,5,6\}$.

Έχουμε, επομένως, τους παρακάτω ορισμούς:

Ορισμοί

Δειγματοχώρος ή Δειγματικός χώρος (Sample Space)
Ονομάζεται το σύνολο το οποίο έχει ως στοιχεία τα δυνατά αποτελέσματα του πειράματος τύχης. Συμβολίζεται, συνήθως, με το γράμμα Ω.

Ενδεχόμενο (Event)
Ονομάζεται κάθε υποσύνολο του δειγματοχώρου.

Απλό ενδεχόμενο ή στοιχειώδες ενδεχόμενο
Ονομάζεται κάθε ενδεχόμενο που έχει ένα μόνο δυνατό αποτέλεσμα.

Σύνθετο ενδεχόμενο
Ονομάζεται κάθε ενδεχόμενο με δύο, τουλάχιστον, δυνατά αποτελέσματα.

Πραγματοποίηση ενδεχομένου
Θα λέμε ότι το ενδεχόμενο A πραγματοποιήθηκε, όταν το αποτέλεσμα του πειράματος τύχης είναι στοιχείο του A.

Βέβαιο ενδεχόμενο
Τον δειγματοχώρο Ω τον ονομάζουμε βέβαιο ενδεχόμενο, γιατί πραγματοποιείται πάντοτε.

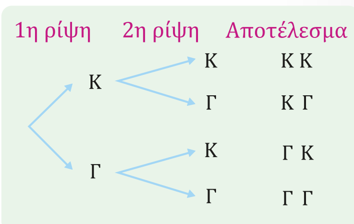
Αδύνατο ενδεχόμενο
Το κενό σύνολο $\varnothing$ το ονομάζουμε αδύνατο ενδεχόμενο, γιατί δεν πραγματοποιείται ποτέ.

Σημείωση: Το πλήθος των στοιχείων ενός ενδεχομένου Α συμβολίζεται με $\nu(A)$.
Για το κενό σύνολο δεχόμαστε ότι $\nu(\varnothing)=0$.

Προσπαθήστε να λύσετε τις ασκήσεις 1 έως 3

Εφαρμογή

Ο Λεωνίδας και ο Σίμος διαφωνούν για το ποιος είναι ο δειγματοχώρος του πειράματος των δύο διαδοχικών ρίψεων ενός νομίσματος. Ο Λεωνίδας κατασκεύασε το διπλανό δεντροδιάγραμμα και πρότεινε στον Σίμο ως δειγματοχώρο το σύνολο $\Omega_1=\{ΚΚ, ΚΓ, ΓΚ, ΓΓ\}$.
Ο Σίμος συμβόλισε με 2Γ το ενδεχόμενο να φέρουν και οι δύο γράμματα, με 1Γ να φέρει μόνο ένας από τους δύο γράμματα και με 0Γ να μην φέρει κανένας γράμματα. Αντιπρότεινε τελικά ως δειγματοχώρο το σύνολο $\Omega_2=\{2Γ, 1Γ, 0Γ\}$.
Ποιος από τους δύο έχει δίκιο; Ποιος δειγματοχώρος περιγράφει καλύτερα το πείραμα;



Λύση

Και οι δύο δειγματοχώροι είναι σωστοί εφόσον περιέχουν όλα τα δυνατά αποτελέσματα του πειράματος. Ο δειγματοχώρος του Λεωνίδα περιγράφει το πείραμα με πιο αναλυτικό τρόπο. Αν, για παράδειγμα, γνωρίζουμε ότι το αποτέλεσμα είναι ΚΓ τότε ξέρουμε ότι ο Λεωνίδας έφερε κεφαλή και ο Σίμος γράμματα. Το αντίστοιχο αποτέλεσμα στον δειγματοχώρο του Σίμου είναι το 1Γ. Είναι φανερό ότι η πληροφορία 1Γ δεν μας δίνει τη δυνατότητα να διακρίνουμε ποιος από τους δύο έφερε γράμματα και ποιος κεφαλή.

Προσπαθήστε να λύσετε τις ασκήσεις 4 έως 11

Πιθανοθεωρητικά Μοντέλα

Με τον όρο Πιθανοθεωρητικά Μοντέλα εννοούμε τις μαθηματικές κατασκευές που χρησιμοποιούμε για να περιγράψουμε τα πειράματα τύχης. Ο δειγματοχώρος ενός πειράματος τύχης, τα ενδεχόμενα του και η πιθανότητα που αποδίδουμε σε καθένα από αυτά, συνιστούν ένα πιθανοθεωρητικό μοντέλο. Τα μοντέλα αυτά έχουν πλήθος εφαρμογών. Ενδεικτικά αναφέρουμε τις παρακάτω:

- Στη Μετεωρολογία χρησιμοποιούνται μοντέλα της θεωρίας πιθανοτήτων για την πρόβλεψη της κατάστασης του καιρού των επόμενων ημερών.
- Στις Επιστήμες Υγείας και στη Βιολογία έχουμε πληθώρα εφαρμογών των μοντέλων πιθανοτήτων, όπως στην εκτίμηση της καταλληλότητας ενός φαρμάκου, στις μεταλλάξεις που συμβαίνουν στο γονιδίωμα των ζωντανών οργανισμών, στους νόμους που διέπουν την κληρονομικότητα, στη μελέτη προσαρμογής των ζώντων οργανισμών σε αβέβαια περιβάλλοντα και πολλές άλλες.
- Στα τυχερά παιχνίδια έχουμε μια εξαιρετική πηγή προβλημάτων για την εφαρμογή πιθανοθεωρητικών μοντέλων. Αποτέλεσαν ιστορικά τη βάση πάνω στην οποία οικοδομήθηκε η θεωρία των πιθανοτήτων.
- Στην Τεχνητή Νοημοσύνη η λειτουργία και εκπαίδευση των μηχανών που προσομοιάζουν την ανθρώπινη νοημοσύνη βασίζεται σε αλγορίθμους που αξιοποιούν πιθανοθεωρητικά μοντέλα.

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Ποια από τα παρακάτω είναι πειράματα τύχης;

α. Επιλέγουμε με κλειστά μάτια ένα από τα 52 φύλλα μιας τράπουλας.

β. Αφήνουμε να πέσει από συγκεκριμένο ύψος ένα σφαιρίδιο και καταγράφουμε την ταχύτητα με την οποία φτάνει στο έδαφος.

γ. Μετράμε πόσοι πελάτες εισέρχονται σε ένα κατάστημα μια συγκεκριμένη ώρα της ημέρας.

δ. Καταγράφουμε το αποτέλεσμα του πολλαπλασιασμού $5 \cdot 3$.

2 Ο δειγματοχώρος ενός πειράματος τύχης είναι το σύνολο $\Omega = \{1,3,5,7,9,11\}$. Ποια από τα παρακάτω σύνολα είναι ενδεχόμενα του πειράματος; Ποια από αυτά είναι απλά και ποια σύνθετα ενδεχόμενα;

$$A = \{1,3\}, \quad \varnothing, \quad B = \{9,12\}, \quad \Omega, \quad \Gamma = \{5\},$$

$$\Delta = \{1,3,5,7,9,11,13\}.$$

3 Στο πείραμα ρίψης ενός ζαριού φέραμε τον αριθμό 5. Ποια από τα παρακάτω ενδεχόμενα πραγματοποιήθηκαν;

$$A = \{1,3,4\}, \quad \varnothing, \quad B = \{4,6\}, \quad \Gamma = \{5\}, \quad \Omega = \{1,2,3,4,5,6\}.$$

4 Ένας μαθητής έγραψε ότι ο δειγματοχώρος του πειράματος τύχης της ρίψης ενός ζαριού είναι ο Ω={άρτιος, περιττός}. Είναι σωστός ο δειγματοχώρος που έγραψε ο μαθητής; Ποιες αδυναμίες εντοπίζετε στον Ω;

Ασκήσεις

5 Ρίχνουμε ένα ζάρι και καταγράφουμε την ένδειξη της άνω έδρας του. Να παραστήσετε με αναγραφή των στοιχείων του το ενδεχόμενο A: «Το αποτέλεσμα να είναι αριθμός που διαιρείται με το 3».

6 Ρίχνουμε ένα ζάρι δύο φορές και καταγράφουμε τις ενδείξεις των άνω εδρών με τη σειρά που ήρθαν.

α. Να βρείτε τον δειγματοχώρο του πειράματος.

β. Να παραστήσετε με αναγραφή των στοιχείων του το ενδεχόμενο A: «Το ζάρι έφερε διπλές».

7 Ένα εργοστάσιο αυτοκινήτων παράγει ένα μοντέλο αυτοκινήτου με τις παρακάτω προδιαγραφές:

- Με τρεις επιλογές ηχοσυστημάτων: H_1, H_2, H_3.
- Με δύο επιλογές χρώματος: X_1, X_2.

Επιλέγουμε ένα αυτοκίνητο του εργοστασίου. Χρησιμοποιώντας ένα δεντροδιάγραμμα να βρείτε έναν κατάλληλο δειγματοχώρο του πειράματος.

8 Ρίχνουμε ένα νόμισμα τρεις φορές και καταγράφουμε τις ενδείξεις των άνω εδρών με τη σειρά που ήρθαν.

α. Να βρείτε τον δειγματοχώρο του πειράματος.

β. Να παραστήσετε με αναγραφή των στοιχείων τους και με ένα διάγραμμα Venn τα παρακάτω ενδεχόμενα:

A: «Να φέρουμε περισσότερες κεφαλές από γράμματα»,

B: «Να φέρουμε τουλάχιστον μία φορά κεφαλή»,

Γ: «Να φέρουμε το πολύ μία φορά κεφαλή».

9 Πατάμε από μία φορά, στην τύχη, το καθένα από τα πλήκτρα Α, Ε και Ν του πληκτρολογίου του υπολογιστή μας και τα καταγράφουμε με τη σειρά που τα πατήσαμε.

α. Ποιος είναι ο δειγματοχώρος του πειράματος;

β. Να παραστήσετε με αναγραφή των στοιχείων του και με ένα διάγραμμα Venn το ενδεχόμενο να σχηματιστεί λέξη που έχει νόημα στην Ελληνική γλώσσα.

10 Σε ένα κουτί υπάρχουν τρεις σφαίρες, μία άσπρη, μία κόκκινη και μία μαύρη. Επιλέγουμε μία σφαίρα από το κουτί και καταγράφουμε το χρώμα της. Έπειτα, χωρίς να βάλουμε την πρώτη σφαίρα στο κουτί, επιλέγουμε άλλη μία σφαίρα και καταγράφουμε το χρώμα της.

α. Να βρείτε τον δειγματοχώρο του πειράματος.

β. Να παραστήσετε με αναγραφή των στοιχείων του και με ένα διάγραμμα Venn το ενδεχόμενο να μην βγει μαύρη σφαίρα.

11 Ρίχνουμε διαδοχικά ένα νόμισμα και ένα ζάρι και καταγράφουμε τις ενδείξεις των άνω εδρών τους.

α. Να βρείτε έναν κατάλληλο δειγματοχώρο του πειράματος.

β. Να παραστήσετε με αναγραφή των στοιχείων τους τα επόμενα ενδεχόμενα και να βρείτε τις πιθανότητές τους:

A: «Το αποτέλεσμα είναι κεφαλή και περιττός αριθμός» και

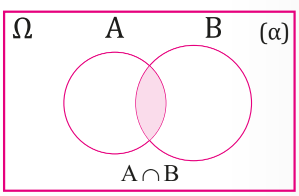
B: « Το αποτέλεσμα είναι γράμματα και αριθμός μεγαλύτερος του 2».

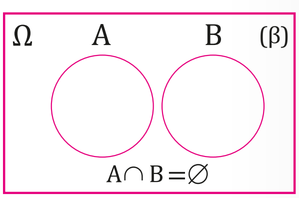
4.2 Μεταγραφή Ενδεχομένων

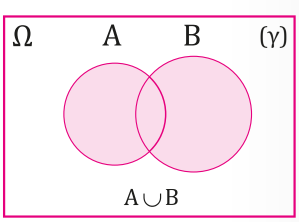
Κλασικός ορισμός της πιθανότητας

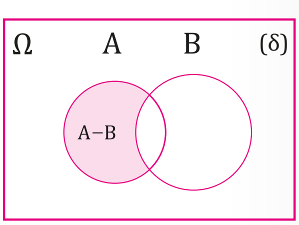
Περιέχονται:
- Πράξεις και σχέσεις ενδεχομένων σε φυσική και συμβολική γλώσσα
- Τομή, ένωση και διαφορά ενδεχομένων
- Συμπλήρωμα ενδεχομένου

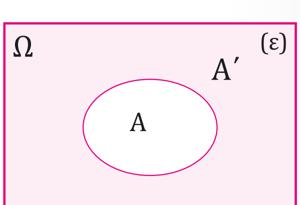
Λέξεις κλειδιά:
Μεταγραφή ενδεχομένου
Κλασικός ορισμός πιθανότητας











Στην ενότητα αυτή θα μάθουμε:
- Να μεταγράφουμε ενδεχόμενα και σχέσεις ενδεχομένων που είναι διατυπωμένες σε φυσική γλώσσα, στη γλώσσα των συνόλων και αντίστροφα.
- Να υπολογίζουμε πιθανότητες ενδεχομένων με τον κλασικό ορισμό.

Πράξεις και σχέσεις ενδεχομένων σε φυσική και συμβολική γλώσσα

Κατά τη μοντελοποίηση ενός προβλήματος πιθανοτήτων είναι απαραίτητο να μεταβούμε από την περιγραφή του σε φυσική γλώσσα στην αντίστοιχη μαθηματική και αντίστροφα. Επειδή τα ενδεχόμενα ενός δειγματοχώρου είναι σύνολα, μπορούμε να αναγνωρίζουμε και να δηλώνουμε πράξεις και σχέσεις μεταξύ ενδεχομένων, οι οποίες είναι αντίστοιχες εκείνων των συνόλων, με κυριότερες την τομή, την ένωση, τη διαφορά και το συμπλήρωμα.

Η τομή δύο ενδεχομένων A και B συμβολίζεται με $A\cap B$ και διαβάζεται « A και B ». Πραγματοποιείται όταν πραγματοποιούνται συγχρόνως τα A και B.
Σε διάγραμμα Venn παριστάνεται με το κοινό μέρος των δύο συνόλων (Σχ. (α)).

Αν τα ενδεχόμενα δεν έχουν κοινά στοιχεία, τότε η τομή τους είναι το κενό σύνολο. Στην περίπτωση αυτή γράφουμε $A\cap B=\varnothing$ και λέμε ότι τα ενδεχόμενα είναι ξένα ή ασυμβίβαστα ή αμοιβαία αποκλειόμενα, αφού όταν πραγματοποιείται το ένα αποκλείεται να πραγματοποιηθεί το άλλο (Σχ. (β)).

Η ένωση δύο ενδεχομένων A και B συμβολίζεται με $A\cup B$ και διαβάζεται «A ή B». Πραγματοποιείται όταν πραγματοποιείται το A ή το B, δηλαδή πραγματοποιείται, όταν πραγματοποιείται τουλάχιστον ένα από τα A και B (Σχ. (γ)).

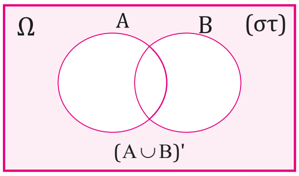
Η διαφορά του ενδεχομένου B από το A συμβολίζεται με $A-B$ και διαβάζεται «A και όχι B». Πραγματοποιείται όταν πραγματοποιείται το A και δεν πραγματοποιείται το B. Σε περιπτώσεις που δεν μπορεί να υπάρξει παρανόηση, λέμε ότι το $A-B$ πραγματοποιείται όταν πραγματοποιείται μόνο το A (Σχ. (δ)). Εύκολα καταλαβαίνουμε ότι οι διαφορές $A-B$ και $B-A$ δεν είναι πάντα ίσες.

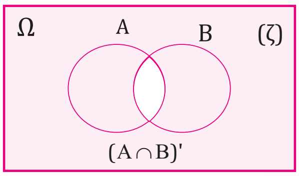
Ειδικά, η διαφορά $\Omega-A$ συμβολίζεται A' και ονομάζεται συμπλήρωμα ή συμπληρωματικό ενδεχόμενο του A και πραγματοποιείται όταν δεν πραγματοποιείται το A (Σχ. (ε)).
Ακόμα διαβάζεται και ως, « όχι A » και γι' αυτό λέγεται και αντίθετο του A.
Εύκολα καταλαβαίνουμε ότι

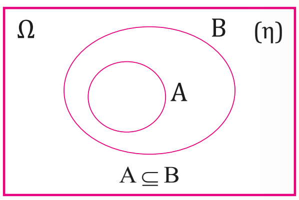
$$A'=\Omega-A,\quad A\cap A'=\varnothing \quad \text{και} \quad A\cup A'=\Omega$$

Επειδή, όταν πραγματοποιείται το A και δεν πραγματοποιείται το B είναι σαν λέμε, ότι πραγματοποιείται το A και πραγματοποιείται το B', είναι:

$$A-B=A\cap B'$$







Πράξεις-σχέσεις ενδεχομένων

Μεταγραφή ενδεχομένων

Περιγραφή πειράματος τύχης

Με την έκφραση «επιλέγουμε τυχαία» ή «επιλέγουμε στην τύχη» ένα αντικείμενο ενός συνόλου θα εννοούμε ότι κάθε αντικείμενο έχει την ίδια πιθανότητα να επιλεχθεί.

Ενδιαφέρον έχουν και τα συμπληρώματα της ένωσης και της τομής.
Το συμπλήρωμα της ένωσης διαβάζεται «όχι (A ή B)», συμβολίζεται με $(A\cup B)'$ και πραγματοποιείται όταν δεν πραγματοποιείται κανένα από τα A και B (Σχ. (στ)).

Το συμπλήρωμα της τομής διαβάζεται «όχι (A και B)», συμβολίζεται με $(A\cap B)'$ και πραγματοποιείται όταν δεν πραγματοποιούνται συγχρόνως τα A και B ή διαφορετικά, όταν πραγματοποιείται το πολύ ένα από τα A και B (Σχ. (ζ)).

Έστω $A\subseteq B$. Όταν πραγματοποιείται το ενδεχόμενο A, τότε πραγματοποιείται και το ενδεχόμενο Β. Λέμε ότι το ενδεχόμενο Α συνεπάγεται το Β (Σχ. (η)).

Παράδειγμα 1

Από τα παιδιά μιας τάξης σκοπεύουμε να επιλέξουμε ένα και να το ρωτήσουμε αν μαθαίνει Αγγλικά ή Γερμανικά. Θεωρούμε τα ενδεχόμενα:
A: «Το παιδί μαθαίνει Αγγλικά» και B: «Το παιδί μαθαίνει Γερμανικά».
Στον παρακάτω πίνακα φαίνεται η αντιστοιχία μεταξύ φυσικής και συμβολικής γλώσσας για διάφορα ενδεχόμενα.

ΦΥΣΙΚΗ ΓΛΩΣΣΑ	ΣΥΜΒΟΛΙΚΗ ΓΛΩΣΣΑ
Το παιδί μαθαίνει Αγγλικά και Γερμανικά.	$A\cap B$
Το παιδί μαθαίνει Αγγλικά ή Γερμανικά.	$A\cup B$
Το παιδί μαθαίνει Αγγλικά και δεν μαθαίνει Γερμανικά.	$A-B$
Το παιδί μαθαίνει Γερμανικά και όχι Αγγλικά.	$B-A$
Το παιδί δεν μαθαίνει Γερμανικά.	B'

Προσπαθήστε να λύσετε τις ασκήσεις 1 έως 7

Παράδειγμα 2

Από τα ψηφία του αριθμού 12345678 επιλέγουμε ένα, στην τύχη, και θεωρούμε τα ενδεχόμενα:

A: Το ψηφίο είναι άρτιος αριθμός.
B: Το ψηφίο είναι περιττός αριθμός.
Γ: Το ψηφίο είναι μικρότερο του 6.

α. Να εξετάσετε αν τα ενδεχόμενα A και B είναι ασυμβίβαστα.
β. Να παραστήσετε λεκτικά και με αναγραφή των στοιχείων τους τα ενδεχόμενα A', B' και $(A\cup B)'$ και να βρείτε τις πιθανότητές τους.

Λύση

α. Τα ενδεχόμενα είναι ασυμβίβαστα, γιατί δεν έχουν κοινά στοιχεία. Αυτό είναι φανερό γιατί ένα ψηφίο δεν μπορεί να είναι και άρτιος και περιττός αριθμός.
β. Είναι $\Omega=\{1, 2, 3, 4, 5, 6, 7, 8\}$

Μεταγραφή ενδεχομένων

και

$$A = \{\ 2,\ 4,\ 6,\ 8\ \} \text{ και } B = \{\ 1,\ 3,\ 5,\ 7\ \}$$

Το A' είναι το ενδεχόμενο «το ψηφίο δεν είναι άρτιος αριθμός». Επομένως είναι το ενδεχόμενο «το ψηφίο είναι περιττός αριθμός», δηλαδή είναι το B:

$$A' = B = \{\ 1,\ 3,\ 5,\ 7\ \}$$

Το B' είναι το ενδεχόμενο «το ψηφίο δεν είναι περιττός αριθμός». Επομένως είναι το ενδεχόμενο «το ψηφίο είναι άρτιος αριθμός», δηλαδή είναι το A:

$$B' = A = \{\ 2,\ 4,\ 6,\ 8\ \}$$

Το $(A \cup B)'$ είναι το ενδεχόμενο «όχι (A ή B)». Πραγματοποιείται, όταν δεν πραγματοποιείται κανένα από τα A και B.

Παρατηρούμε ότι $A \cup B = \Omega$. Επομένως $(A \cup B)' = \varnothing$

Από τον κλασικό ορισμό της πιθανότητας έχουμε:

$$P(A') = \frac{\nu(A')}{\nu(\Omega)} = \frac{4}{8} = \frac{1}{2}\ ,\quad P(B') = \frac{\nu(B')}{\nu(\Omega)} = \frac{4}{8} = \frac{1}{2}\ ,\quad P\big((A \cup B)'\big) = P(\varnothing) = 0$$

Προσπαθήστε να λύσετε τις ασκήσεις 8 και 9

Εφαρμογή

Μια εταιρεία απασχολεί 4 άτομα ως διοικητικό προσωπικό και 40 υπαλλήλους σε δύο τομείς, τον τομέα 1 με 24 εργαζόμενους και τον τομέα 2 με 16 υπαλλήλους, εκ των οποίων οι 5 ανήκουν και στους δύο τομείς, ως συντονιστές του έργου τους. Επιλέγουμε τυχαία έναν υπάλληλο από τους 44. Θεωρούμε τα ενδεχόμενα:

- A: «Ο υπάλληλος εργάζεται στον τομέα 1».
- B: «Ο υπάλληλος εργάζεται στον τομέα 2».

Διάγραμμα Venn με δύο ενδεχόμενα

α. Να διατυπώσετε σε φυσική γλώσσα τα ενδεχόμενα A' και $B - A$ και να υπολογίσετε τις πιθανότητές τους.

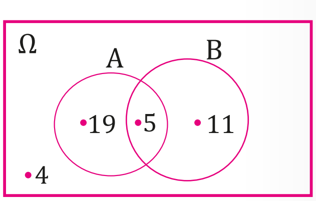
β. Να γράψετε σε συμβολική γλώσσα τα επόμενα ενδεχόμενα και να υπολογίσετε τις πιθανότητές τους:

- Γ: Ο υπάλληλος δεν εργάζεται σε κανέναν από τους τομείς 1 και 2.
- Δ: Ο υπάλληλος εργάζεται το πολύ σε έναν από τους τομείς 1 και 2.

Λύση

α. Το A' είναι το ενδεχόμενο:
«Ο υπάλληλος δεν εργάζεται στον τομέα 1».
Το $\nu(A')$ υπολογίζεται αν από το $\nu(\Omega) = 44$ αφαιρέσουμε το $\nu(A) = 24$, δηλαδή είναι $\nu(A') = 44 - 24 = 20$. Επομένως

$$P(A') = \frac{\nu(A')}{\nu(\Omega)} = \frac{20}{44} = \frac{5}{11}$$



Το $B - A$ είναι το ενδεχόμενο:
«Ο υπάλληλος εργάζεται στον τομέα 2 και δεν εργάζεται στον τομέα 1».
Το $\nu(B - A)$ υπολογίζεται εύκολα, αρκεί από το $\nu(B) = 16$ να αφαιρέσουμε 5 που αντιστοιχεί στους κοινούς υπαλλήλους των δύο τομέων, δηλαδή

$$\nu(B - A) = 16 - 5 = 11$$

Επομένως,

$$P(B-A)=\frac{\nu(B-A)}{\nu(\Omega)}=\frac{11}{44}=\frac{1}{4}$$

β. Το Γ είναι το ενδεχόμενο:
«Ο υπάλληλος να μην εργάζεται σε κανέναν από τους τομείς 1 και 2».
Επομένως είναι το συμπληρωματικό της ένωσης, δηλαδή είναι το $(A\cup B)'$. Επειδή η ένωση έχει 40 ευνοϊκές περιπτώσεις το $(A\cup B)'$ έχει $44-40=4$ ευνοϊκές περιπτώσεις. Επομένως είναι $\nu(\Gamma)=4$ και

$$P(\Gamma)=\frac{\nu(\Gamma)}{\nu(\Omega)}=\frac{4}{44}=\frac{1}{11}$$

Το Δ είναι το ενδεχόμενο:
«Ο υπάλληλος εργάζεται το πολύ σε έναν από τους τομείς 1 και 2».
Επομένως, είναι το συμπλήρωμα $(A\cap B)'$. Επειδή η τομή έχει 5 ευνοϊκές περιπτώσεις, το συμπλήρωμά της έχει $44-5=39$ ευνοϊκές περιπτώσεις. Επομένως,

$$P(\Delta)=\frac{\nu(\Delta)}{\nu(\Omega)}=\frac{39}{44}$$

Προσπαθήστε να λύσετε τις ασκήσεις 10, 11 και 12

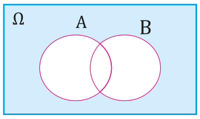
Ασκήσεις και Προβλήματα

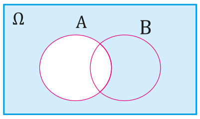
Ερωτήσεις

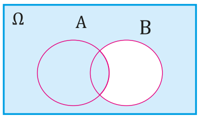
1 Να συμπληρώσετε τον παρακάτω πίνακα.

Ενδεχόμενο	Συμβολισμός
Α και Β	
Α ή Β	
όχι Α	
Α και όχι Β	
όχι Α και όχι Β	
Πραγματοποιείται το πολύ ένα από τα Α και Β	

2 Να εκφράσετε στη γλώσσα των συνόλων και σε φυσική γλώσσα τα παρακάτω σκιαγραφημένα με μπλε χρώμα ενδεχόμενα:

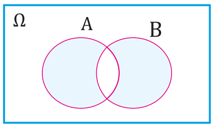






3 **Συμμετρική διαφορά ενδεχομένων**
Το ενδεχόμενο $(A-B)\cup(B-A)$ ονομάζεται συμμετρική διαφορά των Α και Β. Στο παρακάτω διάγραμμα το ενδεχόμενο αυτό έχει σκιαγραφηθεί με μπλε χρώμα. Ποια από τις παρακάτω εκφράσεις ορίζει τη συμμετρική διαφορά των Α και Β;

α. Πραγματοποιείται μόνο το Α και μόνο το Β.
β. Πραγματοποιείται μόνο το Α ή μόνο το Β.
γ. Πραγματοποιείται το Α και το Β.



4 Σε ποια από τα παρακάτω πειράματα τύχης τα ενδεχόμενα Α και Β είναι ασυμβίβαστα;

α. Επιλέγουμε μια Ελληνίδα γυναίκα. Α είναι το ενδεχόμενο να έχει ηλικία μικρότερη των 40 ετών και Β να είναι παντρεμένη περισσότερα από 40 έτη.
β. Επιλέγουμε έναν κάτοικο της Ελλάδας. Α είναι το ενδεχόμενο να είναι Έλληνας και Β να είναι κάτοικος Αθηνών.

Ασκήσεις

5 Ρίχνουμε δύο ζάρια. Με A συμβολίζουμε το ενδεχόμενο το άθροισμα των ενδείξεων να είναι άρτιος αριθμός και με B το ενδεχόμενο το πρώτο ζάρι να φέρει 1.
Να αποδώσετε λεκτικά τα παρακάτω ενδεχόμενα:
α. $A \cap B$ β. $A \cup B$ γ. A'

6 Ρίχνουμε μία φορά ένα ζάρι. Θεωρούμε τα ενδεχόμενα:
A : «Το ζάρι να φέρει 1» και
B : «Το ζάρι να φέρει περιττό αριθμό».
Να εξηγήσετε γιατί το ενδεχόμενο A συνεπάγεται το ενδεχόμενο B.

7 **Τύποι De Morgan** Χρησιμοποιώντας διαγράμματα Venn να επαληθεύσετε τις παρακάτω ισότητες μεταξύ ενδεχομένων και να αποδώσετε λεκτικά το νόημά τους.

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

8 Ρίχνουμε ένα αμερόληπτο νόμισμα μία φορά. Να παραστήσετε με αναγραφή όλα τα ενδεχόμενα του δειγματοχώρου του πειράματος και να υπολογίσετε τις πιθανότητές τους.

9 Ρίχνουμε μία φορά ένα ζάρι. Θεωρούμε τα ενδεχόμενα:
A : «Το αποτέλεσμα είναι πρώτος αριθμός» και
B : «Το αποτέλεσμα είναι άρτιος αριθμός».
Να αποδώσετε λεκτικά και με αναγραφή των στοιχείων τους τα παρακάτω ενδεχόμενα:
α. $A \cap B$ β. $A \cup B$ γ. A'
Στη συνέχεια, αν θεωρήσουμε το ζάρι αμερόληπτο, να υπολογίσετε τις πιθανότητές τους.

10 Ρωτήσαμε 40 παιδιά σε μια κατασκήνωση ποια από τις παρακάτω επιλογές πρωινού προτιμούν:
Α: Δημητριακά με γάλα Β: Ψωμί με μαρμελάδα
Από τις απαντήσεις προέκυψαν τα εξής:
- Σε 15 παιδιά αρέσει μόνο το Α.
- Σε 7 παιδιά αρέσει μόνο το Β.
- Σε 15 παιδιά αρέσουν και οι δύο επιλογές Α και Β.

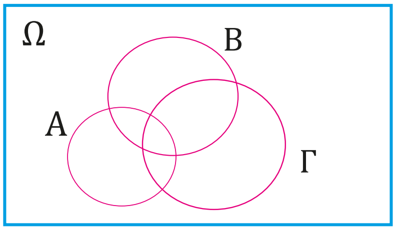
Επιλέγουμε στην τύχη ένα παιδί και θεωρούμε τα ενδεχόμενα:
A: «Στο παιδί να αρέσουν τα δημητριακά με γάλα» και
B: «Στο παιδί να αρέσει το ψωμί με μαρμελάδα».

α. Να παραστήσετε με ένα διάγραμμα Venn τα ενδεχόμενα Α και Β και να βρείτε τις πιθανότητές τους.
β. Να εκφράσετε σε φυσική γλώσσα τα παρακάτω ενδεχόμενα και να υπολογίσετε τις πιθανότητές τους:
A', B', $A \cap B$, $(A \cap B)'$, $(A \cup B)'$

11 Ρίχνουμε ένα συμμετρικό δωδεκάεδρο ζάρι που φέρει τις ενδείξεις 1, 2, 3, ..., 12 και καταγράφουμε την ένδειξη της άνω έδρας του. Θεωρούμε τα ενδεχόμενα
Α: «Η ένδειξη να είναι πολλαπλάσιο του 3»
Β: «Η ένδειξη να είναι πολλαπλάσιο του 4»

Να εκφράσετε πρώτα συμβολικά και έπειτα με αναγραφή των στοιχείων τους τα παρακάτω ενδεχόμενα και να βρείτε τις πιθανότητές τους:
α. Δεν πραγματοποιείται το Β.
β. Πραγματοποιούνται συγχρόνως τα A και B.
γ. Πραγματοποιείται το A και όχι το B.
δ. Δεν πραγματοποιείται κανένα από τα A και B.

12 Τα σύνολα A, B και Γ στο διάγραμμα Venn του σχήματος είναι ενδεχόμενα ενός δειγματοχώρου Ω. Να παραστήσετε τα παρακάτω ενδεχόμενα σχεδιάζοντας κάθε φορά την κατάλληλη περιοχή:
α. Πραγματοποιείται τουλάχιστον ένα από τα τρία ενδεχόμενα.
β. Πραγματοποιείται το Α ή το Β και όχι το Γ.
γ. Πραγματοποιούνται και τα τρία ενδεχόμενα.



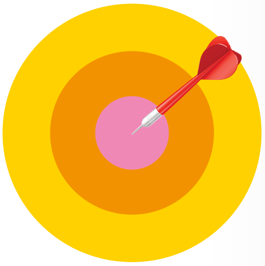
Διάγραμμα Venn με τρία ενδεχόμενα

4.3 Πειράματα τύχης με μη ισοπίθανα απλά ενδεχόμενα

Περιέχονται:

- Μη ισοπίθανα απλά ενδεχόμενα
  - Η ρίψη ενός μη συμμετρικού (μεροληπτικού) νομίσματος
  - Διαφοροποίηση του δειγματοχώρου ανάλογα με το τι θεωρείται αποτέλεσμα του πειράματος από διαφορετικούς παρατηρητές
  - Η ρίψη δύο αμερόληπτων ζαριών

Λέξεις κλειδιά:
Μη συμμετρικό νόμισμα
Ρίψη ενός νομίσματος
Ρίψη δύο ζαριών



Οι όροι "μεροληπτικό νόμισμα" και "μη συμμετρικό νόμισμα " είναι συνώνυμοι.

Στην ενότητα αυτή θα μάθουμε:
Να περιγράφουμε πειράματα τύχης με μη ισοπίθανα απλά ενδεχόμενα.

Δραστηριότητα
Ρίχνουμε ένα βέλος σε έναν κυκλικό στόχο ο οποίος είναι χωρισμένος σε τρεις περιοχές, την κίτρινη (Κ), την πορτοκαλί (Π) και τη ροζ (Ρ). Ως αποτέλεσμα του παραπάνω πειράματος θεωρούμε το χρώμα της περιοχής που θα καταλήξει το βέλος. Μπορούμε να θεωρήσουμε ότι τα δυνατά αποτελέσματα Κ, Π, Ρ είναι ισοπίθανα; Ποια περιοχή του στόχου θεωρείται πιο πιθανή να καταλήξει το βέλος

α. αν παίζουμε το παιχνίδι για πρώτη φορά;

β. αν έχουμε εμπειρία στο συγκεκριμένο παιχνίδι;

Μη ισοπίθανα απλά ενδεχόμενα

Στο Γυμνάσιο μελετήσαμε πειράματα με ισοπίθανα απλά ενδεχόμενα. Υπάρχουν, όμως, πειράματα στα οποία η υπόθεση αυτή δεν ισχύει. Ένα τέτοιο πείραμα είναι αυτό της παραπάνω δραστηριότητας. Αν διαθέτουμε εμπειρία στο συγκεκριμένο παιχνίδι, αναμένουμε ότι το βέλος θα καταλήξει στην πορτοκαλί ή στη ροζ περιοχή. Από την άλλη, αν δοκιμάζουμε τις δυνατότητές μας για πρώτη φορά είναι πιο πιθανό το βέλος να καταλήξει στην κίτρινη περιοχή, η οποία καταλαμβάνει και το μεγαλύτερο μέρος του στόχου. Λέμε ότι τα δυνατά αποτελέσματα δεν είναι ισοπίθανα.

Ένας τρόπος για να αποδώσουμε πιθανότητα σε κάθε δυνατό αποτέλεσμα ενός τέτοιου πειράματος είναι να επαναλάβουμε το πείραμα πολλές φορές και να καταγράψουμε τη σχετική συχνότητα εμφάνισης κάθε αποτελέσματος. Μετά από πολλές επαναλήψεις του πειράματος αναμένουμε, από τον Νόμο των μεγάλων αριθμών, οι σχετικές συχνότητες να έχουν σταθεροποιηθεί γύρω από κάποιους αριθμούς. Καθένας από αυτούς τους αριθμούς θα εκφράζει και την (θεωρητική) πιθανότητα εμφάνισης του αντίστοιχου απλού ενδεχομένου.

Παράδειγμα 1

Η ρίψη ενός μη συμμετρικού (μεροληπτικού) νομίσματος

Ρίχνουμε ένα μη συμμετρικό, με ακανόνιστο σχήμα, νόμισμα και καταγράφουμε την ένδειξή του. Ο δειγματοχώρος του πειράματος είναι ο $\Omega=\{\text{Κ}, \Gamma\}$.
Είναι φανερό ότι δεν μπορούμε να υποθέσουμε εκ των προτέρων ότι τα δυνατά αποτελέσματα Κ και Γ έχουν την ίδια πιθανότητα εμφάνισης.

Για να διαπιστώσουμε ποια είναι η πιθανότητα εμφάνισης κεφαλής και γραμμάτων για το συγκεκριμένο νόμισμα εκτελέσαμε τέσσερες φορές το πείραμα από 1000 ρίψεις την κάθε φορά και πήραμε τα παρακάτω αποτελέσματα:

Στη συνέχεια του βιβλίου για ένα απλό ενδεχόμενο $\{\omega\}$ θα γράφουμε $P(\omega)$ αντί του $P(\{\omega\})$ και $\nu(\omega)$ αντί του $\nu(\{\omega\})$

	1η φορά	2η φορά	3η φορά	4η φορά
Κεφαλή	695	705	720	698
Γράμματα	305	295	280	302
Σύνολο	1000	1000	1000	1000

Παρατηρούμε ότι η σχετική συχνότητα εμφάνισης της κεφαλής πλησιάζει τον αριθμό 0,7 ενώ αυτή των γραμμάτων το 0,3. Είναι λοιπόν λογικό να αποδώσουμε τις εξής πιθανότητες: $P(K)=0{,}7$ και $P(\Gamma)=0{,}3$.

Προσπαθήστε να λύσετε την άσκηση 1

Μη ισοπίθανα απλά ενδεχόμενα

Εφαρμογή 1

Διαφοροποίηση του δειγματοχώρου ανάλογα με το τι θεωρείται αποτέλεσμα του πειράματος από διαφορετικούς παρατηρητές.

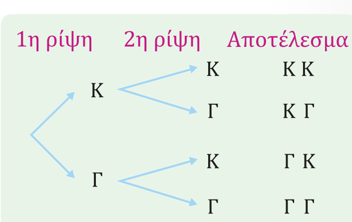
Η Μαίρη θα ρίξει ένα αμερόληπτο νόμισμα δύο φορές. Αφού εκτελέσει το πείραμα θα ανακοινώσει στη Ναταλία κάποιον από τους αριθμούς 0, 1 ή 2 ανάλογα με το αν οι δύο ενδείξεις περιέχουν καμία, μία ή δύο κεφαλές αντίστοιχα.

α. Ποιος είναι ο δειγματοχώρος του πειράματος για την Μαίρη;
β. Ποιος είναι ο δειγματοχώρος του πειράματος για την Ναταλία;
γ. Είναι τα απλά ενδεχόμενα του δειγματοχώρου της Ναταλίας ισοπίθανα;
δ. Ποιες πιθανότητες είναι λογικό να αποδώσουμε σε κάθε απλό ενδεχόμενο των δύο δειγματοχώρων;

Λύση

α. Τα αποτελέσματα που παρατήρησε η Μαίρη είναι τα ΚΚ, ΚΓ, ΓΚ, ΓΓ.
Ο δειγματοχώρος για την Μαίρη είναι το σύνολο:

$$\Omega=\{KK, K\Gamma, \Gamma K, \Gamma\Gamma\}$$



β. Είναι σαφές ότι για τη Ναταλία, ο δειγματοχώρος του πειράματος είναι το σύνολο $\Omega'=\{0,1,2\}$
αφού η μόνη γνώση που μπορεί να έχει είναι το πόσες φορές εμφανίστηκε η κεφαλή.

γ. Τα απλά ενδεχόμενα {0}, {1}, {2} του δειγματοχώρου της Ναταλίας δεν είναι ισοπίθανα. Στο ενδεχόμενο {1} αντιστοιχούν δύο ισοπίθανα δυνατά αποτελέσματα (ΚΓ και ΓΚ) ενώ στα ενδεχόμενα {0} και {2} αντιστοιχούν τα αποτελέσματα ΓΓ και ΚΚ αντίστοιχα.

δ. Επειδή το ζάρι είναι αμερόληπτο η πιθανότητα κάθε απλού ενδεχομένου ω του δειγματοχώρου Ω της Μαίρης είναι $P(\omega)=\frac{1}{4}$.

Αγώνας δρόμου με ζάρια Ι

Για τις πιθανότητες των απλών ενδεχομένων του δειγματοχώρου της Ναταλίας, είναι λογικό να αποδώσουμε τις παρακάτω πιθανότητες:

(Αποτέλεσμα ΚΚ)	(Αποτελέσματα ΚΓ, ΓΚ)	(Αποτέλεσμα ΓΓ)
$P(0)=\frac{1}{4}$	$P(1)=\frac{2}{4}=\frac{1}{2}$	$P(2)=\frac{1}{4}$

Προσπαθήστε να λύσετε τις ασκήσεις 2 έως 6

Πίνακας 1

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

Πίνακας 2

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

Ρίψη δύο αμερόληπτων ζαριών

Αγώνας δρόμου με ζάρια ΙΙ

Εφαρμογή 2

Η ρίψη δύο αμερόληπτων ζαριών

Μια οθόνη έχει ενσωματωμένο έναν υπολογιστή ο οποίος έχει τη δυνατότητα να προσομοιώνει το πείραμα ρίψης δύο αμερόληπτων ζαριών. Αφού εκτελεστεί η προσομοίωση θα εμφανιστεί στην οθόνη το άθροισμα των δύο ενδείξεων. Μπροστά από την οθόνη βρίσκεται ένας παρατηρητής ο οποίος καταγράφει το αποτέλεσμα που εμφανίζεται στην οθόνη.

α. Να βρεθεί ο δειγματοχώρος του πειράματος για τον παρατηρητή.

β. Ποιο αποτέλεσμα είναι πιο πιθανό να παρατηρήσει ο παρατηρητής και ποια είναι η πιθανότητά του;

Λύση

α. Ο παρατηρητής μπορεί να δει στη οθόνη μόνο το άθροισμα των δύο ενδείξεων και όχι την ένδειξη κάθε ζαριού.
Κατασκευάζουμε έναν πίνακα διπλής εισόδου (Πίνακας 1) και σε κάθε κελί σημειώνουμε το άθροισμα των αντίστοιχων ενδείξεων. Έτσι παίρνουμε τον διπλανό πίνακα. Βλέπουμε ότι οι δυνατές τιμές του αθροίσματος είναι οι φυσικοί αριθμοί από το 2 μέχρι το 12, οπότε ο δειγματοχώρος Ω του πειράματος για τον παρατηρητή είναι

$$\Omega = \{2, 3, 4, \dots, 12\}$$

β. Για ευκολία στην οπτική παρατήρηση, χρωματίσαμε τα κελιά του Πίνακα 1 όπως φαίνεται στον Πίνακα 2. Παρατηρούμε ότι το άθροισμα 7 εμφανίζεται τις περισσότερες φορές (6 φορές) και επομένως ο αριθμός 7 είναι το πιο πιθανό αποτέλεσμα. Ο δειγματοχώρος του πειράματος ρίψης δύο ζαριών, ο οποίος συγκροτείται από ισοπίθανα απλά ενδεχόμενα είναι ο

$$\Omega' = \left\{\begin{matrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{matrix}\right\}$$

Το ενδεχόμενο να πάρουμε ένδειξη 7 είναι το

$$A_7 = \{(1,6),\ (2,5),\ (3,4),\ (4,3),\ (5,2),\ (6,1)\}$$

Επομένως $$P(A_7) = \frac{\nu(A_7)}{\nu(\Omega')} = \frac{6}{36} = \frac{1}{6}$$

Προσπαθήστε να λύσετε την άσκηση 7

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Ρίχνουμε κατακόρυφα και προς τα πάνω μια πινέζα όπως αυτές του σχήματος.
Θεωρούμε ότι το πείραμα έχει δύο δυνατά αποτελέσματα ανάλογα με τον τρόπο που η πινέζα θα καταλήξει στο έδαφος:

- με τη μύτη προς τα πάνω ή
- με τη μύτη προς το έδαφος.

α. Θεωρείτε ότι τα δύο δυνατά αποτελέσματα είναι ισοπίθανα;
β. Να προτείνετε μια μέθοδο για την εκτίμηση της πιθανότητας εμφάνισης κάθε δυνατού αποτελέσματος.

2 Σε ποια από τα παρακάτω πειράματα τύχης τα δυνατά αποτελέσματα είναι ισοπίθανα και σε ποια όχι;
α. Ένα αμερόληπτο νόμισμα ρίχνεται δύο φορές. Ως δυνατά αποτελέσματα θεωρούμε τους αριθμούς 0, 1, 2 ανάλογα με το αν έρθουν καμία, μία ή δύο φορές γράμματα, αντίστοιχα.
β. Από τα 52 φύλλα της τράπουλας επιλέγεται ένα στην τύχη. Ως δυνατά αποτελέσματα θεωρούμε τον αριθμό 1, αν το φύλλο είναι φιγούρα, και τον αριθμό 0, αν είναι άσσος ή αριθμός.
γ. Ρίχνουμε ένα βέλος σε κυκλικό στόχο ο οποίος έχει χωριστεί σε δύο ημικύκλια, τα Α και Β. Ως δυνατά αποτελέσματα θεωρούμε τον αριθμό 1, αν το βέλος καταλήξει στο ημικύκλιο Α, και τον αριθμό 0, αν καταλήξει στο ημικύκλιο Β.
δ. Ένα γράμμα από το ελληνικό αλφάβητο επιλέγεται τυχαία. Ως δυνατά αποτελέσματα θεωρούμε τον αριθμό 1, αν το γράμμα είναι φωνήεν, και τον αριθμό 0, αν είναι σύμφωνο.

3 **Το σφάλμα του D'Alembert.** Για τον υπολογισμό της πιθανότητας να εμφανιστεί τουλάχιστον μία κεφαλή σε δύο ρίψεις ενός νομίσματος ο D'Alembert πρότεινε την ακόλουθη λύση:

Ο δειγματοχώρος του πειράματος είναι το σύνολο $\Omega=\{0,1,2\}$ αφού τα δυνατά αποτελέσματα στις δύο ρίψεις του νομίσματος είναι 0 κεφαλές ή 1 κεφαλή ή 2 κεφαλές. Το ενδεχόμενο να εμφανιστεί τουλάχιστον μία κεφαλή στις δύο ρίψεις του νομίσματος είναι το $A=\{1,2\}$. Επειδή τα δυνατά αποτελέσματα είναι 3 και οι ευνοϊκές περιπτώσεις του Α είναι 2, η πιθανότητα εμφάνισης του Α θα είναι:

$$P(A)=\frac{\nu(A)}{\nu(\Omega)}=\frac{2}{3}$$

Να εξηγήσετε γιατί η παραπάνω λύση είναι λανθασμένη.

Ασκήσεις

4 Επιλέγουμε στην τύχη ένα από τα δέκα γράμματα της λέξης «ΠΙΘΑΝΟΤΗΤΑ» και το καταγράφουμε.
α. Ποιος είναι ο δειγματοχώρος του πειράματος; Είναι τα απλά ενδεχόμενα ισοπίθανα;
β. Να υπολογίσετε την πιθανότητα εμφάνισης των γραμμάτων Π, Τ, Α.

5 Ρίχνουμε μία φορά ένα δίκαιο, δηλαδή αμερόληπτο ζάρι, και ως αποτέλεσμα θεωρούμε την λέξη «επιτυχία» αν η ένδειξη διαιρείται με το 3 και τη λέξη «αποτυχία» διαφορετικά.
α. Ποιος είναι ο δειγματοχώρος του πειράματος τύχης;
β. Ποια είναι η πιθανότητα κάθε δυνατού αποτελέσματος;

6 Ο Αλέξης ρίχνει μία φορά ένα συμμετρικό ζάρι του οποίου οι έδρες με αριθμούς 1, 2, 3, 4 έχουν χρωματιστεί κόκκινες ενώ οι υπόλοιπες πράσινες. Αφού ρίξει το ζάρι ανακοινώνει το χρώμα της έδρας που ήρθε στον Αντρέα.
α. Ποιος είναι ο δειγματοχώρος του πειράματος για τον Αντρέα; Είναι τα απλά ενδεχόμενα ισοπίθανα;
β. Να υπολογίσετε την πιθανότητα κάθε απλού ενδεχομένου του δειγματοχώρου του Αντρέα.

7 Ρίχνουμε 2 αμερόληπτα ζάρια, από μία φορά το καθένα, και καταγράφουμε τη διαφορά των ενδείξεων. Ποιο αποτέλεσμα είναι πιο πιθανό να εμφανιστεί;

4.4 Αξιωματικός ορισμός πιθανότητας

Περιέχονται:
- Αξιωματικός ορισμός της πιθανότητας
- Ο Κλασικός ορισμός ως ειδική περίπτωση του Αξιωματικού ορισμού

Λέξεις κλειδιά:
Αξιωματικός ορισμός πιθανότητας

Στον πίνακα με τις πιθανότητες του νόμου του Benford

α. Παρατηρήστε ότι όλες οι πιθανότητες ανήκουν στο κλειστό διάστημα [0,1] και

β. το άθροισμά τους ισούται με 1.

Νόμος του Benford

Στην ενότητα αυτή θα μάθουμε:
- Να διατυπώνουμε τον αξιωματικό ορισμό της πιθανότητας για έναν πεπερασμένο δειγματοχώρο.
- Να αναγνωρίζουμε διαφορές και συνδέσεις μεταξύ του αξιωματικού και του κλασικού ορισμού της πιθανότητας.

Δραστηριότητα

Σύμφωνα με τον **νόμο του Benford** (Benford's law) το πρώτο ψηφίο κάθε αριθμού από αρκετά σύνολα αριθμητικών δεδομένων μπορεί να είναι κάποιος από τους φυσικούς αριθμούς 1 έως 9 με αντίστοιχες πιθανότητες που φαίνονται στον παρακάτω πίνακα:

Πρώτο ψηφίο	1	2	3	4	5	6	7	8	9
Πιθανότητα	0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,051	0,046

Ας υποθέσουμε ότι επιλέγετε τυχαία έναν αριθμό, από ένα τέτοιο σύνολο δεδομένων. Τι πιθανότητα θα αποδώσετε στα παρακάτω ενδεχόμενα;

Α = {το πρώτο ψηφίο του αριθμού είναι άρτιος}

Β = {το πρώτο ψηφίο του αριθμού είναι μικρότερο του 3}

Αξιωματικός ορισμός της πιθανότητας

Σε πολλά πειράματα τύχης, τα δυνατά αποτελέσματα δεν είναι ισοπίθανα, γεγονός που καθιστά ανεπαρκή τον κλασικό ορισμό της πιθανότητας. Ένα χαρακτηριστικό παράδειγμα είναι ο νόμος του Benford, όπου τα πρώτα ψηφία αριθμητικών δεδομένων έχουν άνισες πιθανότητες εμφάνισης.

Παρατηρήσεις όπως η προηγούμενη αναδεικνύουν την ανάγκη για μια πιο ευέλικτη θεώρηση της πιθανότητας. Είναι γεγονός ότι ο κλασικός ορισμός της πιθανότητας δεν μπορεί να καλύψει όλες τις περιπτώσεις αβεβαιότητας. Τέτοιες καταστάσεις απαιτούν έναν γενικότερο ορισμό για να περιγραφούν σωστά οι πιθανότητες των ενδεχομένων. Ο αξιωματικός ορισμός της πιθανότητας παρέχει αυτό το εργαλείο, αντιστοιχίζοντας κάθε ενδεχόμενο ενός πειράματος τύχης σε έναν πραγματικό αριθμό του διαστήματος [0, 1], ο οποίος εκφράζει το μέτρο της ''προσδοκίας'' με την οποία αναμένουμε την πραγματοποίησή του.

Πριν διατυπώσουμε έναν τέτοιο ορισμό, ας δούμε τι απαιτήσεις πρέπει να ικανοποιεί:

• Πρέπει να εφαρμόζεται σε πειράματα, όπως αυτό της δραστηριότητας, στα οποία τα απλά ενδεχόμενα δεν είναι ισοπίθανα.

• Πρέπει να εφαρμόζεται και σε πειράματα με ισοπίθανα απλά ενδεχόμενα. Αυτό σημαίνει ότι ο κλασικός ορισμός θα είναι ειδική περίπτωσή του.

• Πρέπει να είναι συμβατός με τον Νόμο των μεγάλων αριθμών.

Και οι τρεις απαιτήσεις ικανοποιούνται από τον παρακάτω ορισμό.

Ο αξιωματικός ορισμός της πιθανότητας δόθηκε από τον Andrei Kolmogorov το 1933.

Ορισμός (Αξιωματικός ορισμός της πιθανότητας)
Έστω Ω ένας πεπερασμένος δειγματοχώρος. Δεχόμαστε τα επόμενα αξιώματα:
1. Η πιθανότητα ενός ενδεχομένου A του Ω συμβολίζεται $P(A)$ και είναι ένας μη αρνητικός πραγματικός αριθμός, δηλαδή $P(A)\geq 0$.
2. Για τον δειγματοχώρο Ω ισχύει $P(\Omega)=1$.
3. Αν A και B είναι δύο, οποιαδήποτε, ασυμβίβαστα ενδεχόμενα του Ω τότε
$$P(A\cup B)=P(A)+P(B).$$

Ο ορισμός αυτός αποτελεί στην πραγματικότητα μια αξιωματική θεμελίωση της έννοιας της πιθανότητας και μπορεί εδώ να αναφερόμαστε αποκλειστικά για πεπερασμένους δειγματοχώρους, αλλά ο ορισμός μπορεί να επεκταθεί και σε μη πεπερασμένους. Μερικά παραδείγματα πειραμάτων τύχης με μη πεπερασμένο δειγματοχώρο είναι τα επόμενα: η επιλογή ενός πραγματικού αριθμού από το διάστημα [0, 1], η επιλογή ενός σημείου σε έναν κυκλικό δίσκο, η ρίψη ενός ζαριού μέχρι να φέρουμε εξάρι κ.α.

Τρία ή περισσότερα ενδεχόμενα λέμε ότι είναι ανά δύο ξένα, όταν δύο οποιαδήποτε από αυτά είναι ξένα μεταξύ τους.

Σημείωση 1: Το τρίτο αξίωμα ονομάζεται Απλός Προσθετικός Νόμος των πιθανοτήτων και αποδεικνύεται ότι ισχύει για περισσότερα από δύο ενδεχόμενα τα οποία είναι ανά δύο ξένα μεταξύ τους. Για παράδειγμα, αν τα ενδεχόμενα A, B και Γ είναι ανά δύο ξένα, τότε $P(A\cup B\cup \Gamma)=P(A)+P(B)+P(\Gamma)$.

Σημείωση 2: Θα αποδείξουμε στην επόμενη παράγραφο ότι $P(A)\leq 1$ για κάθε ενδεχόμενο A του Ω.
Άμεση συνέπεια του αξιωματικού ορισμού είναι η παρακάτω πρόταση.

Πρόταση: Έστω $\Omega=\{\omega_1,\omega_2,...,\omega_\nu\}$ ένας πεπερασμένος δειγματοχώρος.
1. Η πιθανότητα του κενού συνόλου είναι μηδέν: $P(\varnothing)=0$.
2. Για οποιοδήποτε ενδεχόμενο $A=\{\alpha_1,\alpha_2,...,\alpha_\kappa\}\neq\varnothing$ του Ω, ισχύει:
$$P(A)=P(\alpha_1)+P(\alpha_2)+...+P(\alpha_\kappa).$$
3. Το άθροισμα των πιθανοτήτων των απλών ενδεχομένων του Ω είναι 1:
$$P(\omega_1)+P(\omega_2)+...+P(\omega_\nu)=1.$$

Απόδειξη
1. Είναι $\varnothing\cap\varnothing=\varnothing$, δηλαδή το κενό σύνολο είναι ασυμβίβαστο με τον εαυτό του. Επίσης $\varnothing\cup\varnothing=\varnothing$. Από το τρίτο αξίωμα έχουμε:
$$P(\varnothing\cup\varnothing)=P(\varnothing) \text{ ή } P(\varnothing)+P(\varnothing)=P(\varnothing) \text{ ή } P(\varnothing)=0$$
2. Από τη γενίκευση του απλού προσθετικού νόμου, επειδή τα απλά ενδεχόμενα είναι ανά δύο ξένα, έχουμε:
$$P(A)=P(\{\alpha_1,\alpha_2,...,\alpha_\kappa\})=P(\{\alpha_1\}\cup\{\alpha_2\}\cup...\cup\{\alpha_\kappa\})=P(\alpha_1)+P(\alpha_2)+...+P(\alpha_\kappa).$$
3. Από το προηγούμενο αποτέλεσμα για το $\Omega=\{\omega_1,\omega_2,\dots,\omega_\nu\}$ έχουμε:
$$P(\Omega)=P(\omega_1)+P(\omega_2)+...+P(\omega_\nu)$$
και επειδή $P(\Omega)=1$, συμπεραίνουμε ότι: $P(\omega_1)+P(\omega_2)+...+P(\omega_\nu)=1$

Αξιωματικός ορισμός πιθανότητας

Παράδειγμα

Ένας μαθητής αντέγραψε στο τετράδιό του τον παρακάτω πίνακα που αφορά ένα πείραμα τύχης με τέσσερα δυνατά αποτελέσματα. Ξέχασε, όμως, να συμπληρώσει την πιθανότητα που αντιστοιχεί στο αποτέλεσμα 4. Μπορείτε να τον βοηθήσετε να την υπολογίσει;

Αποτέλεσμα	1	2	3	4
Πιθανότητα	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{4}$	

Λύση

Έστω p η πιθανότητα που αντιστοιχεί στο αποτέλεσμα 4. Επειδή το άθροισμα των πιθανοτήτων όλων των απλών ενδεχομένων ενός πειράματος τύχης είναι 1, έχουμε:

$$P(1)+P(2)+P(3)+P(4)=1 \quad \text{ή} \quad \frac{1}{8}+\frac{1}{12}+\frac{1}{4}+p=1$$

Άρα

$$\frac{11}{24}+p=1 \quad \text{ή} \quad p=\frac{13}{24}$$

Προσπαθήστε να λύσετε τις ασκήσεις 1 έως 7

Αξιωματικός ορισμός πιθανότητας

Εφαρμογή 1

Έστω $\Omega=\{\alpha,\beta,\gamma\}$ ο δειγματοχώρος ενός πειράματος τύχης και τα ενδεχόμενα

$$A=\{\alpha,\beta\},\ B=\{\beta,\gamma\} \text{ και } \Gamma=\{\alpha,\gamma\}.$$

Αν $P(A)=0{,}8$ και $P(B)=0{,}5$, τότε:

α. Να υπολογίσετε τις πιθανότητες $P(\alpha)$, $P(\beta)$ και $P(\gamma)$.

β. Να υπολογίσετε την πιθανότητα $P(\Gamma)$.

Λύση

α. Επειδή το άθροισμα των πιθανοτήτων των απλών ενδεχομένων είναι 1, έχουμε

$$P(\alpha)+P(\beta)+P(\gamma)=1 \qquad (1)$$

Επειδή η πιθανότητα κάθε ενδεχομένου είναι ίση με το άθροισμα των πιθανοτήτων των απλών ενδεχομένων είναι:

$$P(A)=P(\alpha)+P(\beta) \text{ ή } P(\alpha)+P(\beta)=0{,}8 \qquad (2)$$

$$P(B)=P(\beta)+P(\gamma) \text{ ή } P(\beta)+P(\gamma)=0{,}5 \qquad (3)$$

Η (1) λόγω της (2) δίνει: $0{,}8+P(\gamma)=1$ ή $P(\gamma)=0{,}2$

Η (1) λόγω της (3) δίνει: $0{,}5+P(\alpha)=1$ ή $P(\alpha)=0{,}5$

Τέλος, από την (1) βρίσκουμε: $P(\beta)=1-0{,}5-0{,}2=0{,}3$

β. Σύμφωνα με τα παραπάνω είναι $P(\Gamma)=P(\alpha)+P(\gamma)=0{,}5+0{,}2=0{,}7$.

Προσπαθήστε να λύσετε την άσκηση 8

Εφαρμογή 2

Ένα κατάστημα πώλησης κινητών τηλεφώνων παρέχει εγγύηση στα προϊόντα του για ένα έτος. Ο ιδιοκτήτης του καταστήματος έχει παρατηρήσει ότι στη διάρκεια ενός έτους από τη στιγμή της πώλησης, το 15% των τηλεφώνων χρειάζεται μία επισκευή, το 8% χρειάζεται δύο επισκευές ενώ το 1% χρειάζεται τρεις ή παραπάνω επισκευές (3+). Αν αγοράσουμε ένα τηλέφωνο από το συγκεκριμένο κατάστημα, να βρεθεί η πιθανότητα το τηλέφωνο:

α. Να χρειαστεί τουλάχιστον μία επισκευή εντός εγγύησης.
β. Να μην χρειαστεί επισκευή εντός εγγύησης.

Λύση

Ο δειγματοχώρος του πειράματος είναι το σύνολο $\Omega = \{0, 1, 2, 3+\}$.
Τα δεδομένα του προβλήματος παρουσιάζονται στον επόμενο πίνακα:

Αριθμός επισκευών	1	2	3+
Πιθανότητα	0,15	0,08	0,01

α. Ας συμβολίσουμε με A το ενδεχόμενο «το τηλέφωνο να χρειαστεί τουλάχιστον μία επισκευή». Είναι $A = \{1, 2, 3+\}$ και

$$P(A) = P(1) + P(2) + P(3+) = 0{,}15 + 0{,}08 + 0{,}01 = 0{,}24$$

β. Το ενδεχόμενο «Το τηλέφωνο να μην χρειαστεί επισκευή εντός εγγύησης» είναι το $\{0\}$. Επίσης,

$$P(0) + P(1) + P(2) + P(3+) = 1$$

ή

$$P(0) = 1 - P(1) - P(2) - P(3+) = 1 - 0{,}15 - 0{,}08 - 0{,}01 = 0{,}76$$

Το άθροισμα των πιθανοτήτων όλων των απλών ενδεχομένων ενός πειράματος τύχης είναι 1.

Προσπαθήστε να λύσετε την άσκηση 9

Εφαρμογή 3

Σε έναν αγώνα δρόμου συμμετέχουν τρεις αθλητές, ο Α, ο Β και ο Γ. Ο αθλητής Α έχει διπλάσια πιθανότητα από τον Β να κερδίσει το αγώνα και ο Β τριπλάσια από τον Γ. Ποια είναι η πιθανότητα κάθε αθλητή να κερδίσει τον αγώνα;

Λύση

Συμβολίζουμε p_A, p_B, p_Γ τις πιθανότητες νίκης των αθλητών Α, Β και Γ, αντίστοιχα. Ισχύει:

$$p_A + p_B + p_\Gamma = 1 \qquad (1)$$

Από τα δεδομένα του προβλήματος είναι:

$$p_A = 2p_B \quad \text{και} \quad p_B = 3p_\Gamma \qquad (2)$$

Από τις δύο τελευταίες ισότητες συμπεραίνουμε ότι $p_A = 6p_\Gamma$. Αντικαθιστούμε στη σχέση (1) και παίρνουμε:

$$6p_\Gamma + 3p_\Gamma + p_\Gamma = 1 \quad \text{ή} \quad p_\Gamma = \frac{1}{10} \qquad (3)$$

Τέλος, από τις σχέσεις (2) και (3) βρίσκουμε:

$$p_B = 3p_\Gamma = \frac{3}{10} \quad \text{και} \quad p_A = 6p_\Gamma = \frac{3}{5}$$

Προσπαθήστε να λύσετε τις ασκήσεις 10, 11 και 12

Αξιωματικός ορισμός πιθανότητας

Ο Κλασικός ορισμός ως ειδική περίπτωση του Αξιωματικού ορισμού

Έστω ότι ο δειγματοχώρος Ω ενός πειράματος τύχης αποτελείται από ν ισοπίθανα δυνατά αποτελέσματα: $\Omega=\{\omega_1,\omega_2,...,\omega_\nu\}$.

Έστω $A=\{\alpha_1,\alpha_2,...,\alpha_\kappa\}\neq\varnothing$ ένα ενδεχόμενο του Ω. Θα αποδείξουμε ότι η πιθανότητα $P(A)$ δίνεται από τη σχέση:

$$P(A)=\frac{\nu(A)}{\nu(\Omega)}$$

δηλαδή αυτή του κλασικού ορισμού. Πράγματι, είναι

$$P(A)=P(\alpha_1)+P(\alpha_2)+...+P(\alpha_\kappa)$$

Οι πιθανότητες των απλών ενδεχομένων είναι ίσες με $\frac{1}{\nu}$.

Επομένως

$$P(A)=\underbrace{\frac{1}{\nu}+\frac{1}{\nu}+...+\frac{1}{\nu}}_{\kappa\ \text{φορές}}=\frac{\kappa}{\nu}\quad \text{ή}\quad P(A)=\frac{\nu(A)}{\nu(\Omega)}$$

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Να χαρακτηρίσετε καθεμία από τις επόμενες προτάσεις, ως αληθή ή ψευδή:

α. Η πιθανότητα ενός ενδεχομένου A μπορεί να είναι αρνητικός αριθμός.

β. Η πιθανότητα ενός ενδεχομένου A μπορεί να είναι αριθμός μεγαλύτερος του 1.

γ. Το άθροισμα των πιθανοτήτων όλων των απλών ενδεχομένων ενός δειγματοχώρου είναι ίσο με 1.

δ. Αν το σύνολο A είναι ενδεχόμενο ενός δειγματοχώρου Ω, τότε $P(A)\leq P(\Omega)$.

ε. Για κάθε ενδεχόμενο A ισχύει $P(A)\geq P(\varnothing)$.

2 Αν τα ενδεχόμενα A και B ενός δειγματοχώρου είναι ασυμβίβαστα μπορεί να ισχύει: $P(A)=0{,}3$ και $P(B)=0{,}8$; Να αιτιολογήσετε την απάντησή σας.

3 Ένα πείραμα τύχης έχει δειγματοχώρο $\Omega=\{0,1,2\}$. Η Αγγελική, μαθήτρια της Α΄Λυκείου, κατασκεύασε τον παρακάτω πίνακα στον οποίο φαίνονται οι πιθανότητες που απέδωσε σε καθένα από τα δυνατά αποτελέσματα του πειράματος. Συμφωνεί η απόδοση πιθανοτήτων της Αγγελικής με τον αξιωματικό ορισμό;

Αποτέλεσμα	0	1	2
Πιθανότητα	32%	45%	23%

Ασκήσεις

4 Ρίχνουμε μία φορά ένα μη συμμετρικό νόμισμα. Αν η πιθανότητα να φέρουμε κεφαλή είναι 55% να υπολογίσετε την πιθανότητα να φέρουμε γράμματα.

5 Ρίχνουμε μία φορά ένα μη συμμετρικό ζάρι και καταγράφουμε το αποτέλεσμα. Τα δυνατά αποτελέσματα και οι αντίστοιχες πιθανότητες παρουσιάζονται στον παρακάτω πίνακα. Να υπολογίσετε την πιθανότητα p.

Αποτέλεσμα	1	2	3	4	5	6
Πιθανότητα	5%	p	24%	10%	50%	7%

6 Έστω $\Omega=\{\omega_1,\omega_2,\omega_3\}$ ο δειγματοχώρος ενός πειράματος τύχης. Αν

$$P(\omega_1)=0{,}6,\ \ P(\omega_2)=0{,}3,\ \ P(\omega_3)=0{,}1,$$

να υπολογίσετε τις πιθανότητες των ενδεχομένων:

$$A=\{\omega_1,\omega_2\},\ \ B=\{\omega_2,\omega_3\},\ \ \Gamma=\{\omega_1,\omega_3\}$$

7 Έστω $\Omega = \{\omega_1, \omega_2, \omega_3\}$ ο δειγματοχώρος ενός πειράματος τύχης. Αν $P(\omega_1) = \frac{1}{10}$, να υπολογίσετε την πιθανότητα του ενδεχομένου $A = \{\omega_2, \omega_3\}$.

8 Έστω $\Omega = \{\omega_1, \omega_2, \omega_3\}$ ο δειγματοχώρος ενός πειράματος τύχης. Θεωρούμε τα ενδεχόμενά του

$$A = \{\omega_1, \omega_2\} \text{ και } B = \{\omega_2, \omega_3\}$$

με

$$P(A) = \frac{2}{3} \text{ και } P(B) = \frac{3}{5}.$$

Να υπολογίσετε τις πιθανότητες

$$P(\omega_1),\ P(\omega_2) \text{ και } P(\omega_3).$$

9 Ρωτήσαμε 50 μαθητές ενός σχολείου πόσες ώρες αφιερώνουν καθημερινά στη μελέτη των σχολικών τους μαθημάτων. Οι απαντήσεις που πήραμε φαίνονται στον παρακάτω πίνακα.

Ώρες μελέτης	[0,1)	[1,2)	[2,3)	[3,4]
Αριθμός μαθητών	5	30	10	5

Αν επιλέξουμε τυχαία έναν μαθητή, να υπολογίσετε την πιθανότητα:

α. Να μελετάει τουλάχιστον 1 ώρα καθημερινά.

β. Να μελετάει λιγότερο από 2 ώρες καθημερινά.

γ. Να μελετάει τουλάχιστον 1 και λιγότερο από 3 ώρες καθημερινά.

10 Έστω $\Omega = \{\omega_1, \omega_2, \omega_3\}$ ο δειγματοχώρος ενός πειράματος τύχης με $P(\omega_1) = 2P(\omega_2) = 4P(\omega_3)$. Να υπολογίσετε τις πιθανότητες $P(\omega_1)$, $P(\omega_2)$, $P(\omega_3)$.

11 Ρίχνουμε μία φορά ένα μη συμμετρικό ζάρι και καταγράφουμε το αποτέλεσμα. Τα δυνατά αποτελέσματα και οι αντίστοιχες πιθανότητες παρουσιάζονται στον παρακάτω πίνακα.
Η πιθανότητα να φέρει το ζάρι 3 είναι διπλάσια της πιθανότητας να φέρει 5.

Αποτέλεσμα	1	2	3	4	5	6
Πιθανότητα	$\frac{1}{10}$	$\frac{1}{5}$	p	$\frac{3}{10}$	q	$\frac{1}{10}$

α. Να υπολογίσετε τις πιθανότητες p και q.

β. Τι είναι πιο πιθανό να φέρει το ζάρι, άρτιο ή περιττό αριθμό;

12 Ένα πείραμα τύχης έχει δυνατά αποτελέσματα τους αριθμούς 1, 2, 3 και 4. Οι πιθανότητες εμφάνισης κάθε αριθμού δίνονται στον παρακάτω πίνακα:

Αποτέλεσμα	1	2	3	4
Πιθανότητα	20%	40%	p	q

Ο Ιάσονας, μαθητής της Α΄Λυκείου, θέλει να βρει δύο αριθμούς p και q τέτοιους, ώστε η πιθανότητα το αποτέλεσμα να είναι περιττός να είναι διπλάσια της πιθανότητας να είναι άρτιος. Είναι αυτό δυνατό;

4.5 Λογισμός πιθανοτήτων

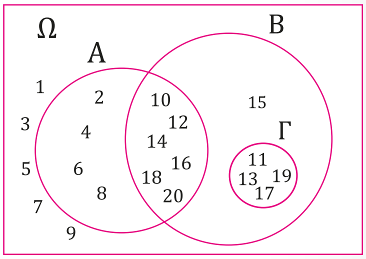
Στην ενότητα αυτή θα μάθουμε:

- Να διατυπώνουμε υποθέσεις για τους κανόνες που αναμένεται να ισχύουν στο λογισμό πιθανοτήτων.
- Να αποδεικνύουμε τους κανόνες λογισμού πιθανοτήτων στα πλαίσια του αξιωματικού ορισμού της πιθανότητας.

Περιέχονται:

- Κανόνες λογισμού πιθανοτήτων
- Εφαρμογές των κανόνων λογισμού πιθανοτήτων

Λέξεις κλειδιά:
Κανόνες λογισμού πιθανοτήτων



Κανόνες λογισμού πιθανοτήτων

Δραστηριότητα

Στο διπλανό διάγραμμα Venn φαίνεται ο δειγματοχώρος

$$\Omega=\{x\in\mathbb{N}\,|\,1\le x\le 20\}$$

ενός πειράματος τύχης με ισοπίθανα απλά ενδεχόμενα και τα σύνθετα ενδεχόμενά του

$$A=\{x\in\Omega\,|\,x \text{ άρτιος}\},$$
$$B=\{x\in\Omega\,|\,10\le x\le 20\} \text{ και}$$
$$\Gamma=\{x\in\Omega\,|\,x \text{ πρώτος μεγαλύτερος του } 10\}.$$

Α. Παρατηρώντας το διάγραμμα να επαληθεύσετε τις παρακάτω ισότητες:

α. $\nu(A)+\nu(A')=\nu(\Omega)$

β. $\nu(A\cup B)=\nu(A)+\nu(B)-\nu(A\cap B)$

γ. $\nu(A-B)=\nu(A)-\nu(A\cap B)$

Β. Βασιζόμενοι στις παραπάνω σχέσεις και στον κλασικό ορισμό της πιθανότητας να διερευνήσετε την αλήθεια των παρακάτω εικασιών:

δ. $P(A)+P(A')=1$

ε. $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

στ. $P(A-B)=P(A)-P(A\cap B)$

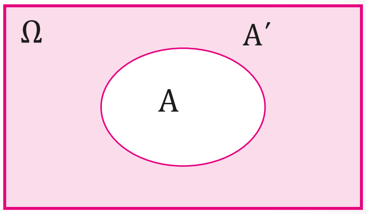
Να εξετάσετε αν οι παραπάνω ισότητες ισχύουν για οποιαδήποτε ενδεχόμενα ενός πεπερασμένου δειγματοχώρου.

Κανόνες λογισμού πιθανοτήτων

Τα συμπεράσματα, τα οποία προκύπτουν από τη δραστηριότητα, μπορούν να γενικευθούν σε κάθε πεπερασμένο δειγματοχώρο με ισοπίθανα ή μη ισοπίθανα δυνατά αποτελέσματα. Συγκεκριμένα ισχύουν οι παρακάτω κανόνες:

Κανόνας 1. Για οποιοδήποτε ενδεχόμενο A ενός δειγματοχώρου Ω ισχύει:

$$P(A)+P(A')=1$$



Απόδειξη

Τα ενδεχόμενα A και A' είναι ασυμβίβαστα και για την ένωσή τους ισχύει: $A\cup A'=\Omega$. Επομένως, από τον απλό προσθετικό νόμο ($3^{ο}$ αξίωμα) έχουμε:

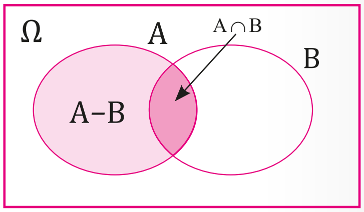
$$P(A\cup A')=P(A)+P(A') \text{ ή } P(\Omega)=P(A)+P(A') \text{ ή } 1=P(A)+P(A')$$

Σημείωση: Ο παραπάνω κανόνας είναι πολύ χρήσιμος σε προβλήματα στα οποία η πιθανότητα ενός ενδεχομένου A είναι δύσκολο να υπολογιστεί άμεσα.

Μπορούμε να υπολογίσουμε έμμεσα την $P(A)$. Υπολογίζουμε την πιθανότητα του συμπληρωματικού ενδεχομένου και εφαρμόζουμε τον κανόνα 1. (βλέπε πρόβλημα 3, ενότητα 4.6)

Κανόνας 2. Για δύο οποιαδήποτε ενδεχόμενα A και B ενός δειγματοχώρου Ω ισχύει:

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



Απόδειξη

Παρατηρούμε ότι τα ενδεχόμενα $A-B$ και $A\cap B$ είναι ασυμβίβαστα και για την ένωσή τους ισχύει: $(A-B)\cup(A\cap B)=A$.

Από τον απλό προσθετικό νόμο έχουμε:

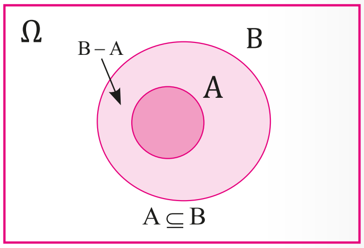
$$P(A)=P\big((A-B)\cup(A\cap B)\big)=P(A-B)+P(A\cap B)$$

ή

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

Κανόνας 3. Για δύο οποιαδήποτε ενδεχόμενα A και B ενός δειγματοχώρου Ω ισχύει:

$$\text{Αν } A\subseteq B \text{ τότε } P(A)\le P(B)$$



Απόδειξη

Παρατηρούμε ότι τα ενδεχόμενα A και $B-A$ είναι ασυμβίβαστα και για την ένωσή τους ισχύει: $A\cup(B-A)=B$.

Από τον απλό προσθετικό νόμο (3ο αξίωμα) έχουμε:

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

Επομένως, $P(B)\ge P(A)$, διότι $P(B-A)\ge 0$.

Μπορούμε τώρα να αποδείξουμε την παρακάτω πρόταση:

Πρόταση: Για κάθε ενδεχόμενο A ενός πεπερασμένου δειγματοχώρου Ω ισχύει:

$$P(A)\le 1$$

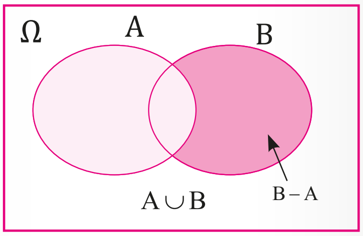
Απόδειξη: Από το 2ο αξίωματου αξιωματικού ορισμού της πιθανότητας έχουμε: $P(\Omega)=1$.

Από τον κανόνα 3 είναι: $P(A)\le P(\Omega)=1$

Κανόνας 4. Για δύο οποιαδήποτε ενδεχόμενα A και B ενός δειγματοχώρου Ω ισχύει:

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

Ο παραπάνω κανόνας ονομάζεται Προσθετικός Νόμος των πιθανοτήτων και αποτελεί γενίκευση του απλού προσθετικού νόμου.



Απόδειξη

Παρατηρούμε ότι τα ενδεχόμενα A και $B-A$ είναι ασυμβίβαστα και για την ένωσή τους ισχύει: $A\cup(B-A)=A\cup B$.

Από τον απλό προσθετικό νόμο έχουμε:

$$P(A\cup B)=P(A\cup(B-A)) \quad \text{ή}$$
$$P(A\cup B)=P(A)+P(B-A) \quad \text{ή, από τον κανόνα 2,}$$
$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

Σημείωση: Αποδείξαμε ακόμα ότι $P(A\cup B)=P(A)+P(B-A)$

Προσπαθήστε να λύσετε τις ασκήσεις 1 έως 3

Εφαρμογές των κανόνων λογισμού πιθανοτήτων

Εφαρμογή 1

Τα σύνολα A και B είναι ενδεχόμενα ενός δειγματοχώρου Ω με $P(A)=0{,}5$, $P(B)=0{,}6$ και $P(A\cap B)=0{,}3$. Να υπολογίσετε τις παρακάτω πιθανότητες:

α. $P(A')$ και $P(B')$ β. $P(A\cup B)$ γ. $P(A-B)$ και $P(B-A)$

Λύση

α. Από τον κανόνα 1 έχουμε:

$$P(A')=1-P(A)=1-0{,}5=0{,}5 \quad \text{και} \quad P(B')=1-P(B')=1-0{,}6=0{,}4.$$

β. Από τον προσθετικό νόμο (κανόνας 4) έχουμε:

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(A\cap B)\\&=0{,}5+0{,}6-0{,}3\\&=0{,}8.\end{aligned}$$

γ. Από τον κανόνα 2 έχουμε:

$$P(A-B)=P(A)-P(A\cap B)=0{,}5-0{,}3=0{,}2$$

και

$$P(B-A)=P(B)-P(A\cap B)=0{,}6-0{,}3=0{,}3.$$

Λογισμός πιθανοτήτων

Λογισμός πιθανοτήτων

Προσπαθήστε να λύσετε τις ασκήσεις 4 έως 6

Εφαρμογή 2

Τα σύνολα A και B είναι ενδεχόμενα ενός δειγματοχώρου Ω με

$$P(A')=0{,}3,\ P(B')=0{,}4 \ \text{και}\ P(A\cup B)=0{,}9$$

Να υπολογίσετε τις παρακάτω πιθανότητες:

α. $P(A)$ και $P(B)$

β. $P(A\cap B)$

Λύση

α. Από τον κανόνα 1 έχουμε:

$$P(A)=1-P(A')=1-0{,}3=0{,}7 \quad \text{και} \quad P(B)=1-P(B')=1-0{,}4=0{,}6$$

β. Από τον προσθετικό νόμο έχουμε:

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

ή

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

Αντικαθιστώντας τις πιθανότητες $P(A)$, $P(B)$ και $P(A\cup B)$ με τις τιμές τους παίρνουμε:

$$P(A\cap B)=P(A)+P(B)-P(A\cup B)\ =0{,}7+0{,}6-0{,}9\ =0{,}4$$

Προσπαθήστε να λύσετε την άσκηση 7

Εφαρμογή 3

Τα σύνολα A και B είναι ενδεχόμενα ενός δειγματοχώρου Ω με $A\subseteq B$,

$$P((A\cap B)')=\frac{3}{4}\quad \text{και}\quad P(A\cup B)=\frac{1}{2}.$$

Να υπολογίσετε τις πιθανότητες:

α. $P(A)$ και $P(B)$ β. $P(B\cap A')$

Λύση

α. Από τον κανόνα 1 έχουμε:

$$P(A\cap B)=1-P((A\cap B)')=1-\frac{3}{4}=\frac{1}{4}.$$

Επιπλέον, επειδή $A\subseteq B$, θα ισχύουν οι ισότητες: $A\cap B=A$ και $A\cup B=B$. Οπότε,

$$P(A)=P(A\cap B)=\frac{1}{4}\quad \text{και}\quad P(B)=P(A\cup B)=\frac{1}{2}.$$

β. Το ενδεχόμενο $B\cap A'$ είναι το $B-A$. Από τον κανόνα 2 και από το προηγούμενο ερώτημα έχουμε:

$$P(B-A)=P(B)-P(A\cap B)=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}.$$

Προσπαθήστε να λύσετε τις ασκήσεις 8 έως 13

Διάγραμμα Venn και πιθανότητα

Λογισμός πιθανοτήτων

Λογισμός πιθανοτήτων

Ασκήσεις και Προβλήματα

Ερωτήσεις

1 Να χαρακτηρίσετε καθεμία από τις επόμενες προτάσεις, ως αληθή ή ψευδή:

α. $P(A')=1-P(A)$.

β. Για δύο οποιαδήποτε ενδεχόμενα Α και Β ενός δειγματοχώρου Ω ισχύει: $P(A\cup B)=P(A)+P(B)$.

γ. Αν $P(A\cap B)=0$, τότε $P(A\cup B)=P(A)+P(B)$.

δ. Είναι $P(A)\le P(A\cup B)$

ε. Είναι $P(B-A)=P(B)-P(A\cap B)$

2 Τα σύνολα Α και Β είναι ενδεχόμενα ενός δειγματοχώρου Ω, με $P(A)=P(B)=0{,}6$
Μπορούμε να συμπεράνουμε, ότι τα ενδεχόμενα είναι ίσα;

3 Για δύο ενδεχόμενα Α και Β ενός δειγματοχώρου Ω ισχύουν $P(A)=0{,}7$ και $P(B)=0{,}6$. Μια μαθήτρια ισχυρίστηκε ότι $P(A\cap B)\le 0{,}6$. Έχει δίκιο; Να αιτιολογήσετε την απάντησή σας.

Ασκήσεις

4 Για δύο ενδεχόμενα Α και Β ενός δειγματοχώρου Ω ισχύουν

$$P(A)=0{,}6,\ P(B)=0{,}4 \text{ και } P(A\cup B)=0{,}8$$

Να υπολογίσετε τις πιθανότητες

$$P(A\cap B) \text{ και } P(B-A).$$

5 Για δύο ενδεχόμενα Α και Β ενός δειγματοχώρου Ω ισχύουν

$$P(A\cup B)=0{,}7,\ P(A\cap B)=0{,}2 \text{ και } P(A)=2P(B)$$

Να υπολογίσετε τις πιθανότητες $P(A)$ και $P(B)$.

6 Αν για ένα ενδεχόμενο Α ενός δειγματοχώρου Ω ισχύει

$$9P(A)=P(A')+4$$

να αποδείξετε ότι τα ενδεχόμενα Α και A' είναι ισοπίθανα.

7 Τα Α και Β είναι ενδεχόμενα ενός δειγματοχώρου Ω με

$$P(A')=0{,}7,\ P(B')=0{,}5 \text{ και } P(A\cup B)=0{,}7$$

Να υπολογίσετε τις παρακάτω πιθανότητες:

$$P(A),\ P(B) \text{ και } P(A\cap B).$$

8 Αν τα σύνολα Α και Β είναι ενδεχόμενα ενός δειγματοχώρου Ω με $A\subseteq B$, να αποδείξετε ότι:

$$P(A\cup B)=P(B)$$

9 Έστω τα ενδεχόμενα Α και Β ενός δειγματοχώρου Ω με:

$$P(A')=0{,}5 \text{ και } P(A\cup B)=0{,}7$$

Να βρείτε την πιθανότητα να πραγματοποιηθεί το Β και να μη πραγματοποιηθεί το Α.

10 Έστω τα ενδεχόμενα Α και Β ενός δειγματοχώρου Ω με:

$$P(B)=0{,}5 \text{ και } P(B-A)=0{,}2$$

α. Να βρείτε την πιθανότητα να πραγματοποιηθεί το Α και το Β.

β. Αν $P(A\cup B)=0{,}7$ να βρείτε ην πιθανότητα να πραγματοποιηθεί το Α.

11 Τα Α και Β είναι ενδεχόμενα ενός δειγματοχώρου Ω με:

$$B\subseteq A,\quad P((A\cap B)')=\frac{4}{5} \text{ και } P((A\cup B)')=\frac{2}{3}.$$

Να υπολογίσετε τις πιθανότητες: $P(A)$, $P(B)$ και $P(A\cap B')$.

12 Για κάθε ενδεχόμενο Α ενός δειγματοχώρου Ω να αποδείξετε ότι:

$$P(A)P(A')\le\frac{1}{4}$$

13 Για τα ενδεχόμενα Α και Β ενός δειγματοχώρου ισχύουν:

$$P(A)=0{,}8 \text{ και } P(B)=0{,}6$$

α. Να εξετάσετε αν τα Α και Β είναι ασυμβίβαστα.

β. Να αποδείξετε ότι:

$$0{,}4\le P(A\cap B)\le 0{,}6$$

4.6 Επίλυση προβλημάτων

Στην ενότητα αυτή θα μάθουμε:
Να επιλύουμε προβλήματα πιθανοτήτων χρησιμοποιώντας και τους κανόνες λογισμού πιθανοτήτων.

Περίεργα ζάρια

Πρόβλημα 1

Έστω η εξίσωση $x^2 - 4x + \alpha = 0$. Ρίχνουμε ένα ζάρι και αντικαθιστούμε το α με την ένδειξη του ζαριού. Ποια είναι η πιθανότητα η εξίσωση να είναι αδύνατη;

Λύση

Ο δειγματοχώρος του πειράματος είναι ο

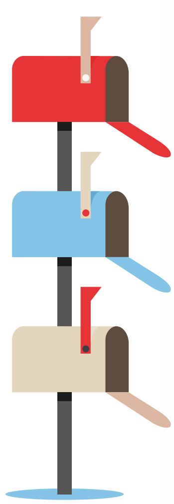
$$\Omega = \{\ 1,\ 2,\ 3,\ 4,\ 5,\ 6\ \}$$

Η εξίσωση είναι αδύνατη όταν $\Delta < 0$. Είναι $\Delta = 16 - 4\alpha$.
Επομένως είναι $\Delta < 0$ όταν

$$16 - 4\alpha < 0 \text{ ή } \alpha > 4.$$

Έτσι το ενδεχόμενο να είναι αδύνατη η εξίσωση είναι το $A = \{5,6\}$.
Επομένως

$$P(A) = \frac{\nu(A)}{\nu(\Omega)} = \frac{2}{6} = \frac{1}{3}$$



Πρόβλημα 2

Σε μία πολυκατοικία μένουν τρείς οικογένειες. Ο ταχυδρόμος τοποθετεί στα γραμματοκιβώτιά τους τον λογαριασμό ρεύματος στην τύχη, χωρίς δηλαδή να κοιτάξει σε ποια οικογένεια αντιστοιχεί κάθε λογαριασμός.

α. Να γράψετε έναν κατάλληλο δειγματοχώρο του πειράματος.

β. Να βρείτε την πιθανότητα κάθε οικογένεια να πάρει τον σωστό λογαριασμό.

Λύση

α. Έστω α ο λογαριασμός της 1ης οικογένειας, β την 2ης και γ της 3ης. Στον επόμενο πίνακα βλέπουμε όλες τις δυνατές περιπτώσεις που μπορεί ο ταχυδρόμος να μοιράσει τους λογαριασμός α, β και γ.

1η οικογένεια	α	α	β	β	γ	γ
2η οικογένεια	β	γ	α	γ	α	β
3η οικογένεια	γ	β	γ	α	β	α

Επομένως ένας κατάλληλος δειγματοχώρος είναι ο

$$\Omega = \{\ \alpha\beta\gamma,\ \alpha\gamma\beta,\ \beta\alpha\gamma,\ \beta\gamma\alpha,\ \gamma\alpha\beta,\ \gamma\beta\alpha\ \}$$

β. Το ενδεχόμενο κάθε οικογένεια να πάρει τον σωστό λογαριασμό είναι το

$$A = \{\alpha\beta\gamma\}$$

Επομένως

$$P(A)=\frac{\nu(A)}{\nu(\Omega)}=\frac{1}{6}$$

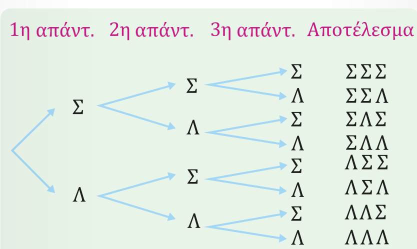
Πρόβλημα 3

Σε ένα διαγώνισμα υπάρχουν 3 ερωτήσεις τύπου «Σωστό – Λάθος». Ο Γιάννης μη γνωρίζοντας τις απαντήσεις, απαντάει στην τύχη.

α. Να γράψετε έναν κατάλληλο δειγματοχώρο με τις δυνατές απαντήσεις που μπορεί να δώσει ο Γιάννης.

β. Να βρείτε την πιθανότητα ο Γιάννης να απαντήσει τουλάχιστον μία ερώτηση σωστά.

γ. Ποια είναι η πιθανότητα να απαντήσει ο Γιάννης σωστά τουλάχιστον μία ερώτηση, αν το διαγώνισμα είχε 4 ερωτήσεις «Σωστό – Λάθος».



Λύση

α. Ένας κατάλληλος δειγματοχώρος είναι ο

$$\Omega=\{\ \Sigma\Sigma\Sigma,\ \Sigma\Sigma\Lambda,\ \Sigma\Lambda\Sigma,\ \Sigma\Lambda\Lambda,\ \Lambda\Sigma\Sigma,\ \Lambda\Sigma\Lambda,\ \Lambda\Lambda\Sigma,\ \Lambda\Lambda\Lambda\ \}$$

β. Έστω Α το ενδεχόμενο να απαντήσει ο Γιάννης τουλάχιστον μία ερώτηση σωστά. Το συμπληρωματικό Α΄ του Α είναι να απαντήσει ο Γιάννης όλες τις ερωτήσεις λανθασμένα, το οποίο περιέχει ένα μόνο στοιχείο του Ω. Επομένως:

$$P(A')=\frac{\nu(A')}{\nu(\Omega)}=\frac{1}{8}$$

και έτσι

$$P(A)=1-P(A')=1-\frac{1}{8}=\frac{7}{8}$$

γ. Αν οι ερωτήσεις είναι τέσσερις, τότε ο δειγματοχώρος Ω_1 περιέχει 16 απλά ενδεχόμενα. Πραγματικά, αυτά προκύπτουν από τα 8 απλά ενδεχόμενα του δειγματοχώρου Ω του (α) ερωτήματος, αν σε αυτά προσθέσουμε για την τέταρτη ερώτηση μία φορά το Σ και μία φορά το Λ.
Έστω Β το ενδεχόμενο να απαντήσει ο Γιάννης τουλάχιστον μία ερώτηση σωστά. Το συμπληρωματικό Β΄ του Β είναι να απαντήσει ο Γιάννης όλες τις ερωτήσεις λανθασμένα. Το Β΄ περιέχει προφανώς μόνο ένα στοιχείο.
Επομένως:

$$P(B')=\frac{\nu(B')}{\nu(\Omega_1)}=\frac{1}{16}$$

και έτσι

$$P(B)=1-P(B')=1-\frac{1}{16}=\frac{15}{16}$$

Πρόβλημα 4

Η πιθανότητα ένας μαθητής/τρια της Α΄ Λυκείου να γράψει στις προαγωγικές εξετάσεις κάτω από τη βάση στο μάθημα της Γλώσσας και των Μαθηματικών είναι 0,2 και 0,3 αντίστοιχα, ενώ η πιθανότητα να γράψει κάτω από τη βάση και στα δύο μαθήματα είναι 0,05. Να υπολογίστε την πιθανότητα ένας τυχαία επιλεγμένος μαθητής/τρια της Α΄ Λυκείου να γράψει στις προαγωγικές εξετάσεις κάτω από τη βάση:

α. Τουλάχιστον σε ένα από τα δύο μαθήματα.

β. Μόνο στο μάθημα της Γλώσσας.

γ. Μόνο στο μάθημα των Μαθηματικών.

Λύση

Ορίζουμε τα ενδεχόμενα:

Α: «Ο μαθητής/τρια να γράψει κάτω από τη βάση στο μάθημα της Γλώσσας»

Β: «Ο μαθητής/τρια να γράψει κάτω από τη βάση στο μάθημα των Μαθηματικών»

Από τα δεδομένα του προβλήματος έχουμε:

$$P(A)=0{,}2 \quad \text{και} \quad P(A\cap B)=0{,}05.$$

α. Το ενδεχόμενο «Ο μαθητής/τρια να γράψει κάτω από τη βάση τουλάχιστον σε ένα από τα δύο μαθήματα» είναι το $A\cup B$. Επομένως, από τον προσθετικό νόμο, η ζητούμενη πιθανότητα είναι:

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)=0{,}2+0{,}3-0{,}05=0{,}45$$

β. Το ενδεχόμενο «Ο μαθητής/τρια να γράψει κάτω από τη βάση μόνο στο μάθημα της Γλώσσας» είναι το $A-B$. Επομένως, από τον κανόνα 2, η ζητούμενη πιθανότητα είναι:

$$P(A-B)=P(A)-P(A\cap B)=0{,}2-0{,}05=0{,}15$$

γ. Το ενδεχόμενο «Ο μαθητής/τρια να γράψει κάτω από τη βάση μόνο στο μάθημα των Μαθηματικών» είναι το $B-A$. Επομένως, από τον κανόνα 2, η ζητούμενη πιθανότητα είναι:

$$P(B-A)=P(B)-P(A\cap B)=0{,}3-0{,}05=0{,}25$$

Ασκήσεις και Προβλήματα

Προβλήματα

1 Σε μια έρευνα που έγινε στο δήμο Αθηναίων ως προς τον τρόπο που ενημερώνονται οι δημότες για τις ειδήσεις προέκυψε ότι το 75% των δημοτών ενημερώνεται από την τηλεόραση, το 40% από το διαδίκτυο και το 20% από την τηλεόραση και το διαδίκτυο. Αν επιλέξουμε τυχαία ένα δημότη, να υπολογίσετε την πιθανότητα:

α. Να ενημερώνεται από την τηλεόραση ή από το διαδίκτυο.

β. Να μην ενημερώνεται από κανένα από τα δύο μέσα.

γ. Να ενημερώνεται από το διαδίκτυο και όχι από την τηλεόραση.

2 Σε μια τάξη με 25 μαθητές και μαθήτριες οι 7 έγραψαν βαθμό πάνω από 17 στο τελευταίο διαγώνισμα Ιστορίας, ενώ οι 15 έγραψαν από 12 έως 16. Επιλέγουμε τυχαία ένα από τα παιδιά της τάξης. Με Α συμβολίζουμε το ενδεχόμενο «το παιδί έγραψε βαθμό πάνω από 17» και με Β το ενδεχόμενο: «το παιδί έγραψε από 12 έως 16».

α. Να εξηγήσετε γιατί τα ενδεχόμενα Α και Β είναι ασυμβίβαστα.

β. Να υπολογίσετε τις πιθανότητες $P(A)$, $P(B)$, $P(A\cup B)$ και $P((A\cup B)')$.

3 Από ιατρικά δεδομένα έχει διαπιστωθεί ότι το 5% των παιδιών θα παρουσιάσει κάποια στιγμή της ανήλικης ζωής του ανεμοβλογιά, το 3% θα παρουσιάσει παρωτίτιδα ενώ το 92,03% δεν θα παρουσιάσει καμία από τις δύο ασθένειες. Αν επιλέξουμε τυχαία ένα νεογέννητο παιδί, να υπολογίσετε την πιθανότητα να παρουσιάσει:

α. Τουλάχιστον μία από τις δύο ασθένειες.
β. Και τις δύο ασθένειες.
γ. Μόνο ανεμοβλογιά ή μόνο παρωτίτιδα.

4 Μια ομάδα μπάσκετ δίνει δύο αγώνες σε ένα τουρνουά.

- Η πιθανότητα να κερδίσει τον πρώτο αγώνα είναι 0,5.
- Η πιθανότητα να κερδίσει το δεύτερο αγώνα είναι 0,4.
- Η πιθανότητα να κερδίσει και τους δύο αγώνες είναι 0,2.

Να υπολογίσετε την πιθανότητα:

α. Η ομάδα να κερδίσει τουλάχιστον έναν από τους δύο αγώνες.
β. Η ομάδα να κερδίσει μόνο τον δεύτερο αγώνα.

Αν η ομάδα χάσει τουλάχιστον ένα από τους δύο αγώνες τότε αποκλείεται από την επόμενη φάση του τουρνουά.

γ. Να υπολογίσετε την πιθανότητα η ομάδα να αποκλειστεί από την επόμενη φάση.

5 Σε μια χοροεσπερίδα ενός σχολείου θα πουληθούν 100 λαχνοί αριθμημένοι από το 1 έως το 100. Στην κλήρωση επιλέγεται τυχαία ένας λαχνός. Χρησιμοποιώντας τους κανόνες λογισμού πιθανοτήτων να βρείτε την πιθανότητα ο κερδοφόρος λαχνός να έχει περιττό αριθμό ή να είναι πολλαπλάσιο του 5.

6 Ρίχνουμε ένα δίκαιο ζάρι μια φορά και αντικαθιστούμε με την ένδειξή που θα έρθει το λ στην εξίσωση:

$$(\lambda^2 - 4\lambda + 3)x = \lambda + 4$$

Να βρείτε την πιθανότητα να προκύψει εξίσωση με μοναδική λύση.

7 Η κυρία Μαρία έχει τρεις γιους και ετοίμασε για την Πρωτοχρονιά ένα δώρο για τον καθένα τους το οποίο τοποθέτησε σε πανομοιότυπα κουτιά χωρίς να γράφει επάνω το όνομα του γιου για τον οποίο προοριζόταν. Αν κάθε γιος επιλέξει ένα κουτί στην τύχη, ποια είναι η πιθανότητα καθένας να πάρει το σωστό δώρο;

8 Σε μία άσκηση αντιστοίχισης υπάρχουν τρεις ερωτήσεις 1, 2 και 3 και οι απαντήσεις τους α, β και γ γραμμένες σε τυχαία σειρά. Αν ένας μαθητής απαντήσει στην τύχη ποια είναι η πιθανότητα να απαντήσει σωστά σε τουλάχιστον μία ερώτηση;

9 Ρίχνουμε τρεις φορές ένα αμερόληπτο ζάρι και καταγράφουμε τις ενδείξεις. Κατασκευάζοντας ένα συνοπτικό δεντροδιάγραμμα να συμπεράνετε ότι $\nu(\Omega) = 6^3$, όπου Ω ο δειγματοχώρος του πειράματος. Να υπολογίσετε την πιθανότητα τουλάχιστον μία από τις τρεις ενδείξεις να είναι 6.

Πιθανότητες και διαδίκτυο

Η θεωρία πιθανοτήτων διαδραματίζει έναν κρίσιμο ρόλο στη λειτουργία των μηχανών αναζήτησης ιστοσελίδων του διαδικτύου. Οι μηχανές αναζήτησης εκτιμούν τη σημαντικότητα μιας ιστοσελίδας μέσω της πιθανότητας που έχει ένας χρήστης του διαδικτύου να επιλέξει την ιστοσελίδα αυτή κατά την περιήγησή του στο διαδίκτυο. Χρησιμοποιώντας τη θεωρία πιθανοτήτων οι αλγόριθμοι των μηχανών αναζήτησης αναγνωρίζουν ότι οι ιστοσελίδες με περισσότερους συνδέσμους προς αυτές από άλλες ιστοσελίδες έχουν υψηλότερη πιθανότητα να επιλεγούν από έναν χρήστη και τις κατατάσσουν στις πρώτες θέσεις. Με τον τρόπο αυτό, η θεωρία πιθανοτήτων παρέχει τη θεωρητική βάση για την αξιολόγηση της σημαντικότητας των ιστοσελίδων στον κόσμο του διαδικτύου.

4.7 Ανακεφαλαίωση

Στην ενότητα αυτή:
Περιλαμβάνονται έργα για περαιτέρω αναζητήσεις και διερευνήσεις των μαθητών/μαθητριών, με στόχο τη διεύρυνση της μαθησιακής διαδικασίας που οδηγεί στα Προσδοκώμενα Μαθησιακά Αποτελέσματα του κεφαλαίου.

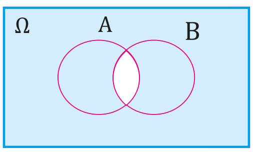
1 Τα σύνολα A και B είναι ενδεχόμενα του δειγματοχώρου Ω ενός πειράματος τύχης. Να αντιστοιχίσετε καθεμιά από τις παρακάτω προτάσεις με το διπλανό διάγραμμα Venn που την εκφράζει.

α. Δεν είναι εφικτό να πραγματοποιηθούν συγχρόνως το A και το B.

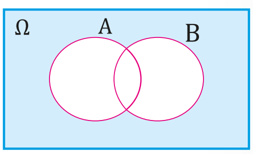
β. Δεν πραγματοποιείται το A.

γ. Πραγματοποιείται το A' ή το B'.

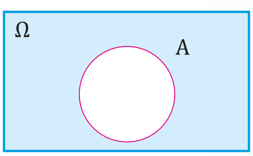
δ. Πραγματοποιείται το A' και το B'.



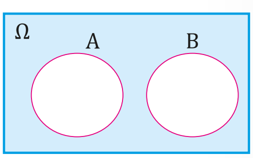
Σχ. 1



Σχ. 2



Σχ. 3



Σχ. 4

2 Η βιβλιοθήκη ενός σχολείου περιέχει 4 αριθμημένα αντίτυπα ενός λογοτεχνικού βιβλίου από τα οποία τα 2 πρώτα, A_1 και A_2, είναι αντίτυπα της πρώτης έκδοσης του βιβλίου ενώ τα υπόλοιπα, B_1 και B_2, είναι της δεύτερης έκδοσης. Επιλέγουμε διαδοχικά και χωρίς να ξανατοποθετήσουμε πίσω στη βιβλιοθήκη βιβλία μέχρι να πετύχουμε ένα βιβλίο δεύτερης έκδοσης.

α. Να βρείτε το δειγματοχώρο του πειράματος.

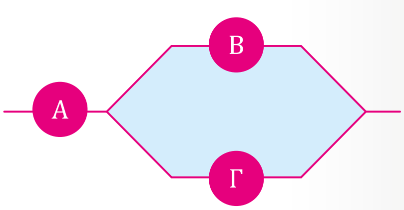
β. Να παραστήσετε με αναγραφή των στοιχείων του το ενδεχόμενο A: «Να χρειαστούμε τουλάχιστον 3 προσπάθειες μέχρι να ολοκληρωθεί η διαδικασία».

3 Σε ένα ηλεκτρικό σύστημα είναι ενσωματωμένη μια συσκευή η οποία μπορεί να λειτουργεί ακόμα και αν κάποια μέρη της πάψουν να λειτουργούν. Μια τέτοια συσκευή φαίνεται στο διπλανό σχήμα. Τα μέρη Β και Γ της συσκευής είναι συνδεδεμένα παράλληλα έτσι ώστε η συσκευή να λειτουργεί αν τουλάχιστον ένα από τα Β και Γ λειτουργεί. Η συσκευή παύει να λειτουργεί αν το μέρος Α χαλάσει. Κωδικοποιούμε κάθε μέρος της συσκευής με τον αριθμό 1, αν αυτό λειτουργεί, και με τον αριθμό 0 αν αυτό έχει χαλάσει. Για παράδειγμα, γράφοντας 110 θα εννοούμε ότι τα μέρη Α και Β λειτουργούν ενώ το μέρος Γ έχει χαλάσει. Ελέγχουμε μια χρονική στιγμή τη συσκευή ως προς τη λειτουργικότητά της.

α. Να βρείτε το δειγματοχώρο του πειράματος.

β. Να βρείτε το ενδεχόμενο να λειτουργούν ακριβώς δύο μέρη της συσκευής.

γ. Να βρείτε το ενδεχόμενο η συσκευή να πάψει να λειτουργεί.



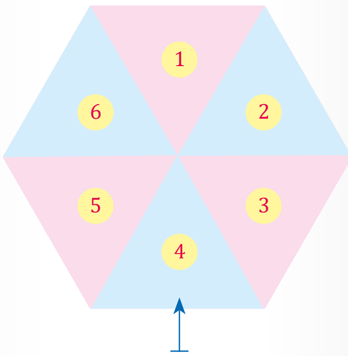
4 Η πιθανότητα να πάει μια οικογένεια για διακοπές στην Κύπρο είναι 0,2, να πάει στην Ελλάδα είναι 0,3 και η πιθανότητα να πάει τουλάχιστον σε μία από τις δύο χώρες είναι 0,4. Να βρείτε την πιθανότητα η οικογένεια :

α. Να πάει και στις δύο χώρες.

β. Να πάει μόνο στην Ελλάδα.

Οδηγός επανάληψης-Πιθανότητες

5 Η βραδινή έξοδος που είναι προγραμματισμένη για αύριο θα αναβληθεί αν ο καιρός είναι συννεφιασμένος ή βροχερός. Η πρόβλεψη καιρού αναφέρει ότι η πιθανότητα βροχής αύριο είναι 30%, η πιθανότητα συννεφιάς είναι 60% και η πιθανότητα ταυτόχρονης συννεφιάς και βροχής είναι 20%. Ποια είναι η πιθανότητα η βραδινή έξοδος να μην αναβληθεί;

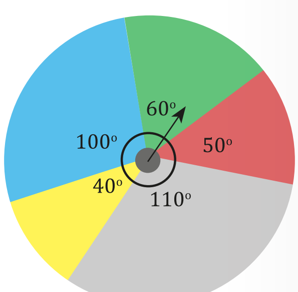


6 Ένας περιστρεφόμενος τροχός έχει σχήμα κανονικού εξαγώνου και είναι κατασκευασμένος από ομοιογενές υλικό. Χωρίσαμε τον τροχό σε έξι ίσα κομμάτια τριγωνικού σχήματος και πάνω σε κάθε κομμάτι σημειώσαμε τους φυσικούς αριθμούς από το 1 μέχρι το 6. Αν στρίψουμε με όλη μας τη δύναμη τον τροχό και, επιπλέον, υποθέσουμε ότι ο δείκτης είναι αδύνατο να σταματήσει πάνω στη διαχωριστική γραμμή δύο οποιονδήποτε κομματιών, πόσο πιθανό θεωρείται ο δείκτης να σταματήσει:

α. Στο τριγωνικό κομμάτι που φέρει τον αριθμό 5;
β. Σε τριγωνικό κομμάτι που φέρει περιττό αριθμό;
γ. Σε τριγωνικό κομμάτι μπλε χρώματος;

7 Ρίχνουμε διαδοχικά δύο αμερόληπτα ζάρια και καταγράφουμε το αποτέλεσμα των ρίψεων. Να υπολογίσετε την πιθανότητα του ενδεχομένου η διαφορά των δύο ενδείξεων να είναι:

α. 0.
β. Μικρότερη του 2.
γ. Μεγαλύτερη του 3.



Περιστρεφόμενος τροχός

8 **Γεωμετρική πιθανότητα** Στο διπλανό σχήμα φαίνεται ένας δίσκος, χωρισμένος σε 5 χρωματισμένους κυκλικούς τομείς. Ο δίσκος χρησιμεύει σε ένα παιχνίδι 5 παικτών με τους εξής κανόνες:

- Κάθε παίκτης ή παίκτρια επιλέγει ένα χρώμα και έπειτα ένας ή μία από τους συμμετέχοντες επιλέγετε τυχαία για να περιστρέψει με δύναμη το βέλος. Νικητής ή νικήτρια του παιχνιδιού είναι αυτός/ή που επέλεξε τον τομέα που σταμάτησε το βέλος.
- Θεωρούμε ότι το βέλος δεν δύναται να σταματήσει μεταξύ δύο τομέων.

Να ορίσετε κατάλληλο δειγματοχώρο για το παραπάνω πείραμα τύχης και να αποδώσετε πιθανότητα νίκης σε κάθε παίκτη ή παίκτρια.

9 Για ένα ζάρι η πιθανότητα να φέρει άρτιο αριθμό είναι ίση με την πιθανότητα να φέρει περιττό αριθμό. Ένας μαθητής ισχυρίστηκε ότι το ζάρι αυτό αναγκαστικά είναι αμερόληπτο. Να εξηγήσετε γιατί ο ισχυρισμός του μαθητή είναι λάθος δίνοντας ένα παράδειγμα ζαριού στο οποίο τα δυνατά αποτελέσματα είναι μη ισοπίθανα αλλά τα ενδεχόμενα να φέρουμε άρτιο και να φέρουμε περιττό να έχουν την ίδια πιθανότητα πραγματοποίησης.

10 **Ομάδες αίματος.** Σύμφωνα με το σύστημα ΑΒΟ το ανθρώπινο αίμα μπορεί να ταξινομηθεί σε τέσσερις ομάδες: Ο, Α, Β, ΑΒ. Οι παραπάνω ομάδες μπορούν να διαχωριστούν περαιτέρω βάσει της παρουσίας ή απουσίας στο αίμα ενός παράγοντα που ονομάζεται Ρέζους (Rh). Η παρουσία του παράγοντα στο αίμα δηλώνεται με το σύμβολο «+» και η απουσία με το σύμβολο «–». Από μία ιατρική μελέτη προέκυψε ο παρακάτω πίνακας για τα ποσοστά κάθε ομάδας αίματος στον ελληνικό πληθυσμό.

O+	O–	A+	A–	B+	B–	AB+	AB–
37,8%	6,6%	32,2%	5,7%	11%	2%	4%	0,7%

Επιλέγουμε τυχαία ένα άτομο από τον ελληνικό πληθυσμό και καταγράφουμε την ομάδα αίματος που έχει και αν αυτή φέρει τον παράγοντα Rh ή όχι.

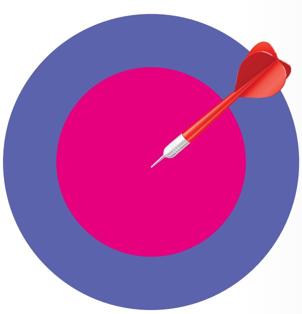
α. Να βρείτε έναν κατάλληλο δειγματοχώρο για το παραπάνω πείραμα. Είναι ισοπίθανα τα απλά ενδεχόμενα του δειγματοχώρου;

β. Ποια ομάδα αίματος είναι πιο πιθανό αυτός να έχει;

γ. Κάθε άνθρωπος με ομάδα αίματος B μπορεί να λάβει αίμα μόνο από άνθρωπο που ανήκει στις κατηγορίες O ή B, ανεξάρτητα από τον παράγοντα Ρέζους τον οποίο φέρει. Ο Αχιλλέας έχει ομάδα αίματος B. Να υπολογίσετε την πιθανότητα το άτομο που θα επιλεχθεί να μπορεί να γίνει δωρητής αίματος για τον Αχιλλέα.

Γεωμετρική πιθανότητα

11 **Γεωμετρική πιθανότητα** Η Κατερίνα δοκιμάζει για πρώτη φορά τις δυνατότητές της στο παιχνίδι του στόχου. Διαθέτει ένα κυκλικό στόχο ακτίνας $\rho = 1$ cm (όπως του διπλανού σχήματος) και θέλει, πριν ρίξει το βέλος, να τον χωρίσει σε δύο περιοχές, έναν εσωτερικό κυκλικό δίσκο με κέντρο το κέντρο του στόχου και έναν εξωτερικό κυκλικό δακτύλιο, ώστε το βέλος να έχει την ίδια πιθανότητα να καταλήξει σε μία από τις δύο περιοχές. Ποια πρέπει να είναι η ακτίνα του κυκλικού δίσκου και ποιο το πάχος του κυκλικού δακτυλίου;



12 Ένα κουτί περιέχει όμοιους βώλους από τους οποίους οι κ είναι κόκκινοι και οι μ μαύροι. Επιλέγουμε τυχαία έναν βώλο από το κουτί. Η πιθανότητα να είναι κόκκινος βώλος είναι $\frac{1}{3}$.

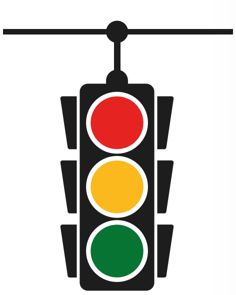
α. Να αποδείξετε ότι οι μαύροι βώλοι είναι διπλάσιοι από τους κόκκινους.

β. Αν προσθέσουμε 5 επιπλέον κόκκινους βώλους στο κουτί τότε η πιθανότητα να επιλεγεί κόκκινος βώλος γίνεται $\frac{1}{2}$. Να βρεθούν τα κ και μ.

13 Η Αλίκη παίρνει κάθε πρωί το αυτοκίνητό της για να πάει στη δουλειά. Κατά τη διαδρομή συναντάει τρία φανάρια. Οι παρακάτω πιθανότητες αντιστοιχούν στα ενδεχόμενα η Αλίκη να σταματήσει σε 0, 1, 2 και 3 κόκκινα φανάρια πριν φτάσει στη δουλειά:

$P(0 \text{ κόκκινα φανάρια}) = 0{,}13$ $\quad$ $P(1 \text{ κόκκινο φανάρι}) = 0{,}37$

$P(2 \text{ κόκκινα φανάρια}) = 0{,}33$ $\quad$ $P(3 \text{ κόκκινα φανάρια}) = 0{,}17$



α. Να υπολογίσετε την πιθανότητα να σταματήσει τουλάχιστον σε ένα κόκκινο φανάρι κατά τη διαδρομή προς τη δουλειά.

β. Να υπολογίσετε την πιθανότητα να σταματήσει σε δύο, το πολύ κόκκινα φανάρια.

Υποδείξεις - Απαντήσεις ασκήσεων

ΚΕΦΑΛΑΙΟ 1

Ενότητα 1.1

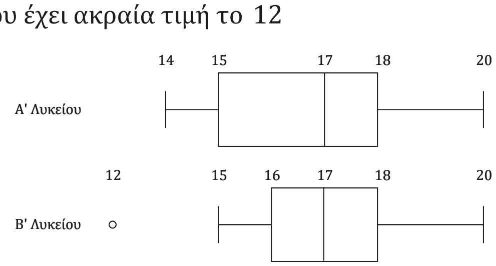
1. (α) Πληθυσμός: τα αυτοκίνητα της πόλης, δείγμα: τα αυτοκίνητα που διήλθαν από το κέντρο της πόλης τη συγκεκριμένη Δευτέρα
(β) Μεταβλητές: χρώμα, φύλο, επιβάτες, αριθμός κυκλοφορίας
Ποιοτικές: χρώμα, φύλο και αριθμός κυκλοφορίας,
Ποσοτικές: πλήθος επιβατών
2. (α) Ποσοτική συνεχής
(β) Ποσοτική διακριτή
(γ) Κατηγορική ονομαστική
(δ) Ποσοτική διακριτή
(ε) Κατηγορική διατάξιμη
(στ) Κατηγορική διατάξιμη
(ζ) Ποσοτική συνεχής
(η) Κατηγορική διατάξιμη
(θ) Κατηγορική ονομαστική
(ι) Ποσοτική διακριτή

Ενότητα 1.2

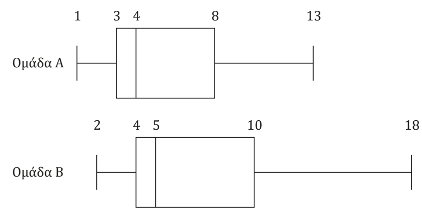
1. Ψ, Α, Α, Α, Α
2. (α) 10,2 (β) 14,9
3. (α) $Q_1 = 30$, $Q_3 = 50$, (β) Κάτω οριακή τιμή το 10 και άνω οριακή τιμή το 60, (γ) ναι
(δ) Ακραία τιμή το 90, (ε) ναι
4. (α) $Q_1 = 3$, $Q_3 = 9$ και $Q = 6$, (β) -6 και 18,
(γ) $L = 2$ και $U = 12$, (δ) δεν υπάρχουν
5. Ακραία μόνο το 32
6. (α) $Q_3 = 22$, (β) $[Q_1 - 1{,}5Q, Q_3 + 1{,}5Q] = [17{,}25]$
(γ) το 17 είναι οριακή τιμή
7.

	Q_1	δ	Q_3	**Q**	**L**	**U**	Ακραίες τιμές
Γ΄ Γυμν.	1,5	3	4	2,5	0	7,5	8
Α΄ Λυκ.	1	2	3	2	0	6	7 και 8

8. (α) Η Α΄ Λυκείου δεν έχει ακραία τιμή, η Β΄ Λυκείου έχει ακραία τιμή το 12



9. (α) Δεν υπάρχουν ακραίες σε καμία ομάδα
(β) Και τα δύο θηκογράμματα παρουσιάζουν θετική ασυμμετρία.



Ενότητα 1.3 (Ανακεφαλαίωση)

2. Α, Α, Ψ, Α, Ψ
3. (α) $Q_1 = 60$, $Q_3 = 80$ και $\delta = 75$, (β) άνω φράγμα το 110 και κάτω φράγμα 30, (γ) Άνω οριακή τιμή το 100 και κάτω οριακή τιμή το 35
(δ) Α, Ψ, Α, Α
4. (α)

	min	Q_1	δ	Q_3	max	Q
Πόλη Α	50	70	80	120	130	50
Πόλη Β	20	60	90	110	140	50

(β) $R_A = 80$ και $R_B = 120$, (γ) 25 (δ) 25
5. Ναι, τα θηκογράμματα των παρακάτω συνόλων διαφέρουν
Σύνολο Α: 2, 8, 8, 10, 16, 16, 20
Σύνολο Β: 2, 6, 6, 10, 14, 14, 20

ΚΕΦΑΛΑΙΟ 2

Ενότητα 2.1

1. (β) απόλυτη απόκλιση 3 και $s^2 = 9$, (γ) η ομάδα Α
3. $MAD = 7/3$, $s^2 = 22/3$ και $s \approx 2{,}71$
4. $\bar{x} = 4{,}5$, $MAD = 3/4$, $s^2 = 0{,}75$ και $s \approx 0{,}87$
6. (α) μέση τιμή 14, τυπική απόκλιση 3 (β) μέση τιμή 17,6, τυπική απόκλιση 3,3
7. (α) $s^2 = 0{,}64$ και $s = 0{,}8$ (β) η διασπορά παραμένει ίδια
8. (α) $s^2 = 8$, (β) 30, 21, 27, 24, 18 και η διασπορά τους θα είναι 18
9. (α) $MAD = 1{,}4$ (β) παραμένει η ίδια
10. (α) $s \approx 5{,}1$ (β) παραμένει η ίδια
11. Μέση τιμή 19 χρόνια και 3 μήνες, τυπική απόκλιση 2 μήνες
12. $\bar{x} = 1$, $s^2 = 0$ και $s = 0$

Ενότητα 2.2

1. Ψ, Ψ, Ψ, Α

2. (α) $\bar{x}_A = \bar{x}_\Gamma = 8$, (β) τα δεδομένα στις γυναίκες παρουσιάζουν μεγαλύτερη μεταβλητότητα

(γ) $s_A^2 = 3{,}6$, $s_A \approx 1{,}9$, $Q_A = 4$, $s_\Gamma^2 \approx 21{,}67$, $s_\Gamma \approx 4{,}66$ και $Q_\Gamma = 8$

3 (α) $\bar{x}_M = 6$ και $\bar{x}_\Phi = 5$ (β) $s_M = s_\Phi = \sqrt{2} \approx 1{,}41$

4. (α) $\bar{x}_A = 31$, $\delta_A = 31$, $\bar{x}_B = 32{,}4$ και $\delta_B = 32{,}5$ (β) $Q_A = 2$, $s_A^2 = 1$, $s_A = 1$, $Q_B = 3$, $s_B^2 = 2{,}01$ και $s_B \approx 1{,}42$ (γ) Οι συσκευασίες της εταιρείας Α παρουσιάζουν μικρότερη διασπορά και ενδοτεταρτημοριακό εύρος, που σημαίνει ότι είναι πιο αξιόπιστη για σταθερό περιεχόμενο. Η εταιρεία Β έχει μεγαλύτερη μέση τιμή και διάμεσο, κάτι που υποδηλώνει ότι οι συσκευασίες της περιέχουν, κατά μέσο όρο, περισσότερα αμύγδαλα. Ωστόσο, η διαφορά στη μέση τιμή και τη διάμεσο μεταξύ των δύο εταιρειών είναι πολύ μικρή. Αν η σταθερότητα είναι πιο σημαντική από τη μικρή αυτή αύξηση, η εταιρεία Α μπορεί να θεωρηθεί προτιμότερη.

5. (α) $\bar{x}_A \approx 5{,}29$, $\bar{x}_B \approx 4{,}57$, $\delta_A = 5$ και $\delta_B = 5$, (β) $Q_A = 4$, $s_A^2 \approx 7{,}06$, $s_A \approx 2{,}66$, $Q_B = 2$, $s_B^2 \approx 2{,}53$ και $s_B \approx 1{,}59$
(γ) Τις πιο θετικές αξιολογήσεις έχει το παγωτό Α. Επειδή $Q_A > Q_B$ και $s_A > s_B$ μεγαλύτερη μεταβλητότητα παρουσιάζει το παγωτό Α

6. (α) $\bar{x}_A \approx 7{,}67$, $\bar{x}_B \approx 8{,}33$, $\delta_A = 5{,}5$ και $\delta_B = 3$, (β) $Q_A = 7{,}5$, $s_A^2 \approx 78{,}72$, $s_A \approx 8{,}87$, $Q_B = 13{,}5$, $s_B^2 \approx 110{,}06$ και $s_B \approx 10{,}49$

Ενότητα 2.3

1. Ψ, Ψ, Α, Α, Α
2. (α) Όχι, (β) ο συντελεστής μεταβλητότητας, (γ) $CV \le 10\%$, (δ) ναι
3. $CV = 20\%$
4. $s = 2$
5. $\bar{x} = 200$ ή $\bar{x} = -200$. Είναι ομοιογενής
6. (α) Όχι, (β) του κ. Γιώργου
7. $CV_M = 20\%$ και $CV_X \approx 11{,}7\%$ άρα $CV_X < CV_M$. Μεγαλύτερη ομοιογένεια έχουμε στο μάθημα της Χημείας
8. Μεγαλύτερη ομοιογένεια παρουσιάζει η ομάδα 3
9. $CV \approx 44\%$ το δείγμα δεν είναι ομοιογενές
10. $CV_B < CV_A$ άρα μεγαλύτερη ομοιογένεια παρουσιάζουν οι βολές του παίκτη Β

Ενότητα 2.4

1. Α, Α, Ψ, Α
2. Η ομάδα Α
3. Στην ύπαρξη ενός ή περισσότερων πολύ ψηλών αγοριών
4. (α) Το 7 είναι ακραία τιμή (β) Η διάμεσος παραμένει σταθερή, ενώ η μέση τιμή ελαττώνεται
5. (α) $s^2 \approx 2{,}33$ και $R = 5$ (β) θα μειωθούν και το εύρος και η διασπορά με $R = 2$, και $s^2 = 0{,}64$
6. Η διάμεσος και το ενδοτεταρτημοριακό εύρος λόγω ύπαρξης της ακραίας τιμής 50
7. (α) $R = 9$, $Q = 1{,}5$, ακραίες τιμές οι 0 και 9, (β) χωρίς την ακραία τιμή αριστερά $R = 6$ και $Q = 2$, χωρίς την ακραία τιμή δεξιά $R = 5$ και $Q = 1$
8. Αρχική $MAD = 70/9 \approx 7{,}78$
Τελική $MAD = 0{,}88$. Όταν απομακρύναμε την ακραία τιμή η MAD ελαττώθηκε σημαντικά
9. (α) Ναι, οι 8 και 142. Η τιμή 8 πρέπει να απομακρυνθεί
(β) αφύσικη ακραία 8, οπότε $\bar{x}_A \approx 86{,}29$, $\delta = 85$, $R = 83$, $Q = 24$ και $s \approx 25{,}13$
(γ) η διάμεσος και ενδοτεταρτημοριακό εύρος
10. (α) δ (β) $\bar{x}$ (γ) $\bar{x}$ (δ) $\bar{x}$ ή δ

Ενότητα 2.5 (Ανακεφαλαίωση)

1.

α	β	γ	δ	ε	στ	ζ	η	θ	ι	ια	ιβ	ιγ	ιδ
Α	Α	Ψ	Ψ	Α	Ψ	Ψ	Α	Α	Ψ	Ψ	Α	Α	Ψ

2. (α) $\bar{x}_A = \bar{x}_B = \bar{x}_\Gamma$ και $\delta_A = \delta_B = \delta_\Gamma$ (β) $s_\Gamma < s_A < s_B$
3. Η νέα, μέση τιμή είναι 11,57 εκ. ευρώ και η τυπική απόκλιση είναι 0,9 εκ. ευρώ. Η επιχείρηση βελτίωσε την οικονομική της θέση.
4. (α) $\bar{x} \approx 91{,}82$ και $\delta = 95$, (β) το 230 ακραία τιμή αφού $230 > Q_3 + 1{,}5Q = 197{,}5$
(γ) $\bar{x} = 78$ και $\delta = 88{,}5$ άρα η ακραία τιμή επηρεάζει σε μεγαλύτερο βαθμό την μέση τιμή από τη διάμεσο
(δ) Αρχικά $s^2 \approx 2716{,}15$ και $Q = 59$. Μετά την απομάκρυνση της τιμής 230 προκύπτει ότι $s^2 \approx 887{,}4$ και $Q = 50$. Η ακραία τιμή επηρεάζει σε μεγαλύτερο βαθμό τη διασπορά απ' ό,τι το ενδοτερτημοριακό εύρος
5. (α) Ο άνθρωπος που ακολούθησε την διατροφή Α είχε πρόσληψη ενός κιλού
(β) $\bar{x}_A \approx 2{,}67$, $\delta_A = 3$, $\bar{x}_B = 4$ και $\delta_B = 4$,
(γ) $s_A^2 \approx 2{,}89$, $s_A \approx 1{,}7$, $Q_A = 2{,}5$, $s_B^2 = 1{,}33$,

$s_B \approx 1{,}15$, $Q_B = 2$
Είναι $CV_A \approx 64\%$ και $CV_B \approx 29\%$ οπότε μεγαλύτερη μεταβλητότητα παρουσιάζει η διατροφή Α.
(δ) αποτελεσματικότερη είναι η διατροφή Β

6. (α) $\delta = 16{,}5$ και $\bar{x} = 16$, (β) $s^2 = 4$
(γ$_1$) $\bar{x} = 16{,}5, \delta = 17, s^2 = 4,\ s = 2$
(γ$_2$) οι νέες βαθμολογίες έχουν μεγαλύτερη ομοιογένεια
7. $CV_B > CV_A$ άρα μεγαλύτερη ομοιογένεια παρουσιάζουν οι μισθοί του τμήματος Α
8. (α) Γιατί είναι διαφορετική η μονάδα μέτρησης. Καταλληλότερο μέτρο είναι ο συντελεστής μεταβλητότητας
(β) $CV_E < CV_A$ άρα μεγαλύτερη ομοιογένεια παρουσιάζει το δείγμα της ευρωπαϊκής εταιρείας
(γ) Α: η s δεν θα μεταβληθεί, ο CV θα ελαττωθεί. Ε: η s θα αυξηθεί, ο CV δεν θα μεταβληθεί
9. (α) Μας βοηθάει να συγκρίνουμε την αξία των σπιτιών στα δύο προάστια
(β) Στο προάστιο Α υπάρχουν αναλογικά φθηνότερα σπίτια σε σχέση με το προάστιο Β (σύγκριση διαμέσων). Η μέση τιμή $\bar{x}_A$ είναι παραπλανητική λόγω ύπαρξης σπιτιών με πολύ μεγάλη αξία σε σχέση με την πλειοψηφία των σπιτιών (σύγκριση $\bar{x}_A$, δ_A). Η μεταβλητότητα στα δύο προάστια είναι περίπου ίδια (σύγκριση Q_A, Q_B)

ΚΕΦΑΛΑΙΟ 3

Ενότητα 3.1

1. (α) $\delta_A > \delta_B$, (β) στο Β
2. (α) Β, (β) Α, Β (γ) Α, αρνητική
3. (α) $\delta_K - \delta_A = 5$ (β) $R_A > R_K$ (γ) $Q_K = 7$ και $Q_A = 10$ (δ) για τα αγόρια είναι 50% και για τα κορίτσια 75%
4. (α) 40 έως 45 (β) 35 έως 50
5. (α) Η ομάδα Α (β) η ομάδα Β (γ) δεν μπορούμε να συμπεράνουμε, διότι αν και η ομάδα Β πέτυχε λιγότερους πόντους μπορεί να πέτυχε περισσότερες νίκες στο πρωτάθλημα
6. (α) Στο τμήμα A_2 παρατηρήθηκε και ο χαμηλότερος και ο υψηλότερος βαθμός
(β) στο A_2 (γ) 13 μαθητές
7. $\delta_\Delta > \delta_\Sigma = \delta_K$
$Q_K > Q_\Delta = Q_\Sigma$
$U_K - L_K > U_\Sigma - L_\Sigma > U_\Delta - L_\Delta$

Δ: Αρνητική ασυμμετρία, Σ: Θετική ασυμμετρία, Κ: συμμετρική κατανομή
Παρατηρείτε μία μόνο ακραία τιμή (μικρή) στα δεδομένα της Δενδρότσιχλας.

Ενότητα 3.2

1. (α) Υπάρχει σχέση εξάρτησης, (β) όχι,
(γ)π.χ. διαφορετικές περιοχές με διαφορετική αγοραστική δύναμη
2. (α) Υπάρχει σχέση εξάρτησης, (β) όχι,
(γ)π.χ. καιρικές συνθήκες
3. (α) Υπάρχει σχέση εξάρτησης, (β) όχι,
(γ)π.χ. γενετικοί λόγοι

Ενότητα 3.3 (Ανακεφαλαίωση)

1. Η ομάδα Β παρουσιάζει μικρότερους χρόνους παρακολούθησης καθώς και μικρότερη μεταβλητότητα
2. Κατά μέσο όρο η πόλη Β έχει ελαφρώς μεγαλύτερο ύψος βροχόπτωσης αλλά παρουσιάζει μικρότερη μεταβλητότητα
3. (α) Μεγαλύτερη μεταβλητότητα η μπαταρία Α,
(β) δεν γνωρίζουμε (γ) το κινητό Α
4. (α)i) Γ, Β, Α ii) Γ, Β, Α iii) $Q_\Gamma = Q_B > Q_A$, (β) δεν έχουμε αρκετές πληροφορίες
5. (α) Όχι (β) π.χ. λόγω προγραμματισμού γεννήσεων ώστε μία οικογένεια να γεννήσει την άνοιξη

ΚΕΦΑΛΑΙΟ 4

Ενότητα 4.1

3. Γ και Ω
5. $A = \{3,6\}$
6 (α) $\Omega = \{(1,1),(1,2),(1,3),\ldots,(2,1),(2,2),\ldots,(6,6)\}$
(β) $A = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$
7.
$\Omega = \{(H_1,X_1),(H_1,X_2),(H_2,X_1),(H_2,X_2),(H_3,X_1),(H_3,X_2)\}$
8. (α)
$\Omega = \{ΚΚΚ, ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΚΚ, ΓΚΓ, ΓΓΚ, ΓΓΓ\}$

(β) $A = \{ΚΚΚ, ΚΚΓ, ΚΓΚ, ΓΚΚ\}$,
$B = \{ΚΚΚ, ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΚΚ, ΓΚΓ, ΓΓΚ\}$,
$\Gamma = \{ΚΓΓ, ΓΚΓ, ΓΓΚ, ΓΓΓ\}$
9. (α) $\Omega = \{ΝΕΑ, ΝΑΕ, ΑΕΝ, ΑΝΕ, ΕΑΝ, ΕΝΑ\}$
(β) $\{ΝΕΑ, ΝΑΕ, ΕΑΝ, ΕΝΑ\}$
10. (α) $\Omega = \{ΑΚ, ΑΜ, ΚΑ, ΚΜ, ΜΑ, ΜΚ\}$
(β) $\{ΑΚ, ΚΑ\}$

11. (α) $\Omega=\{1K,1\Gamma,2K,2\Gamma,3K,3\Gamma,4K,4\Gamma,5K,5\Gamma,6K,6\Gamma\}$,
(β) $A=\{1K,3K,5K\}$, $B=\{3\Gamma,4\Gamma,5\Gamma,6\Gamma\}$

Ενότητα 4.2

2. (α) $(A\cup B)'$, (β) A', (γ) $(B-A)'$
3. Το β
4. Το α
6. $A=\{1\}$, $B=\{1,3,5\}$. Είναι $A\subseteq B$
8. $P(K)=1/2$, $P(\Gamma)=1/2$, $P(\Omega)=1$, $P(\varnothing)=0$
9. (α) $A\cap B=\{2\}$, (β) $A\cup B=\{2,3,4,5,6\}$,
(γ) $A'=\{1,4,6\}$, $P(A\cap B)=1/6$,
$P(A\cup B)=5/6$, $P(A')=1/2$,
10. (α) $P(A)=3/4$, $P(B)=11/20$,
(β) $P(A')=1/4$, $P(B')=9/20$,
$P(A\cap B)=3/8$, $P((A\cap B)')=5/8$,
$P((A\cup B)')=3/4$
11. (α) $B'=\{1,2,3,5,6,7,9,10,11\}$, $P(B')=3/4$,
(β) $A\cap B=\{12\}$, $P(A\cap B)=1/12$,
(γ) $A-B=\{3,6,9\}$, $P(A-B)=1/4$,
(δ) $A'\cap B'=\{1,2,5,7,10,11\}$, $P(A'\cap B')=1/2$

Ενότητα 4.3

1. (α) Όχι (β) επαναλαμβάνουμε πολλές φόρες το πείραμα και εκτιμούμε την πιθανότητα από τη σχετική συχνότητα
4. (α) $\Omega=\{\Pi,I,\Theta,A,O,T,N,H\}$, (β) $P(\Pi)=1/10$, $P(T)=1/5$, $P(A)=1/5$
5. (α) $\Omega=\{E,A\}$, (β) $P(E)=1/3$, $P(A)=2/3$
6. (α) $\Omega=\{K,\Pi\}$, (β) $P(K)=2/3$, $P(\Pi)=1/3$

Ενότητα 4.4

1. Ψ, Ψ, Α, Α, Α
2. Όχι
3. Ναι, διότι $P(0)+P(1)+P(2)=1$
4. (α) 45%
5. $p=0{,}04$
6. $P(A)=0{,}9$, $P(B)=0{,}4$ και $P(\Gamma)=0{,}7$
7. $P(A)=9/10$
8. $P(\omega_1)=2/5$, $P(\omega_2)=4/15$ και $P(\omega_3)=1/3$
9. (α) $P(A)=90\%$ (β) $P(B)=70\%$
(γ) $P(\Gamma)=80\%$
10. $P(\omega_1)=4/7$, $P(\omega_2)=2/7$ και $P(\omega_3)=1/7$
11. (α) $p=1/5$ και $q=1/10$ (β) άρτιο
12. Όχι

Ενότητα 4.5

1. Α,Ψ, Α, Α, Α
3. Ναι $P(A\cap B)\le P(B)$
4. $P(A\cap B)=0{,}2$, $P(B-A)=0{,}2$
5. $P(B)=0{,}3$, $P(A)=0{,}6$
7. $P(A)=0{,}3$, $P(B)=0{,}5$ και $P(A\cap B)=0{,}1$
9. $P(B-A)=0{,}2$
10. (α) $P(B\cap A)=0{,}3$, (β) $P(A)=0{,}5$
11. $P(A)=1/3$, $P(B)=1/5$ και $P(A\cap B')=2/15$

Ενότητα 4.6

1. (α) 0,95 (β) 0,05 (γ) 0,2
2. (β) $P(A)=7/25$, $P(B)=3/5$,
$P(A\cup B)=22/25$, $P((A\cup B)')=3/25$
3. (α) 0,0797 (β) 0,0003 (γ) 0,0794
4. (α) 0,7 (β) 0,2 (γ) 0,8
5. 0,6
6. $2/3$
7. $1/6$
9. $1-(5/6)^3$

Ενότητα 4.7 (Ανακεφαλαίωση)

1. (α) Σχ.4 (β) Σχ.3 (γ) Σχ.1 (δ) Σχ.2
2. (α) $\Omega=\{A_1A_2B_1, A_1A_2B_2, A_2A_1B_1, A_2A_1B_2, A_1B_1, A_1B_2, A_2B_1, A_2B_2, B_1, B_2\}$
(β) $A=\{A_1A_2B_1, A_1A_2B_2, A_2A_1B_1, A_2A_1B_2\}$
3. (α) $\Omega=\{111,110,101,100,000,011,001,010\}$
(β) $\{110,101,011\}$ (γ) $\{100,000,001,010,011\}$
4. (α) 0,1 (β) 0,2
5. 0,3
6. (α) $1/6$ (β) $1/2$ (γ) $1/2$
7. (α) $1/6$ (β) $4/9$ (γ) $1/6$
8. Ω = { μπλε, κόκκινο, κίτρινο, πράσινο, γκρι }
P(μπλε) = $5/18$, P(κόκκινο) = $5/36$,
P(κίτρινο) = $1/9$, P(πράσινο) = $1/6$,
P(γκρι) = $11/36$,
9. $P(1)=1/4$, $P(3)=1/6$, $P(5)=1/12$ και
$P(2)=1/6$, $P(4)=1/4$, $P(6)=1/12$
10. (α) $\Omega=\{O+, O-, A+, A-, B+, B-, AB+, AB-\}$.
Δεν είναι ισοπίθανα (β) O+ (γ) 57,4%
11. $\sqrt{2}/2$, $1-\sqrt{2}/2$
12. (β) 5 κόκκινοι και 10 μαύροι βώλοι
13. (α) 0,87 (β) 0,83

Ορολογία